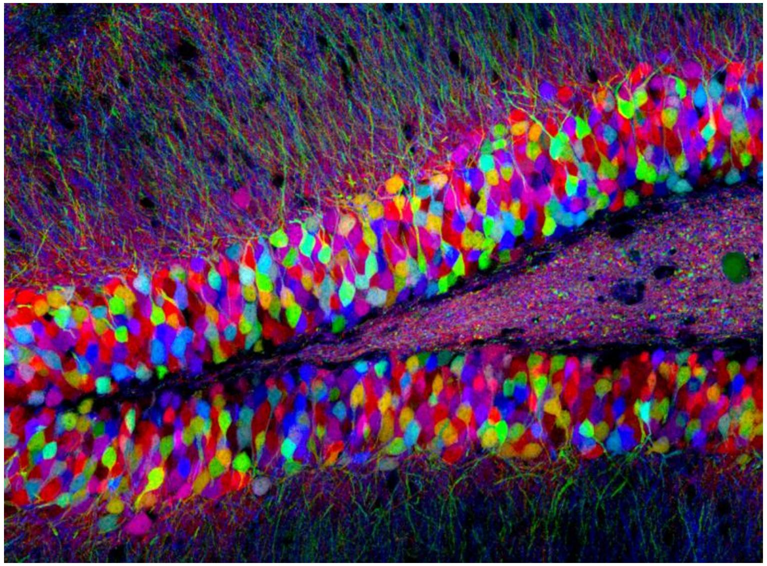
Compressed sensing

The Group

20 July 2015

High dimensional spaces

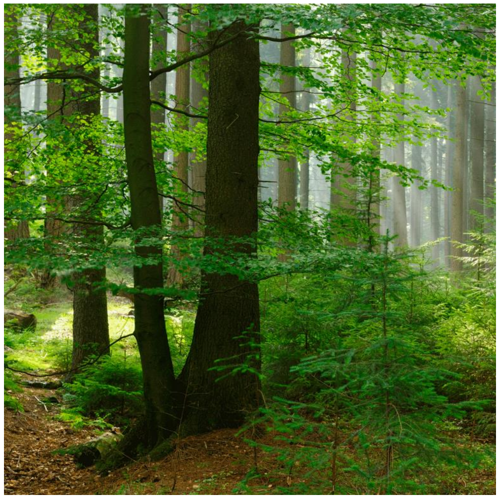
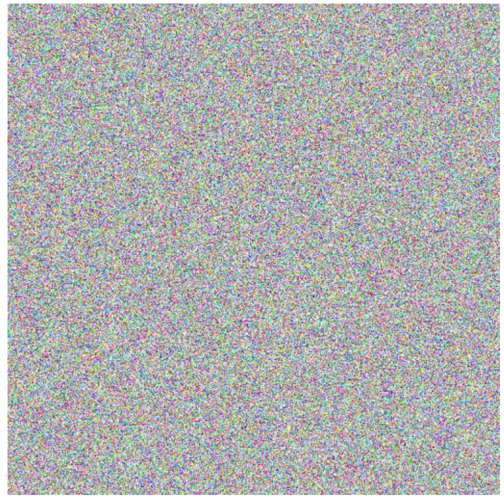
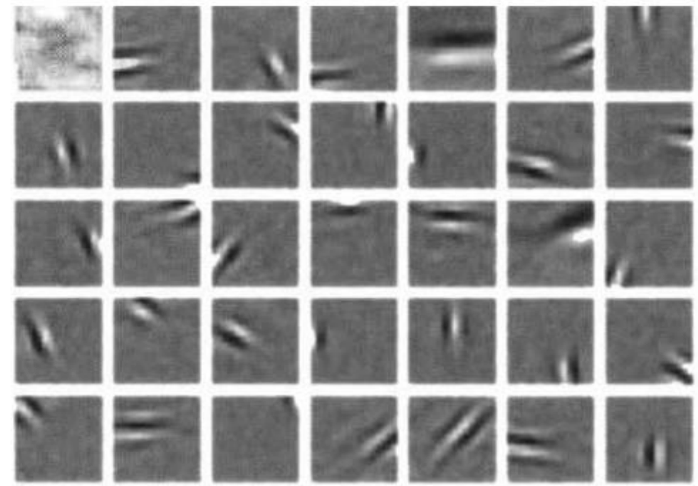


Curse of dimensionality:

- Undersampling (M << N)
- Difficult to do statistics
- Counter-intuitive

For an entertaining illustration, see
Bushdid, Science (2014)
M Meister, eLife (2015)

Sparsity

How to take advantage of sparsity?

Want to:

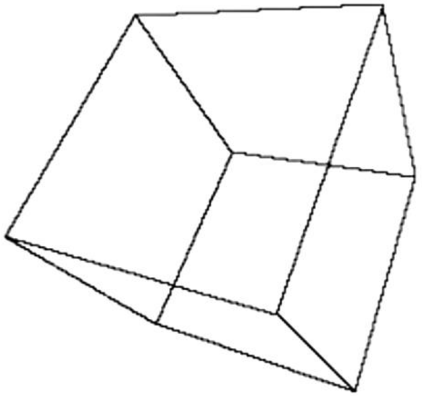
- Extract relevant dimensions
- Reduce measurement redundancy

Random projections

N dimensions
K-sparse signal

Project along
M random directions

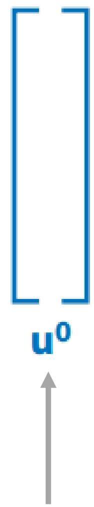
M dimensions



$$M > O(K \log(N/K))$$

random projections guarantee
perfect signal reconstruction

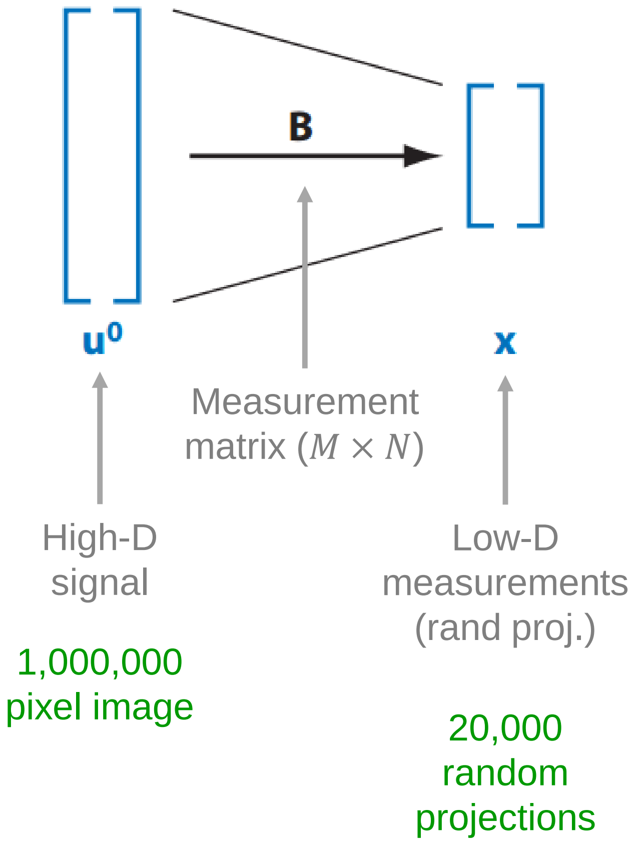
CS framework

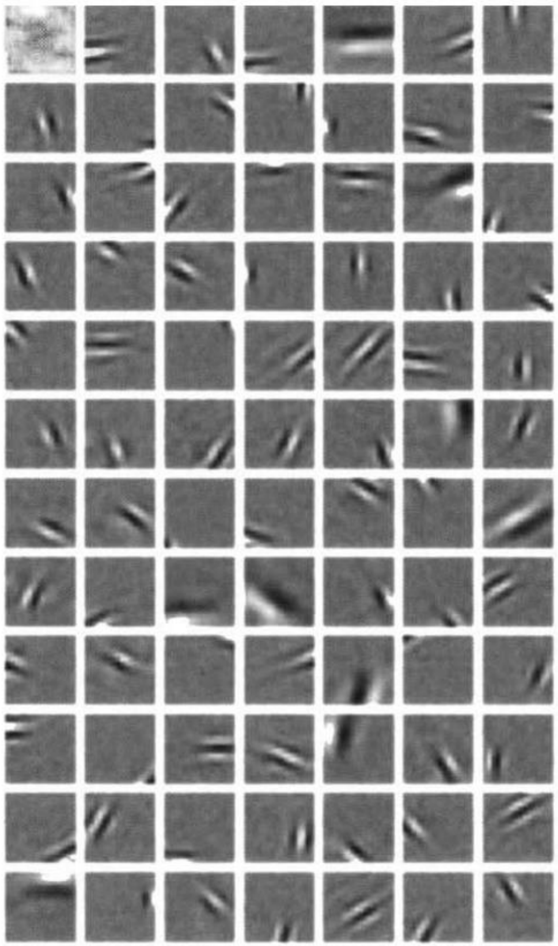


Take measurements:

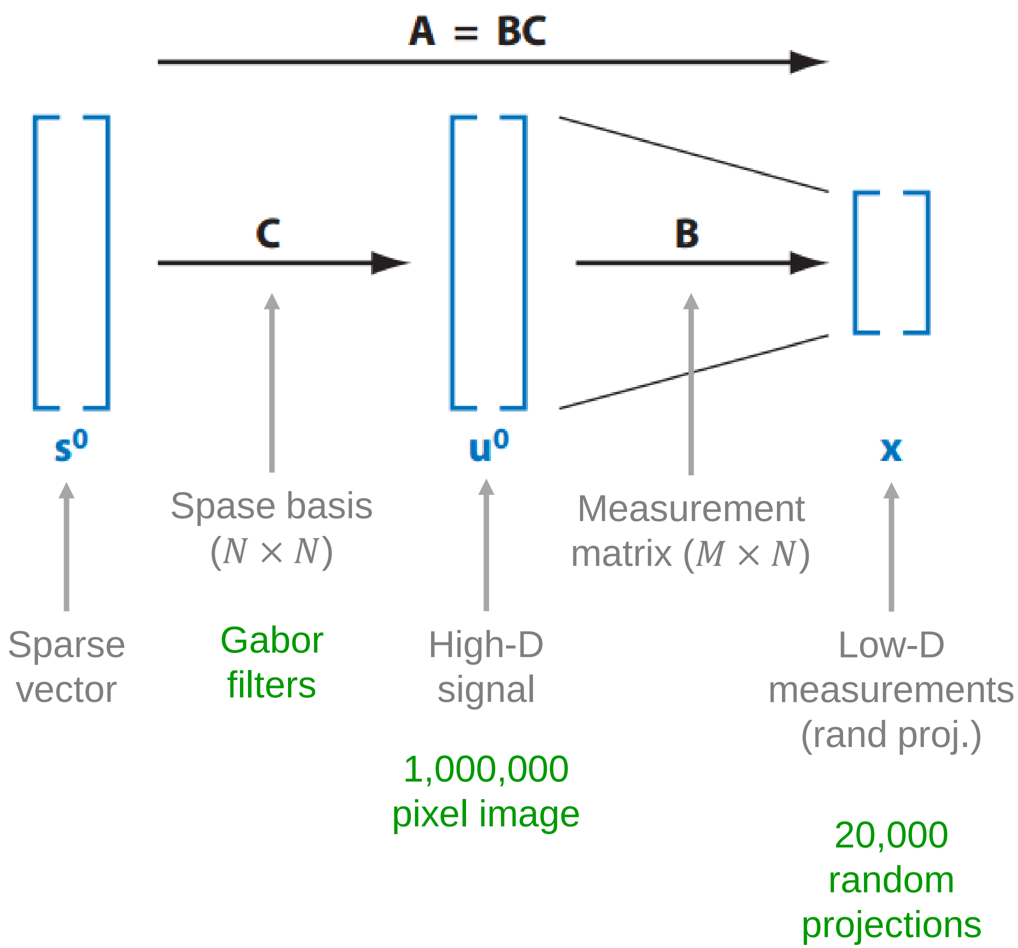
$$x_\mu = \mathbf{b}^\mu \cdot \mathbf{u}^0$$
$$\mu = 1, \ldots, M$$

CS framework



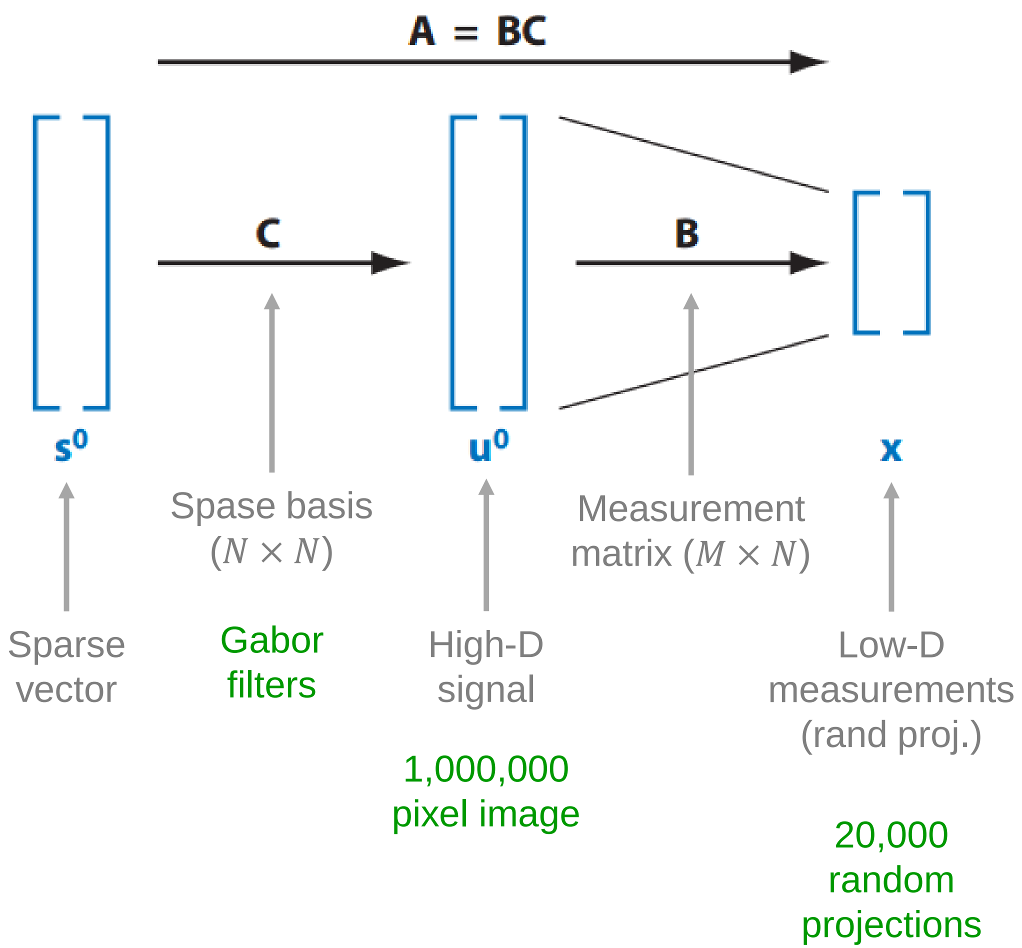


CS framework



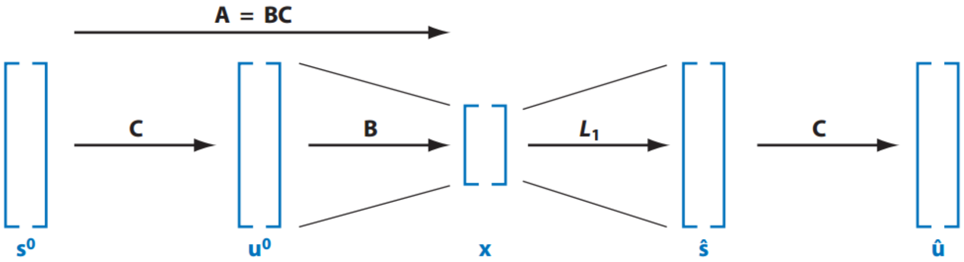
B must be incoherent with respect to C (random will do)

CS framework

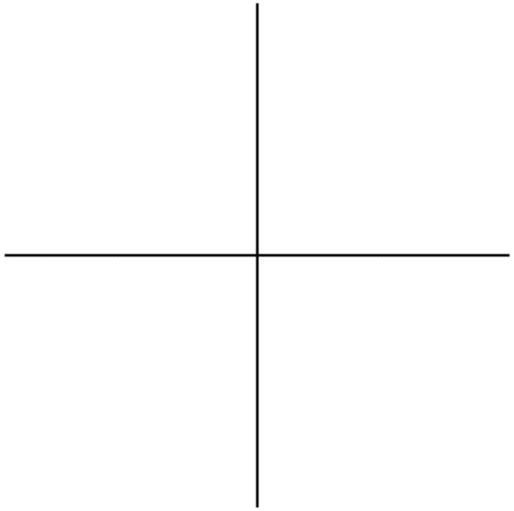


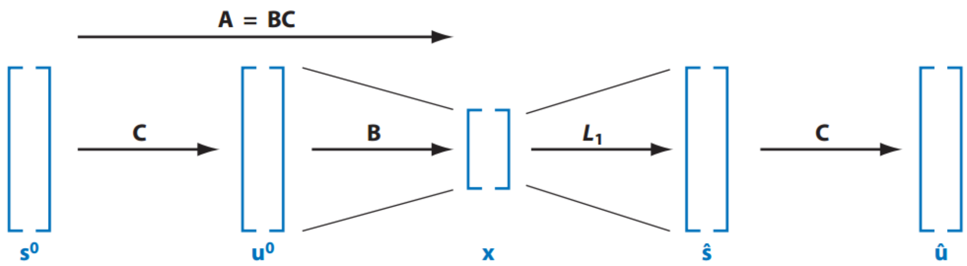
Given the short measurement vector $\mathbf{x}$
how can we reconstruct $\mathbf{s}^0$?

$\mathbf{x} = \mathbf{A}\mathbf{s}^0$
(N unknowns, M equations, N>>M)

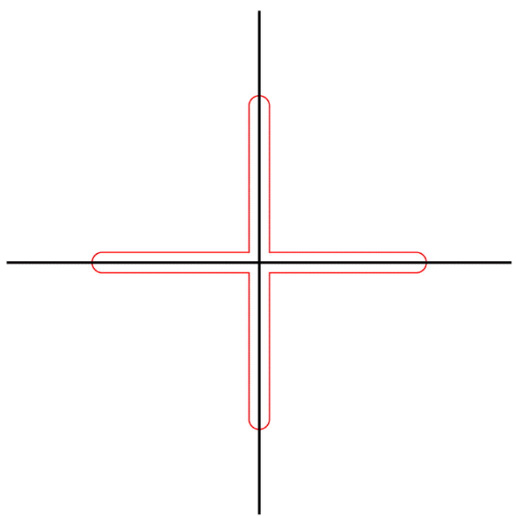


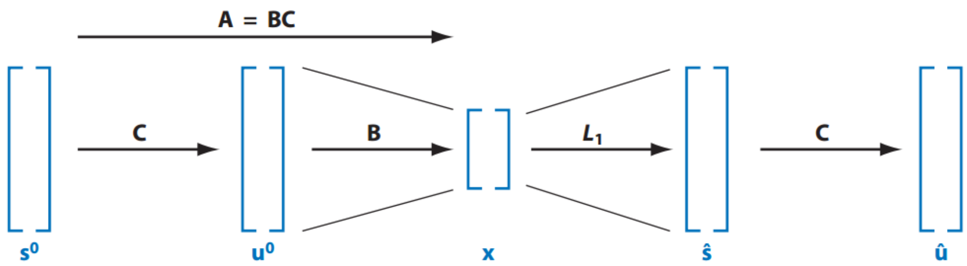
Basis space **A**
N = 2 dimensions



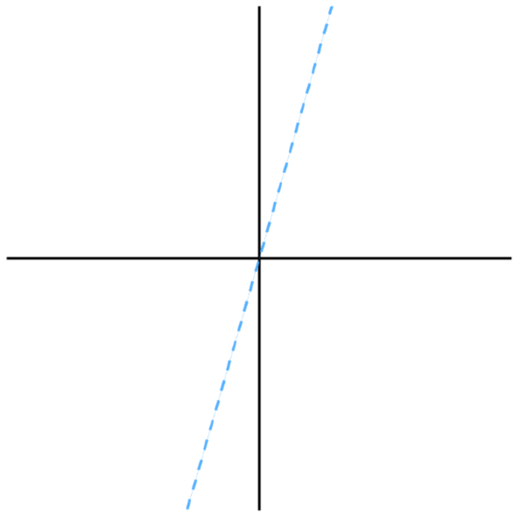


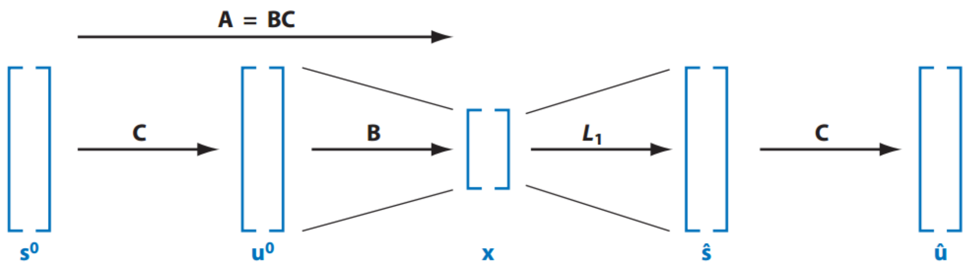
Sparse signal $\mathbf{s^0}$
lives on axes



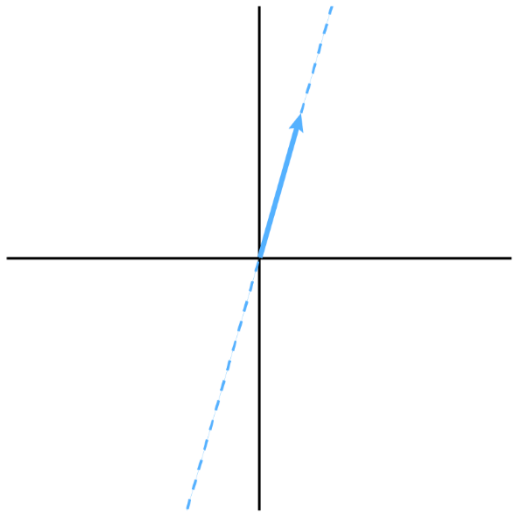


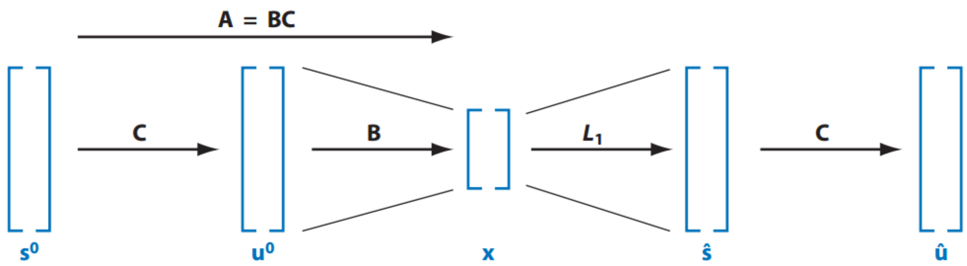
One random measurement
M = 1 dimension



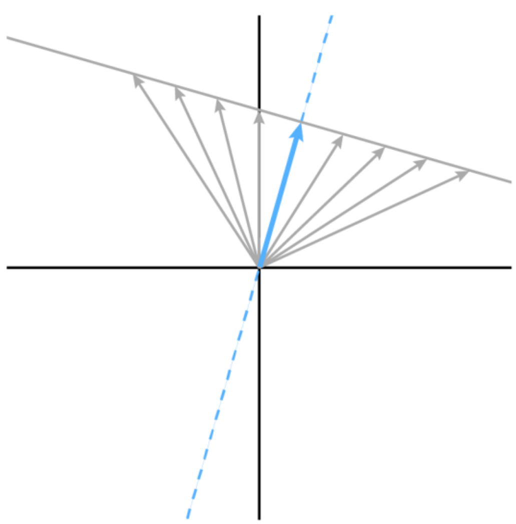


Measurement **x**:
Projection of 2-D signal
onto 1-D measurement
vector (M × N matrix **B**)





$\mathbf{x} = \mathbf{A}\mathbf{s}^0$
is under-constrained
(N=2 unknowns, M=1 equations)

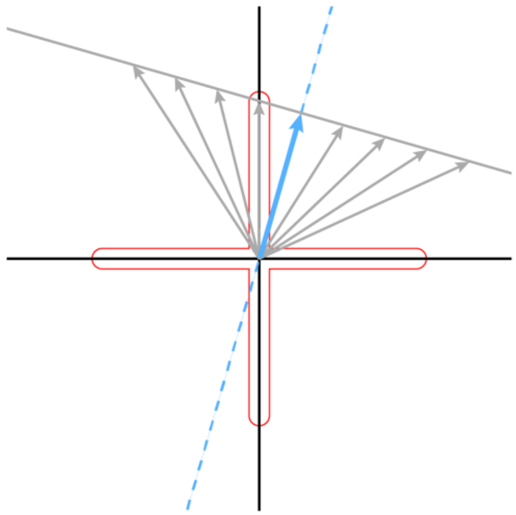


→ Need to solve optimization problem:

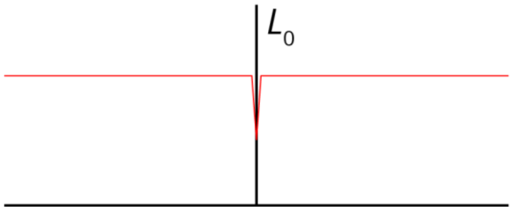
$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \sum_{i=1}^{N} V(s_i) \quad \text{subject to } \mathbf{x} = \mathbf{A}\mathbf{s}$$

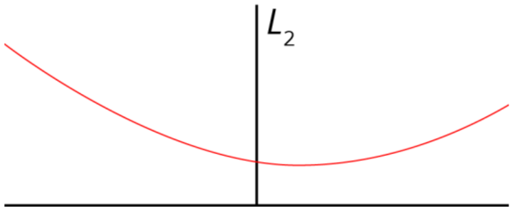
where V(s) is a penalty for non-sparsity.

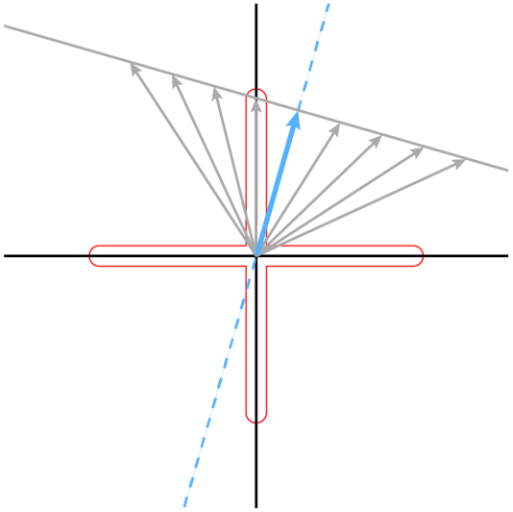
L_p-norm: $\|x\|_p = (|x_1|^p + |x_2|^p + \ \ldots + |x_n|^p)^{1/p}$

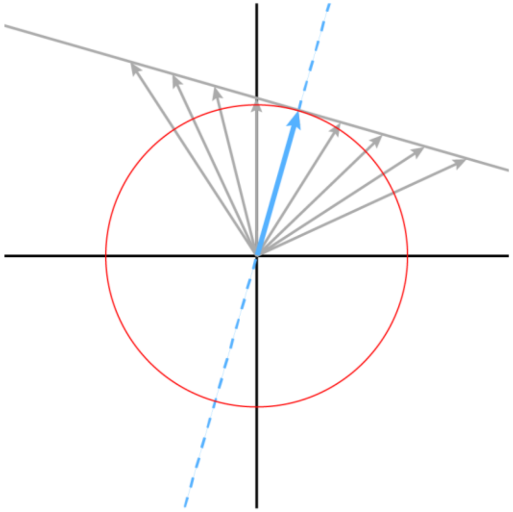


L_p-norm: $\|x\|_p = (|x_1|^p + |x_2|^p + \ \ldots + |x_n|^p)^{1/p}$









Matlab: `s = A\x`

Approximate sparsity and noise

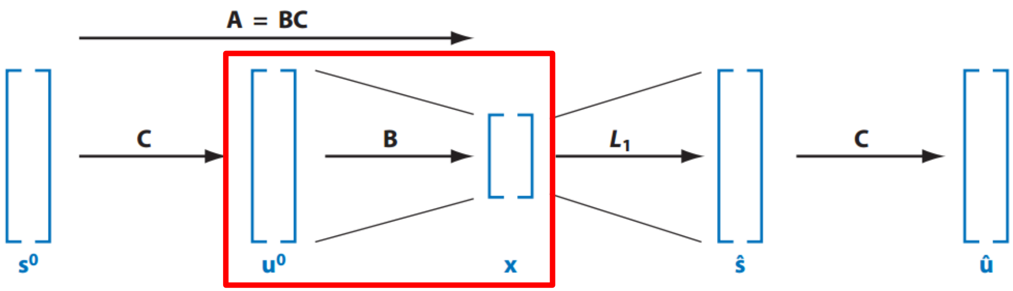
Exact: $$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \sum_{i=1}^{N} |s_i| \quad \text{subject to } \mathbf{x} = \mathbf{A}\mathbf{s}$$

With noise: $$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \left\{ \|\mathbf{x} - \mathbf{A}\mathbf{s}\|^2 + \lambda \sum_{i=1}^{N} |s_i| \right\}$$

Least absolute shrinkage and selection operator

Matlab: `s = lasso(A, x, 'lambda', lambda)`

How many measurements?



Johnson-Lindenstrauss lemma:

Given $0 < \varepsilon < 1$ and a set **s** of K points in R^N, there is a linear map $f : R^N \rightarrow R^M$ such that

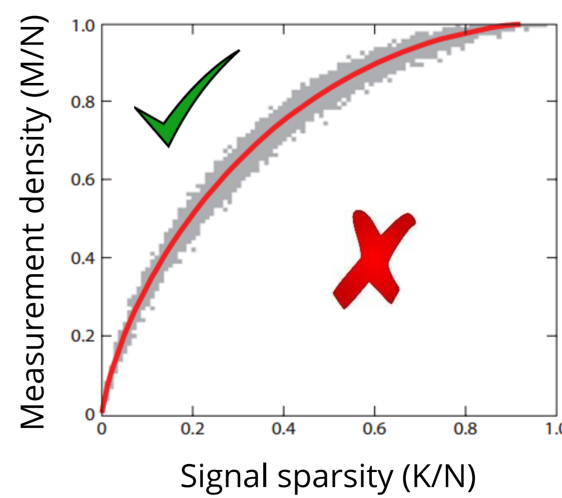
$$(1-\epsilon)\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\epsilon)\|u-v\|^2$$

for all u, v ∈ **s**, **if $M > 8\ln(K)/\varepsilon^2$**.

→ M > O(K log(N/K))

random projections ~guarantee perfect signal reconstruction

Phase transition to perfect reconstruction



How is this useful?

- Reconstruction of inaccessible high-D signal (break Nyquist bound)

- Statistics easier in dense low-D that sparse high-D
  - Regression
  - Classification
  - Clustering/Nearest-neighbor-finding
  - (Everything that relies on point distances)