Modern Quantization Strategies for Compressive Sensing and Acquisition Systems

Petros Boufounos

petros@boufounos.com

Laurent Jacques

laurent.jacques@uclouvain.be

Signal Acquisition Pipeline

- Typically linear (at least today's discussion)
- Can be designed to be invertible (e.g. Nyquist theorem, compressive sensing, signals with finite rate of innovation, etc...)

- Classical: Linear reconstruction
- Modern: Non-linear, heavy computation (e.g., compressive sensing, finite rate of innovation)

- Highly non-linear
- Not invertible $\Longleftrightarrow$ loss of information
- Design to minimize loss

Today

Quantizer Design

Interaction with measurement system

Optimal reconstruction

Today’s Topics

1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

Today's Topics

1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

SIGNAL REPRESENTATION

Linear Measurement and Reconstruction Model

Linear measurement: $\mathbf{y} = \mathbf{A}\mathbf{x}$
(analysis, sampling)

Linear reconstruction: $\widehat{\mathbf{x}} = \mathbf{S}\mathbf{y}$
(synthesis)

Signal: belongs in a vector space
(e.g., $\mathbb{R}^N$, $\mathbb{C}^N$, bandlimited signals)

Quantized
Measurements

$\mathbf{A}$: Basis expansion (critically sampled) or frame expansion (oversampled)

In absence of quantization: $\mathbf{S} = \mathbf{A}^{-1}$ or $\mathbf{S} = \mathbf{A}^{\dagger}$

Biorthogonal (dual) basis

Dual frame

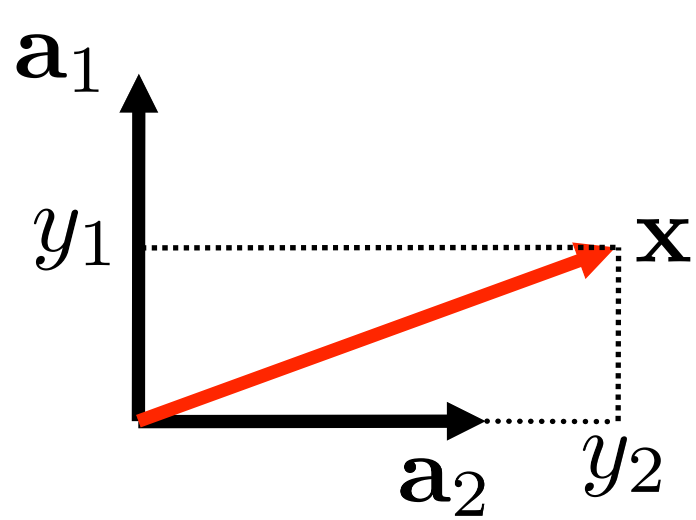
Basis Expansions

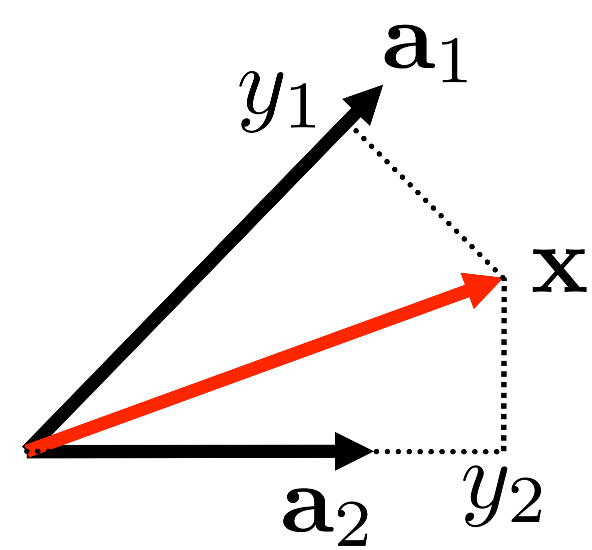
Orthonormal | Oblique

Analysis (Measurement)

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

$$y_i = \langle \mathbf{a}_i, \mathbf{x} \rangle$$



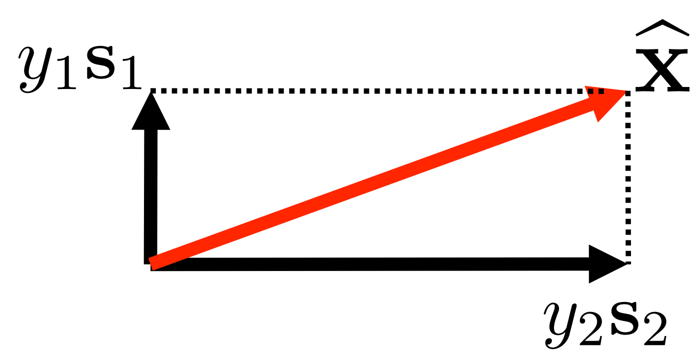


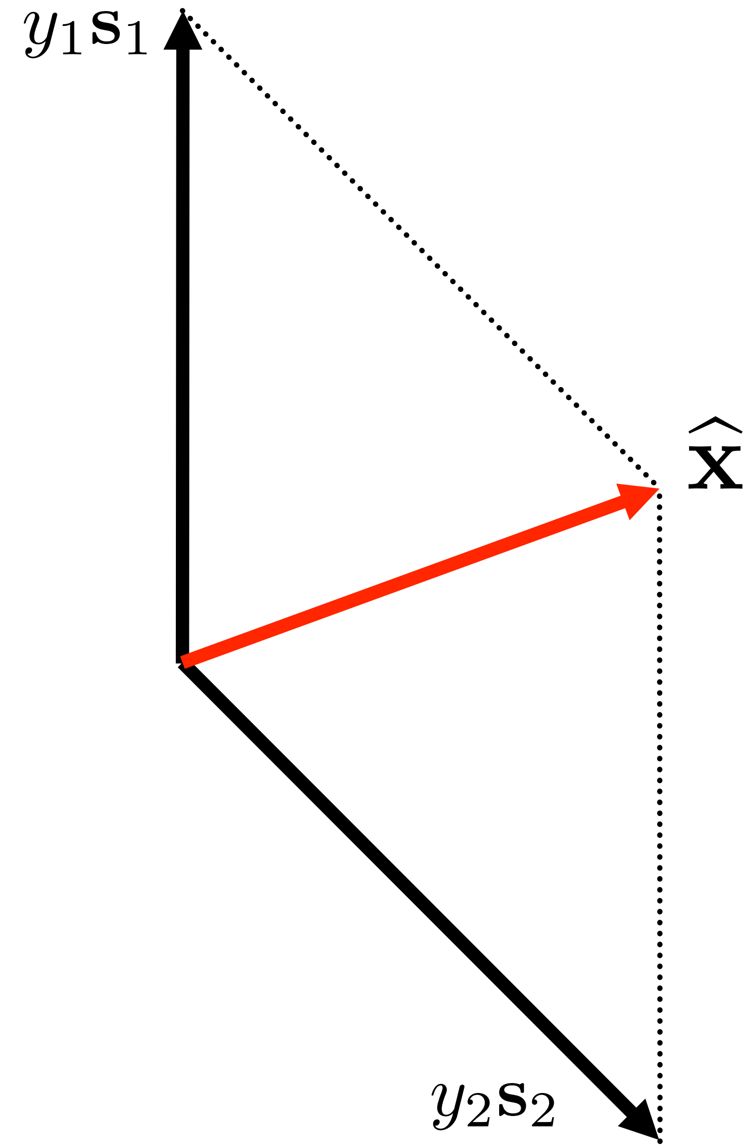
Synthesis (Reconstruction)

$$\widehat{\mathbf{x}} = \mathbf{S}\mathbf{y}$$

$$\widehat{\mathbf{x}} = \sum_i y_i \mathbf{s}_i$$

$$\mathbf{S} = \mathbf{A}^{-1}$$



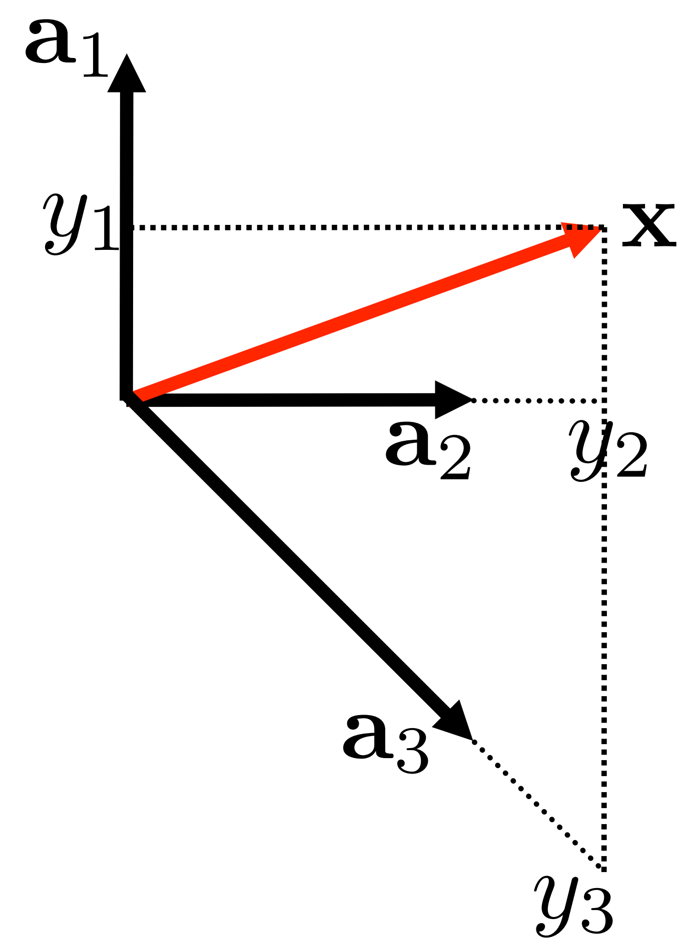


Frame Representations and Oversampling

Analysis (Measurement)

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

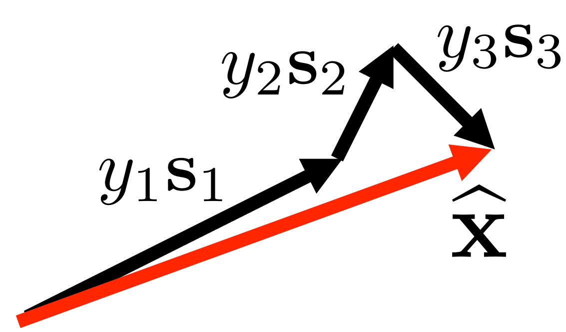
$$y_i = \langle \mathbf{a}_i, \mathbf{x} \rangle$$

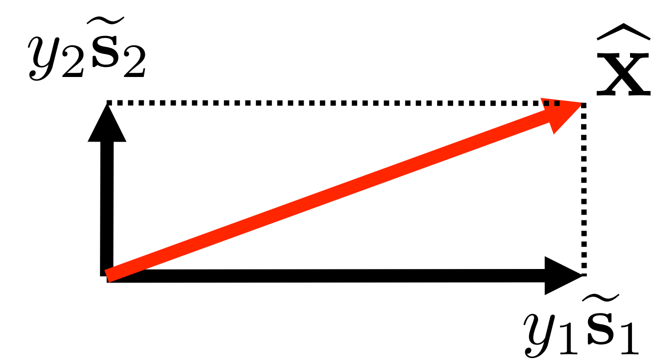


Synthesis (Reconstruction)

$$\widehat{\mathbf{x}} = \mathbf{S}\mathbf{y}$$

$$\widehat{\mathbf{x}} = \sum_i y_i \mathbf{s}_i$$





$\mathbf{S} = \mathbf{A}^{\text{“}-1\text{”}}$, i.e. $\mathbf{S}\mathbf{A} = \mathbf{I}$

$\mathbf{S} = \mathbf{A}^{\dagger}$

Canonical Dual Frame

$\widetilde{\mathbf{S}} = \mathbf{A}_{1,2}^{-1}$

Examples of Frames and Frame Expansions

Matrix Operations in $\mathbb{R}^{M\times N}$

Analysis
(Measurement)

$$\begin{bmatrix} -\mathbf{a}_1- \\ \vdots \\ -\mathbf{a}_M- \end{bmatrix}\begin{bmatrix} | \\ \mathbf{x} \\ | \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} \Leftrightarrow y_i = \langle \mathbf{a}_i, \mathbf{x}\rangle$$

Redundancy
$r=M/N$

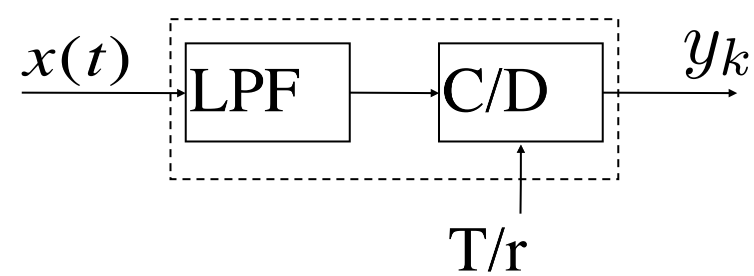
Synthesis
(Reconstruction)

$$\begin{bmatrix} | & & | \\ \mathbf{s}_1 & \cdots & \mathbf{s}_M \\ | & & | \end{bmatrix}\begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} | \\ \mathbf{x} \\ | \end{bmatrix} \Leftrightarrow \mathbf{x} = \sum_i y_i \mathbf{s}_i$$

r-times Oversampling:

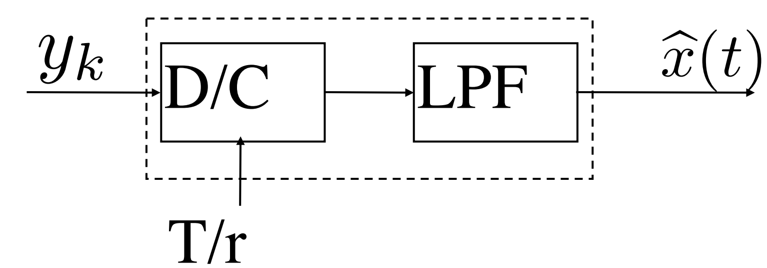
Analysis
(Measurement)

$$y_i = \int_{-\infty}^{+\infty} x(t)\frac{1}{rT}\mathrm{sinc}\left(\frac{r}{T}t - i\right)dt \Leftrightarrow y_i = \langle \mathbf{a}_i, \mathbf{x}\rangle$$

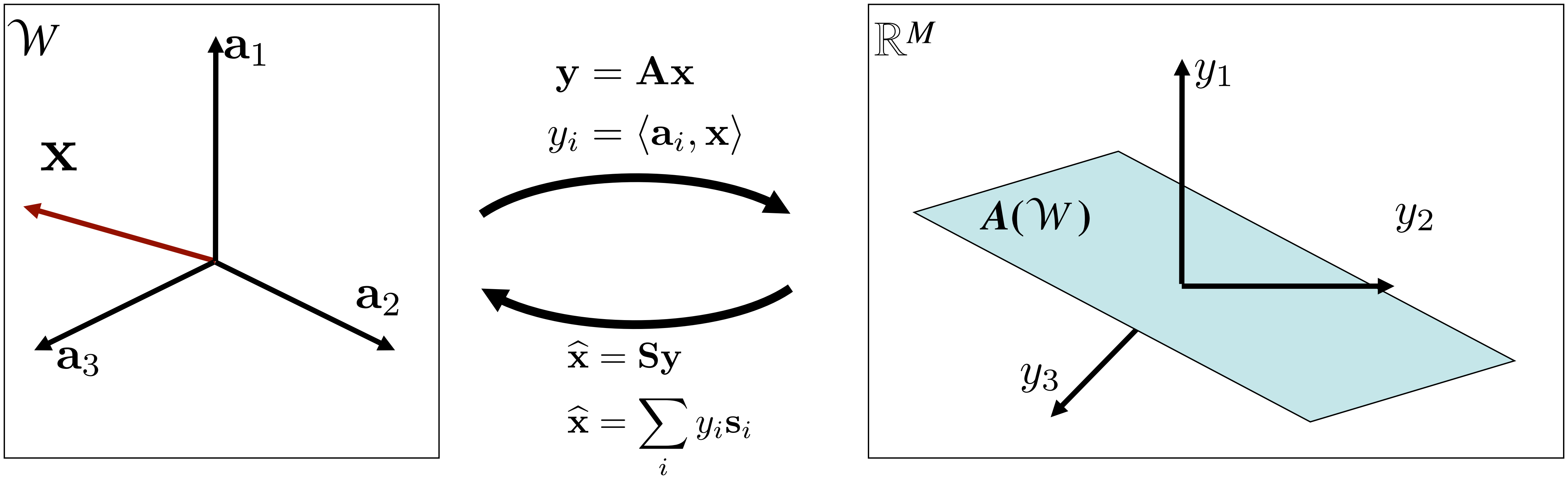


Synthesis
(Reconstruction)

$$x(t) = \sum_i y_i \mathrm{sinc}\left(\frac{r}{T}t - i\right) \Leftrightarrow \mathbf{x} = \sum_i y_i \mathbf{s}_i$$



Frame Expansion/Oversampling: Subspace Mapping



Signal Space $\mathcal{W}$

Frame: $\{\mathbf{a}_i, i = 1, \ldots, M \mid \mathbf{a} \in \mathcal{W}\}$

$\dim(\mathcal{W}) = N$

Coefficient/Measurement Space $\mathbb{R}^M$

Image is N-dimensional

$\dim(\mathbf{A}(\mathcal{W})) \leq \operatorname{rank}(\mathbf{A}) \leq N < M$

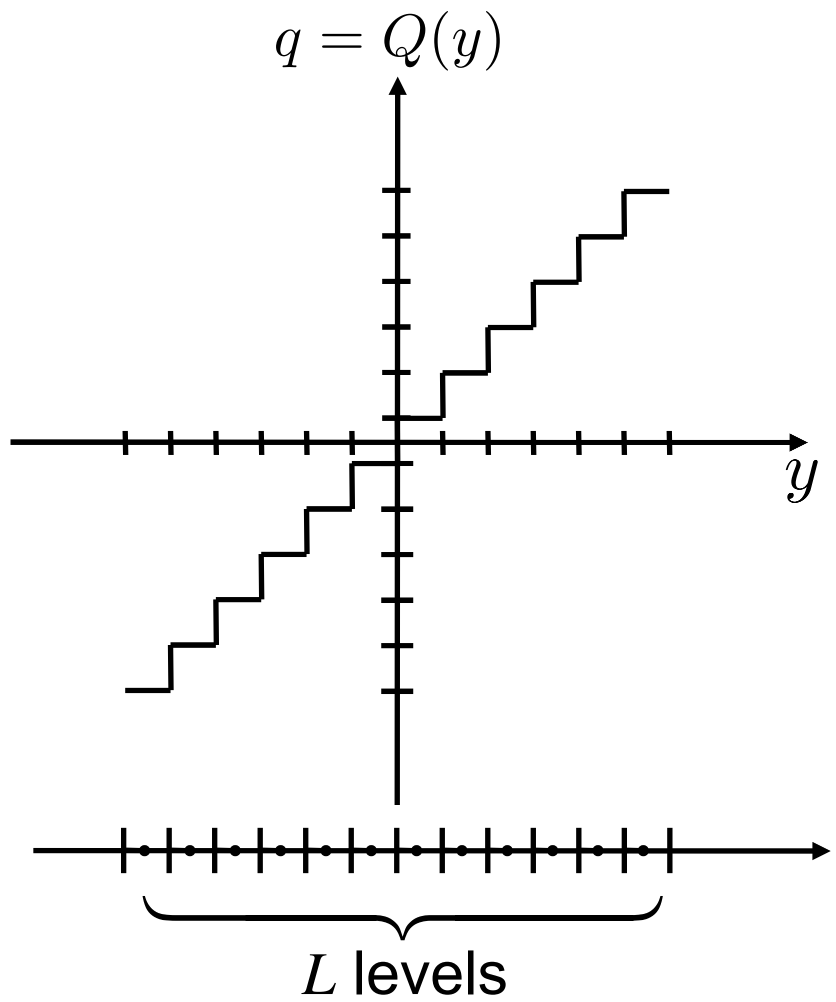
Frames provide **redundancy**

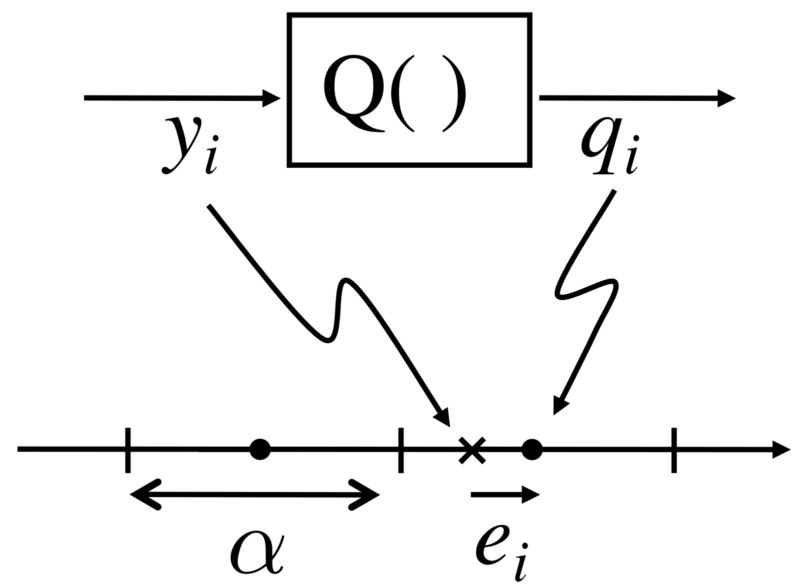
Mechanism: nullspace of synthesis operator.

Redundancy can be exploited for quantization **robustness**

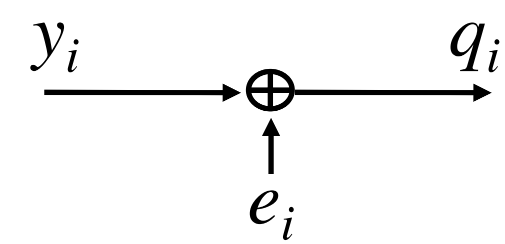
SCALAR QUANTIZATION

Scalar Quantization





=

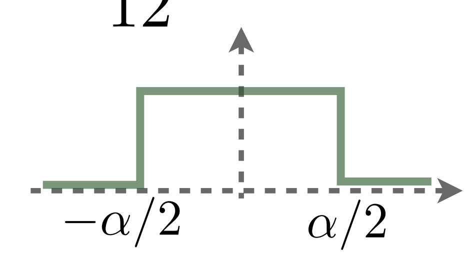


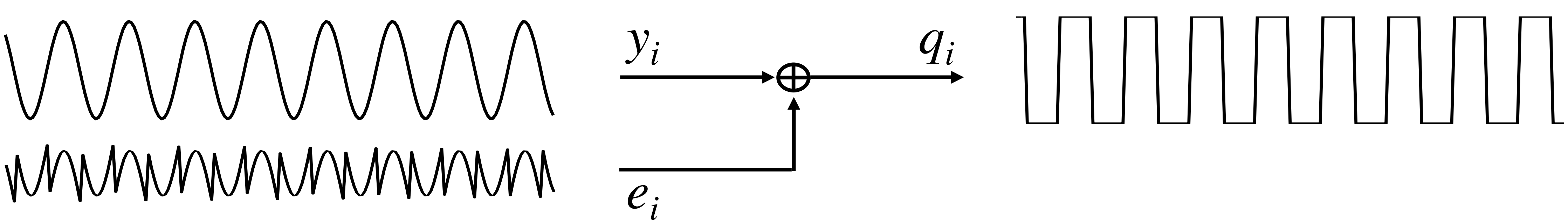
$$|e_i| \leq \frac{\alpha}{2}$$

L level quantizer: $B \sim \log_2(L)$ bits per coefficient

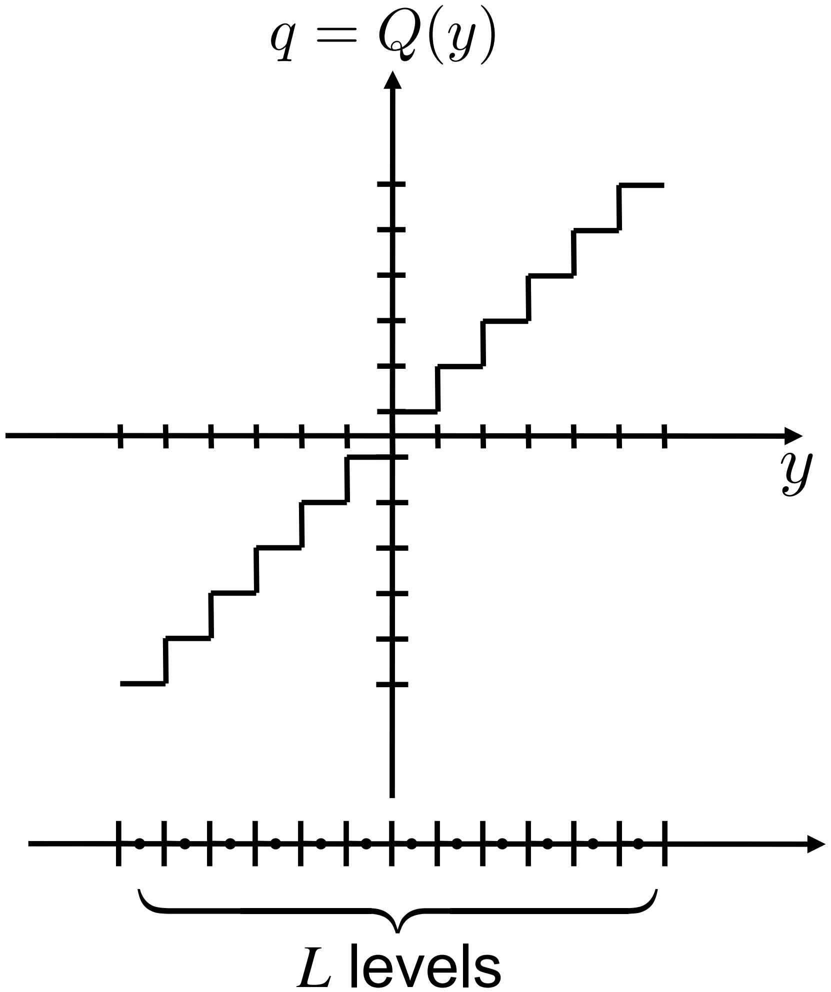
Additive noise model: e_i uncorrelated, uniform in $\pm\frac{\alpha}{2}$

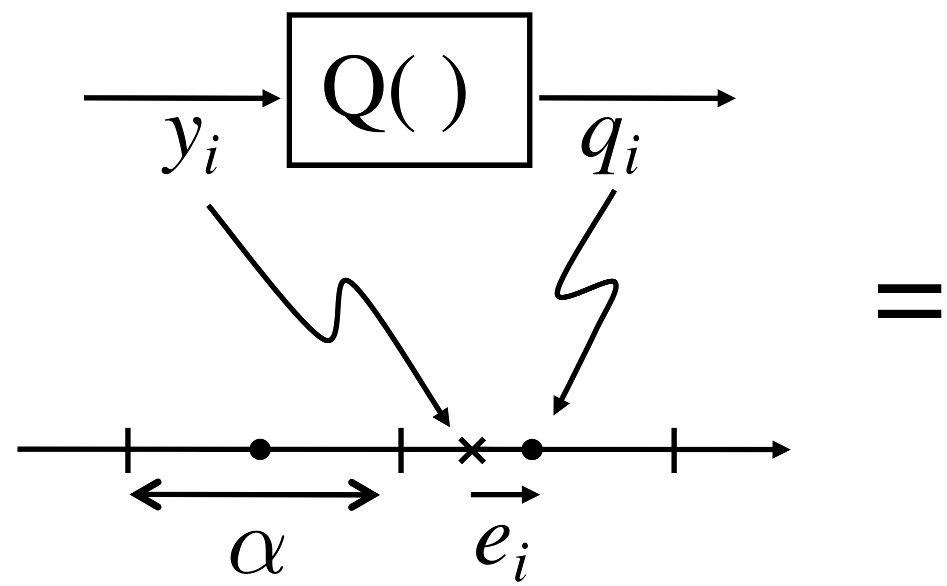
$$\sigma_e^2 = \frac{\alpha^2}{12} \propto 2^{-B+1}$$

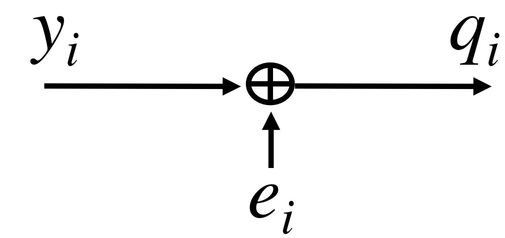




Scalar Quantization





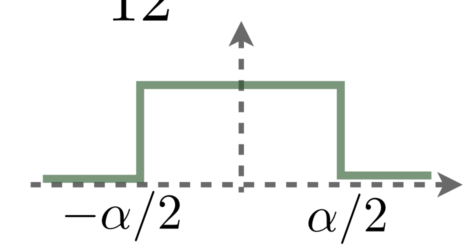


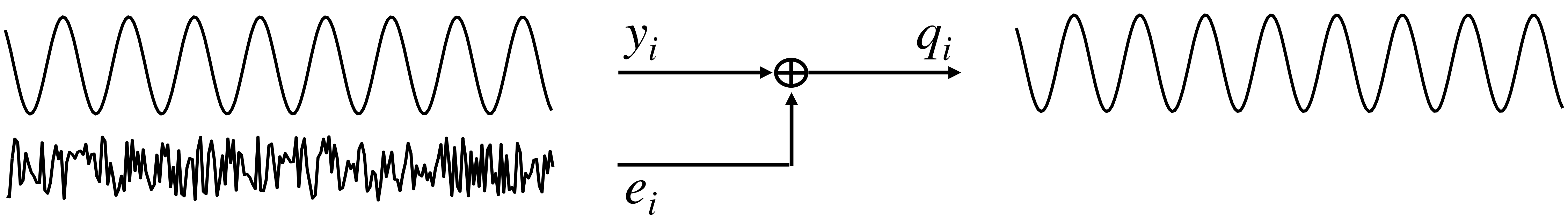
$$|e_i| \leq \frac{\alpha}{2}$$

L level quantizer: $B \sim \log_2(L)$ bits per coefficient

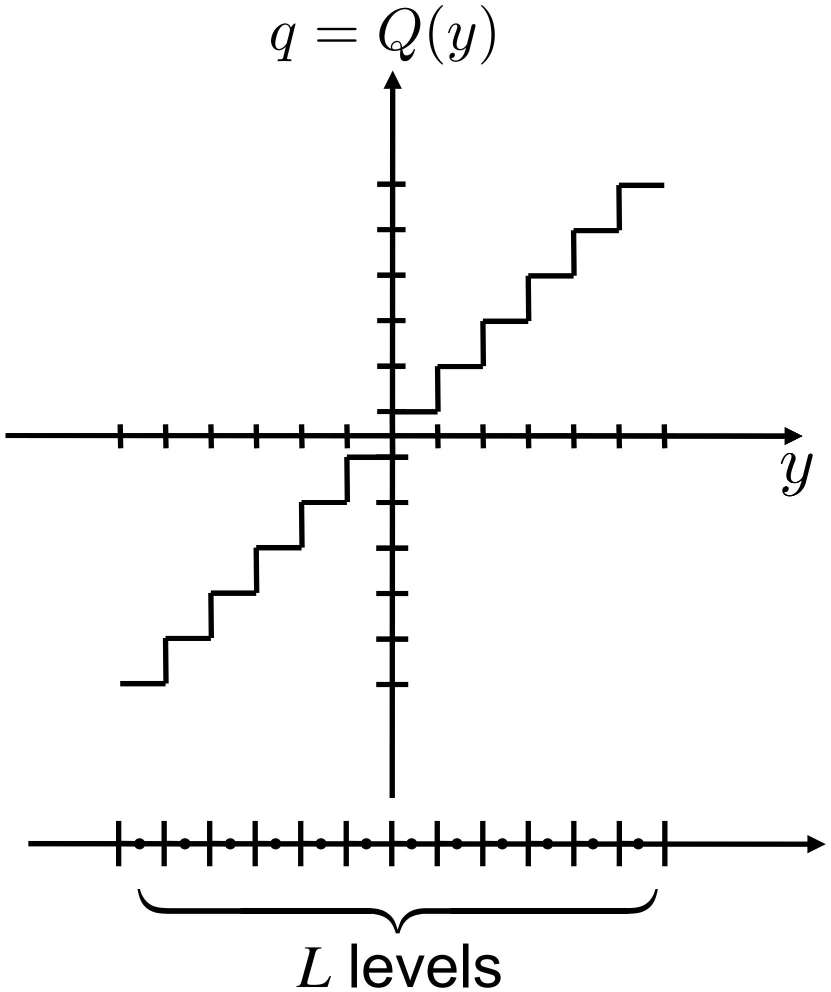
$$\sigma_e^2 = \frac{\alpha^2}{12} \propto 2^{-B+1}$$

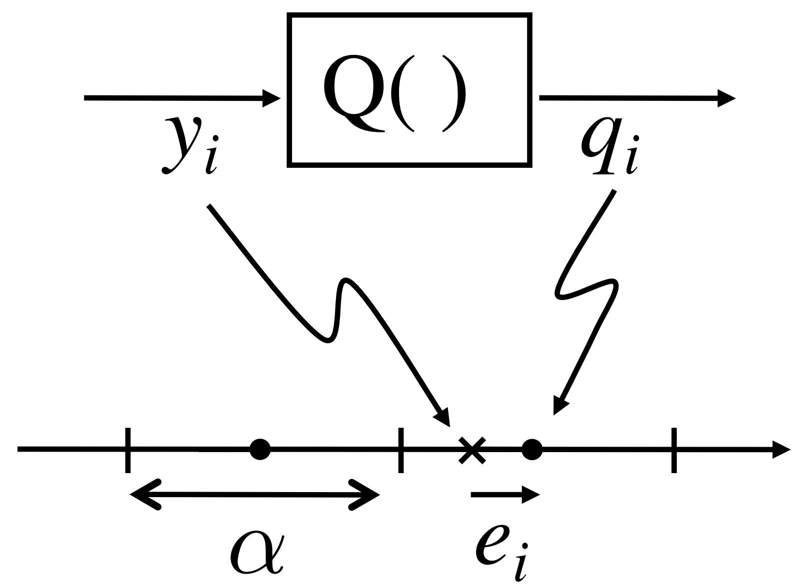
Additive noise model: e_i uncorrelated, uniform in $\pm\frac{\alpha}{2}$



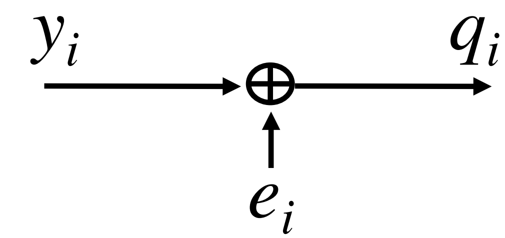


Scalar Quantization





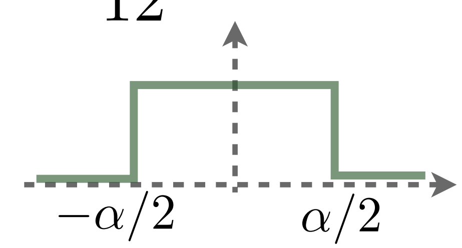
=



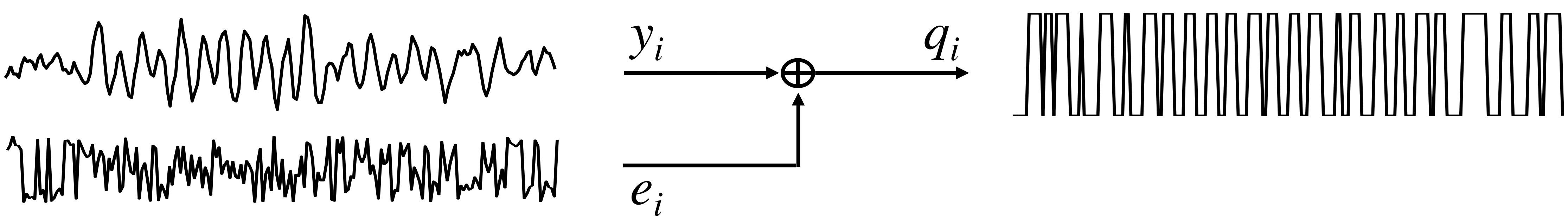
$$|e_i| \leq \frac{\alpha}{2}$$

L level quantizer: $B \sim \log_2(L)$ bits per coefficient

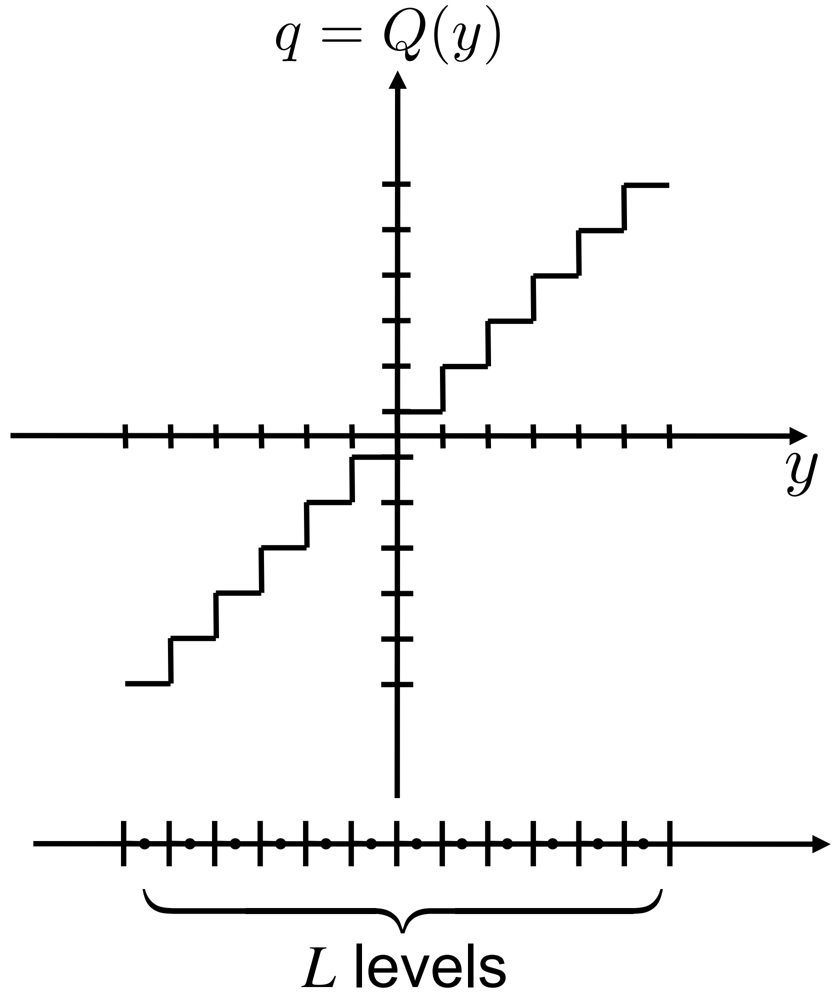
$$\sigma_e^2 = \frac{\alpha^2}{12} \propto 2^{-B+1}$$

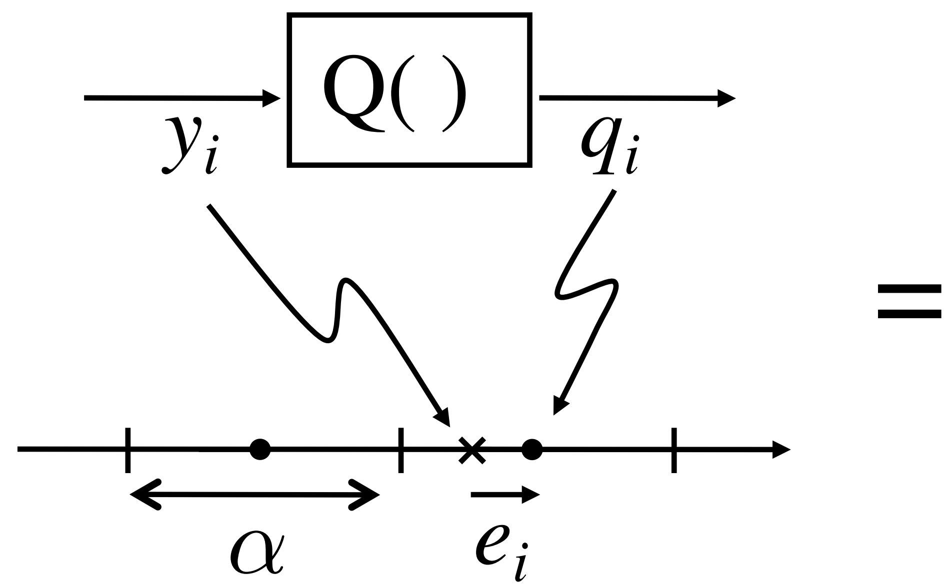


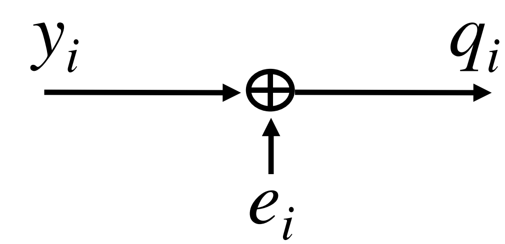
Additive noise model: e_i uncorrelated, uniform in $\pm\frac{\alpha}{2}$



Scalar Quantization





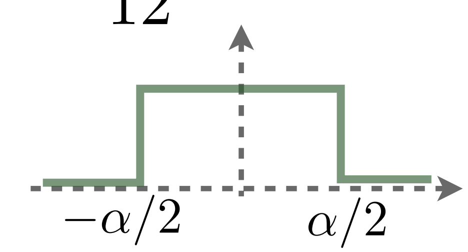


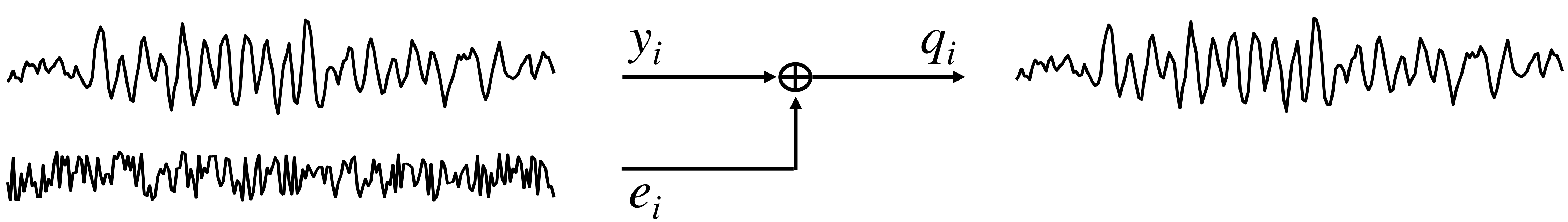
$$|e_i| \le \frac{\alpha}{2}$$

L level quantizer: $B \sim \log_2(L)$ bits per coefficient

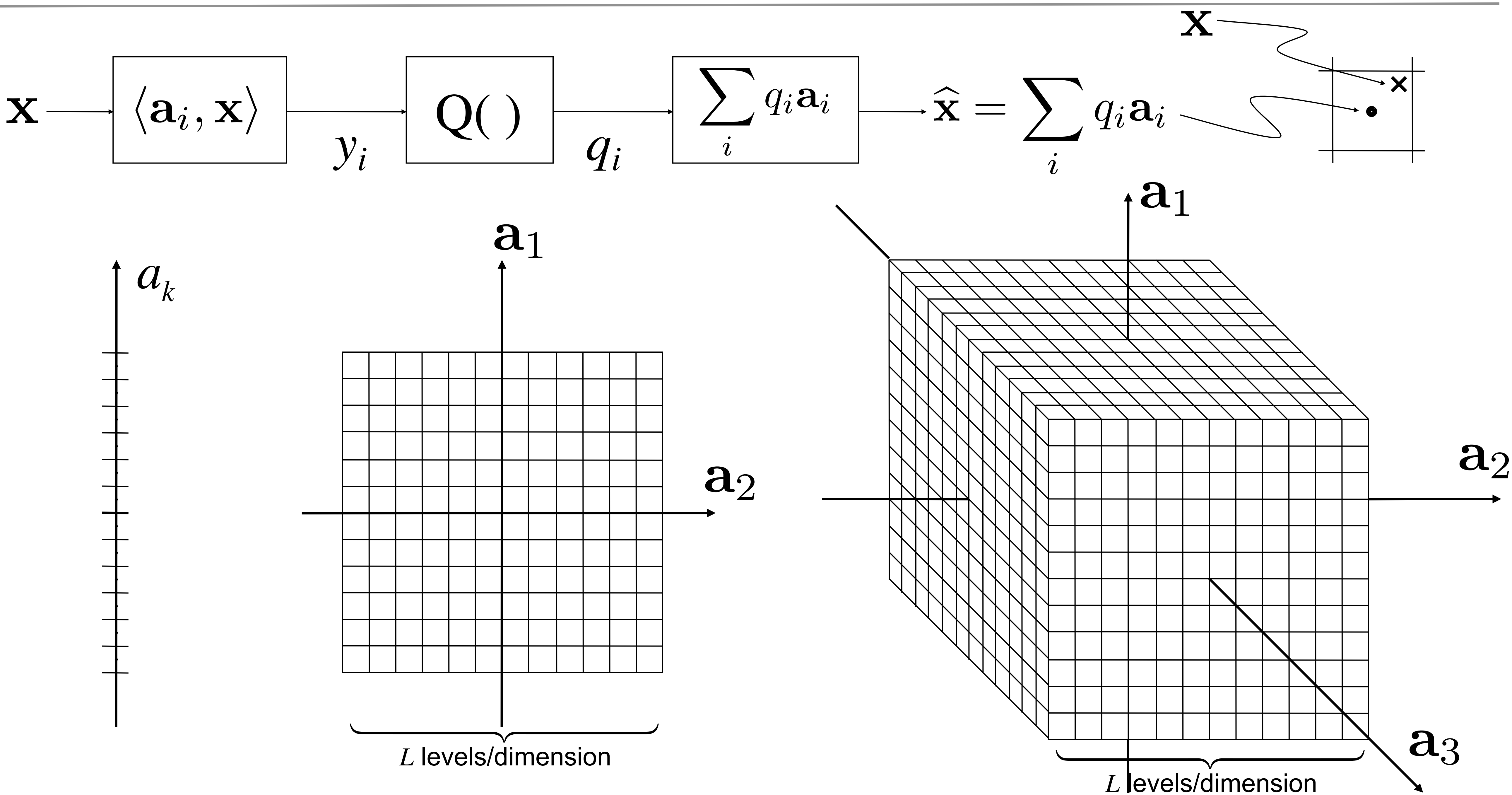
Additive noise model: e_i uncorrelated, uniform in $\pm\dfrac{\alpha}{2}$

$$\sigma_e^2 = \frac{\alpha^2}{12} \propto 2^{-B+1}$$





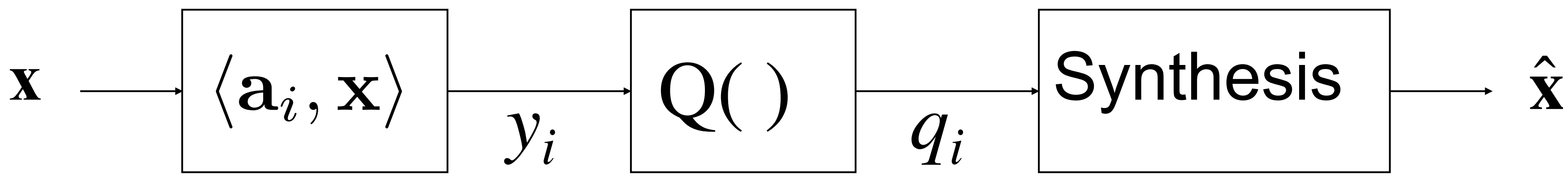
Quantization of Orthonormal Basis Expansions

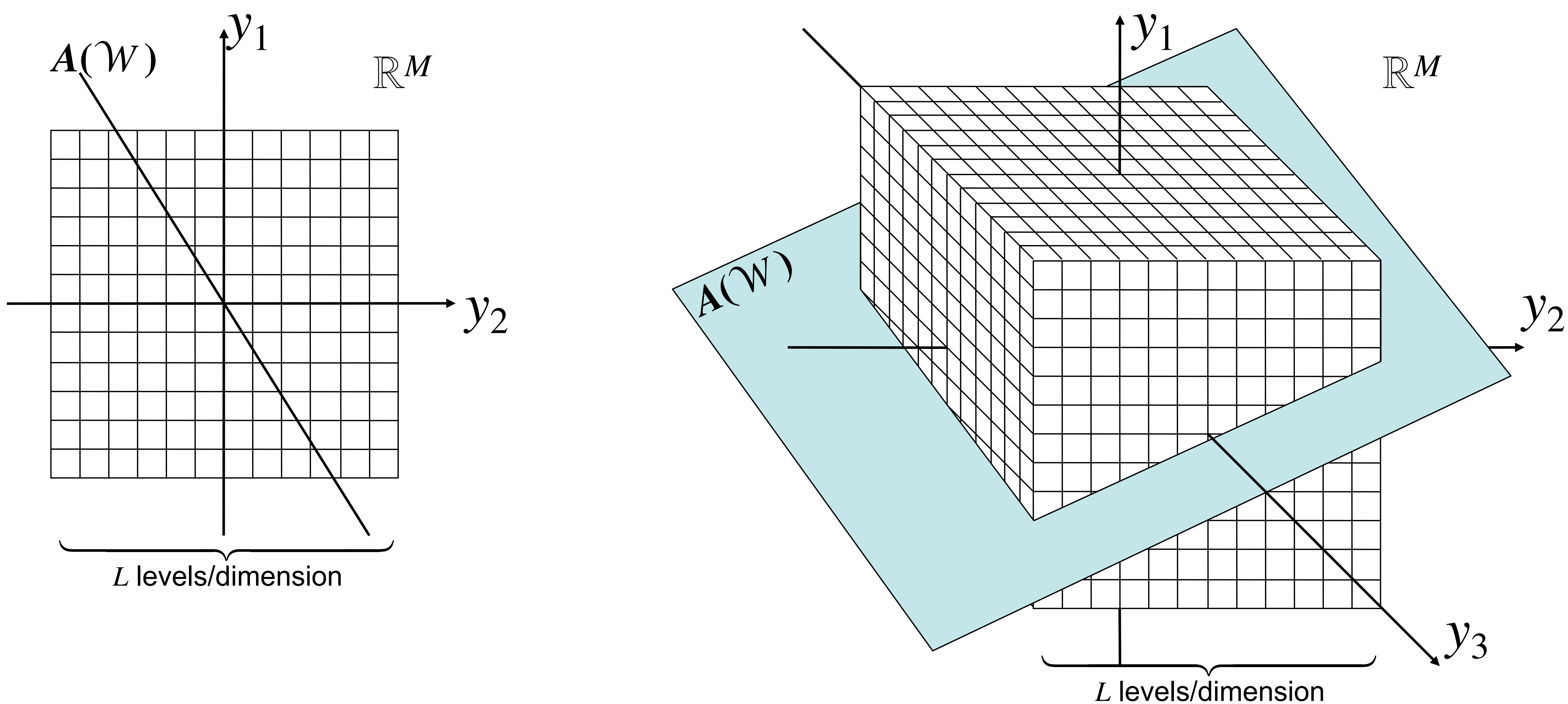


$B = \log_2 L$ bits per coefficient
M expansion coefficients $\Rightarrow R = MB = M \log_2 L$ bits used (rate)

Total Error $\varepsilon = O(c^{-R})$

Quantization of Frame Representations

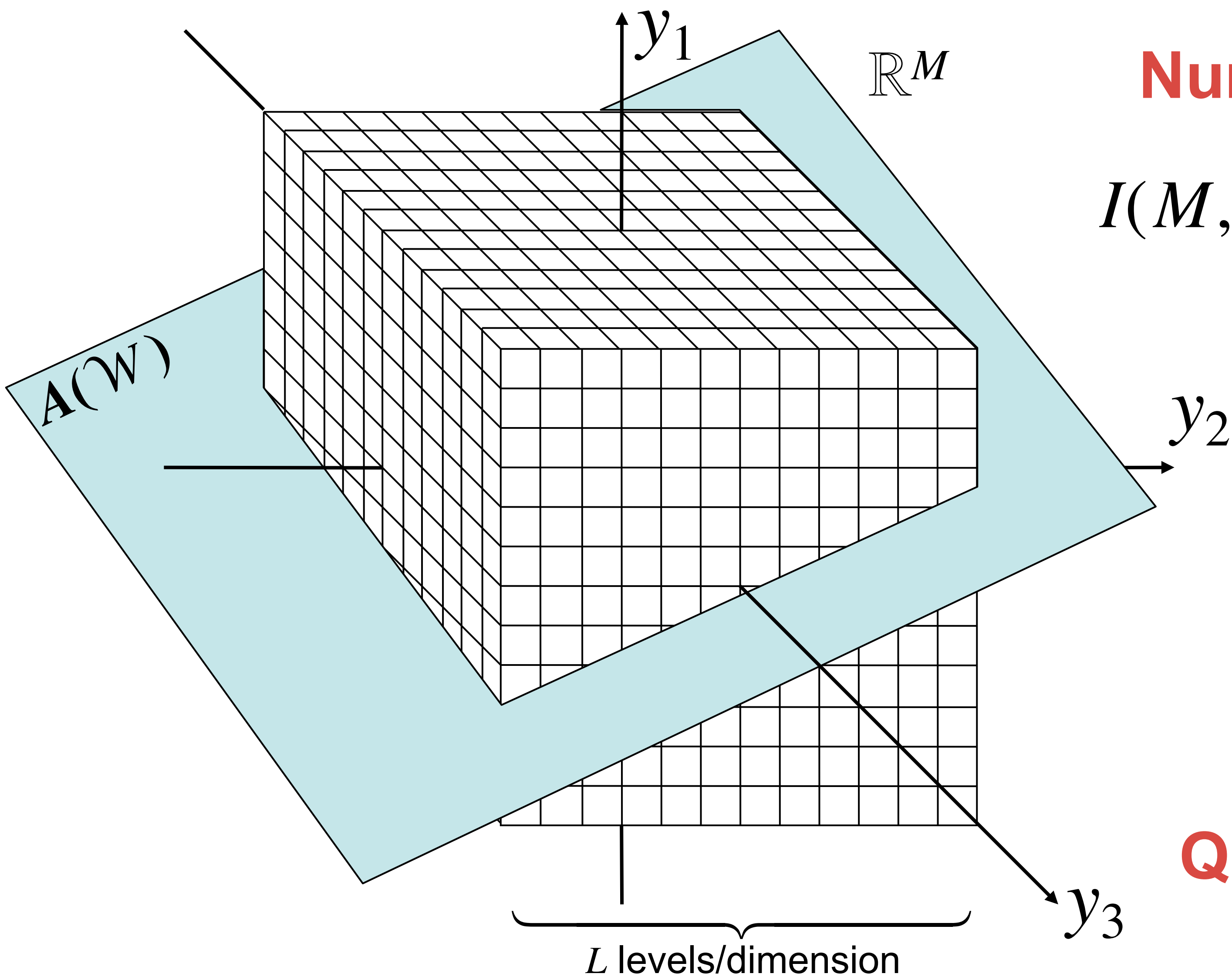




Very few quantization cells are intersected!

Oversampling provides **robustness**,
but also introduces **inefficiency**

Bounds on Scalar Quantization



Number of cells intersected $I(M,N,L)$:

$$I(M,N,L) \le (2L)^N \binom{M}{N} \le \left(\frac{2LMe}{N}\right)^N = (2Lre)^N$$

Bit-use efficiency:

$$\frac{\text{bits necessary}}{\text{bits used}} \le \frac{\log_2\left(I(M,N,L)\right)}{M\log_2 L} \le \frac{\log_2(2Lre)}{r\log_2 L},$$

Quantization Error Reduction Rate:

$$\varepsilon^2 = \Omega(r^{-2})$$

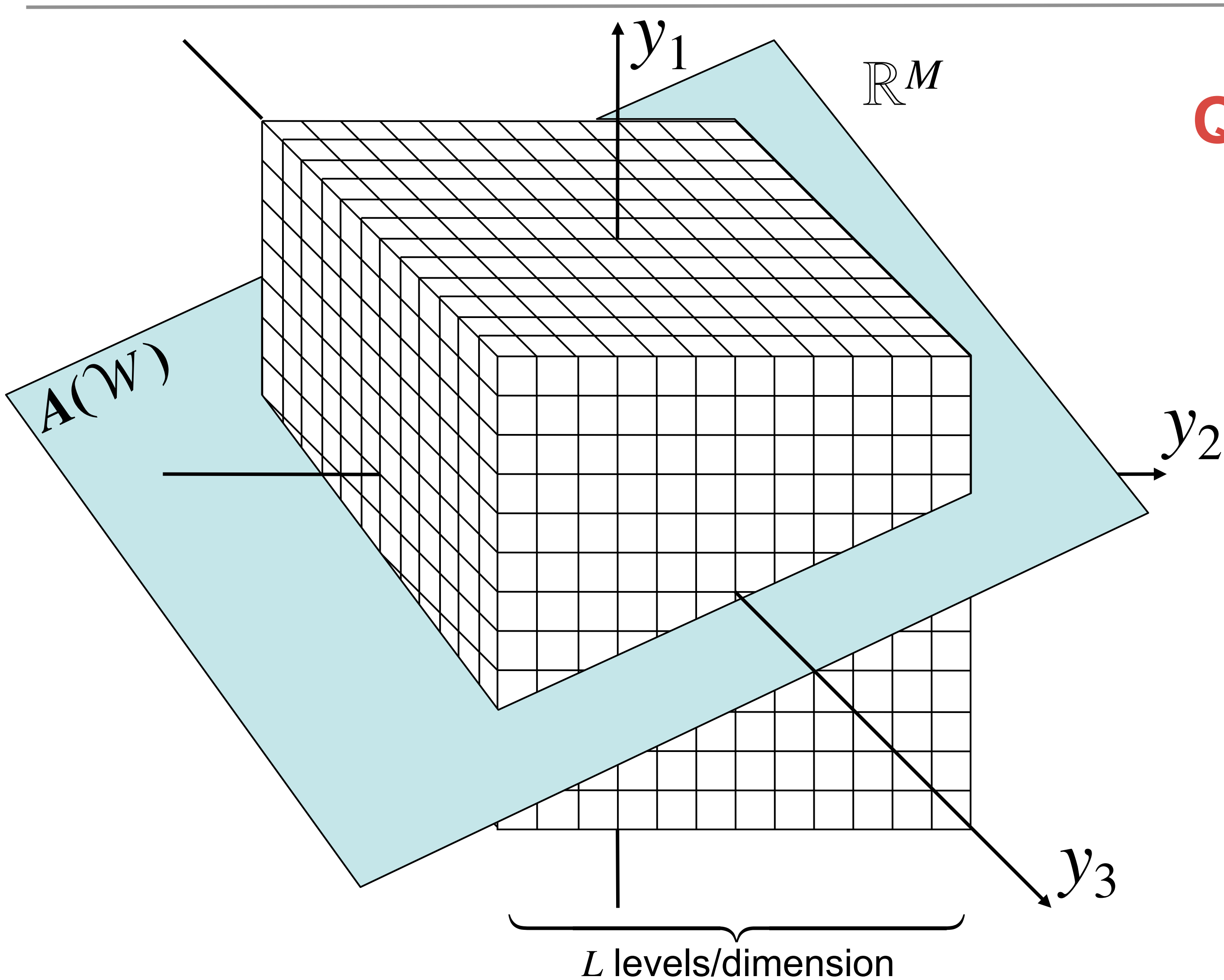
Total error $\varepsilon = \Omega(1/R)$

(vs. $\varepsilon = O(c^{-R})$ for basis expansions)

Oversampled scalar quantization is inefficient!

- Thao N. T. and Vetterli M., "Lower bound on the mean-squared error in oversampled quantization of periodic signals using vector quantization analysis," *IEEE Trans. Info. Theory*, vol. 42, no. 2, pp. 469–479, Mar. 1996.
- Boufounos P. T., "Quantization and erasures in frame representations," *MIT D.Sc. Thesis*, Cambridge, MA, January 2006.

Bounds on Scalar Quantization



Quantization Error Reduction Rate:

$$\varepsilon^2 = \Omega(r^{-2})$$

But: Can we achieve it?

Linear reconstruction

$$\mathbf{q} = Q(\mathbf{A}\mathbf{x})$$
$$\Rightarrow \widehat{\mathbf{x}} = \mathbf{A}^{\dagger}\mathbf{q}$$
$$\Rightarrow \varepsilon^2 = \Omega(r^{-1})$$

Solution: "Consistent reconstruction"

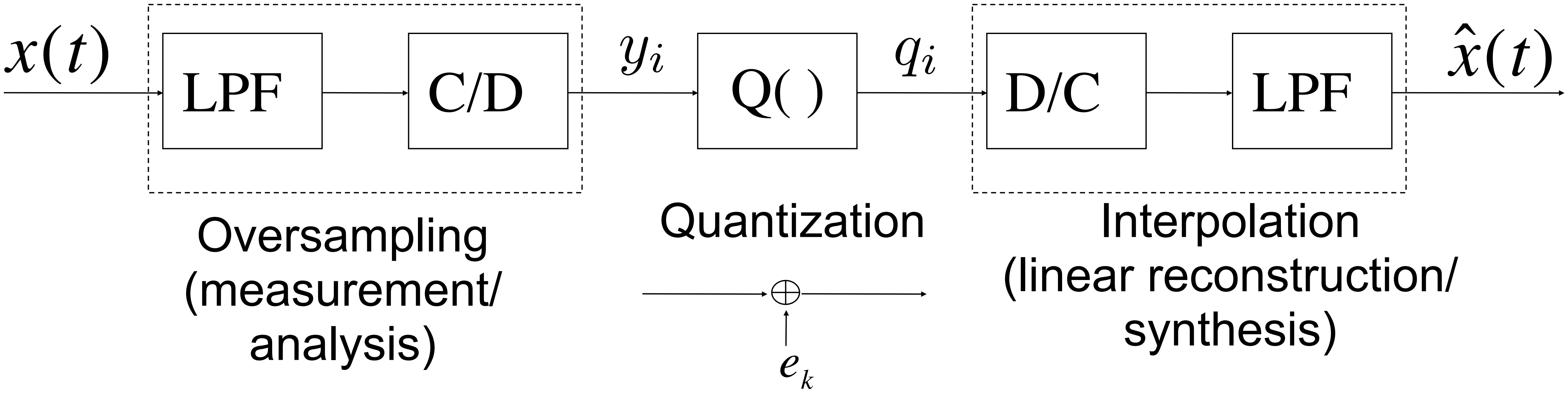
Reconstruct a signal that explains *quantized* measurements

$$\widehat{\mathbf{x}} \quad \text{s.t. } \mathbf{q} = Q(\mathbf{A}\widehat{\mathbf{x}})$$

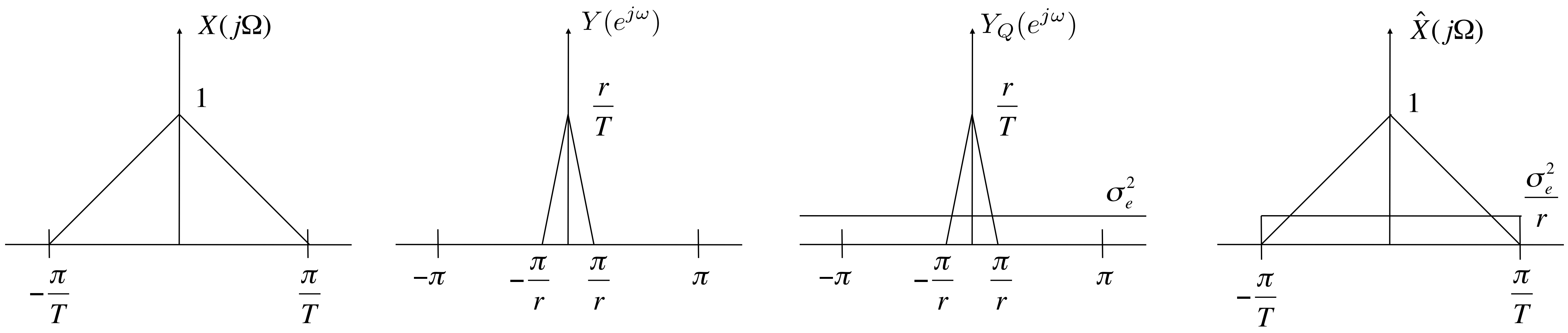
$$\text{i.e. } q_i - \frac{\alpha}{2} \leq \langle \mathbf{a}_i, \widehat{\mathbf{x}} \rangle \leq q_i + \frac{\alpha}{2}$$

- Thao N. and Vetterli M.,"Reduction of the MSE in R-times oversampled A/D conversion O(1/R) to O(1/R^2)," *IEEE Trans. Signal Processing*, vol. 42, no. 1, pp. 200–203, Jan 1994.

Oversampling and Quantization



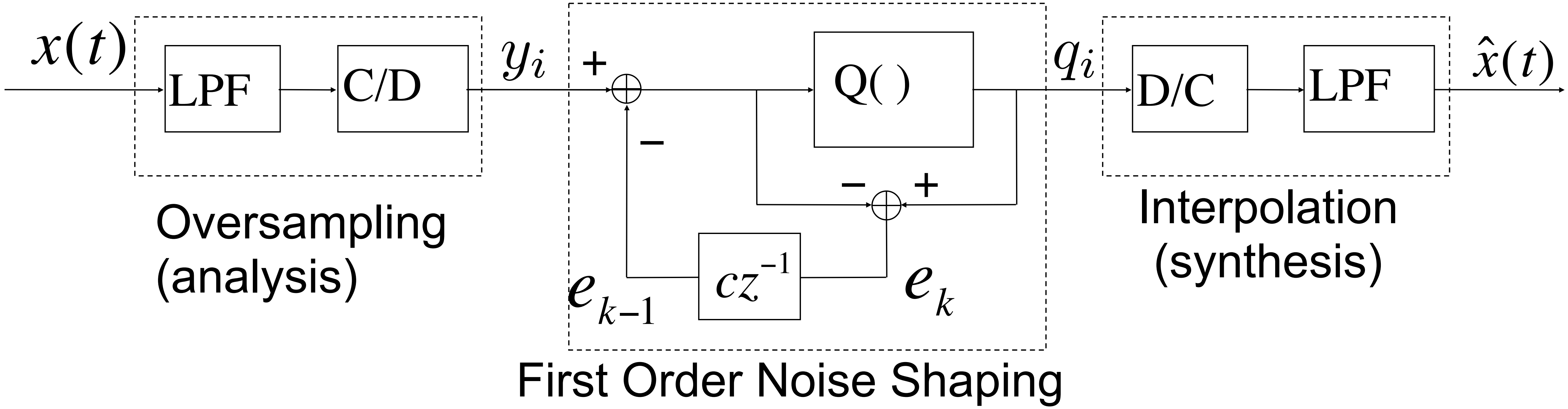
Using the additive noise model, e_k uncorrelated, uniform in $\pm\frac{\Delta}{2}$

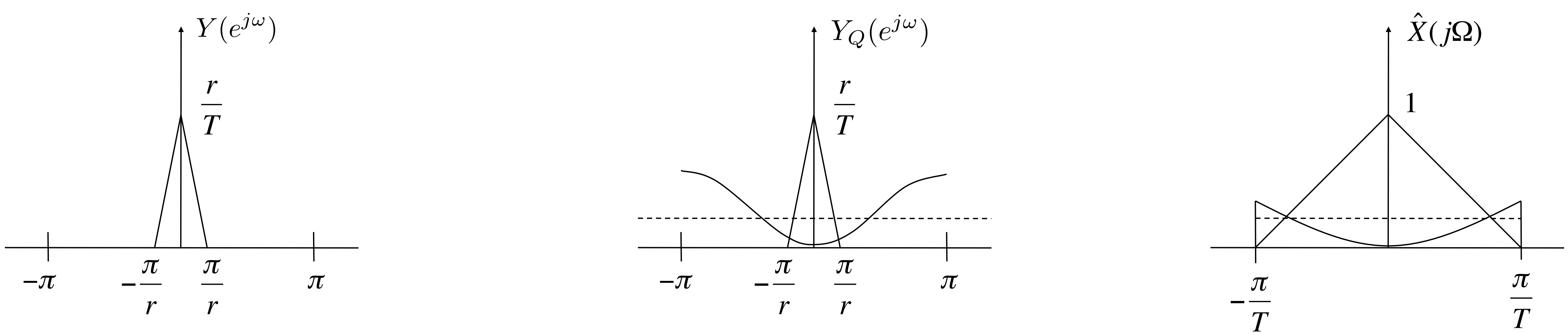


Tradeoff: Gain **1 bit for** each **4 times oversampling**

Quantization error $\varepsilon^2 \sim \Omega(1/r)$

First Order Noise Shaping



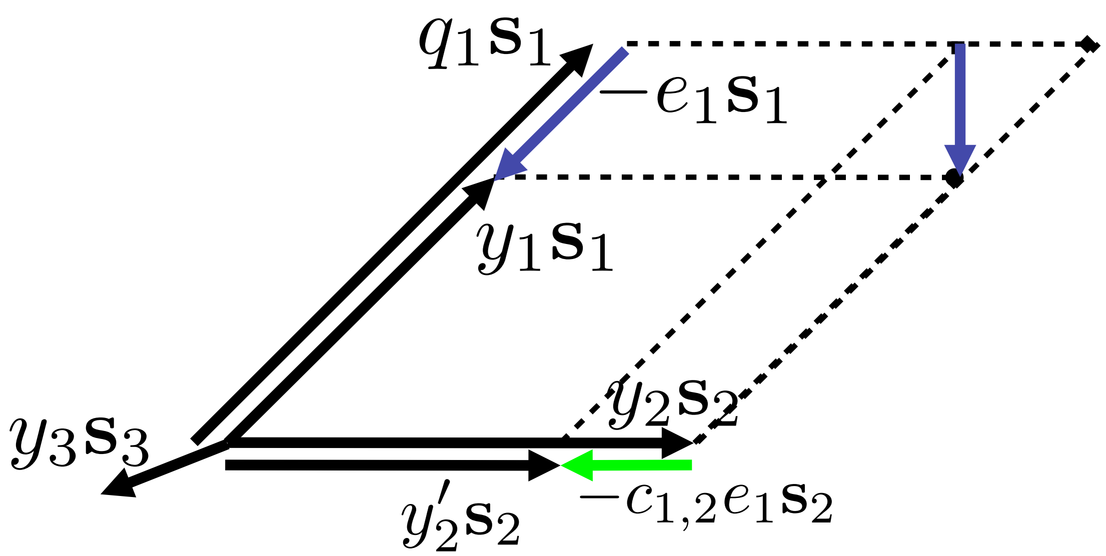


Optimal choice $c = \mathrm{sinc}\left(\frac{\pi}{r}\right)$ (≈ 1 for $r \geq 4$)

Can we extend noise shaping to arbitrary frames?

Error compensation using projections

$$\mathbf{x} = y_1\mathbf{s}_1 + y_2\mathbf{s}_2 + y_3\mathbf{s}_3$$



1. Quantization

$$\mathbf{x} = q_1\mathbf{s}_1 + y_2\mathbf{s}_2 + y_3\mathbf{s}_3$$

2. Compensation using projection

$$y_2' = y_2 - e_1 c_{1,2}$$

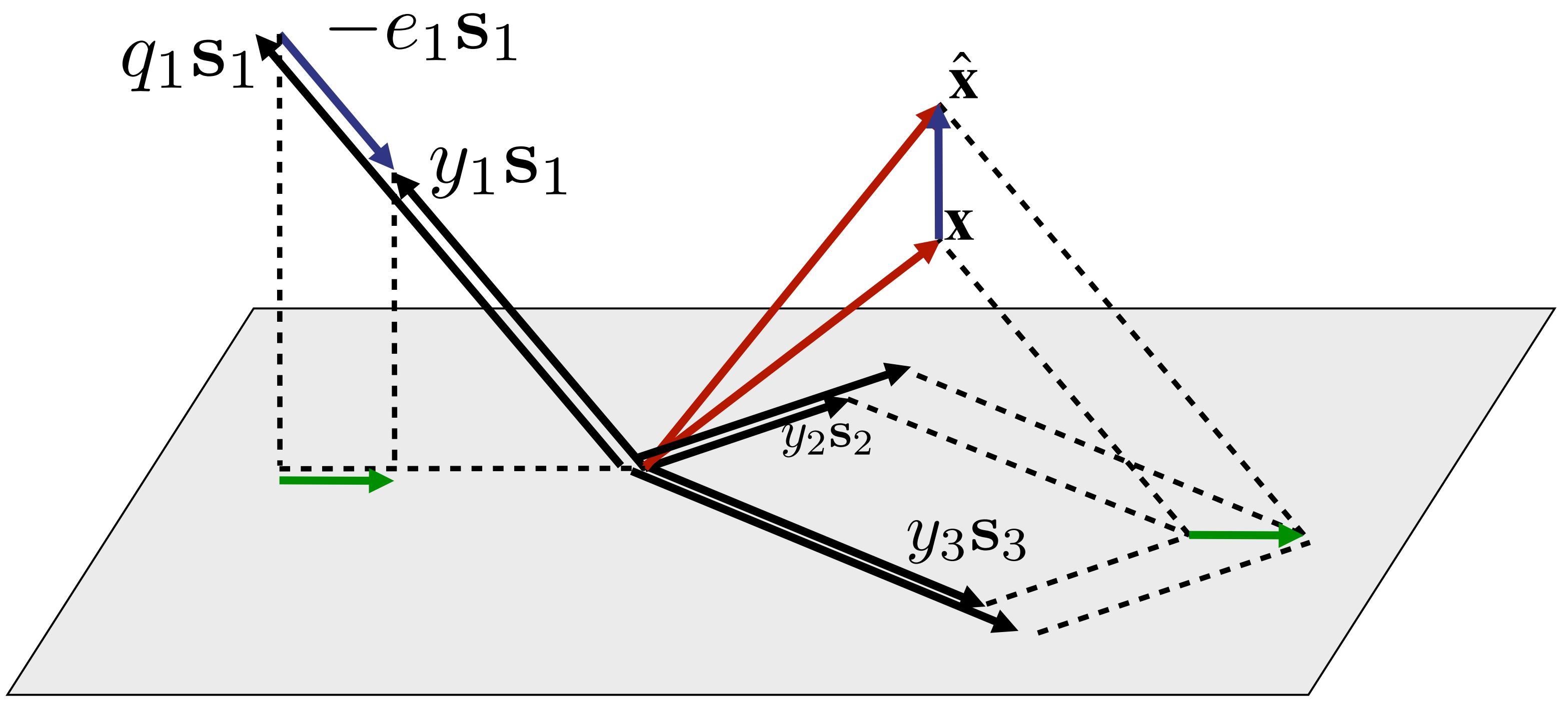
$$\mathbf{x} = q_1\mathbf{s}_1 + y_2'\mathbf{s}_2 + y_3\mathbf{s}_3 - e_1(\mathbf{s}_1 - c_{1,2}\mathbf{s}_2)$$

$$\text{Incremental error :} \quad -e_1(\mathbf{s}_1 - c_{1,2}\mathbf{s}_2) \Rightarrow c_{1,2} = \frac{\langle \mathbf{s}_1, \mathbf{s}_2 \rangle}{\|\mathbf{s}_2\|^2}$$

Compensation linear in the error.
Coefficients can be pre-computed.

- Boufounos P. and Oppenheim A. V., "Quantization noise shaping on arbitrary frame expansions," *EURASIP Journal on Applied Signal Processing, Special issue on Frames and Overcomplete Representations in Signal Processing, Communications, and Information Theory*, vol. 2006, pp. Article ID 53 807, 12 pages, DOI:10.1155/ASP/2006/53 807, 2006.

Higher Order Projections



$$\mathbf{x} = y_1\mathbf{s}_1 + y_2\mathbf{s}_2 + y_3\mathbf{s}_3$$

1. Quantization:

$$q_i = Q(y_i') = y_i' + e_i$$

2. Projection:

$$y_{i+1}' = y_{i+1} - c_{i,i+1}e_i$$

$$\vdots$$

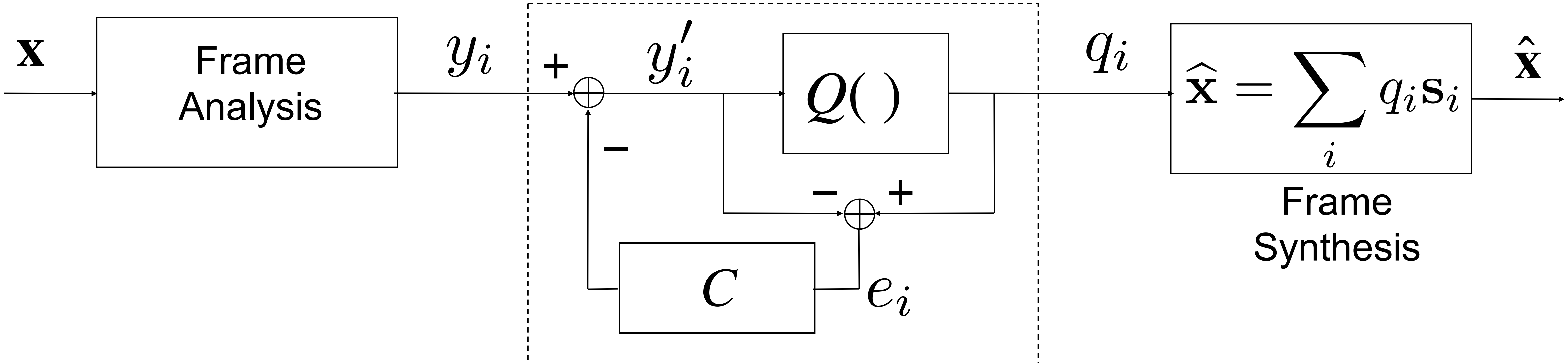
$$y_{i+p}' = y_{i+p} - c_{i,i+p}e_i$$

Projection coefficients $c_{i,i+k}$ designed to reduce or minimize

$$\left\| \mathbf{s}_i - \sum_{k=1}^{p} c_{i,i+k}\mathbf{s}_{i+k} \right\|_2$$

- Boufounos P. and Oppenheim A. V., "Quantization noise shaping on arbitrary frame expansions," *EURASIP Journal on Applied Signal Processing, Special issue on Frames and Overcomplete Representations in Signal Processing, Communications, and Information Theory*, vol. 2006, pp. Article ID 53 807, 12 pages, DOI:10.1155/ASP/2006/53 807, 2006.
- Benedetto J. J., Powell A. M., and Yilmaz O., "Sigma-Delta quantization and finite frames," *IEEE Trans. Info. Theory*, vol. 52, no. 5, pp. 1990–2005, May 2006.
- Deift, P., Krahmer, F. and Güntürk, C. S. (2011), "An optimal family of exponentially accurate one-bit Sigma-Delta quantization schemes." *Comm. Pure Appl. Math.*, 64: 883–919. doi: 10.1002/cpa.20367

System Description



1. Quantization:

$$q_i = Q(y_i') = y_i' + e_i$$

2. Projection (Coefficient Update):

$$y_{i+1}' = y_{i+1} - c_{i,i+1} e_i$$

$$\vdots$$

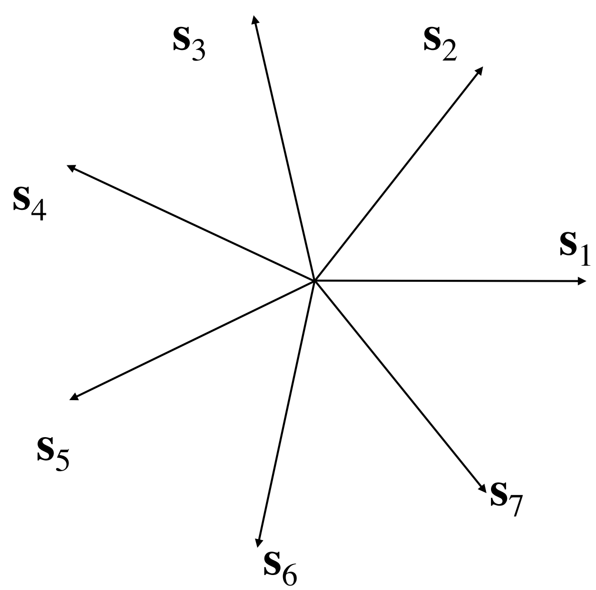
$$y_{i+p}' = y_{i+p} - c_{i,i+p} e_i$$

Achievable error decay $\varepsilon = O(r^{p+1})$

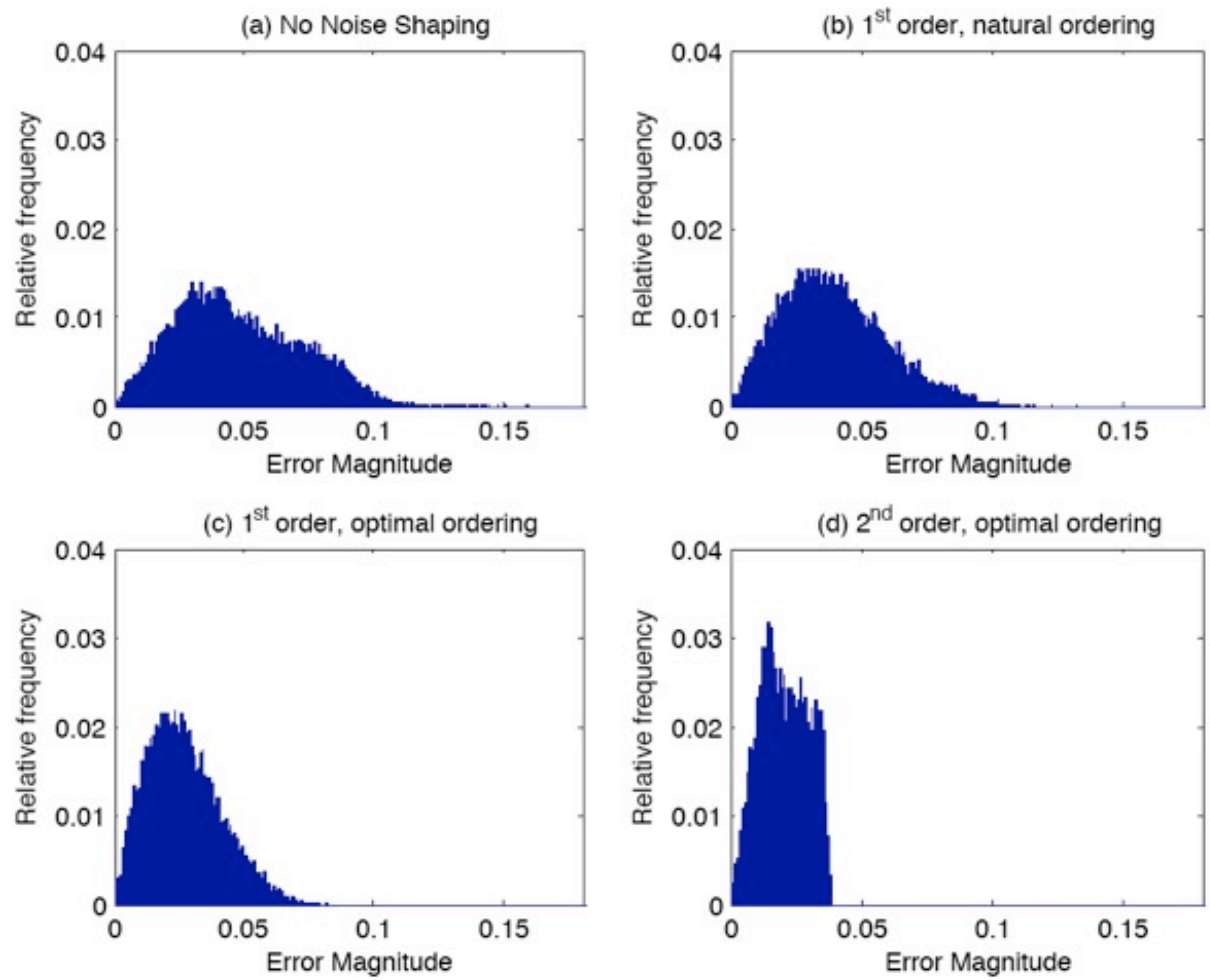
- Benedetto J. J., Powell A. M., and Yilmaz O., "Sigma-Delta quantization and finite frames," *IEEE Trans. Info. Theory*, vol. 52, no. 5, pp. 1990–2005, May 2006.

Example: Simulation Results

Frame: 7th roots of unity



Histogram of the Error Magnitude



- Random points on the plane, uniform inside the unit circle.
- Quantization points: (-7/8, -5/8, -3/8, -1/8, 1/8, 3/8, 5/8, 7/8)
- Optimal ordering (one of many) is: $\left(\mathbf{s}_1, \mathbf{s}_4, \mathbf{s}_7, \mathbf{s}_3, \mathbf{s}_6, \mathbf{s}_2, \mathbf{s}_5\right)$

Further Reading

- Thao N. and Vetterli M.,"Reduction of the MSE in R-times oversampled A/D conversion O(1/R) to O(1/R^2)," *IEEE Trans. Signal Processing*, vol. 42, no. 1, pp. 200–203, Jan 1994.

- Thao N. T. and Vetterli M., "Lower bound on the mean-squared error in oversampled quantization of periodic signals using vector quantization analysis," *IEEE Trans. Info. Theory*, vol. 42, no. 2, pp. 469–479, Mar. 1996.

- Thao N. T., "Vector quantization analysis of ΣΔ modulation," *IEEE Trans. Signal Processing*, vol. 44, no. 4, pp. 808–817, Apr. 1996.

- Goyal V. K.,Vetterli M., and Thao N. T., "Quantized overcomplete expansions in R^N : Analysis, synthesis, and algorithms," *IEEE Trans. Info. Theory*, vol. 44, no. 1, pp. 16–31, Jan. 1998.

- Boufounos P. and Oppenheim A.V., "Quantization noise shaping on arbitrary frame expansions," *EURASIP Journal on Applied Signal Processing, Special issue on Frames and Overcomplete Representations in Signal Processing, Communications, and Information Theory*, vol. 2006, pp. Article ID 53 807, 12 pages, DOI:10.1155/ASP/2006/53 807, 2006.

- Boufounos P. T., "Quantization and erasures in frame representations," *MIT D.Sc. Thesis*, Cambridge, MA, January 2006.

- Benedetto J. J., Powell A. M., and Yilmaz O., "Sigma-Delta quantization and finite frames," *IEEE Trans. Info. Theory*, vol. 52, no. 5, pp. 1990–2005, May 2006.

- Deift, P., Krahmer, F. and Güntürk, C. S. (2011), "An optimal family of exponentially accurate one-bit Sigma-Delta quantization schemes." *Comm. Pure Appl. Math.*, 64: 883–919. doi: 10.1002/cpa.20367

Today’s Topics

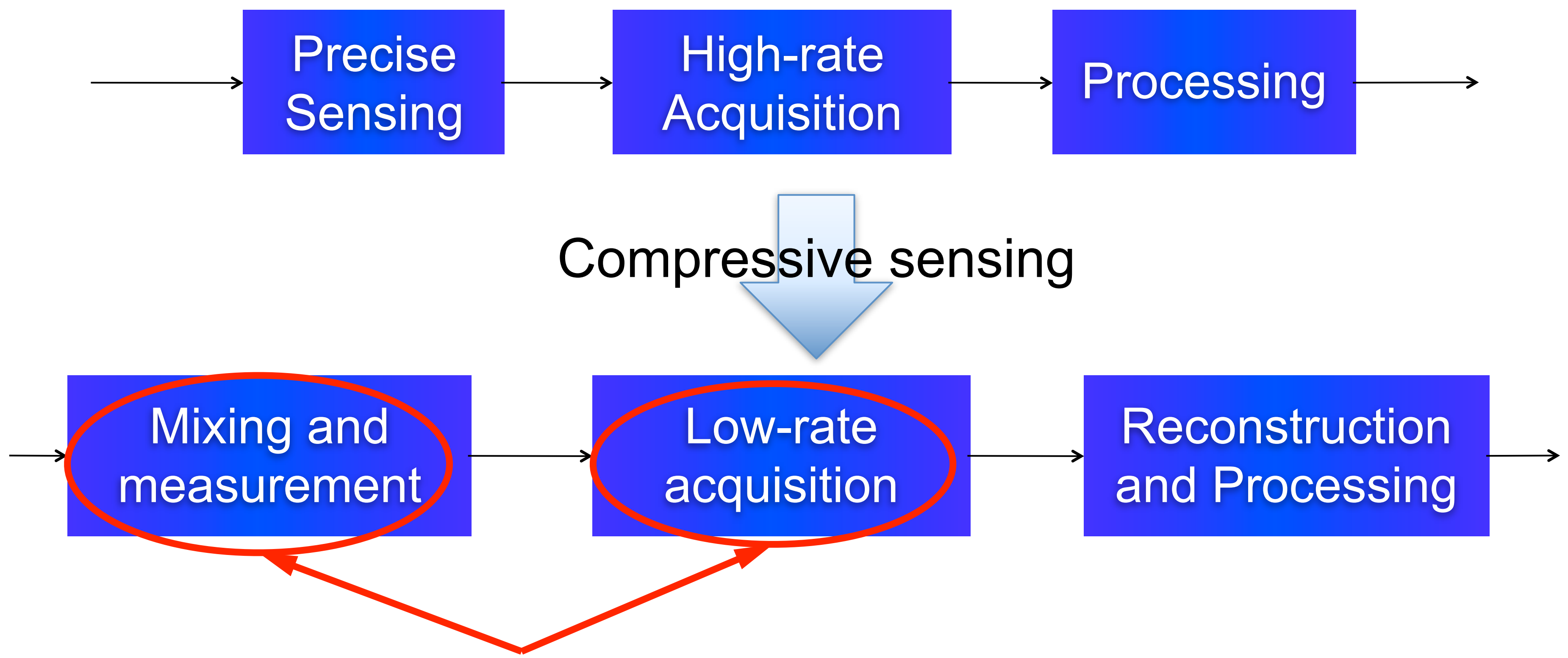
1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

Today's Topics

1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

- Candès, E., Romberg, J., and Tao, T., "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
- Donoho D. , "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Sept. 2006.

Sensing Pipeline Paradigm Change

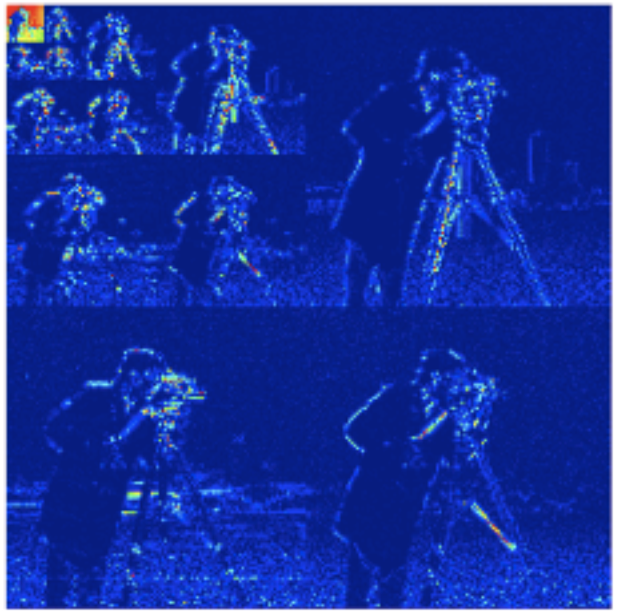


Goal: exploit mixing to **simplify sensor or improve sensor** specifications (e.g., sensor speed, A/D conversion rate, measured bandwidth/resolution)

- Compressive sensing has significantly improved our sensing capability
- Two fundamental Compressive Sensing research aspects
  - Hardware modifications for efficient acquisition
  - Signal/scene models and processing algorithms

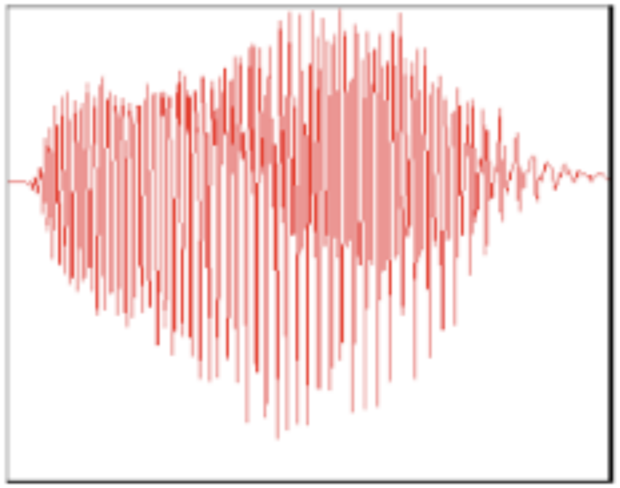
Signal Structure: Sparsity

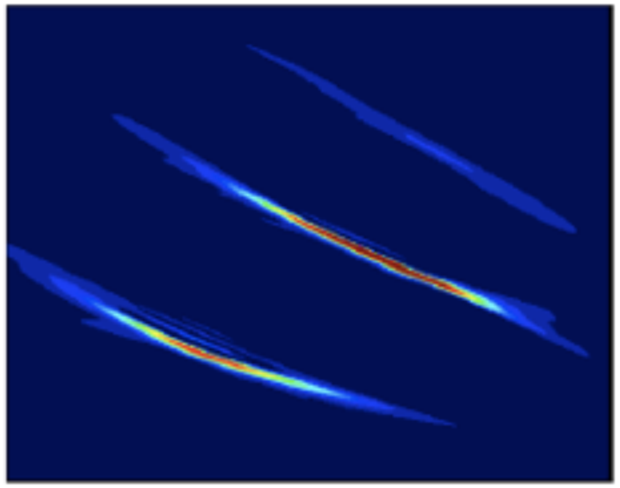
N
pixels



$K \ll N$
large
wavelet
coefficients

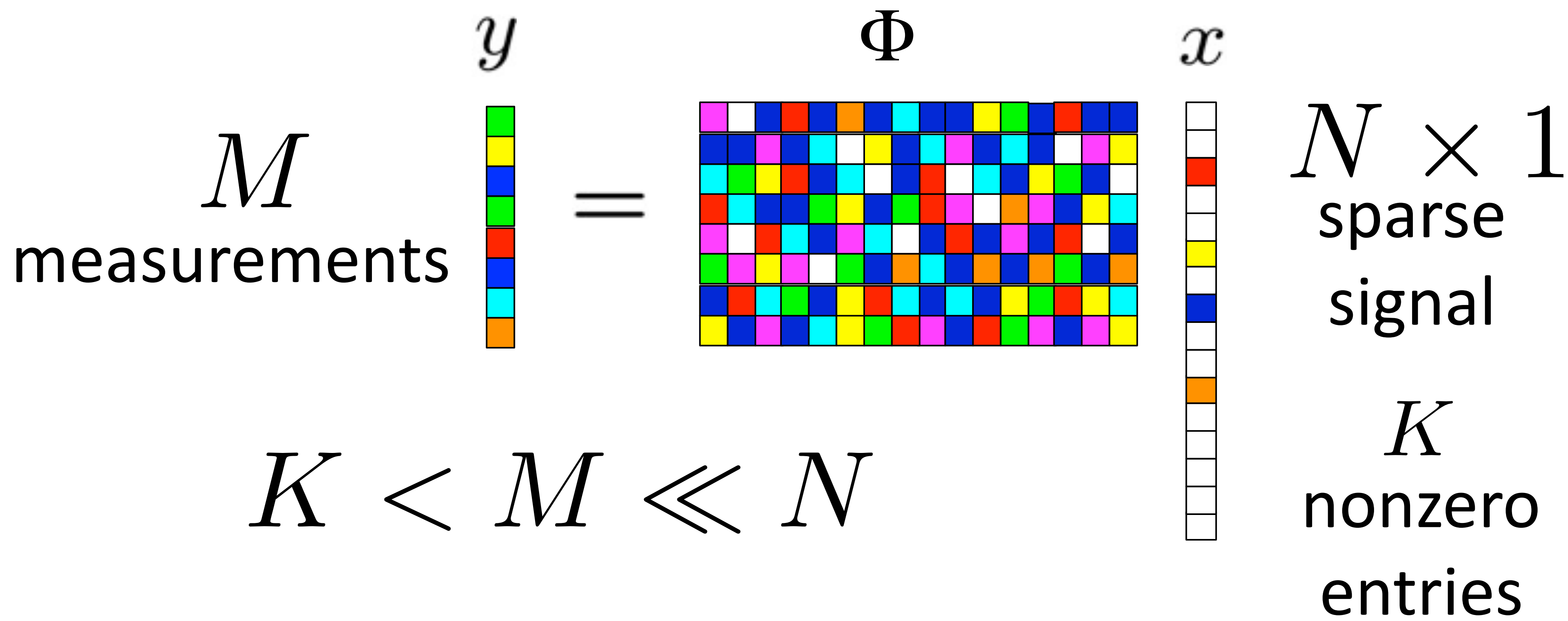
N
wideband
signal
samples





$K \ll N$
large
Gabor
coefficients

Measurement Model: Incoherence [Candes et al]



- x is K-sparse or K-compressible
- Φ **random**, satisfies a *restricted isometry property (RIP)*

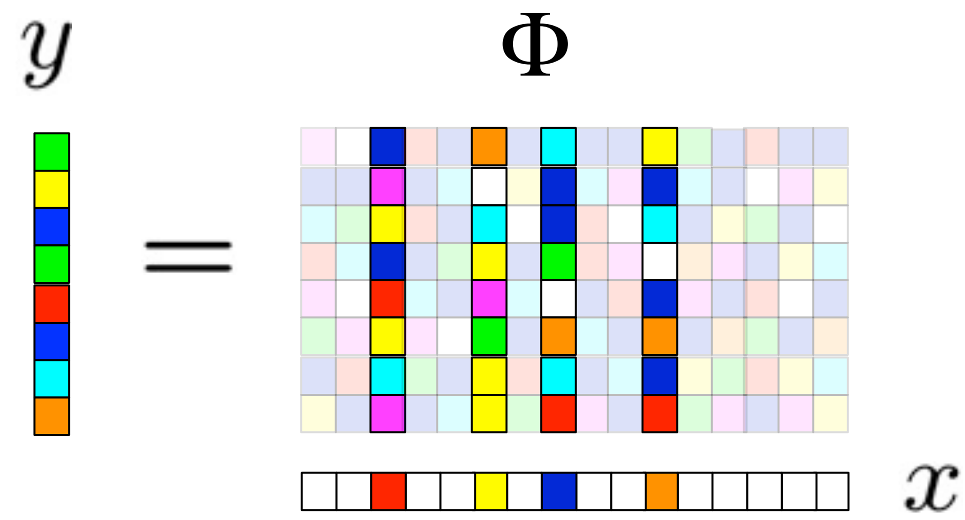
  Φ has RIP of order $2K$ with constant δ

  If there exists δ s.t. for all 2K-sparse x:

  $$(1-\delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1+\delta)\|x\|_2^2$$

- M=O(Klog$N$/K)
- Φ also has small *coherence* $\mu \triangleq \max_{i \neq j} |\langle \phi_i, \phi_j \rangle|$

Measurement Model: Incoherence [Candes et al]



- x is K-sparse or K-compressible
- Φ **random**, satisfies a *restricted isometry property (RIP)*

Φ has RIP of order $2K$ with constant δ

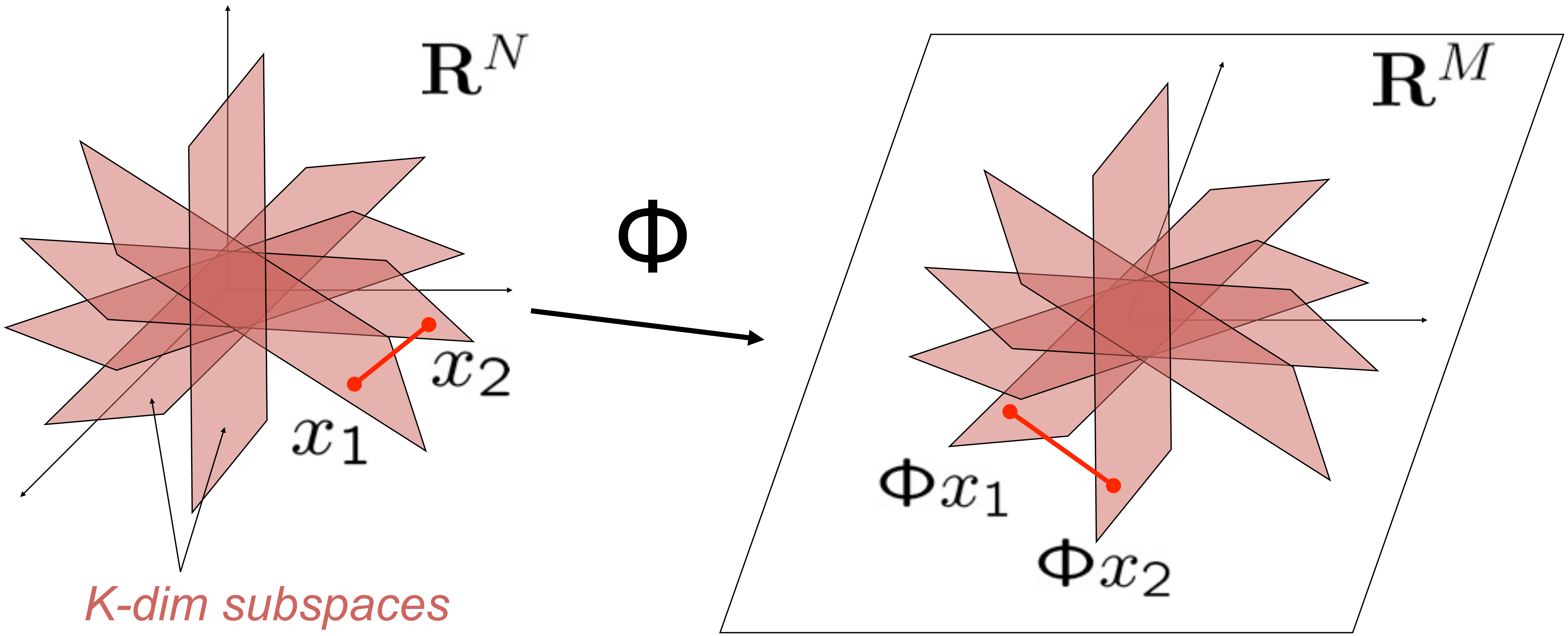
If there exists δ s.t. for all $2K$-sparse x:

$$(1-\delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1+\delta)\|x\|_2^2$$

- $M = O(K \log N/K)$
- Φ also has small *coherence* $\mu \triangleq \max_{i \neq j} |\langle \phi_i, \phi_j \rangle|$

RIP/Stable Embedding

- An information preserving projection $\mathbf{A}$ preserves the **geometry** of the set of sparse signals



Restricted Isometry Property

$$(1-\delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1+\delta)\|x\|_2^2$$

Reconstruction: Non-linear, Enforcing Structure

- Reconstruction using **sparse approximation**:
  - Find sparsest $\mathbf{x}$ such that $\mathbf{y} \approx \mathbf{A}\mathbf{x}$

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{y} \approx \mathbf{\Phi}\mathbf{x}$$

- **Convex optimization** approach:
  - Minimize l_1 norm: e.g.,

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } \mathbf{y} \approx \mathbf{\Phi}\mathbf{x}$$

- **Greedy algorithms** approach:
  - Minimize $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2$ such that $\mathbf{x}$ is sparse

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 \text{ s.t. } \|\mathbf{x}\|_0 \leq K$$

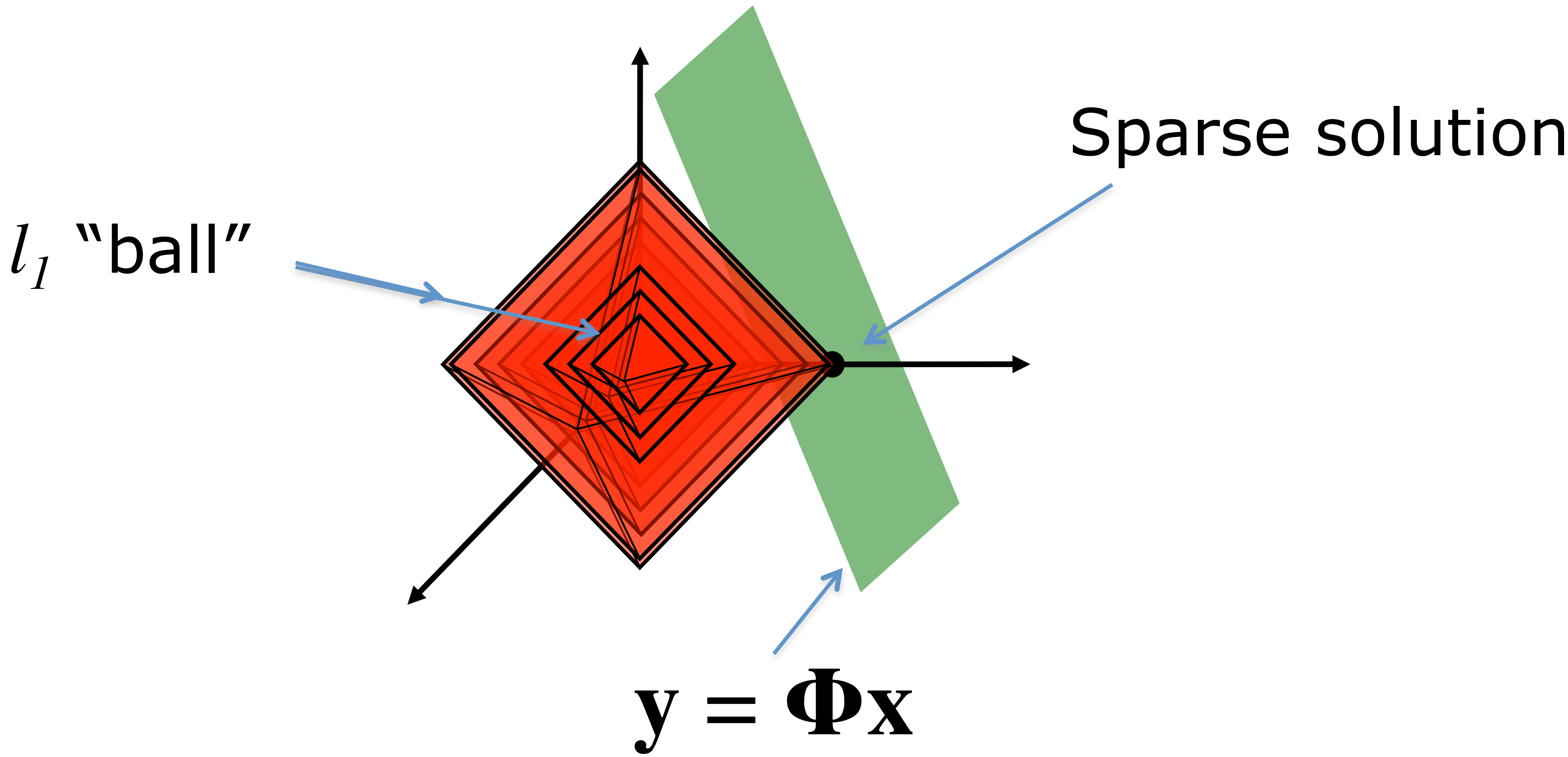
  - MP, OMP, ROMP, StOMP, CoSaMP, SP, ALPS, PYAMP (Pick Your Acronym Matching Pursuit)

- **More general cost functions,**
  - GraSP, generalization of CoSaMP

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} f(\mathbf{x}) \text{ s.t. } \|\mathbf{x}\|_0 \leq K$$

Why l_1 relaxation works

$$\min \ \|\mathbf{x}\|_1 \ \text{s.t.} \ \mathbf{y} \approx \mathbf{\Phi x}$$

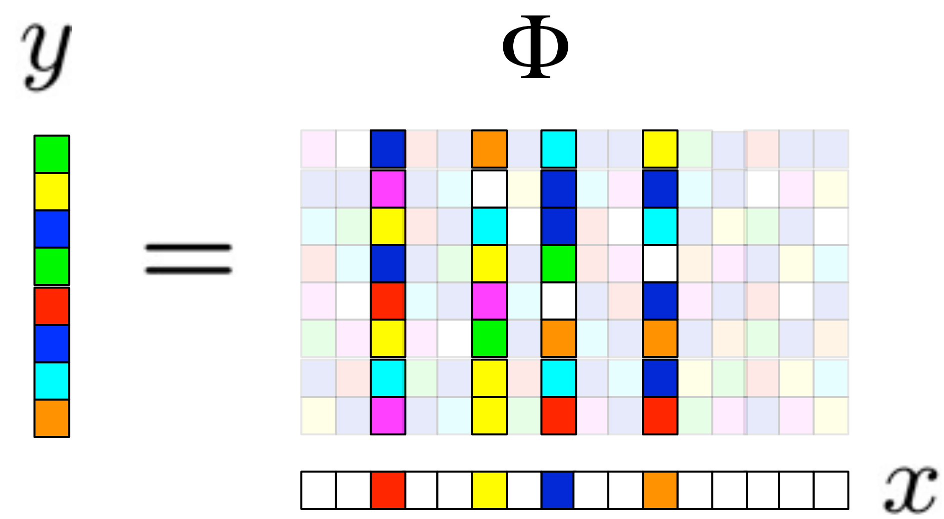


K-term approximation error

If Φ satisfies the RIP: $\|\widehat{\mathbf{x}} - \mathbf{x}\|_2 \leq c_1 \frac{\|\mathbf{x} - \mathbf{x}_K\|_1}{\sqrt{K}} + c_2 \epsilon$

Measurement error

Greedy Pursuits Core Idea



- $\boldsymbol{y}$ highly correlated with $\boldsymbol{\Phi}$ at locations where $\boldsymbol{x}$ is high
- $\boldsymbol{\Phi}^T\boldsymbol{y}$ provides a good idea of these locations
  - This is why low coherence is important

$$\mu \triangleq \max_{i \neq j} \left| \langle \phi_i, \phi_j \rangle \right|$$

  - $\boldsymbol{\Phi}^T\boldsymbol{y}$ referred to as *proxy* for $\boldsymbol{x}$
- General Strategy:
  - Identify locations
  - Invert the system only on those locations

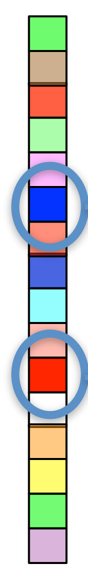
GraSP (Gradient Subspace Pursuit)

State Variables: Signal estimate, $\hat{\mathbf{x}}$ support estimate: T

Initialize estimate and support: $\hat{\mathbf{x}}=0$, $T=\text{supp}(\hat{\mathbf{x}})$

Compute Gradient at Current Estimate

$$\nabla f(\widehat{\mathbf{x}}) = \mathbf{g}$$



Select location of largest *2K* gradient directions

$$\text{supp}(\mathbf{g}|_{2K})$$

Add to support set

$$\Omega = \text{supp}(\mathbf{g}|_{2K}) \cup T$$

Minimize over support

$$\mathbf{b} = \arg\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t. } \mathbf{x}_{\Omega^c} = 0$$

Truncate result

$$\widehat{\mathbf{x}} = \mathbf{b}|_K$$

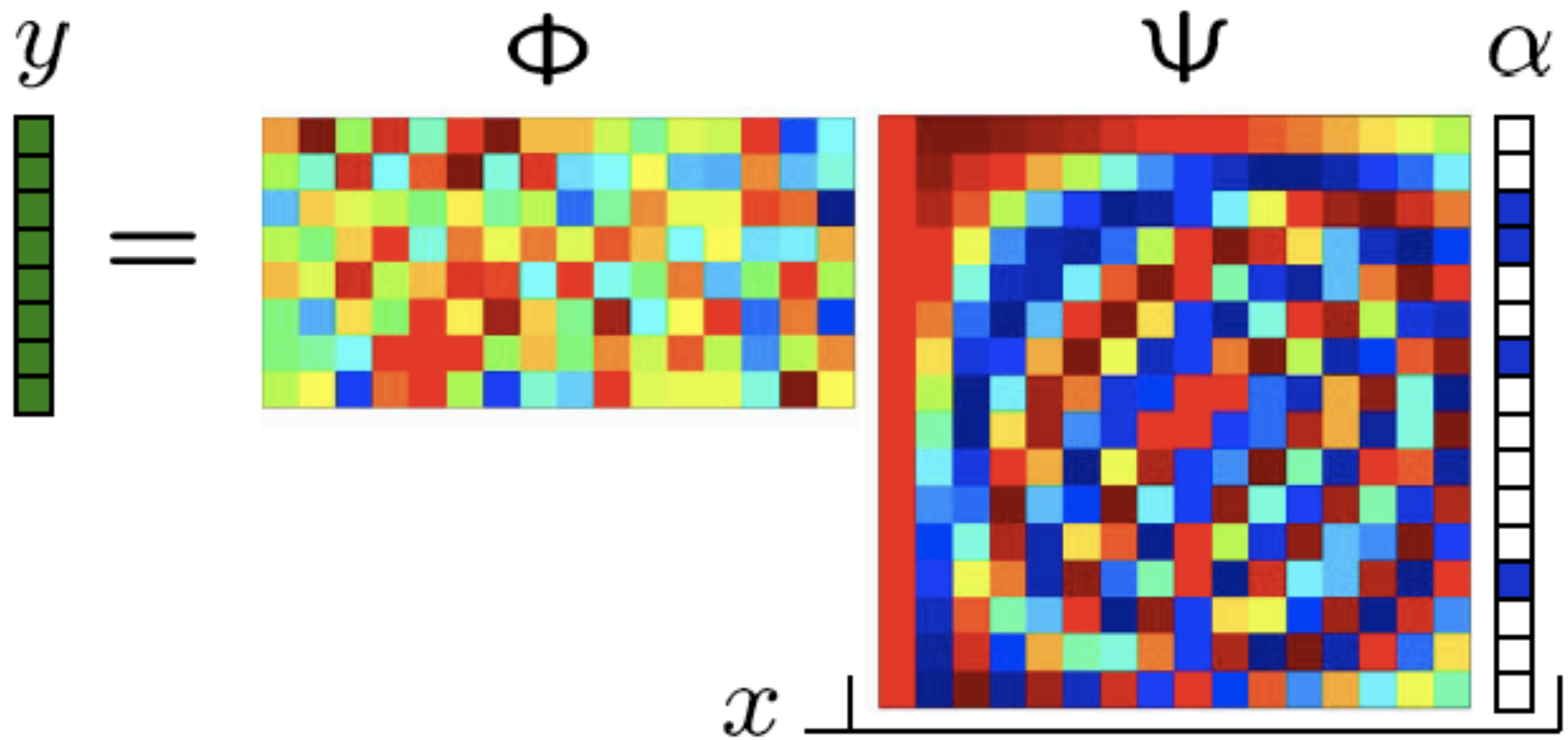
$$T = \text{supp}\left(\mathbf{b}|_K\right)$$

Iterate using residual

- S. Bahmani, B. Raj, and P. T. Boufounos, "Greedy Sparsity-Constrained Optimization," *Journal of Machine Learning Research*, v. 14, pp. 807-841, March, 2013.

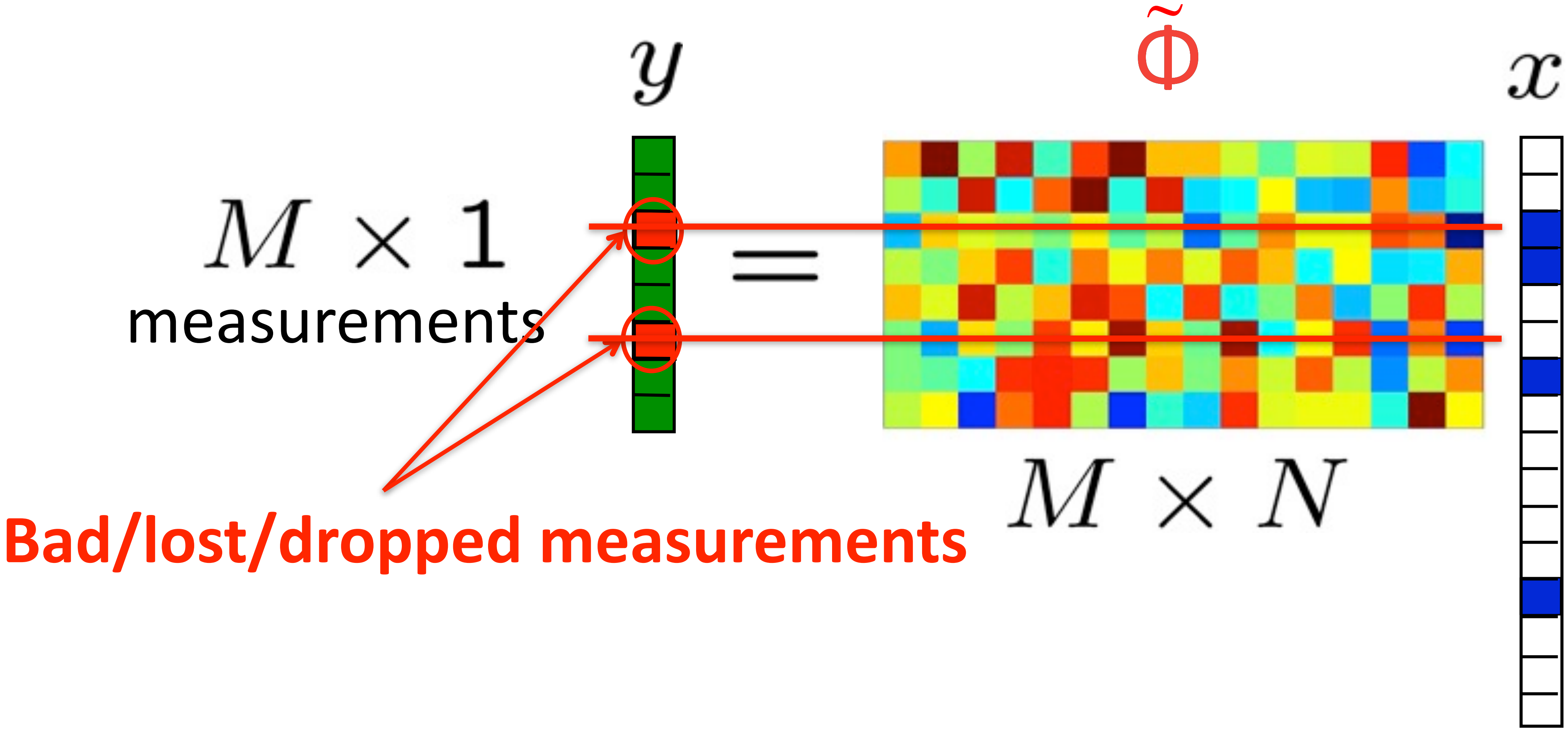
Universality

- Gaussian white noise basis is incoherent with ***any*** fixed orthonormal basis (with high probability)
- Signal sparse in frequency domain: $\Psi = \mathrm{idct}$



- Product $\Phi\Psi$ remains Gaussian white noise

Democracy



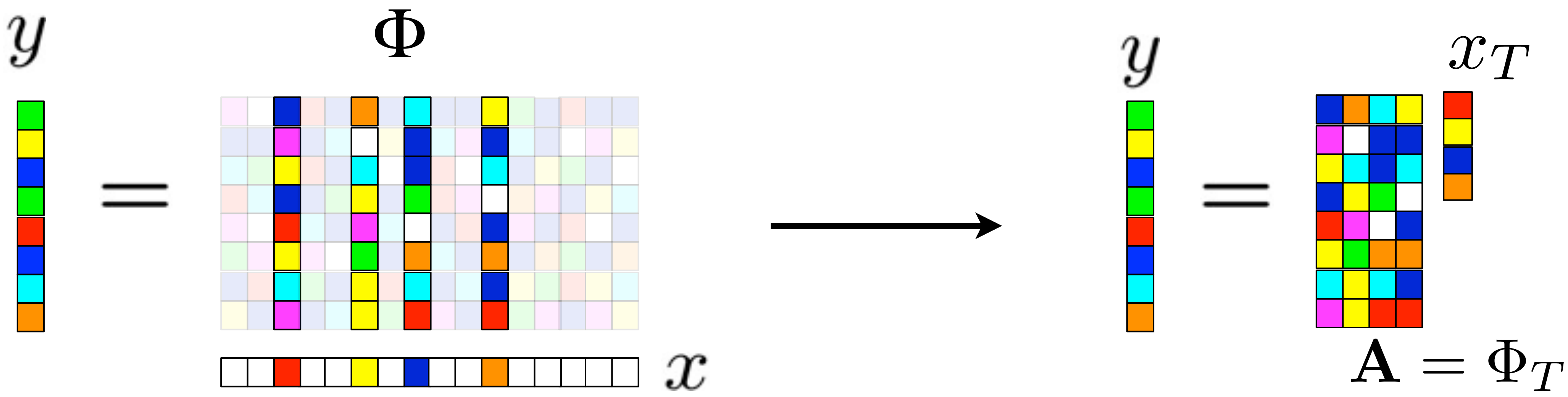
- Measurements are **democratic** [Davenport, Laska, Boufounos, Baraniuk]
  - They are all equally important
  - We can **loose** some **arbitrarily**
    (i.e. an adversary can choose which ones)

- The $\tilde{\Phi}$ still satisfies RIP (as long as we don't drop too many)

- M. A. Davenport, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "A simple proof that random matrices are democratic," *Rice University ECE Department Technical Report TREE-0906*, Houston, TX, November, 2009.

Compressive Sensing and Oversampling

Given support of signal T



Resulting system is oversampled: $\mathbf{A} \in \mathbb{R}^{M \times K}$

$$M = O(K \log N)$$

$$\Rightarrow r = O(\log N)$$

Oversampling Rate

Oversampling provides **robustness**, but introduces **inefficiency**

- Boufounos P., Baraniuk R. G., "Quantization of Sparse Representations." *Rice University ECE Department Technical Report 0701*. Summary appears in *Proc. of the Data Compression Conference (DCC '07)*, March 27-29 2007, Snowbird, UT.

Further Reading

- Candès, E., Romberg, J., and Tao, T., “Stable signal recovery from incomplete and inaccurate measurements,” *Comm. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
- Donoho D. , “Compressed sensing,” *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Sept. 2006.
- Candès, Emmanuel J. “Compressive sampling.” *Proceedings oh the International Congress of Mathematicians: invited lectures,* August 22-30, 2006, Madrid, Spain.
- Candes, E.J.; Wakin, M.B., "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol.25, no. 2, pp.21,30, March 2008
- M. A. Davenport, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, “A simple proof that random matrices are democratic,” *Rice University ECE Department Technical Report TREE-0906*, Houston, TX, November, 2009.
- Needell D. and Tropp J.A., “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301-321, May 2009.
- S. Bahmani, B. Raj, and P. T. Boufounos, “Greedy Sparsity-Constrained Optimization,” *Journal of Machine Learning Research*, v. 14, pp. 807-841, March, 2013.

Today’s Topics

1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

Today's Topics

1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

- Candès, E., Romberg, J., and Tao, T., "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
- Donoho D. , "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Sept. 2006.

Part III:

When quantization meets compressed sensing

Laurent Jacques, UCL, Belgium
Petros Boufounos, MERL, USA

Outline:

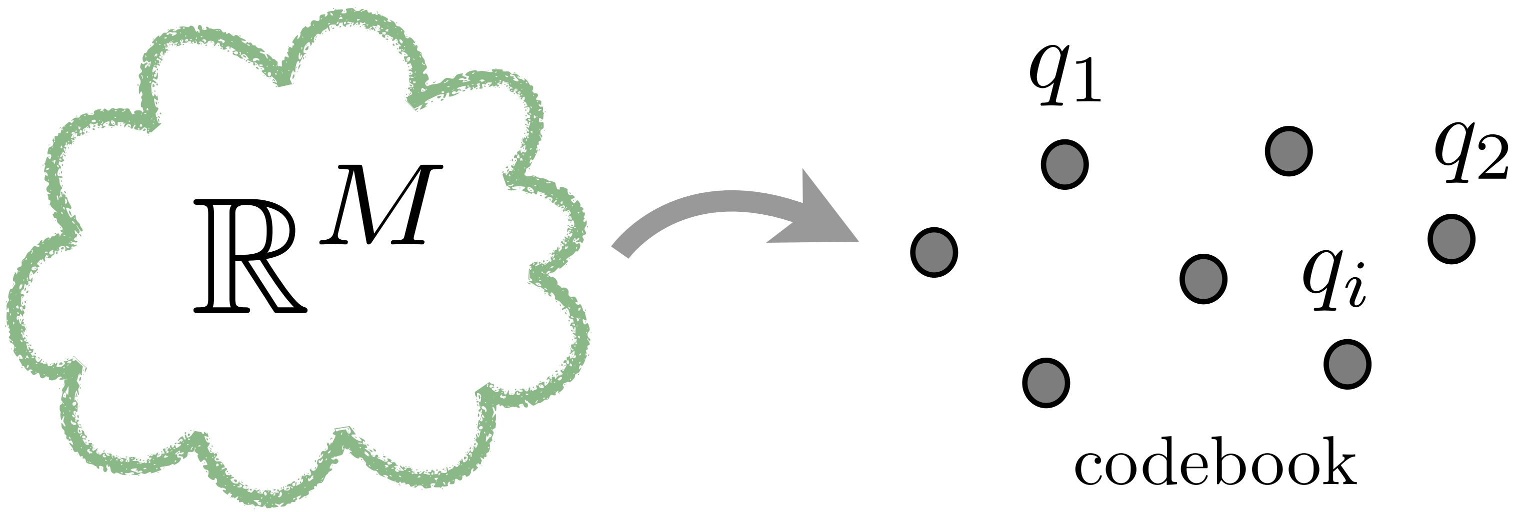
1. Context
2. Former QCS methods and performance limits
3. Consistent Reconstructions
4. Sigma-Delta quantization in CS
5. To saturate or not? And how much?

1. Context

What is quantization?

- <u>Generality</u>:

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



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

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

Example 1: scalar quantization

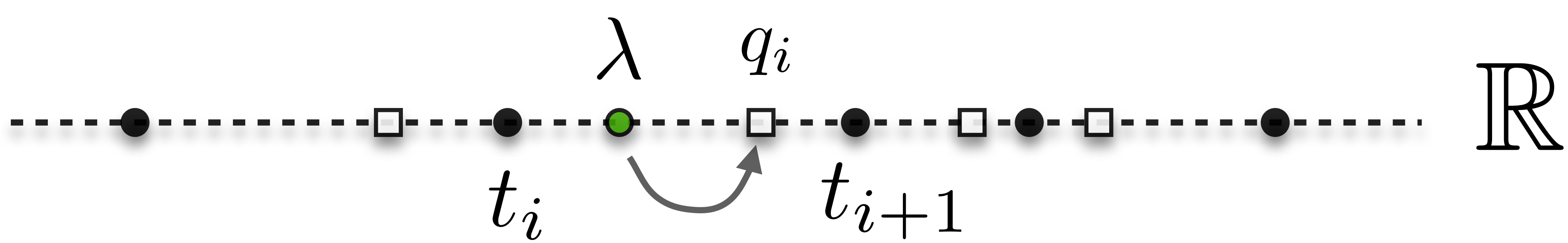
- In $\mathbb{R}^M$, on each component of M- dimensional vectors:

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

$$\mathcal{T} = \{t_i \in \overline{\mathbb{R}} : 1 \leqslant i \leqslant 2^B + 1, t_i \leqslant t_{i+1}\} \quad \text{(thresholds)} \quad \bullet$$

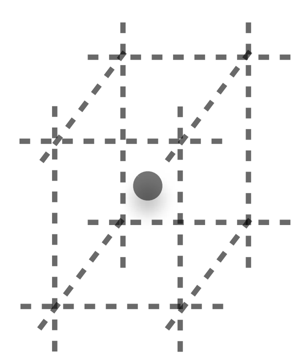
$$\forall \lambda \in \mathbb{R}, \qquad \mathcal{Q}[\lambda] = q_i \;\Leftrightarrow\; \lambda \in \mathcal{R}_i \triangleq [t_i, t_{i+1}), \;\; \text{1-D quantization cell}$$

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



- Globally:

$$\mathcal{Q}[\boldsymbol{z}] = \boldsymbol{q} \in \Omega^M \;\Leftrightarrow\; \boldsymbol{z} \in$$



M - D quantization cell

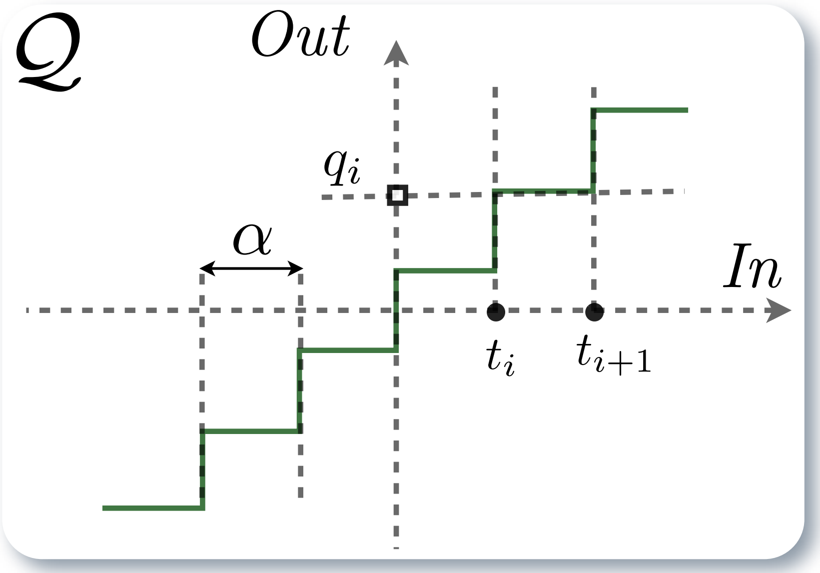
$$\mathcal{R}_{i_1} \times \mathcal{R}_{i_2} \times \cdots \times \mathcal{R}_{i_M}$$

$$:= \mathcal{Q}^{-1}[\boldsymbol{q}]$$

Example 1: scalar quantization

- Regular uniform

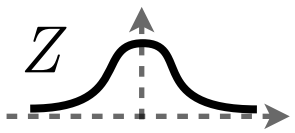
$$q_k = (k + 1/2)\alpha$$
$$t_k = k\alpha$$



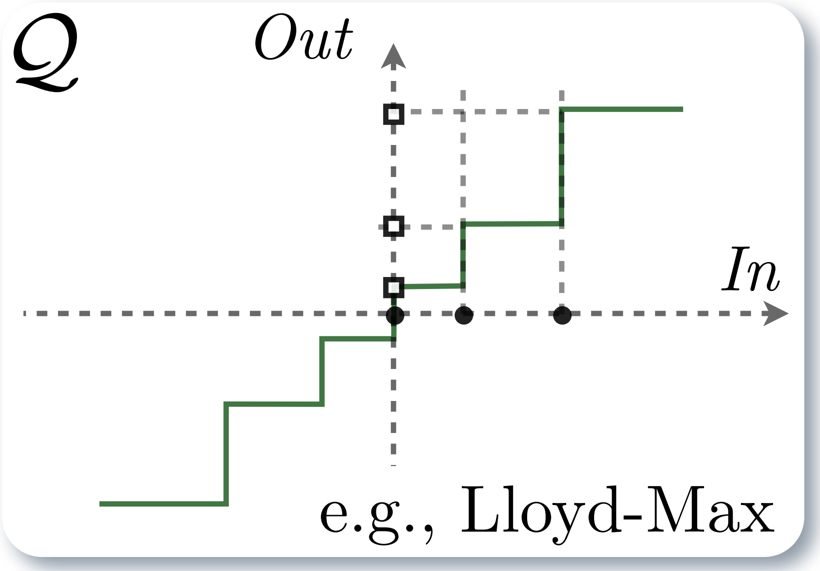
- Regular non-uniform

Ω and $\mathcal{T}$ optimized

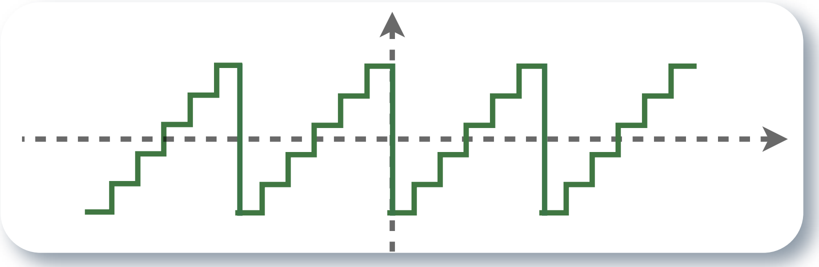
e.g., wrt an input distribution Z
find minimum distortion, *i.e.*,



$$\underset{\mathcal{T},\Omega}{\text{argmin}}\ \mathbb{E}_Z\|Z - \mathcal{Q}[Z]\|^2$$



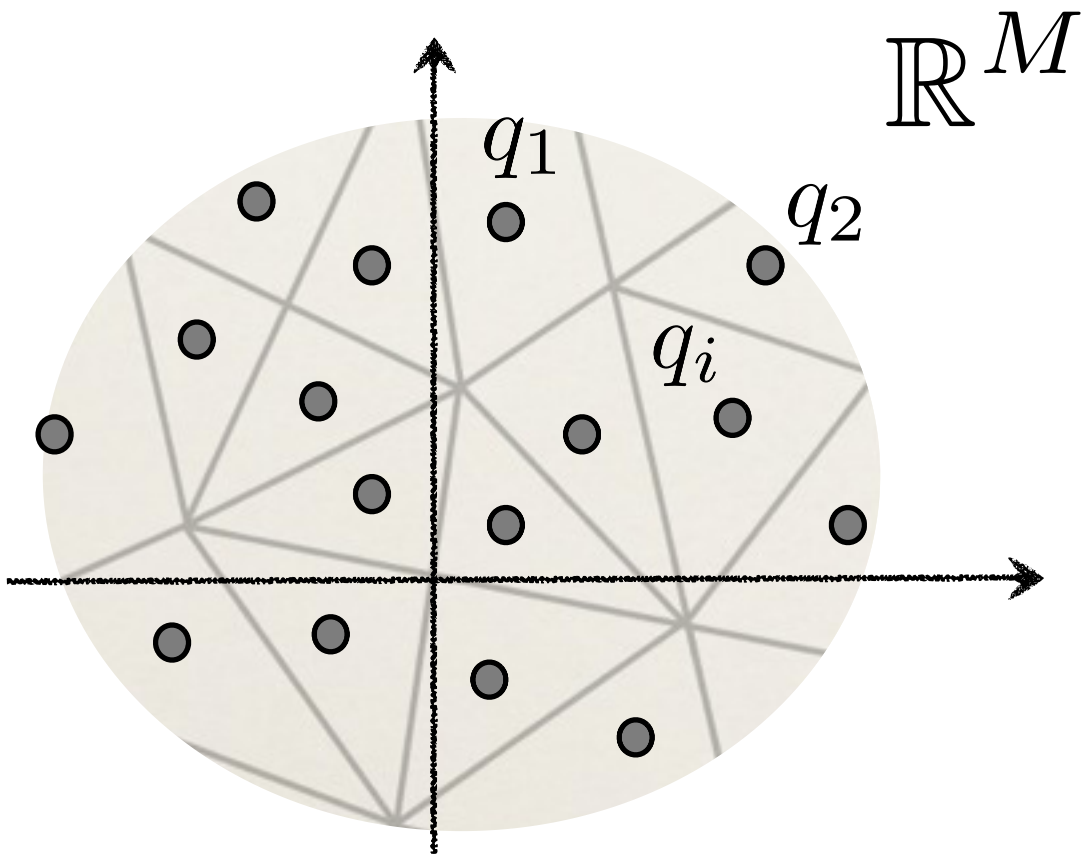
- Non-regular $\rightarrow$ Petros, Part V



Example 2: vector quantization

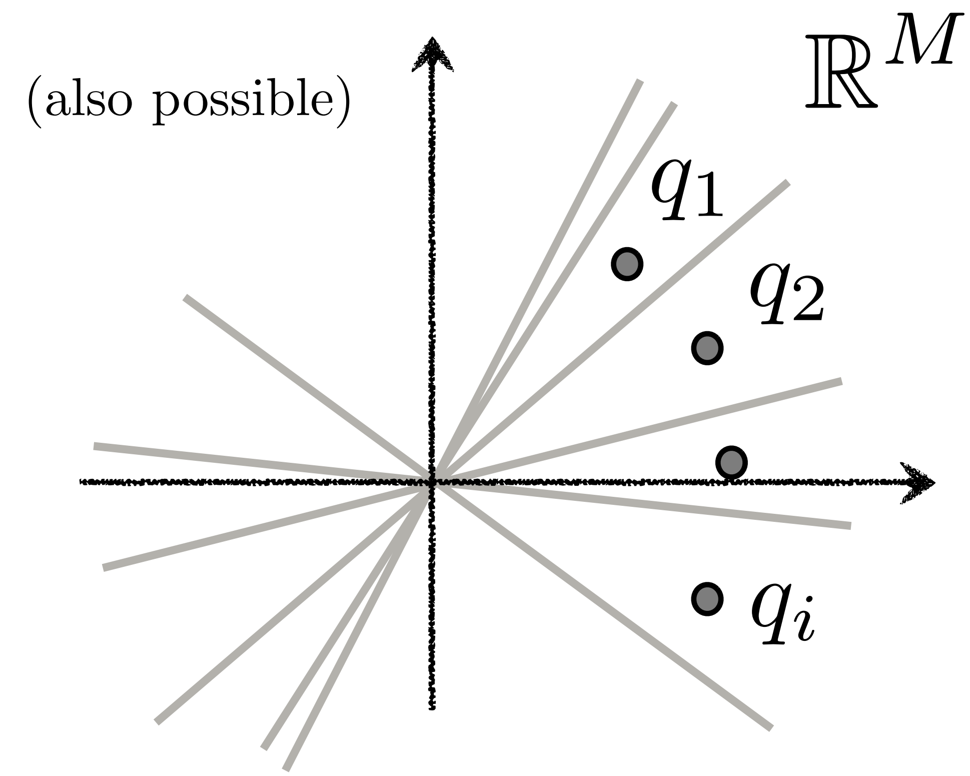
(caveat: not really covered in this tutorial, ... except $\Sigma\Delta$, see later)

$$\text{Quantization} = \text{codebook } \boldsymbol{\Omega} + \text{quantization cells } \boldsymbol{\mathcal{R}} = \{\mathcal{R}_i \subset \mathbb{R}^M\}$$



(non-separable quantization)

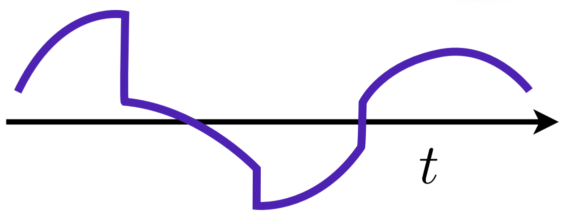
e.g., $\underset{\boldsymbol{\Omega},\boldsymbol{\mathcal{R}}}{\operatorname{argmin}}\ \mathbb{E}_Z \|\boldsymbol{Z} - \mathcal{Q}[\boldsymbol{Z}]\|^2$

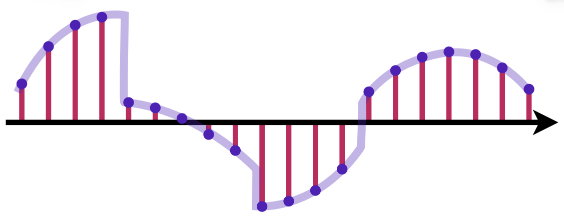


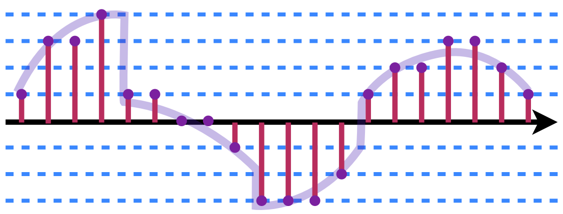
e.g., encoding components ordering + sign
(*permutation frame quantization*)
(Nguyen et al, Goyal, ...)

Classical Sampling and Quantization

signal $\xrightarrow{x(t)}$ Anti-alias (measurement) $\xrightarrow{T}$ $x[n]$ → Quantizer $\xrightarrow{q[n]}$ bits







<u>For reconstruction:</u> $\xrightarrow{q[n]}$ Low-pass Filter (linear reconstruction) $\xrightarrow{\widehat{x}(t)}$

Sampling: discretization in time

Lossless at the Nyquist rate

Quantization: discretization in amplitude

Always lossy

Need both for digital data acquisition

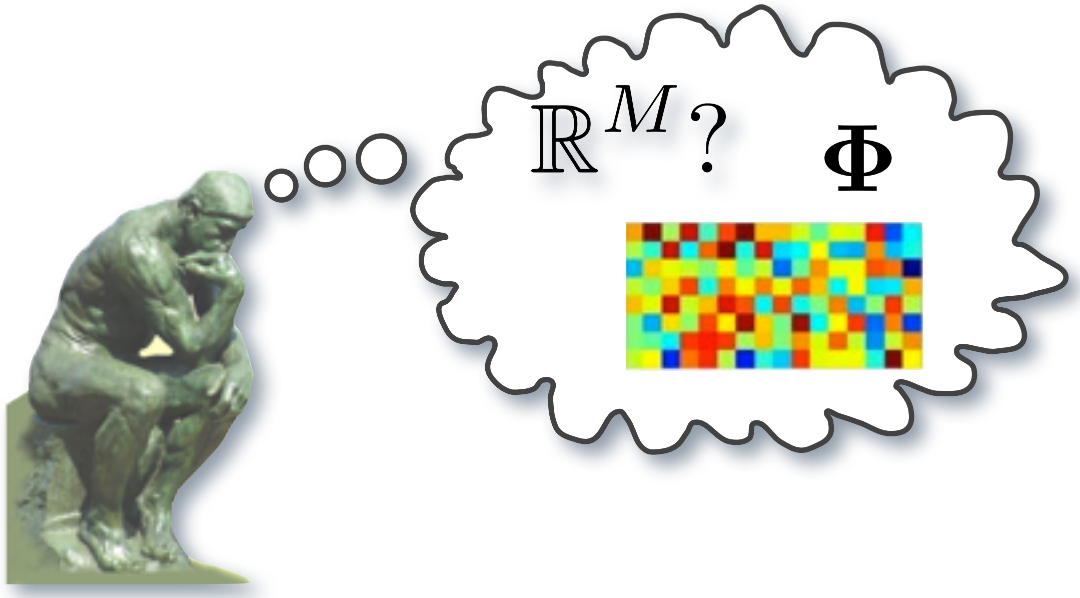
Compressive Sampling and Quantization

Compressed sensing theory says:

"Linearly sample a signal

at a rate function of

its intrinsic dimensionality"

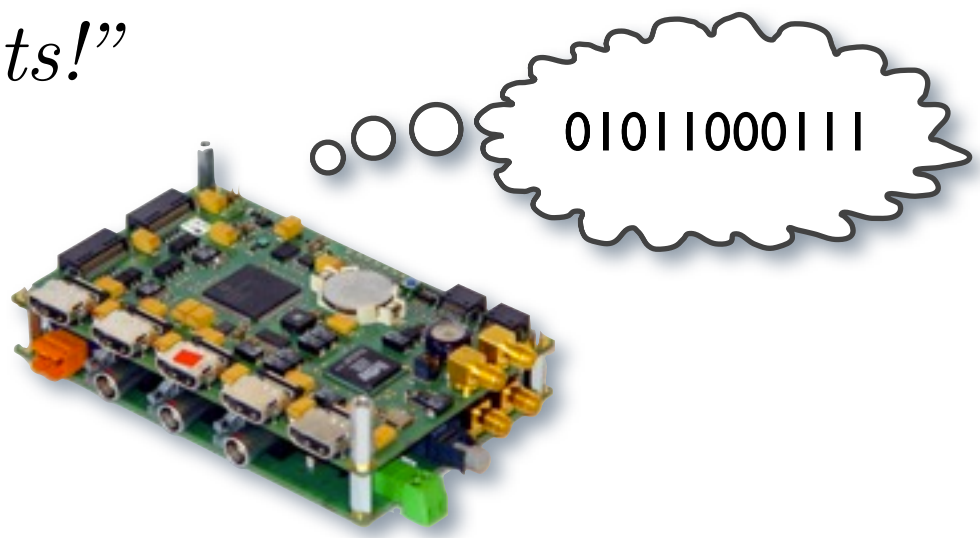


Information theory and sensor designer say:

"Okay, but I need to

quantize/digitize my measurements!"

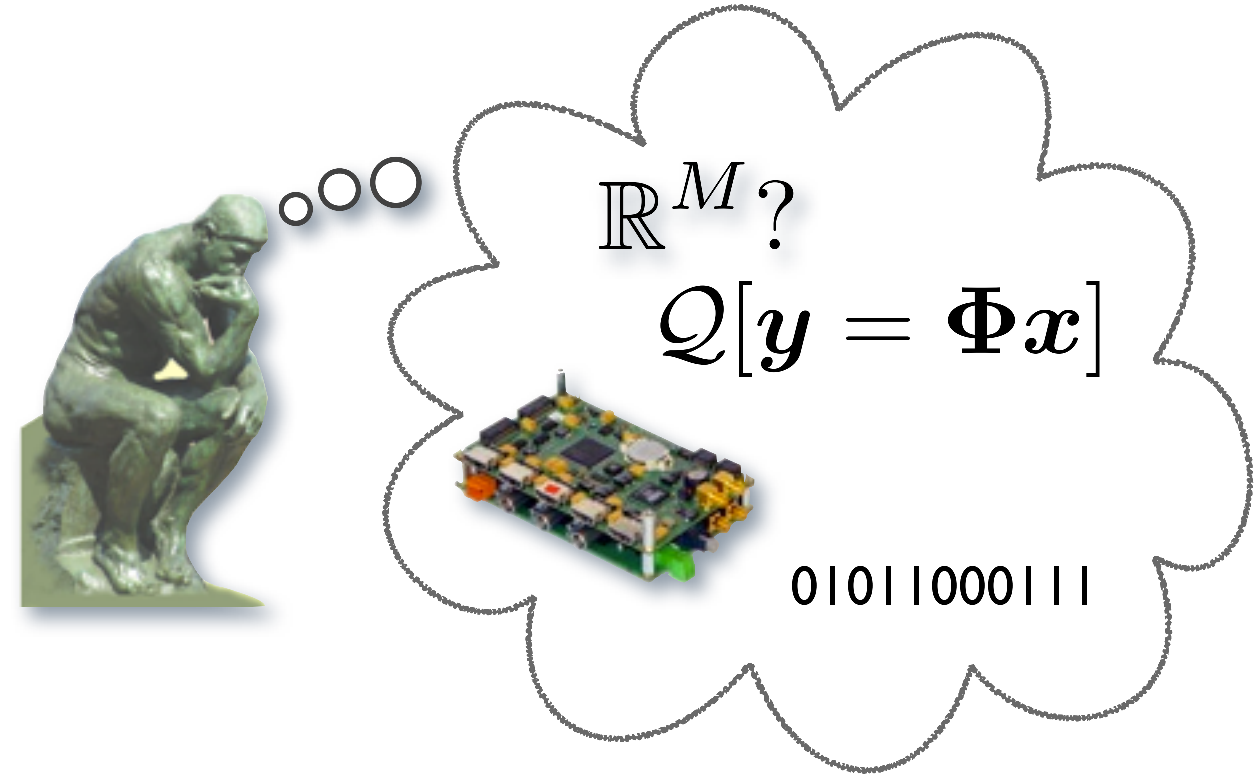
(e.g., in ADC)



The Quantized CS Problem (QCS)

Natural questions:

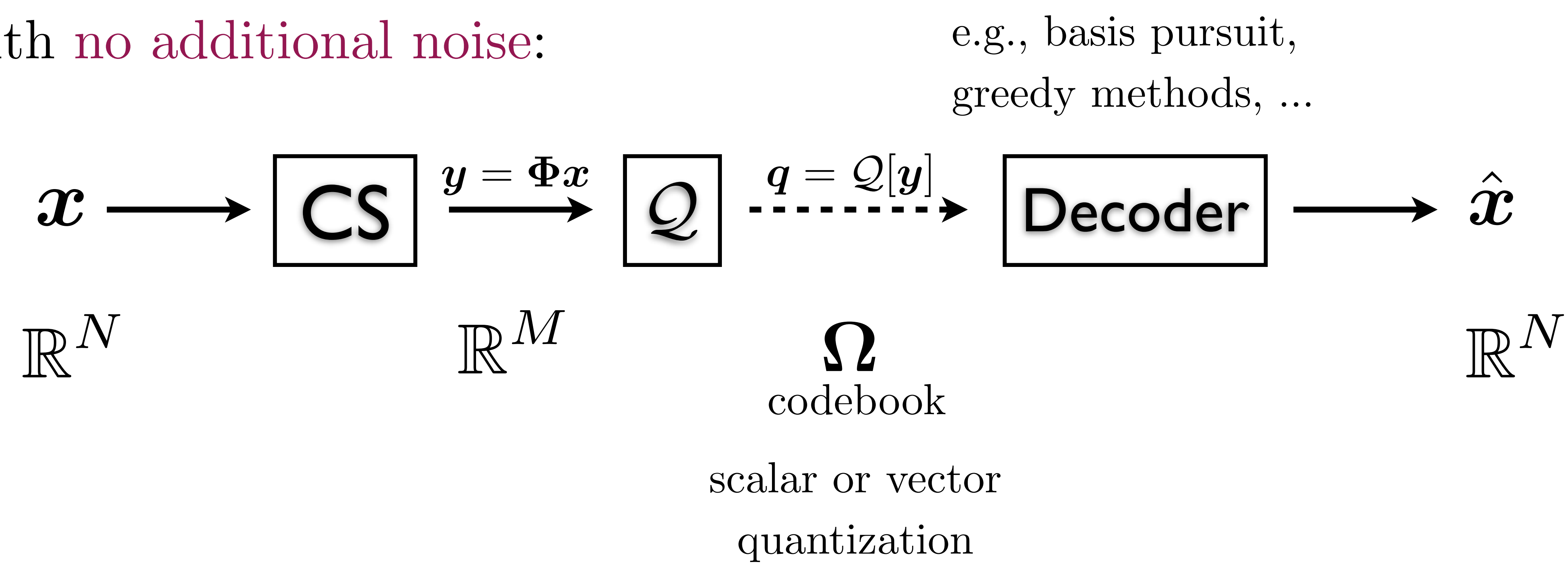
- How to integrate quantization in CS?



- What do we loose?
- Are they some theoretical limitations?
  (related to information theory? geometry?)
- How to minimize quantization effects in the reconstruction?

QCS: a system view

With no additional noise:



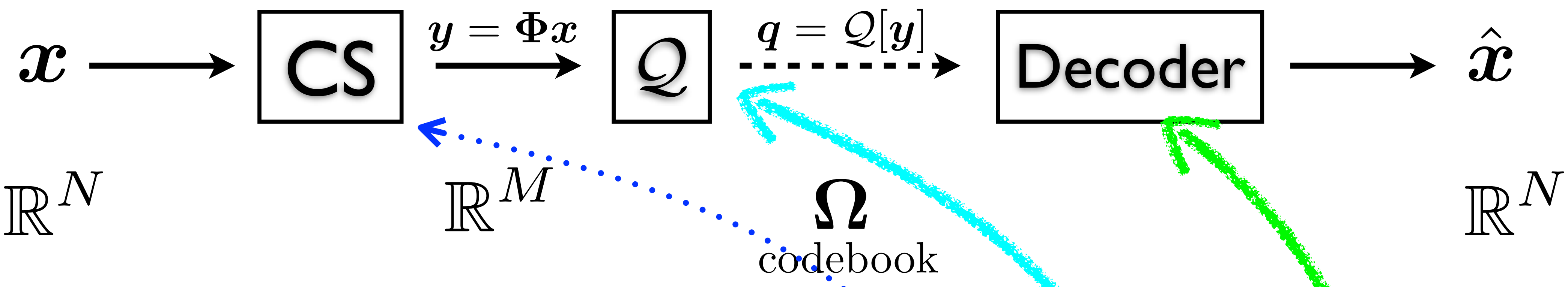
Finite codebook $\Rightarrow \hat{\boldsymbol{x}} \neq \boldsymbol{x}$

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

QCS: a system view

With no additional noise:

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



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

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

given a certain number of:
bits, measurements, or bits/meas.

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

2. Former QCS methods and performance limits

Scalar quantization in CS

Turning measurements into bits $\rightarrow$ scalar quantization

$$q_i = \mathcal{Q}[(\boldsymbol{\Phi x})_i] = \mathcal{Q}[\langle \boldsymbol{\phi}_i, \boldsymbol{x} \rangle] \ \in \Omega \subset \mathbb{R}$$
$$\boldsymbol{q} \ = \ \mathcal{Q}\big[\boldsymbol{\Phi x}\big] \ \in \boldsymbol{\Omega} = \Omega^M,$$

Important points:

- Definition of $\boldsymbol{\Phi}$ independent of M (e.g., $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$)
  $\rightarrow$ preserves measurement dynamic!
- B bits per measurement
- Total bit budget: $R = BM$
- No further encoding (e.g., entropic)

Former solution (Candès, Tao, ...)

- Quantization is like a noise

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

quantization distortion

Former solution (Candès, Tao, ...)

- Quantization is like a noise

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

and 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

How to find it?

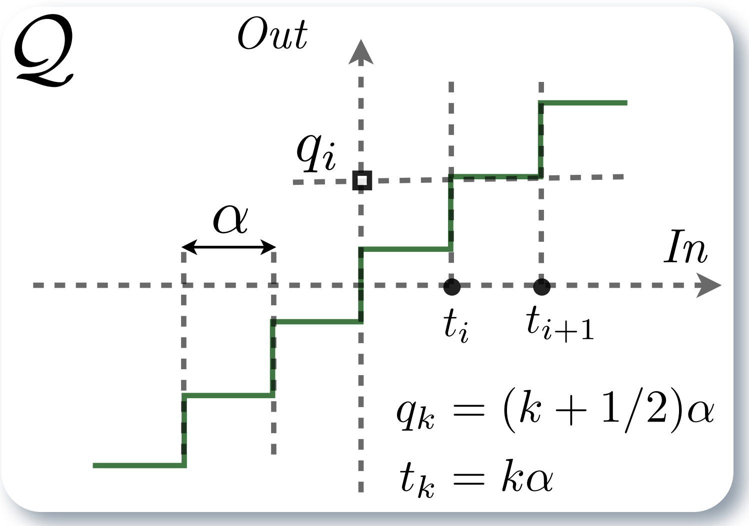
$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \;\leqslant\; C\frac{\epsilon}{\sqrt{M}} + D\, e_0(K),$$

for some $C, D > 0$ and $e_0(K) = \|\boldsymbol{x} - \boldsymbol{x}_K\|_1/\sqrt{K}$.

Former solution (Candès, Tao, ...)

1. For uniform quantization, by construction:

ϵ?



$$
\begin{aligned}
n_i &= \mathcal{Q}[(\boldsymbol{\Phi x})_i] - (\boldsymbol{\Phi x})_i \\
&\in q_{k_i} - \mathcal{R}_{k_i} = [-\alpha/2, \alpha/2] \\
&\Rightarrow \|\boldsymbol{n}\|_\infty \le \alpha/2
\end{aligned}
$$

$$\Rightarrow \|\boldsymbol{n}\|^2 \leqslant M\|\boldsymbol{n}\|_\infty^2 \leqslant M\alpha^2/4$$

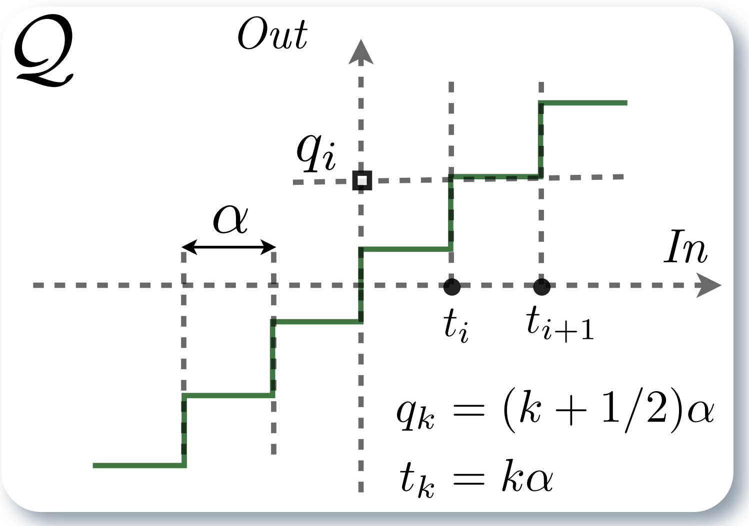
and plug this upper bound in BPDN

can be improved!

Former solution (Candès, Tao, ...)

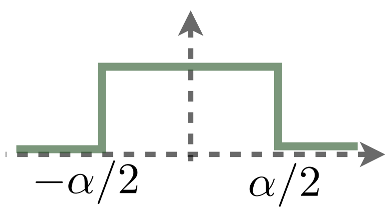
2. For uniform quantization, uniform model!

ϵ?



$$n_i = \mathcal{Q}[(\boldsymbol{\Phi x})_i] - (\boldsymbol{\Phi x})_i$$
$$\in q_{k_i} - \mathcal{R}_{k_i} = [-\alpha/2, \alpha/2]$$
$$\sim_{\text{iid}} \text{Uniform}([-\alpha/2, \alpha/2])$$

(HRA - high resolution assumption)



$$\Rightarrow \mathbb{E}|n_i|^2 = \alpha^2/12$$

$$\Rightarrow \|\boldsymbol{n}\|^2 \leqslant \mathbb{E}\|\boldsymbol{n}\|^2 + \kappa\sqrt{\text{Var}\|\boldsymbol{n}\|^2}$$ (Chernoff-Hoeffding, bounded RVs)

$$\boxed{\leqslant M\frac{\alpha^2}{12} + \kappa\sqrt{M}\frac{\alpha^2}{6\sqrt{5}} = \epsilon_2^2 \simeq M\frac{\alpha^2}{12}}$$

with $\Pr > 1 - e^{-2\kappa^2}$

and plug this upper bound in BPDN

Former solution (Candès, Tao, ...)

- Therefore, from BPDN $\ell_2 - \ell_1$ instance optimality:

$$\boxed{\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \lesssim C\,\alpha + D\,e_0(K),}$$

for $C,D > 0$

(for BPDN with ϵ_2, under prev. cond.)

- Assuming :
  - bounded dynamics: $\|\boldsymbol{\Phi x}\|_\infty = \max_j |(\boldsymbol{\Phi x})_i| \leqslant \rho$ (e.g., by discarding saturation) (see later)
  - B bits per measurements $\Rightarrow \alpha \simeq \rho 2^{1-B}$

$$\boxed{\Rightarrow \text{BPDN RMSE} \lesssim C'\ 2^{-B} + D\,e_0(K)}$$

for $C',D > 0$

as soon as RIP holds: $M = O(K \log N/K)$

- Equivalently: BPDN RMSE $\simeq O(2^{-R/M}) + e_0(K)$

  for a rate $R = BM$ bits (total "bid budget" for all meas.)

RMSE Lower bound?

- Let a fixed K-sparse $\boldsymbol{x} \in \mathbb{R}^N$
- Oracle: you know $T = \operatorname{supp} \boldsymbol{x}$
- Noisy measurements (random noise):

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

$$\boldsymbol{y} = \boldsymbol{\Phi}_T \boldsymbol{x} + \boldsymbol{n}, \quad \text{with} \quad \mathbb{E}\, \boldsymbol{n}\boldsymbol{n}^T = \sigma^2 \mathbf{Id}_{M \times M}$$

- Assume: $\frac{1}{\sqrt{M}} \boldsymbol{\Phi}$ is $\mathrm{RIP}(K, \delta_K)$ and $\mathrm{RIP}(1, \delta_1)$
- Compute LS solution: $\hat{\boldsymbol{x}}_T = \underbrace{\boldsymbol{\Phi}_T^\dagger}_{\text{pseudo-inverse}} \boldsymbol{y} = (\boldsymbol{\Phi}_T^* \boldsymbol{\Phi}_T)^{-1} \boldsymbol{\Phi}_T^* \boldsymbol{y}$, $\quad \hat{\boldsymbol{x}}_{T^c} = 0$
- Then: $\boxed{\mathrm{MSE} = \ \mathbb{E}_{\boldsymbol{n}} \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 \geqslant r^{-1}\, \sigma^2 \left(\frac{1-\delta_1}{1+\delta_K}\right)}$

(as for BPDN) & $\mathrm{MSE} \leqslant \frac{1}{1-\delta_K} \sigma^2$ from [Needell, Tropp, 08]

for oversampling factor $r = M/K$

- for QCS: $\boxed{\Rightarrow \mathrm{RMSE} = \ \Omega(r^{-1/2}\, 2^{-B})}$ & $\mathrm{RMSE} = O(2^{-B})$

3. Consistent Reconstructions

Consistent reconstructions in CS?

- Problem in previous case: if $\hat{\boldsymbol{x}}$ solution of BPDN,
  - no Quantization Consistency (QC): $\mathcal{Q}[\boldsymbol{\Phi}\hat{\boldsymbol{x}}] \neq \mathcal{Q}[\boldsymbol{\Phi}\boldsymbol{x}]$

$$\underbrace{\|\boldsymbol{\Phi}\hat{\boldsymbol{x}} - \mathcal{Q}[\boldsymbol{\Phi}\boldsymbol{x}]\| \leqslant \epsilon_2}_{\text{(from BPDN constraint)}} \quad \not\Rightarrow \mathcal{Q}[\boldsymbol{\Phi}\hat{\boldsymbol{x}}] = \mathcal{Q}[\boldsymbol{\Phi}\boldsymbol{x}]$$

$\Rightarrow$ sensing information is fully not exploited!

  - ℓ_2 constraint $\approx$ Gaussian distribution (MAP - cond. log. lik.)
- But why looking for consistency ?

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:*

$$\mathrm{RMSE} = (\mathbb{E}\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2)^{1/2} \gtrsim r^{-1}\alpha,$$

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

This bound is achieved, at least, for $\boldsymbol{\Phi}_T = \mathrm{DFT} \in \mathbb{R}^{M\times K}$, *when* $\mathrm{Dec}()$ *is* **consistent.**

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

In quest of consistency...

$\ell_2 \to \ell_\infty$

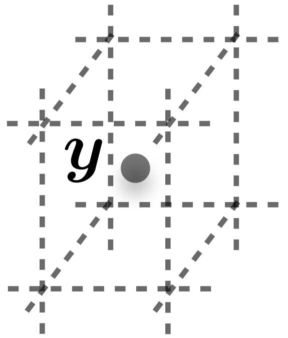
- Modify BPDN [W. Dai, O. Milenkovic, 09]

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{u}\in\mathbb{R}^N}{\mathrm{argmin}} \|\boldsymbol{u}\|_1 \text{ s.t. } \mathcal{Q}[\boldsymbol{\Phi u}] = \boldsymbol{q}$$

$+$ modified greedy algo:
"*subspace pursuit*"

$$\Leftrightarrow \boldsymbol{\Phi u} \in \mathcal{Q}^{-1}[\boldsymbol{q}]$$

convex set in $\mathbb{R}^M$



$$\Leftrightarrow \|\boldsymbol{\Phi u} - \boldsymbol{q}\|_\infty \leq \alpha/2$$

(if uniform quant.)

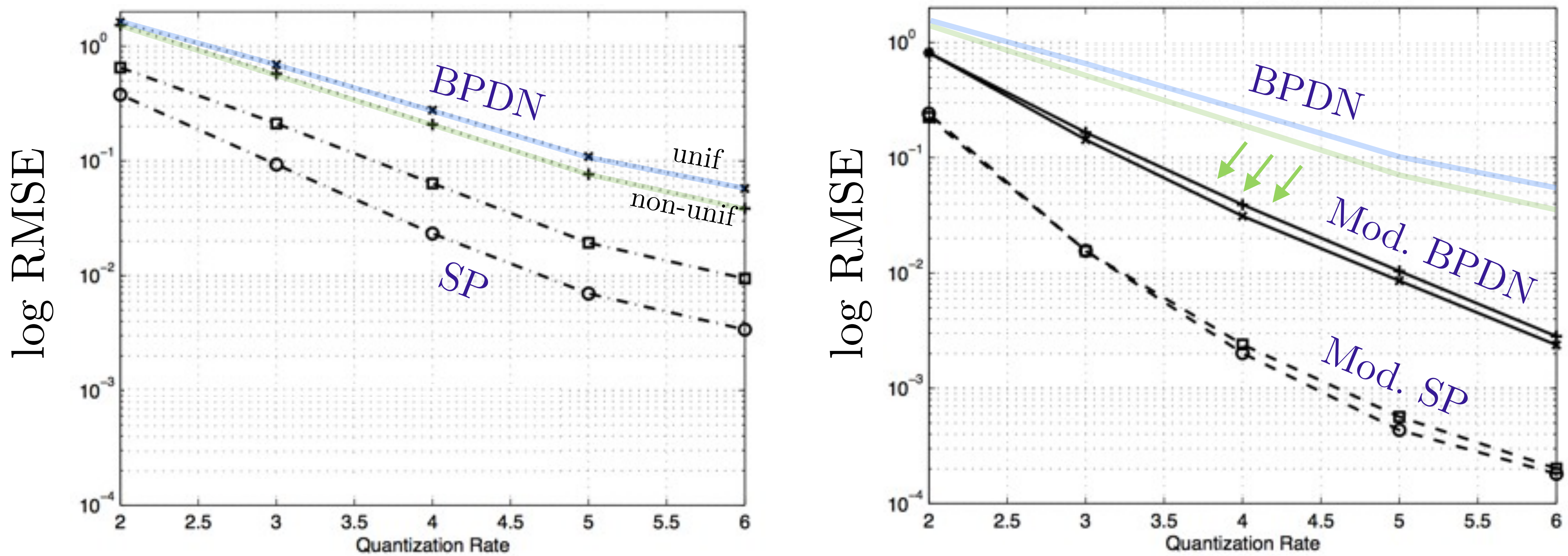
$\exists$ numerical methods

In quest of consistency... $\ell_2 \to \ell_\infty$

- Modify BPDN [W. Dai, O. Milenkovic, 09]

$$\hat{\boldsymbol{x}} = \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_1 \text{ s.t. } \mathcal{Q}[\boldsymbol{\Phi u}] = \boldsymbol{q}$$

Simulations: $M = 128, N = 256, K = 6, 1000$ trials $\Rightarrow \lambda \simeq 20$

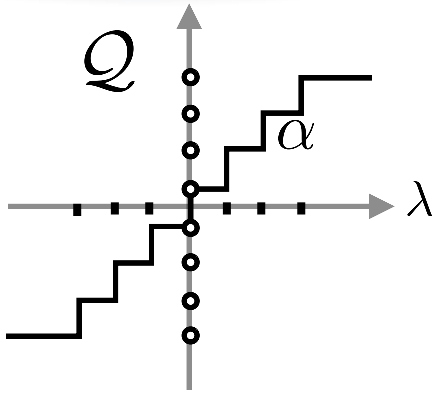


W. Dai, H. V. Pham, and O. Milenkovic, "Quantized Compressive Sensing", preprint, 2009

Dequantizing CS?

[LJ, Hammond, Fadili, 2009, 2011]

$\ell_2 \to \ell_p \; (p \geq 2)$



- Distortion model:

$$\boldsymbol{q} = \mathcal{Q}[\boldsymbol{\Phi x}] = \boldsymbol{\Phi x} + \boldsymbol{n}, \quad n_i \sim U(-\tfrac{\alpha}{2}, \tfrac{\alpha}{2})$$

- Observation: $\|\boldsymbol{\Phi x} - \boldsymbol{q}\|_\infty \leq \alpha/2$
- Reconstruction: Generalizing BPDN with BPDQ

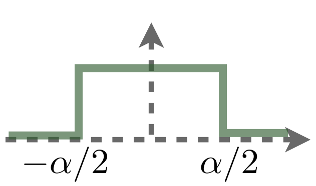
$$\boxed{\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{u} \in \mathbb{R}^N} \|\boldsymbol{u}\|_1 \text{ s.t. } \|\boldsymbol{q} - \boldsymbol{\Phi u}\|_p \leq \epsilon_p}$$

Towards $p = \infty$

Related to GGD MAP

How to find it? again, uniform model:

$$\begin{aligned} n_i &= \mathcal{Q}[(\boldsymbol{\Phi x})_i] - (\boldsymbol{\Phi x})_i \\ &\in q_{k_i} - \mathcal{R}_{k_i} = [-\alpha/2, \alpha/2] \\ &\sim_{\text{iid}} \text{Uniform}([-\alpha/2, \alpha/2]) \end{aligned} \quad \Longrightarrow$$



Estimating p^{th} moment:

$$\epsilon_p(\alpha) = \frac{\alpha}{2(p+1)^{1/p}} \big(M + \kappa(p+1)\sqrt{M}\,\big)^{1/p}$$

works with $\Pr \geq \; 1 - e^{-2\kappa^2}$

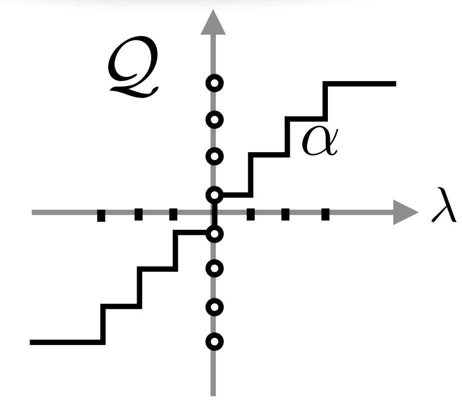
Note: $\epsilon_p(\alpha) \underset{p \to \infty}{\longrightarrow} \frac{\alpha}{2} = \text{QC}\,!$

Dequantizing CS?

$\ell_2 \to \ell_p \ (p \geq 2)$

[LJ, Hammond, Fadili, 2009, 2011]

- Distortion model:



$$\boldsymbol{q} = \mathcal{Q}[\boldsymbol{\Phi x}] = \boldsymbol{\Phi x} + \boldsymbol{n}, \quad n_i \sim U(-\tfrac{\alpha}{2}, \tfrac{\alpha}{2})$$

- Observation: $\|\boldsymbol{\Phi x} - \boldsymbol{q}\|_\infty \leq \alpha/2$
- Reconstruction: Generalizing BPDN with BPDQ

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

Towards $p = \infty$

Related to GGD MAP

BPDQ Stability ?

If $\boldsymbol{\Phi}$ is RIP$_p$ of order K, *i.e.*,

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

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

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

Gain over BPDN (for tight $\epsilon_p(\alpha, M)$)

$\Rightarrow \|\boldsymbol{x} - \hat{\boldsymbol{x}}\| = O(\epsilon_p/\mu_p)$

$$\Rightarrow \|\boldsymbol{x} - \hat{\boldsymbol{x}}\| = O(\alpha/\sqrt{p+1})$$

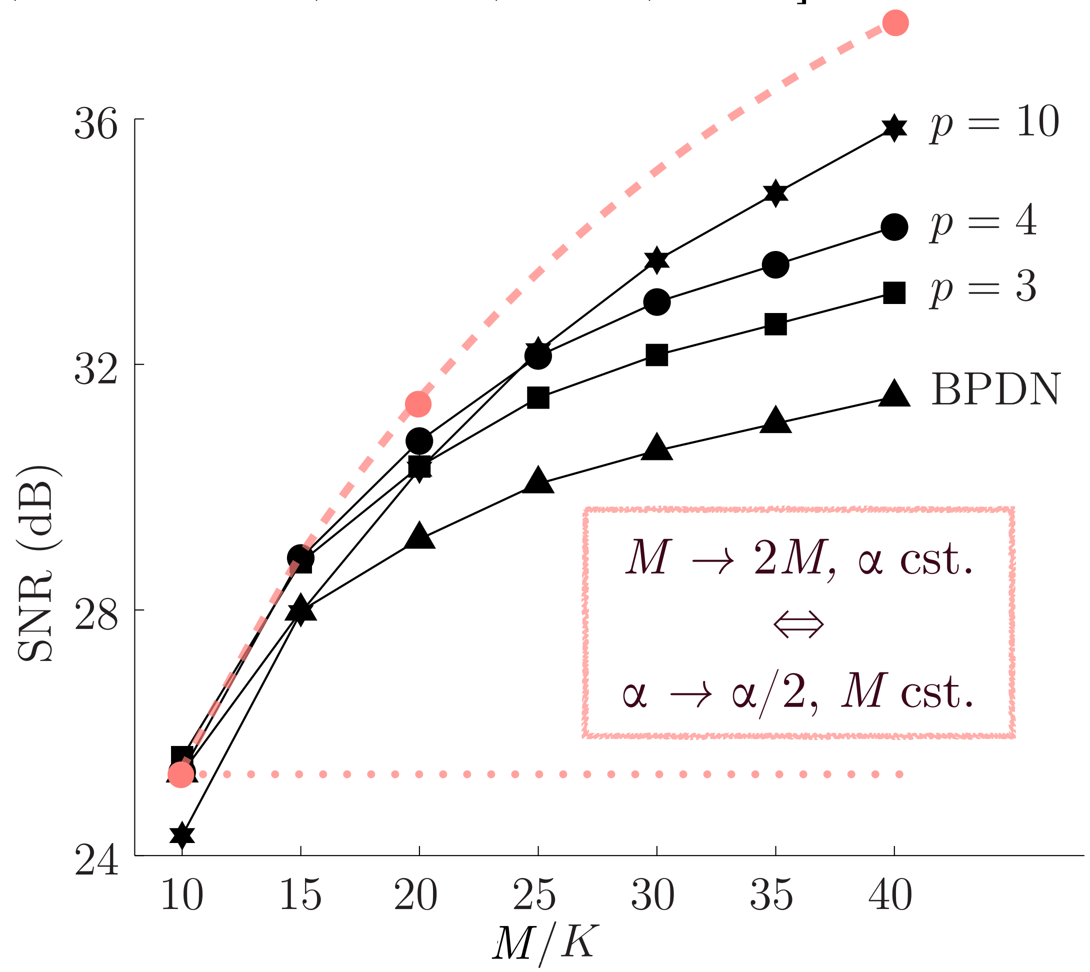
But no free lunch: for $\boldsymbol{\Phi}$ Gaussian

$$M = O\big(\,(K \log N/K\,)^{p/2}\,\big)$$

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

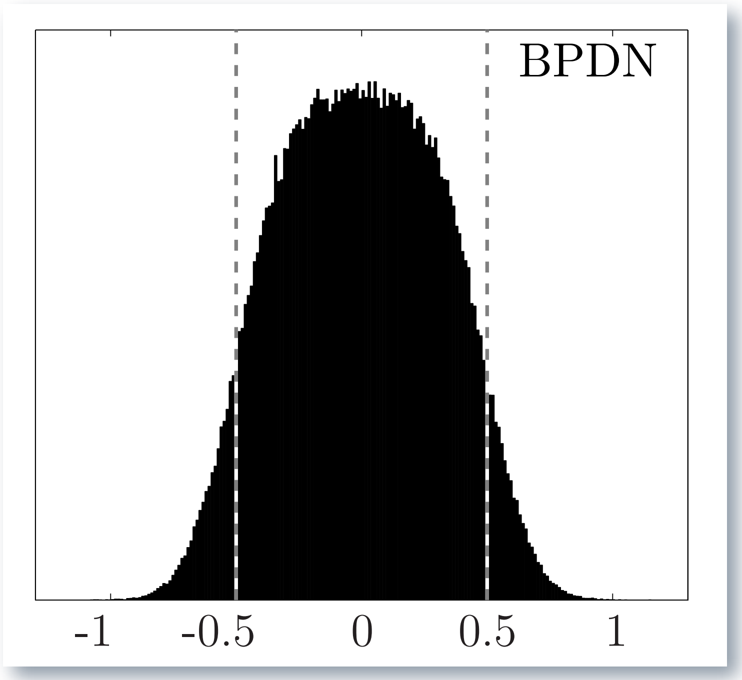
Dequantizing CS?

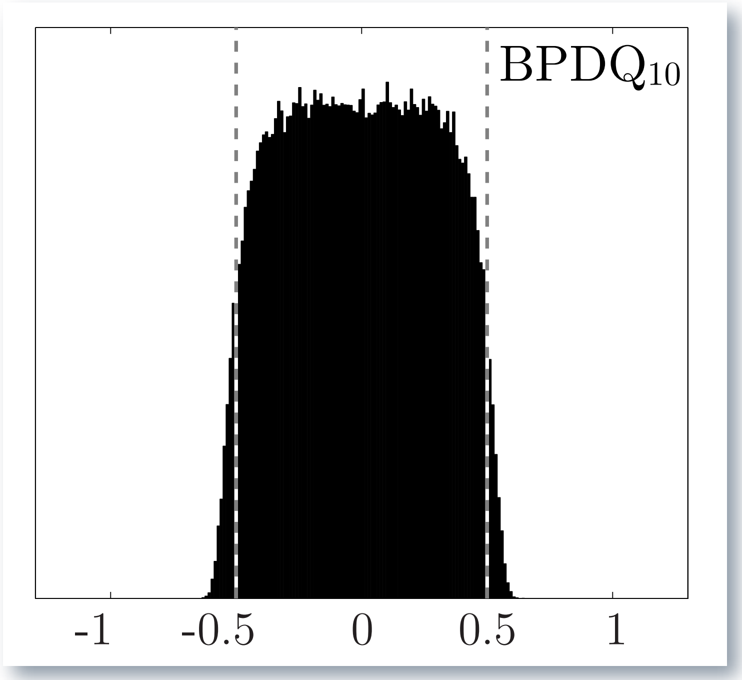
[LJ, Hammond, Fadili, 2009, 2011]



Histograms of

$$\alpha^{-1}(\boldsymbol{q} - \boldsymbol{\Phi}\hat{\boldsymbol{x}})_i$$





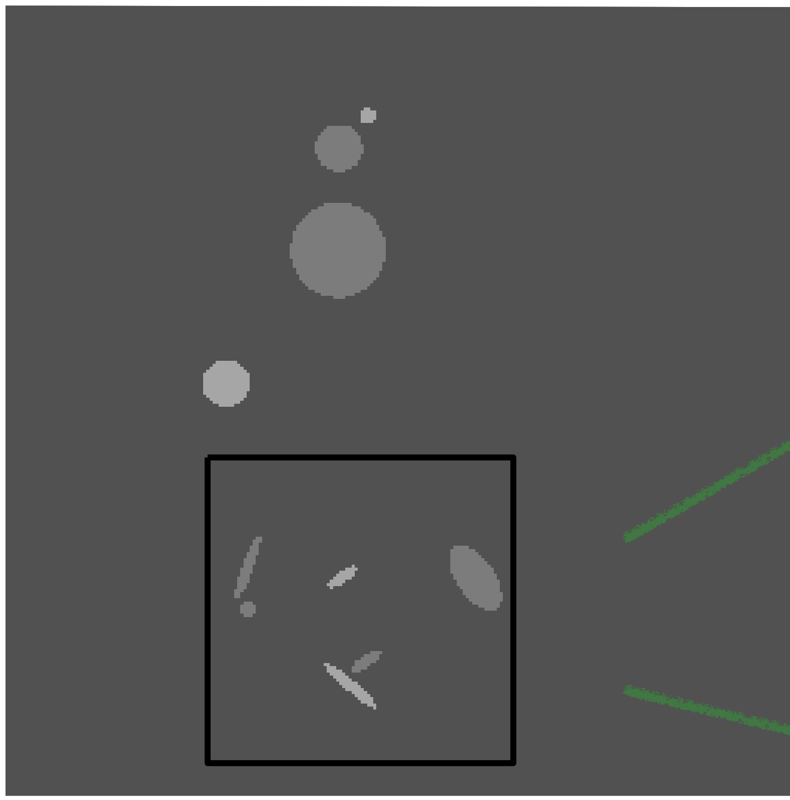
- N=1024, K=16, Gaussian $\boldsymbol{\Phi}$
- 500 K-sparse (canonical basis)
- Non-zero components follow $\mathcal{N}(0, 1)$
- Quantiz. bin width $\alpha = \|\boldsymbol{\Phi x}\|_\infty/40$

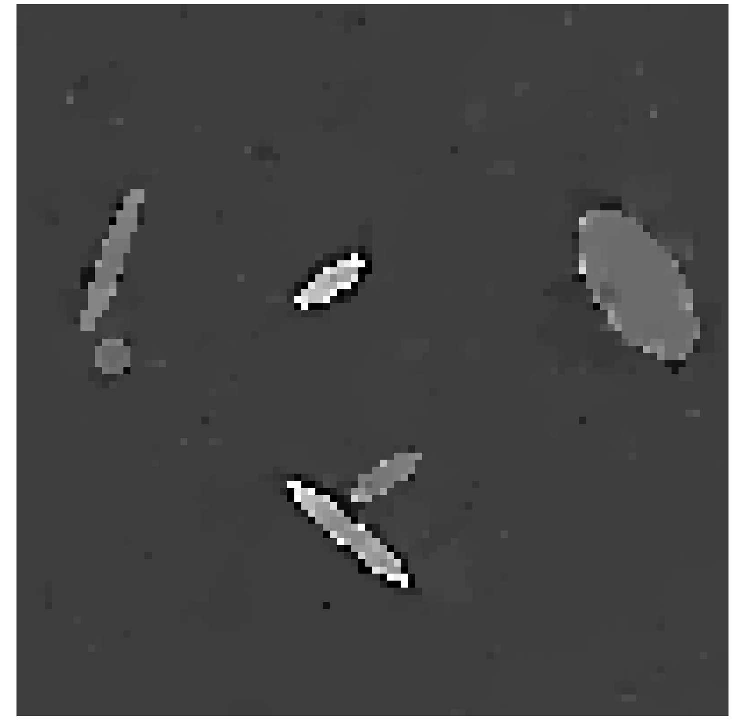
LJ, D. Hammond, J. Fadili "Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine." Information Theory, IEEE Transactions on, 57(1), 559-571.

Dequantizing CS?

[LJ, Hammond, Fadili, 2009, 2011]

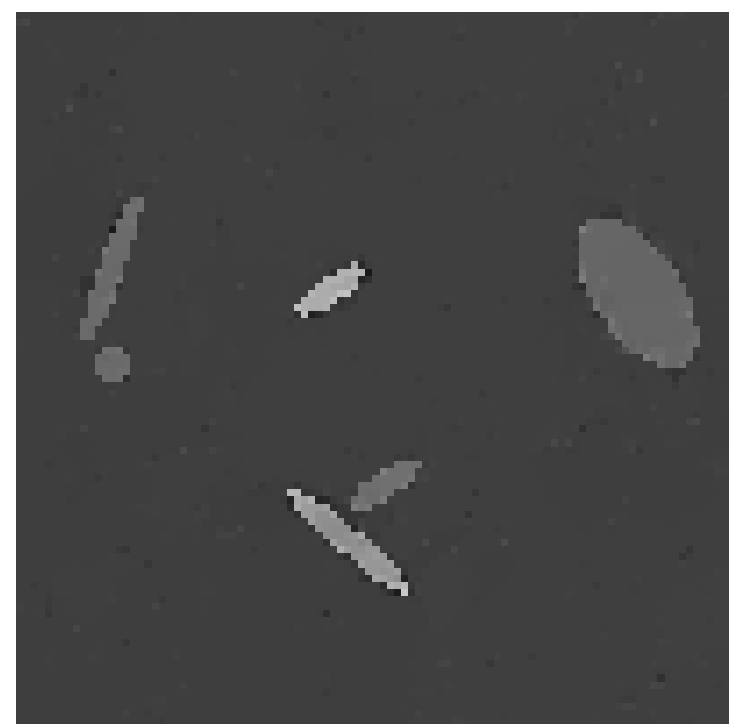
A bit outside the theory...





BPDN-TV

SNR: 8.96 dB



BPDQ$_{10}$ -TV

SNR: 12.03 dB

* Synthetic Angiogram [Michael Lustig 07, SPARCO],
* $\boldsymbol{\Phi}$: **Random Fourier Ensemble**
* $N/M = 8$
* Decoder: $\Delta_{TV,p}(y, \epsilon_p)$
* Quantiz. bin width $= 50$ (i.e. 12 bins)

LJ, D. Hammond, J. Fadili "Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine." Information Theory, IEEE Transactions on, 57(1), 559-571.

Non-uniform dequantization?

Possible!

Under High Resolution Assumption (HRA)

$= \text{high } B$

1. Use compander formalism:

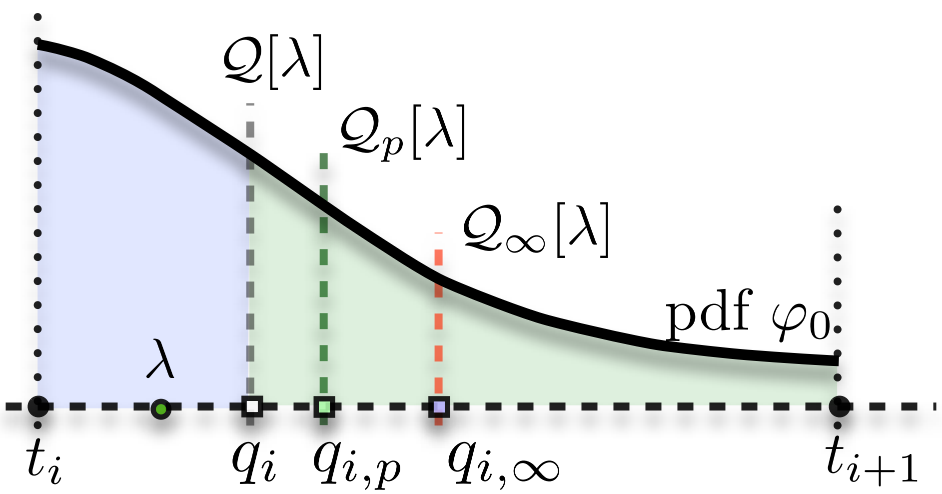
$$\boxed{\mathcal{Q} = \mathcal{G}^{-1} \circ \mathcal{Q}_\alpha \circ \mathcal{G}}$$

Expander Compressor

2. Redefine $\boldsymbol{q}$: (post-sensing)

$$\boldsymbol{q} \to \boldsymbol{q}_p = \mathcal{Q}_p[\boldsymbol{q}] = \mathcal{Q}_p[\boldsymbol{\Phi x}]$$

with $(q_p)_i$ minimizing p^{th} moment in each bin.



3. Reweight the bins: $\|\cdot\|_p \to \|\text{diag}(\boldsymbol{w}) \cdot\|_p =: \|\cdot\|_{p,\boldsymbol{w}}$

with: $w_i(p) := \mathcal{G}'\big((q_p)_i\big)^{(p-2)/p}$

$\to$ kind of noise stabilization operation ("equi-p-distortion")

4. Solve:

$$\boxed{\hat{\boldsymbol{x}} \;=\; \arg\min_{\boldsymbol{u} \in \mathbb{R}^N} \|\boldsymbol{u}\|_1 \text{ s.t. } \|\boldsymbol{q} - \boldsymbol{\Phi u}\|_{p,\boldsymbol{w}} \;\le \epsilon_p}$$

LJ, D. Hammond, J. Fadili, "Stabilizing Nonuniformly Quantized Compressed Sensing with Scalar Companders", arXiv:1206.6003 (2012).

Non-uniform dequantization?

- Stability? Well ... need a more general RIP ☺

$\mathrm{RIP}(\ell_{p,\boldsymbol{w}}, \ell_2 | K, \delta, \mu)$

$\exists\, \mu > 0,\ \delta \in (0,1)$

$$\sqrt{1-\delta}\,\|\boldsymbol{v}\|_2 \ \leqslant\ \tfrac{1}{\mu}\|\boldsymbol{\Phi v}\|_{p,\boldsymbol{w}} \ \leqslant\ \sqrt{1+\delta}\,\|\boldsymbol{v}\|_2$$

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

$$\Rightarrow M = O((\theta(\boldsymbol{w}) K \log N/K)^{p/2})$$

with: $\theta(\boldsymbol{w}) \simeq M^{2/p} \|\boldsymbol{w}\|_\infty^2 / \|\boldsymbol{w}\|_p^2$ ($= 1$ if w_i =cst)

Then,

Given $\mathcal{Q}_p[\cdot]$, $\boldsymbol{w}(p) \in \mathbb{R}_+^M$ and ϵ_p as before, GBPDN robustness provides:

$$\|\boldsymbol{x}^* - \boldsymbol{x}\| \ \underset{B,M}{\lesssim}\ 4\,c'\,\frac{2^{-B}}{\sqrt{p+1}} \ +\ 2\,e_0(K),$$

with $c' = (9/8)(e\pi/3)^{1/2} < 1.8981$.

4. Sigma-Delta quantization in CS

Context:

- **Former attempts:** (see prev. slides)

  CS + uniform scalar quantization (or pulse code modulation - PCM)

For K-sparse signals: $\|\mathcal{Q}_\alpha[\boldsymbol{\Phi x}] - \boldsymbol{\Phi x}\|_2 \leqslant c\sqrt{M}\alpha \;\Rightarrow\; \|\boldsymbol{x}^* - \boldsymbol{x}\| \leqslant C\alpha$ (with RIP)

and for high λ, $\|\mathcal{Q}_\alpha[\boldsymbol{\Phi x}] - \boldsymbol{\Phi x}\|_p \leqslant cM^{1/p}\alpha \;\Rightarrow\; \|\boldsymbol{x}^* - \boldsymbol{x}\| \leqslant C\alpha/\sqrt{p+1}$ (with RIP$_p$)

- **No improvement if M increases!**
- **Can we do better?**

  Can we have $\|\boldsymbol{x}^* - \boldsymbol{x}\| \leqslant O(r^{-s}\alpha)$ for some $s > 0$?

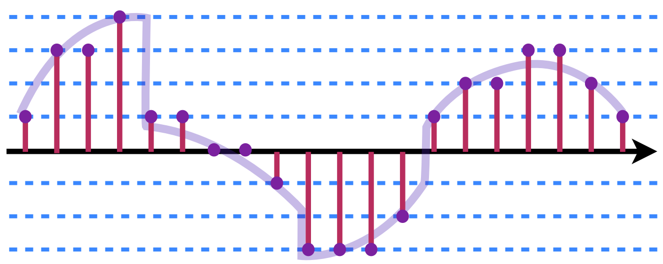
- **Staying with PCM,** $s \leqslant 1$ (Goyal-Vetterli-Thao lower bound)
- **Solution: replacing PCM by $\Sigma\Delta$ quantization!**

  [S. Güntürk, A. Powell, R. Saab, Ö. Yılmaz]

$\Sigma\Delta$ quantization (reminder)

- PCM: Signal sensing + unif. quantization (step α)

$$\boldsymbol{x} \in \mathbb{R}^K \quad \longrightarrow \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \in \mathbb{R}^M$$



$\boldsymbol{q} = \mathcal{Q}_{\rm PCM}[\boldsymbol{y}]$ with

$$q_k = \mathcal{Q}_{\rm PCM}[y_k] := \operatorname*{argmin}_{u \in \alpha\mathbb{Z}} |y_k - u|, \quad 1 \leqslant k \leqslant M$$

Let $\boldsymbol{A}^\#$, a left inverse of $\boldsymbol{A}$, *i.e.*, $\boldsymbol{A}^\#\boldsymbol{A} = \mathbf{Id}$.

$$\hat{\boldsymbol{x}} := \boldsymbol{A}^\#\boldsymbol{q} \Rightarrow \|\boldsymbol{x} - \hat{\boldsymbol{x}}\| = \|\boldsymbol{A}^\#\underbrace{(\boldsymbol{y} - \boldsymbol{q})}_{\text{quant. noise}}\|$$

Taking (Moore-Penrose) pseudo-inverse: $\underbrace{\boldsymbol{A}^\# = \boldsymbol{A}^\dagger = (\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*}_{\text{(or canonical dual of the frame } \boldsymbol{A})}$

minimize $\|\boldsymbol{A}^\#(\boldsymbol{y} - \boldsymbol{q})\|$! (least square solution)

- In CS, this could be used if signal support was known (see before)

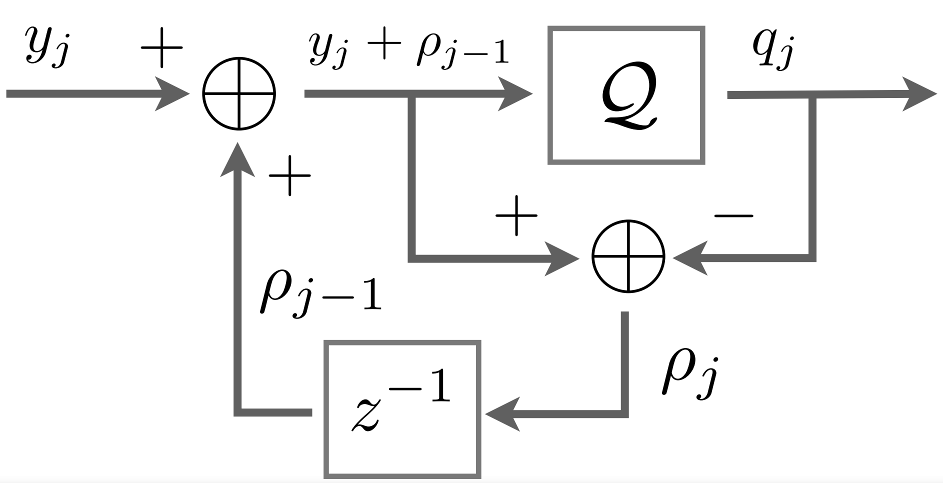
$\Sigma\Delta$ quantization (reminder)

- $\Sigma\Delta \equiv$ noise shaping! Enjoy of:
  - freedom to pick $\boldsymbol{q} \in \alpha\mathbb{Z}^M$
  - freedom to take another left inverse $\boldsymbol{A}^{\#}$
- 1^{st} order $\Sigma\Delta$: (in 1-D) Quantizing the sequence $\{y_j : j \geqslant 0\}$

Use of state variables $\{\rho_j\}$ (1-step memory):

find q_j: $q_j = \mathcal{Q}^{(1)}_{\Sigma\Delta}[y_j] := \mathrm{argmin}_{u \in \alpha\mathbb{Z}} |\rho_{j-1} + y_j - u|$ $= \mathcal{Q}_{\text{PCM}}[\rho_{j-1} + y_j]$

find ρ_j: $(\Delta\rho)_j = \rho_j - \rho_{j-1} = y_j - q_j$ (difference eq.)



with: $|\rho_j| \leqslant \alpha$

$|y_j - q_j| \leqslant 2\alpha$

bigger than α but still $O(\alpha)$

$\Sigma\Delta$ quantization (reminder)

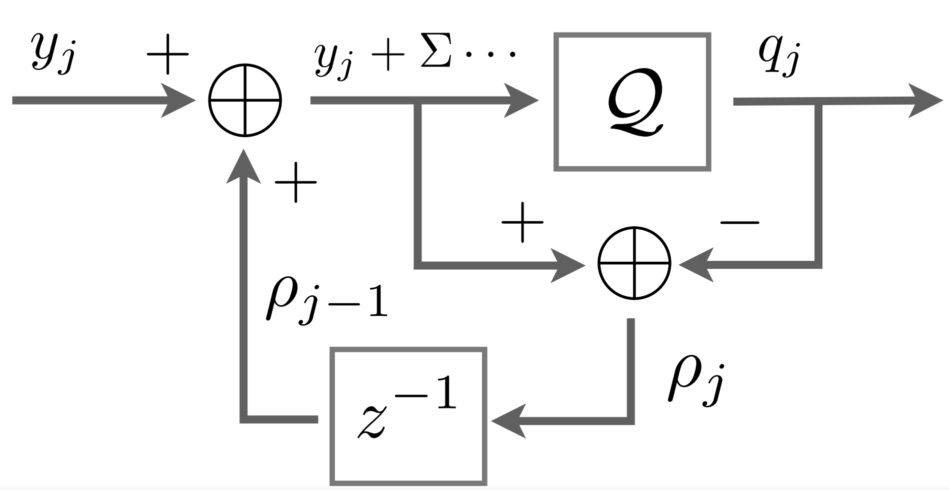
- $\Sigma\Delta \equiv$ noise shaping! Enjoy of:
  - freedom to pick $\boldsymbol{q} \in \alpha\mathbb{Z}^M$
  - freedom to take another left inverse $\boldsymbol{A}^{\#}$
- s^{th} order $\Sigma\Delta$: (in 1-D) Quantizing the sequence $\{y_j : j \geqslant 0\}$

  Use of state variables $\{\rho_j\}$ (s-step memory):

Remark:
PCM is
0^{th} order $\Sigma\Delta$

find q_j: $\quad q_j = \mathcal{Q}^{(s)}_{\Sigma\Delta}[y_j] := \text{argmin}_{u\,\in\,\alpha\mathbb{Z}}\,\big|\sum_{i=1}^{s}(-1)^{i-1}\binom{s}{i}\rho_{j-n} + y_j - u\big|$

find ρ_j: $\quad (\Delta^s\rho)_j = y_j - q_j \quad (s^{\text{th}}$ order difference eq.)



with: $|\rho_j| \leqslant \alpha$

$|y_j - q_j| \leqslant 2^{s-1}\alpha$

bigger than α but still $O(\alpha)$

$\Sigma\Delta$ quantization (reminder)

- $\Sigma\Delta \equiv$ noise shaping! Enjoy of:
  - freedom to pick $\boldsymbol{q} \in \alpha\mathbb{Z}^M$
  - freedom to take another left inverse $\boldsymbol{A}^{\#}$
- s^{th} order $\Sigma\Delta$:

Most important fact: $(\Delta^s \rho)_j = y_j - q_j \;\Leftrightarrow\; \boldsymbol{D}^s \boldsymbol{\rho} = \boldsymbol{y} - \boldsymbol{q}$

$$\hat{\boldsymbol{x}} := \boldsymbol{A}^{\#}\boldsymbol{q} \Rightarrow \|\boldsymbol{x} - \hat{\boldsymbol{x}}\| = \|\boldsymbol{A}^{\#}\boldsymbol{D}^s(\boldsymbol{y} - \boldsymbol{q})\|$$

minimize $\|\boldsymbol{A}^{\#}\boldsymbol{D}^s(\boldsymbol{y} - \boldsymbol{q})\|$!

~~Pseudo-inverse~~

$\cancel{\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*}$

Sobolev duals

$\boldsymbol{A}_{\mathrm{sob},s} = (\boldsymbol{D}^{-s}\boldsymbol{A})^{\dagger}\boldsymbol{D}^{-s}$

$\Sigma\Delta$ quantization (reminder)

- $\Sigma\Delta \equiv$ noise shaping! Enjoy of:
  - freedom to pick $\boldsymbol{q} \in \alpha\mathbb{Z}^M$
  - freedom to take another left inverse $\boldsymbol{A}^{\#}$
- s^{th} order $\Sigma\Delta$:

Most important fact: $(\Delta^s \rho)_j = y_j - q_j \;\Leftrightarrow\; \boldsymbol{D}^s \boldsymbol{\rho} = \boldsymbol{y} - \boldsymbol{q}$

$$\boxed{\hat{\boldsymbol{x}} = \boldsymbol{A}_{\mathrm{sob},s}\boldsymbol{q}} \qquad \boldsymbol{A}_{\mathrm{sob},s} = (\boldsymbol{D}^{-s}\boldsymbol{A})^{\dagger}\boldsymbol{D}^{-s}$$

Proposition *Let* $\boldsymbol{A} \in \mathbb{R}^{M\times K}$ *with* $A_{ij} \sim_{\mathrm{iid}} \mathcal{N}(0,1)$. *For any* $\kappa \in (0,1)$, *if* $r := M/K \geqslant c(\log M)^{1/(1-\kappa)}$, *then with* $Pr > 1 - e^{-c' M/r^{\kappa}}$,

$$\boxed{\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \;\leqslant\; C_s\, r^{-\kappa(s-\frac{1}{2})}\alpha,}$$

for some $c, c', C_s > 0$.

proof: show that
$\sigma_{\min}(\boldsymbol{D}^{-s}\boldsymbol{A}) > C'_s\, r^{\kappa(s-\frac{1}{2})}\sqrt{M}$

$\Sigma\Delta$ quantization in CS

$$\boldsymbol{x} \in \Sigma_K \subset \mathbb{R}^N \quad \longrightarrow \quad \boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} \in \mathbb{R}^M \quad \longrightarrow \quad \boldsymbol{q} = \mathcal{Q}^{(s)}_{\Sigma\Delta}[\boldsymbol{y}]$$

$$\|\boldsymbol{y} - \boldsymbol{q}\| \leqslant 2^{s-1}\,\alpha\,\sqrt{M}$$

Two-steps procedure:

1. find the support T of $\boldsymbol{x}$: coarse approx. with BPDN
2. compute $\hat{\boldsymbol{x}} := (\boldsymbol{\Phi}_T)_{\mathrm{sob},s}\boldsymbol{q} = (\boldsymbol{D}^{-s}\boldsymbol{\Phi}_T)^{\dagger}\boldsymbol{D}^{-s}\boldsymbol{q}$

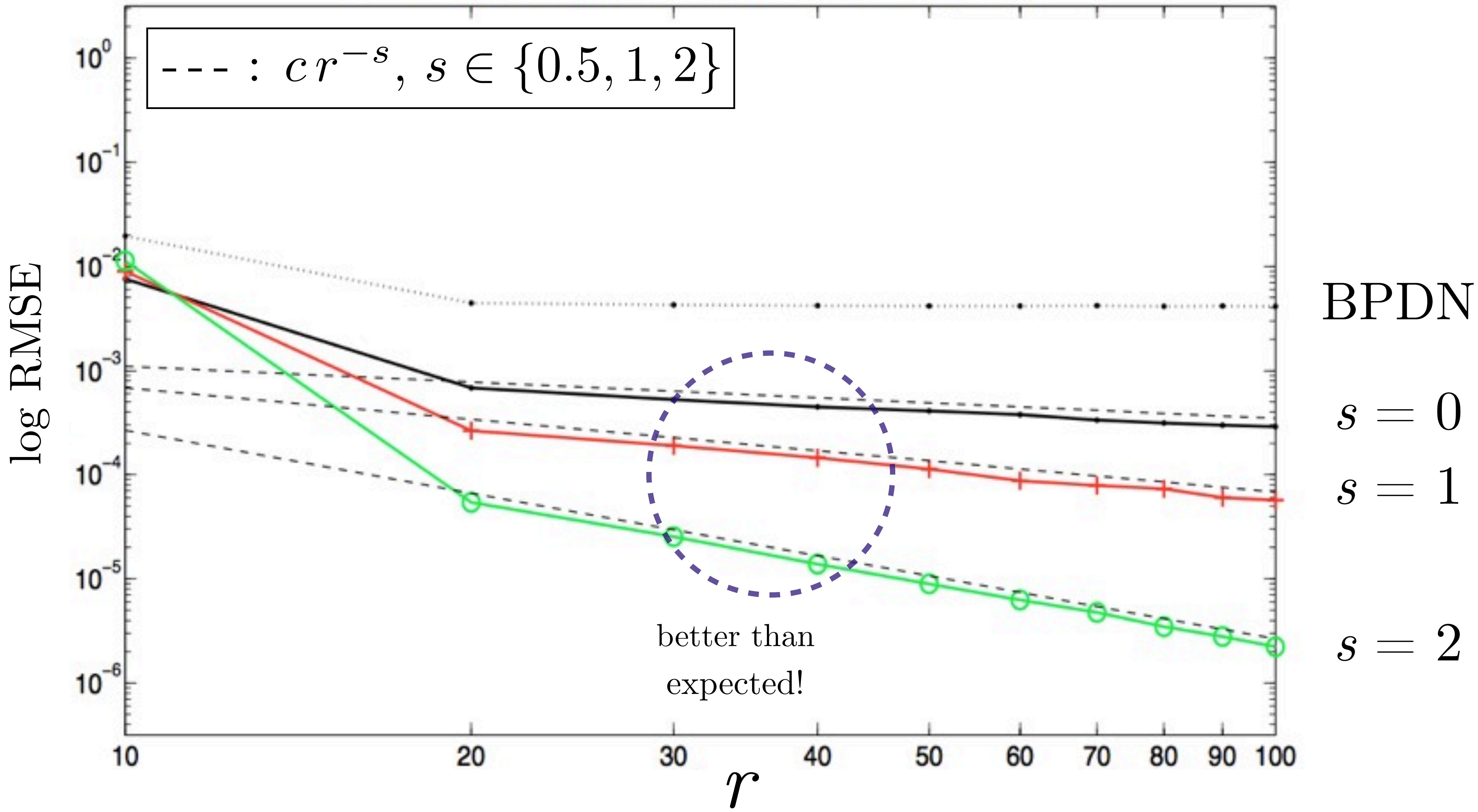
Proposition *Let* $\boldsymbol{\Phi} \in \mathbb{R}^{M\times K}$ *with* $\Phi_{ij} \sim_{\mathrm{iid}} \mathcal{N}(0,1)$. *Suppose* $\kappa \in (0,1)$ *and* $r := M/K \geqslant c(\log M)^{1/(1-\kappa)}$ *for* $c > 0$. *Then,* $\exists\, c', C, C_s > 0$ *such that, with* $Pr > 1 - e^{-c' M/r^{\kappa}}$, *for all* $\boldsymbol{x} \in \Sigma_K$ *s.t.* $\min_{i\in\operatorname{supp}\boldsymbol{x}} |x_i| \geqslant C\alpha$,

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \;\leqslant\; C_s\, r^{-\kappa(s-\frac{1}{2})}\alpha.$$

proof: Union bound on any K-column subset of $\boldsymbol{\Phi}$ + proba having good support.

ΣΔ quantization in CS (Simulations)

$M \in \{100, 200, \cdots, 1000\}$, $K = 10$ and 1000 trials ($x_i \in \{0, \pm 1/\sqrt{K}\}$, $\|\boldsymbol{x}\| \simeq 1, \alpha = 10^{-2}$)



Güntürk, C. S., Lammers, M., Powell, A. M., Saab, R., & Yılmaz, Ö. (2013). **Sobolev duals for random frames and ΣΔ quantization of compressed sensing measurements**. Foundations of Computational Mathematics, 13(1), 1-36.

5. To saturate or not? And how much?

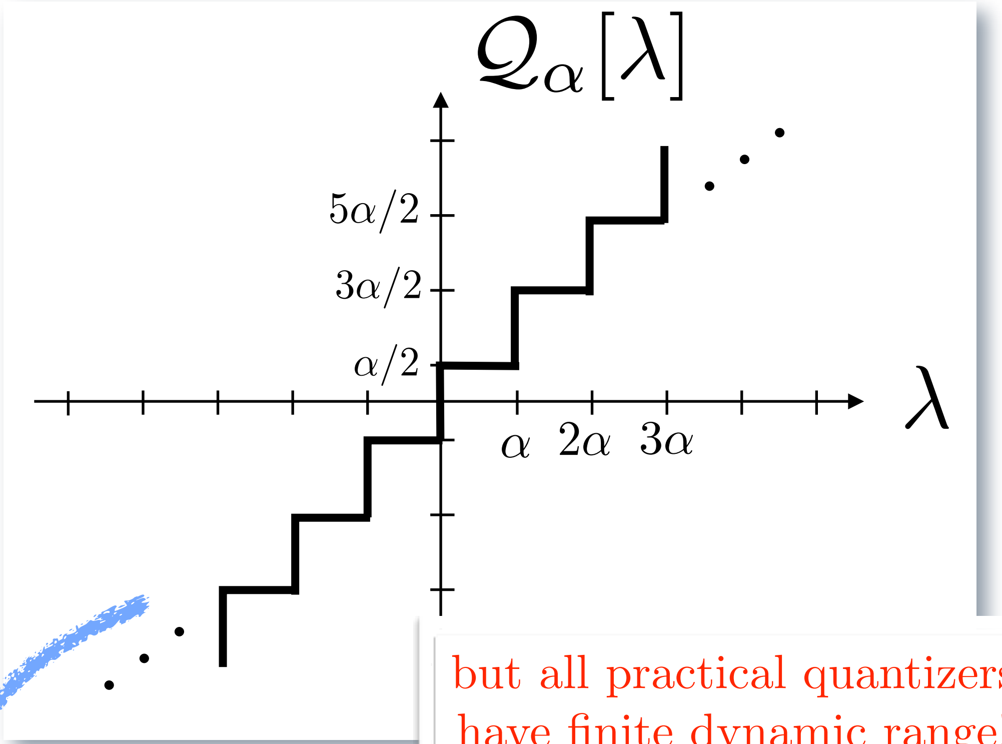
ICASSP 2013
Vancouver Convention & Exhibition Centre
May 26 - 31, 2013 • Vancouver, Canada

Saturation phenomenon:

Uniform quantization:

- α quantization interval
- error per measurement bounded:

$$|\lambda - \mathcal{Q}_\alpha[\lambda]| \leqslant \alpha/2$$

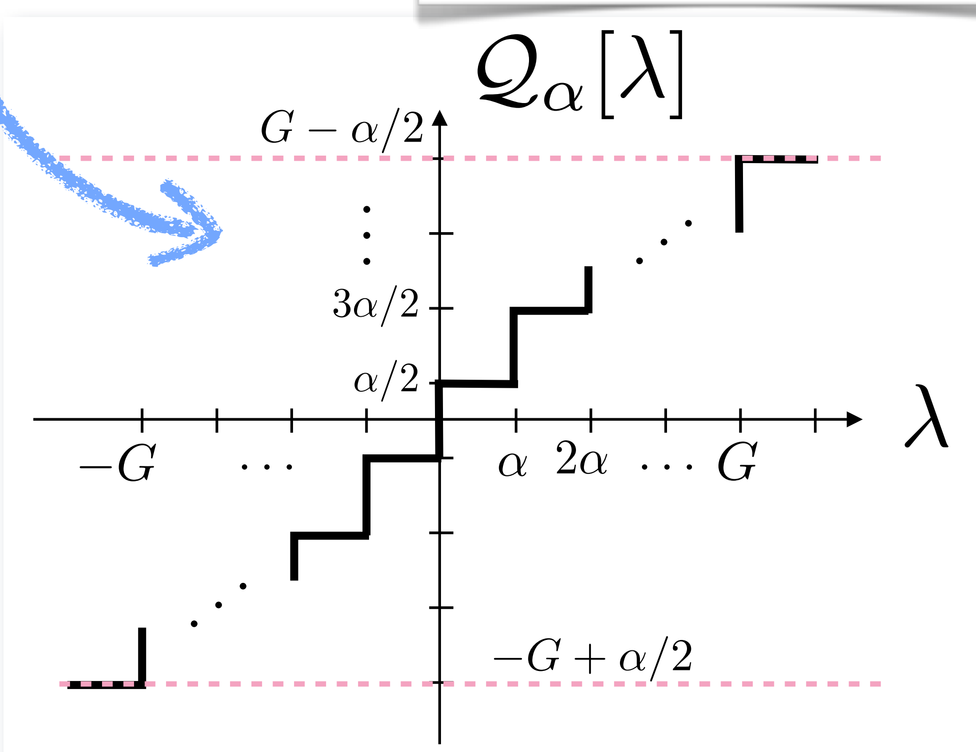


but all practical quantizers have finite dynamic range!

Finite Dynamic Range Quantization:

- G "saturation level"
- B bit rate (bits per measurement)
- quantization interval is $\alpha = 2^{-B+1}G$
- measurements above G saturate
- saturation error is *unbounded*

CS guarantees are for bounded errors only!

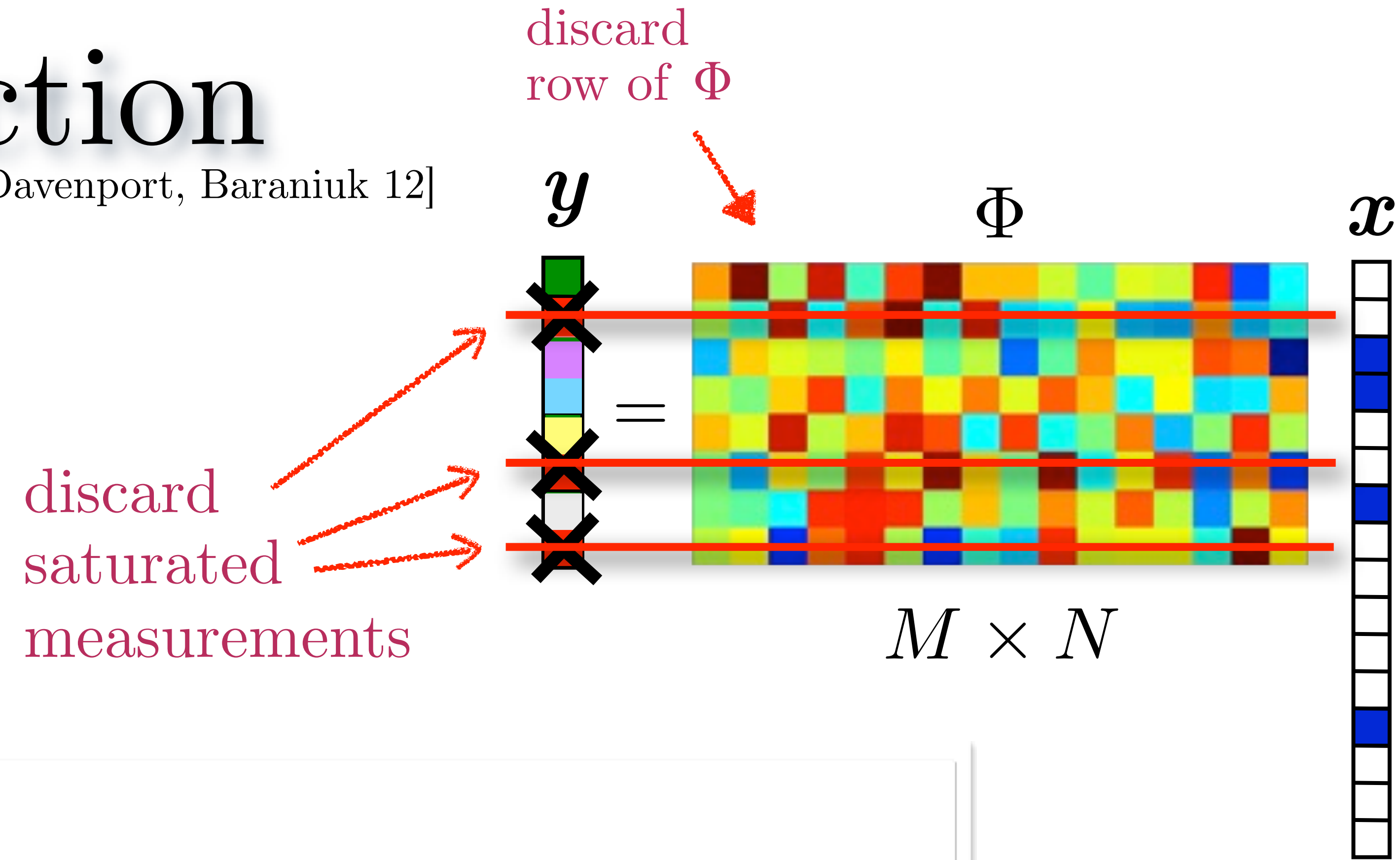


Democracy in Action

[Laska, Boufounos, Davenport, Baraniuk 12]

(i) <u>Saturation Rejection</u>:

Simply discard saturated measurements and corresponding rows of Φ

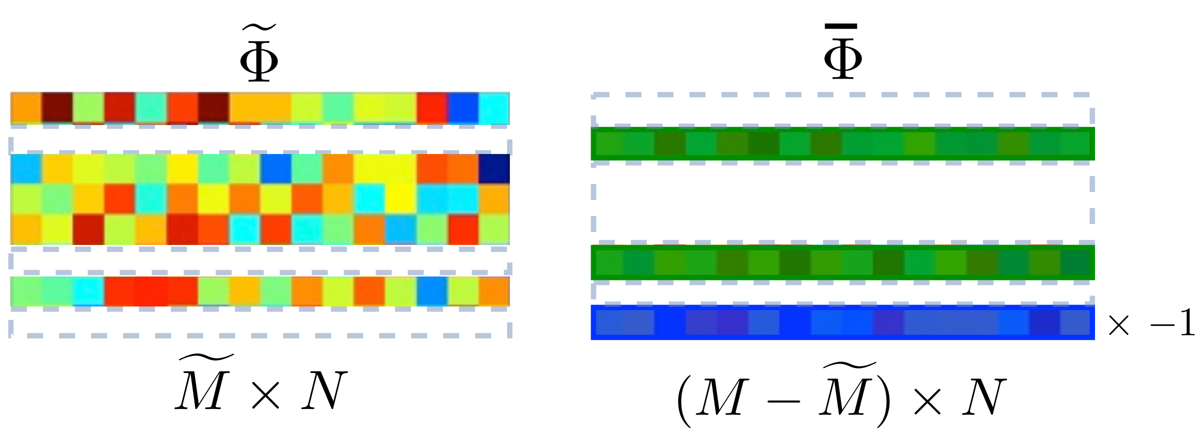


> *"democratic measurements"*
> each measurement has roughly same amount of information

RIP holds on row subsets of Φ

(ii) <u>Saturation Consistency:</u>

Include saturated measurements as inequality constraint



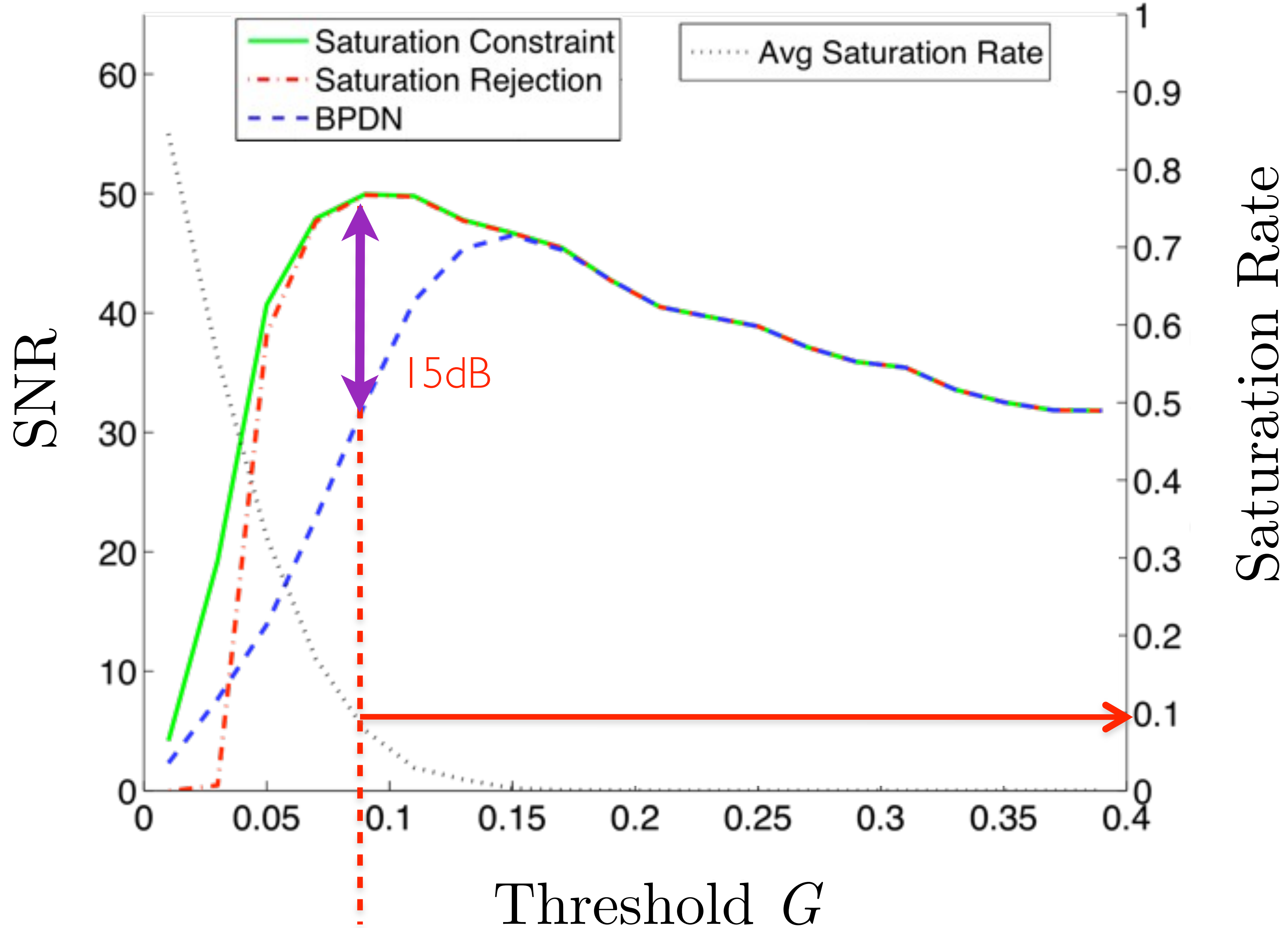
$$\widehat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\widetilde{\Phi}\mathbf{x} - \widetilde{\mathbf{y}}\|_2 < \epsilon$$

Measurement error term (quantization)

$$\text{and} \quad \bar{\Phi}\mathbf{x} \geq G \cdot \mathbf{1}$$

Saturation consistency constraint

Experimental Results

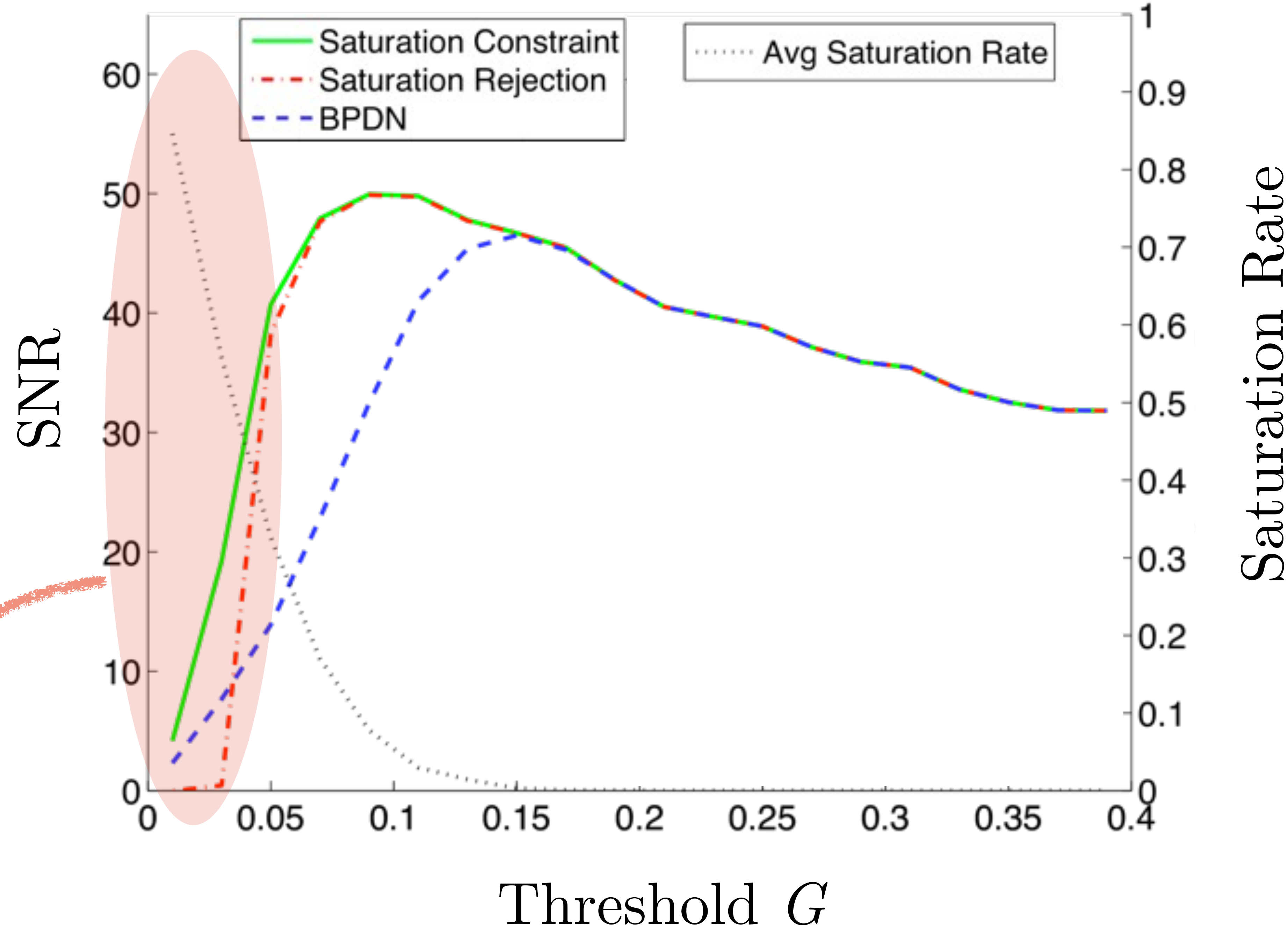


Note: optimal performance requires 10% saturation

J.N. Laska, P.T. Boufounos, M.A. Davenport, R.G.Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing". *Applied and Computational Harmonic Analysis*, *31*(3), 429-443. (2011)

Experimental Results

The "saturation gap"

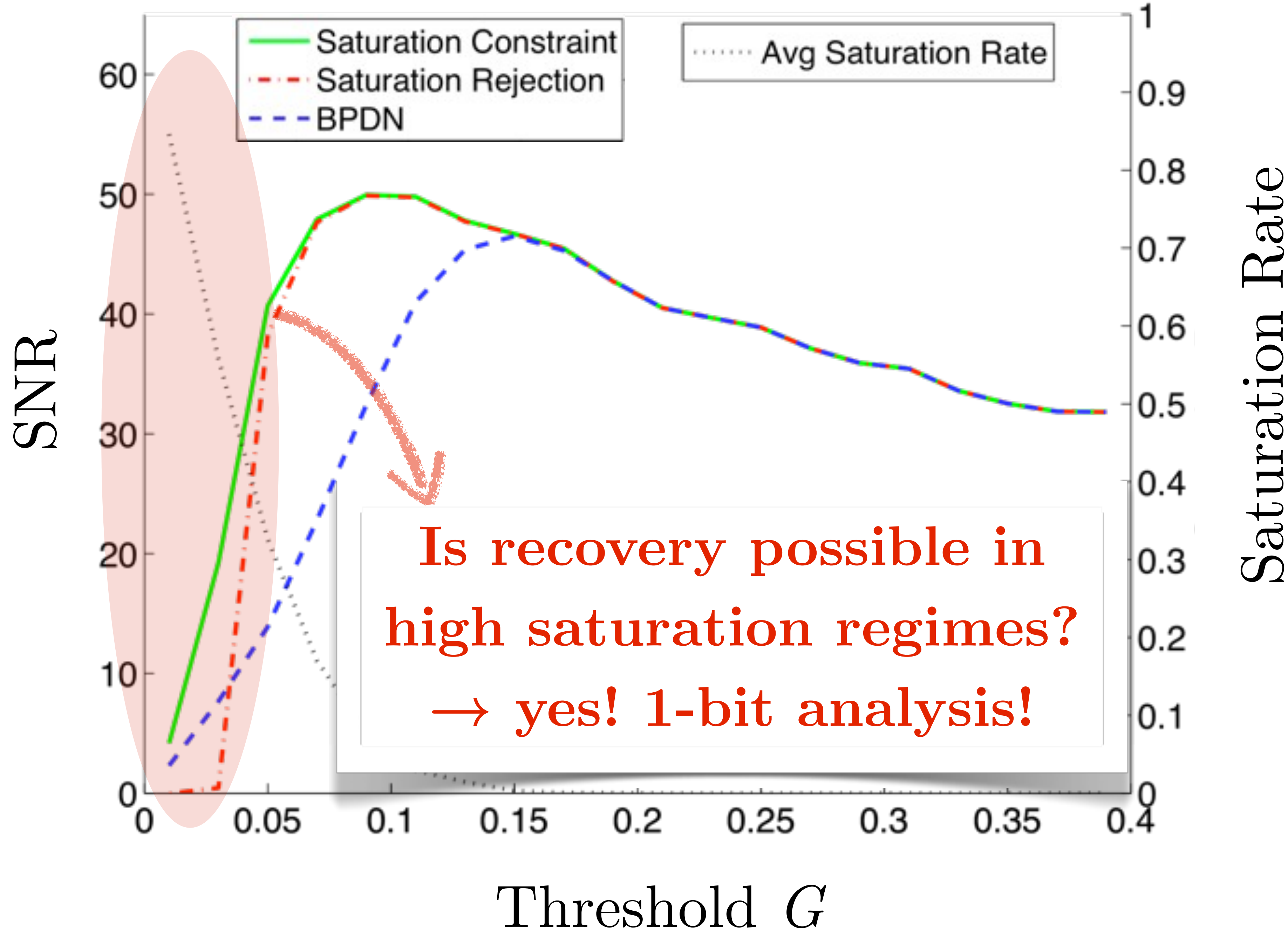


- Majority of measurements saturate
- Recovery fails

J.N. Laska, P.T. Boufounos, M.A. Davenport, R.G.Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing". *Applied and Computational Harmonic Analysis*, *31*(3), 429-443. (2011)

Experimental Results

The "saturation gap"



- Majority of measurements saturate
- Recovery fails

J.N. Laska, P.T. Boufounos, M.A. Davenport, R.G.Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing". *Applied and Computational Harmonic Analysis*, *31*(3), 429-443. (2011)

Further Reading

- V. K Goyal, M. Vetterli, N. T. Thao, "Quantized Overcomplete Expansions in RN: Analysis, Synthesis, and Algorithms", *IEEE Trans. Info. Theory*, 44(1), 1998

- P. T. Boufounos and R. G. Baraniuk, "Quantization of sparse representations," *Rice University ECE Department Technical Report 0701*. Summary appears in *Proc. Data Compression Conference (DCC)*, Snowbird, UT, March 27-29, 2007

- W. Dai, H. V. Pham, and O. Milenkovic, "Quantized Compressive Sensing", preprint, 2009

- L. Jacques, D. Hammond, J. Fadili "Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine." *IEEE Transactions on Information Theory*, 57(1), 559-571, 2011

- J.N. Laska, P.T. Boufounos, M.A. Davenport, R.G.Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing". *Applied and Computational Harmonic Analysis*, 31(3), 429-443, 2011

- L. Jacques, D. Hammond, J. Fadili, "Stabilizing Nonuniformly Quantized Compressed Sensing with Scalar Companders", arXiv:1206.6003, 2012

- Güntürk, C. S., Lammers, M., Powell, A. M., Saab, R., & Yılmaz, Ö. "Sobolev duals for random frames and $\Sigma\Delta$ quantization of compressed sensing measurements". *Foundations of Computational Mathematics*, 13(1), 1-36, 2013

Part IV:

Extreme quantization: 1-bit compressed sensing

Laurent Jacques, UCL, Belgium

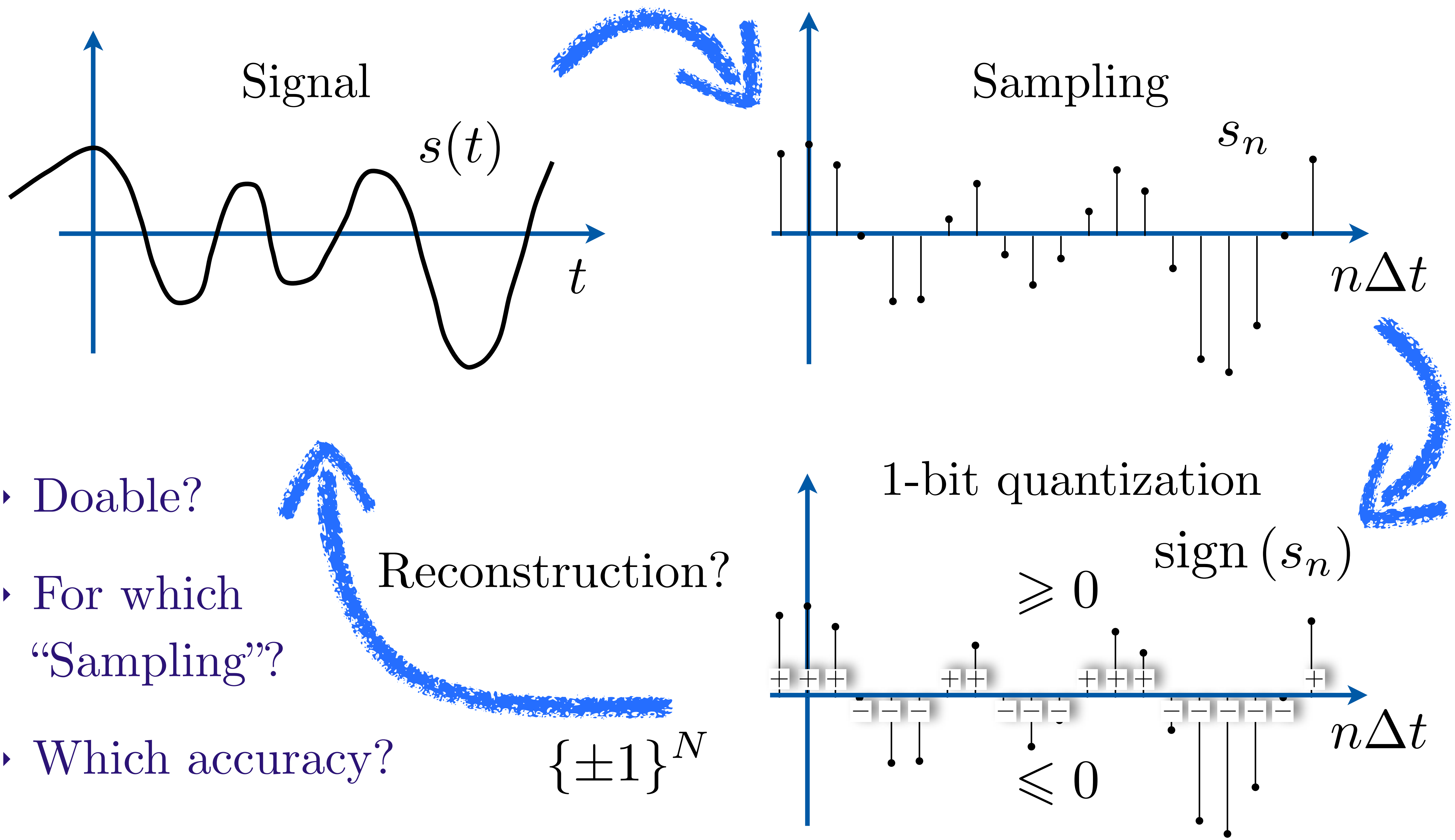
Petros Boufounos, MERL, USA

Outline:

1. Context
2. Theoretical performance limits
3. Stable embeddings: angles are preserved
4. Generalized Embeddings
5. 1-bit CS Reconstructions?
6. Playing with thresholds in 1-bit CS

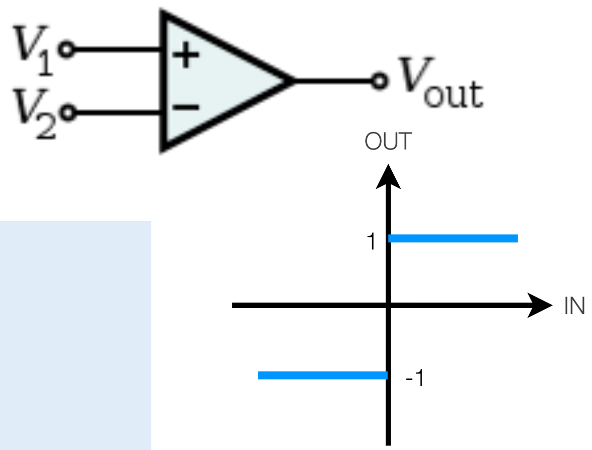
1. Context

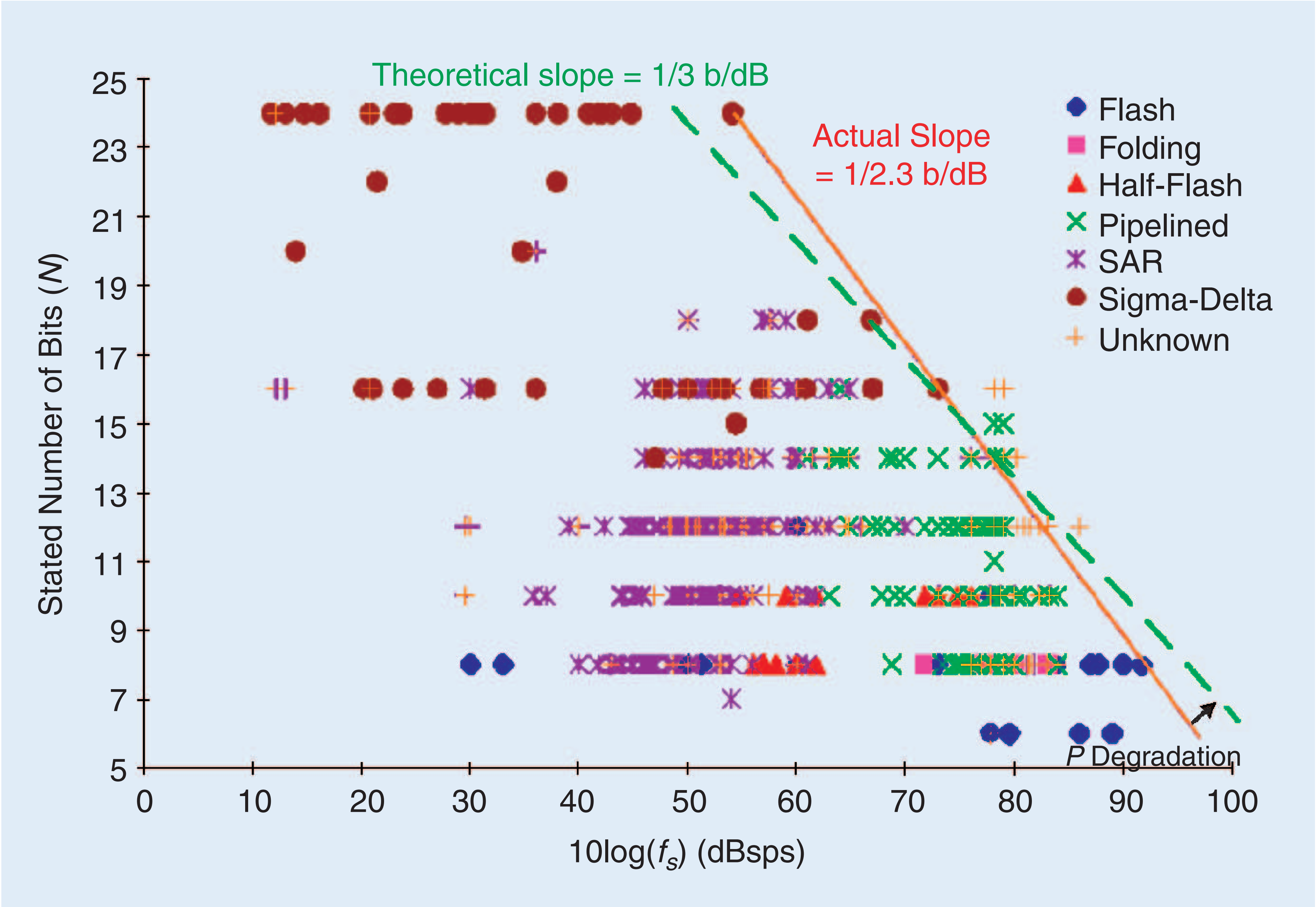
Central question: 1-bit sampling?



- Doable?
- For which "Sampling"?
- Which accuracy?

Why 1-bit? Very Fast Quantizers!

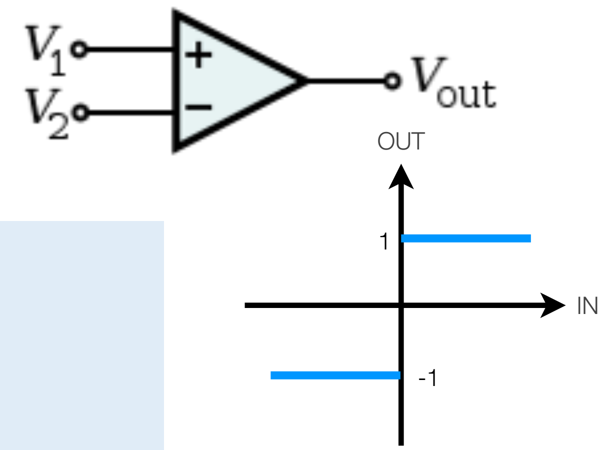


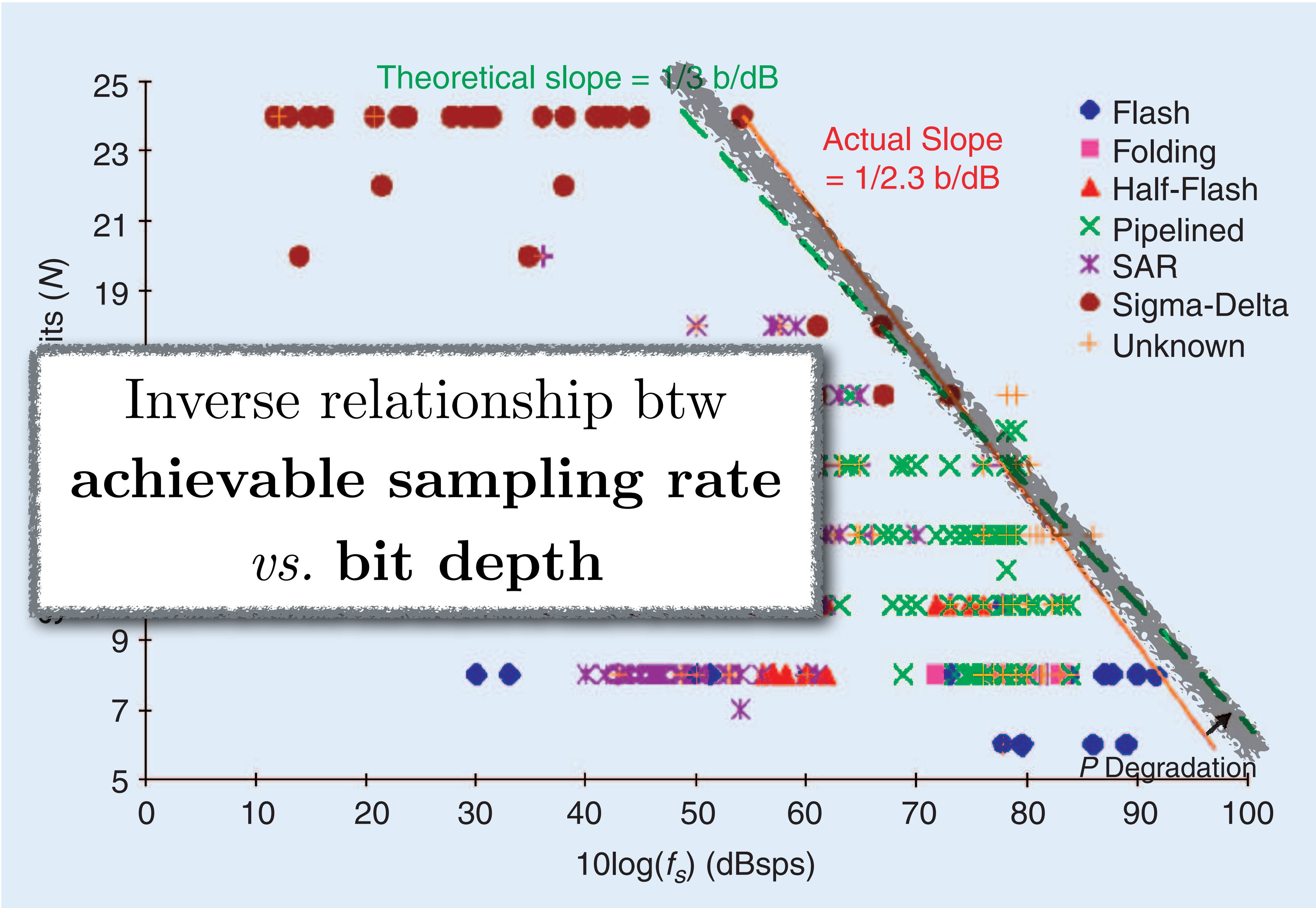


[FIG1] Stated number of bits versus sampling rate.

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

Why 1-bit? Very Fast Quantizers!



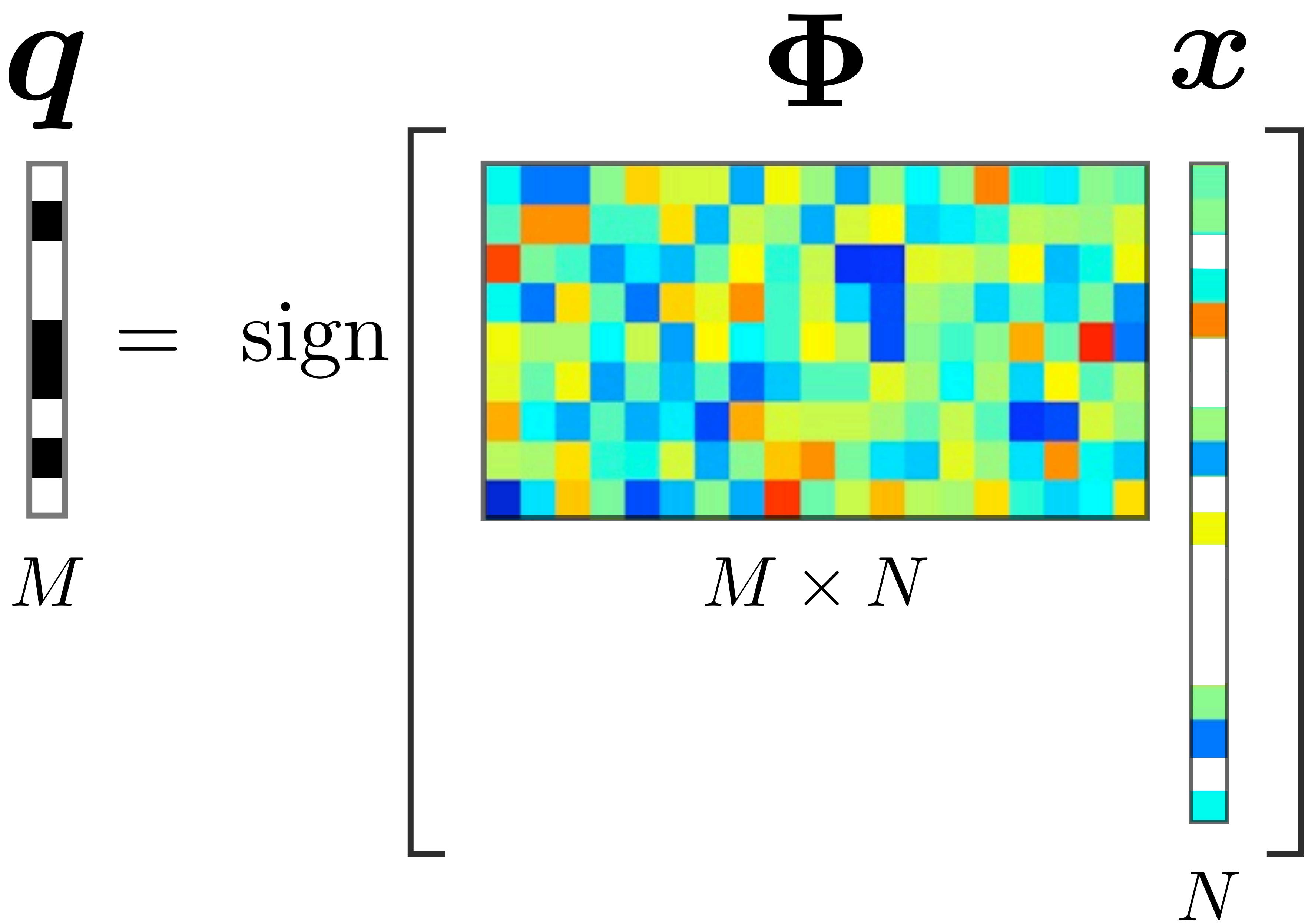


Inverse relationship btw **achievable sampling rate** *vs.* **bit depth**

[FIG1] Stated number of bits versus sampling rate.

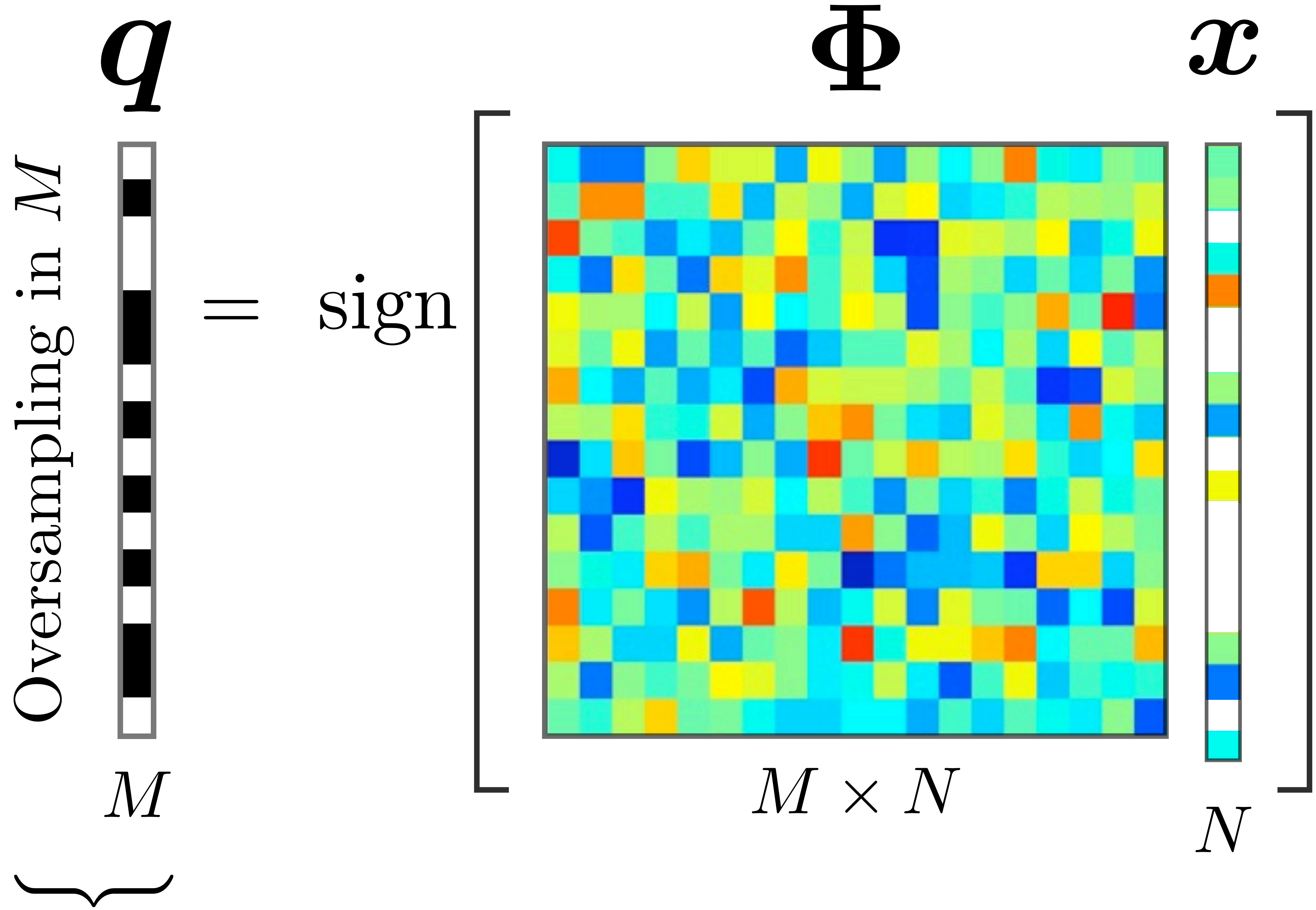
[From "Analog-to-digital converters" B. Le, T.W. Rondeau, J.H. Reed, and C.W.Bostian, IEEE Sig. Proc. Magazine, Nov 2005]

1-bit Compressed Sensing



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

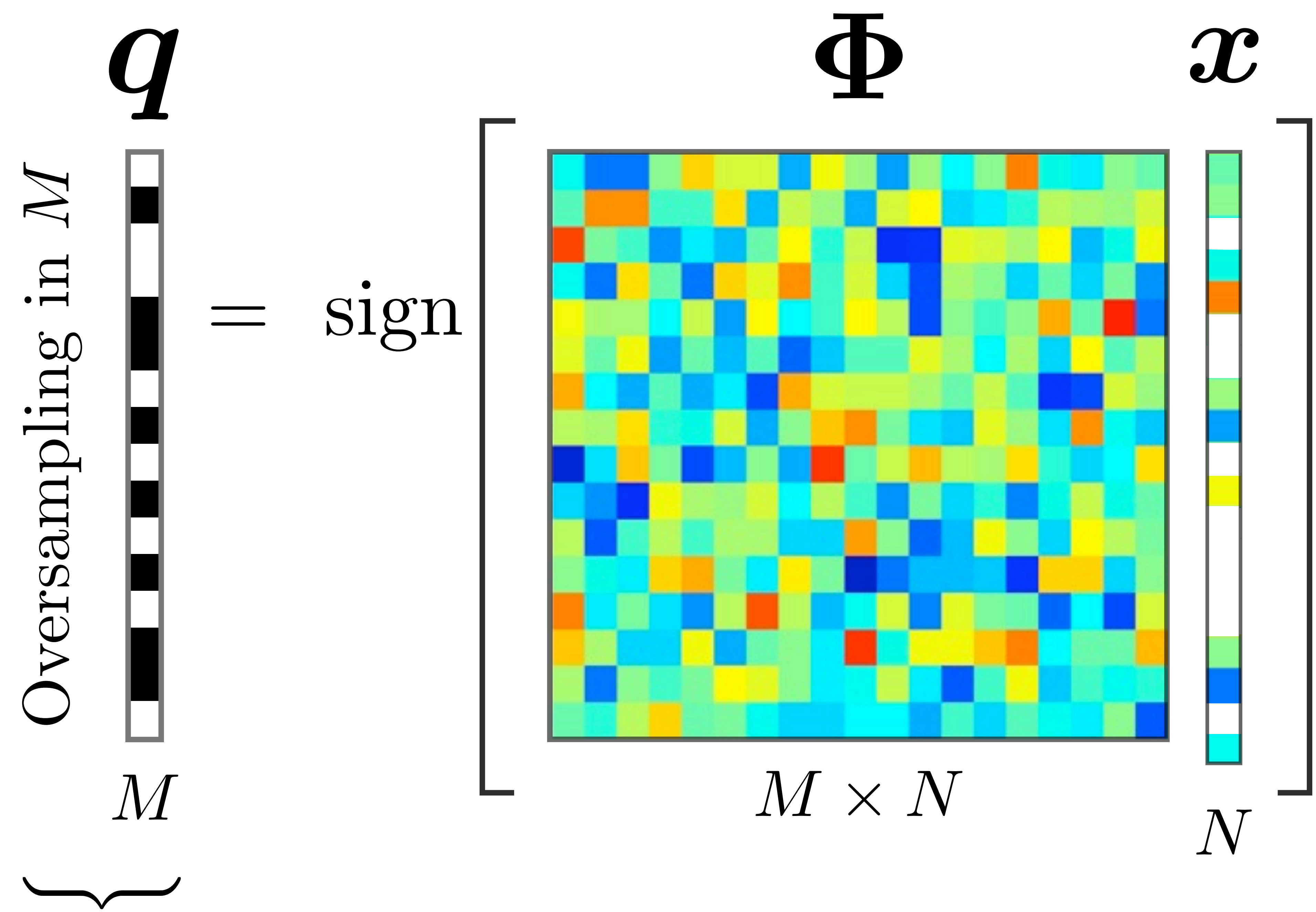
1-bit Compressed Sensing



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

1-bit Computational ~~Compressed~~ Sensing

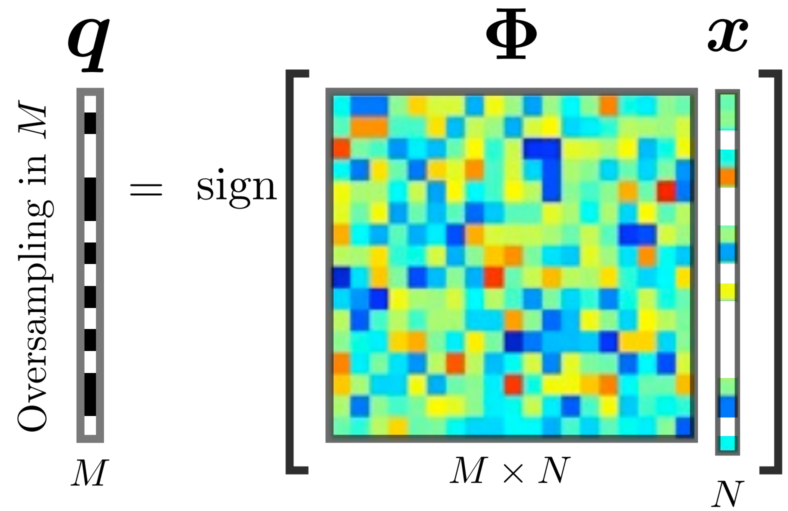
$\boxed{\text{bits matter!}}$



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

1-bit ~~Compressed~~ Computational Sensing

bits matter!



Warning 1: signal amplitude is lost!

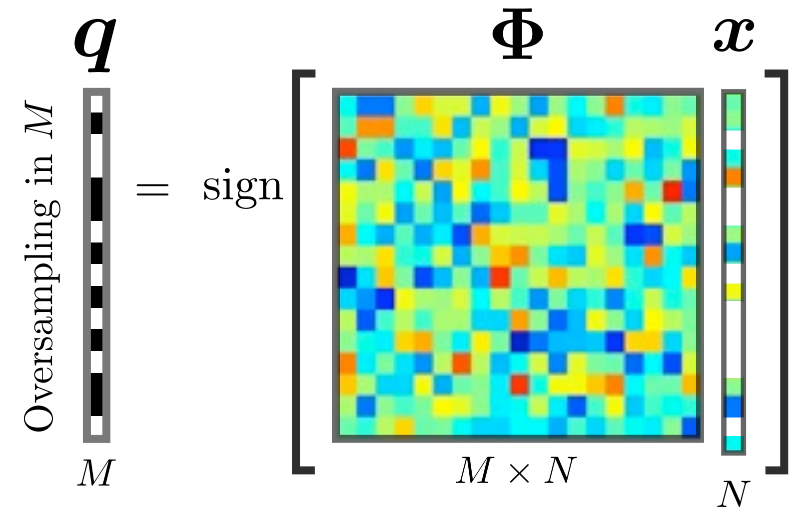
$$\boldsymbol{q} = \operatorname{sign}\left(\boldsymbol{\Phi}(\lambda \boldsymbol{x})\right) = \operatorname{sign}\left(\boldsymbol{\Phi}\boldsymbol{x}\right), \quad \forall \lambda > 0$$

$\Rightarrow$ Amplitude is arbitrarily fixed

Examples : $\|\boldsymbol{x}\| = 1$ or $\|\boldsymbol{\Phi}\boldsymbol{x}\|_1 = 1$

1-bit ~~Compressed~~ Computational Sensing

bits matter!



[Plan, Vershynin, 11]

Warning 2: $\exists$ forbidden sensing!

Let $\boldsymbol{x}_\lambda := (1, \lambda, 0, \cdots, 0)^T \in \mathbb{R}^N$

and $\boldsymbol{\Phi} \in \{\pm 1\}^{M \times N}$ (*e.g.*, Bernoulli).

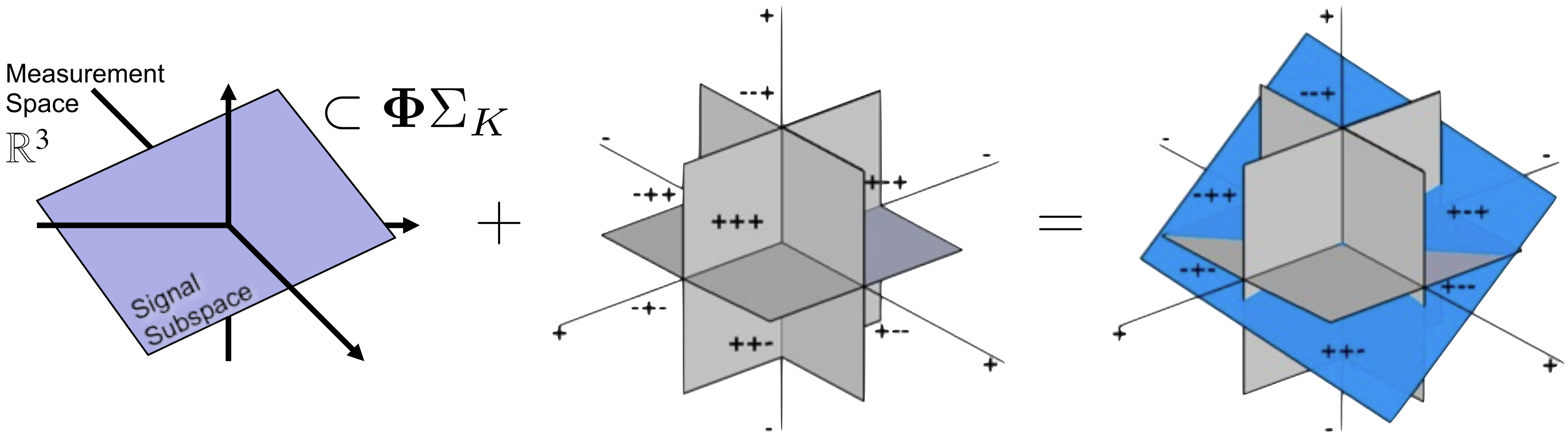
We have $\|\boldsymbol{x}_0 - \boldsymbol{x}_\lambda\| = \lambda$

but $\boldsymbol{q} = \operatorname{sign}(\boldsymbol{\Phi}\boldsymbol{x}_0) = \operatorname{sign}(\boldsymbol{\Phi}\boldsymbol{x}_\lambda),\ \forall |\lambda| < 1$

$\Rightarrow$ No hope to distinguish them by increasing M!

2. Theoretical performance limits

Lower bound: cell intersection viewpoint

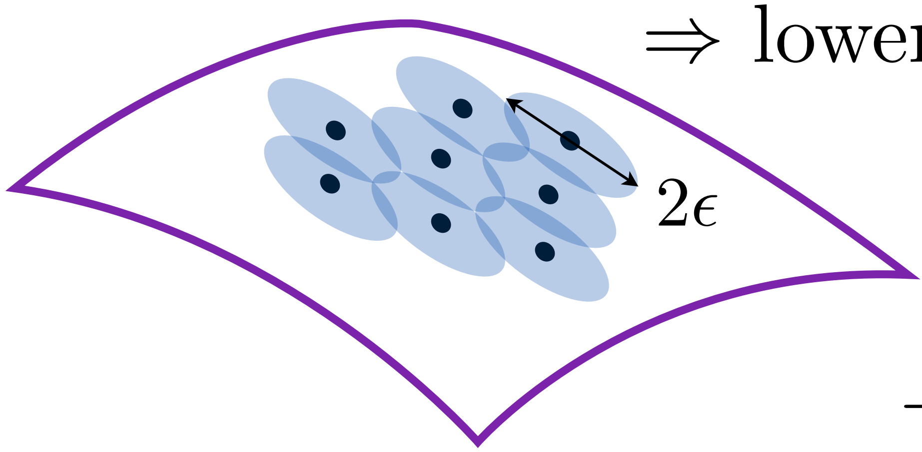


Not all quantization cells intersected!

no more than $C = 2^K \binom{N}{K}\binom{M}{K}$

Most efficient ϵ-covering of $S^{N-1} \cap \Sigma_K$ with ϵ-caps

$\Rightarrow$ lower bound on C by "$\mathrm{vol}(S^{N-1} \cap \Sigma_K)/\mathrm{vol}(\epsilon\text{-cap})$"

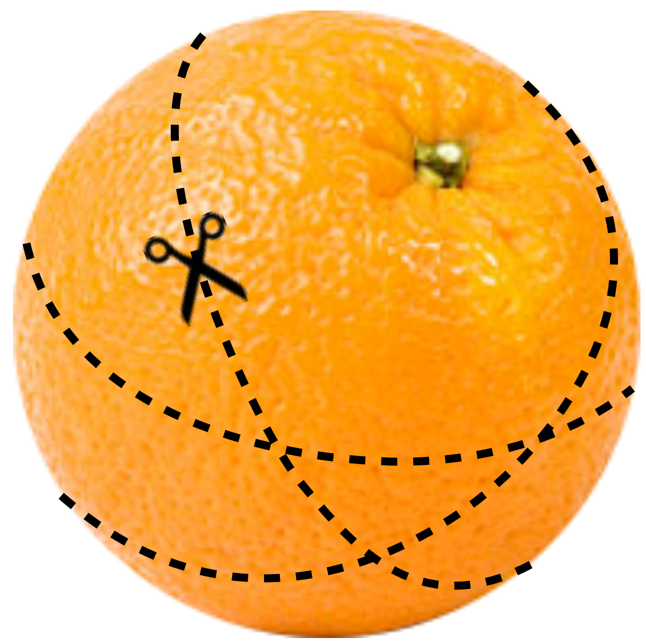


$$\Rightarrow \epsilon = \Omega(K/M)$$

$\rightarrow$ Lower bound on any 1-bit reconstruction error

Reaching this bound ?

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com



Reaching this bound ?

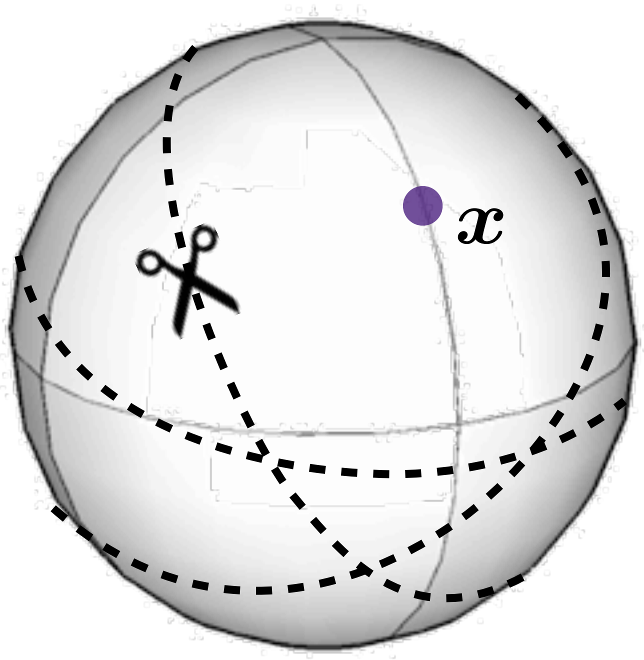
Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

$\boldsymbol{x}$ on S^2

M vectors:

$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

iid Gaussian



Reaching this bound ?

$\boldsymbol{x}$ on S^2

M vectors:

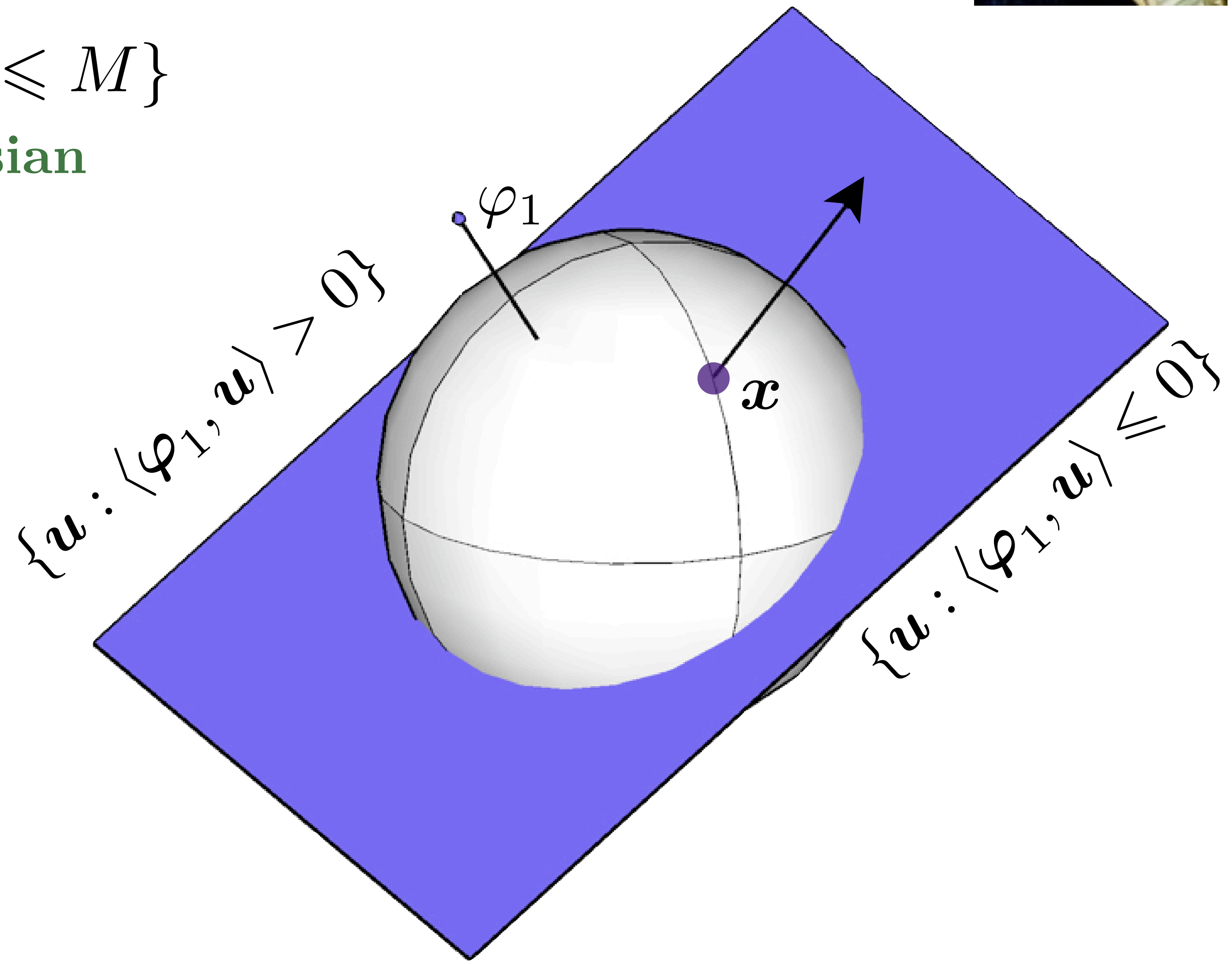
$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

iid Gaussian

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

1-bit Measurements

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



Reaching this bound ?

$\boldsymbol{x}$ on S^2

M vectors:

$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

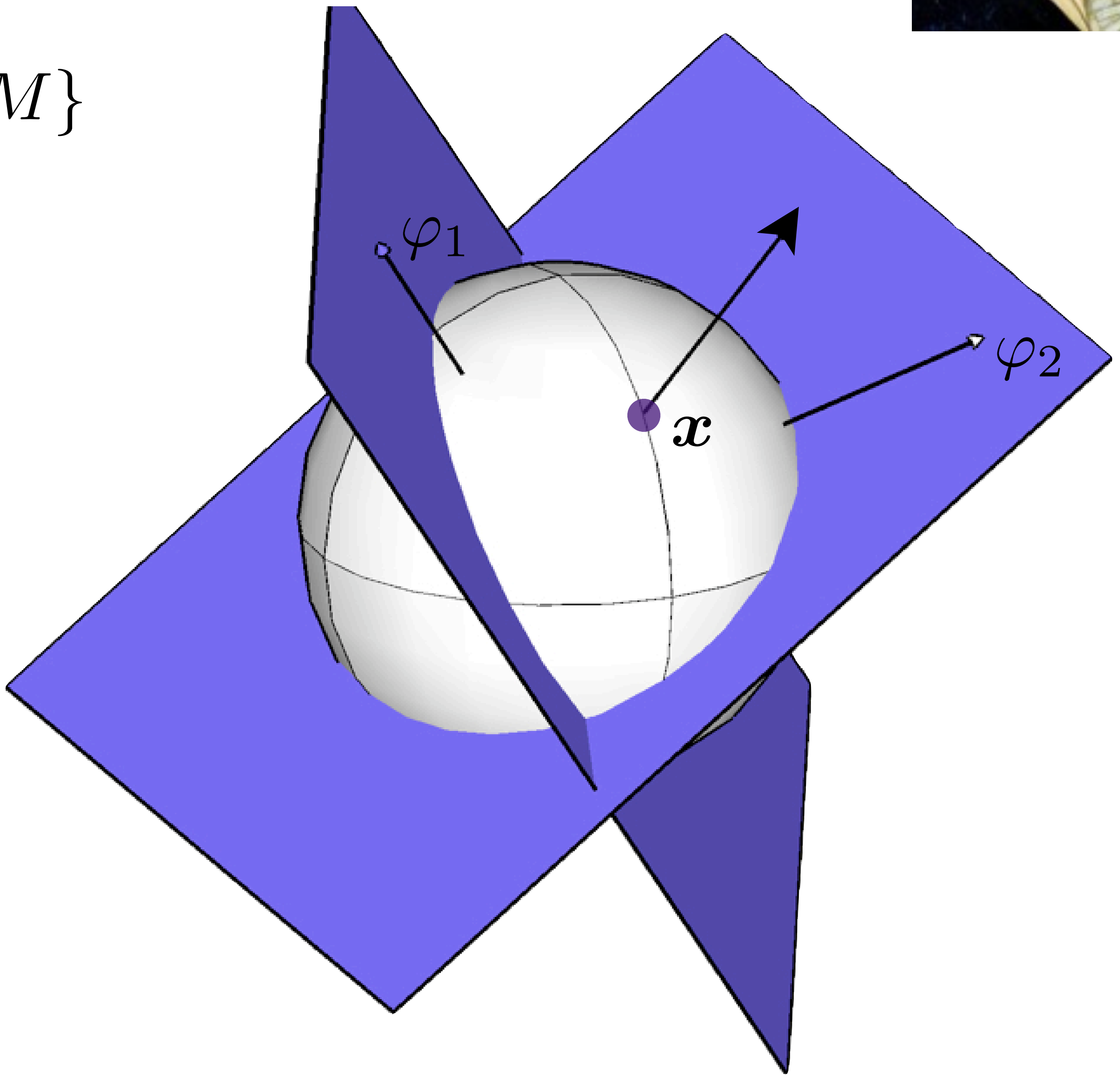
iid Gaussian

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

1-bit Measurements

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

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



Reaching this bound ?

$\boldsymbol{x}$ on S^2

M vectors:

$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

iid Gaussian

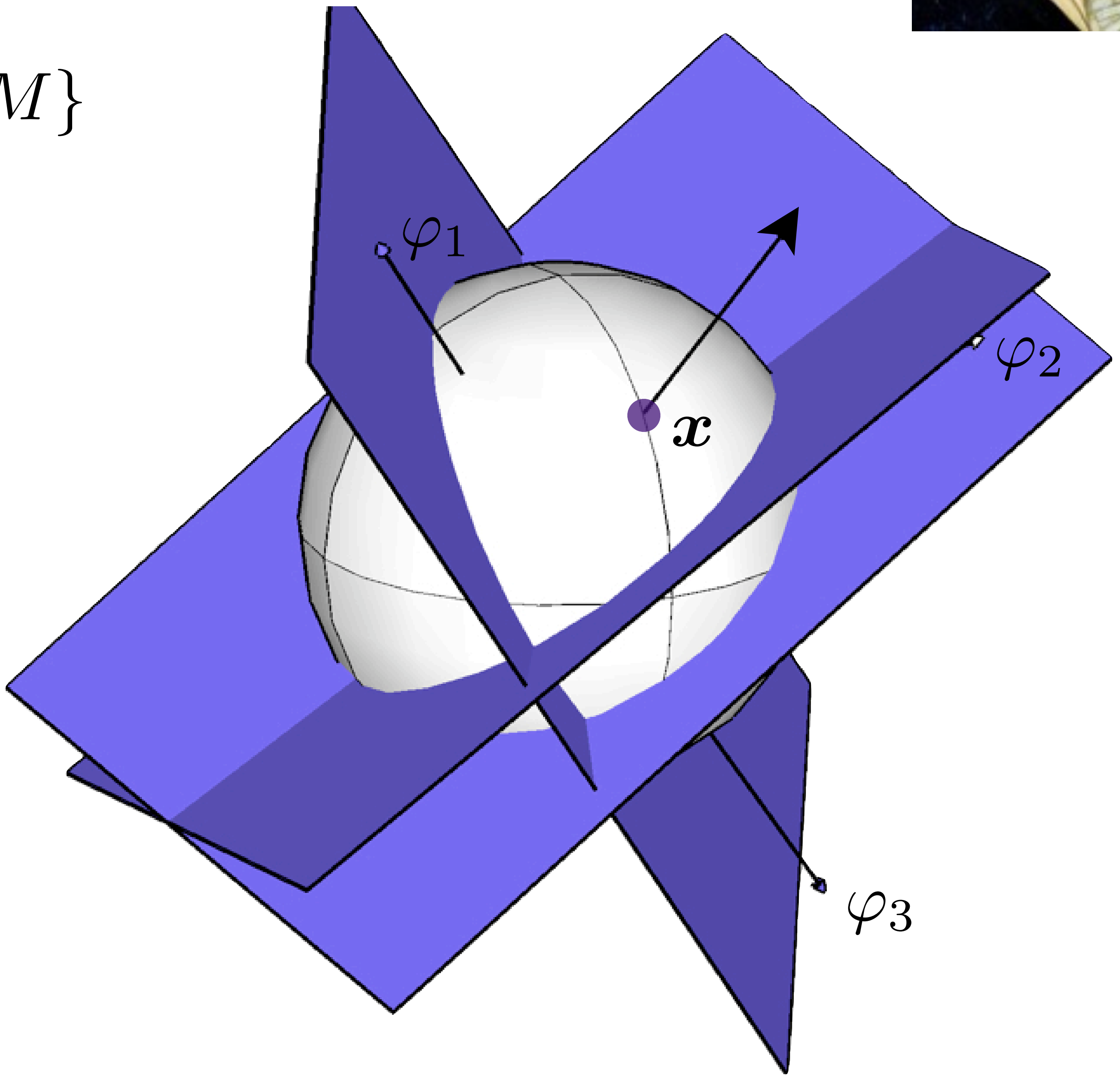
Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

1-bit Measurements

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

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

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



Reaching this bound ?

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

$\boldsymbol{x}$ on S^2

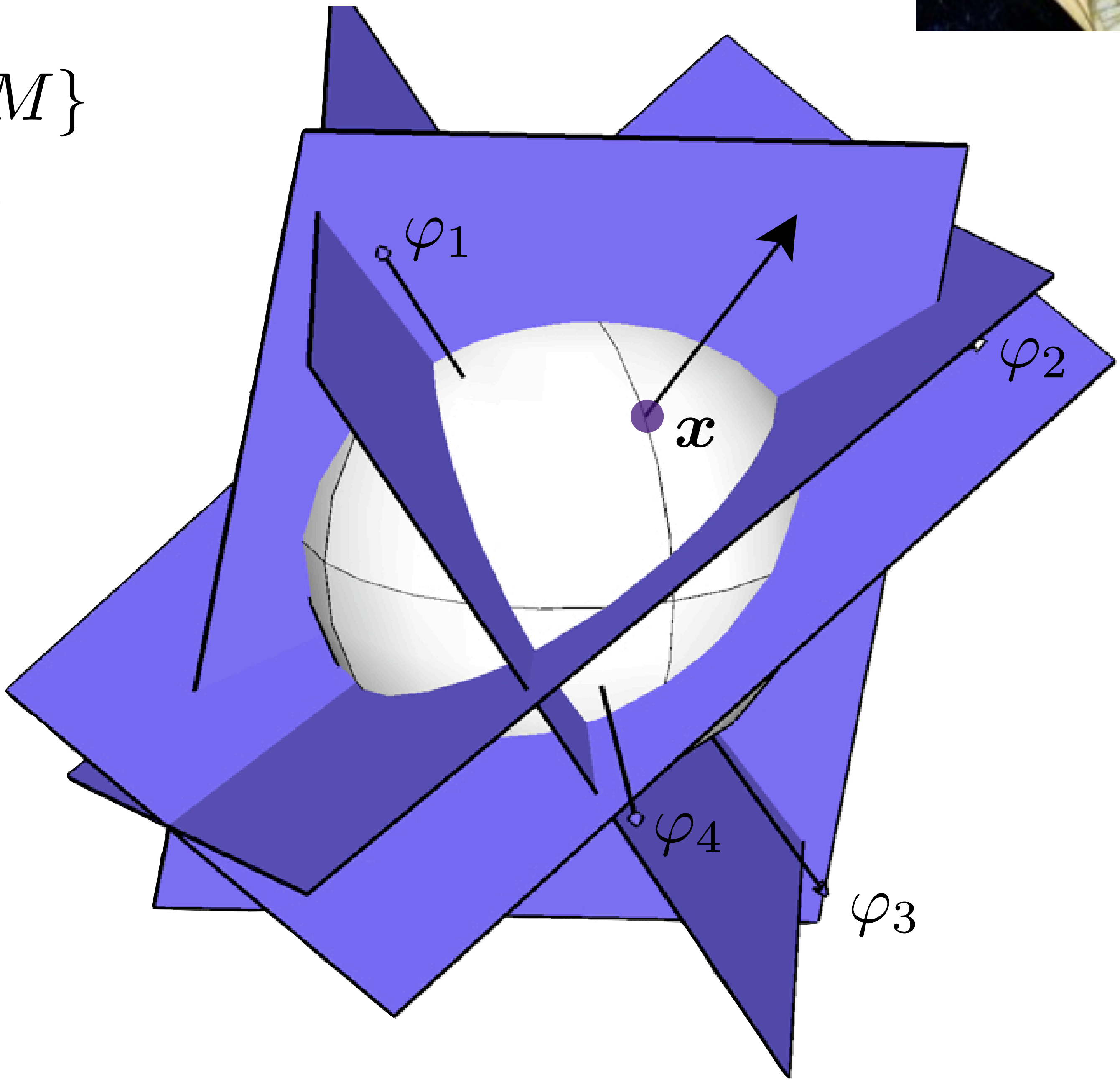
M vectors:

$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

iid Gaussian

1-bit Measurements

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



Reaching this bound ?

$\boldsymbol{x}$ on S^2

M vectors:

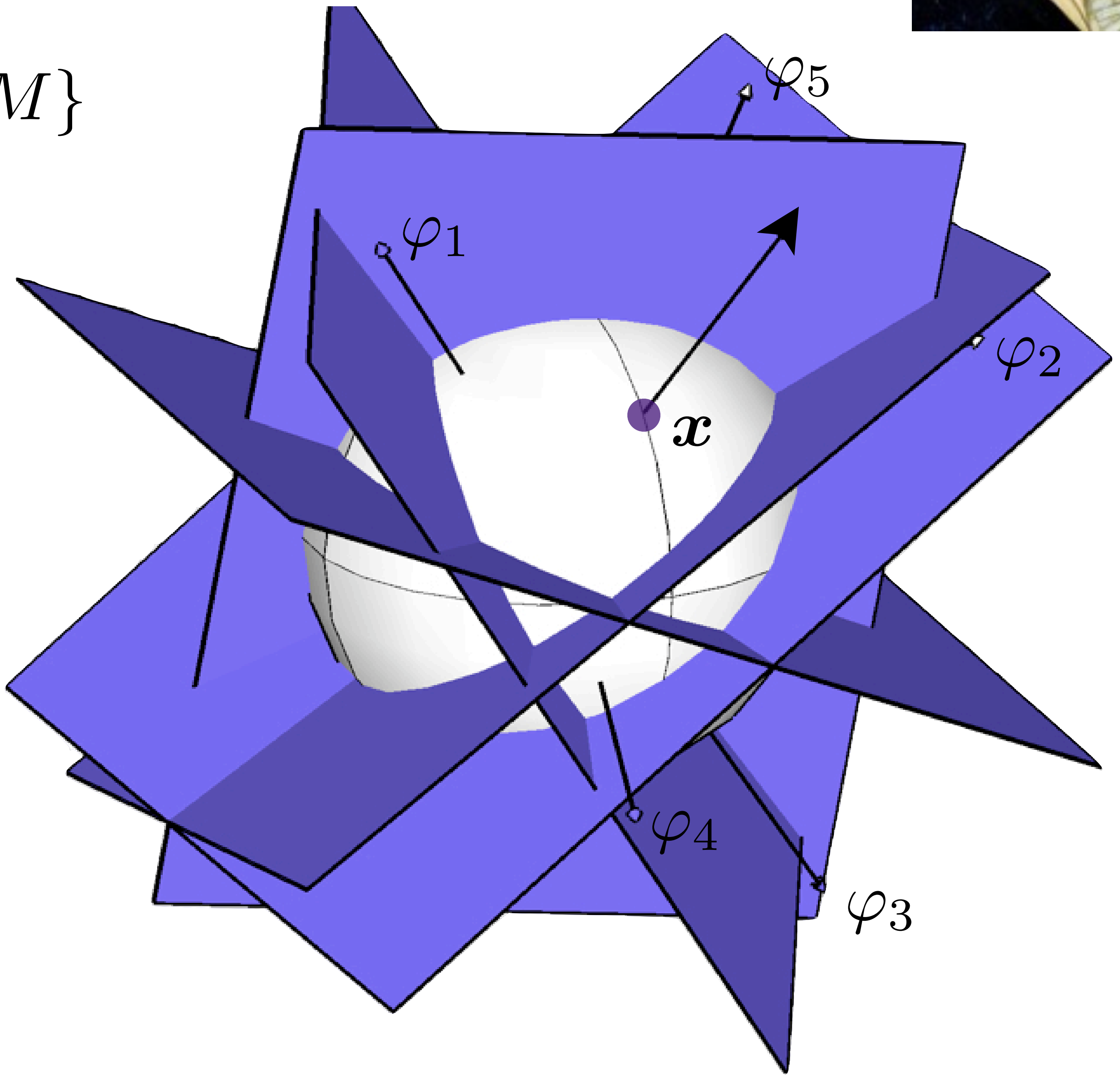
$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

iid Gaussian

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

1-bit Measurements

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



Reaching this bound ?

$\boldsymbol{x}$ on S^2

M vectors:

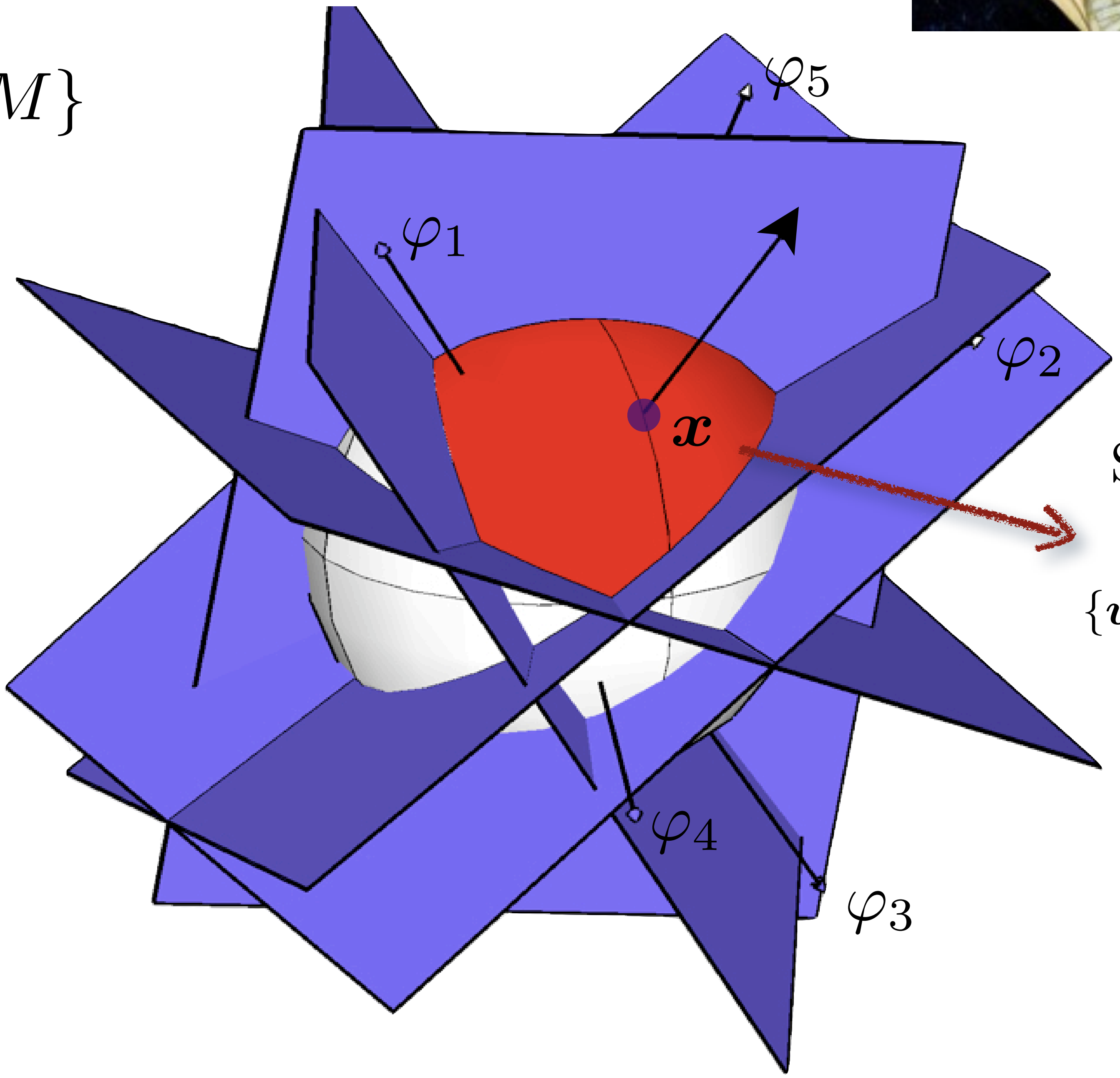
$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

iid Gaussian

1-bit Measurements

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

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com



Smaller and smaller when M increases

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

Reaching this bound ?

$\boldsymbol{x}$ on S^2

M vectors:

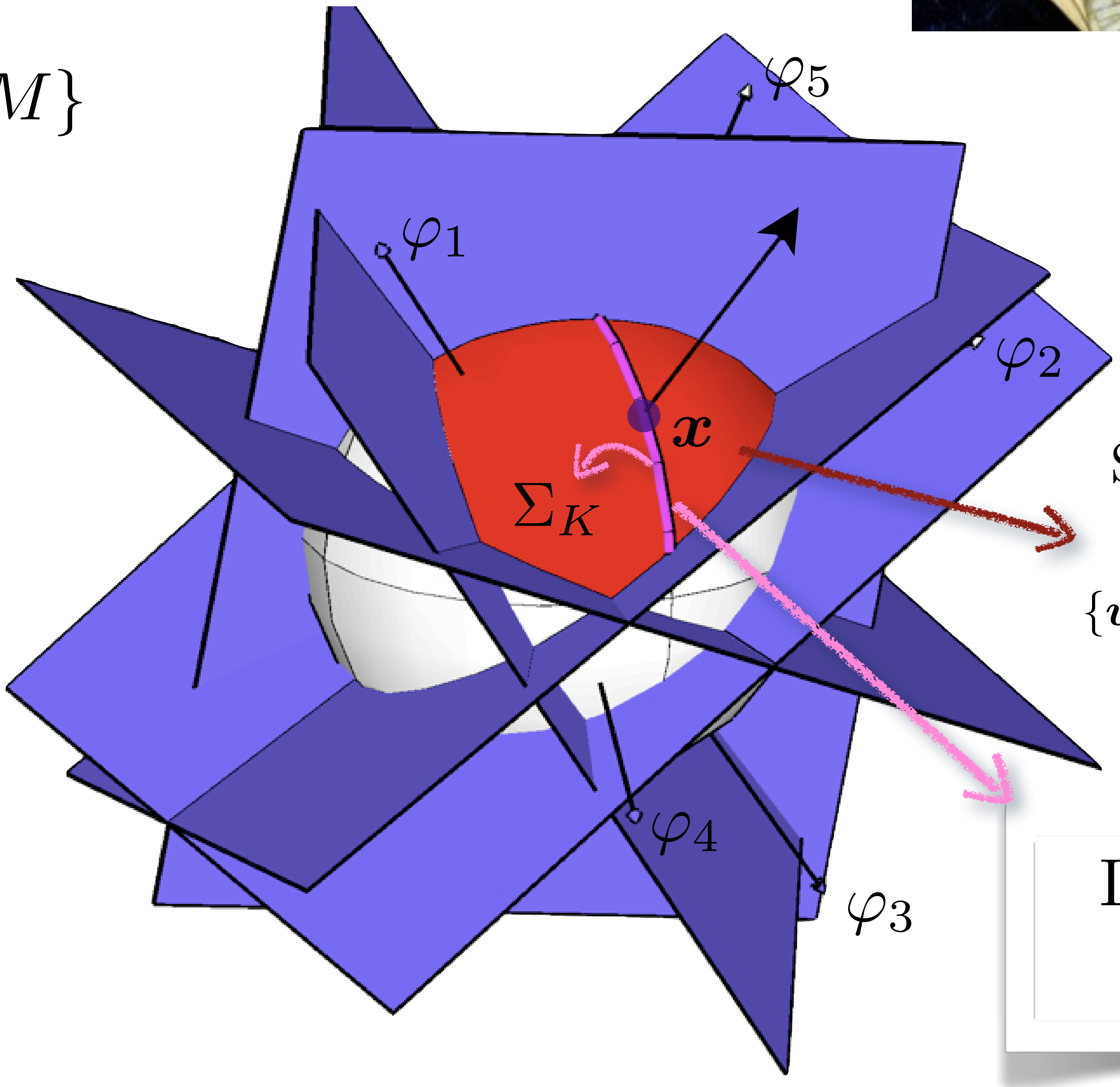
$\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

iid Gaussian

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

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

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

Lower bound on this width?

Reaching this bound ?

Carl Friedrich Gauss:
"1-bit CS? I solved it at breakfast by randomly slicing my orange!"
http://www.gaussfacts.com

Let $A(\cdot) := \mathrm{sign}\,(\boldsymbol{\Phi}\,\cdot)$ with $\boldsymbol{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$.

If $M = O(\epsilon^{-1} K \log N)$, then, w.h.p,

for any two unit K-sparse vectors $\boldsymbol{x}$ and $\boldsymbol{s}$,

$$A(\boldsymbol{x}) = A(\boldsymbol{s}) \quad \Rightarrow \quad \|\boldsymbol{x} - \boldsymbol{s}\| \leq \epsilon$$

$$\Leftrightarrow \epsilon = O\big(\tfrac{K}{M} \log \tfrac{MN}{K}\big)$$

almost optimal

Note: You can even afford a small error, *i.e.*,

if only b bits are different between $A(\boldsymbol{x})$ and $A(\boldsymbol{s})$ $\Rightarrow \|\boldsymbol{x} - \boldsymbol{s}\| \leqslant \frac{K+b}{K}\,\epsilon$

3. Stable embeddings: angles are preserved

Starting point: Hamming/Angle Concentration

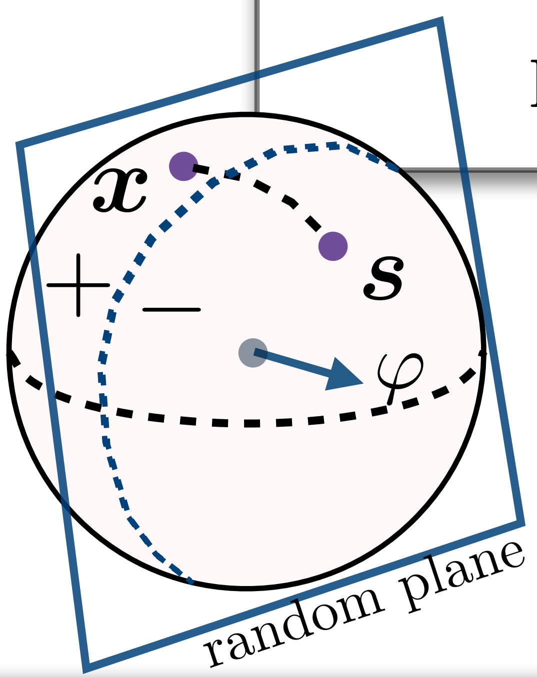
- Metrics of interest:

$$d_H(\boldsymbol{u}, \boldsymbol{v}) = \tfrac{1}{M} \textstyle\sum_i (u_i \oplus v_i) \quad \text{(norm. Hamming)}$$
$$d_{\rm ang}(\boldsymbol{x}, \boldsymbol{s}) = \tfrac{1}{\pi} \arccos(\langle \boldsymbol{x}, \boldsymbol{s} \rangle) \quad \text{(norm. angle)}$$

- Known fact: if $\boldsymbol{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$ [*e.g.*, Goemans, Williamson 1995]

Let $\boldsymbol{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$, $A(\cdot) = \mathrm{sign}\,(\boldsymbol{\Phi}\,\cdot) \in \{-1,1\}^M$ and $\epsilon > 0$.
For any $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$, we have

$$\mathbb{P}_{\boldsymbol{\Phi}}\Big[\, \big|\, d_H\big(A(\boldsymbol{x}), A(\boldsymbol{s})\big) \;-\; d_{\rm ang}(\boldsymbol{x}, \boldsymbol{s}) \big| \;\leqslant\; \epsilon \,\Big] \;\geqslant\; 1 - 2\, e^{-2\epsilon^2 M}.$$



Thanks to $A(.)$, Hamming distance concentrates around vector angles!

Binary ϵ Stable Embedding (BϵSE)

A mapping $A : \mathbb{R}^N \to \{\pm 1\}^M$ is a **binary ϵ-stable embedding** (BϵSE) of order K for sparse vectors if

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

for all $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$ with $\boldsymbol{x} \pm \boldsymbol{s}$ K-sparse.

kind of "binary restricted (quasi) isometry"

- *Corollary*: for any algorithm with output $\boldsymbol{x}^*$ jointly K-sparse and consistent (*i.e.*, $A(\boldsymbol{x}^*) = A(\boldsymbol{x})$),

$$d_{\rm ang}(\boldsymbol{x}, \boldsymbol{x}^*) \leqslant 2\epsilon!$$

- If limited binary noise, $d_{\rm ang}$ still bounded
- If not exactly sparse signals (but almost), $d_{\rm ang}$ still bounded

BϵSE existence? Yes!

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

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

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

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

Proof sketch:

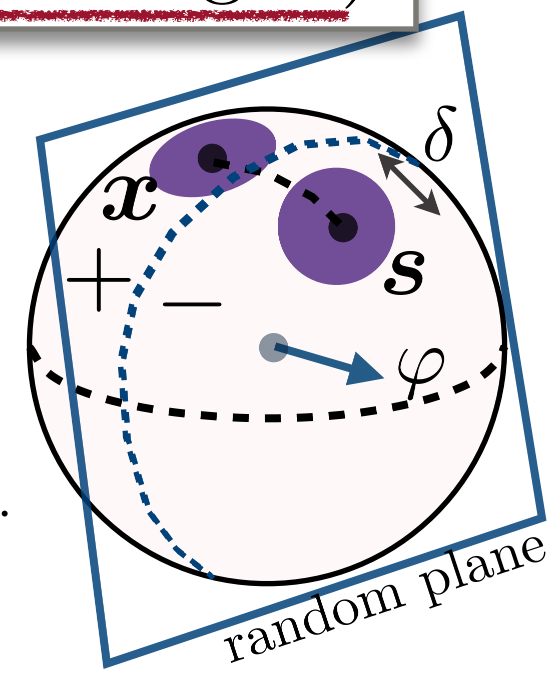
1) Generalize

$$\mathbb{P}_{\mathbf{\Phi}}\,\big[\,\big|\,d_H\big(A(\boldsymbol{x}), A(\boldsymbol{s})\big) \;-\; d_{\rm ang}(\boldsymbol{x},\boldsymbol{s})\big| \;\leqslant\; \epsilon\,\big] \;\geqslant\; 1 - 2\,e^{-2\epsilon^2 M}.$$

to

$$\mathbb{P}_{\mathbf{\Phi}}\,\Big[\,\big|\,d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big) \;-\; d_{\rm ang}(\boldsymbol{x},\boldsymbol{s})\big| \;\leqslant\; \epsilon + (\tfrac{\pi}{2}D)^{1/2}\,\delta\,\Big] \;\geqslant\; 1 - 2\,e^{-2\epsilon^2 M}.$$

for $\boldsymbol{u}, \boldsymbol{v}$ in a D-dimensional neighborhood of width δ around $\boldsymbol{x}$ and $\boldsymbol{s}$ resp.



2) Covers the space of "K-sparse signal pairs" in $\mathbb{R}^N$ by

$$O(\tbinom{N}{K}\delta^{-2K}) = O((\tfrac{eN}{K\delta^2})^K) \text{ neighborhoods.}$$

3) Apply Point 1 with union bound, and "stir until the proof thickens"

BϵSE existence? Yes!

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

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

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

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

$\Rightarrow$

BϵSE consistency "width":

$$\epsilon \;=\; O\big((\tfrac{K}{M}\log\tfrac{MN}{K})^{1/2}\big)$$

not as optimal but stronger result!

$d_H \leftrightarrow d_{\rm ang}$

4. Generalized Embeddings

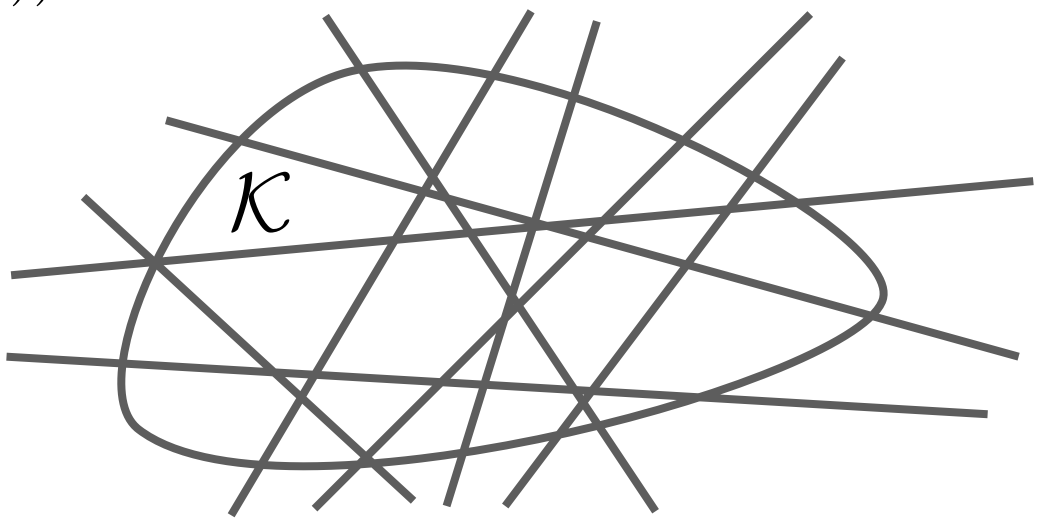
Beyond strict sparsity ...

Let $\mathcal{K} \subset S^{N-1}$ (*e.g.*, compressible signals s.t. $\|\boldsymbol{x}\|_2/\|\boldsymbol{x}\|_1 \leqslant \sqrt{K}$)
$\neq \Sigma_K$

What can we say on $d_H(A(\boldsymbol{x}), A(\boldsymbol{s}))$ for $\boldsymbol{x}, \boldsymbol{s} \in \mathcal{K}$?

Uniform tesselation: [Plan, Vershynin, 11]

$$\mathrm{P}\big(\underbrace{\#\text{ random hyperplanes btw } \boldsymbol{x}}_{d_H(A(\boldsymbol{x}), A(\boldsymbol{s}))} \text{ and } \boldsymbol{s} \propto d_{\mathrm{ang}}(\boldsymbol{x}, \boldsymbol{s})\big)\ ?$$

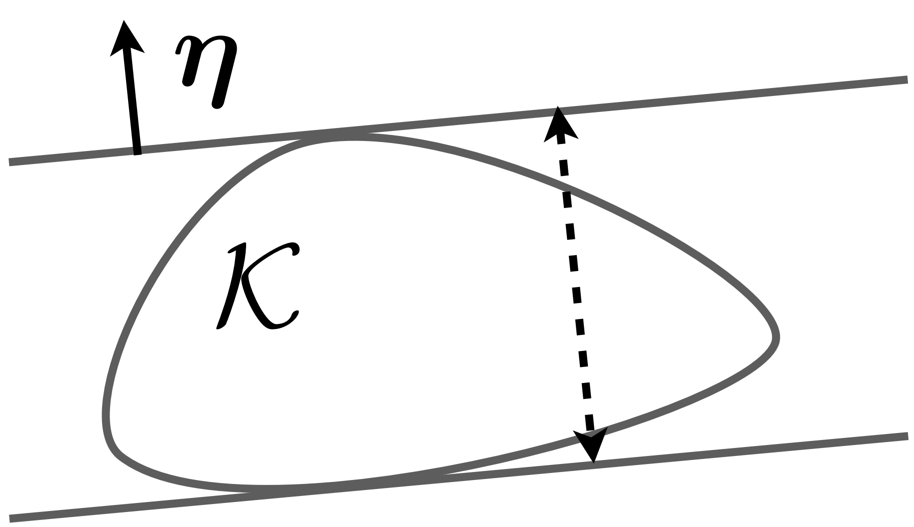


Y. Plan, R. Vershynin, "Dimension reduction by random hyperplane tessellations", 2011, arXiv:1111.4452
Y. Plan, R. Vershynin, "Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach", IEEE TIT 2012, arXiv:1202.1212.

Beyond strict sparsity ...

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

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



width in direction $\boldsymbol{\eta}$

<u>Examples</u>:

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

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

$$w^2(\mathcal{K}) \leqslant L \quad \text{if subspace with } \dim \mathcal{K} = L$$

$$w^2(\Sigma_K) \simeq K \log(2N/K)$$

$$\vdots$$

Y. Plan, R. Vershynin, "Dimension reduction by random hyperplane tessellations", 2011, arXiv:1111.4452

Y. Plan, R. Vershynin, "Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach", IEEE TIT 2012, arXiv:1202.1212.

Beyond strict sparsity ...

Proposition *Let* $\mathbf{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$ *and* $\mathcal{K} \subset \mathbb{R}^N$*. Then, for some* $C, c > 0$*, if*

$$M \geqslant C\epsilon^{-6} w^2(\mathcal{K}),$$

not as optimal but stronger result!

then, with $Pr \geqslant 1 - e^{-c\epsilon^2 M}$*, we have*

$$d_{\rm ang}(\boldsymbol{x}, \boldsymbol{s}) - \epsilon \leqslant d_H(A(\boldsymbol{x}), A(\boldsymbol{s})) \leqslant d_{\rm ang}(\boldsymbol{x}, \boldsymbol{s}) - \epsilon, \quad \forall \boldsymbol{x}, \boldsymbol{s} \in \mathcal{K}.$$

Generalize BϵSE to more general sets.

In particular, to

$$\mathcal{C}_K = \{\boldsymbol{u} \in \mathbb{R}^N : \|\boldsymbol{u}\|_2/\|\boldsymbol{u}\|_1 \leqslant \sqrt{K}\} \supset \Sigma_K$$

with $w^2(\mathcal{C}_K) \leqslant cK \log N/K$.

$\Rightarrow$ Extension to "1-bit Matrix Completion" possible!

i.e., $w^2(r\text{-rank } N_1 \times N_2 \text{ matrix}) \leqslant c\, r(N_1 + N_2)$!

Y. Plan, R. Vershynin, "Dimension reduction by random hyperplane tessellations", 2011, arXiv:1111.4452

Y. Plan, R. Vershynin, "Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach", IEEE TIT 2012, arXiv:1202.1212.

5. 1-bit CS Reconstructions?

Dumbest 1-bit reconstruction

Fact: If $M = O(\epsilon^{-2} K \log N/K)$ (for $\boldsymbol{x} \in \Sigma_K$ fixed, $\forall\, \boldsymbol{s} \in \Sigma_K$)
or, if $M = O(\epsilon^{-6} K \log N/K)$ ($\forall\, \boldsymbol{x}, \boldsymbol{s} \in \Sigma_K$), then, w.h.p,

$$\left|\tfrac{\sqrt{\pi/2}}{M} \langle \mathrm{sign}\,(\boldsymbol{\Phi x}), \boldsymbol{\Phi s}\rangle - \langle \boldsymbol{x}, \boldsymbol{s}\rangle\right| \leq \epsilon \qquad \text{[Plan, Vershynin, 12]}$$

- Implication? [LJ, Degraux, De Vleeschouwer, 13]

Let $\boldsymbol{x} \in \Sigma_K \cap S^{N-1}$ and $\boldsymbol{q} = \mathrm{sign}(\boldsymbol{\Phi x})$.
Compute

$$\boxed{\hat{\boldsymbol{x}} = \tfrac{\pi}{2M}\, \mathcal{H}_K(\boldsymbol{\Phi}^* \boldsymbol{q})}$$

Then, if previous property holds,

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq 2\epsilon.$$

Non-uniform case ($\boldsymbol{x}$ given):
$\Rightarrow \epsilon \;=\; O\big((\tfrac{K}{M} \log \tfrac{MN}{K})^{1/2}\big)$
Uniform case:
$\Rightarrow \epsilon \;=\; O\big((\tfrac{K}{M} \log \tfrac{MN}{K})^{1/6}\big)$

Y. Plan, R. Vershynin, "Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach", IEEE TIT 2012, arXiv:1202.1212.
LJ, K. Degraux, C. De Vleeschouwer, "Quantized Iterative Hard Thresholding: Bridging 1-bit and High-Resolution Quantized Compressed Sensing", SAMPTA2013

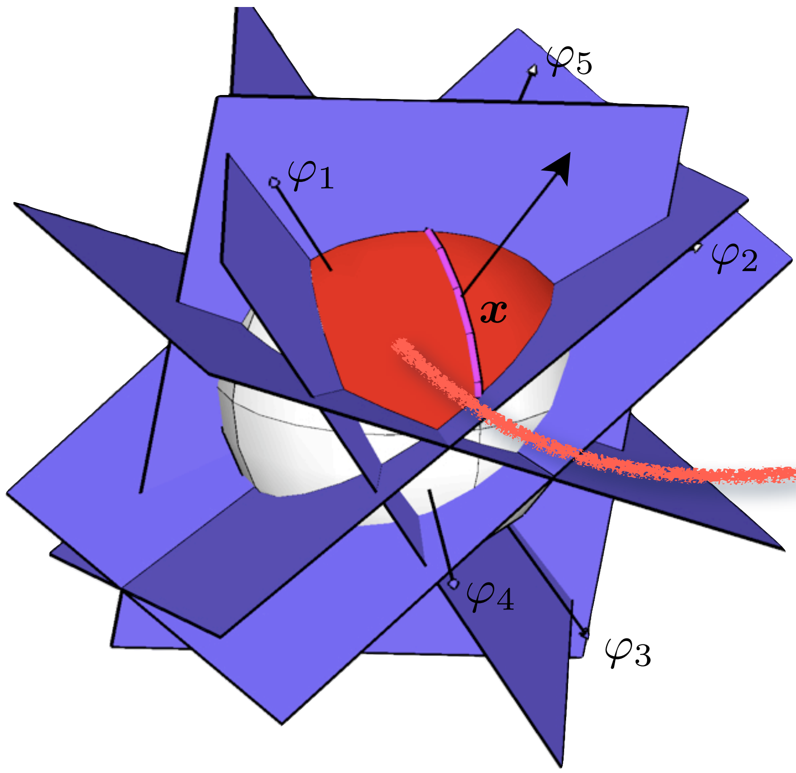
Initial approach

- Let $\boldsymbol{q} = \mathrm{sign}\,(\boldsymbol{\Phi x}) =: A(\boldsymbol{x})$
- Initially: [Boufounos, Baraniuk 2008]

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

<u>**Non-convex**</u>! 2 numerical choices :

1. relax + projection on $S^{N\text{-}1}$
2. *"trust region methods"*
   $\rightarrow$ *Restricted-Step Shrinkage (RSS)*



<u>**Consistency constraint**</u>:

$$\{\boldsymbol{u} \in \mathbb{R}^N \cap S^{N-1} : \boldsymbol{q} = A(\boldsymbol{u})\}$$

$$\Leftrightarrow \{\boldsymbol{u} \in \mathbb{R}^N \cap S^{N-1} : \mathrm{diag}(\boldsymbol{q})\boldsymbol{\Phi u} > 0\}$$

$$\ni \boldsymbol{x}$$

Initial approach

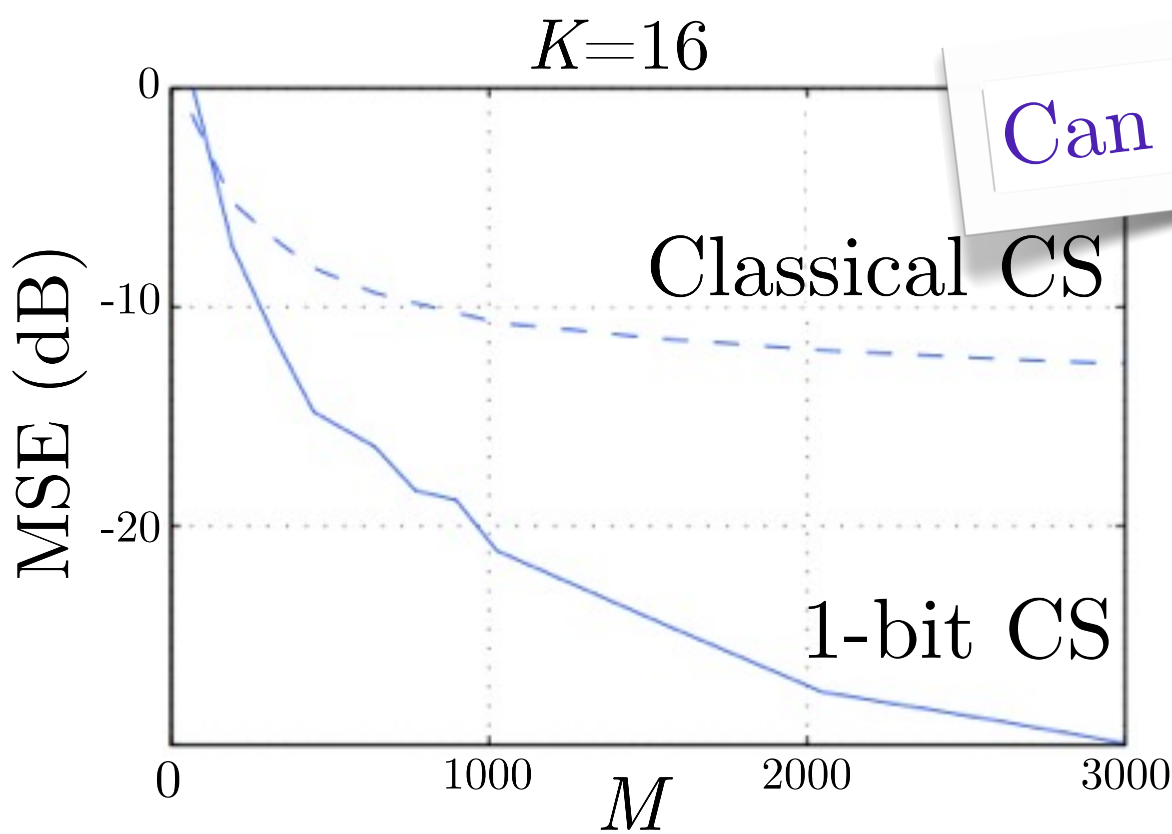
- Let $\boldsymbol{q} = \mathrm{sign}\,(\boldsymbol{\Phi x}) =: A(\boldsymbol{x})$
- Initially: [Boufounos, Baraniuk 2008]

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

(*e.g.*, take the 1st choice)

(relaxed) $\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{u}} \|\boldsymbol{u}\|_1 + \lambda \|(\mathrm{diag}(\boldsymbol{q})\,\boldsymbol{\Phi u})_-\|^2 \quad \text{s.t.} \quad \|\boldsymbol{u}\|_2 = 1$

$\to$ Solved by projected gradient descent



Can we do better?

gain brought by (almost) consistency

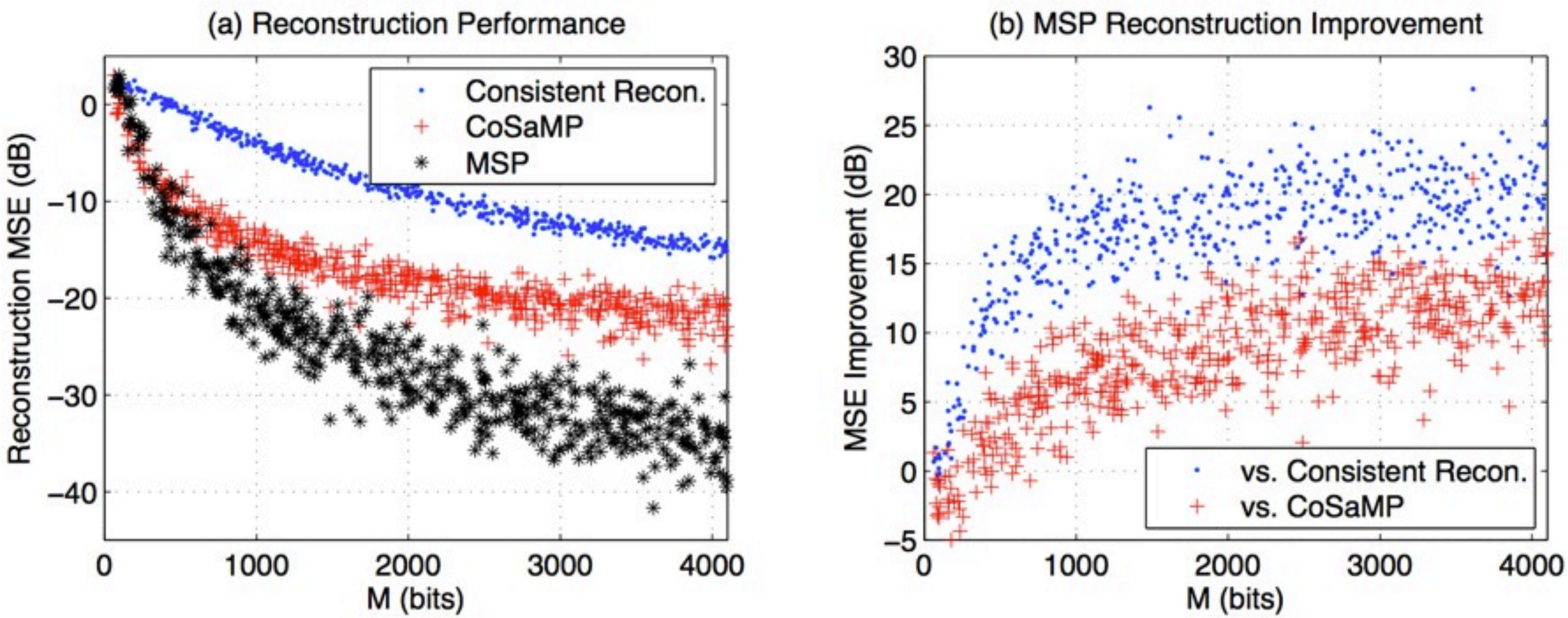
Other methods:

- Matching Sign Pursuit [Boufounos]
- Restricted-Step Shrinkage (RSS) [Laska, We, Yin, Baraniuk]
- Binary Iterative Hard Thresholding [Jacques, Laska, Boufounos, Baraniuk]
- Convex Optimization [Plan, Vershynin]
- …

Matching Sign Pursuit (MSP)

- Iterative greedy algorithm, similar to CoSaMP [Needell, Tropp, 08]
- Maintains running signal estimate and its support T.
- <u>MSP iteration</u>:
  - Identify sign violations $\longrightarrow \boldsymbol{r} = (\mathrm{diag}(\boldsymbol{y})\,\boldsymbol{\Phi}\widehat{\boldsymbol{x}})_-$
  - Compute proxy $\longrightarrow \boldsymbol{p} = \boldsymbol{\Phi}^T\boldsymbol{r}$
  - Identify support $\longrightarrow \Omega = \mathrm{supp}\,\boldsymbol{p}|_{2K} \cup T$
  - Consistent Reconstruction over support estimate:
  $$\boldsymbol{b}|_\Omega = \arg\min_{\boldsymbol{u}\in\mathbb{R}^N} \|(\mathrm{diag}(\boldsymbol{y})\boldsymbol{\Phi}\boldsymbol{u})_-\|_2^2 \text{ s.t } \|\boldsymbol{u}\|_2 = 1 \text{ and } \boldsymbol{u}|_{T^c} = 0$$
  - Truncate, normalize, and update estimate: $\widehat{\boldsymbol{x}} \leftarrow \boldsymbol{b}|_K \,/\, \|\boldsymbol{b}|_K\|_2$

Matching Sign Pursuit (MSP)



Boufounos, P. T. (2009, November). “Greedy sparse signal reconstruction from sign measurements”.
In Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on (pp. 1305-1309). IEEE.

Binary Iterative Hard Thresholding

Given $\boldsymbol{q} = A(\boldsymbol{x})$ and K, set $l = 0$, $\boldsymbol{x}^0 = 0$:

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

("gradient" towards consistency)

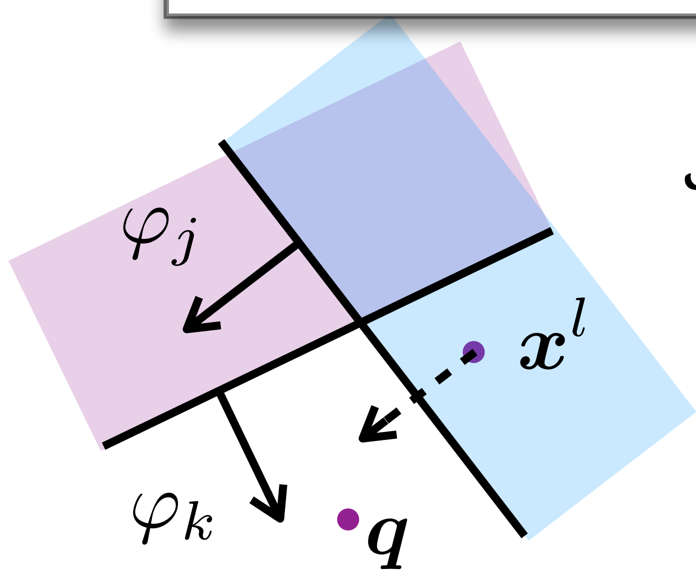
($\tau > 0$ controls gradient descent)

(proj. K-sparse signal set)

with $\mathcal{H}_K(\boldsymbol{u}) = K$-term hard thresholding

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

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



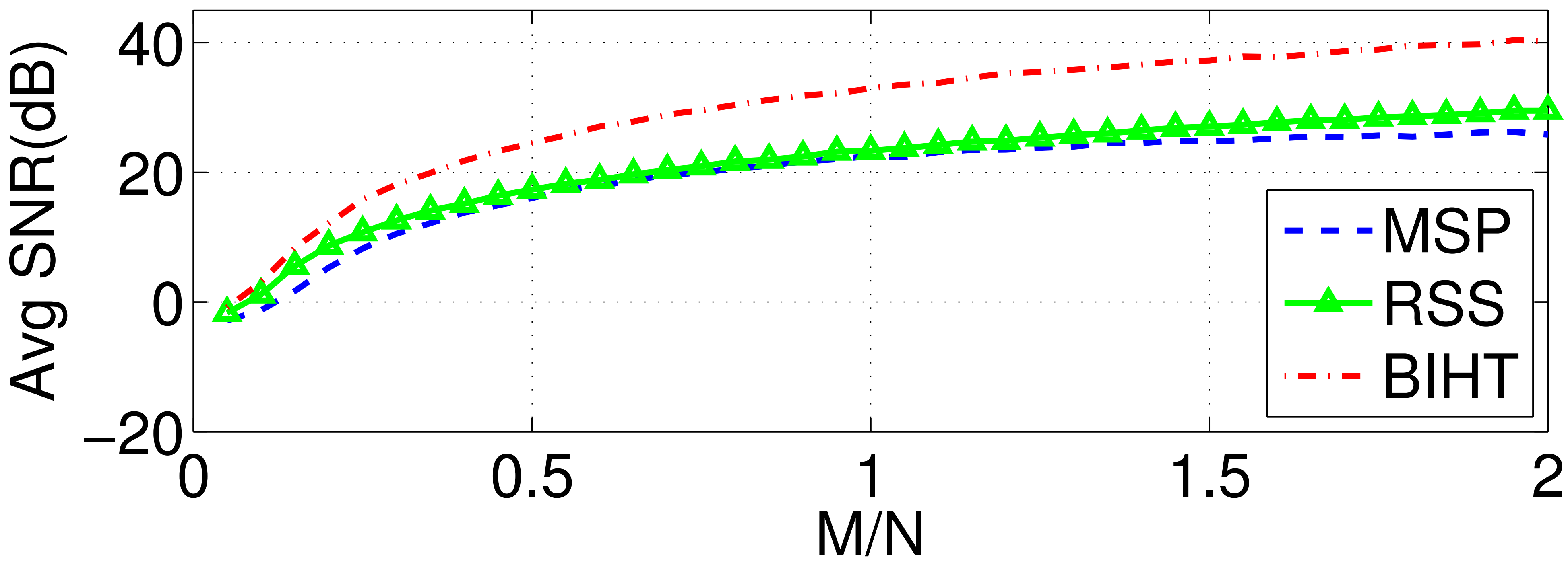
$$\mathcal{J}(\boldsymbol{x}') = \sum_{j=1}^{M} \Big|\big(\overbrace{\mathrm{sign}\,(\langle \boldsymbol{\varphi}_j, \boldsymbol{x}\rangle)}^{q_j}\,\langle \boldsymbol{\varphi}_j, \boldsymbol{x}'\rangle\big)_-\Big|$$

$$q_k - A(\boldsymbol{x}^l)_k = 0$$

$$q_j - A(\boldsymbol{x}^l)_j > 0$$

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

Binary Iterative Hard Thresholding

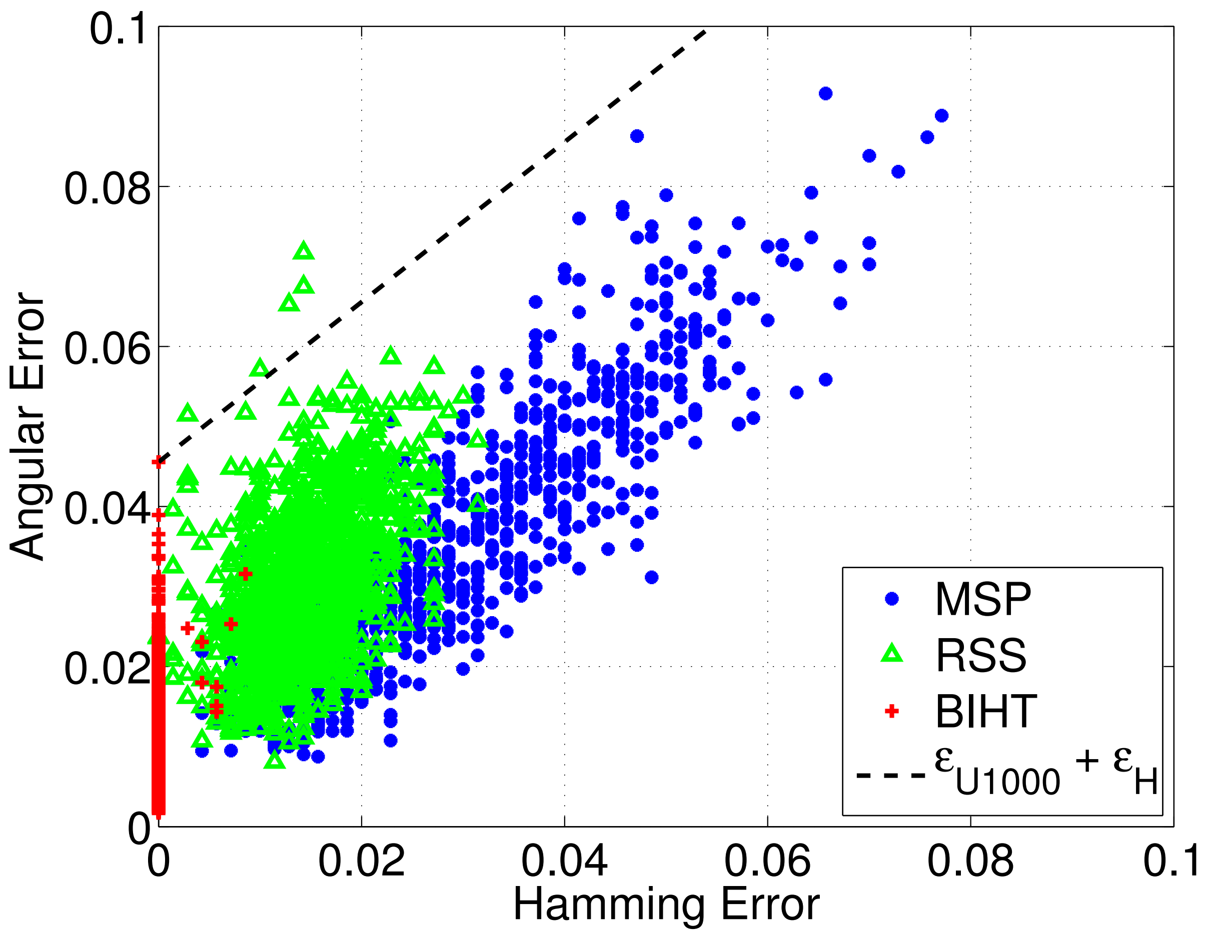


$N = 1000,\ K = 10$
Bernoulli-Gaussian model
normalized signals
1000 trials

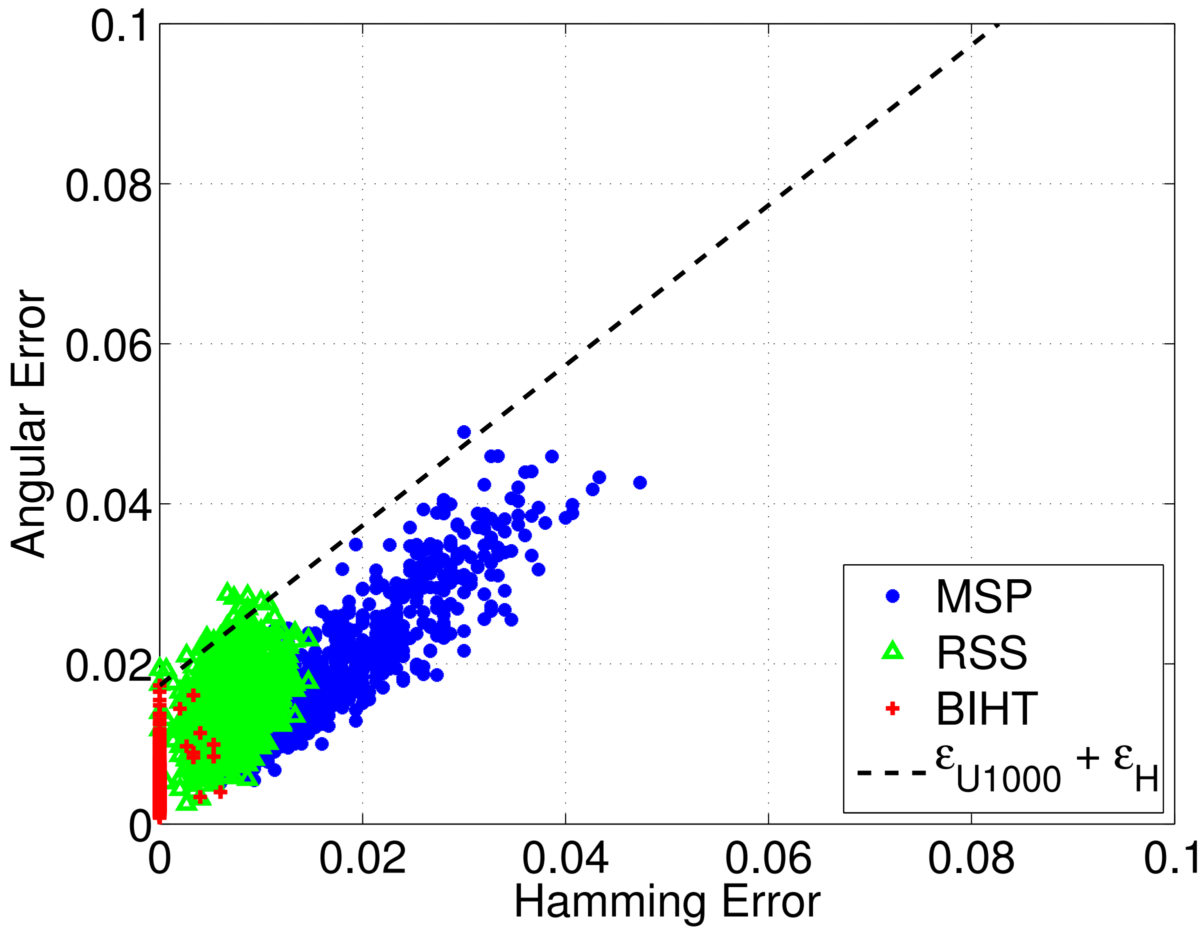
Matching Sign pursuit (MSP)
Restricted-Step Shrinkage (RSS)
Binary Iterative Hard Thresholding (BIHT)

Binary Iterative Hard Thresholding

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

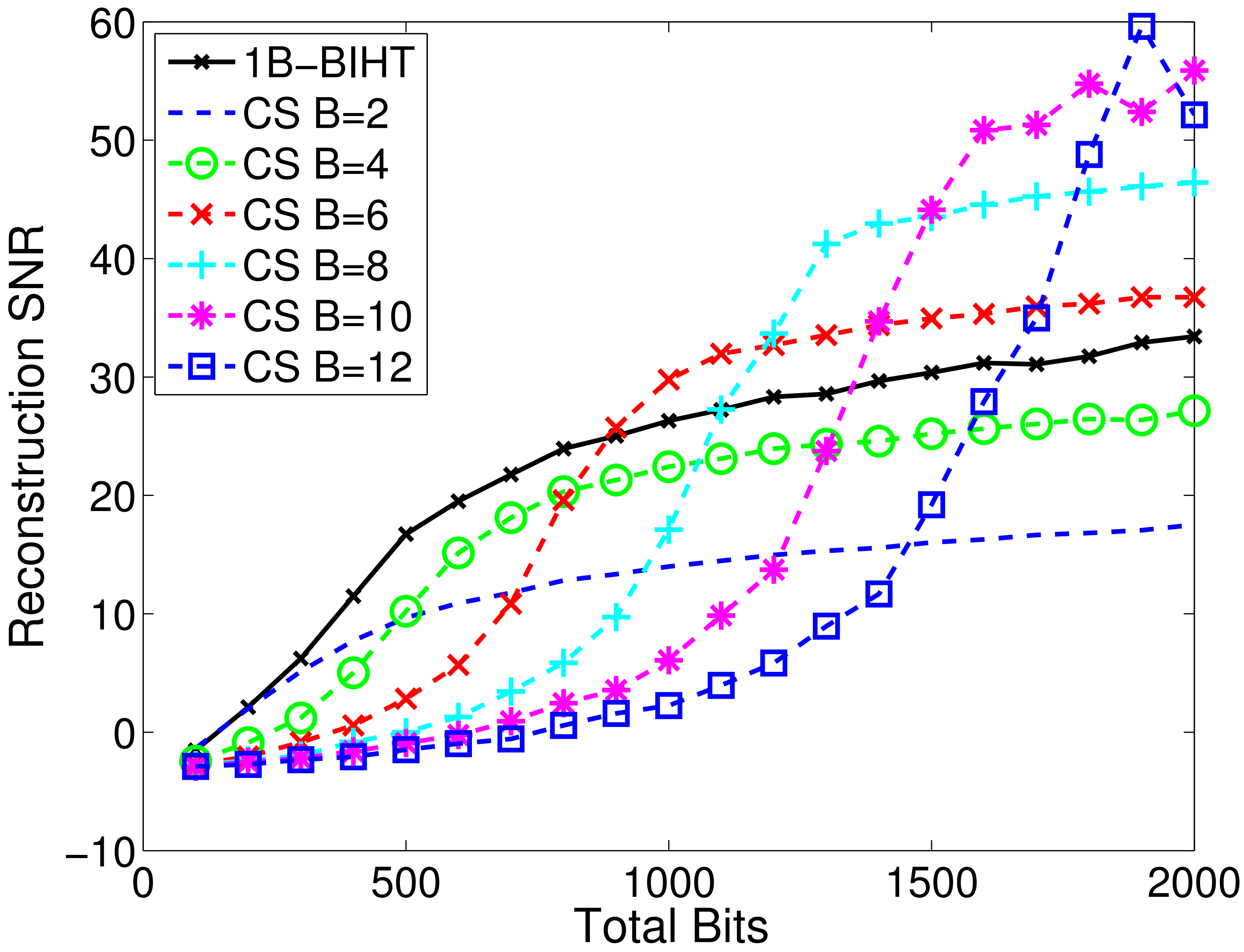


$M/N = 0.7$



$M/N = 1.5$

Remark: CS vs bits/meas.



$N = 2000,\ K = 20$

Bernoulli-Gaussian model

normalized signals

B bits/measurement

$B = 1, \ldots, 12$

$M = \text{Total Bits}/B$

1000 trials

Convex Optimization [Plan, Vershynin, 12]

Let $\boldsymbol{q} = \mathrm{sign}\,(\boldsymbol{\Phi x})$ for some signal $\boldsymbol{x} \in \mathcal{K} \subset B_2^N$

e.g., sparse, compressible, low-rank matrix

Compute

$$\boxed{\hat{\boldsymbol{x}} \;=\; \arg\max_{\boldsymbol{u} \in \mathbb{R}^N} \; \boldsymbol{q}^T \boldsymbol{\Phi u} \quad \text{s.t.} \quad \boldsymbol{u} \in \mathcal{K}}$$

maximize consistency

Convex problem if $\mathcal{K}$ convex!
No ambiguous amplitude definition
($\boldsymbol{u} = 0$ avoided)

Remark: (PV-L0 problem) [Bahmani, Boufounos, Raj, 13]

$$\hat{\boldsymbol{x}} \;=\; \frac{1}{\|\mathcal{H}_K(\boldsymbol{\Phi}^* \boldsymbol{q})\|}\, \mathcal{H}_K(\boldsymbol{\Phi}^* \boldsymbol{q}) \text{ if } \mathcal{K} = \Sigma_K \text{ !!}$$

S. Bahmani, P.T. Boufounos, B. Raj, "Robust 1-bit Compressive Sensing via Gradient Support Pursuit", arxiv:1304.6626

Convex Optimization [Plan, Vershynin, 12]

Let $\boldsymbol{q} = \operatorname{sign}(\boldsymbol{\Phi x})$ for some signal $\boldsymbol{x} \in \mathcal{K} \subset B_2^N$

e.g., sparse, compressible, low-rank matrix

Compute $\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{u} \in \mathbb{R}^N} \boldsymbol{q}^T \boldsymbol{\Phi u}$ s.t. $\boldsymbol{u} \in \mathcal{K}$

maximize consistency

-2 if $\boldsymbol{x}$ is fixed

Proposition *(assuming $\|\boldsymbol{x}\| = 1$) For some $C, c > 0$, if $M \geqslant C\epsilon^{-6} w^2(\mathcal{K})$, then, with $Pr \geqslant 1 - e^{-c\epsilon^2 M}$, we have $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 \leqslant \sqrt{\frac{\pi}{2}}\,\epsilon$.*

Convex Optimization [Plan, Vershynin, 12]

Let $\boldsymbol{q} = \operatorname{sign}(\boldsymbol{\Phi x})$ for some signal $\boldsymbol{x} \in \mathcal{K} \subset B_2^N$

Compute $\boxed{\hat{\boldsymbol{x}} \ = \ \arg\max_{\boldsymbol{u} \in \mathbb{R}^N} \ \boldsymbol{q}^T \boldsymbol{\Phi u} \quad \text{s.t.} \quad \boldsymbol{u} \in \mathcal{K}}$

Proposition *(assuming $\|\boldsymbol{x}\| = 1$) For some $C, c > 0$, if $M \geqslant C\epsilon^{-6} w^2(\mathcal{K})$, then, with $Pr \geqslant 1 - e^{-c\epsilon^2 M}$, we have $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 \leqslant \sqrt{\frac{\pi}{2}}\,\epsilon$.*

+ Robust to noise:

noise (bit flip) noise power

Let $\boldsymbol{q}_{\mathrm{n}} = \operatorname{diag}(\boldsymbol{\eta})\,\boldsymbol{q}$ with $\eta_i \in \{\pm 1\}^M$, and assume $d_H(\boldsymbol{q}, \boldsymbol{q}_{\mathrm{n}}) \leqslant p$

(under the same conditions)

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 \ \leqslant \ \epsilon\sqrt{\log e/\epsilon} \ + \ 11\,p\,\sqrt{\log e/p}$$

Note: if $M = O(\epsilon^{-2}(p - 1/2)^{-2} K \log N/K)$
this term disappear if $\eta_i = \pm 1$ are iid RVs (with $P(\eta_i = 1) = p$)

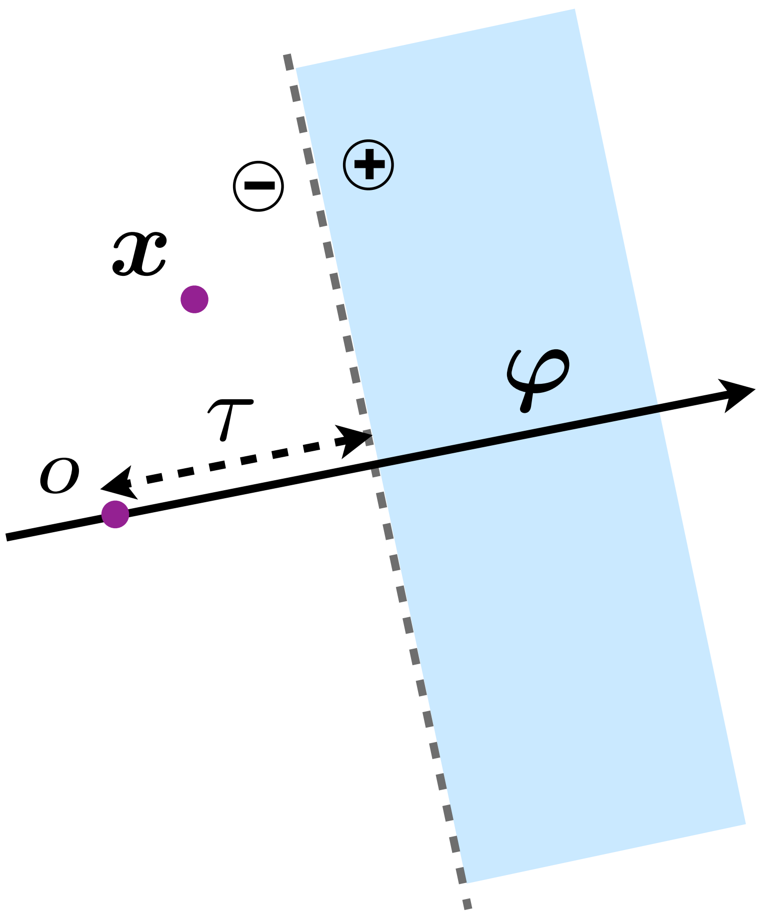
5. Playing with thresholds in 1-bit CS

Thresholds?

- Given $\boldsymbol{x} \in \mathbb{R}^N$ (*e.g.*, sparse)
  Is there an interest in sensing

$$\boxed{\operatorname{sign}\left(\langle \boldsymbol{\varphi}, \boldsymbol{x} \rangle - \tau\right)}$$

  for some (random) $\boldsymbol{\varphi}$ and $\tau \in \mathbb{R}$?

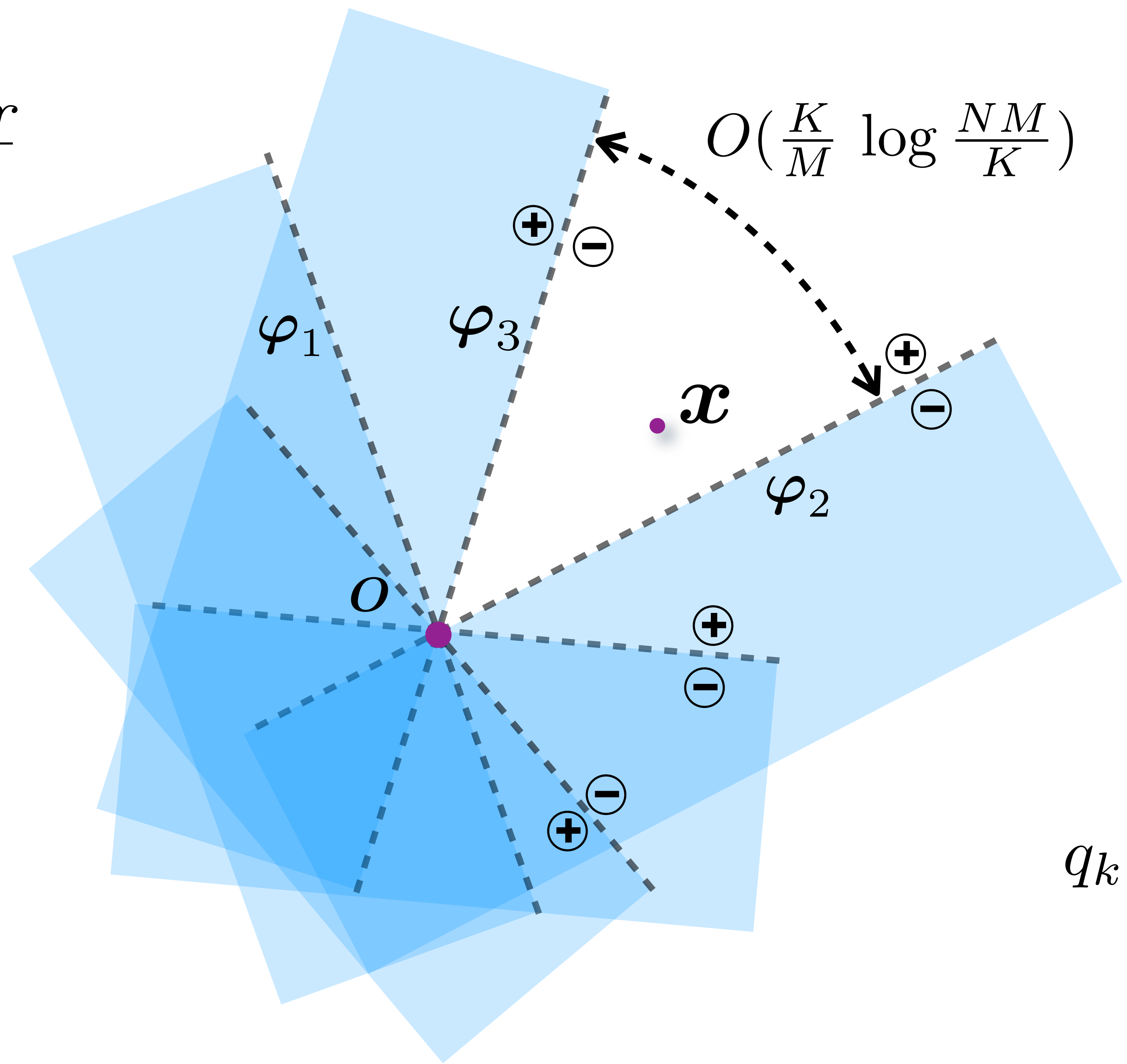


- Two recent applications:
  - adaptive thresholds [Kamilov, Bourquard, Amini, Unser, 12]
  - bridging 1-bit and B-bits QCS [LJ, Degraux, De Vleeschouwer, 13]

1-bit CS with adaptive thresholds

Non-adaptive 1-bit CS $(\tau = 0)$

Reminder



$$q_k = \mathrm{sign}\left(\langle \boldsymbol{\varphi}_k, \boldsymbol{x} \rangle\right)$$

1-bit CS with adaptive thresholds

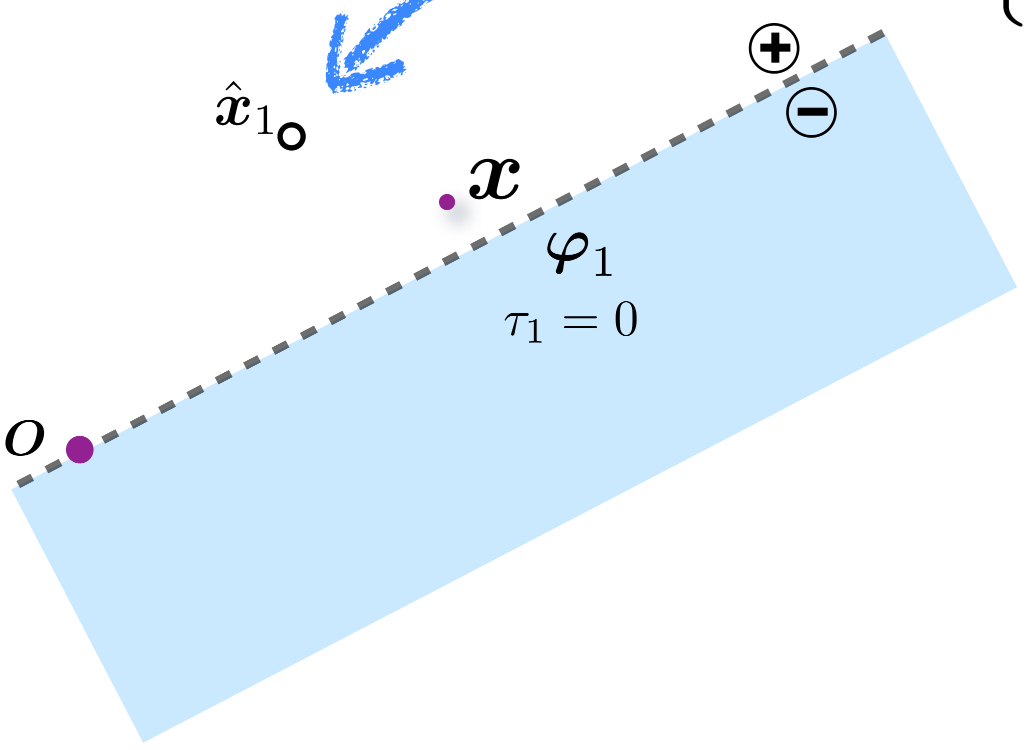
Adaptive 1-bit CS [Kamilov, Bourquard, Amini, Unser, 12]

Given a decoder Rec()

adapted from prev. meas.

$$q_k = \mathrm{sign}\left(\langle \boldsymbol{\varphi}_k, \boldsymbol{x}\rangle - \tau_k\right)$$

$$\begin{cases} \hat{\boldsymbol{x}}_k := \mathrm{Rec}(y_1, \cdots, y_k, \boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_k, \tau_1, \cdots, \tau_k) \\ \tau_{k+1} \text{ s.t. } \langle \boldsymbol{\varphi}_{k+1}, \hat{\boldsymbol{x}}_k\rangle - \tau_{k+1} = 0 \end{cases}$$



U.S. Kamilov, A. Bourquard, A. Amini, M. Unser,
"One-bit measurements with adaptive thresholds". Signal Processing Letters, IEEE, 19(10), 607-610.

1-bit CS with adaptive thresholds

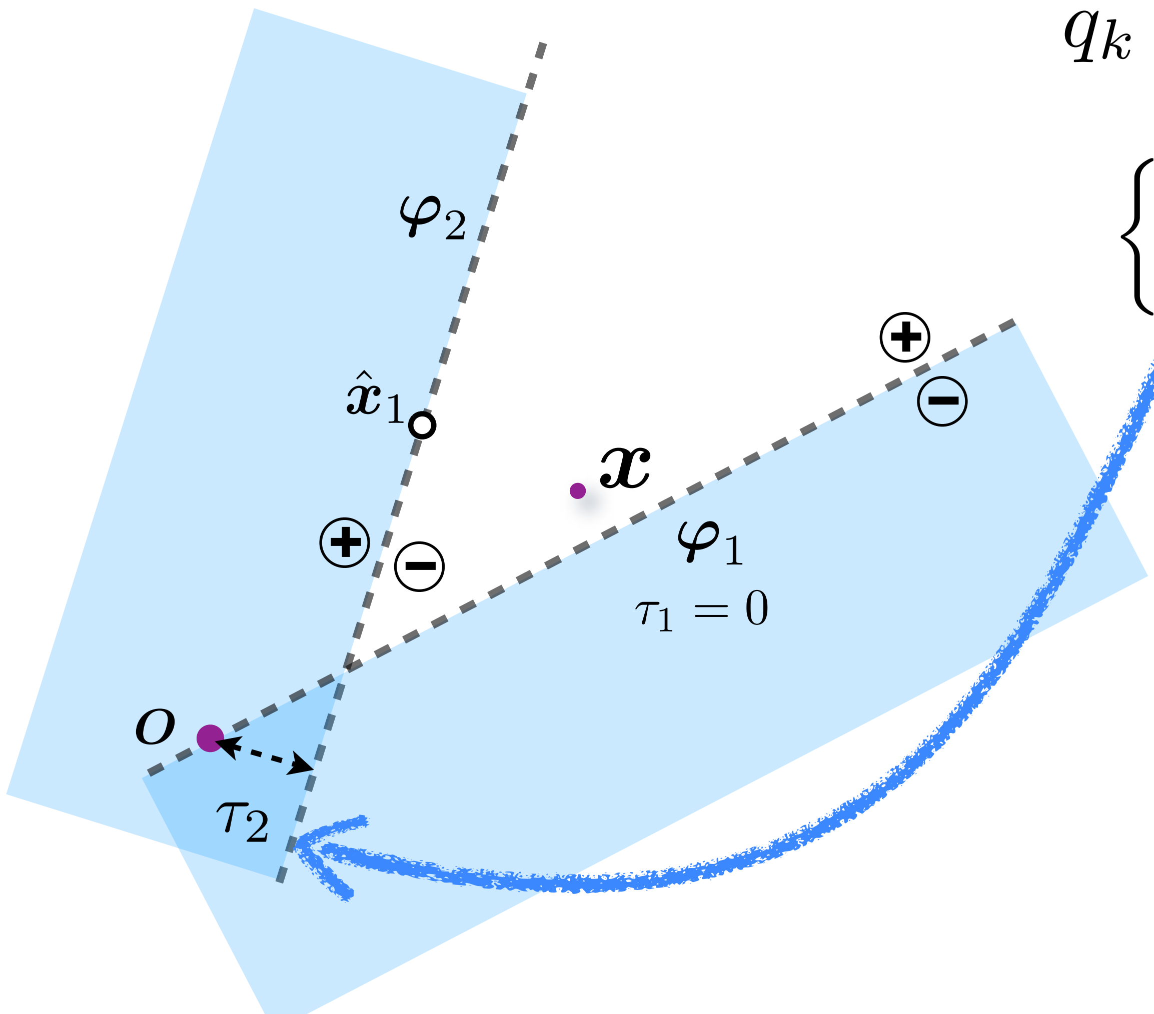
Adaptive 1-bit CS [Kamilov, Bourquard, Amini, Unser, 12]

Given a decoder Rec()

adapted from prev. meas.

$$q_k = \mathrm{sign}\left(\langle \boldsymbol{\varphi}_k, \boldsymbol{x}\rangle - \tau_k\right)$$

$$\begin{cases} \hat{\boldsymbol{x}}_k := \mathrm{Rec}(y_1, \cdots, y_k, \boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_k, \tau_1, \cdots, \tau_k) \\ \tau_{k+1} \text{ s.t. } \langle \boldsymbol{\varphi}_{k+1}, \hat{\boldsymbol{x}}_k\rangle - \tau_{k+1} = 0 \end{cases}$$



U.S. Kamilov, A. Bourquard, A. Amini, M. Unser,
"One-bit measurements with adaptive thresholds". Signal Processing Letters, IEEE, 19(10), 607-610.

1-bit CS with adaptive thresholds

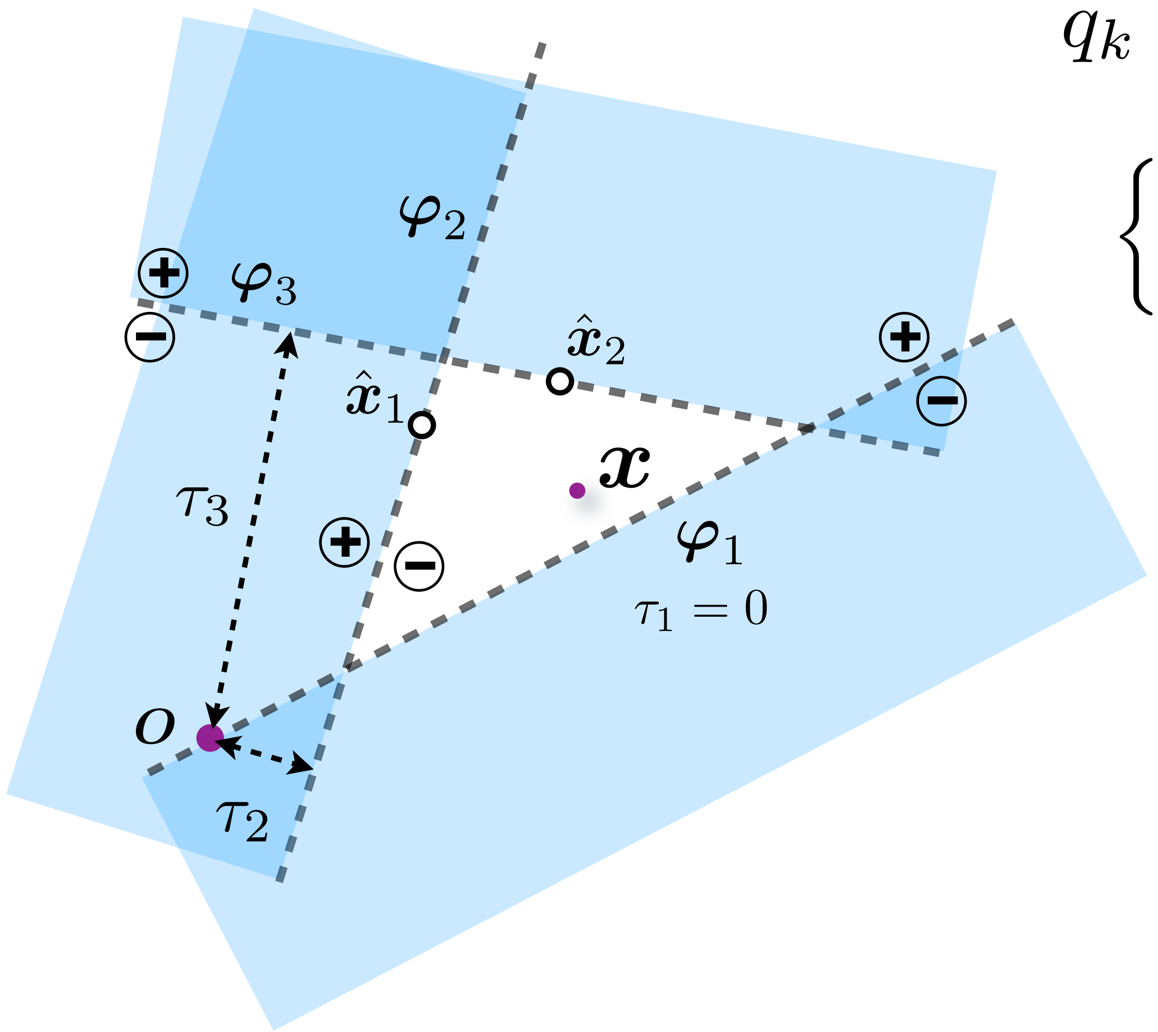
Adaptive 1-bit CS [Kamilov, Bourquard, Amini, Unser, 12]

Given a decoder Rec()

adapted from prev. meas.

$$q_k = \mathrm{sign}\left(\langle \boldsymbol{\varphi}_k, \boldsymbol{x}\rangle - \tau_k\right)$$

$$\begin{cases} \hat{\boldsymbol{x}}_k := \mathrm{Rec}(y_1, \cdots, y_k, \boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_k, \tau_1, \cdots, \tau_k) \\ \tau_{k+1} \text{ s.t. } \langle \boldsymbol{\varphi}_{k+1}, \hat{\boldsymbol{x}}_k\rangle - \tau_{k+1} = 0 \end{cases}$$



U.S. Kamilov, A. Bourquard, A. Amini, M. Unser,
"One-bit measurements with adaptive thresholds". Signal Processing Letters, IEEE, 19(10), 607-610.

1-bit CS with adaptive thresholds

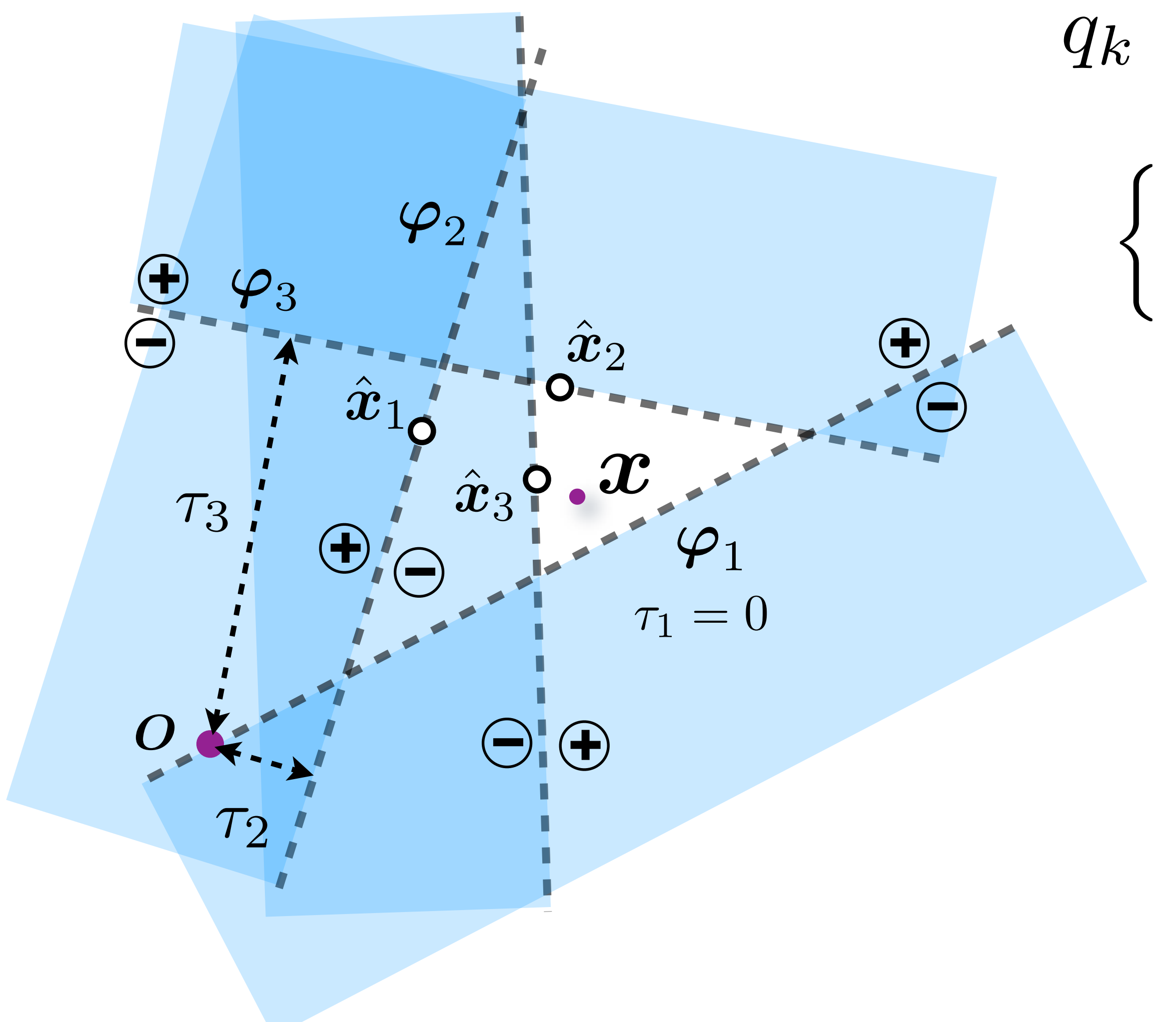
Adaptive 1-bit CS [Kamilov, Bourquard, Amini, Unser, 12]

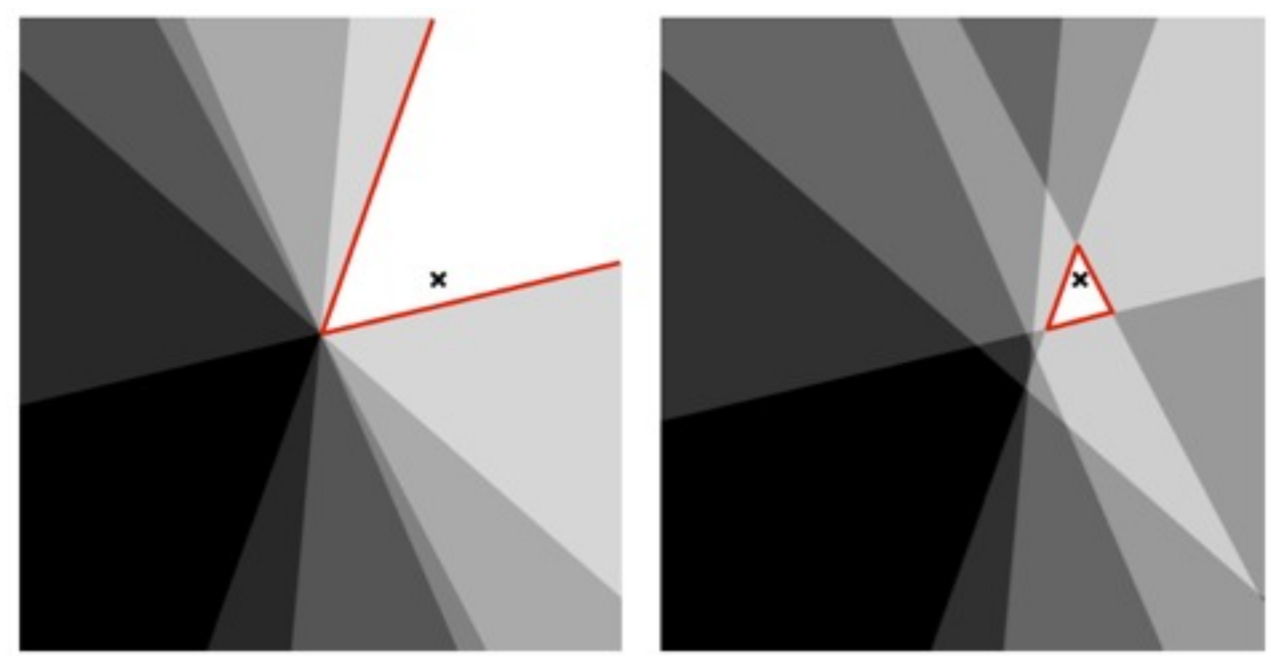
Given a decoder Rec()

adapted from prev. meas.

$$q_k = \mathrm{sign}\left(\langle \boldsymbol{\varphi}_k, \boldsymbol{x}\rangle - \tau_k\right)$$

$$\begin{cases} \hat{\boldsymbol{x}}_k := \mathrm{Rec}(y_1, \cdots, y_k, \boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_k, \tau_1, \cdots, \tau_k) \\ \tau_{k+1} \text{ s.t. } \langle \boldsymbol{\varphi}_{k+1}, \hat{\boldsymbol{x}}_k\rangle - \tau_{k+1} = 0 \end{cases}$$

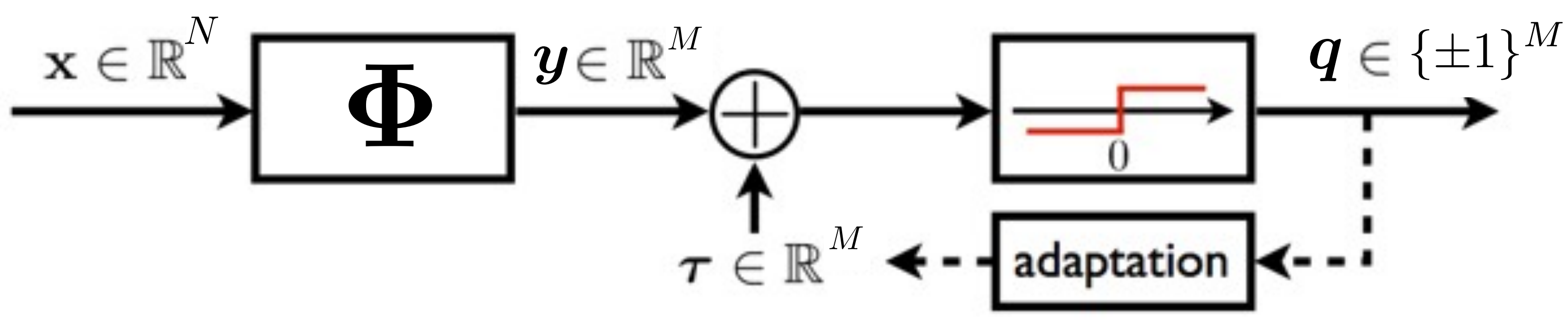




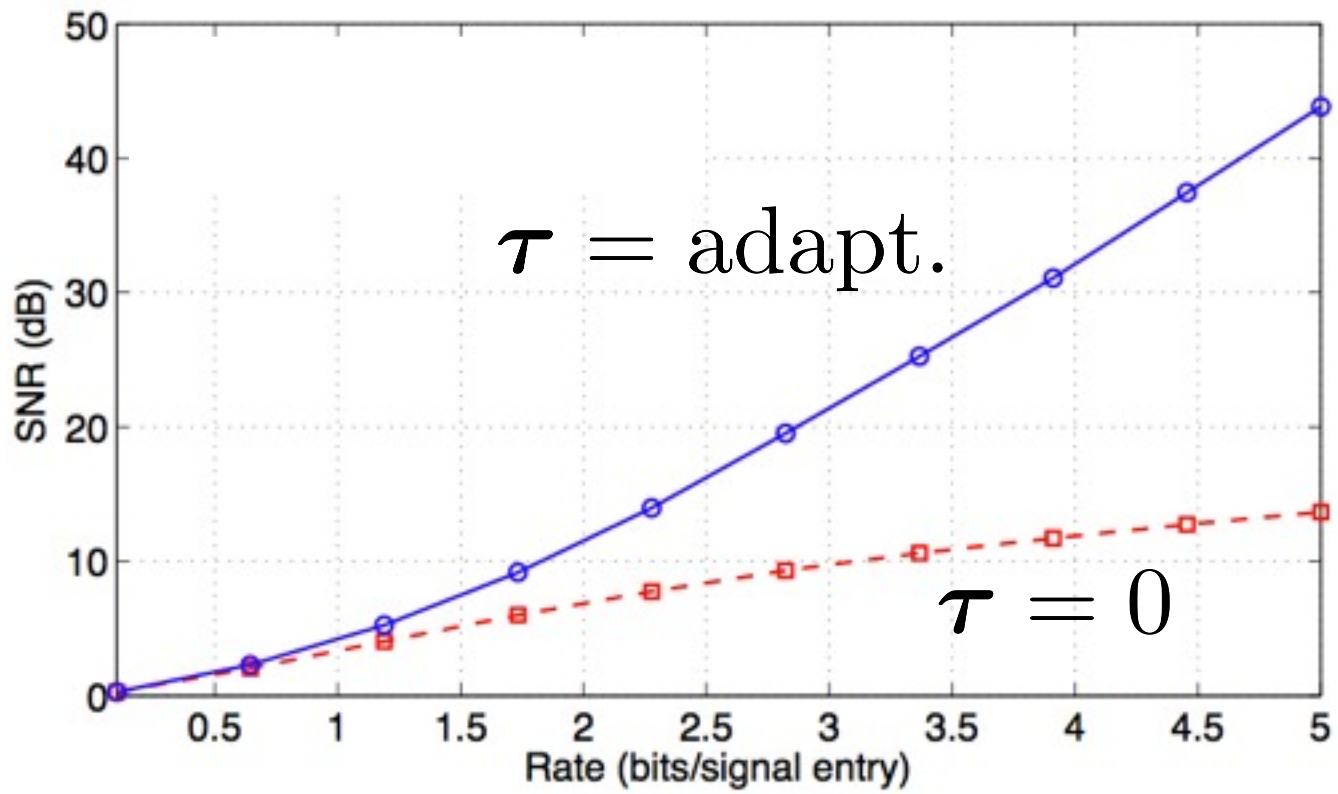
U.S. Kamilov, A. Bourquard, A. Amini, M. Unser,
"One-bit measurements with adaptive thresholds". Signal Processing Letters, IEEE, 19(10), 607-610.

1-bit CS with adaptive thresholds

System view:



Kind of
$\Sigma\Delta$ loop

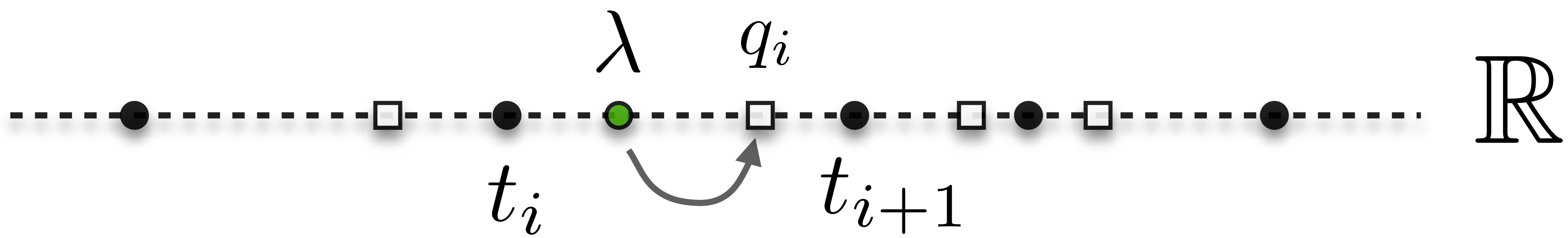


Rec() set to
Generalized
Approximate
Message
Passing

U.S. Kamilov, A. Bourquard, A. Amini, M. Unser,
"One-bit measurements with adaptive thresholds". Signal Processing Letters, IEEE, 19(10), 607-610.

Bridging 1-bit & B-bit CS?

- B-bit quantizer defined with thresholds:



$$\lambda \in \mathcal{R}_i = [t_i, t_{i+1}) \Leftrightarrow \mathrm{sign}\,(\lambda - t_i) = +1 \ \& \ \mathrm{sign}\,(\lambda - t_{i+1}) = -1$$

- Can we combine multiple thresholds in 1-bit CS?

Bridging 1-bit & B-bit CS?

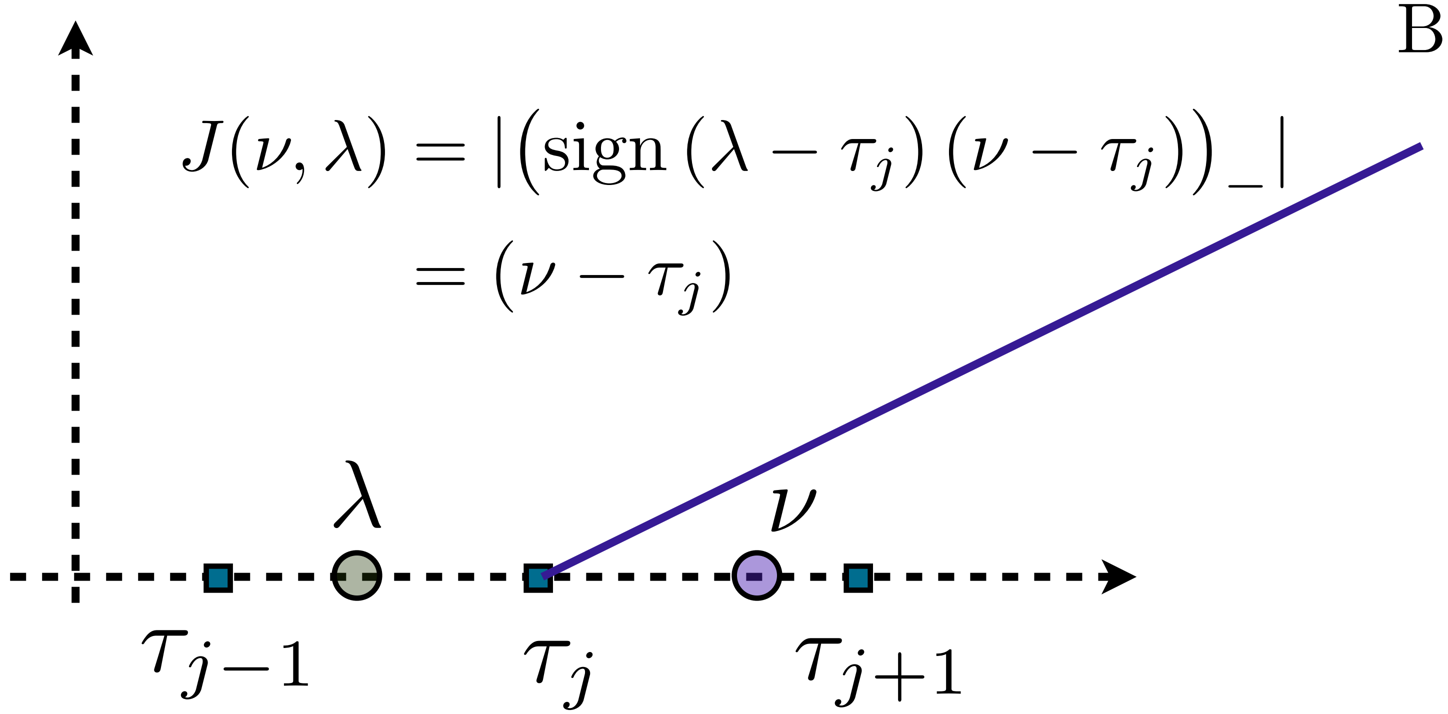
Given $\mathcal{T} = \{\tau_j\}$ and $\Omega = \{q_j\}$ ($|\mathcal{T}| = 2^B + 1 = |\Omega| + 1$), let's define

$$\boxed{J(\nu, \lambda) = \sum_{j=2}^{2^B} w_j \left|\big(\mathrm{sign}\,(\lambda - \tau_j)\,(\nu - \tau_j)\big)_-\right|,}$$

with $w_j = q_j - q_{j-1}$.

<u>Illustration:</u> $\lambda \in [\tau_{j-1}, \tau_j)$, $\nu \in [\tau_j, \tau_{j+1})$

"delocalized"
BIHT ℓ_1-sided norm



(for $w_j = 1$)

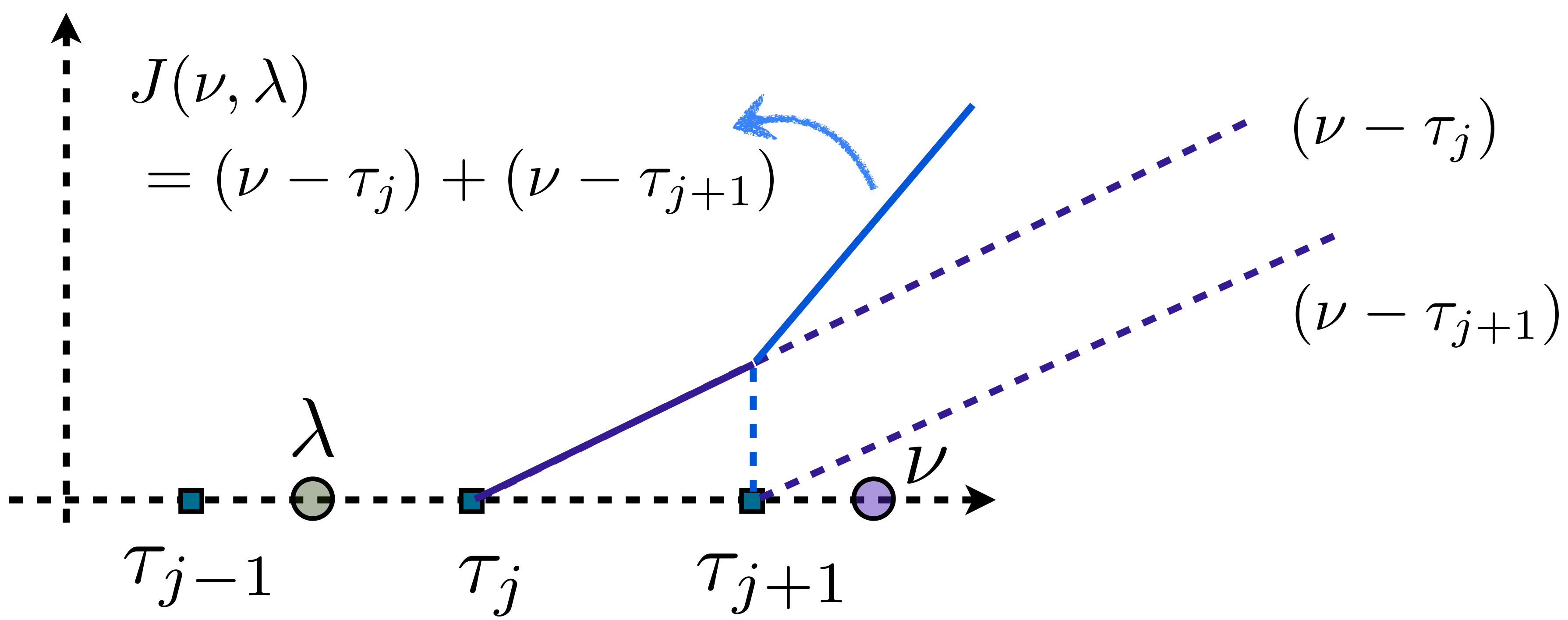
Bridging 1-bit & B-bit CS?

Given $\mathcal{T} = \{\tau_j\}$ and $\Omega = \{q_j\}$ ($|\mathcal{T}| = 2^B + 1 = |\Omega| + 1$), let's define

$$\boxed{J(\nu, \lambda) = \sum_{j=2}^{2^B} w_j \left|\left(\mathrm{sign}\,(\lambda - \tau_j)\,(\nu - \tau_j)\right)_-\right|,}$$

with $w_j = q_j - q_{j-1}$.

<u>Illustration:</u> $\lambda \in [\tau_{j-1}, \tau_j),\ \nu \in [\tau_{j+1}, \tau_{j+2})$



(for $w_j = 1$)

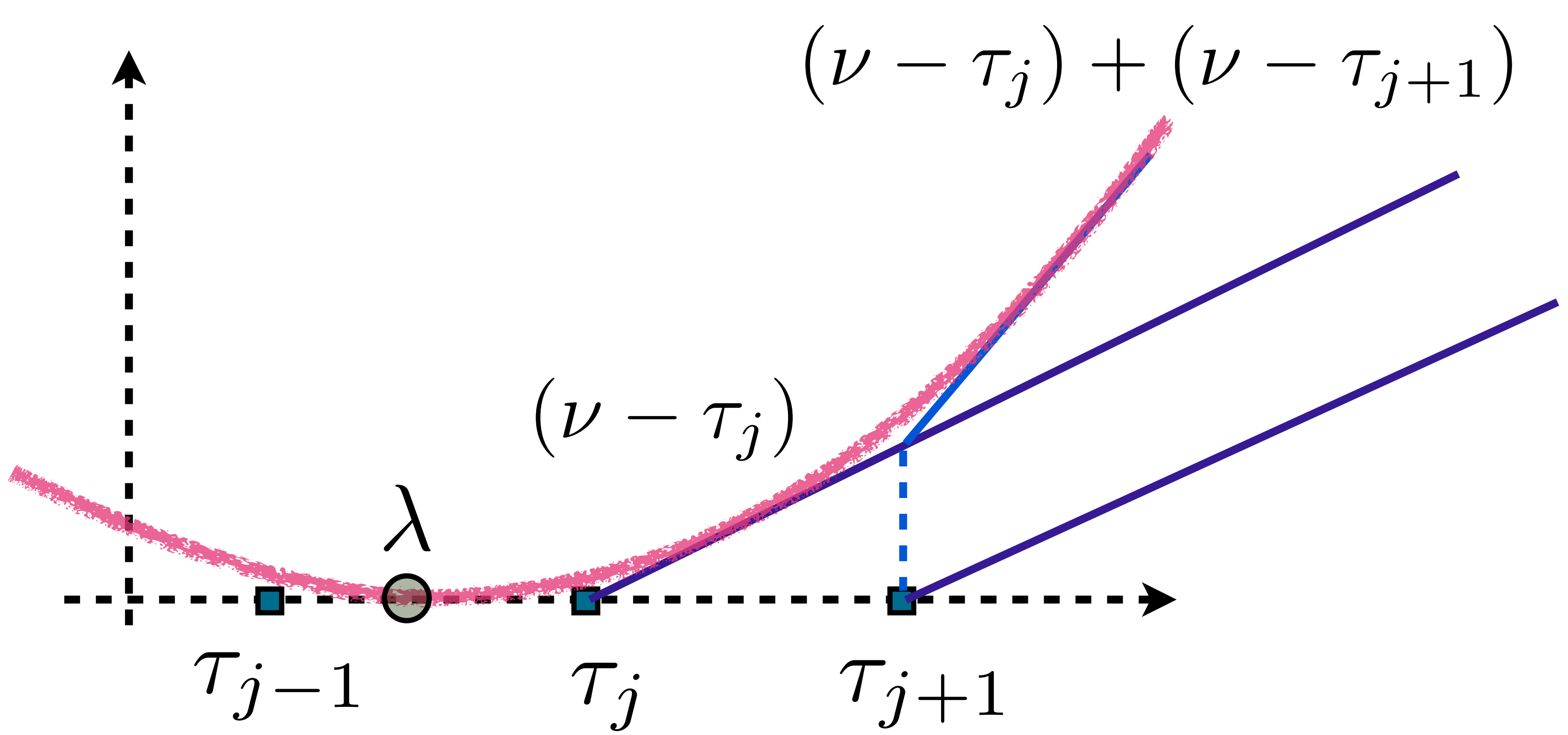
Bridging 1-bit & B-bit CS?

Given $\mathcal{T} = \{\tau_j\}$ and $\Omega = \{q_j\}$ ($|\mathcal{T}| = 2^B + 1 = |\Omega| + 1$), let's define

$$\boxed{J(\nu, \lambda) = \sum_{j=2}^{2^B} w_j \left|\left(\operatorname{sign}(\lambda - \tau_j)(\nu - \tau_j)\right)_-\right|,}$$

with $w_j = q_j - q_{j-1}$.

Illustration:



(for $w_j = 1$)

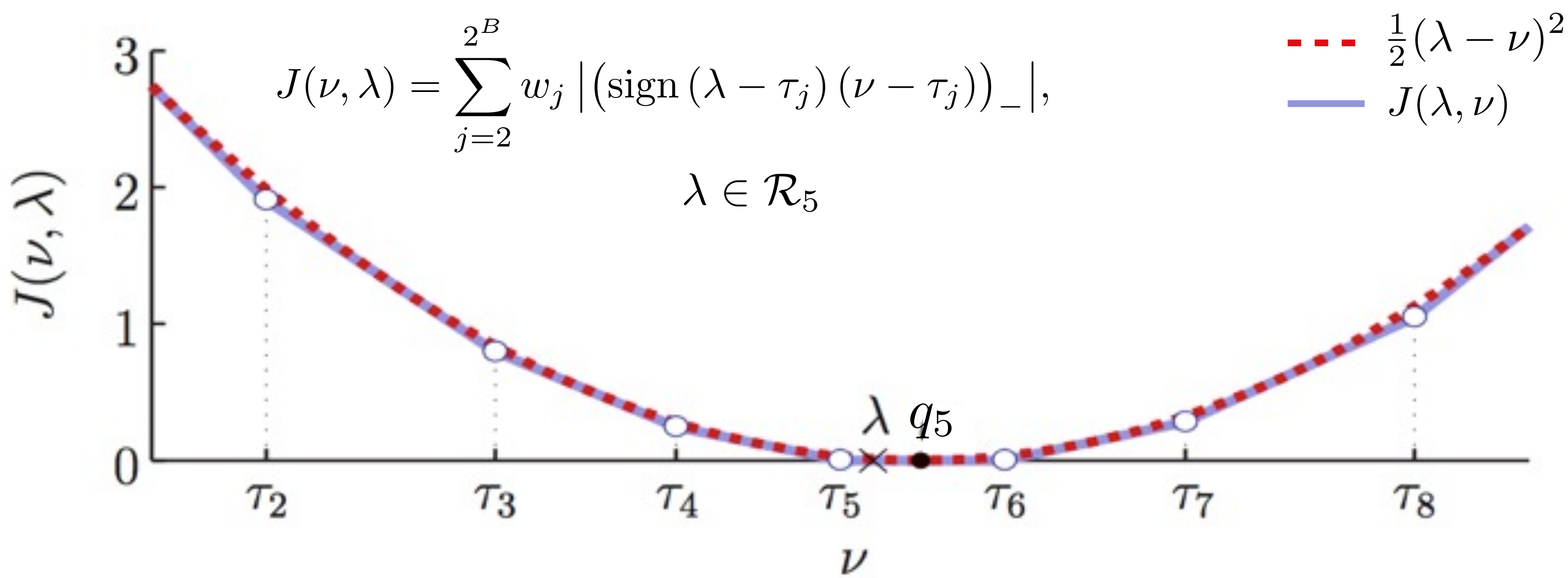
Bridging 1-bit & B-bit CS?

Given $\mathcal{T} = \{\tau_j\}$ and $\Omega = \{q_j\}$ ($|\mathcal{T}| = 2^B + 1 = |\Omega| + 1$), let's define

$$\boxed{J(\nu,\lambda) = \sum_{j=2}^{2^B} w_j \left|\big(\mathrm{sign}\,(\lambda - \tau_j)\,(\nu - \tau_j)\big)_-\right|,}$$

with $w_j = q_j - q_{j-1}$.

<u>Illustration:</u> *more bins*



Bridging 1-bit & B-bit CS?

Given $\mathcal{T} = \{\tau_j\}$ and $\Omega = \{q_j\}$ ($|\mathcal{T}| = 2^B + 1 = |\Omega| + 1$), let's define

$$\boxed{J(\nu, \lambda) = \sum_{j=2}^{2^B} w_j \left|\left(\mathrm{sign}\,(\lambda - \tau_j)\,(\nu - \tau_j)\right)_-\right|,}$$

with $w_j = q_j - q_{j-1}$.

For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^M$: $\boxed{\mathcal{J}(\boldsymbol{u}, \boldsymbol{v}) := \sum_{k=1}^M J(u_k, v_k)}$

Remarks:

- J is convex in ν
- For $B = 1$ ($j = 2$ only):

  $\mathcal{J}(\boldsymbol{u}, \boldsymbol{v}) \propto \|(\mathrm{sign}\,(\boldsymbol{v}) \odot \boldsymbol{u})_-\|_1 \rightarrow \ell_1$-sided 1-bit energy
- For $B \gg 1$:

  $J(\nu, \lambda) \rightarrow \frac{1}{2}(\nu - \lambda)^2$ and $\mathcal{J}(\boldsymbol{u}, \boldsymbol{v}) \rightarrow \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{v}\|^2$ (quadratic energy)

Bridging 1-bit & B-bit CS?

- Let's define an *inconsistency* energy:

$$\boxed{\mathcal{E}_B(\boldsymbol{u}) := \mathcal{J}(\boldsymbol{\Phi u}, \boldsymbol{q})} \text{ with } \boldsymbol{q} = \mathcal{Q}_B[\boldsymbol{\Phi x}] \text{ and } \mathcal{E}\_B(\boldsymbol{x}) = 0$$

- Idea: Minimize it in Σ_K (as for Iterative Hard Thresholding)

[Blumensath, Davies, 08]

$$\boxed{\min_{\boldsymbol{u}\in\mathbb{R}^N} \mathcal{E}_B(\boldsymbol{u}) \text{ s.t. } \|\boldsymbol{u}\|_0 \leqslant K,}$$

- NP Hard but greedy solution (as for IHT):

$$\boxed{\boldsymbol{x}^{(n+1)} = \mathcal{H}_K[\boldsymbol{x}^{(n)} - \mu\, \underbrace{\partial\, \mathcal{E}_B(\boldsymbol{x}^{(n)})}_{\text{(sub) gradient}}] \text{ and } \boldsymbol{x}^{(0)} = 0.}$$

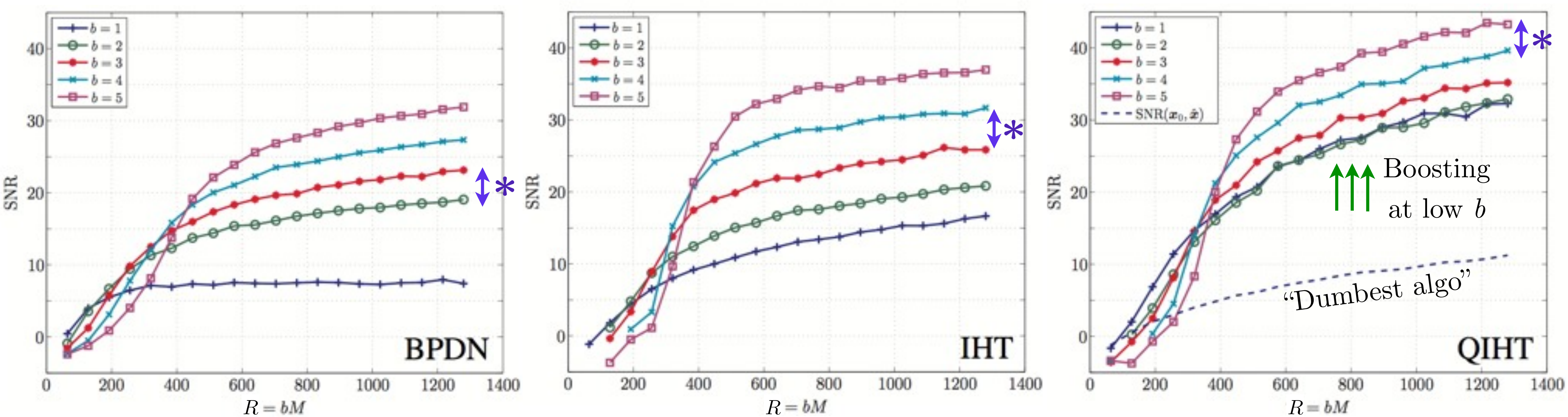
$$\boldsymbol{\Phi}^*(\mathrm{sign}\,(\boldsymbol{\Phi u}) - \mathrm{sign}\,(\boldsymbol{\Phi x})) \xleftarrow[B=1]{} \boxed{\partial\, \mathcal{E}_B(\boldsymbol{u}) = \boldsymbol{\Phi}^*(\mathcal{Q}_B(\boldsymbol{\Phi u}) - \boldsymbol{q})} \xrightarrow[B \gg 1]{} \boldsymbol{\Phi}^*(\boldsymbol{\Phi u} - \boldsymbol{q})$$

BIHT! Quantized IHT (QIHT) IHT!

T. Blumensath, M.E. Davies, "Iterative thresholding for sparse approximations". *Journal of Fourier Analysis and Applications*, *14*(5-6), 629-654. (2008).
LJ, K. Degraux, C. De Vleeschouwer, "Quantized Iterative Hard Thresholding: Bridging 1-bit and High-Resolution Quantized Compressed Sensing", SAMPTA2013

Bridging 1-bit & B-bit CS?

$N = 1024,\ K = 16,\ R = BM \in \{64, 128, \cdots, 1280\}$, 100 trials (+ Lloyd-Max Gauss. Q.)



R: total bit budget (BM)

*: almost "6dB per bit" gain

$\mu = \frac{1}{M}(1 - \sqrt{2K/M})$

Adjusted by limit case analysis: BIHT and IHT

Note: entropy could be computed instead of B ($e.g.$, for further efficient coding)

LJ, K. Degraux, C. De Vleeschouwer, "Quantized Iterative Hard Thresholding: Bridging 1-bit and High-Resolution Quantized Compressed Sensing", SAMPTA2013

Bridging 1-bit & B-bit CS?

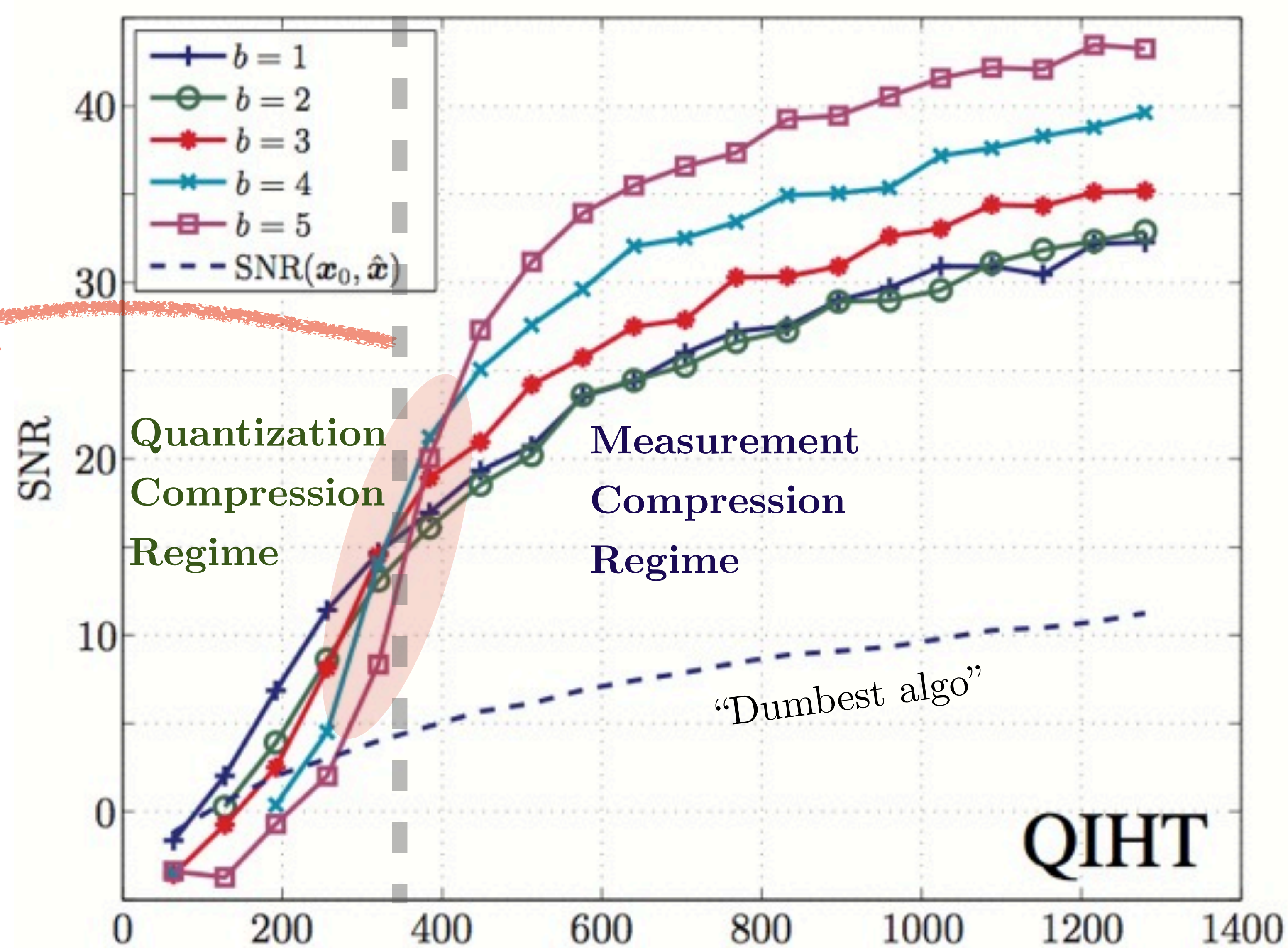
$N = 1024,\ K = 16,\ R = BM \in \{64, 128, \cdots, 1280\}$, 100 trials

Interesting transition at

$R_0 \simeq 375$

"**Regime Change**?"
[Laska, Baraniuk, 12]

R_0 could increase with input noise power.



$R = BM$

J. N. Laska, R. G. Baraniuk, 'Regime change: Bit-depth versus measurement-rate in compressive sensing", *Signal Processing, IEEE Transactions on*, *60*(7), 3496-3505. (2012)

Further Reading

- T. Blumensath, M.E. Davies, “Iterative thresholding for sparse approximations”. *Journal of Fourier Analysis and Applications*, 14(5-6), pp. 629-654, 2008
- P. T. Boufounos and R. G. Baraniuk, “1-Bit compressive sensing,” *Proc. Conf. Inform. Science and Systems (CISS)*, Princeton, NJ, March 19-21, 2008.
- Boufounos, P. T. (2009, November). “Greedy sparse signal reconstruction from sign measurements”. In *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, 2009
- Y. Plan, R. Vershynin, “Dimension reduction by random hyperplane tessellations”, arXiv:1111.4452, 2011.
- Y. Plan, R. Vershynin, “Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach”, *IEEE Trans. Info. Theory*, arXiv:1202.1212, 2012.
- J. N. Laska, R. G. Baraniuk, ‘Regime change: Bit-depth versus measurement-rate in compressive sensing”, *IEEE Trans. Signal Processing*, 60(7), pp. 3496-3505, 2012.
- U.S. Kamilov, A. Bourquard, A. Amini, M. Unser, “One-bit measurements with adaptive thresholds”. *IEEE Signal Processing Letters*, 19(10), pp. 607-610, 2012
- L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, “Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors,” *IEEE Trans. Info. Theory*, 59(4), 2013.
- L. Jacques, K. Degraux, C. De Vleeschouwer, “Quantized Iterative Hard Thresholding: Bridging 1-bit and High-Resolution Quantized Compressed Sensing", SAMPTA 2013, to appear.

Today's Topics

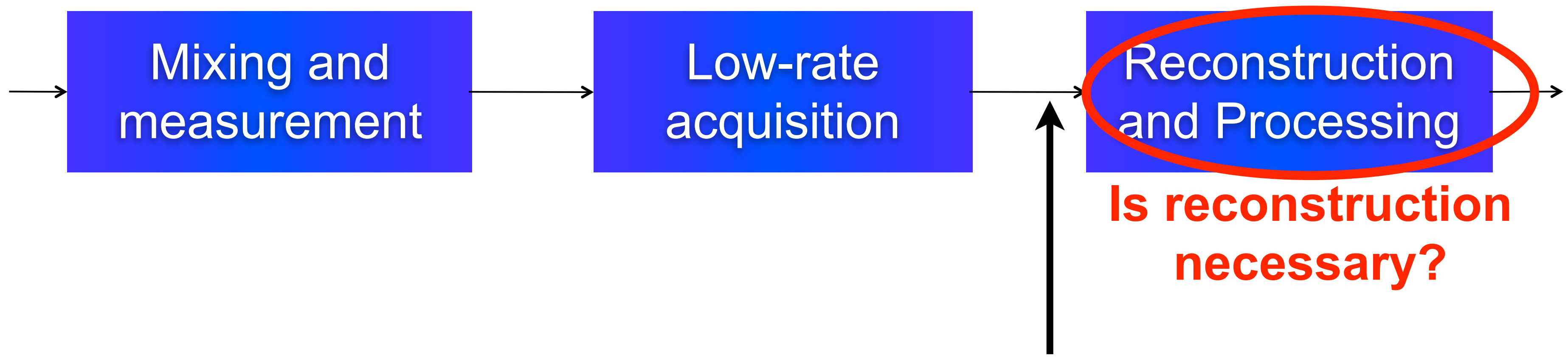
1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

Today's Topics

1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

INFORMATION EMBEDDING

Compressive Domain Processing



No: operate in compressive domain

Moreover: signal does not have to be sparse
(as long at it has some structure)

Compressive operations: **detection, estimation, filtering**

Randomized projection **embeds** signal **information**.

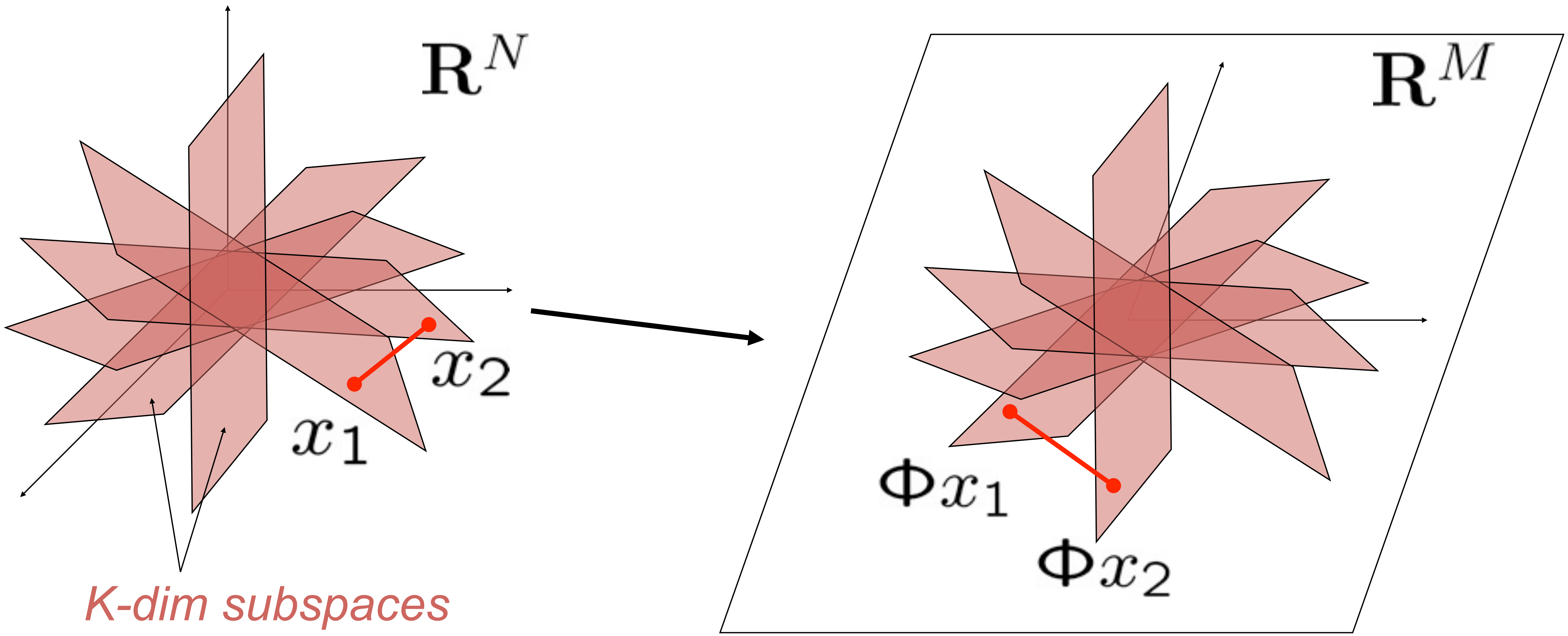
Main benefits: Computation, Memory

Questions:
What information is embedded?
How to best embed information?

- Davenport M. A., Boufounos P. T., Wakin M. B., and Baraniuk R. G., "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, v. 4, no. 2, pp. 445-460, April, 2010.

RIP/Stable Embedding

- An information preserving projection $\mathbf{A}$ preserves the **geometry** of the set of sparse signals



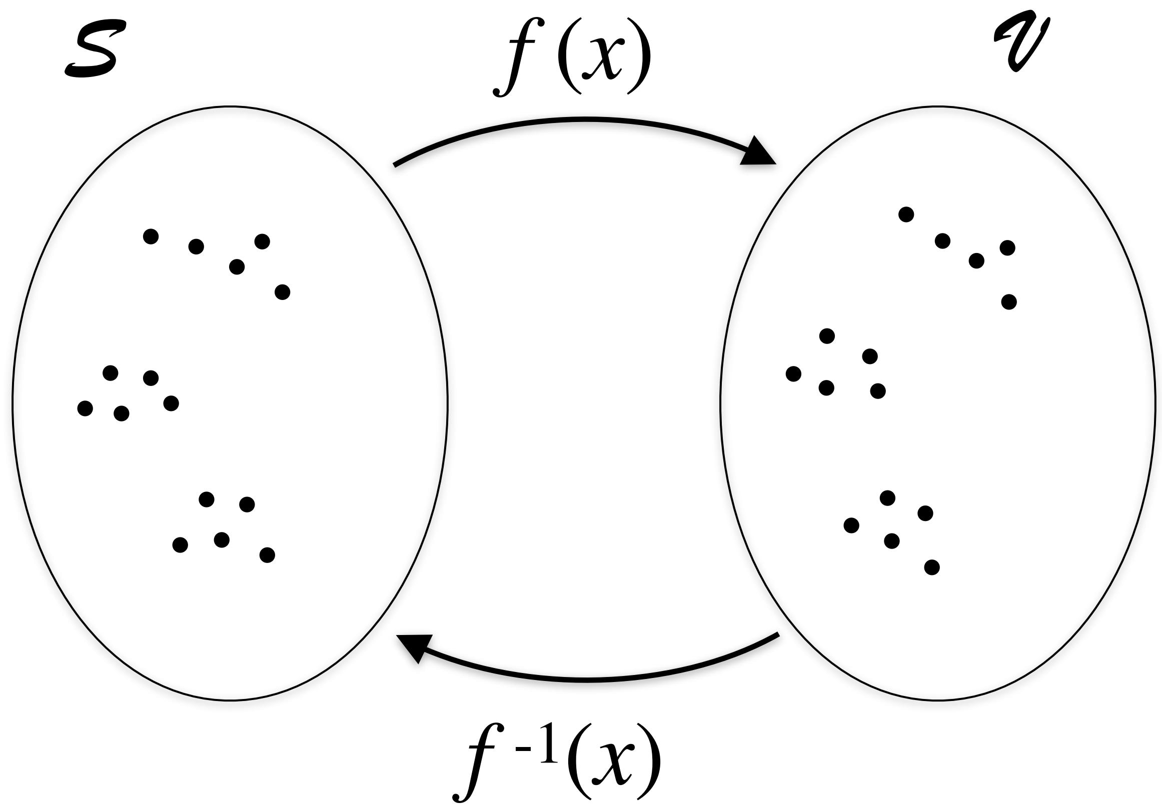
Restricted Isometry Property

$$(1-\delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1+\delta)\|x\|_2^2$$

GEOMETRY-PRESERVING EMBEDDINGS

Isometric (approximate) embeddings

Original space
high-dimensional and
expensive to work with



Embedding space
lower dimension or
easier to work with
(hopefully)

Transformations that preserve distances

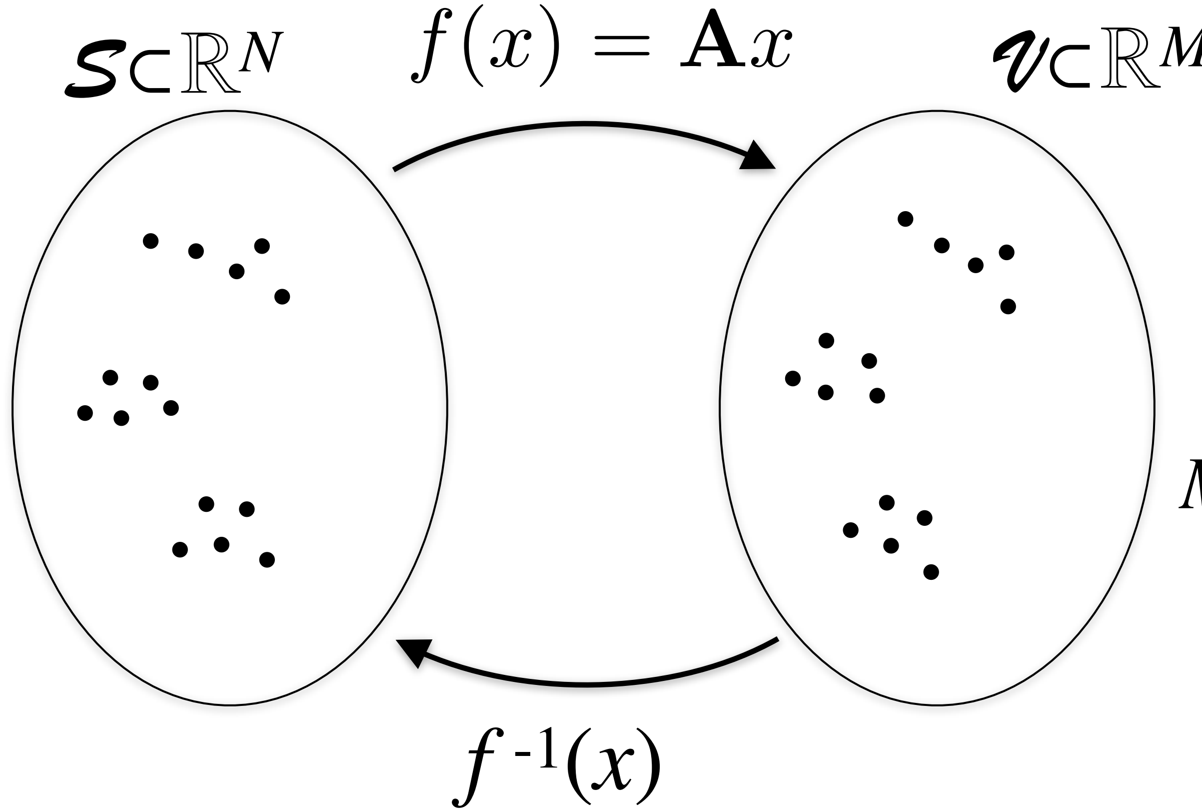
For all x,y in $\mathcal{S}$: $d_{\mathcal{S}}(x,y) \approx d_{\mathcal{V}}(f(x),f(y))$

Johnson-Lindenstrauss embeddings

Original space

Distance metric: ℓ_2

P points in $\mathbb{R}^N$



Embedding space

Embed in $\mathbb{R}^M$

Distance metric: ℓ_2

$M = O(\log P)$ dimensions

Transformations that preserve distances

For all x,y in $\mathcal{S}$:

$$(1-\epsilon)\|x-y\|_2^2 \leq \|f(x)-f(y)\|_2^2 \leq (1+\epsilon)\|x-y\|_2^2$$

- Johnson W. and Lindenstrauss J., "Extensions of Lipschitz mappings into a Hilbert space," *Contemporary Mathematics*, vol. 26, pp. 189 – 206, 1984.

Johnson-Lindenstrauss Lemma

Consider $\mathcal{S} \subset \mathbb{R}^N$ containing P points.

We can embed $\mathcal{S}$ in $\mathbb{R}^M$ such that for all x, y in $\mathcal{S}$:

$$(1-\epsilon)\|x-y\|_2^2 \leq \|f(x)-f(y)\|_2^2 \leq (1+\epsilon)\|x-y\|_2^2$$

using only $M = O\left(\frac{\log P}{\epsilon^2}\right)$ dimensions

Later results

$f(x)$ can be linear $f(x) = \mathrm{A}x$, randomized A achieves bound (e.g., entries Gaussian, +1/-1 Bernoulli, etc.)

Bound (almost) tight: $M = O\left(\frac{\log P}{\epsilon^2 \log \frac{1}{\epsilon}}\right)$ dimensions necessary

BUT: Quantization is necessary for transmission!

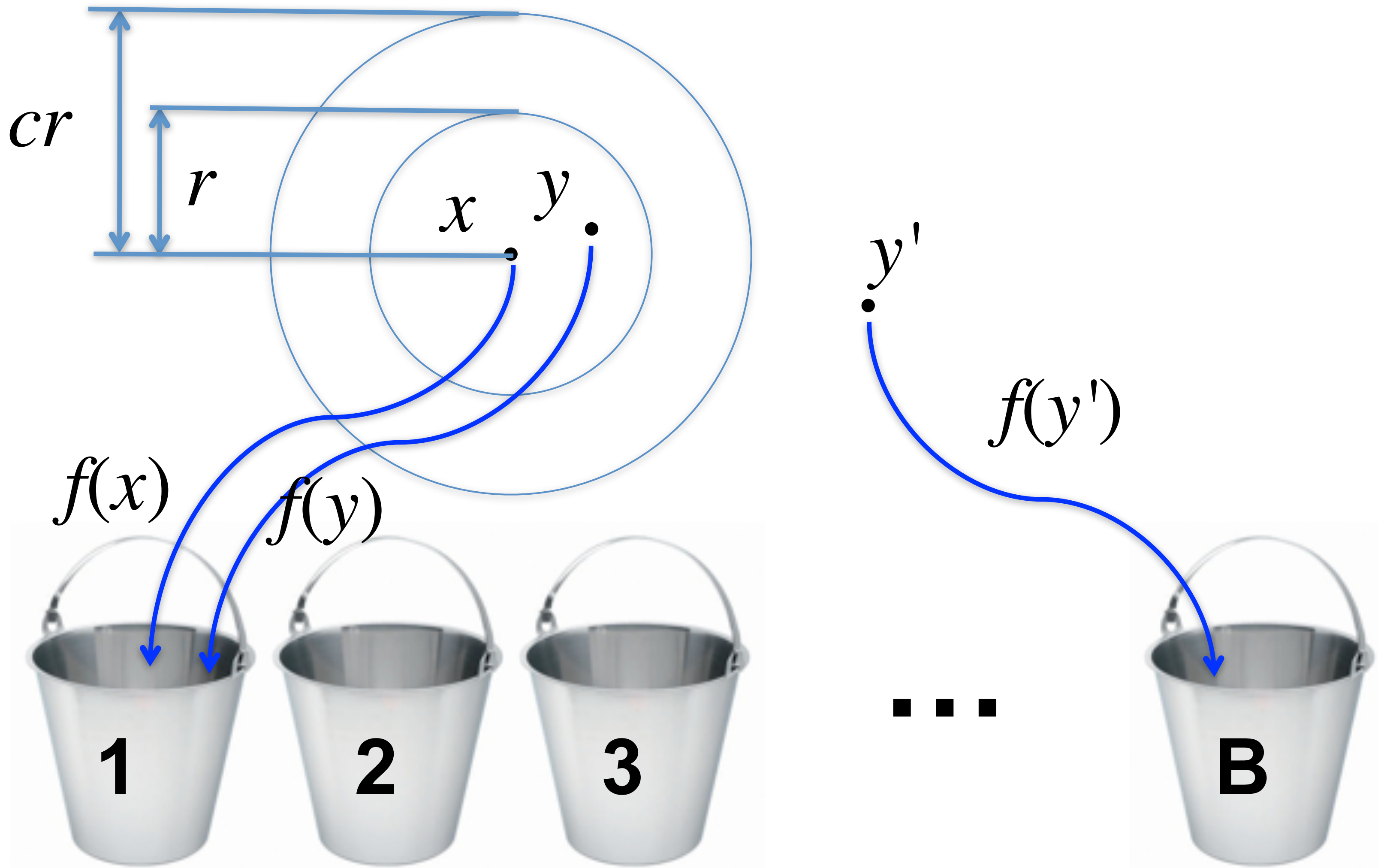
Are J-L Embeddings still appropriate?

Locality Sensitive Hashing

Randomized signal hash $f\colon \mathbb{R}^N \to \mathbb{N}$ such that:

$d(x,y) \le r \Rightarrow f(x) = f(y)$ with high probability

$d(x,y) \ge cr \Rightarrow f(x) \ne f(y)$ with high probability

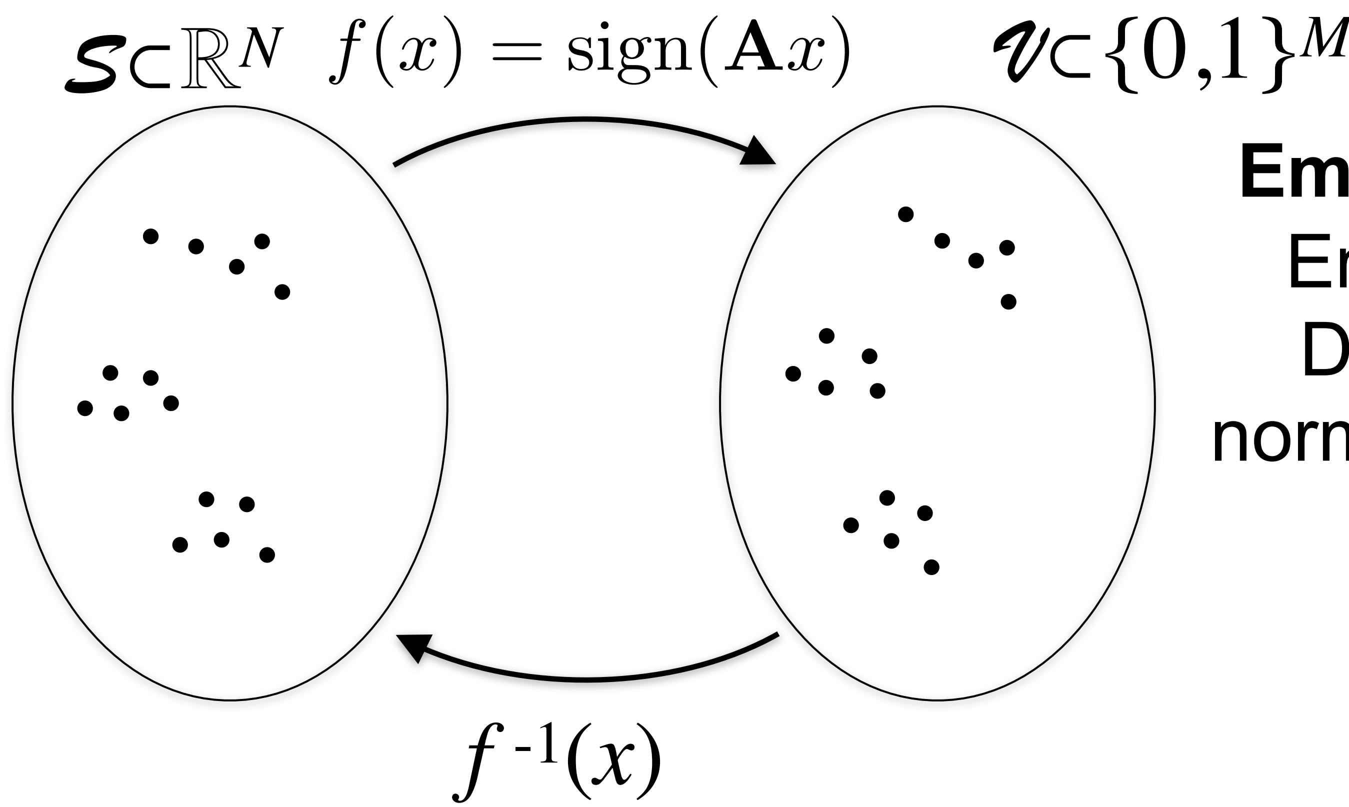


(One) LSH approach: random projection and quantization, i.e., Quantized Johnson-Lindenstrauss

- Andoni, A. and Indyk, P., "Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions," *Commun. ACM*, vol. 51, no. 1, pp. 117–122, 2008.
- Datar M., Immorlica N., Indyk P., and Mirrokni V., "Locality-Sensitive Hashing Scheme Based on p-Stable Distributions," *Proc. Symposium on Computational Geometry*, 2004

Binary Stable Embedding

Original space
Distance metric:
angle between
K-sparse unit norm
vectors in $\mathbb{R}^N$



Embedding space
Embed in $\{0,1\}^M$
Distance metric
normalized hamming
distance

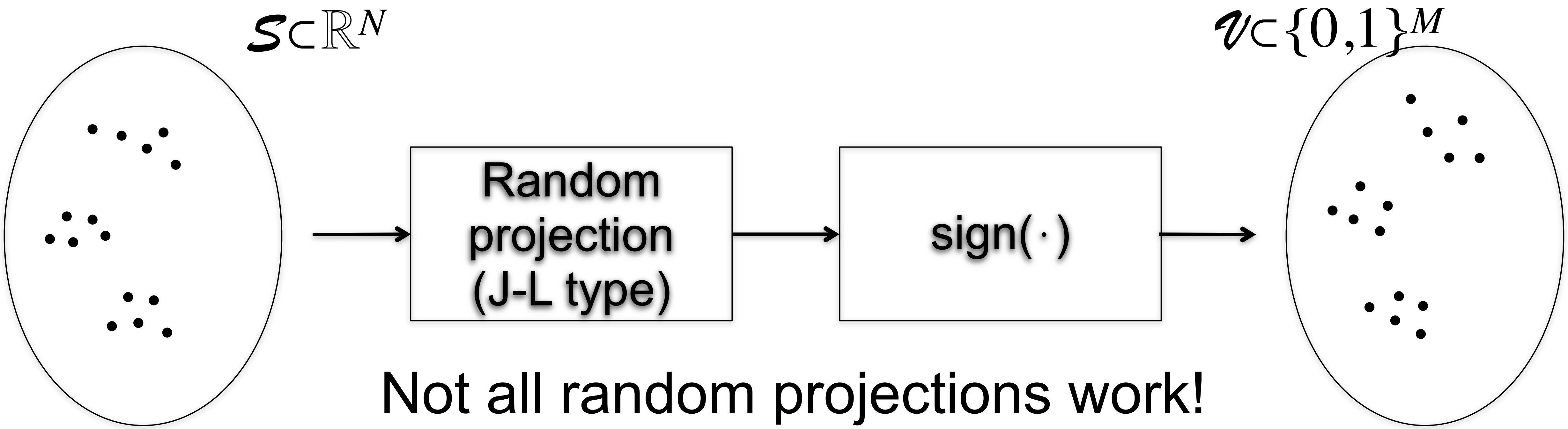
Embedding preservers angles for all K-sparse x,y in $\mathbb{R}^N$:

$$\arccos\left(\frac{\langle x, y\rangle}{\|x\|\|y\|}\right) - \epsilon \leq d_H(f(x), f(y)) \leq \arccos\left(\frac{\langle x, y\rangle}{\|x\|\|y\|}\right) + \epsilon$$

using only $M = O\left(\frac{1}{\epsilon^2}\left(K\log N + K\log\frac{1}{\epsilon}\right)\right)$ measurements

- Jacques L., Laska J. N., Boufounos P. T., Baraniuk R. J., "Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors," *IEEE Trans. Info. Theory*, v. 59, no. 4, April, 2013.

Binary Stable Embedding



Not all random projections work!

Matrices w/ i.i.d. Gaussian entries work

w/ i.i.d. Bernoulli they don’t always

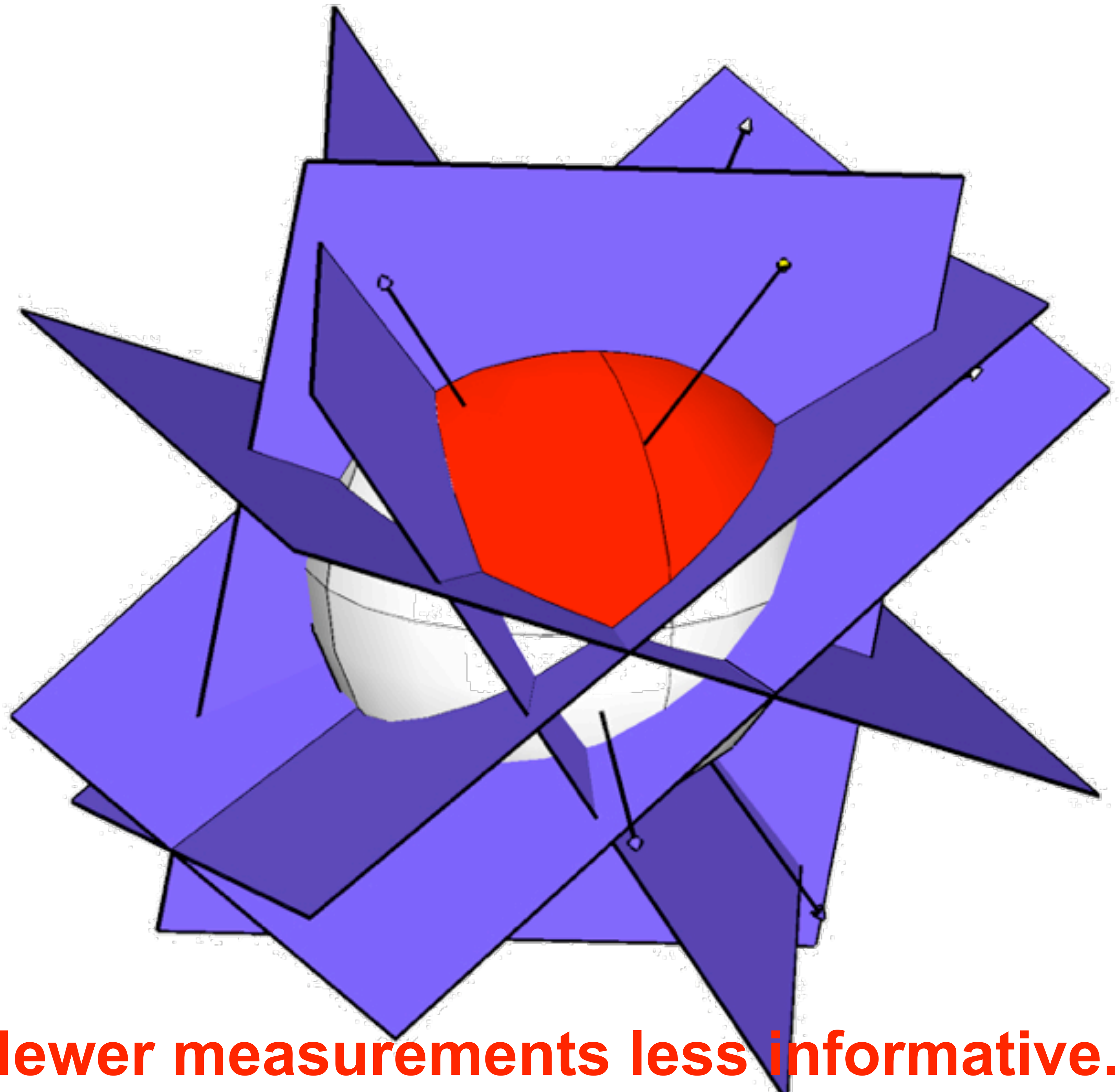
Sufficient information for sparse recovery (previous part)

Embedding does not preserve amplitudes

Is embedding rate-efficient?

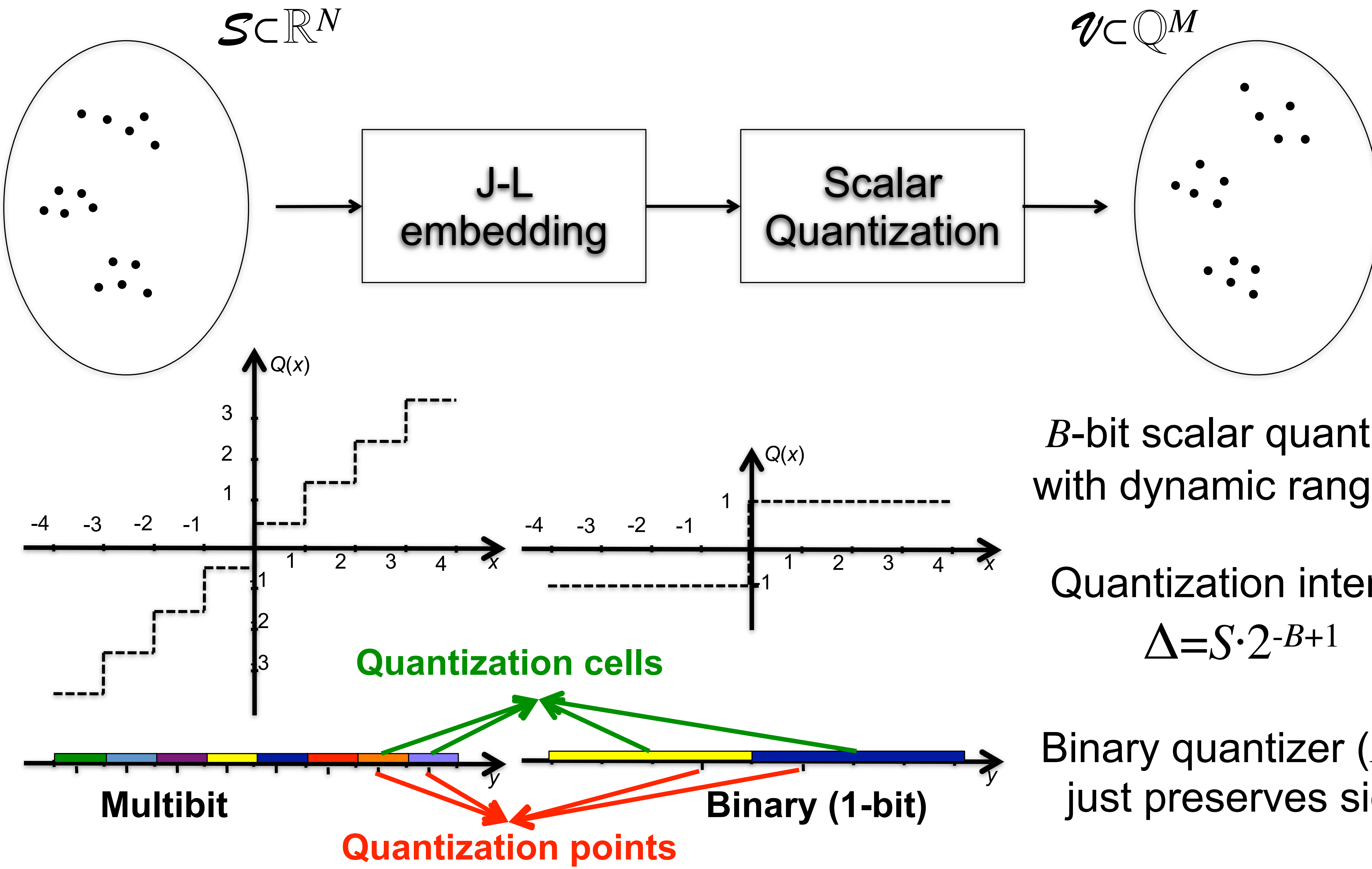
- Plan, Y. and Vershynin, R., “Dimension reduction by random hyperplane tessellations,” preprint, arXiv:1111.4452, 2011.
- Ai, A., Lapanowski, A., Plan, Y., Vershynin, R, “One-bit compressed sensing with non-Gaussian measurements”, *Linear Algebra and Applications*, to appear.

Information in 1-bit Measurements



Newer measurements less informative.
Chance of intersection increasingly smaller
Embedding rate-inefficient!

Quantized J-L Embeddings



B-bit scalar quantizer with dynamic range $\pm S$

Quantization interval $\Delta = S\cdot 2^{-B+1}$

Binary quantizer ($B=1$) just preserves sign

- Li M., Rane S., and Boufounos P. T., "Quantized embeddings of scale-invariant image features for mobile augmented reality," *IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*, Banff, Canada, Sept. 17-19, 2012

Johnson-Lindenstrauss With Quantization [w/ Li, Rane]

Consider $\mathcal{S} \subset \mathbb{R}^N$ containing P points.

We can embed $\mathcal{S}$ in $\mathbb{R}^M$ such that for all x,y in $\mathcal{S}$:

$$
\begin{aligned}
(1-\epsilon)\|x-y\|_2 - 2^{-B+1}S \leq \\
\|Q(f(x)) - Q(f(y))\|_2 \\
\leq (1+\epsilon)\|x-y\|_2 + 2^{-B+1}S
\end{aligned}
$$

using only $M = O\left(\frac{\log P}{\epsilon^2}\right)$ dimensions

and B bits per dimension

(with appropriate normalizations/saturation levels)

Total rate: R=BM

Quantized J-L at Fixed Rate

Given total rate: $R=MB$

How to assign B and M? More M or more B?

Larger B, less quantization distortion

$$2^{-B+1}S$$

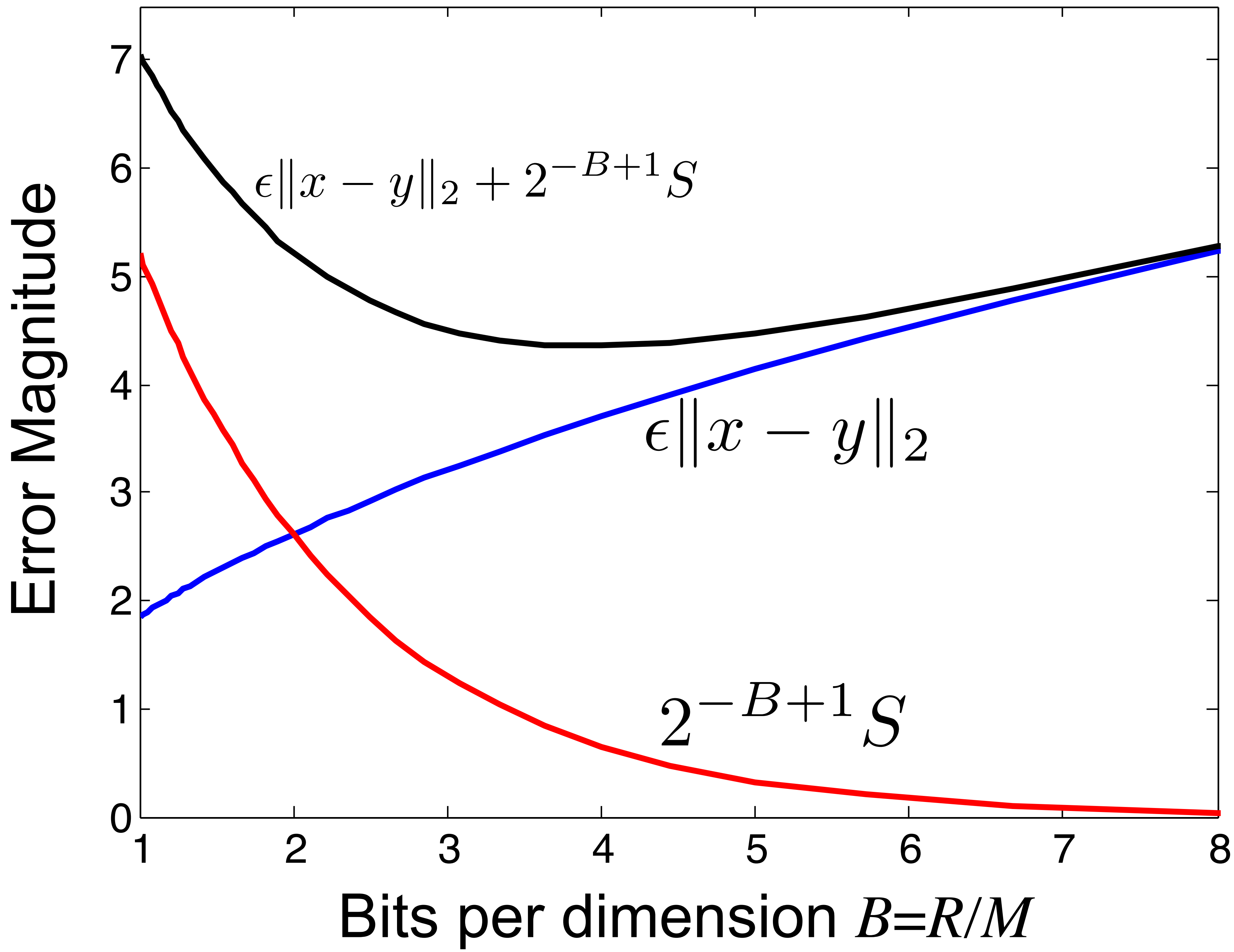
$$(1-\epsilon)\|x-y\|_2 - 2^{-\frac{R}{M}+1}S \leq \|Q(f(x)) - Q(f(y))\|_2 \leq (1+\epsilon)\|x-y\|_2 + 2^{-\frac{R}{M}+1}S$$

Larger M, less J-L type distortion ϵ

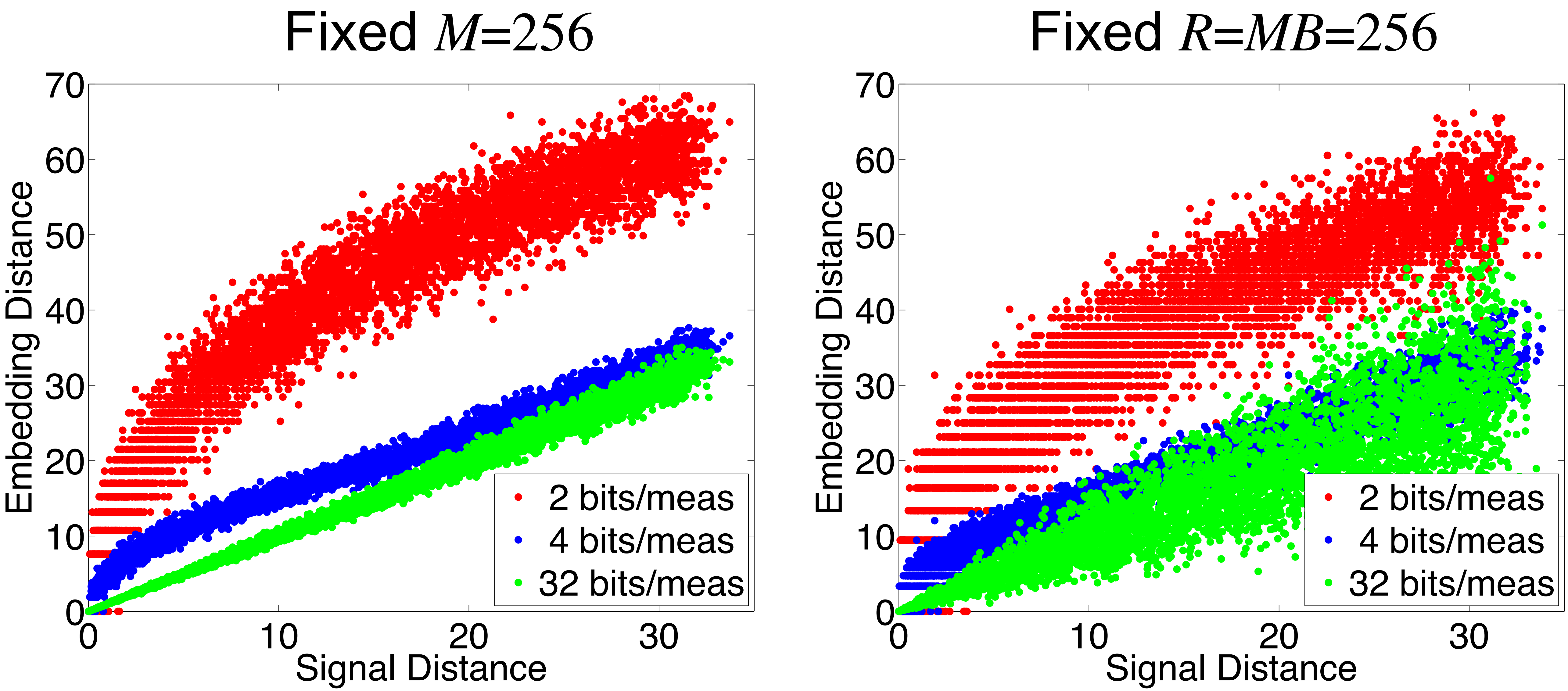
$$\epsilon = O(1/\sqrt{M})$$

Design tradeoff:
Number of projections vs. bits per projection

Exploring the Design Trade-off



Exploring the Design Trade-off



IN PRACTICE

The Augmented Reality Problem

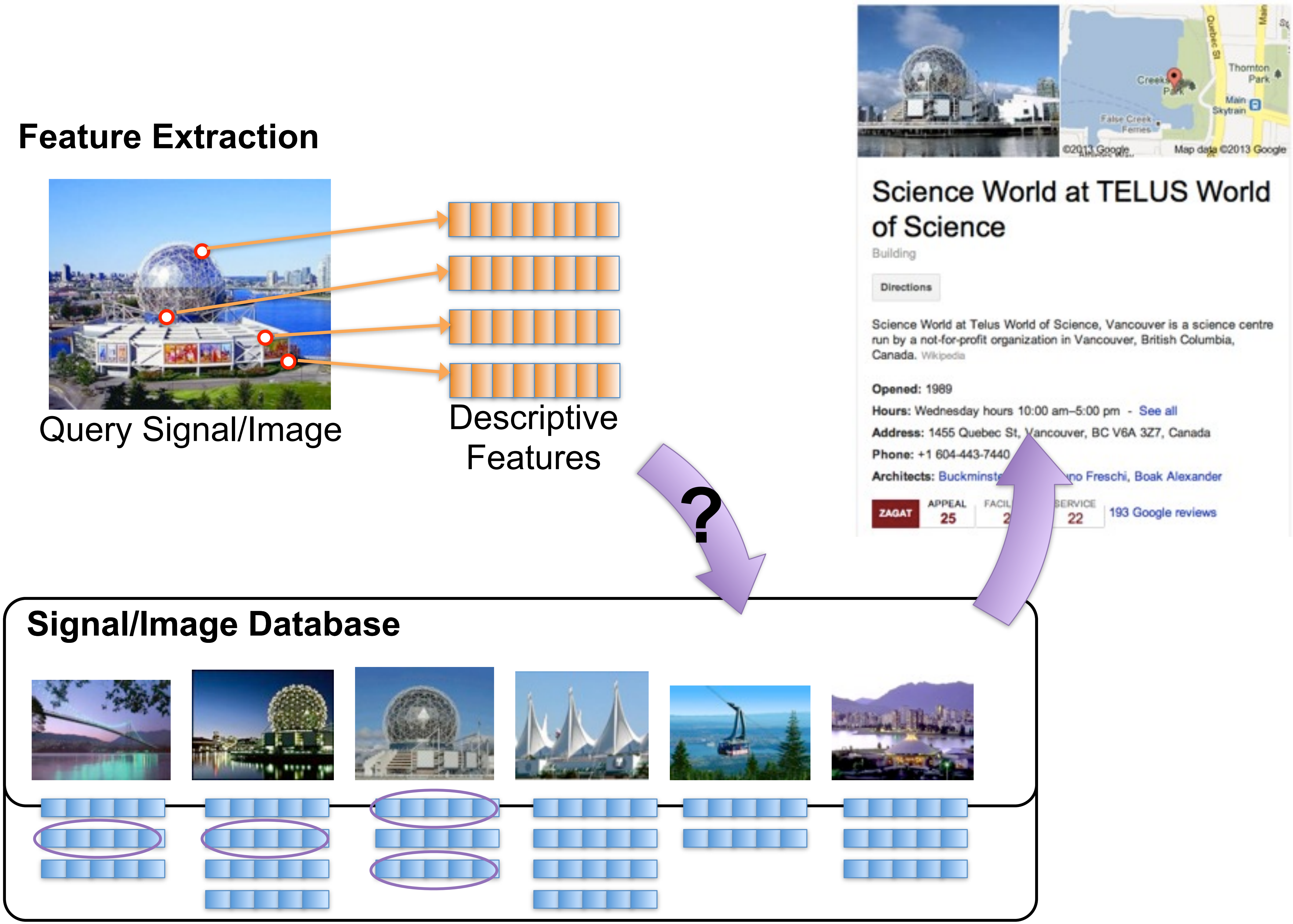
Server-side processing increasingly important (e.g. cloud computing, augmented reality)

Compression is necessary

Goal: detection; not image transmission

Q: Should we transmit the signal?
Can we reduce the rate?

Signal/Image-based Retrieval



Detection/Classification Pipeline (typical)

Feature Extraction

Detection/ Classification

"Golden Gate Bridge"

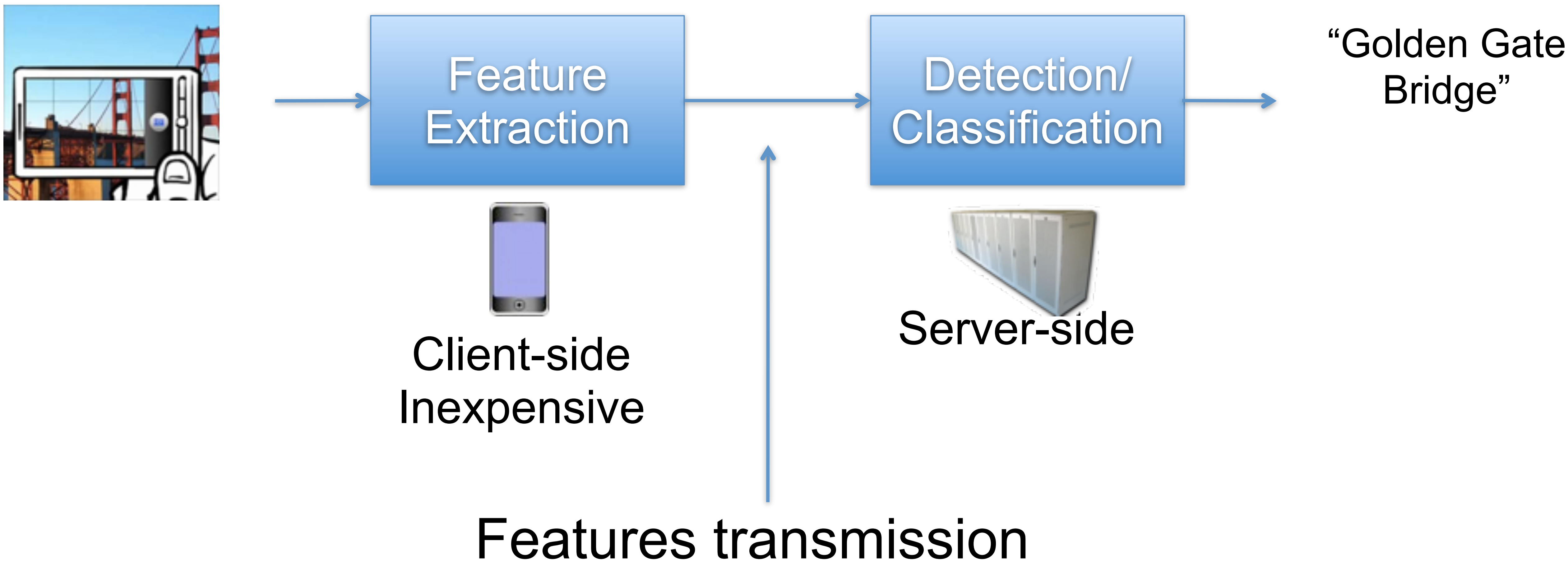
Server-side
Requires reconstruction

Server-side

Image transmission

Detection/Classification: Based on distance/inner product

Detection/Classification Pipeline (efficient)

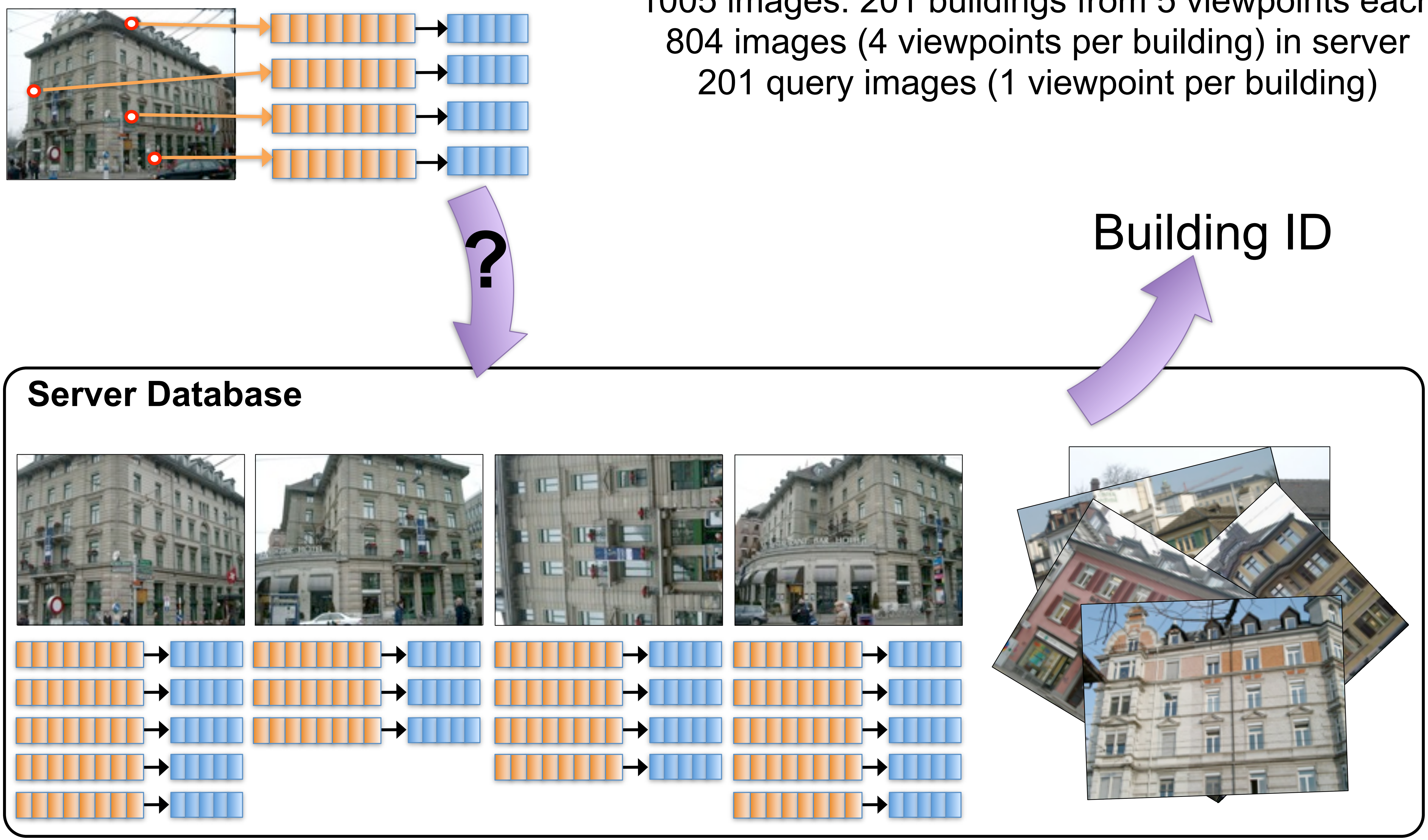


Detection/Classification: Based on distance/inner product

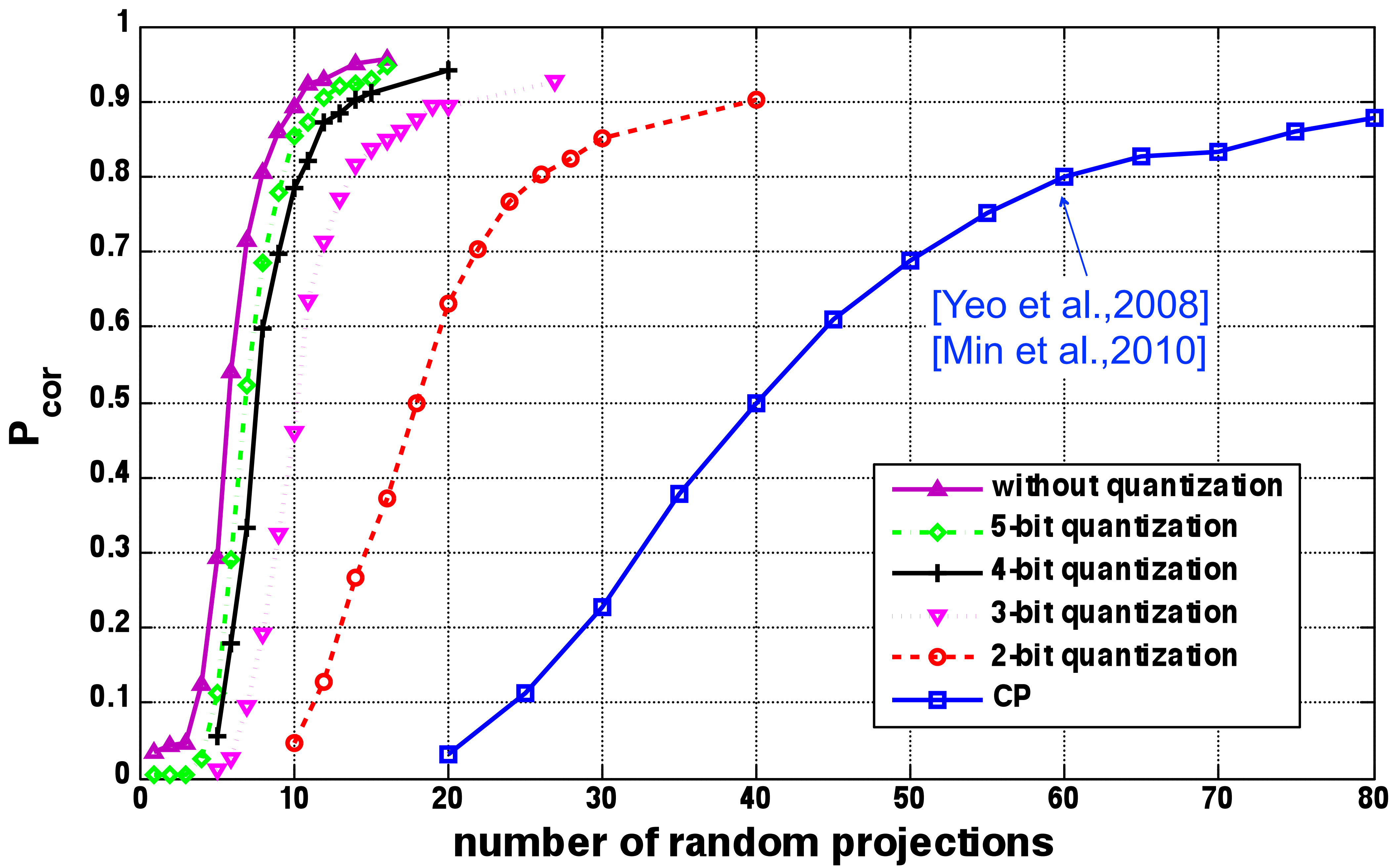
Goal: rate-efficient distance-preserving transmission

ZuBuD: Zurich Buildings Database

1005 images: 201 buildings from 5 viewpoints each
804 images (4 viewpoints per building) in server
201 query images (1 viewpoint per building)

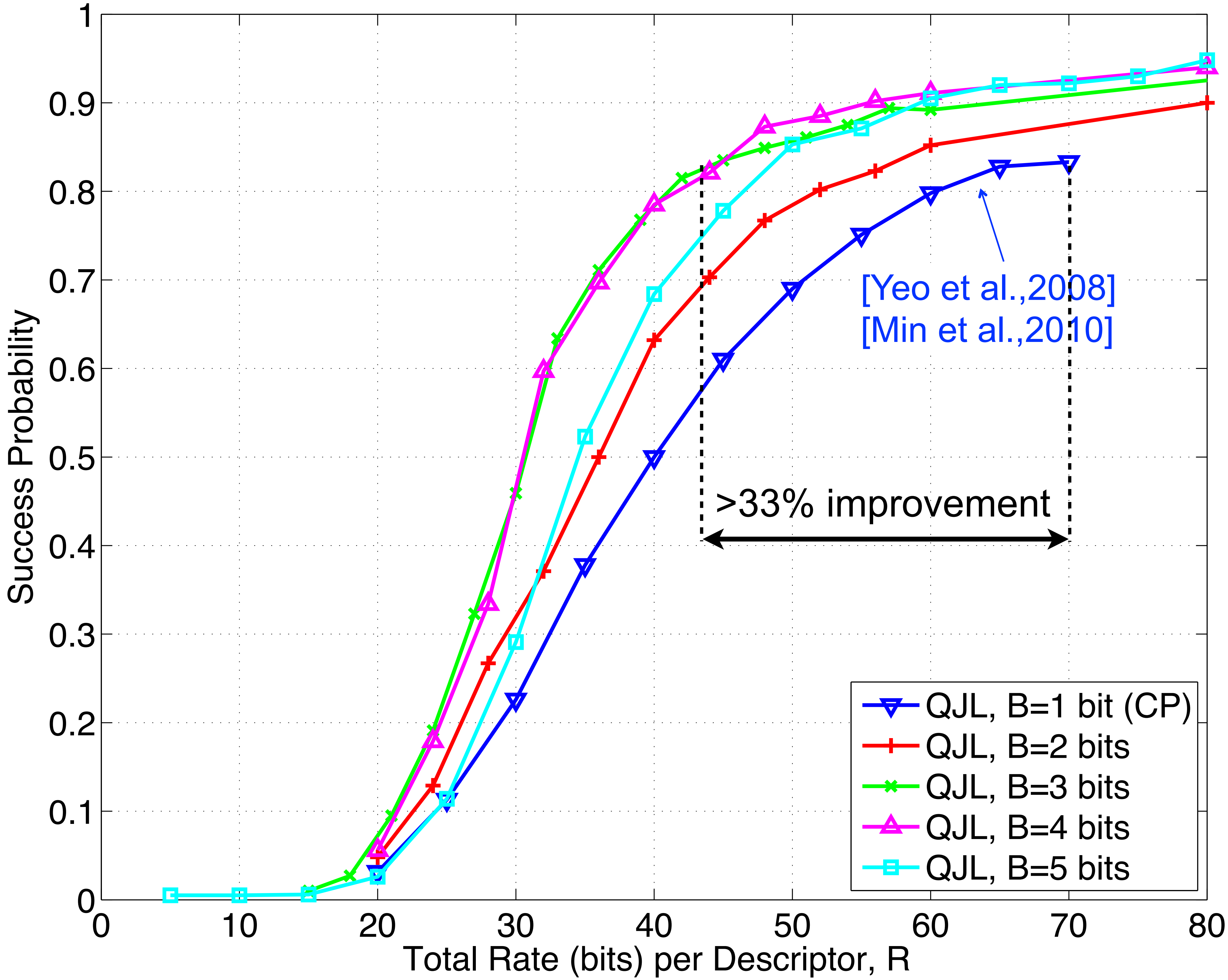


Success Probability

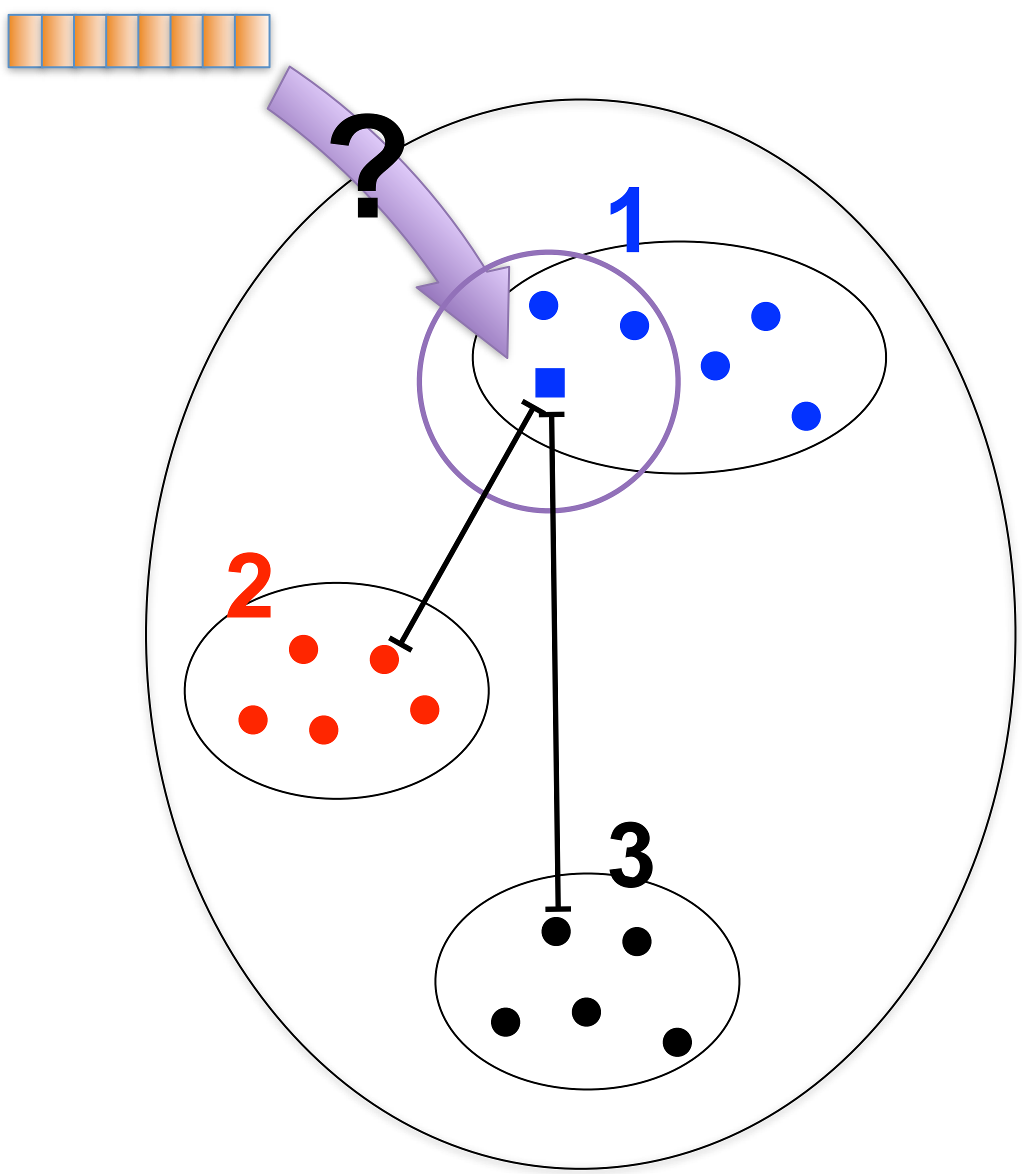


- Yeo C., Ahammad P., and Ramchandran K., “Coding of image feature descriptors for distributed rate-efficient visual correspondences,” *International Journal of Computer Vision*, vol. 94, pp. 267–281, 2011, 10.1007/s11263-011-0427-1.
- Min K., Yang L., Wright J., Wu L., Hua X.-S., and Ma Y., “Compact projection: Simple and efficient near neighbor search with practical memory requirements,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2010.

Success Probability with Fixed Rate



Information Scalability

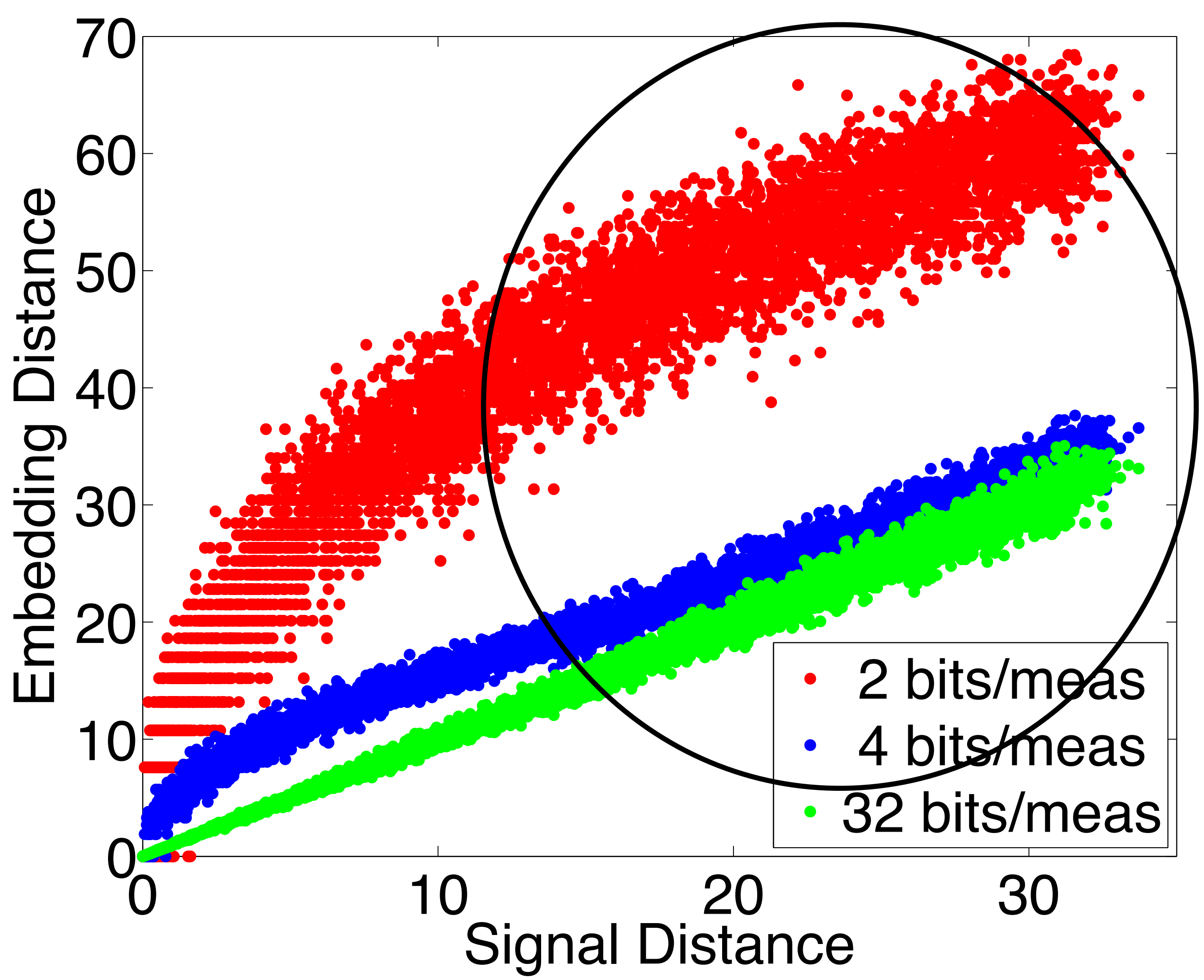


Inference relies on clusters of signals

Large distances not necessary to determine clusters and nearest neighbors

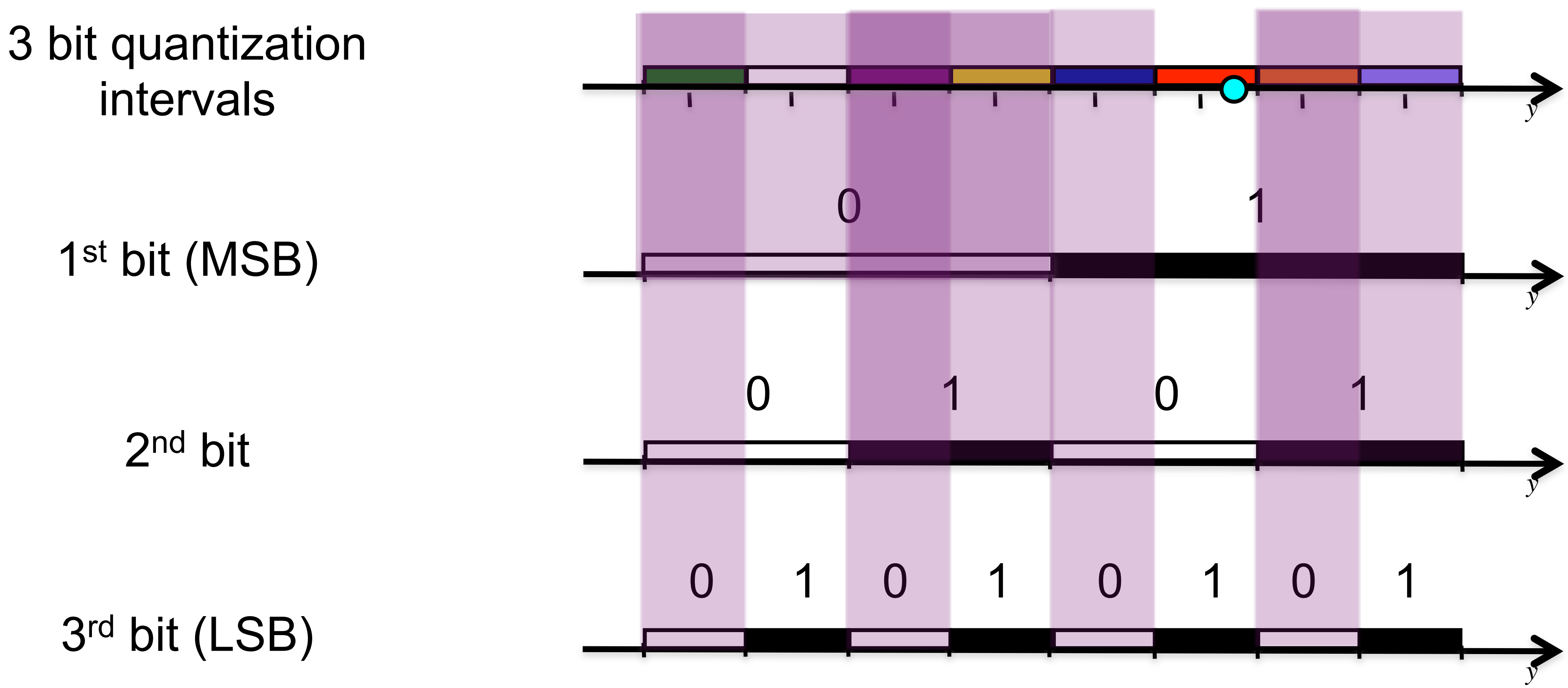
Should not spend bits encoding large distances!

But how?

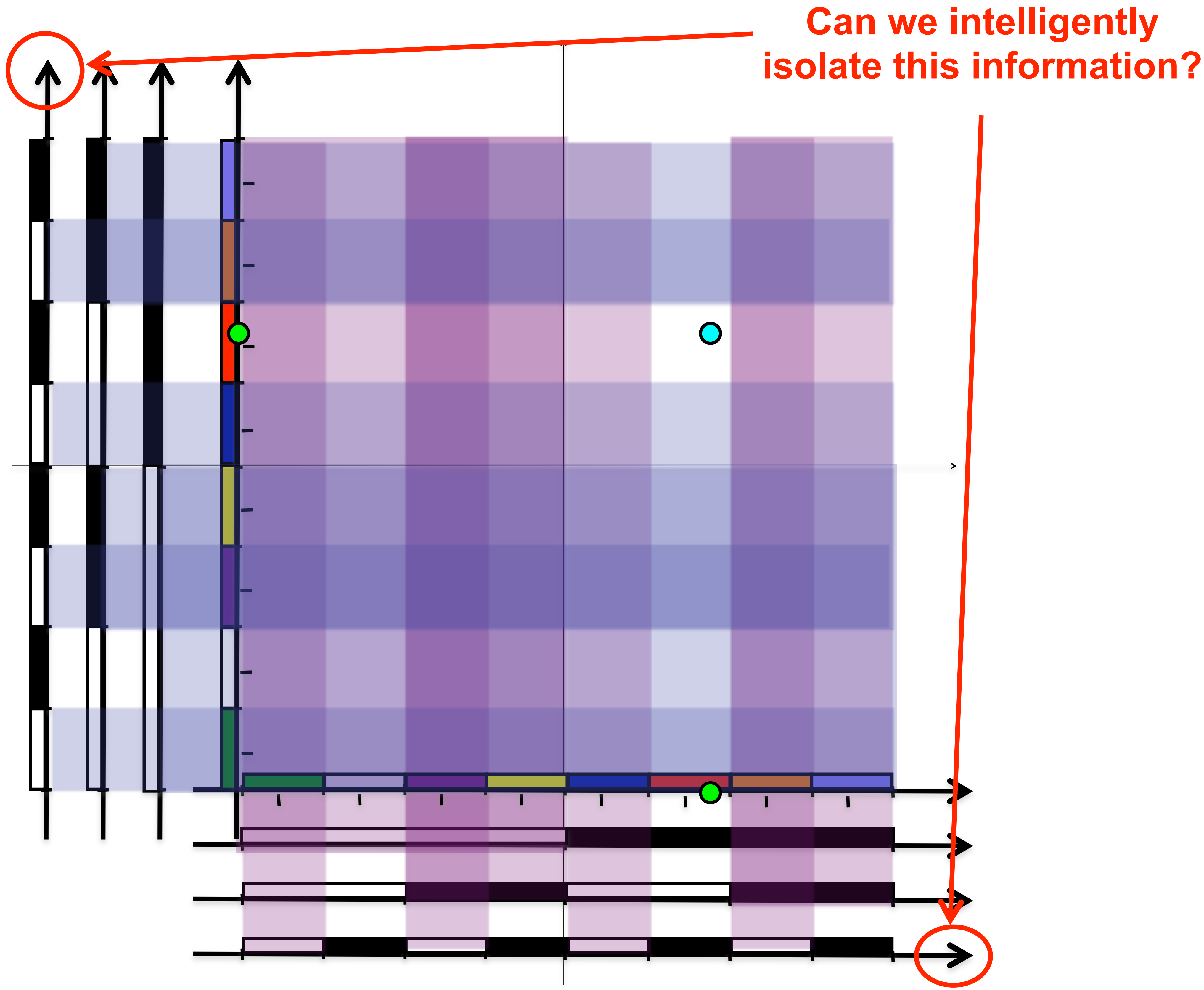


UNIVERSAL QUANTIZED EMBEDDINGS

What can a bit tell us?

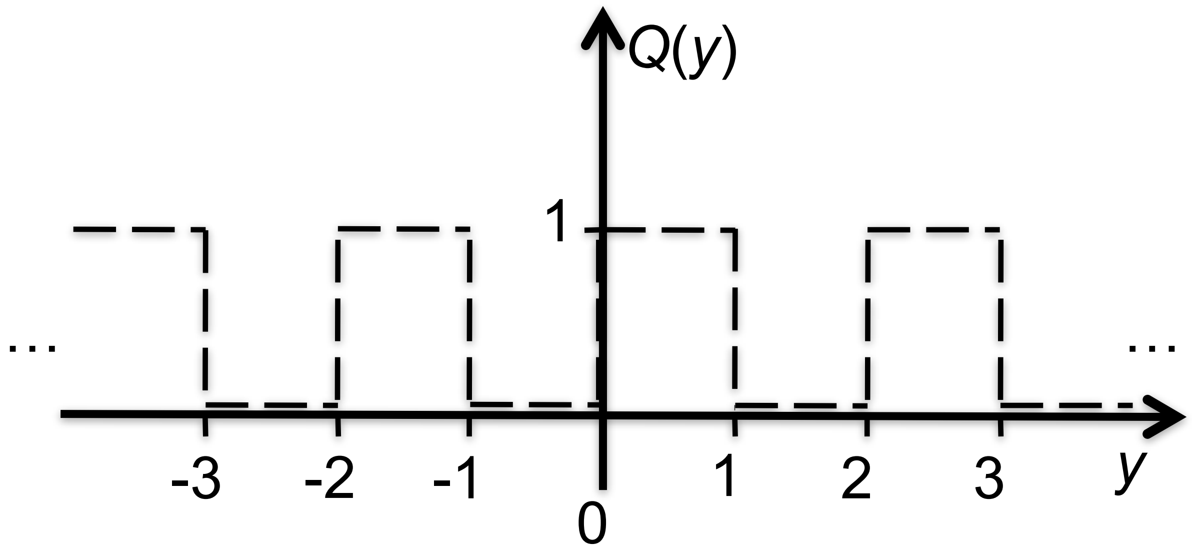


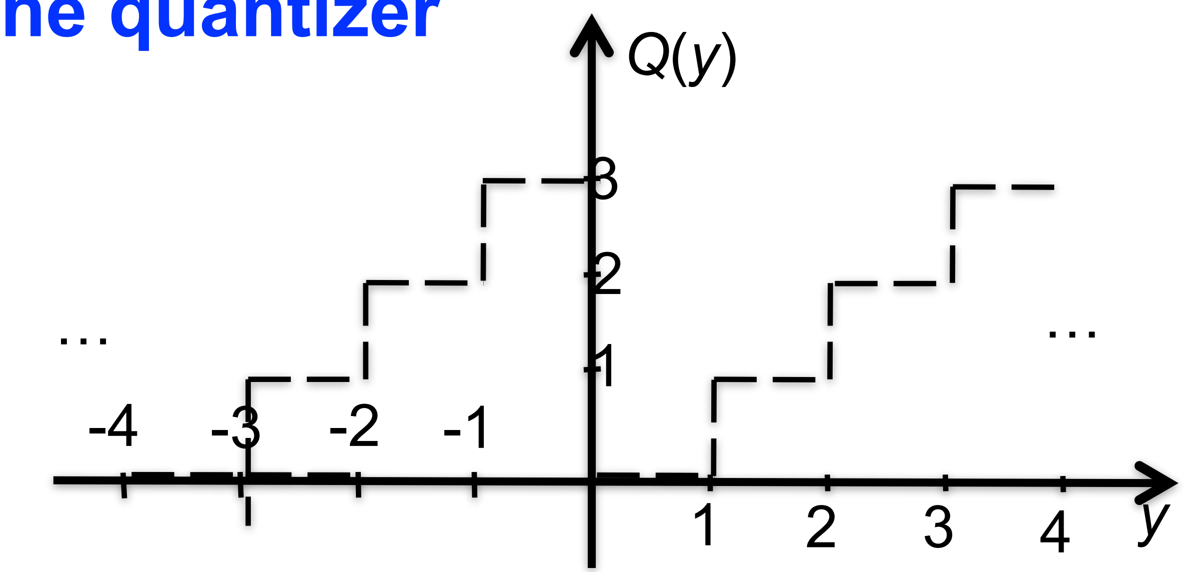
What can a bit tell us?



Rate-Efficient Scalar Quantization

Solution: **Modify the quantizer**





Non-monotonic quantizer: Multiple intervals quantize to same value

(Focus on 1-bit quantizer today)

measurements
(w/ i.i.d. gaussian matrix)

dither
(i.i.d. uniform)

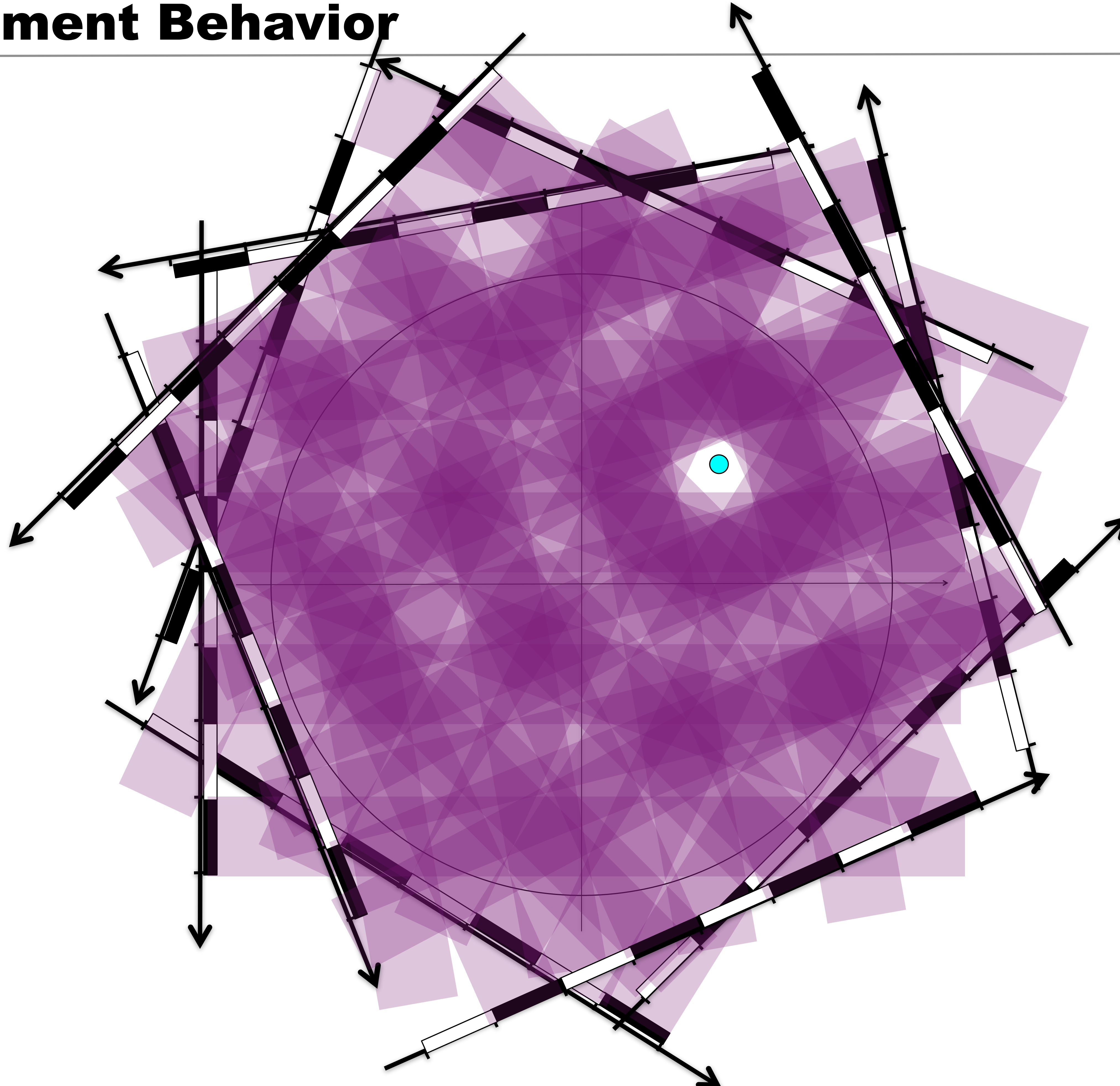
$$q_m = Q\left(\frac{\langle \mathbf{x}, \mathbf{a}_m \rangle + w_m}{\Delta_m}\right), \qquad \mathbf{q} = Q(\Delta^{-1}(\mathbf{A}\mathbf{x} + \mathbf{w}))$$

scalar quantizer
(non-monotonic)

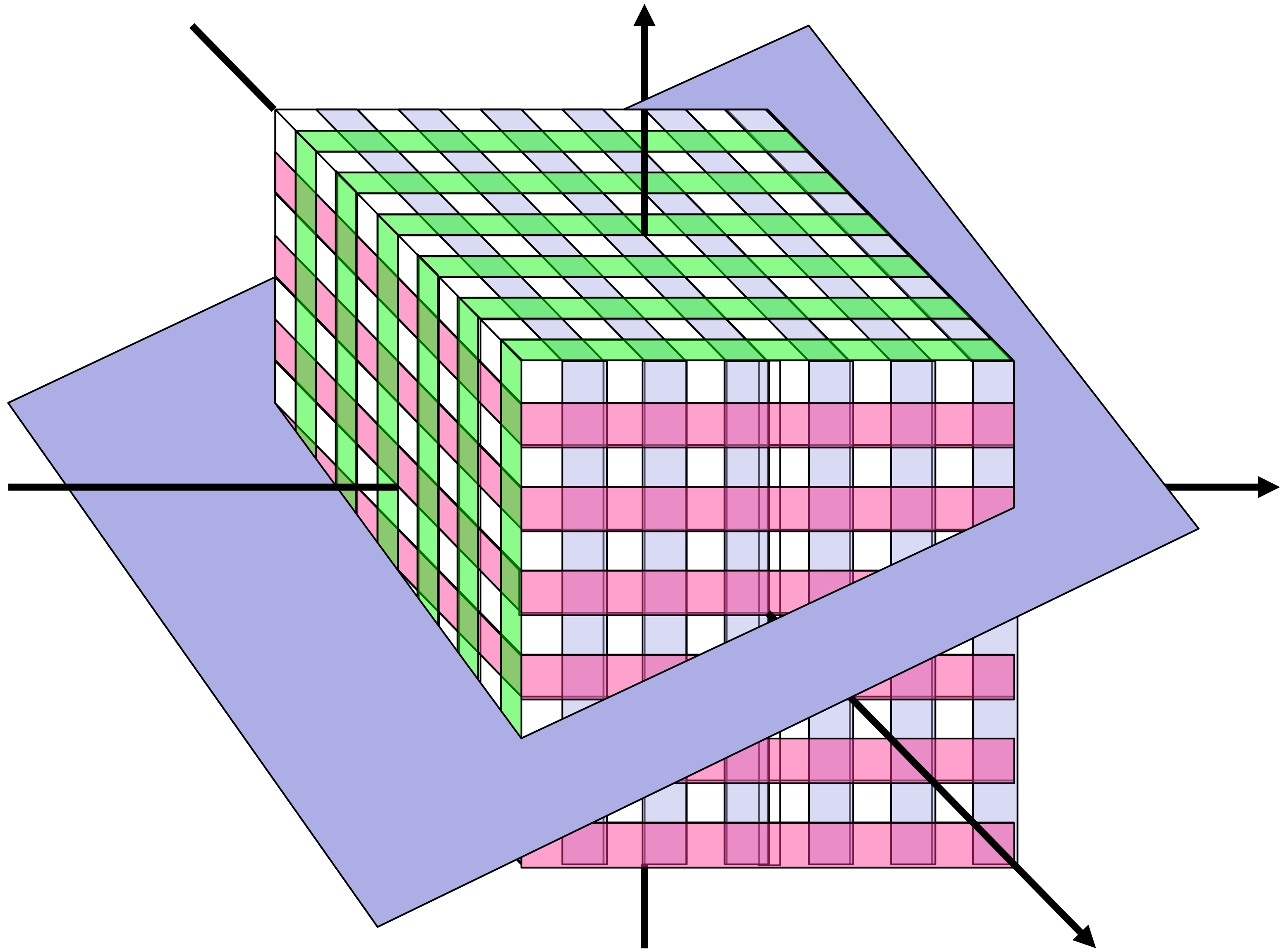
scaling/precision parameter
(Δ_m=Δ, same for all measurements)

- Boufounos P. T., "Universal Rate-Efficient Scalar Quantization," *IEEE Trans. Info. Theory*, v. 58, no. 3, pp. 1861-1872, March, 2012.

Measurement Behavior



Quantizer Geometry (1 bit)



Quantization cells are **not** continuous
Signal subspace intersects **most** of them

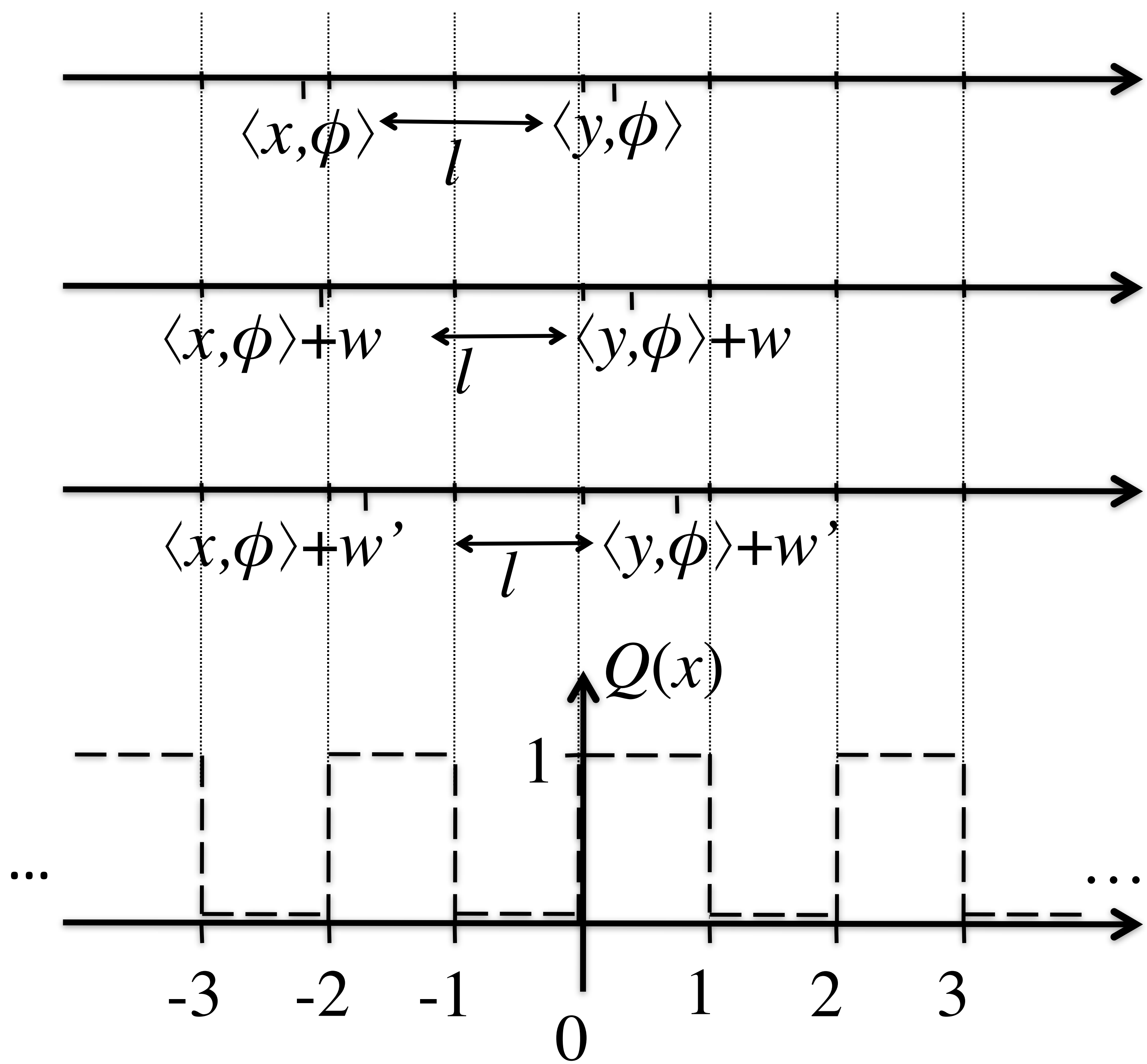
Pairs of Signals, Single Measurement

Projection (measurement)

Projection with dither

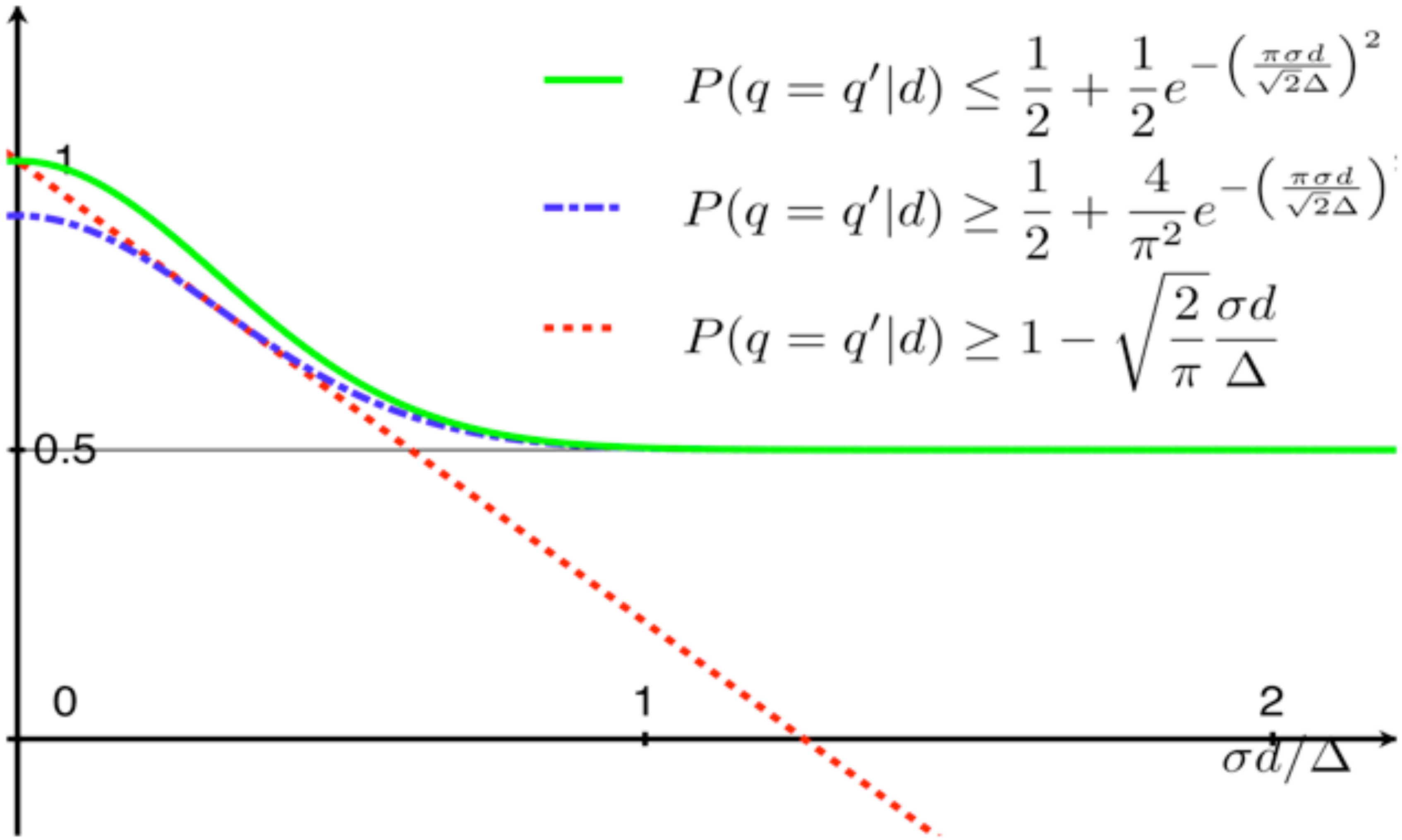
Projection with different dither

Quantization function



Pairs of Signals, Single Measurement

P(*q*=*q'*) : probability that a single measurement is consistent for a pair of signals, given their distance *d*

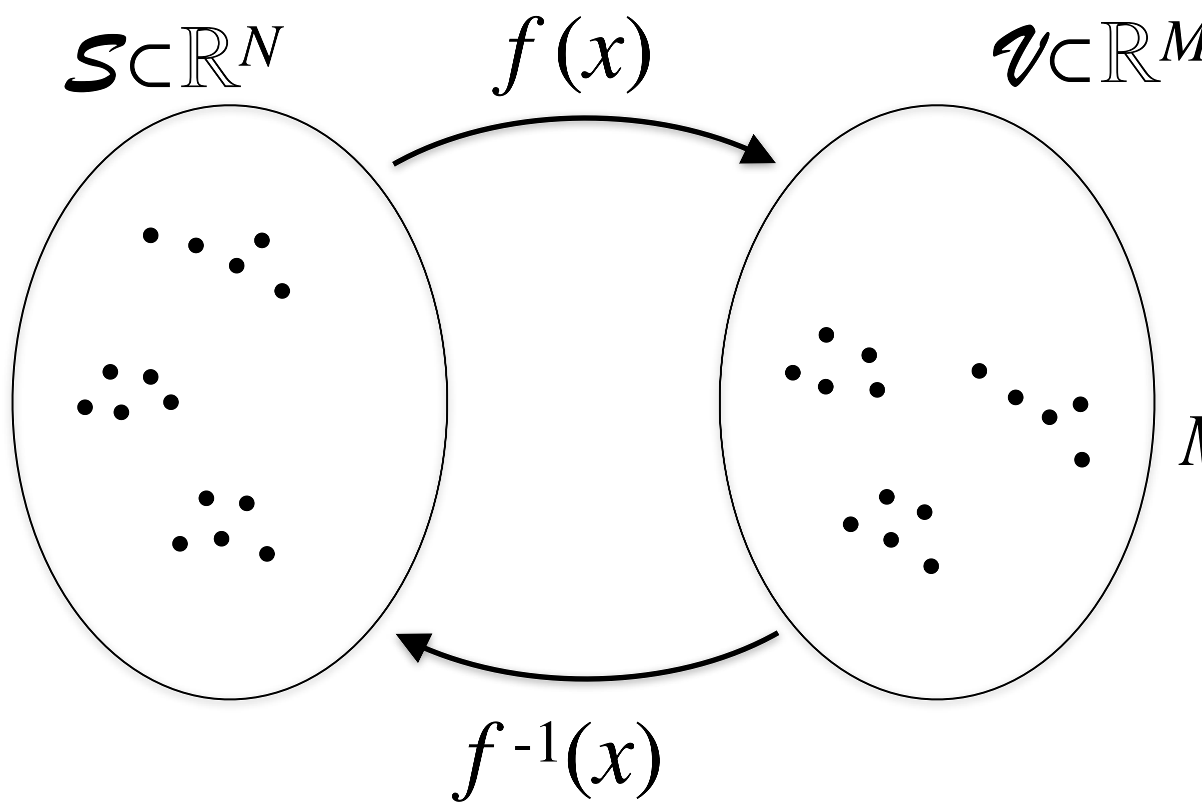


In other words:

Hamming distance of embedding is **proportional to ℓ_2 distance**

up to a point

Embedding Properties



Original space

Distance metric: ℓ_2

P points in $\mathbb{R}^N$

Embedding space

Embed in $\{0,1\}^M$

Hamming distance

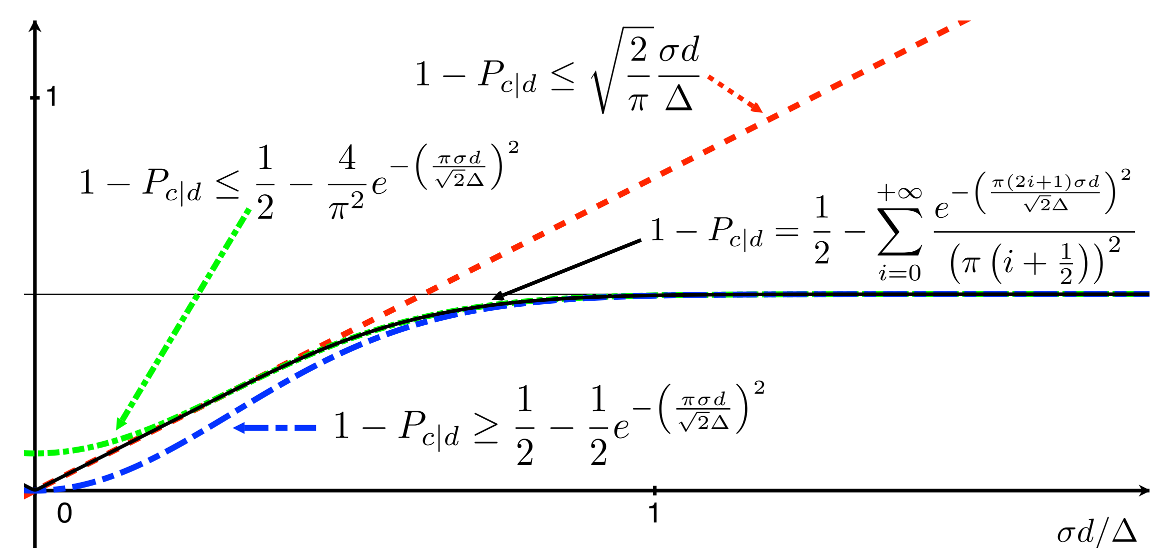
$M = O(\log P)$ dimensions

For all x,y in $\mathcal{S}$:

$$g(d) - \epsilon \leq d_H\left(f(x), f(y)\right) \leq g(d) + \epsilon$$

$$g(d) = 1 - P_{c|d}$$

as long as $M = O\left(\frac{1}{\epsilon^2}\log P\right)$

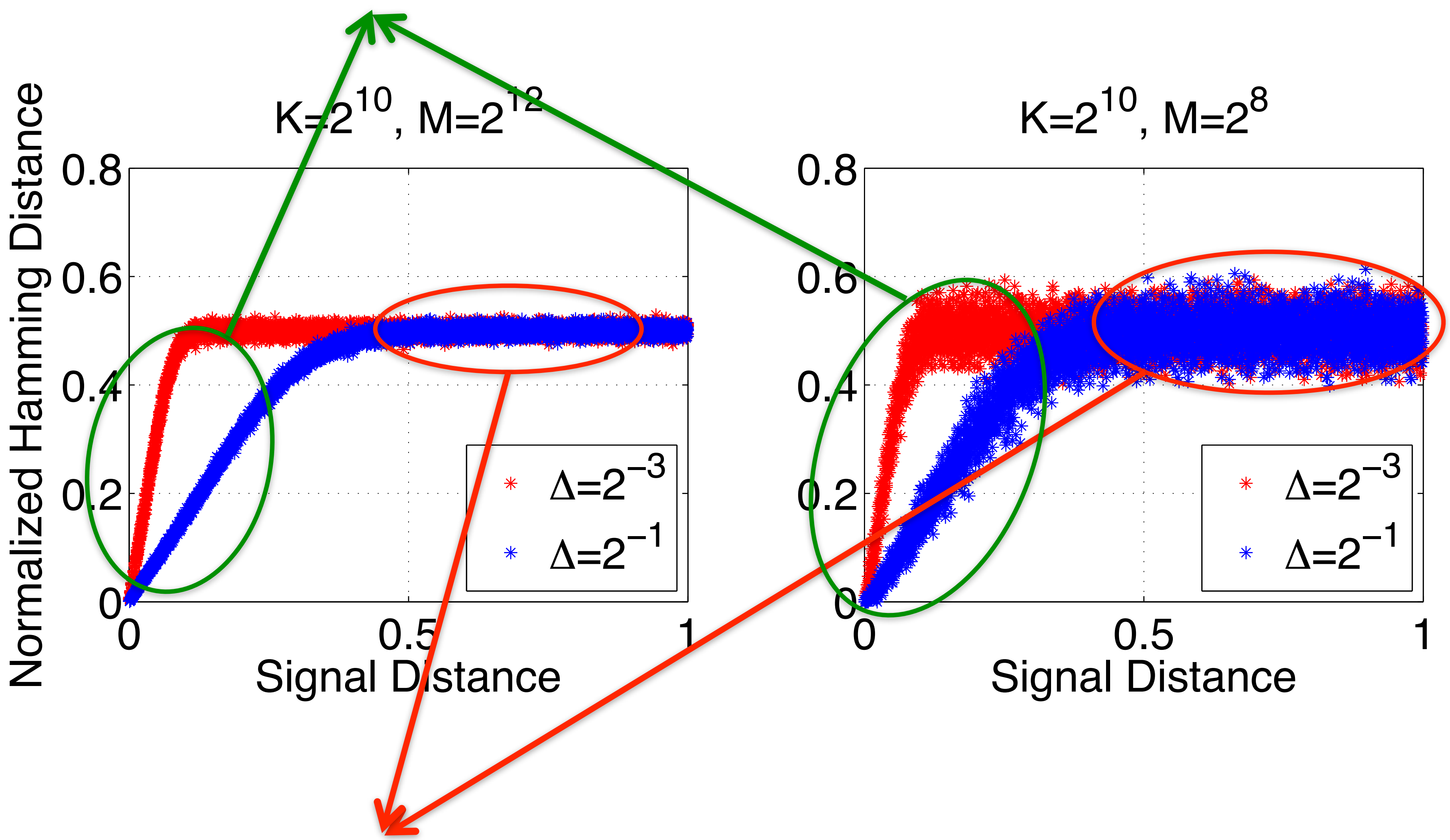


- Boufounos P. T. and Rane S., “Secure Binary Embeddings for Privacy Preserving Nearest Neighbors,” *Proc. Workshop on Information Forensics and Security (WIFS)*, Foz do Iguaçu, Brazil, November 29 – December 2, 2011.

Error Behavior

$$g(d) - \epsilon \leq d_H\left(f(x), f(y)\right) \leq g(d) + \epsilon$$

“Linear” region: $\ell_2 \propto d_H$, slope controlled by Δ



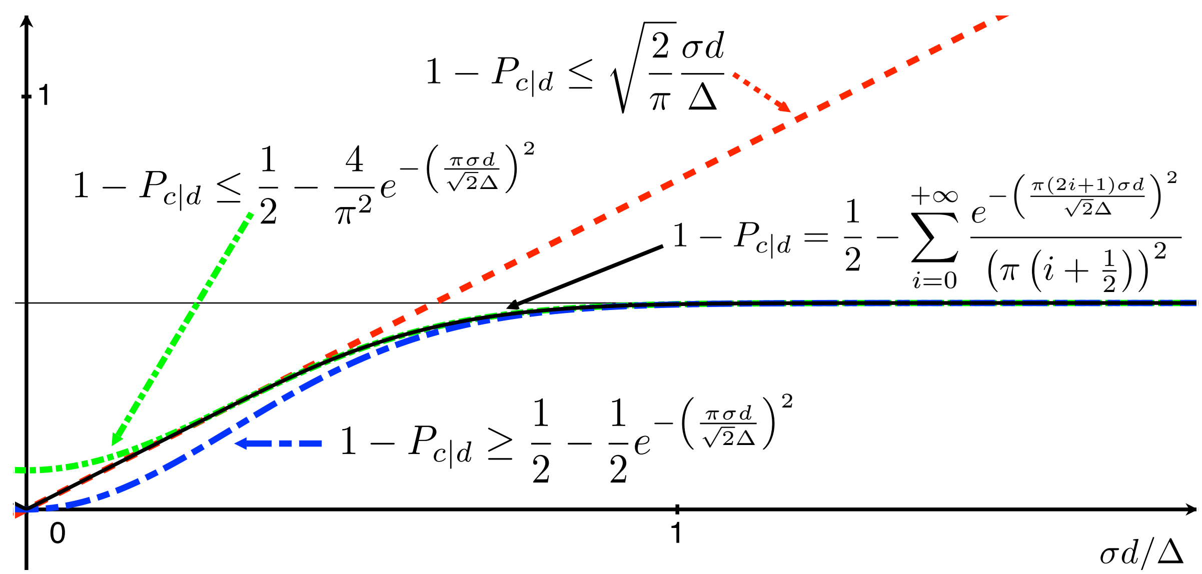
“Flat” region: no distance information

Error Behavior

$$g(d) - \epsilon \le d_H\left(f(x), f(y)\right) \le g(d) + \epsilon$$

$$M = O\left(\frac{1}{\epsilon^2}\log P\right)$$

Similar trade-off as J-L
but on $g(d)=1-P_{c|d}$



Distance estimate: $\widehat{d} = g^{-1}\left(d_H\left(f(x), f(y)\right)\right)$

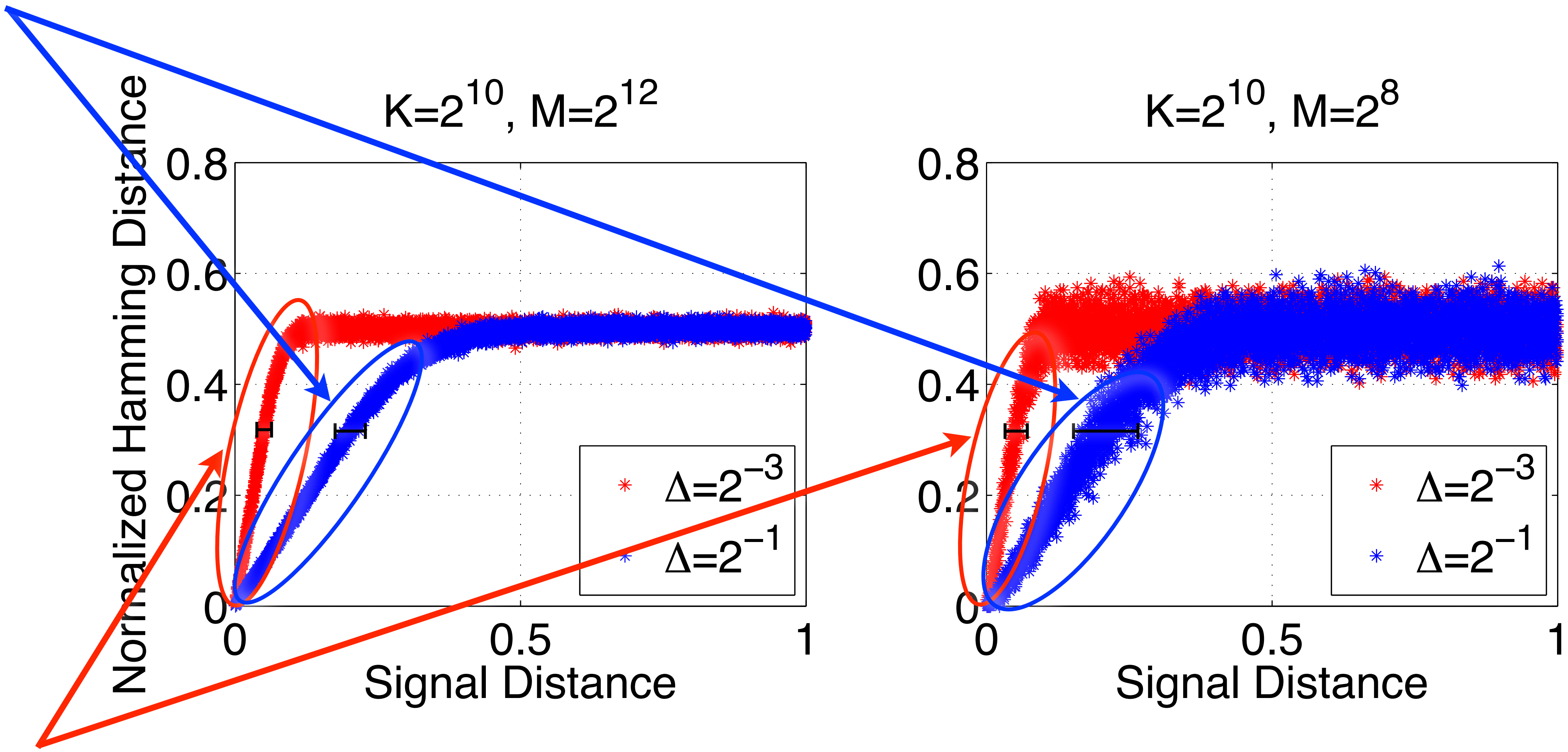
Estimate ambiguity: $\widehat{d} - \frac{\epsilon}{g'(\widehat{d})} \lesssim d \lesssim \widehat{d} + \frac{\epsilon}{g'(\widehat{d})}$

Properties (slope) controlled by choice of Δ

Error Behavior

$$g(d) - \epsilon \leq d_H\left(f(x), f(y)\right) \leq g(d) + \epsilon$$

Large Δ: small slope, more ambiguity, preserves larger distances

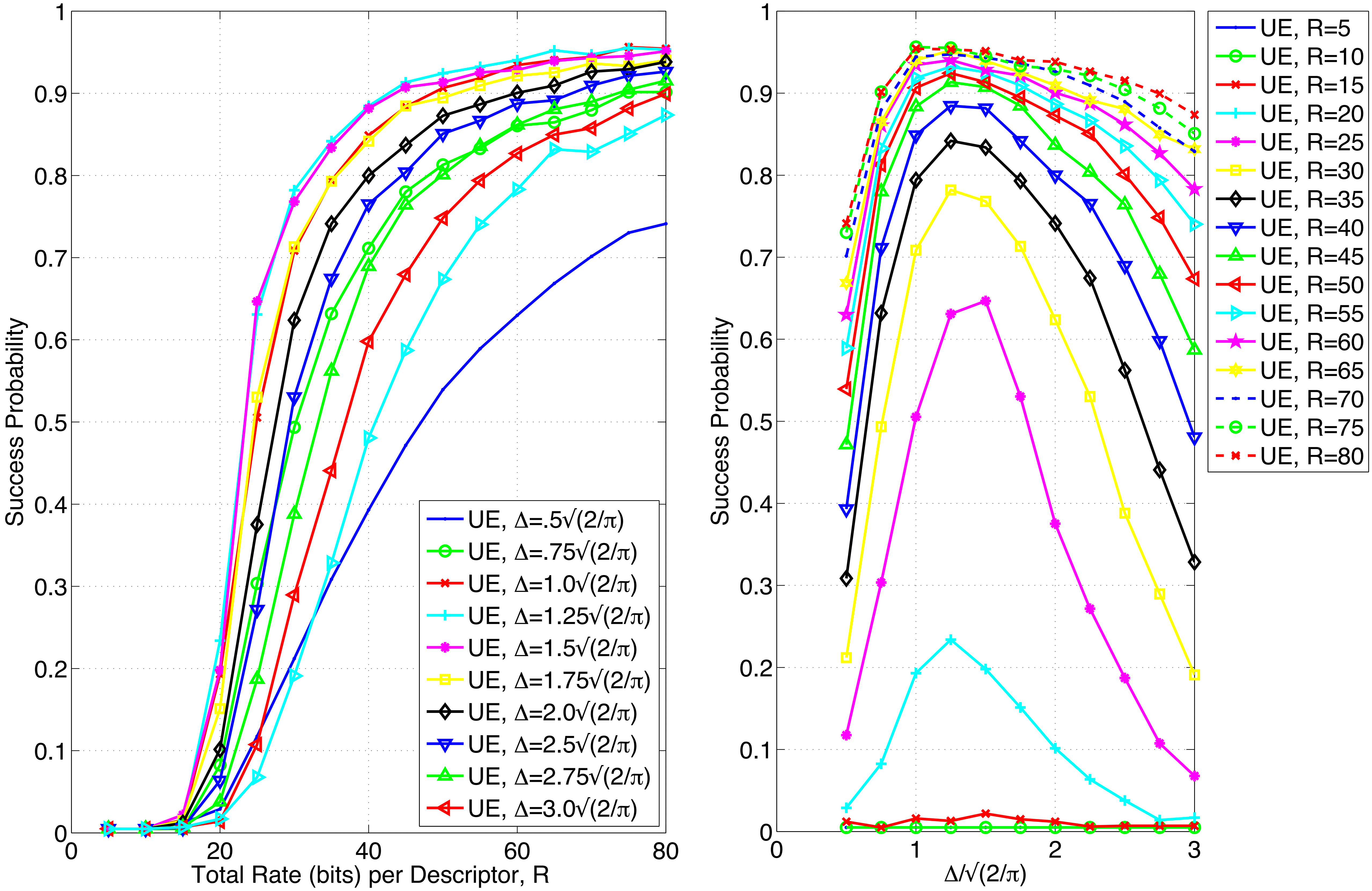


Small Δ: large slope, less ambiguity, preserves smaller distances

- Boufounos P. T. and Rane S., "Efficient Coding of Signal Distances Using Universal Quantized Embeddings," *Proc. Data Compression Conference (DCC)*, Snowbird, UT, March 20-22, 2013.

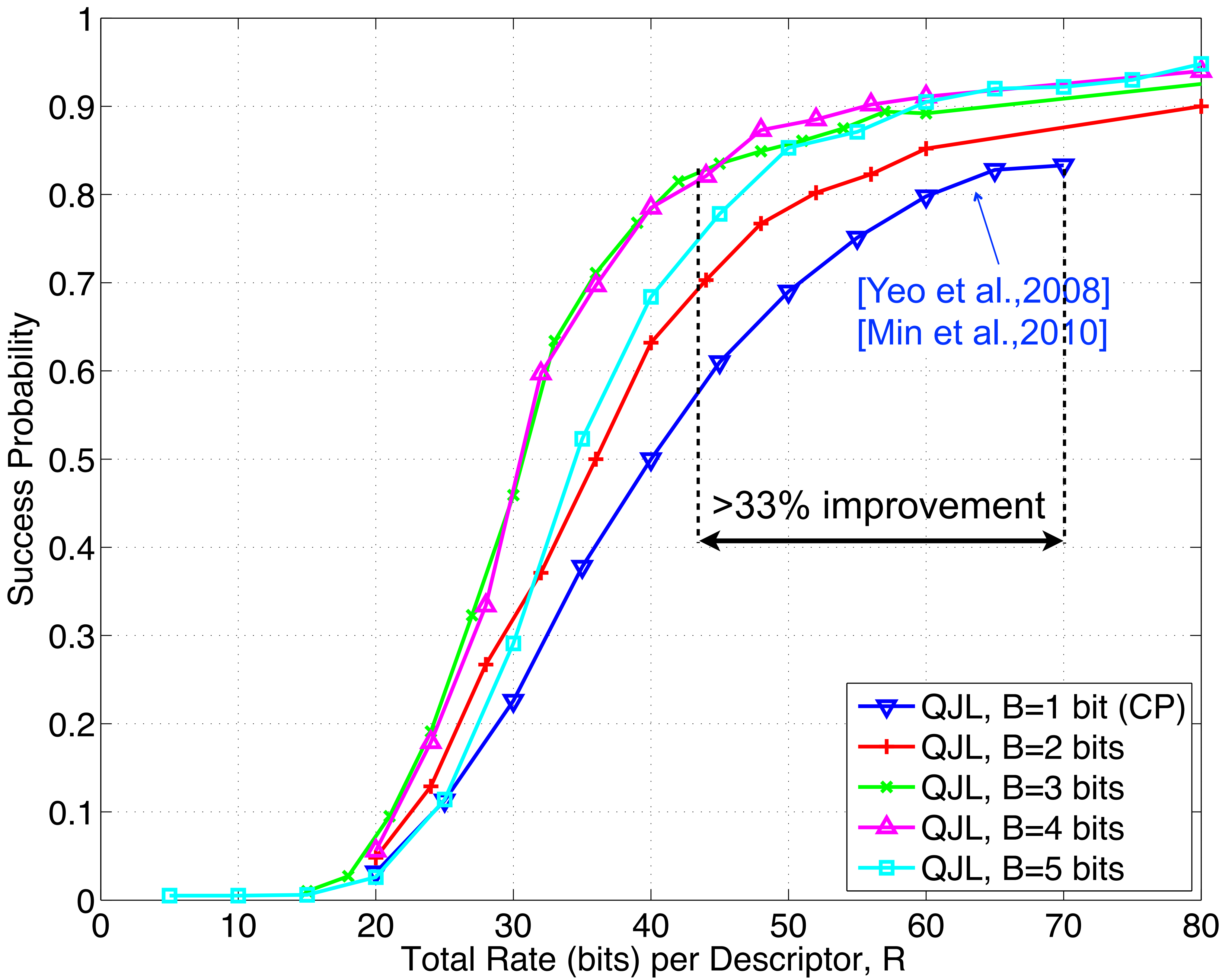
IN PRACTICE

In practice



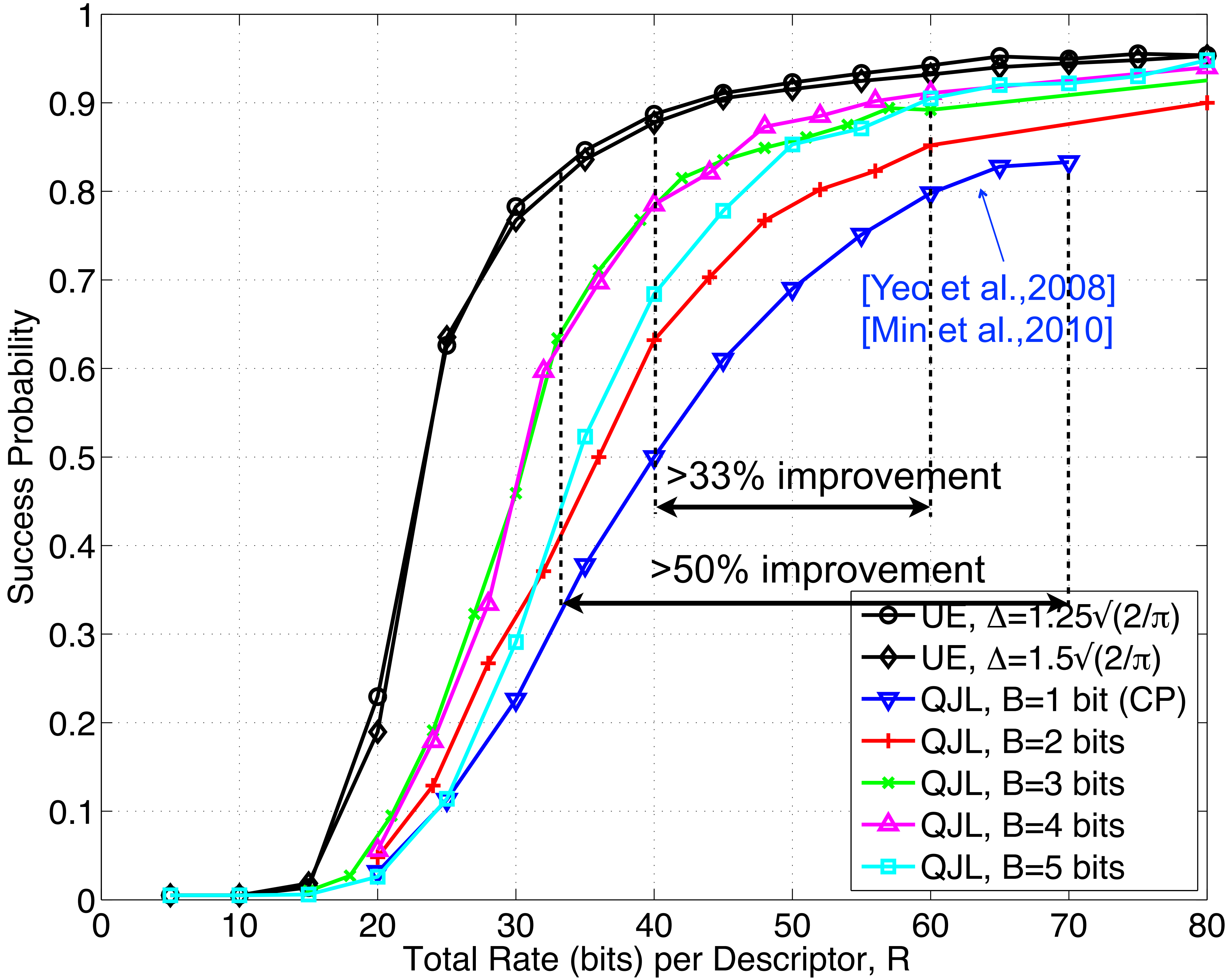
Success Probability
Total Rate (bits) per Descriptor, R
UE, Δ=.5√(2/π)
UE, Δ=.75√(2/π)
UE, Δ=1.0√(2/π)
UE, Δ=1.25√(2/π)
UE, Δ=1.5√(2/π)
UE, Δ=1.75√(2/π)
UE, Δ=2.0√(2/π)
UE, Δ=2.5√(2/π)
UE, Δ=2.75√(2/π)
UE, Δ=3.0√(2/π)
Success Probability
Δ/√(2/π)
UE, R=5
UE, R=10
UE, R=15
UE, R=20
UE, R=25
UE, R=30
UE, R=35
UE, R=40
UE, R=45
UE, R=50
UE, R=55
UE, R=60
UE, R=65
UE, R=70
UE, R=75
UE, R=80

In practice



Success Probability
Total Rate (bits) per Descriptor, R
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
10
20
30
40
50
60
70
80
[Yeo et al.,2008]
[Min et al.,2010]
>33% improvement
QJL, B=1 bit (CP)
QJL, B=2 bits
QJL, B=3 bits
QJL, B=4 bits
QJL, B=5 bits

In practice


Success Probability
Total Rate (bits) per Descriptor, R
>33% improvement
>50% improvement
[Yeo et al.,2008]
[Min et al.,2010]
UE, Δ=1.25√(2/π)
UE, Δ=1.5√(2/π)
QJL, B=1 bit (CP)
QJL, B=2 bits
QJL, B=3 bits
QJL, B=4 bits
QJL, B=5 bits

BEYOND EMBEDDINGS

Reconstruction

- **Consistent reconstruction**: find a signal that quantizes to same bits, i.e.,

$$\widehat{\mathbf{x}} \quad \text{s.t. } \mathbf{q} = Q\left(\Delta^{-1}\left(\Phi\widehat{\mathbf{x}} + \mathbf{w}\right)\right)$$

- Very **good theoretical guarantees**
  - **Exponential error decay** with number of bits $\varepsilon = O(c^{-B})$

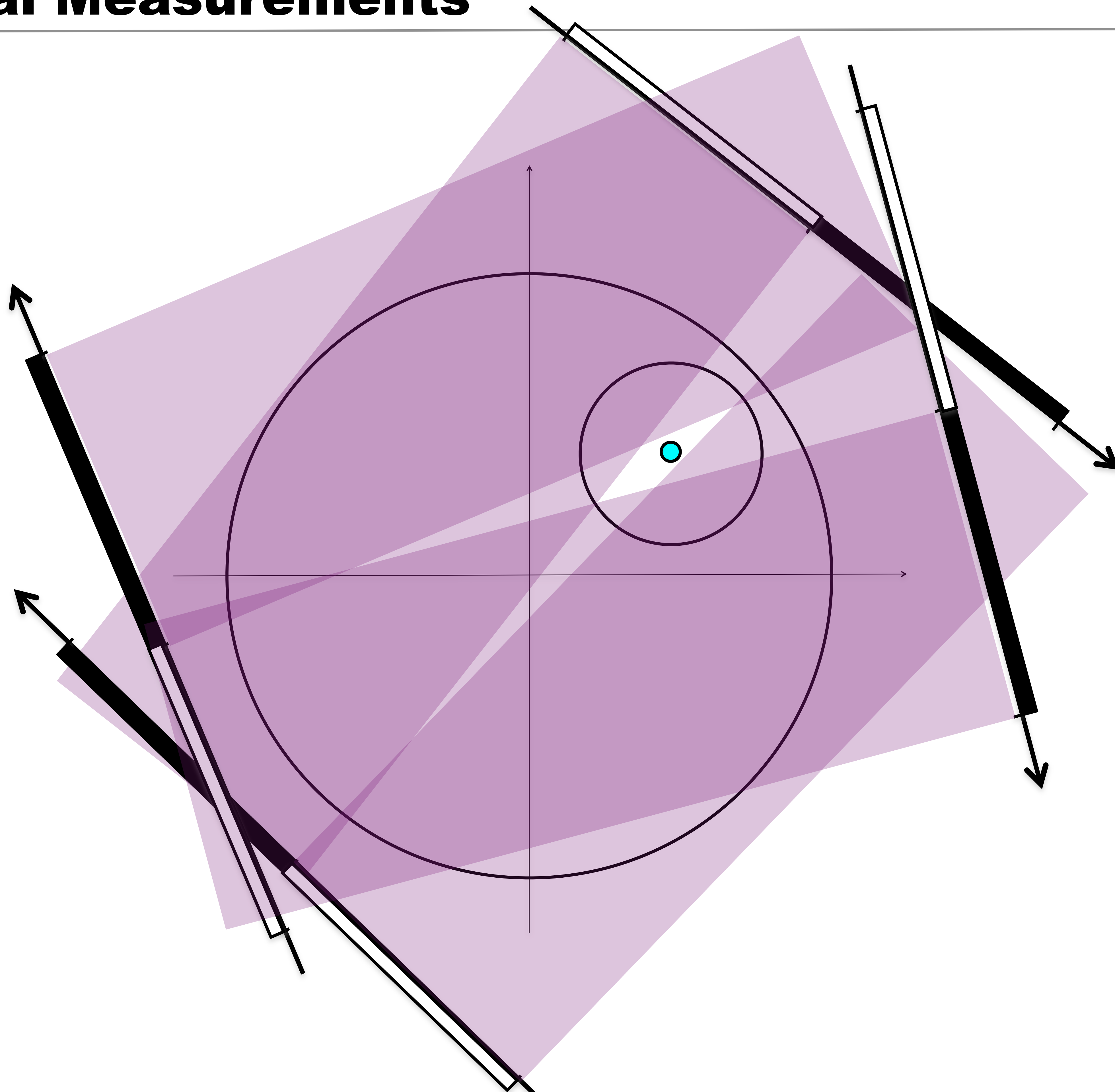
- Reconstruction is a **very hard** problem
  - Seems to have combinatorial complexity
  - Probably NP

- Need to **enable efficient reconstruction**
  - Classical methods **exploit bit hierarchy** to make problem convex
  - Should **maintain theoretical guarantees**

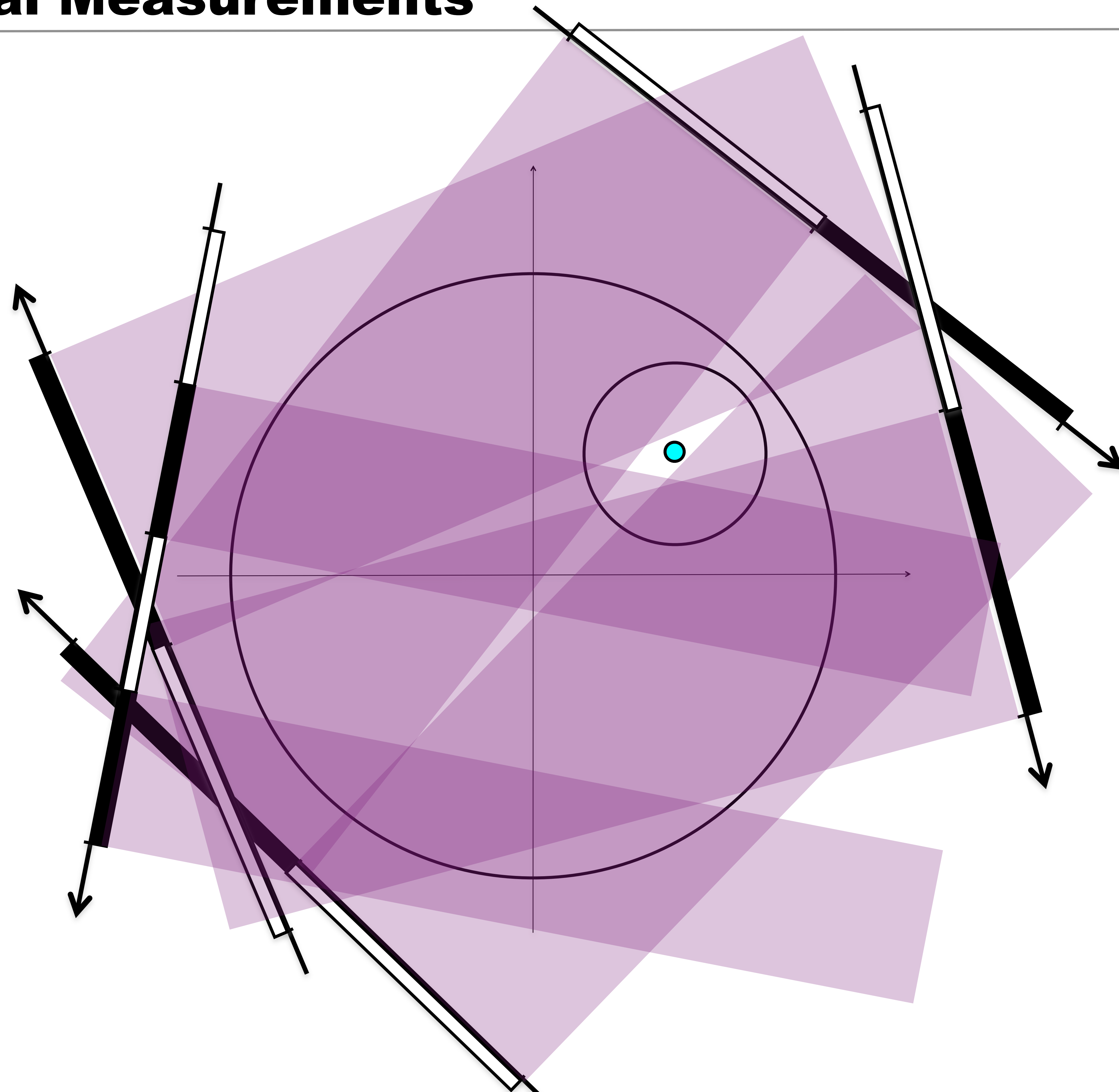
- **Solution**: Construct **bit hierarchy**; **sub-problems** become **convex**

- Boufounos, P.T., "Hierarchical Distributed Scalar Quantization", *Proc. International Conference on Sampling Theory and Applications (SampTA)*, Singapore, May 2011

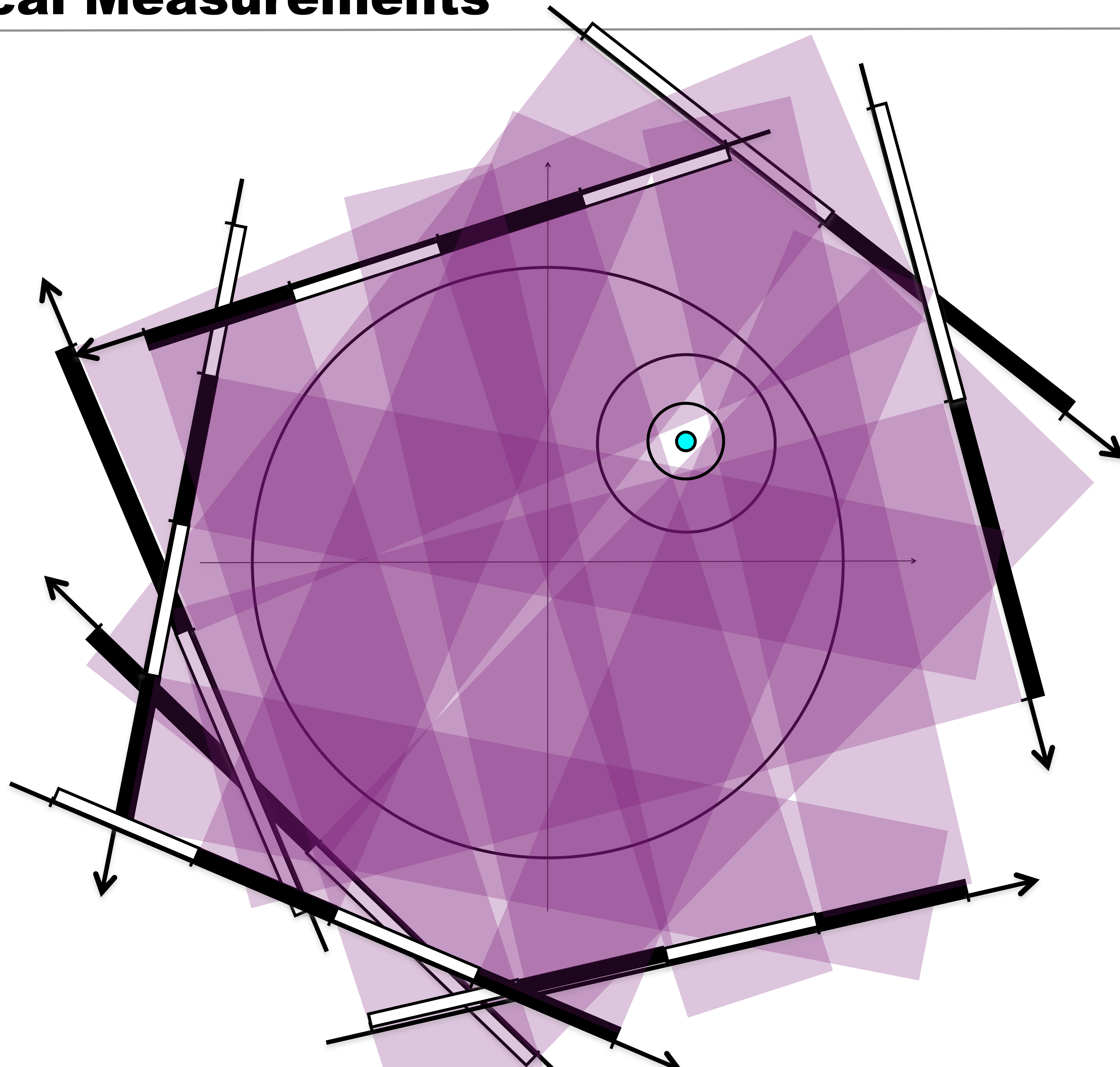
Hierarchical Measurements



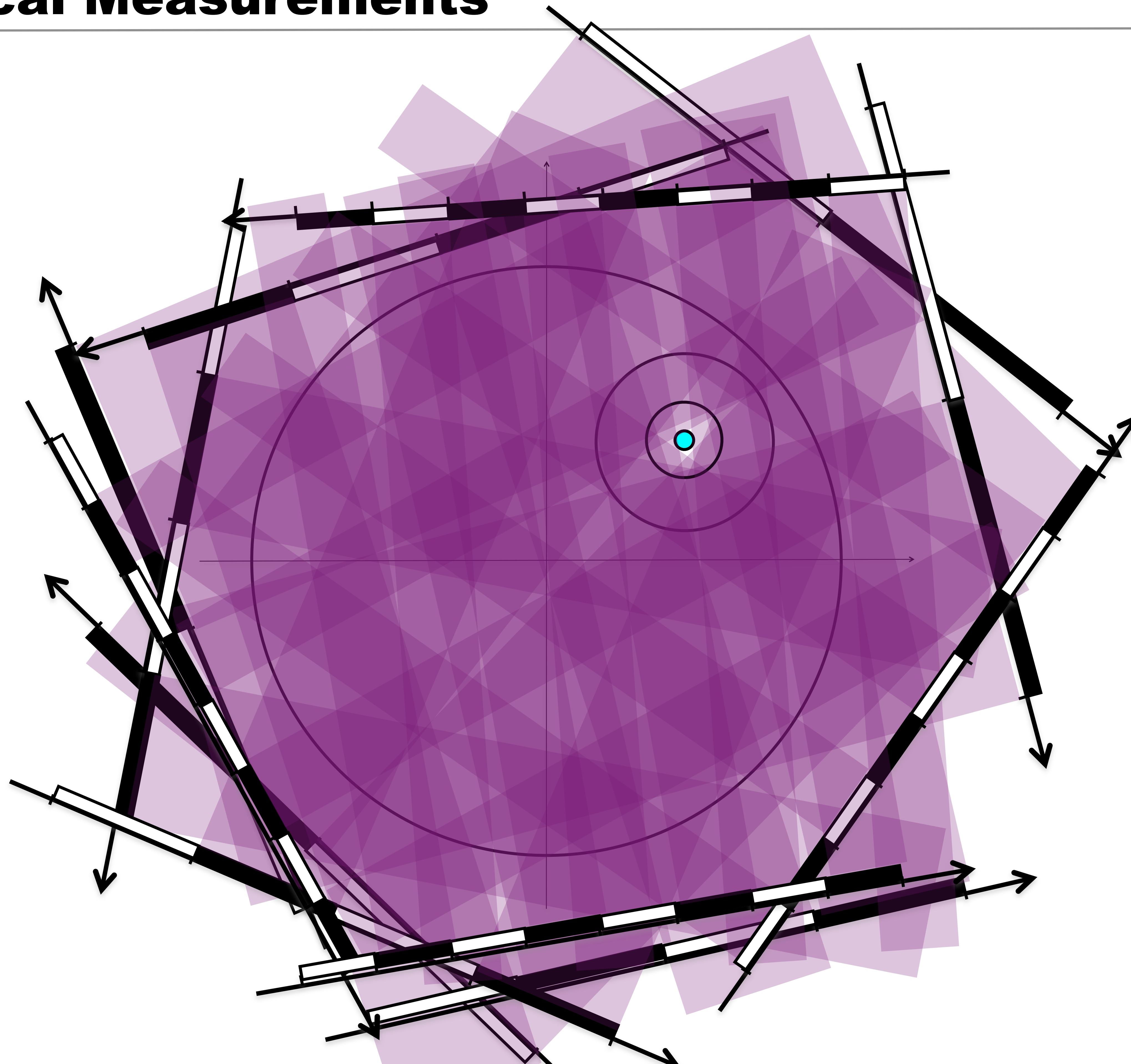
Hierarchical Measurements



Hierarchical Measurements



Hierarchical Measurements



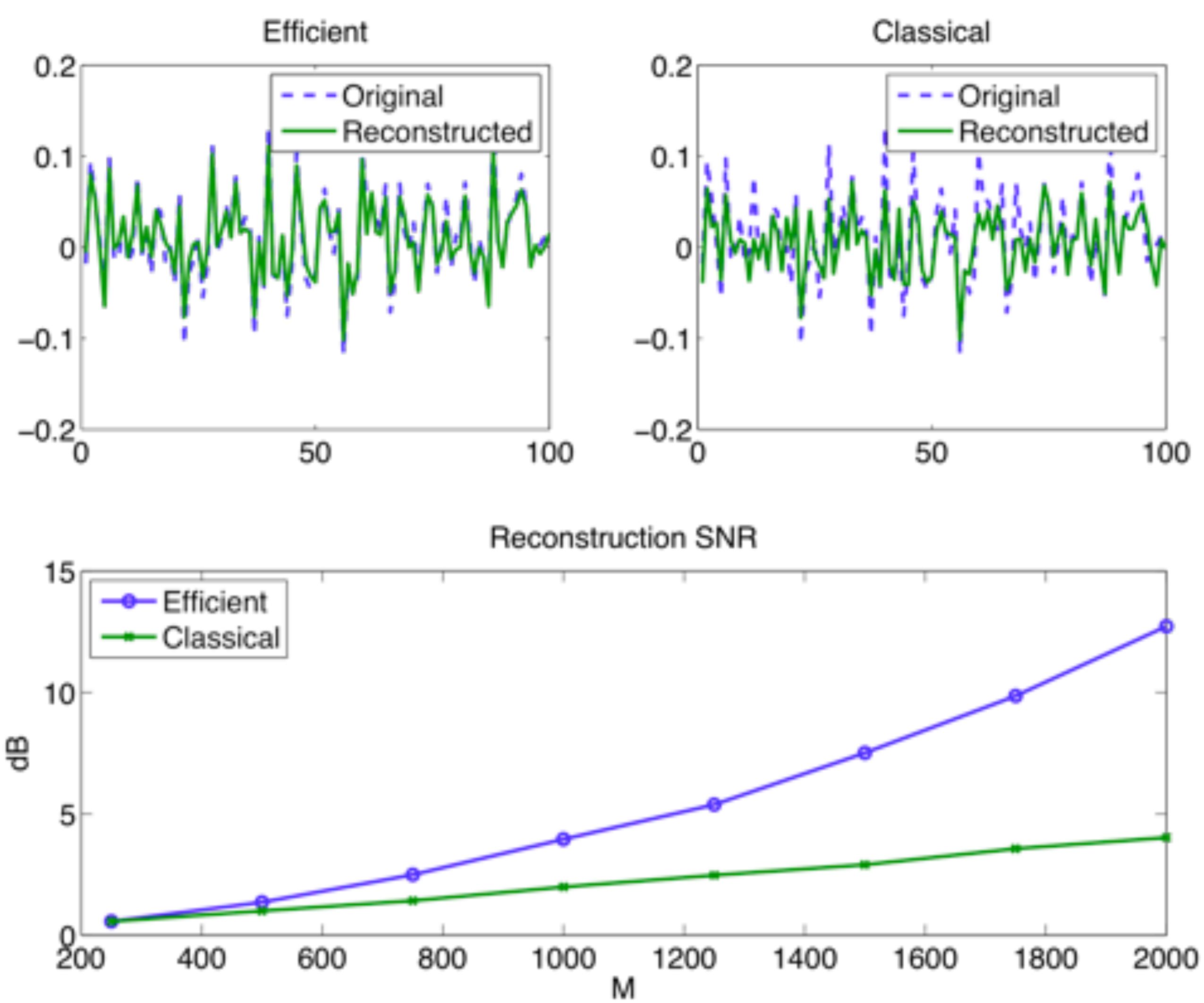
Reconstruction

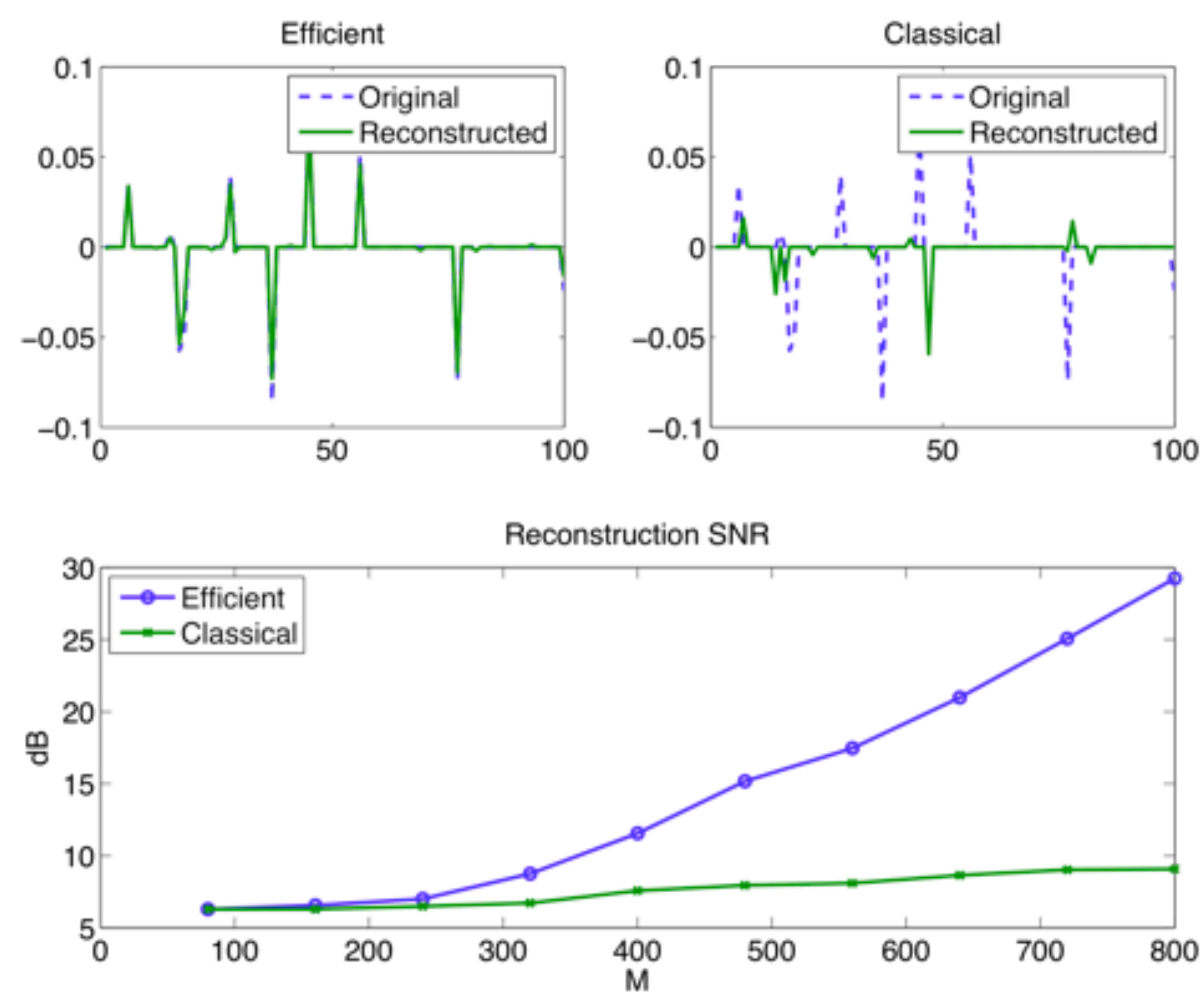
Sampling

- Given uncertainty pick Δ such that reconstruction is convex
- Take enough measurements to scale uncertainty by $\alpha<1$
- Scale $\Delta \leftarrow \alpha\Delta$ for next set of measurements
- Iterate until desired precision

Reconstruction

- For first set of measurements formulate convex reconstruction
- Solve for consistency
- Use solution to incorporate next set of measurements and determine consistency constraints
- Iterate until all measurement sets are incorporated





Privacy Preserving Properties

Assume we have encoding of two signals x,y, but not $\mathbf{A}$ and $\mathbf{w}$
What does the encoding reveal about their relationship?

$$I(f(x); f(y)|d) \leq 10Me^{-\left(\frac{\pi\sigma d}{\Delta}\right)^2}$$

Mutual information decays very fast with d.

Information theoretic privacy-preserving guarantee:

**When signals are far apart,
encoding reveals nothing about their relationship!**

Very useful for security applications
(e.g., privacy-preserving nearest neighbors,
secure biometric authentication)

- Boufounos P. T. and Rane S., "Secure Binary Embeddings for Privacy Preserving Nearest Neighbors," *Proc. Workshop on Information Forensics and Security (WIFS)*, Foz do Iguaçu, Brazil, November 29 – December 2, 2011.

Further Reading

- Johnson W. and Lindenstrauss J., "Extensions of Lipschitz mappings into a Hilbert space," *Contemporary Mathematics*, vol. 26, pp. 189 –206, 1984.
- Andoni, A. and Indyk, P., "Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions," *Comm. ACM*, vol. 51, no. 1, pp. 117–122, 2008.
- Datar M., Immorlica N., Indyk P., and Mirrokni V., "Locality-Sensitive Hashing Scheme Based on p-Stable D istributions," *Proc. Symposium on Computational Geometry*, 2004
- Jacques L., Laska J. N., Boufounos P. T., Baraniuk R. J., "Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors," *IEEE Trans. Info. Theory*, v. 59, no. 4, April, 2013.
- Plan, Y. and Vershynin, R., "Dimension reduction by random hyperplane tessellations," preprint, arXiv:1111.4452, 2011.
- Ai, A., Lapanowski, A., Plan, Y., Vershynin, R, "One-bit compressed sensing with non-Gaussian measurements", *Linear Algebra and Applications*, to appear.
- Boufounos P. T., "Universal Rate-Efficient Scalar Quantization," *IEEE Trans. Info. Theory*, v. 58, no. 3, pp. 1861-1872, March, 2012.
- Boufounos P. T. and Rane S., "Secure Binary Embeddings for Privacy Preserving Nearest Neighbors," *Proc. Workshop on Information Forensics and Security (WIFS)*, Foz do Iguaçu, Brazil, November 29 – December 2, 2011.
- Li M., Rane S., and Boufounos P. T., "Quantized embeddings of scale-invariant image features for mobile augmented reality," *IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*, Banff, Canada, Sept. 17-19, 2012
- Boufounos P. T. and Rane S., "Efficient Coding of Signal Distances Using Universal Quantized Embeddings," *Proc. Data Compression Conference (DCC)*, Snowbird, UT, March 20-22, 2013.
- Yeo C., Ahammad P., and Ramchandran K., "Coding of image feature descriptors for distributed rate-efficient visual correspondences," *International Journal of Computer Vision*, vol. 94, pp. 267–281, 2011, 10.1007/s11263-011-0427-1.
- Min K., Yang L., Wright J., Wu L., Hua X.-S., and Ma Y., "Compact projection: Simple and efficient near neighbor search with practical memory requirements," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2010.

Final Thoughts/Discussion

- Quantization very important in signal processing
  - Signal acquisition systems
  - Information embedding/transmission
  - Information hiding, security, privacy

- Not your college-level quantization
  - High-dimensional geometrical problem
  - Additive noise model inadequate
  - Tight bounds with better models
  - Consistency is important
  - Saturation can be useful
  - Non-linear concentration of measure occurs

- Quantization is a very active area of research
  - 1-bit CS/Quantized CS
  - Sigma-Delta for compressive and non-compressive systems
  - Geometry of non-linear inverse problem solving
  - Quantized Embeddings
  - Vector Quantization (a whole other tutorial)

Still Open and Interesting (a small sampling…)

- Oversampling and Quantization
  - Beyond consistency: quantization with additive noise

- Quantized CS
  - Interaction between sparsity/sensing/quantization (signal/measurement model)
  - 1-bit CS algorithmic convergence guarantees (e.g. BIHT, RSS, MSP)
  - Consistent QCS theory for any bitdepth (1-bit to high-res)
  - Optimal quantizer design for non-gaussian measurements
  - Rate-distortion performance: CS for compression
  - Sigma-Delta CS for 1-bit quantization
  - Vector Quantization of CS measurements

- Universal Quantization and Embeddings
  - General reconstruction algorithms: is reconstruction possible?
  - Embedding guarantees for more general embeddings (e.g. multi-bit)
  - Embedding behavior design
  - Tighter connections with LSH
  - Other security/privacy-preserving properties

Today's Topics

1. Modern Scalar Quantization
2. Compressive Sensing Overview
3. Compressive Sensing and Quantization
4. 1-bit Compressive Sensing
5. Locality Sensitive Hashing and Universal Quantization

For more:

Repository: http://www.boufounos.com/resources-on-quantization/

http://dsp.rice.edu/1bitCS/
http://nuit-blanche.blogspot.com
http://nuit-blanche.blogspot.com/search/label/1bit
http://nuit-blanche.blogspot.com/search/label/QuantCS
http://www.boufounos.com/research/quantization/

Questions/Comments?

petros@boufounos.com
http://boufounos.com

laurent.jacques@uclouvain.be
http://perso.uclouvain.be/laurent.jacques/