Compressed Sensing: Challenges and Emerging Topics

Mike Davies
Edinburgh Compressed Sensing research group (E-CoS)
Institute for Digital Communications
University of Edinburgh

Compressed sensing

Engineering Challenges in CS:

- What is the right signal model?

  Sometimes obvious, sometimes not. When can we exploit additional structure?

- How can/should we sample?

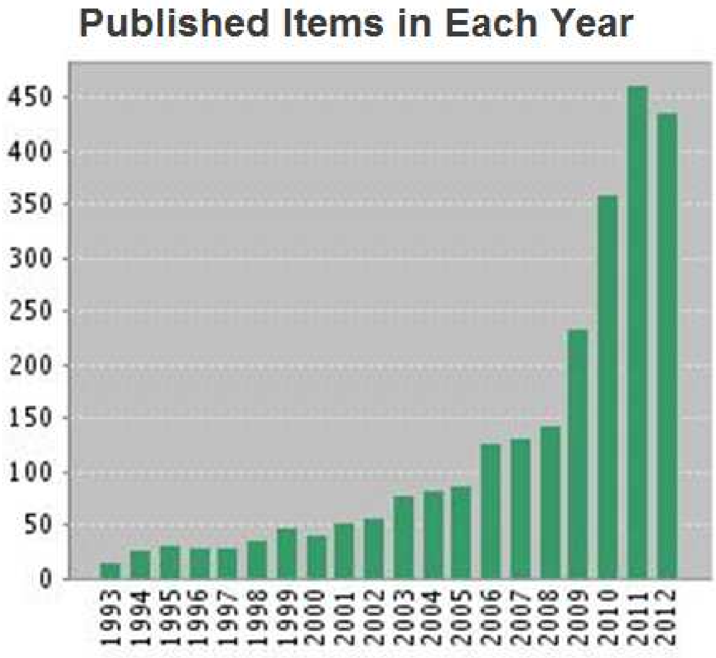
  Physical constraints; can we sample randomly; effects of noise; exploiting structure; how many measurements?

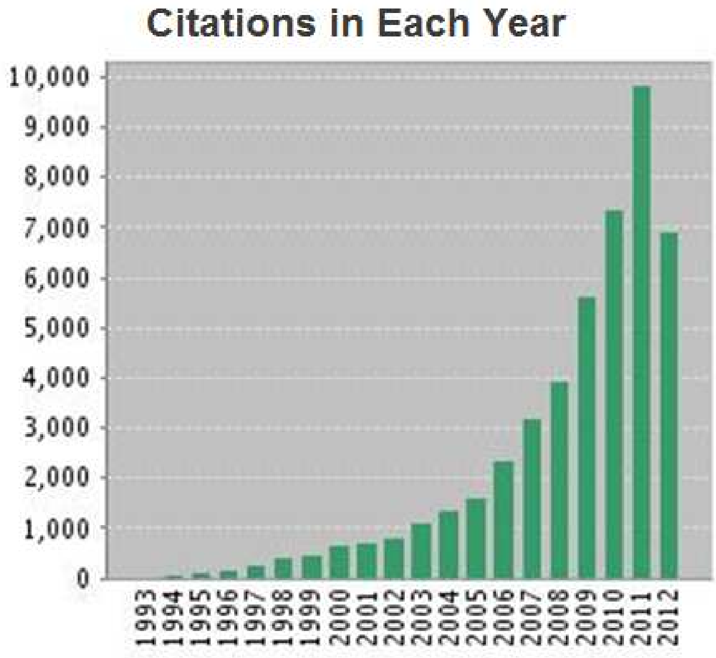
- What are our application goals?

  Reconstruction? Detection? Estimation?

CS today – the hype!

Papers published in Sparse Representations and CS [Elad 2012]





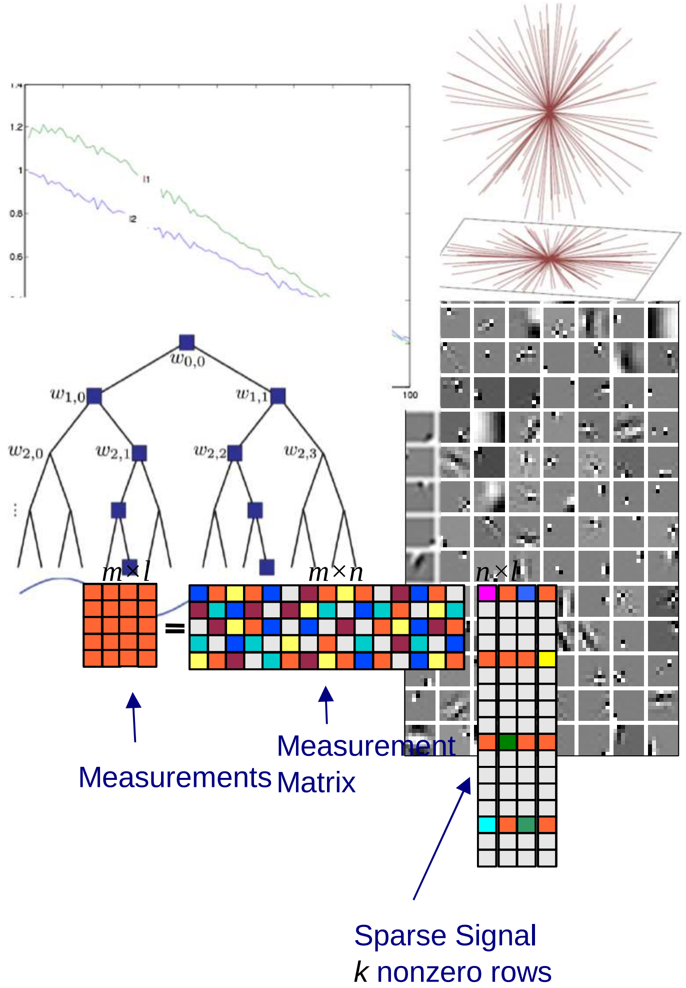
Lots of papers….. lots of excitement…. lots of hype….

Google About 4,500,000 results (0.37 seconds)

CS today: - new directions & challenges

There are many new emerging directions in CS and many challenges that have to be tackled.

- Fundamental limits in CS
- Structured sensing matrices
- Advanced signal models
- Data driven dictionaries
- Effects of quantization
- Continuous (off the grid) CS
- Computationally efficient solutions
- Compressive signal processing



Compressibility and Noise Robustness

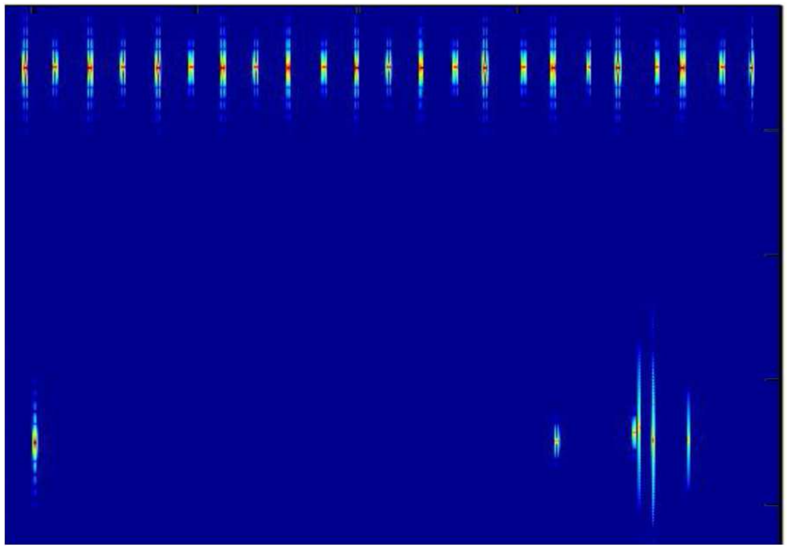
Noise/Model Robustness

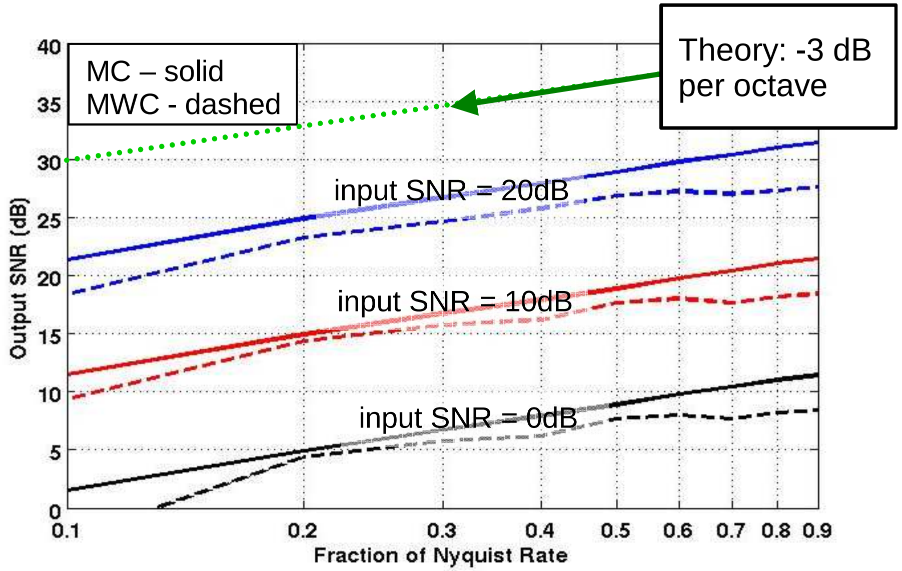
CS is robust to measurement noise (through RIP). What about signal errors, $\Phi(x+e)=y$, or when x is not exactly sparse?

No free lunch!

Wideband spectral sensing

- Detecting signals through wide band receiver noise: noise folding!
  - 3dB SNR loss per factor of 2 undersampling [Treichler et al 2011]

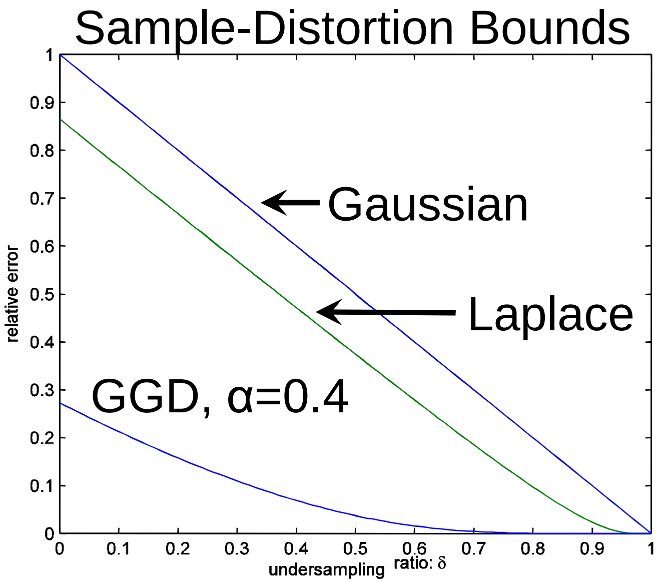




Noise/Model Robustness

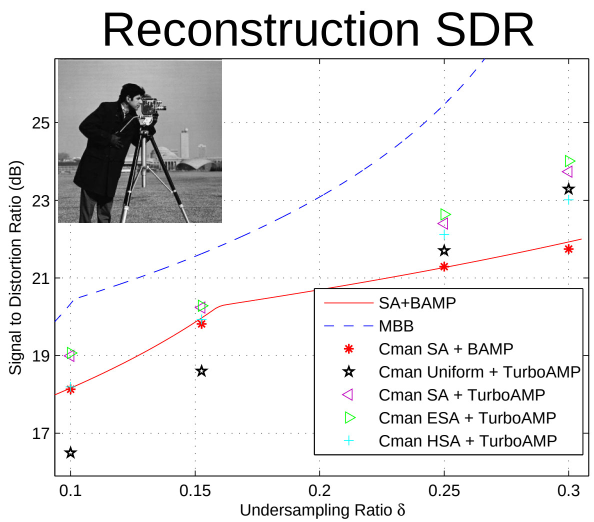
Compressible distributions

- Heavy tailed distributions may not be well approximated by low dimensional models
- Fundamental limits in terms of compressibility of the probability distribution [D. & Guo. 2011; Gribonval et al 2012]



Implications for Compressive Imaging

- Wavelet coefficients not exactly sparse
- Limits CS imaging performance



Adaptive sensing can retrieve lost SNR [Haupt et al 2011]

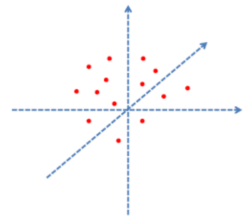
Sensing matrices

Generalized Dimension Reduction

Information preserving matrices can be used to preserve information beyond sparsity. Robust embeddings (RIP for difference vectors):

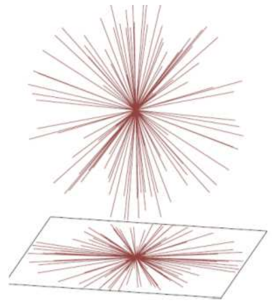
$$(1-\delta)\|x-x'\|_2 \leq \|\Phi(x-x')\|_2 \leq (1+\delta)\|x-x'\|_2$$

hold for many low dimensional sets.



- Sets of n points [Johnston and Lindenstrauss 1984]

$$m \sim \mathcal{O}(\delta^{-2} \log n)$$

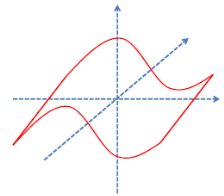


- d-dimensional affine subspaces [Sarlos 2006]

$$m \sim \mathcal{O}(\delta^{-2} d)$$

- Arbitrary Union of L k-dimensional subspaces [Blumensath and D. 2009]

$$m \sim \mathcal{O}(\delta^{-2}(k + \log L))$$



- Set of r-rank $\mathrm{n} \times l$ matrices [Recht et al 2010]

$$m \sim \mathcal{O}(\delta^{-2} r(n+l) \log nl)$$

- d-dimensional manifolds [Baraniuk and Wakin 2006, Clarkson 2008]

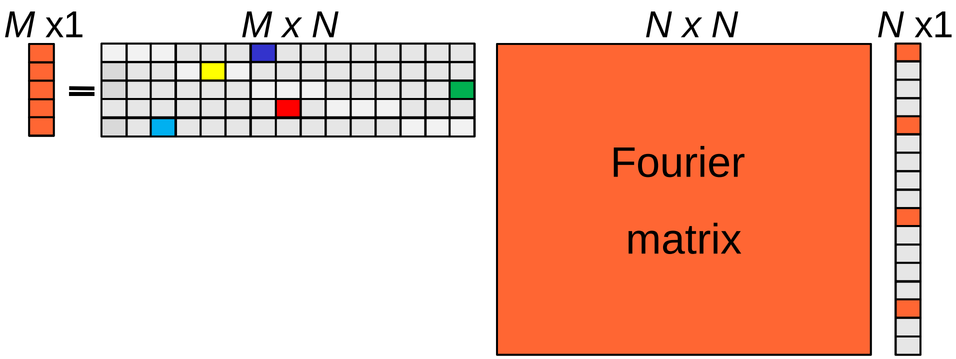
$$m \sim \mathcal{O}(\delta^{-2} d)$$

Structured CS sensing matrices

i.i.d. sensing matrices are really only of academic interest.

Need to consider wider classes, e.g.:

- Random rows of DFT [Rudelson & Vershynin 2008]



δ-RIP of order k with high probability if:

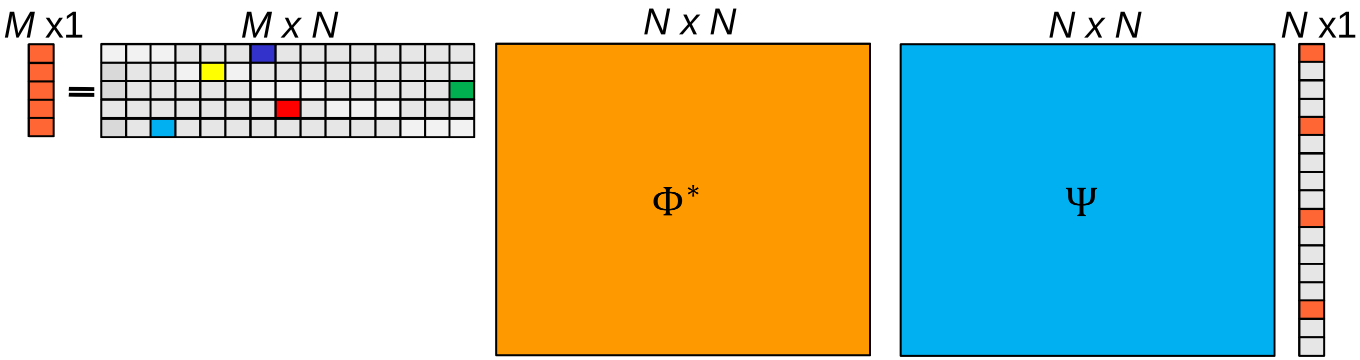
$$m \sim \mathcal{O}(k\, \delta^{-2} \log^4 N)$$

Structured CS sensing matrices

i.i.d. sensing matrices are really only of academic interest.

Need to consider wider classes, e.g.:

- Random samples of a bounded orthogonal system [Rauhut 2010]



Also extends to continuous domain signals.

δ-RIP of order k with high probability if:

$$m \sim \mathcal{O}(kN\mu(\Phi,\Psi)^2\delta^{-2}\log^4 N)$$

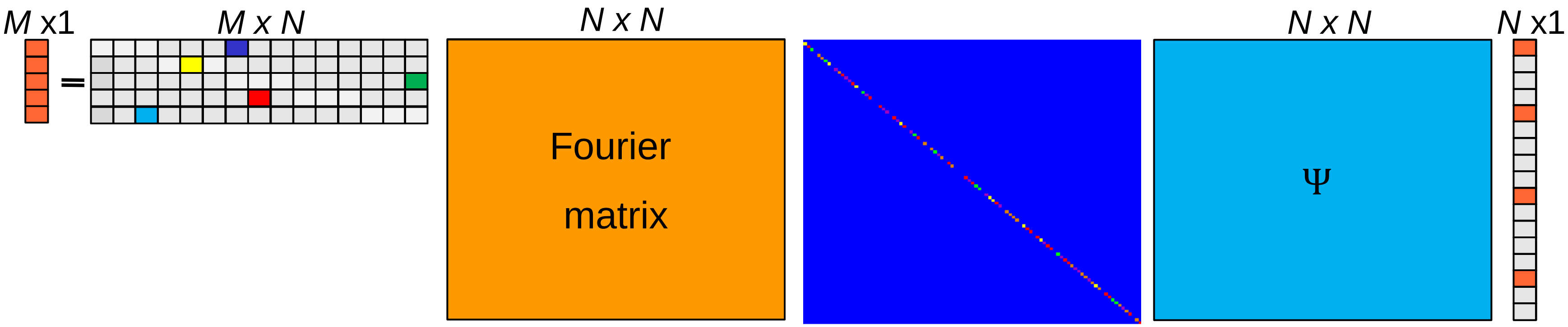
where $\mu(\Phi,\Psi) = \max_{1\le i<j\le N} \left|\langle \Phi_i, \Psi_j \rangle\right|$ is called the mutual coherence

Structured CS sensing matrices

i.i.d. sensing matrices are really only of academic interest.

Need to consider wider classes, e.g.:

- Universal Spread Spectrum sensing [Puy et al 2012]



Sensing matrix is random modulation followed by partial Fourier matrix. δ-RIP of order k with high probability if:

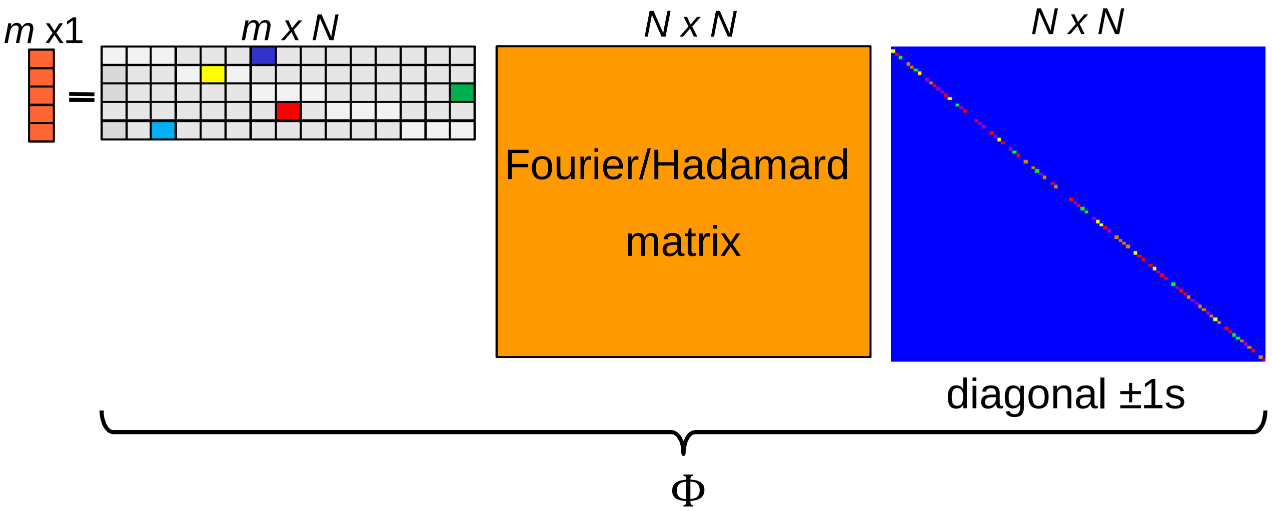
$$m \sim \mathcal{O}(k\,\delta^{-2} \log^5 N)$$

Independent of basis Ψ!

Fast Johnston Lindenstrauss Transform (FJLT)

Can generate computationally fast dimension reducing transforms [Alon & Chazelle 2006]

- The FJLT provides optimal JL dimension reduction with computation of $\mathcal{O}(N \log N)$



- Enables fast approx. nearest neighbour search
- Used in related area of sketching…

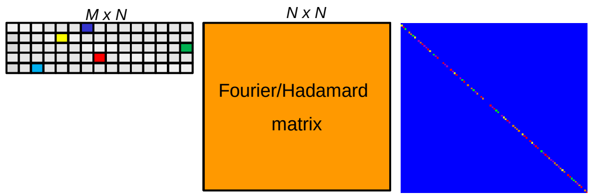
Related ideas of Sketching

e.g. want to solve l_2-regression problem [Sarlos 06]:

$$x^\star = \underset{x}{\operatorname{argmin}} \; \|Ax - y\|_2$$

with $y \in \mathbb{R}^n$, $\mathrm{A} \in \mathbb{R}^{n \times d}$.

Computational cost using normal equations: $\mathcal{O}(nd^2)$



Instead use Fast JL transform $\mathrm{S} \in \mathbb{R}^{r \times n}$ to solve:

$$\hat{\mathrm{x}} = \underset{x}{\operatorname{argmin}} \; \|(SA)x - Sy\|_2$$

If $r \sim d/\epsilon^2$ then this guarantees:

$$\|A\hat{x} - y\|_2 \leq (1+\epsilon) \; \|Ax - y\|_2$$

with high probability and at a computational cost of: $\mathcal{O}(nd \log d + \mathrm{poly}(d/\epsilon))$

– Many other sketching results possible including for constrained LS, approximate SVD, etc…

Advanced signal models & algorithms

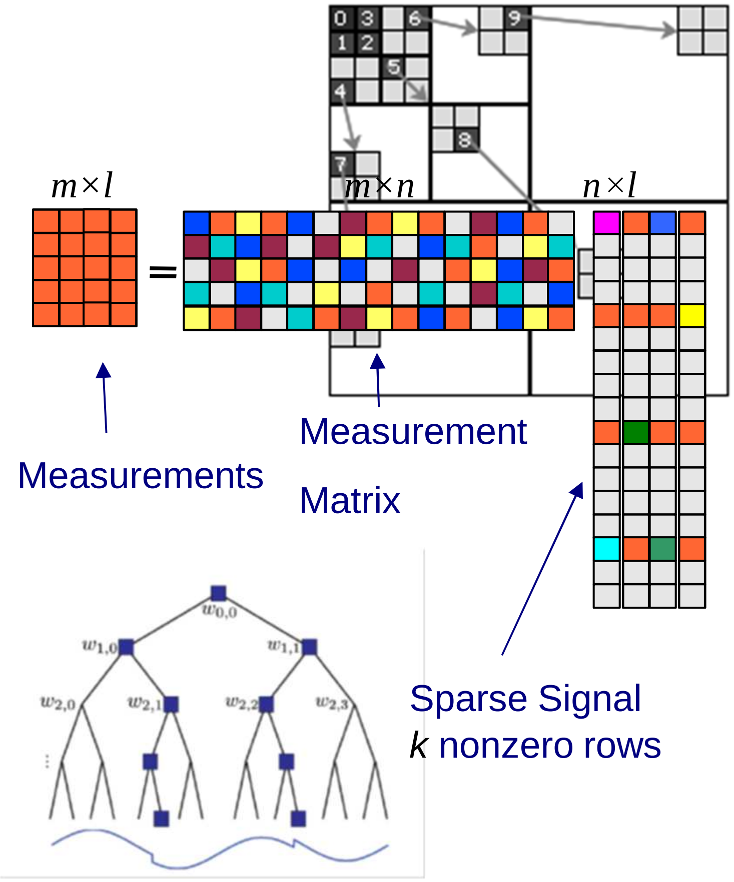
CS with Low Dimensional Models

What about sensing with other low dimensional signal models?

- Matrix completion/rank minimization
- Phase retrieval
- Tree based sparse recovery
- Group/Joint Sparse recovery
- Manifold recovery

… towards a general model-based CS?

[Baraniuk et al 2010, Blumensath 2011]



Matrix Completion/Rank minimization

Retrieve the unknown matrix $X \in \mathbb{R}^{N \times L}$ from a set of linear observations

$$y = \Phi(X),\ y \in \mathbb{R}^m \text{ with } m < NL.$$

Suppose that X is rank r.

Relax!

as with L_1 min., we convexify: replace $\mathrm{rank}(X)$ with the nuclear norm $\|X\|_* = \sum_i \sigma_i$, where σ_i are the singular values of X.

$$\hat{X} = \underset{X}{\operatorname{argmin}} \|X\|_* \text{ subject to } \Phi(X) = y$$

Random measurements (RIP) $\longrightarrow$ successful recovery if

$$m \sim \mathcal{O}(r(N+L)\log NL)$$

e.g. the Netflix prize

– rate movies for individual viewers.

Phase retrieval

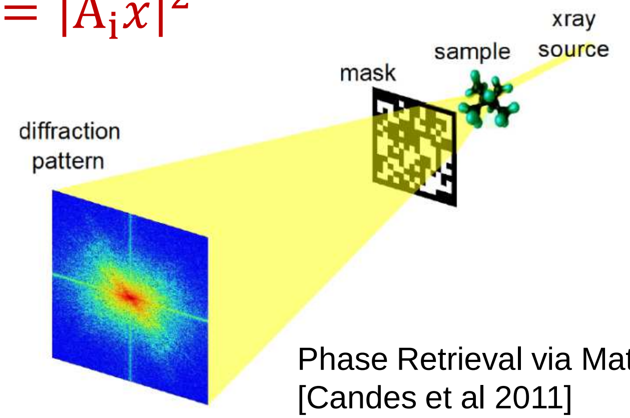
Generic problem:

Unknown $x \in \mathbb{C}^n$,

magnitude only observations: $y_i = |\mathrm{A_i} x|^2$

Applications

- X-ray crystallography
- Diffraction imaging
- Spectrogram inversion



Phase Retrieval via Matrix Completion [Candes et al 2011]

Phaselift

Lift quadratic ⟶ linear problem using rank-1 matrix $X = xx^H$

Solve: $\hat{X} = \underset{X}{\mathrm{argmin}} \|X\|_*$ subject to $\mathcal{A}(X) = y$

Provable performance but lifting space is huge! … surely more efficient solutions? Recent results indicate nonconvex solutions better.

Tree Structured Sparse Representations

Sparse signal models are type of "union of subspaces" model [Lu & Do 2008, Blumensath & Davies 2009] with an exponential number of subspaces.

$$\#\ \text{subspaces} \approx \left(\frac{N}{k}\right)^k \quad \text{(Stirling approx.)}$$

Tree structure sparse sets have far fewer subspaces

$$\#\ \text{subspaces} \approx \frac{(2e)^k}{k+1} \quad \text{(Catalan numbers)}$$

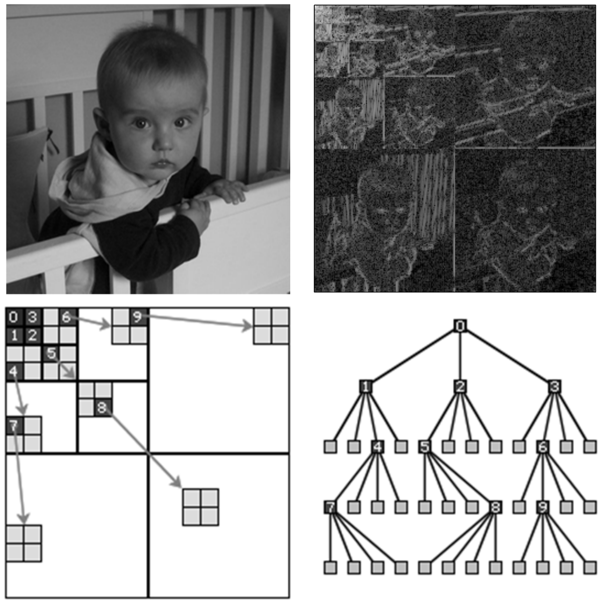
Example exploiting wavelet tree structures

Classical compressed sensing: stable inverses exist when

$$m \sim \mathcal{O}(k \log(N/k))$$

With tree-structured sparsity we only need [Blumensath & D. 2009]

$$m \sim \mathcal{O}(k)$$



Algorithms for model-based recovery

Baraniuk et al. [2010] adapted CoSaMP & IHT to construct provably good ‘model-based’ recovery algorithms.

original | sparse reconstruction | Tree sparse reconstruction

Blumensath [2011] adapted IHT to reconstruct <u>any</u> low dimensional model from RIP-based CS measurements:

$$x^{n+1} = \mathcal{P}_{\mathcal{A}}\left(x^n + \mu\Phi^{\mathrm{T}}(\mathrm{y} - \Phi x^n)\right)$$

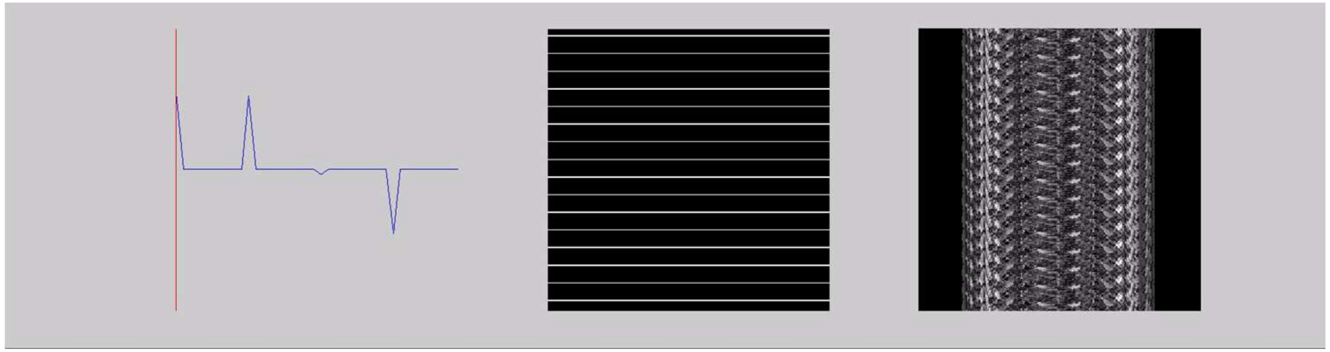
where $\mu \sim N/m$ is the step size, $\mathcal{P}_{\mathcal{A}}$ is the projection onto the signal model.

Requires a computationally efficient $\mathcal{P}_{\mathcal{A}}$ operator.

Model based CS for Quantitative MRI

[Davies et al. SIAM Imag. Sci. 2014]

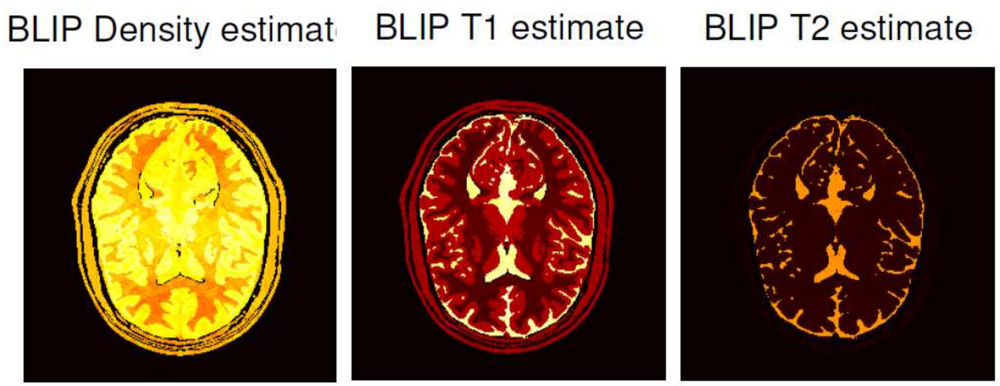
Proposes new excitation and scanning protocols based on the Bloch model



Random RF pulses | random uniform subsampling | Individual aliased images

Quantitative Reconstruction

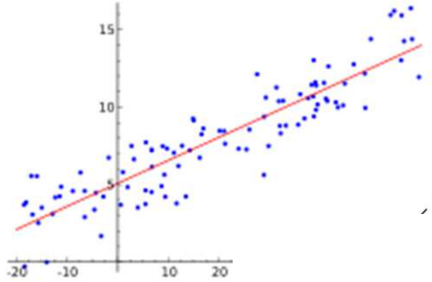
Use Projected gradient algorithm with a discretized approximation of the Bloch response manifold.

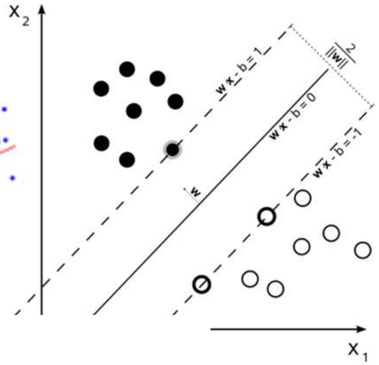


Compressed Signal Processing

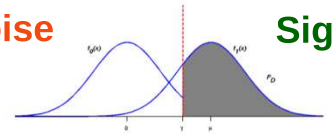
Compressed Signal Processing

There is more to life than signal reconstruction:





- Detection
- Classification
- Estimation
- Source separation



$$\mathcal{H}_0 : y = \Phi n$$
$$\mathcal{H}_1 : y = \Phi(s + n)$$

May not wish to work in large ambient signal space, e.g. ARGUS-IS Gigapixel camera

CS measurements can be information preserving (RIP)… offers the possibility to do all your DSP in the compressed domain!

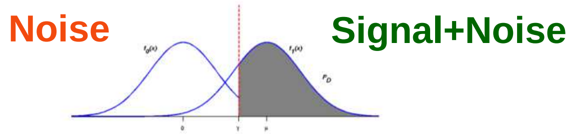
Without reconstruction what replaces Nyquist?

Compressive Detection

The ~~Matched~~ Smashed Filter [Davenport et al 2007]

Detection can be posed as the following hypothesis test:

$$\mathcal{H}_0 : z = hn$$
$$\mathcal{H}_1 : z = h(s+n)$$



The optimal (in Gaussian noise) matched filter is $h = s^H$

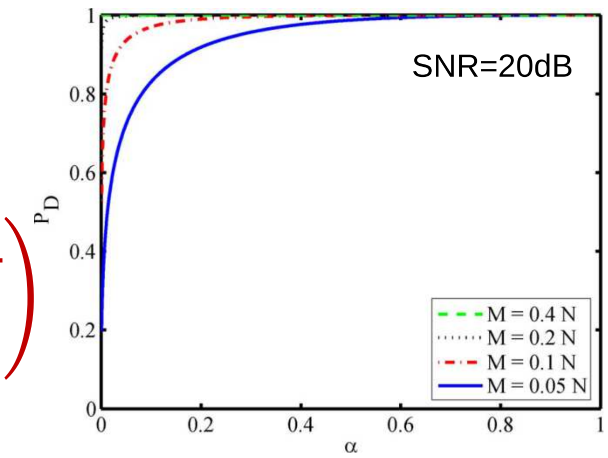
Given CS measurements: $y = \Phi s$, the matched filter (applied to y) is:

$$h = s^H \Phi (\Phi \Phi^H)^{-1}$$

Then

$$P_D \approx Q\left(Q^{-1}(\alpha) - \sqrt{\frac{m}{N}}\sqrt{SNR}\right)$$

Q - the Q-function, α – Prob. false alarm rate



[Davenport et al 2010]

Joint Recovery and Calibration

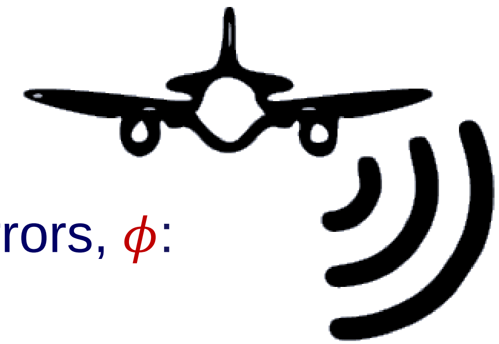
Estimation and recovery, e.g. on-line calibration.

Compressed Calibration

Real Systems often have unknown parameters θ that need to be estimated as part of signal reconstruction.

$$y = \Phi(\theta)x$$

Can we simultaneously estimate x and θ?



Example – Autofocus in SAR

Imperfect estimation of scene centre leads to phase errors, ϕ:

$$Y = \mathrm{diag}\left(e^{j\phi}\right)h(X)$$

X- scene reflectivity matrix, Y- observed phase histories, $h(\cdot)$- sensing operator.

Uniqueness conditions from dictionary learning theory [Kelly et al. 2012].

Joint Recovery and Calibration

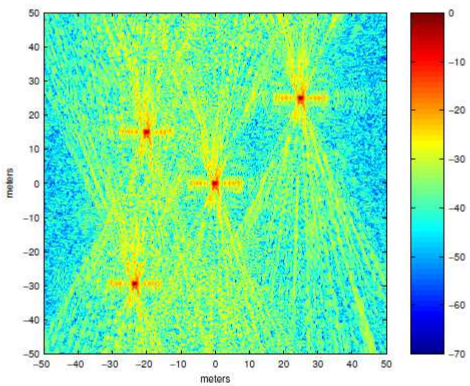
Compressed Autofocus:

Perform joint estimation and reconstruction (not convex):

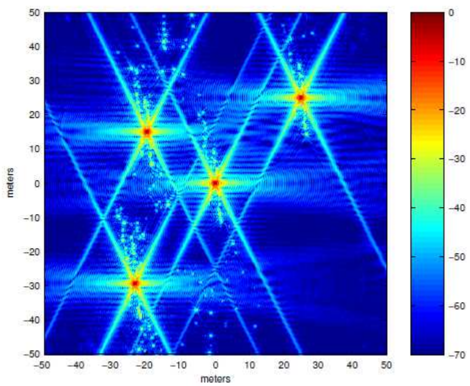
$$\min_{X,d} \|X\|_1 \quad \text{subject to } \|Y - \mathrm{diag}(d)h(X)\|_F \leq \epsilon$$

$$\text{and} \quad d_i d_i^* = 1, i = 1, \ldots, N$$

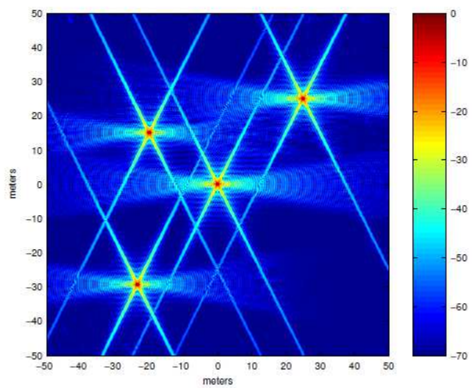
- Fast alternating optimization schemes available
- Provable performance? **Open**



No phase correction



Post-recon. autofocus



Compressive autofocus

Summary

Compressive Sensing (CS)

- combines sensing, compression, processing
- exploits low dimensional signal models and incoherent sensing strategies
- Related notion of `Sketching` in computer science allows faster computations

Still lots to do…

- Developing new and better model-based CS algorithms and acquisition systems
- Emerging field of compressive signal processing
- Exploit dimension reduction in signal processing computation: randomized linear algebra,… big data!

References

Compressibility and SNR loss

J. R. Treichler, M. A. Davenport, J. N. Laska, and R. G. Baraniuk, "Dynamic range and compressive sensing acquisition receivers," in Proc. 7th U.S. / Australia Joint Workshop on Defense Applications of Signal Processing (DASP), 2011.

M. E. Davies and C. Guo, "Sample-Distortion functions for Compressed Sensing", 49th Annual Allerton Conference on Communication, Control, and Computing, pp 902 – 908, 2011.

R. Gribonval, V. Cehver and M. Davies, "Compressible Distributions for high dimensional statistics." IEEE Trans Information Theory, vol. 58(8), pp. 5016 – 5034, 2012.

J. Haupt, R. Castro, and R. Nowak, "Distilled sensing: Adaptive sampling for sparse detection and estimation," IEEE Trans. on Inf. Th., vol. 57, no. 9, pp. 6222-6235, 2011.

Structured Sensing matrices

M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," Comm. Pure Appl. Math, vol. 61, no. 8, pp. 1025–1045, Aug. 2008.

H. Rauhut, Compressive sensing and structured random matrices. Radon Series Comp. Appl. Math., vol. 9, pp. 1-92, 2010.

G. Puy, P. Vandergheynst, R. Gribonval and Y. Wiaux, Universal and ecient compressed sensing by spread spectrum and application to realistic Fourier imaging techniques. EURASIP Journal on Advances in Signal Processing, 2012, 2012:6.

References

Information Preserving Dimension Reduction

W. B. Johnson and J. Lindenstrauss, Extensions of Lipschitz maps into a Hilbert space, Contemp Math 26, pp.189–206, 1984.

R. G. Baraniuk and M. B. Wakin, Random Projections of Smooth Manifolds. Foundations of Computational Mathematics, vol. 9(1), pp. 51-77, 2009.

K. Clarkson, Tighter Bounds for Random Projections of Manifolds. Proceedings of the 24th annual symposium on Computational geometry (SCG'08), pp. 39-48, 2008.

B. Recht, M. Fazel, and P. A. Parrilo, Guaranteed Minimum Rank Solutions to Linear Matrix Equations via Nuclear Norm Minimization.. SIAM Review. Vol 52, no 3, pages 471-501. 2010.

T. Sarlos. Improved approximation algorithms for large matrices via random projections. In FOCS2006: Proc. 47th Annual IEEE Symp. on Foundations of Computer Science, pp. 143–152, 2006.

Structured Sparsity & Model-based CS

Baraniuk, R.G., Cevher, V., Duarte, M.F. & Hegde, C., Model-based compressive sensing. IEEE Trans. on Information Theory 56:1982-2001, 2010.

T. Blumensath, Sampling and Reconstructing Signals From a Union of Linear Subspaces. IEEE Trans. Inf. Theory, vol. 57(7), pp. 4660-4671, 2011.

M. E. Davies, Y. C. Eldar: Rank Awareness in Joint Sparse Recovery. IEEE Transactions on Information Theory 58(2): 1135-1146 (2012).

References

Compressed Signal Processing

M. Davenport, M. Duarte, M. Wakin, J. Laska, D. Takhar, K. Kelly and R. Baraniuk, “The smashed filter for compressive classification and target recognition,” in Proc. SPIE Symp. Electron. Imaging: Comput. Imaging, San Jose, CA, Jan. 2007.

M. Davenport, P. T. Boufounos, M. Wakin and R. Baraniuk, “Signal Processing with Compressive Measurements,” IEEE J. of Sel. Topics in SP, vol. 4(2), pp. 445-460, 2010.

S. I. Kelly, M. Yaghoobi, and M. E. Davies, “Auto-focus for Compressively Sampled SAR,” 1st Int. Workshop on CS Applied to Radar. May 2012.