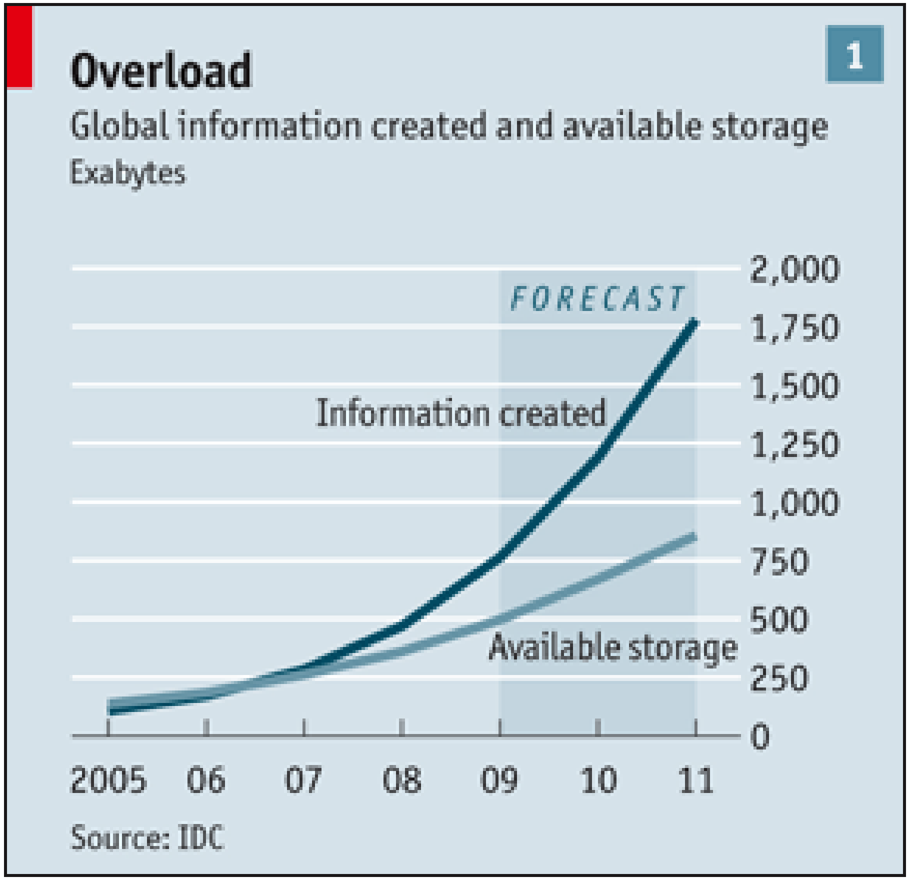
Corruption, Justice, and Democracy in Compressive Sensing

Mark Davenport

Rice University
ECE Department

Data Deluge



In 2007 digital data *generated > total storage*
by 2011, ½ of digital universe will have no home

[The Economist – March 2010]

Data Deluge

How can we extract as much information as possible from a limited amount of data?

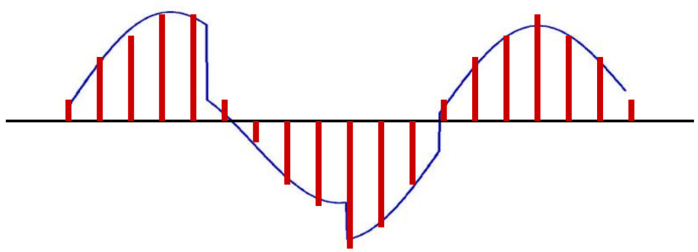
↓

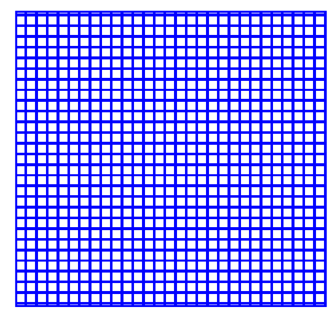
How can we extract any information at all from a massive amount of high-dimensional data?

Digital Revolution

- Foundation: ***Shannon sampling theorem***

Must sample at twice the highest frequency of the signal (Nyquist rate)



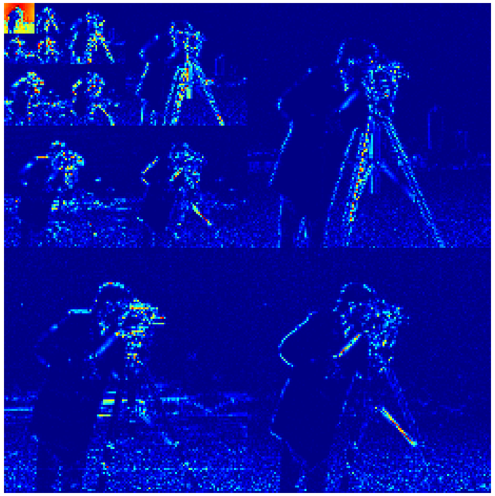


- High-frequency content = *lots of samples…*
- We typically try to compress the data
- Compression relies on *low-dimensional models*

Sparsity

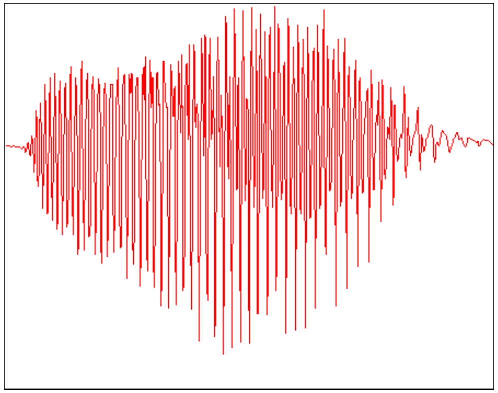
Many signals can be ***compressed*** in some representation/basis (Fourier, wavelets, …)

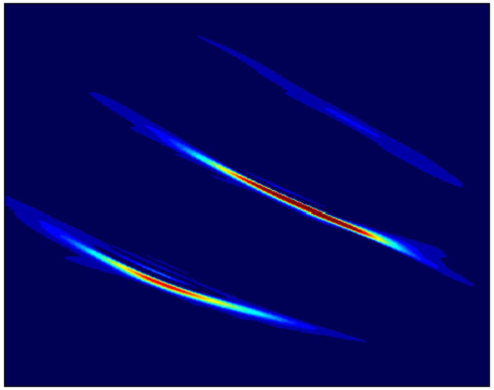
N
pixels



$K \ll N$
large wavelet coefficients

N
wideband signal samples

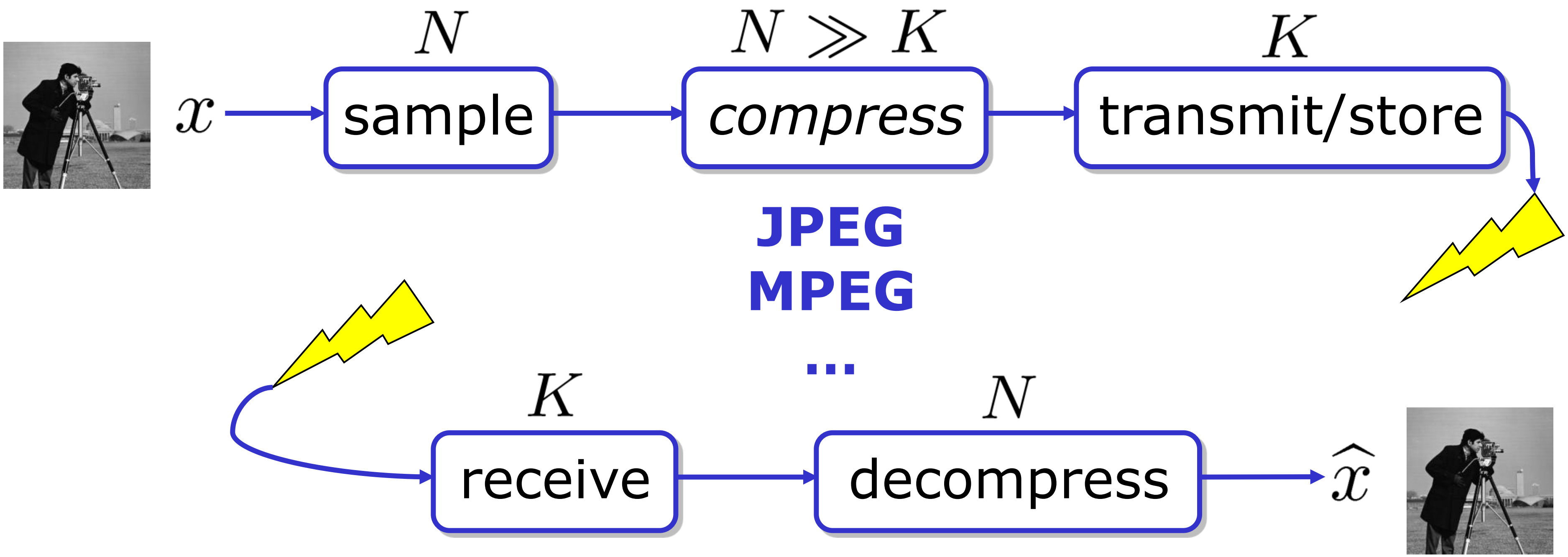




$K \ll N$
large Gabor coefficients

Sample-Then-Compress Paradigm

- Standard paradigm for digital data acquisition
  - ***sample*** data (ADC, digital camera, …)
  - ***compress*** data (signal-dependent, nonlinear)

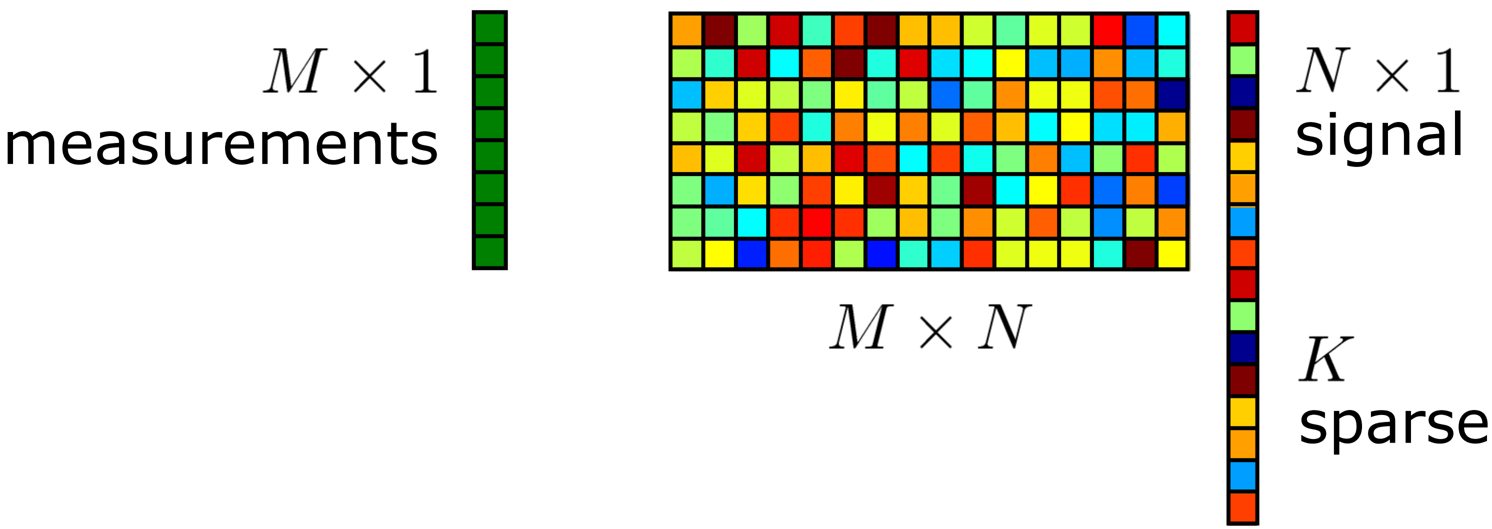


- Sample-and-compress paradigm is ***wasteful***
  - samples cost $$$ and/or time

Compressive Sensing

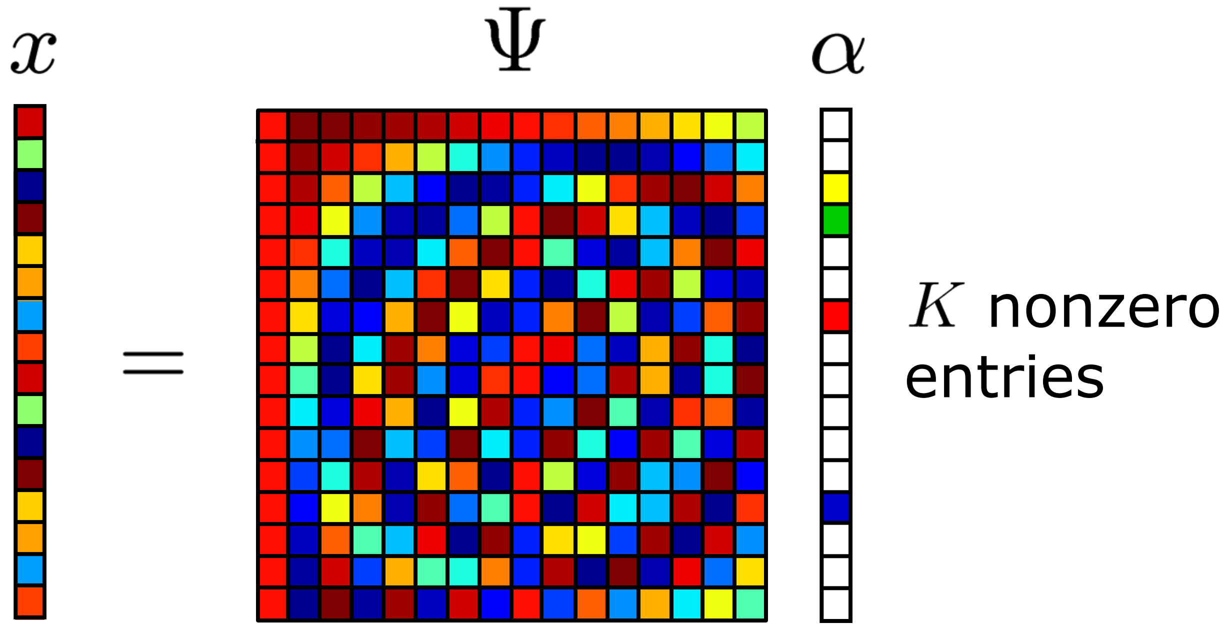
Replace samples with *linear measurements*

$$y = \Phi x$$



$$K < M \ll N$$

Sparsity

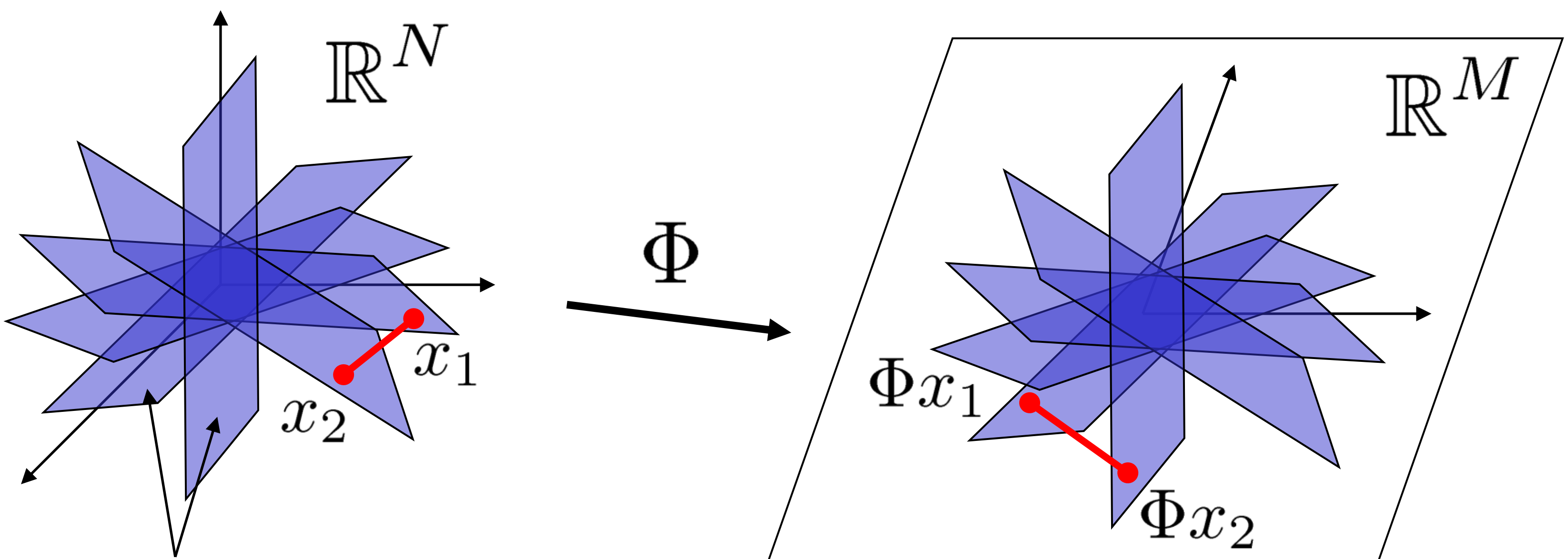


For now: Assume $\Psi = I$

Restricted Isometry Property (RIP)

- Preserve the structure of sparse signals
- For all K-sparse x_1 and x_2

$$(1-\delta_{2K}) \leq \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \leq (1+\delta_{2K})$$



K-dimensional subspaces

Matrices Satisfying the RIP

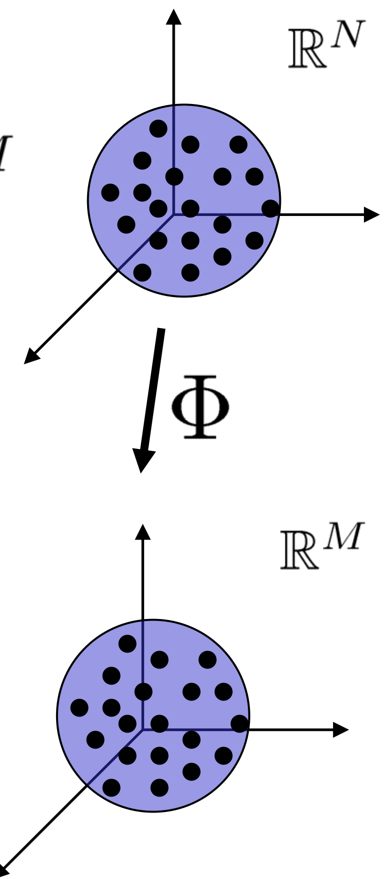
- Pick Φ at *random* using a *sub-Gaussian* distribution

$$\mathbb{E}\left(e^{Xt}\right) \leq e^{c^2 t^2/2}$$

- For any fixed x

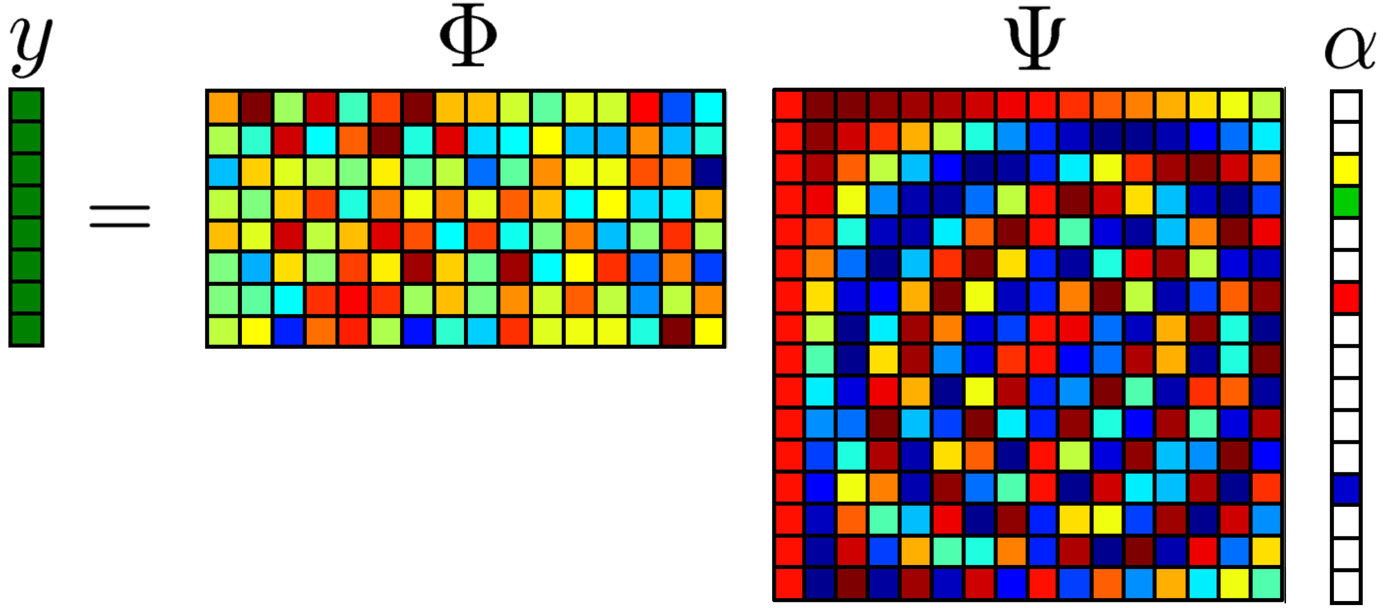
$$\mathbb{P}\left(\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| \geq \epsilon \|x\|_2^2\right) \leq e^{-\widetilde{c}M}$$

- If $M \geq CK \log(N/K)$, then with high probability, Φ will satisfy the RIP
  - fix a $2K$-dimensional subspace
  - pick a finite sampling of points on the sphere
  - repeat for all $\binom{N}{2K}$ subspaces
  - argue that Φ preserves the norm of each point
  - extend from point set to entire sphere

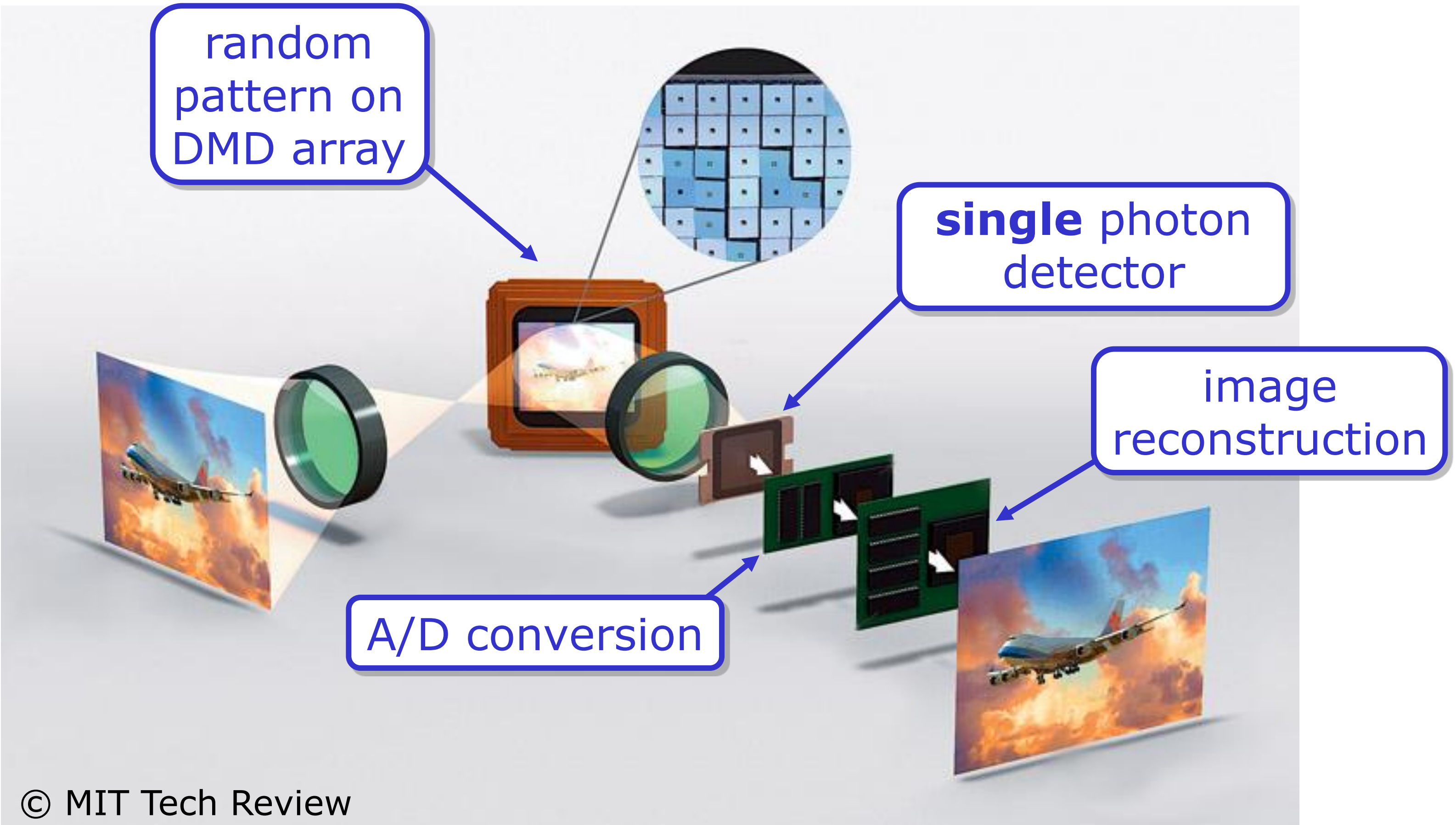


Universality

Random matrix will work with ***any*** fixed orthonormal basis (with high probability)



"Single-Pixel" CS Camera

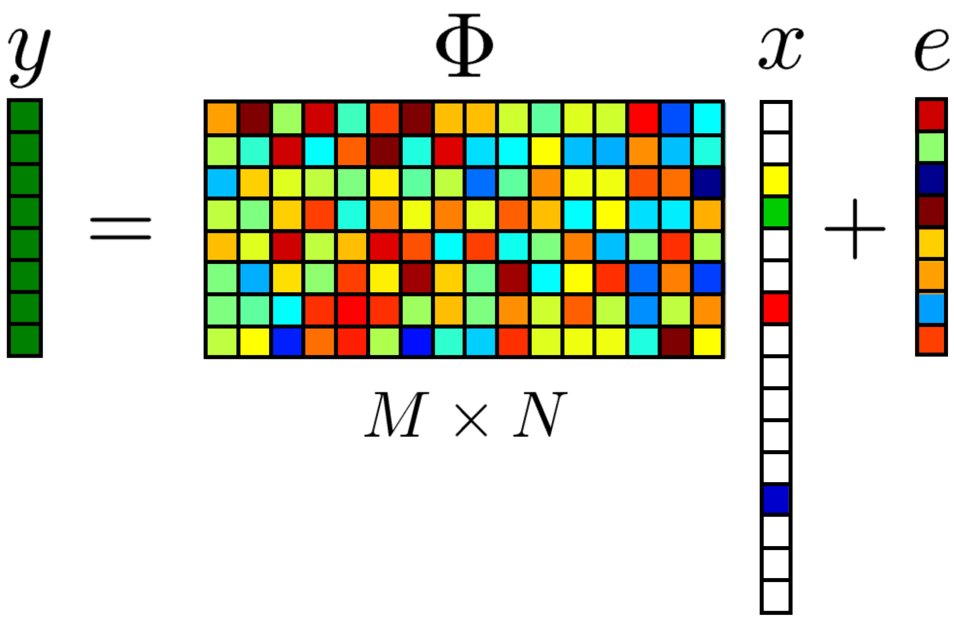


Signal Recovery

given $y = \Phi x + e$

find x

ill-posed inverse problem



Signal Recovery in Noise

- Optimization-based methods

$$\widehat{x} = \arg\min_{x \in \mathbb{R}^N} \|x\|_1 \quad \text{s.t.} \quad \|y - \Phi x\|_2 \leq \epsilon$$

- Greedy/Iterative algorithms
  - OMP, StOMP, ROMP, CoSaMP, Thresh, SP, IHT

$$\|\widehat{x} - x\|_2 \leq C_0 \|e\|_2 + C_1 \frac{\|x - x_K\|_1}{\sqrt{K}}$$

Corruption

$$y = \Phi x + e$$

- What if e represents corruption or *structured noise*, rather than an arbitrary perturbation?

- Structured signal noise:

$$y = \Phi x_S + \Phi x_I$$

- Structured measurement noise:

$$y = \Phi x + \Omega e$$

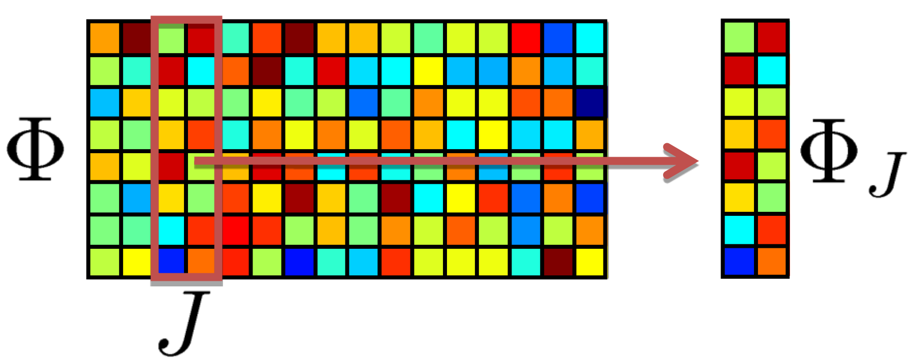
Interference Cancellation

Suppose $x = x_S + x_I$ where x_S is sparse with *unknown* support and x_I is sparse with *known* support J

Goal: Design an $M \times M$ matrix P such that

$$\|P(\Phi x_I)\|_2 \approx 0$$

$$\|P(\Phi x_S)\|_2 \approx \|\Phi x_S\|_2$$



$$P = I - \underbrace{\Phi_J \Phi_J^\dagger}$$

Projection onto $\mathcal{R}(\Phi_J)$

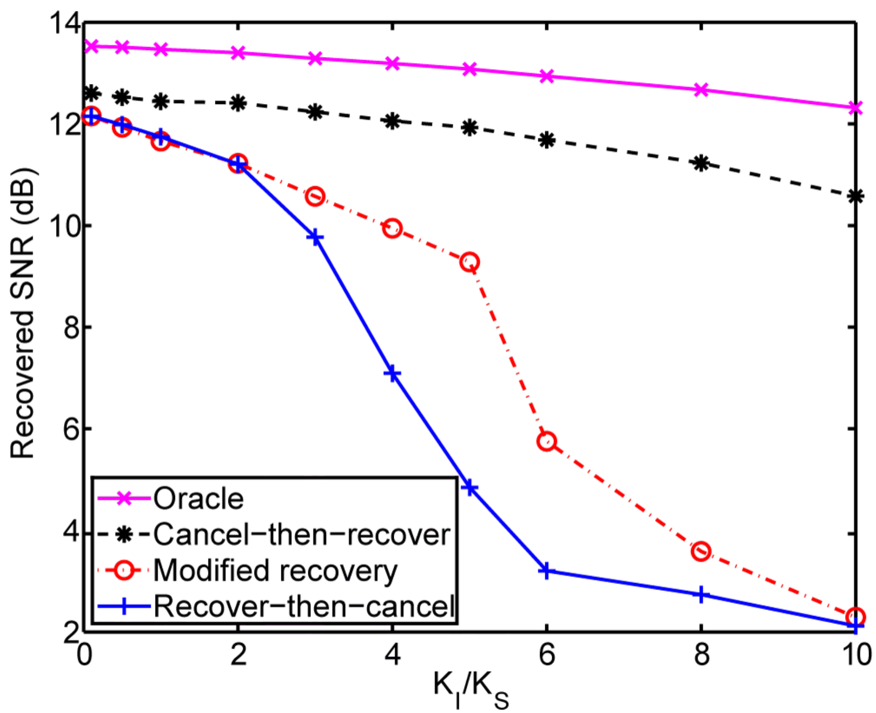
$$P\Phi_J = 0$$

Interference Cancellation

If Φ satisfies the RIP of order $2K_S + K_I$, then $P\Phi$ satisfies

$$\left(1 - \frac{\delta}{1-\delta}\right)\|x\|_2^2 \le \|P\Phi x\|_2^2 \le (1+\delta)\|x\|_2^2$$

for all x such that $\|x\|_0 \le 2K_S$ and $\mathrm{supp}(x) \cap J = \emptyset$.

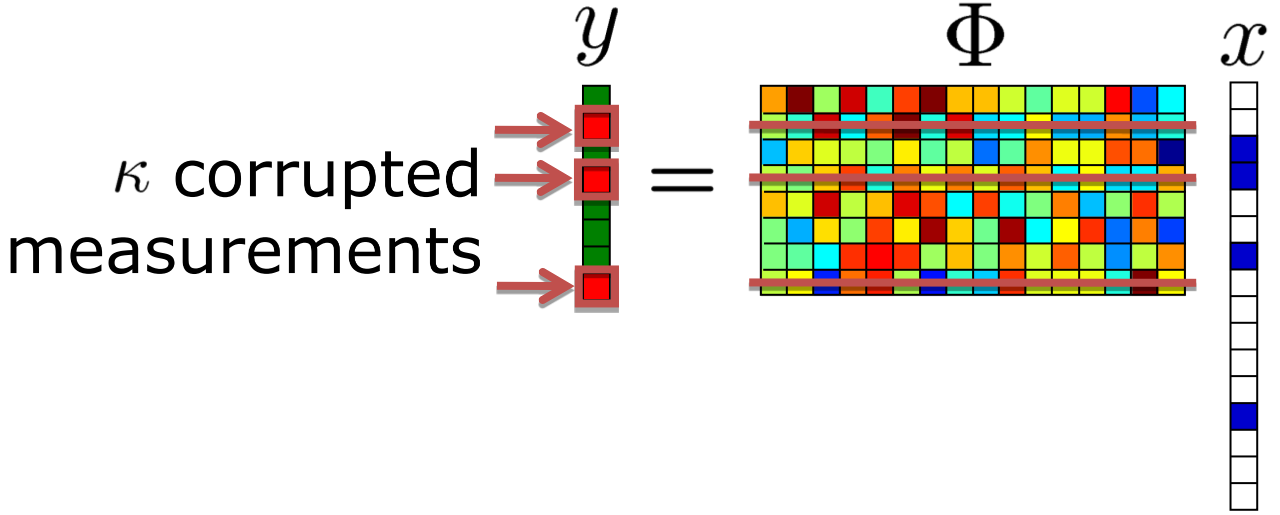


Structured Measurement Noise

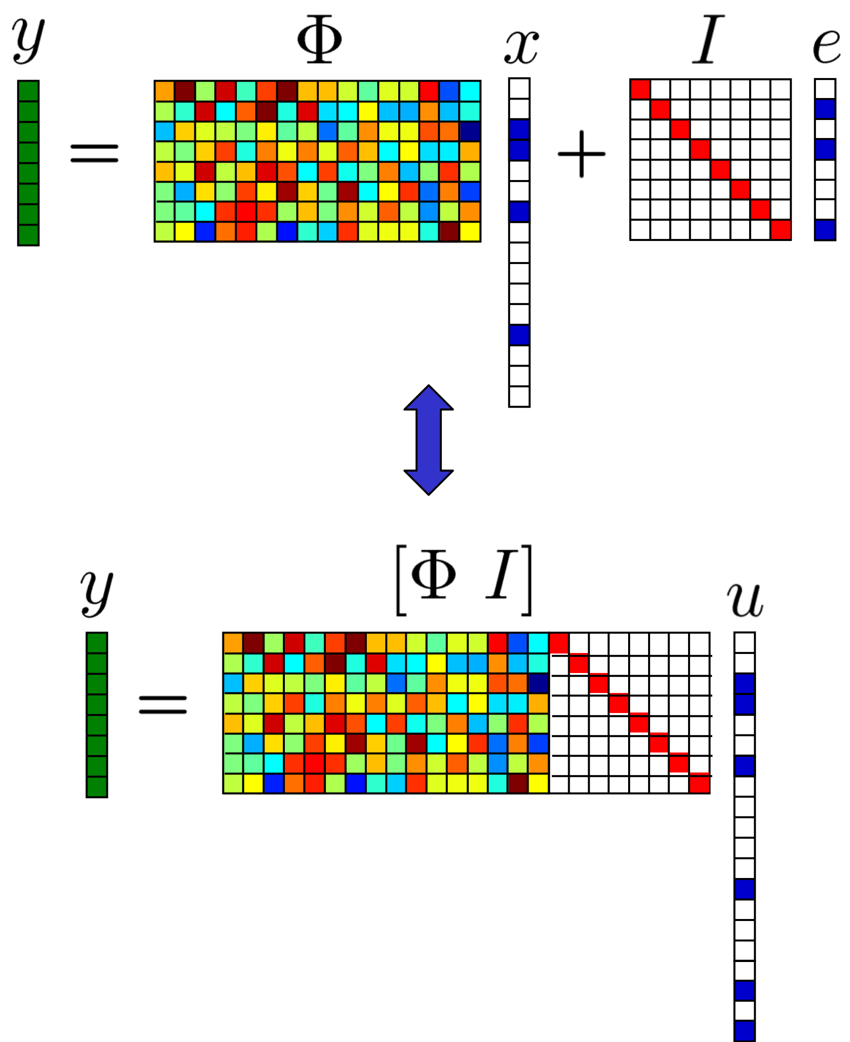
- We have already seen that we can be robust to certain kinds of structured signal noise

$$y = \Phi x_S + \Phi x_I$$

- What about structured measurement noise?



Sparse Noise Model



Justice Pursuit

$$\widehat{u} = \arg\min_u \|u\|_1$$
$$\text{s.t.} \quad y = [\Phi \ I]\, u$$

Does this matrix satisfy the RIP?

Theorem: If Φ is a sub-Gaussian matrix with

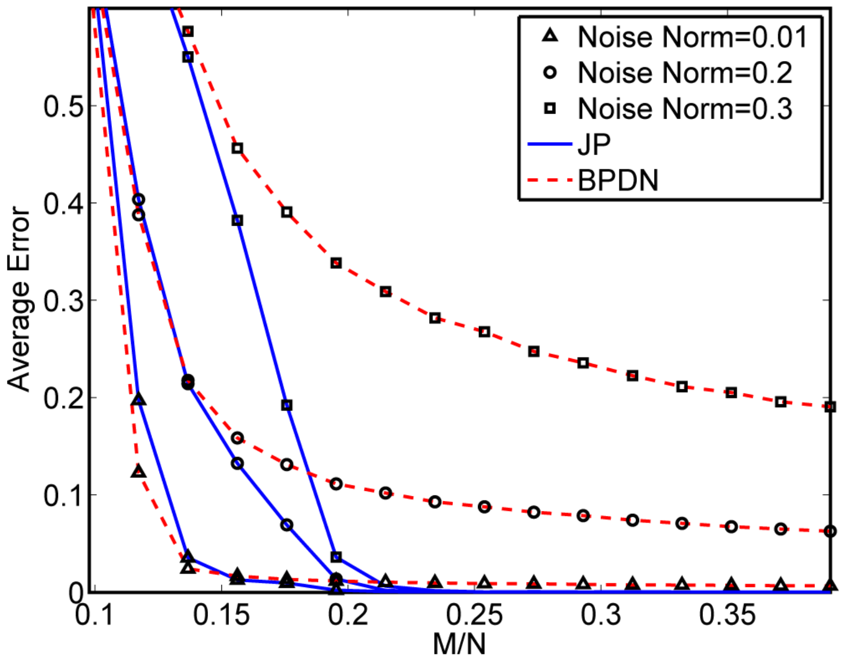
$$M = O\left((K+\kappa)\log\left(\frac{N+M}{K+\kappa}\right)\right)$$

then $[\Phi \ I]$ satisfies the RIP of order $(K+\kappa)$ with probability at least $1 - 3e^{-CM}$.

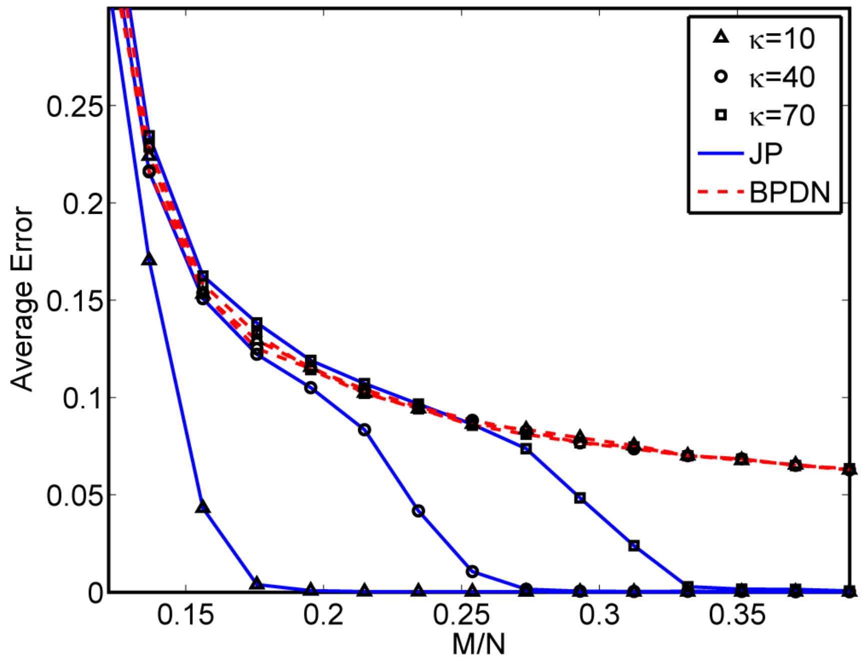
Justice Pursuit

We can recover sparse signals *exactly* in the presence of *unbounded* sparse noise

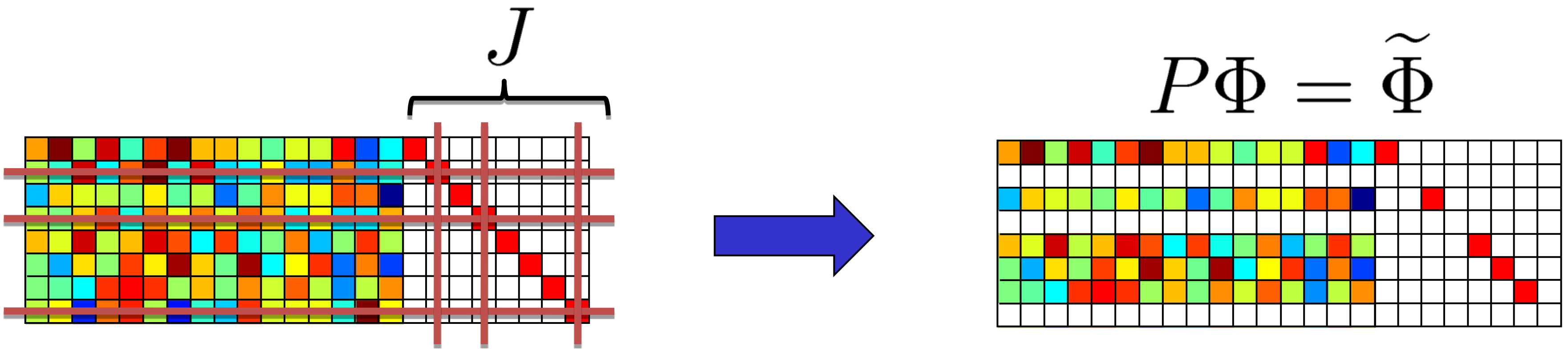
Fixed $\kappa = 10$



Fixed $\|e\|_2 = 0.1$



Justice and Democracy



- The fact that $[\Phi \;\; I]$ satisfies the RIP also implies that we can delete arbitrary rows of Φ and retain the RIP

- Random matrices satisfy a *very strong* ***adversarial*** form of democracy

Conclusions

- Corruption and Justice
  - If the *signal noise* is sparse with known support, it can be cancelled prior to recovery
  - If the *measurement noise* is sparse with potentially unknown support, it can be identified and cancelled
- Justice and Democracy
  - Radom measurements have benefits beyond the RIP and universality
  - Concentration of measure can be a powerful tool

dsp.rice.edu/cs