Compressed Sensing of Continous Index Signals and of Low Rank Matrices

Yoram Bresler

Coordinated Science Laboratory and Department of ECE

University of Illinois at Urbana-Champaign

Acknowledgement:

NSF Grant CCF 06-35234

Outline

- Part I: compressive sensing of continuous-index signals (work with Ping Feng and Raman Venkataramani)

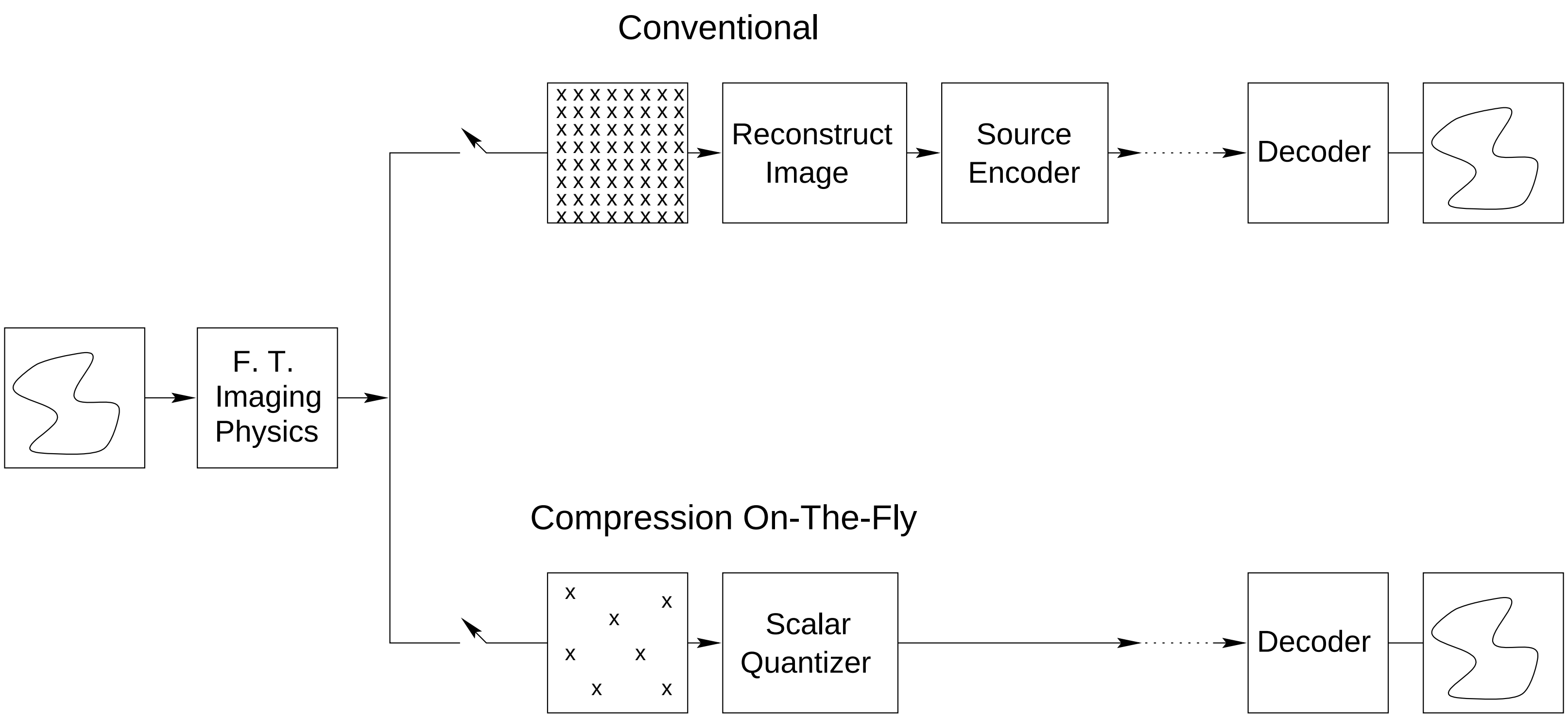
- Part II: Guaranteed low-rank matrix approximation from linear observations (work with Kiryung Lee)

Applications:

Real-world Signals are continuous indexed:

- Communications
- ELINT
- Fourier imaging: MRI, CT, SAR
  - Sparse image $X(f) \quad f \in \mathbb{R}^d$
  - Sampling in $t \in \mathbb{R}^d$

Compression On-The-Fly



1998 - IT Workshop

Spectrum-Sparse Continuous-Index Signals

- 1-D continuous-time signals

$$x(t) \leftrightarrow X(f) \quad t, f \in \mathbb{R}$$

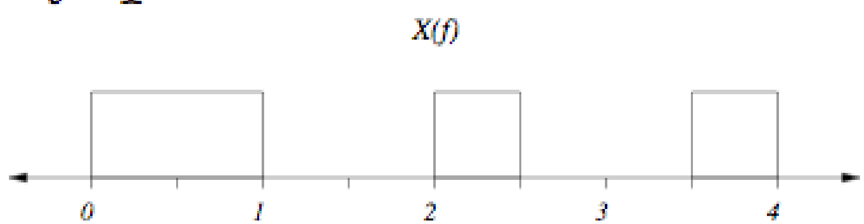
- 1-D discrete-time signals (infinite sequences)

$$x(n) \leftrightarrow X(f) \quad n \in \mathbb{Z}, \quad f \in \mathbb{R}$$

Spectrum-Sparse:

$$X(f) = 0,\ f \notin \mathcal{F}, \quad \mathcal{F} = \bigcup_{i=1}^{n} [a_i, b_i) \subseteq [f_{\min}, f_{\max}]$$

$$\frac{\lambda(\mathcal{F})}{f_{\max} - f_{\min}} \leq \Omega < 1$$



Questions

- Sampling rate requirements ?
  - Known $\mathcal{F}$
  - Unknown $\mathcal{F}$
- Sampling at the minimum rate ?
  - Design of sampling scheme
  - Reconstruction
- How achieve:
  - Universal (non-adaptive sampling)
  - Perfect (or robust) reconstruction
  - Computation linear in the data size
- Relationship to compressive sensing?

Sampling Rate

Throw in the towel (sufficient condition)

- Nyquist sampling: $f_{\mathrm{Nyq}} = f_{\max} - f_{\min}$

Necessary conditions

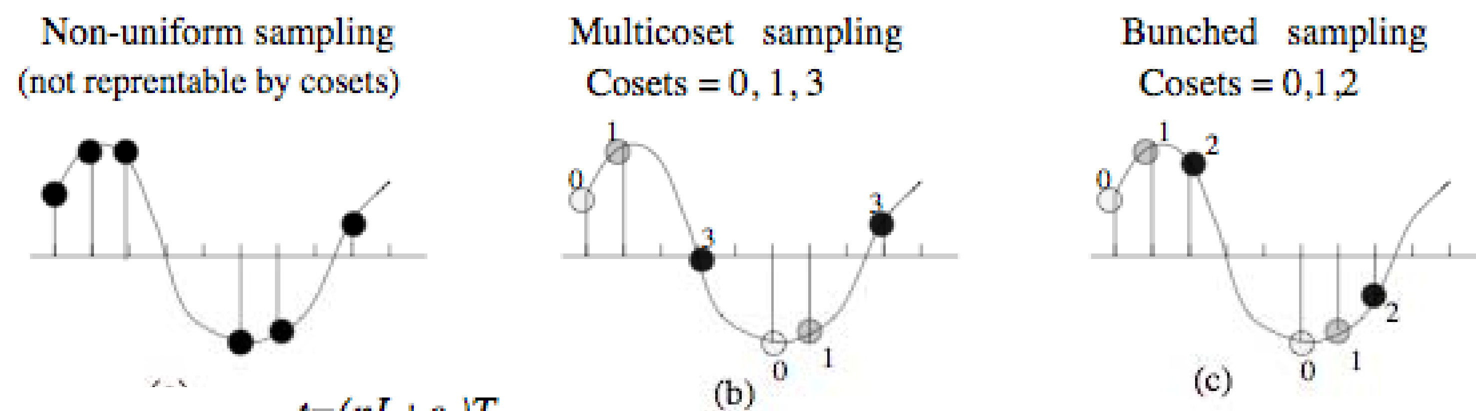
- Landau lower bound (1967) $D^- \geq \lambda(\mathcal{F})$
  - Arbitrary pointwise sampling
  - Sufficient for known $\mathcal{F}$ packable or not (Kahn & Liu, 1965)

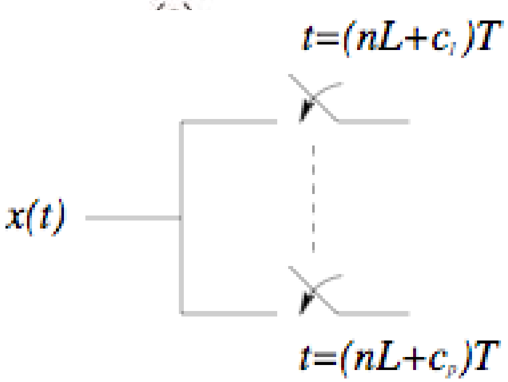
$$D^- \approx \lambda(\mathcal{F}) \leq \Omega f_{\mathrm{Nyq}}$$

Sufficient conditions, unknown $\mathcal{F}$ (Feng & Bresler, 1996):

$$D^- \approx \begin{cases} 2\Omega f_{\mathrm{Nyq}} & \text{for all signals} \\ \Omega f_{\mathrm{Nyq}} & \text{for almost all signals} \end{cases}$$

Periodic Sampling

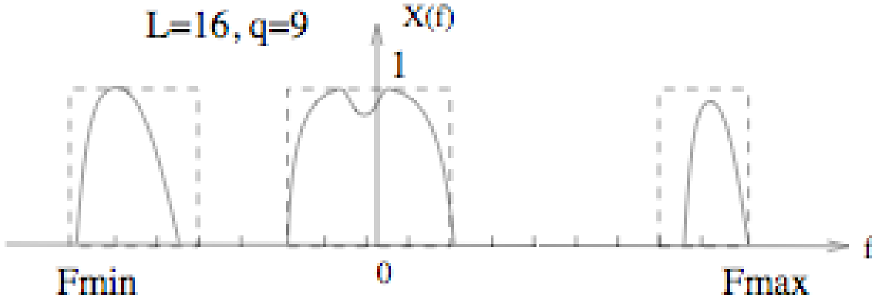




$$\frac{1}{T} \geq f_{\mathrm{Nyq}}$$

- (L, p) sampling pattern: $\mathcal{C} = \{c_i : 1 \leq i \leq p\}$
- Period = L

Spectral support at resolution L



Active spectral cells:

$$\mathcal{F}_i \cap \mathcal{F} \neq \emptyset, i = 1, 2, \ldots q$$

$$\Omega_L \triangleq \frac{q}{L} \approx \Omega$$

$$\mathcal{F}_0 = [0, \frac{1}{LT}], \quad \mathcal{F}_l \triangleq \mathcal{F}_0 + l/L, \quad l = 1, 2, \ldots, L-1,$$

Vectorized signal spectrum

$$X_l(f) \triangleq X\left(f + \tfrac{l}{LT}\right) \chi(f; \mathcal{F}_0)$$

$$\boldsymbol{\zeta}(f) \triangleq [X_1(f), \ldots X_L(f)]', \ f \in \mathcal{F}_0$$

Vectorized sample spectrum

$$x_i(m) = x(Lm + c_i), \ m \in \mathbb{Z} \leftrightarrow X_i(f)$$

$$\boldsymbol{y}(f) \triangleq LT[e^{-j2\pi c_1 fT} X_1(LTf), \ldots, e^{-j2\pi c_1 fT} X_p(LTf)]',$$

$$\boldsymbol{y}(f) = \boldsymbol{A}\boldsymbol{\zeta}(f), \ f \in \mathcal{F}_0, \quad \boldsymbol{A} \in \mathbb{C}^{p \times L} =$$ Submatrix of DFT $L \times L$

Blind Reconstruction & Compressive Sensing

$$\boldsymbol{y}(f) = \boldsymbol{A}\boldsymbol{\zeta}(f),\ f \in \mathcal{F}_0, \quad \boldsymbol{A} \in \mathbb{C}^{p \times L} =$$ Submatrix of DFT

$$\|\boldsymbol{\zeta}(f)\|_0 \leq \Omega_L L,\ f \in \mathcal{F}_0$$

$$\textbf{P0:}\ \text{for each } f \in \mathcal{F}_0 \quad \min_{\boldsymbol{\zeta}(f)} \|\boldsymbol{\zeta}(f)\|_0 \ \text{ subject to } \boldsymbol{y}(f) = \boldsymbol{A}\boldsymbol{\zeta}(f)$$

- Common sparsity pattern for all $f \in \mathcal{F}_0$
- Simultaneous sparse approximation (Rao *et al* 2005, Chen & Huo 2006, Tropp *et al, 2006)*
- *But, uncountably infinite number of simultaneous measurement vectors*

Spectral Support Recovery

- Active spectral cell indices:

$$\boldsymbol{k} = [k_1, k_2, ..., k_q]' \; : \; \mathcal{F}_{k_r} \cap \mathcal{F} \neq \emptyset,$$

$$k_r \in \{1, 2, ..., L\} \quad r = 1, ..., q,$$

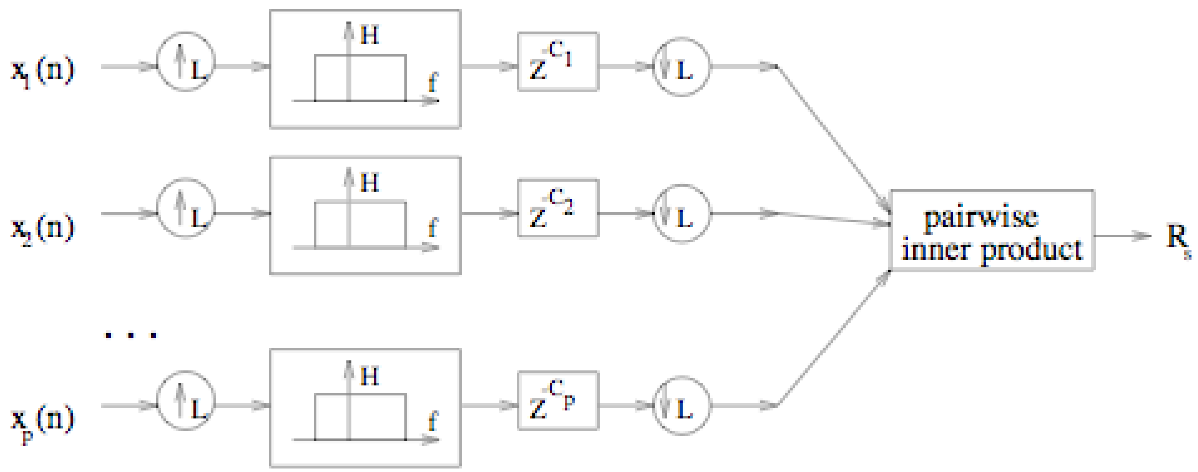
$$\textbf{P1:} \; \hat{\boldsymbol{k}} = \arg \min_{\|\boldsymbol{k}\|_0 = q} \int_{f \in \mathcal{F}_0} \|\boldsymbol{P}_{\text{Range}^{\perp}(\boldsymbol{A_k})} \boldsymbol{y}(f)\|_2^2 df$$

$$\begin{aligned} \textbf{P2:} \; \hat{\boldsymbol{k}} &= \arg \min_{\|\boldsymbol{k}\|_0 = q} \operatorname{tr}(\boldsymbol{P}_{\text{Range}^{\perp}(\boldsymbol{A_k})} \boldsymbol{R}) \\ \boldsymbol{R} &\triangleq \int_{f \in \mathcal{F}_0} \boldsymbol{y}(f) \boldsymbol{y}^*(f) df \\ (R)_{kl} &= \langle x_k(n - \frac{c_k}{L}), x_l(n - \frac{c_l}{L}) \rangle . \end{aligned}$$

- Finite dimensional problem!

Finite-dimensional optimization

$$\textbf{P2:}\ \hat{\boldsymbol{k}} = \arg \min_{q, \|\boldsymbol{k}\|_0 = q} \operatorname{tr}\left[\boldsymbol{P}_{\text{Range}^{\perp}(\boldsymbol{A}_{\boldsymbol{k}})} \boldsymbol{R}\right]$$



- Solution of P2 using greedy algorithm (Feng & Bresler, 1996; Venkataramani & Bresler, 1998)

More equivalent problems

P3: Given $\boldsymbol{R} \in \mathbb{C}^{p\times p}$ find the smallest integer q and spectral cell index vector $\boldsymbol{k}$ of length q such that

$$\boldsymbol{R} = \boldsymbol{A}_k \boldsymbol{Z} \boldsymbol{A}_k^* \text{ for some } \boldsymbol{Z} \in \mathbb{C}^{q\times q},\ \boldsymbol{Z} \succeq 0$$

- Let $\boldsymbol{R} = \boldsymbol{U}_s \Lambda_s \boldsymbol{U}_s^*, \quad \boldsymbol{U} \in \mathbb{C}^{p\times r}, r = \text{rank}(\boldsymbol{R}) = \text{rank}(\boldsymbol{Z})$

P4: Given ***U*** find the smallest integer q and spectral cell index vector $\boldsymbol{k}$ of length q such that

$$\boldsymbol{U} = \boldsymbol{A}_k \boldsymbol{Q} \text{ for some } \boldsymbol{Q} \in \mathbb{C}^{q\times q}$$

- **Proposition:** Problems P0 -- P4 are equivalent

- P4 is the (now) classical MMV problem!
- Mishali+Eldar (2007) proposed to solve it using MMV compressed sensing methods.

Sampling Conditions

Theorem (Feng & Bresler, 1996). Let $\mathcal{X}$ consist of signals with spectral support of spectral occupancy at most Ω_L at resolution *L*. Then

(1) all signals in $\mathcal{X}$ can be uniquely reconstructed from their samples on a periodic pattern with average rate $\approx 2\Omega_L f_{\max}$

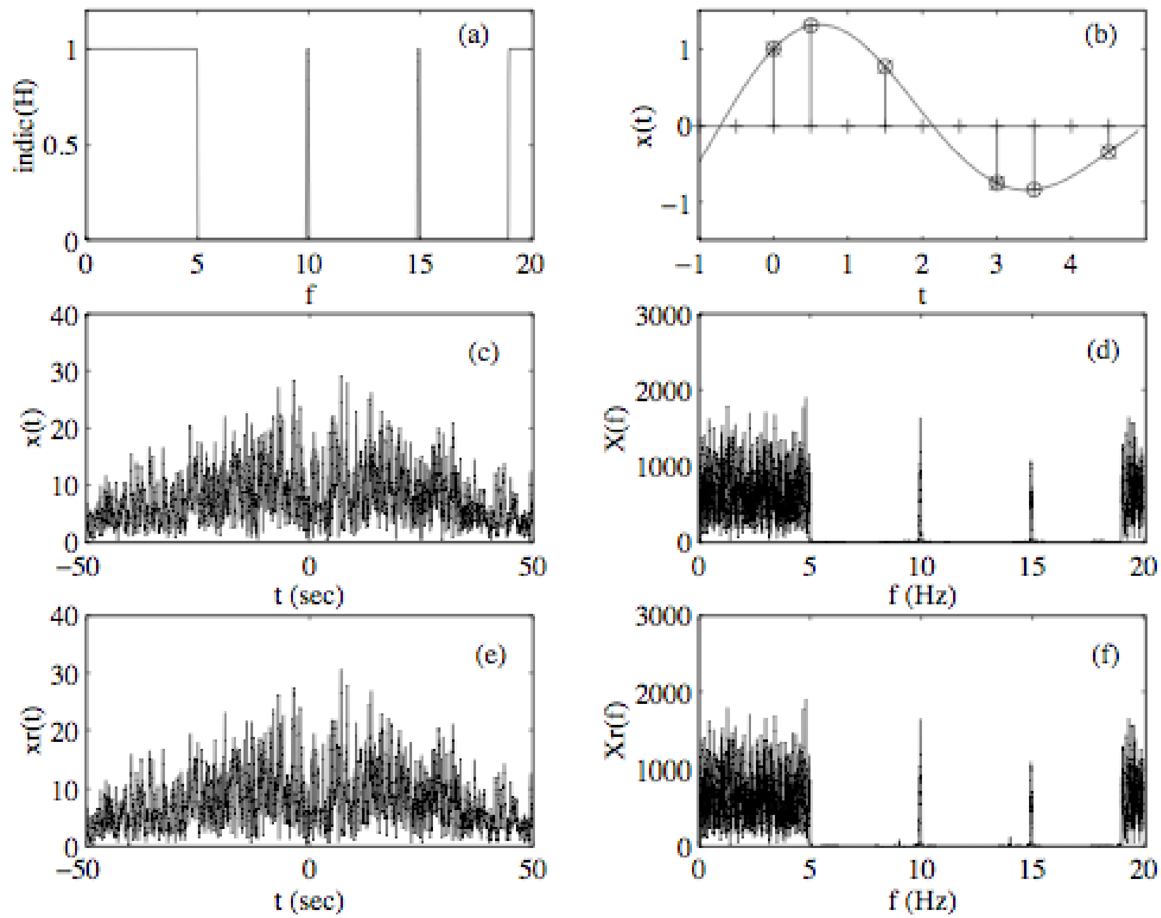
(2) almost all signals in $\mathcal{X}$ can be reconstructed from samples at rate $\approx \Omega_L f_{\max}$.

Proof: (1) For bunched pattern with $p \geq 2q+1$

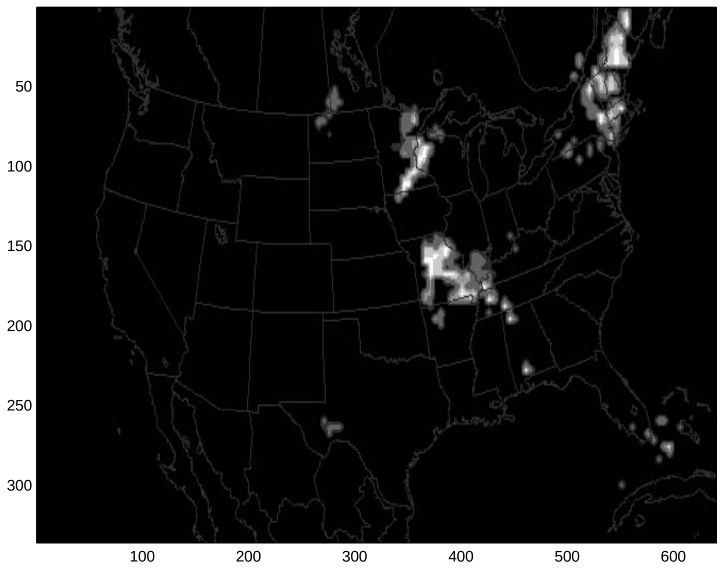
$$\mathrm{Spark}(\boldsymbol{A}) \geq 2q+1$$

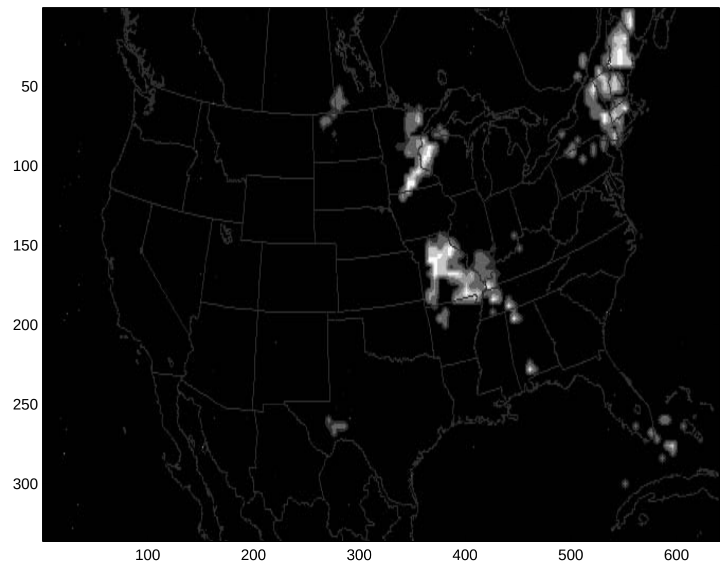
(b) Almost all ***Z*** have full rank.

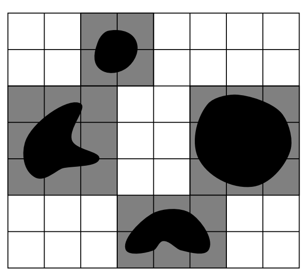
Numerical Experiments



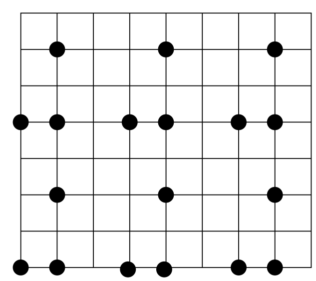
2D Spectrum-Blind Sampling*







(a)



(b)

*ICIP 1998

Conclusions

- Spectrum-Blind sampling generalizes Compressed Sensing, and reduces the infinite dimensional problem to a finite dimensional problem without discretization.
- Compressed sensing with DFT measurement and time-sparse signals in Euclidean spaces is a special case of spectrum-blind sampling.
- Compressed Sensing theory provides new insights and provable, efficient techniques to solve the spectrum-blind sampling problem.

Guaranteed Low-Rank Matrix Approximation from Linear Observations

Work with Kiryung Lee

ECE/CSL UIUC

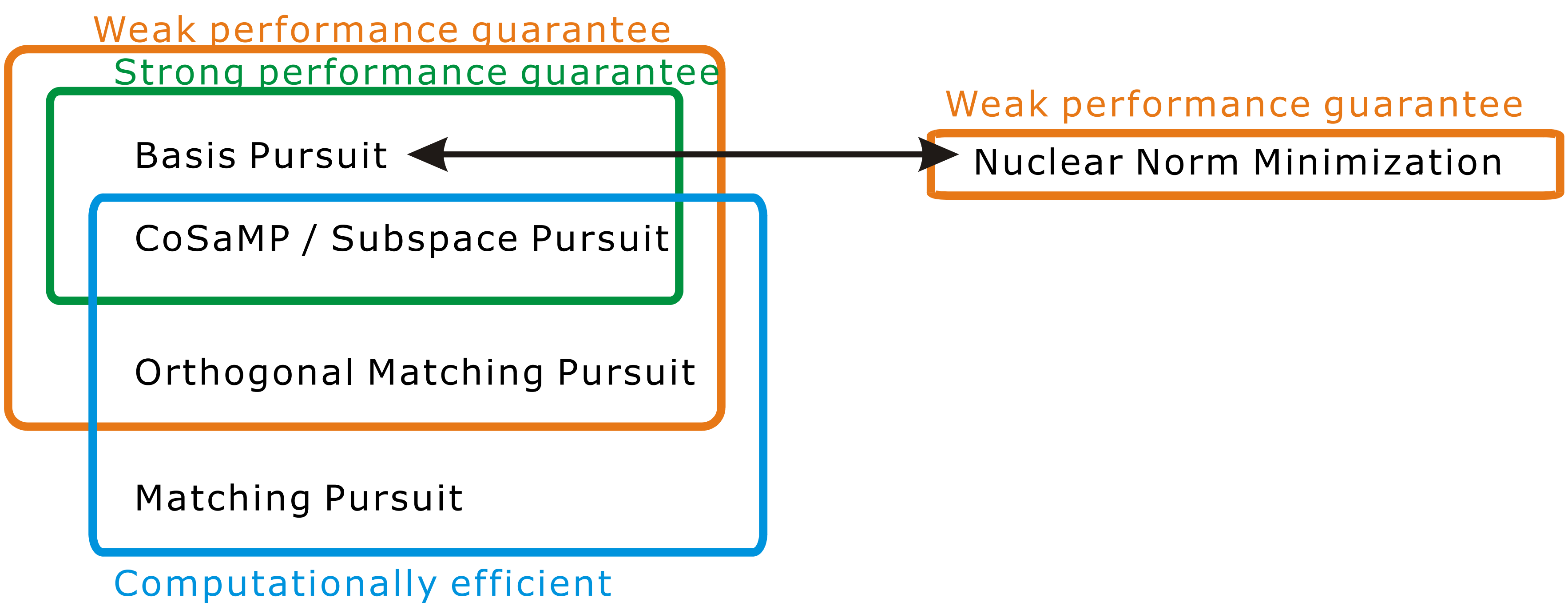
$$\mathcal{A} : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{p}, X \in \mathbb{C}^{m \times n}, y \in \mathbb{C}^{p}$$

$$y = \mathcal{A}X + v,$$

$$\min \operatorname{rank}(X), \text{subject to} \left\| y - \mathcal{A}X \right\|_2 < \varepsilon$$

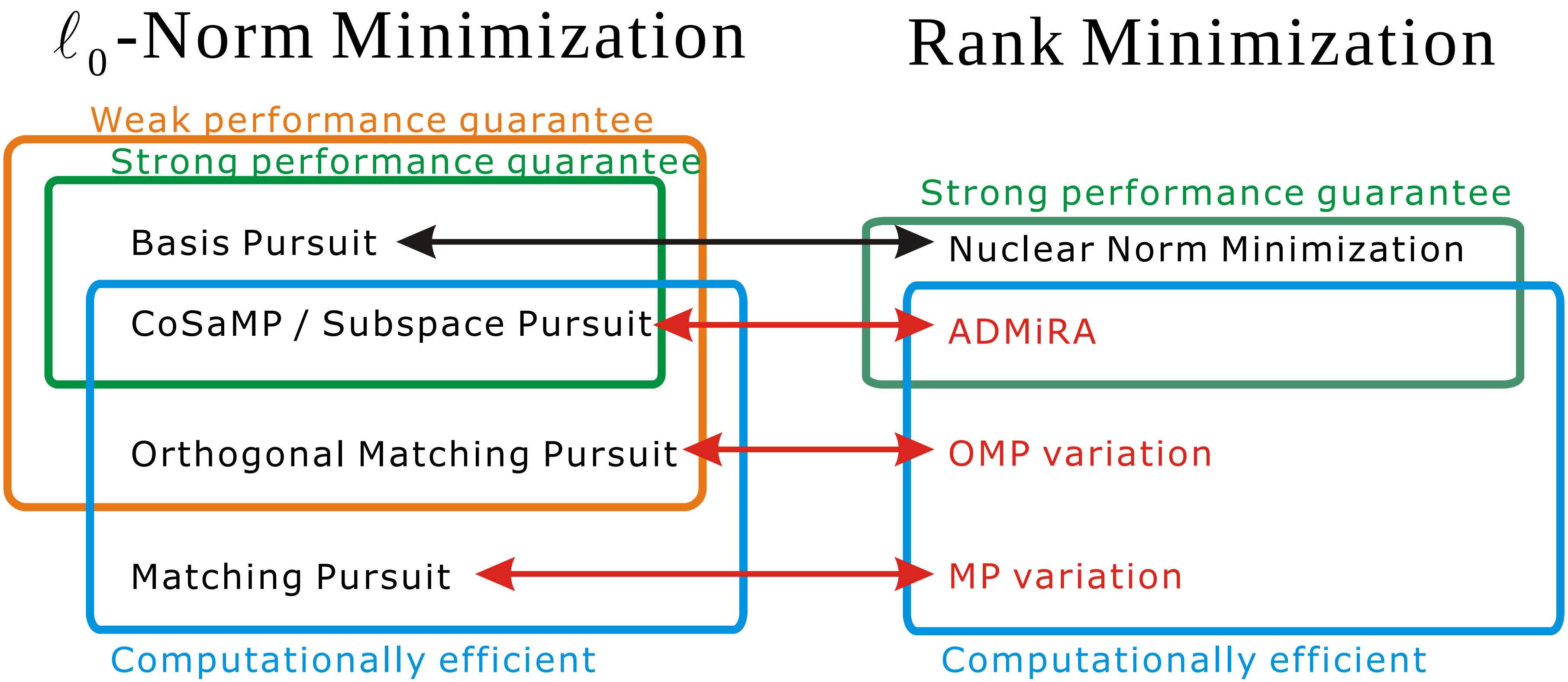
Rank Minimization:
Analogy – Previous Works

ℓ_0-Norm Minimization | Rank Minimization



Performance guarantee	Sparsity (or low-rank)	Measurements
Weak	Exact	Exact - noiseless
Strong	Approximate	Noisy

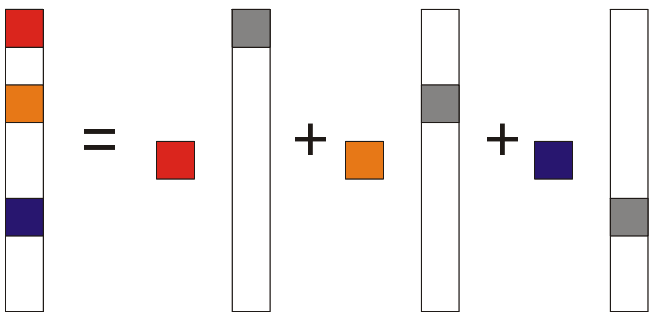
Rank Minimization: Analogy – New Contribution



Performance guarantee	Sparsity (or low-rank)	Measurements
Weak	Exact	Exact - noiseless
Strong	Approximate	Noisy

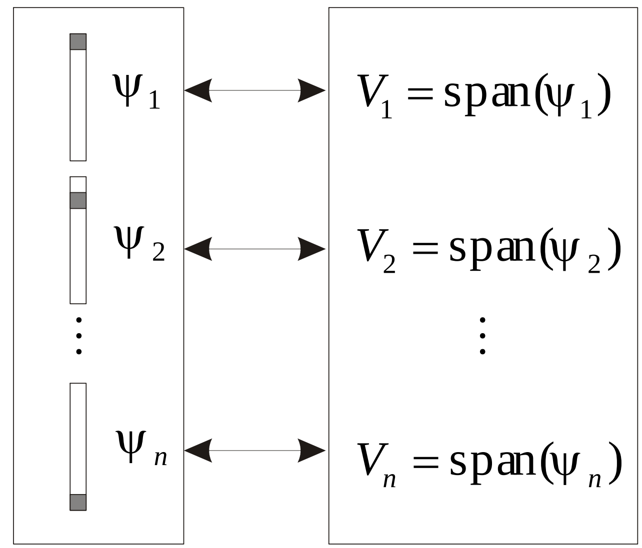
Vector Atomic Decomposition

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



$$x = c_1\psi_i + c_2\psi_j + c_3\psi_k$$

$\mathbb{O}$: set of atoms $\mathbb{A}$: set of atomic spaces



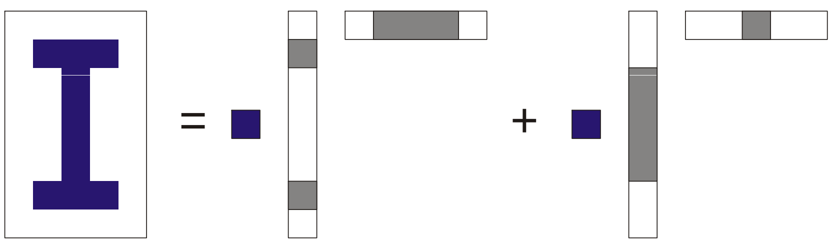
finite

$\|x\|_0 \le s$ *iff* x is spanned by s atoms

$$x \in \bigcup_{i=1}^{n} \left(\sum_{k=1}^{s} V_{i_k} \right) = \bigcup_{i=1}^{n} \mathrm{span}\left(\{\psi_{i_k}\}_{k=1}^{s} \right)$$

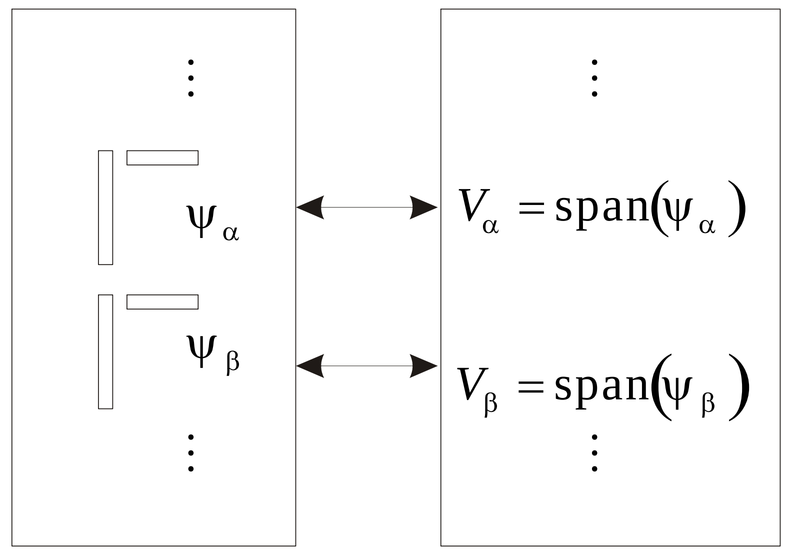
Matrix Atomic Decomposition

$X \in \mathbb{C}^{m\times n}$



$$X = c_1\psi_\alpha + c_2\psi_\beta$$

$\mathbb{O}$: set of atoms $\mathbb{A}$: set of atomic spaces



uncountably infinite

$\mathrm{rank}(X) \le r$ *iff* X is spanned by r atoms

$$X \in \bigcup_{|\Gamma|\le r}\left(\sum_{\alpha\in\Gamma} V_\alpha\right) = \bigcup_{\Psi\subset\mathbb{O},|\Psi|\le r} \mathrm{span}(\Psi)$$

Vector Correlation Maximization

$A \in \mathbb{C}^{p \times n}, y \in \mathbb{C}^p$

a_k : k-th column of A

e_k : k-th column of I_n ($n \times n$ identity matrix)

P_{e_k} : projection onto span(e_k)

$$|\langle y, a_k \rangle| = |\langle y, Ae_k \rangle| = |\langle A^H y, e_k \rangle| = \left\| P_{e_k} A^H y \right\|_2$$

Find a column a_k that maximizes the correlation

$\Leftrightarrow$ Find an atom e_k that maximizes $\left\| P_{e_k} A^H y \right\|_2$

Matrix Correlation Maximization

$\mathcal{A} : \mathbb{C}^{m\times n} \to \mathbb{C}^p, y \in \mathbb{C}^p$

$\mathcal{A}^* :$ adjoint of $\mathcal{A}$

$\psi :$ an atom

$P_\psi :$ projection onto span(ψ)

$|\langle y, \mathcal{A}\psi \rangle| = |\langle \mathcal{A}^* y, \psi \rangle| = \left\| P_\psi \mathcal{A}^* y \right\|_F$

Find an atom ψ that maximizes the correlation $\left| \langle y, \mathcal{A}\psi \rangle \right|$

$\Leftrightarrow$ Find an atom ψ that maximizes $\left\| P_\psi \mathcal{A}^* y \right\|$

- Can be done by SVD of matrix $\mathcal{A}^* y$

Matrix Correlation Maximization

- Greedy algorithms (MP,OMP,CoSaMP,SP) are extended from the vector case to the matrix case
- In particular, we propose **ADMiRA**

  (= **A**tomic **D**ecomposition for **MI**nimum **R**ank **A**pproximation)

that is **efficient** in computation and has a **performance guarantee**.

ADMiRA

- Analogous to CoSaMP/SP
- Iterative subspace selection
- Assumptions
  - Target rank is fixed to *r*
  - Linear operator satisfies RIP: $\delta_{7r} < 0.056$
  - Measurement is obtained by $b = \mathcal{A}X + \nu$

ADMiRA

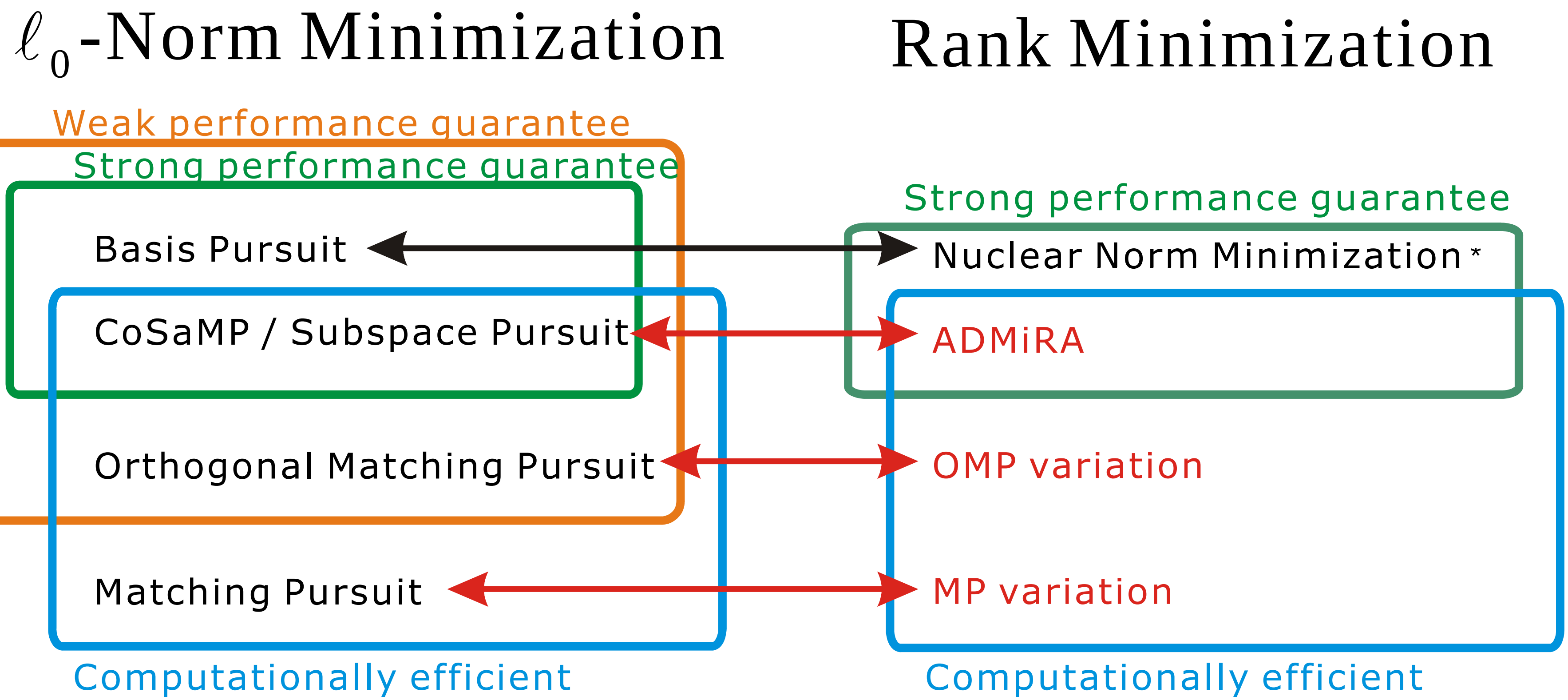
- Performance guarantee
  – After 6(r+1) iterations, ADMiRA provides a rank-r approximation satisfying

$$\left\| X - \hat{X} \right\|_F \leq 20 \left(\left\| X - X_r \right\|_F + \frac{1}{\sqrt{r}} \left\| X - X_r \right\|_* + \left\| \nu \right\|_2 \right)$$

X_r : best rank-r approximation, $\| \ \|_*$: nuclear norm

- Efficient computation
  – most computation is (truncated) SVD
  – Randomized algorithms are applicable

Rank Minimization:
Analogy – New Contribution



Performance guarantee	Sparsity (or low-rank)	Measurements
Weak	**Exact**	**Exact - noiseless**
Strong	**Approximate**	**Noisy**

* Note: after the presentation we learned about the paper by M. Fazel, E. Cand`es, B. Rechtx, & P. Parrilo, "Compressed Sensing and Robust Recovery of Low Rank Matrices," Asilomar Conference, Nov. 2008, which also presented a strong performance guarantee for nuclear norm minimization, although with different constants than ours.