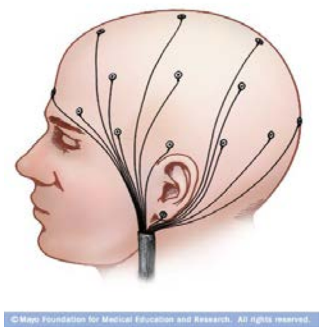
SENSING MATRIX DESIGN CRITERIA FOR ADAPTIVE COMPRESSED SENSING

Valerio Cambareri
Ph.D. Student, DEI – University of Bologna
valerio.cambareri@unibo.it

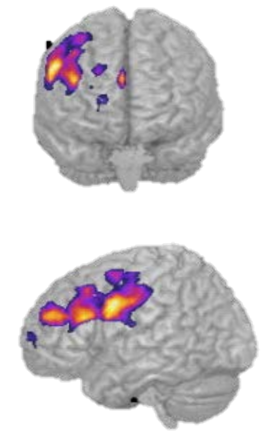
Supervisors:
Prof. Riccardo Rovatti, Prof. Gianluca Setti
SSIGPRO – Statistical Signal Processing Group
ARCES – Advanced Research Center on Electronic Systems
University of Bologna

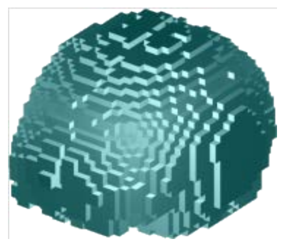
- **A Brief Introduction to Compressed Sensing**
  - Linear Inverse Problems
  - Compressed Sensing in a Nutshell
  - The Sensing Matrix Design Problem
  - (Energy) Localization
- ***Rakeness-based* Random Sensing Matrices**
  - Maximum Energy Projections
  - Rakeness: Definition and Problem Statement
  - A Rakeness-based Design Flow
  - Synthesis with Bernoulli Random Matrices
- ***Maximum-Entropy* Deterministic Sensing Matrices**
  - Deterministic Ensembles
  - Localized Signals and Correlated Measurements
  - The Maximum-Entropy Principle
  - A Maximum-Determinant Heuristic
  - Experimental Results
- **Conclusion**



Noise

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\nu}$$





Measurements

Examples:

- Scalp EEG measurements
- Photons counted at detector
- Observations

Sensing Operator (matrix)

Examples:

- Brain tissue model
- Imaging channel/field propagation model
- Design matrix

Input Signal

Examples:

- Intracranial current density sources
- Radiating object
- Parameters

$$\mathbf{y}_{(M \times 1)} = \mathbf{A}_{(M \times N)} \mathbf{x}_{(N \times 1)}$$

- Measurements have a **cost** we would like to minimize: *make M as small as possible*
  - Underdetermined case: when $M < N$, there is *a whole subspace* of solutions
  - For some systems, the solution may be **sparse**
- **Sparsity**: let $\mathbf{x} = \mathbf{D}_{(N \times P)}\ \mathbf{s}_{(P \times 1)}$, $\mathbf{s}$ be K-sparse, with

$$K = |\mathrm{supp}(\mathbf{s})| \stackrel{\text{def}}{=} \|\mathbf{s}\|_0,\ K \ll P$$

- Then $\mathbf{y} = \mathbf{W}_{(M \times P)}\ \mathbf{s}_{(P \times 1)}$, $\mathbf{W} = \mathbf{A}\ \mathbf{D}$ has a K-sparse solution

$$\mathbf{s} = \arg\min_{\boldsymbol{\xi} \in \mathbb{R}^n} \|\boldsymbol{\xi}\|_0 \ \ s.t.\ \ \mathbf{y} = \mathbf{W}\boldsymbol{\xi}$$

- $\|\cdot\|_0$ is *nonconvex* $\Rightarrow$ the previous problem is *hard*.
- Consider the *convex* problem:

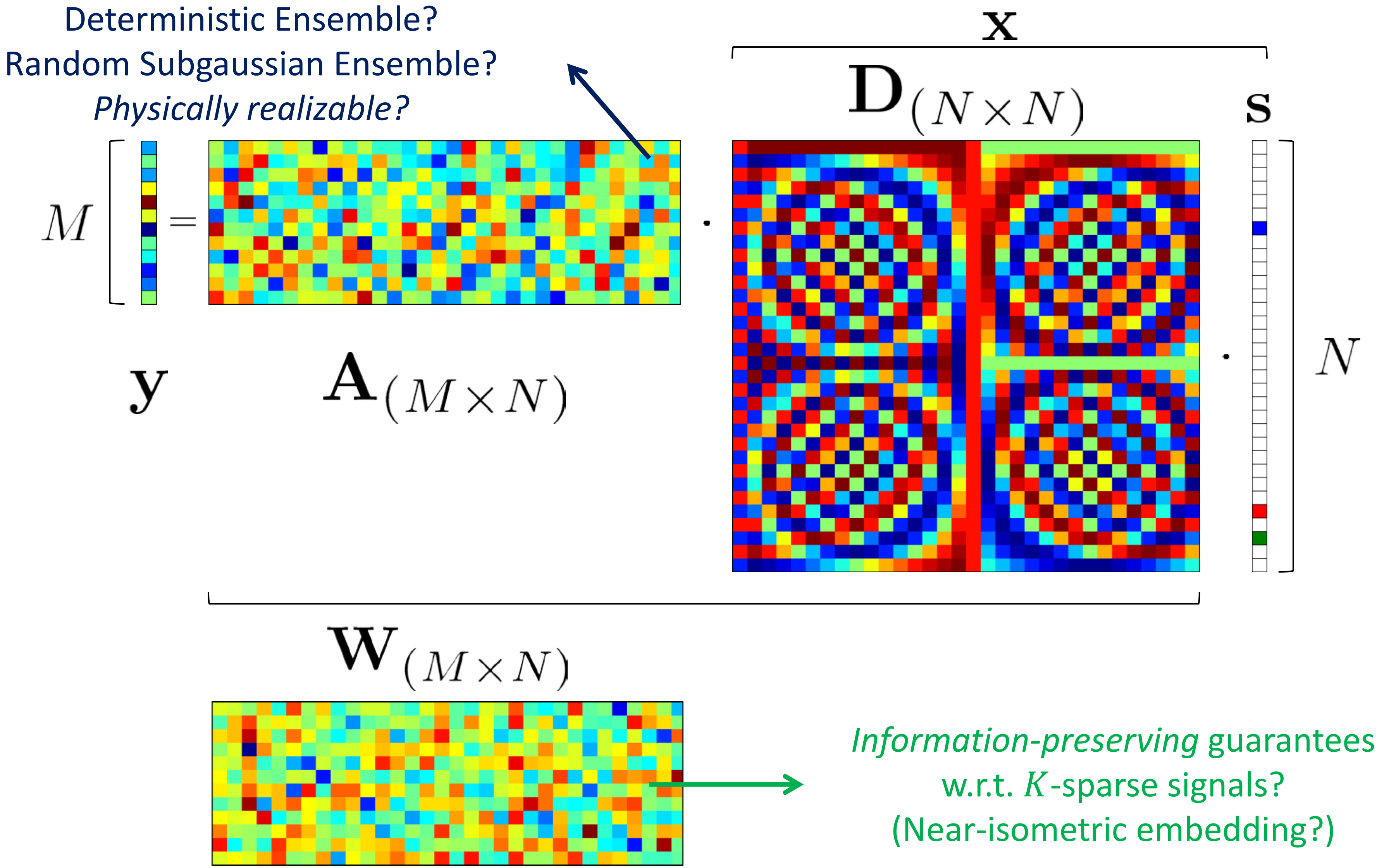
$$\hat{\mathbf{s}} = \arg\min_{\boldsymbol{\xi}\in\mathbb{R}^n} \|\boldsymbol{\xi}\|_1 \;\; s.t. \;\; \mathbf{y} = \mathbf{W}\boldsymbol{\xi}$$

- $\hat{\mathbf{s}} = \mathbf{s}$ if (<u>sufficient</u> condition) $\mathbf{W}$ has the *restricted isometry property* [CT,2005] w.r.t. K-sparse $\mathbf{s}$.
- Under similar hypotheses, *noisy* measurements and *approximately sparse* signals the recovery

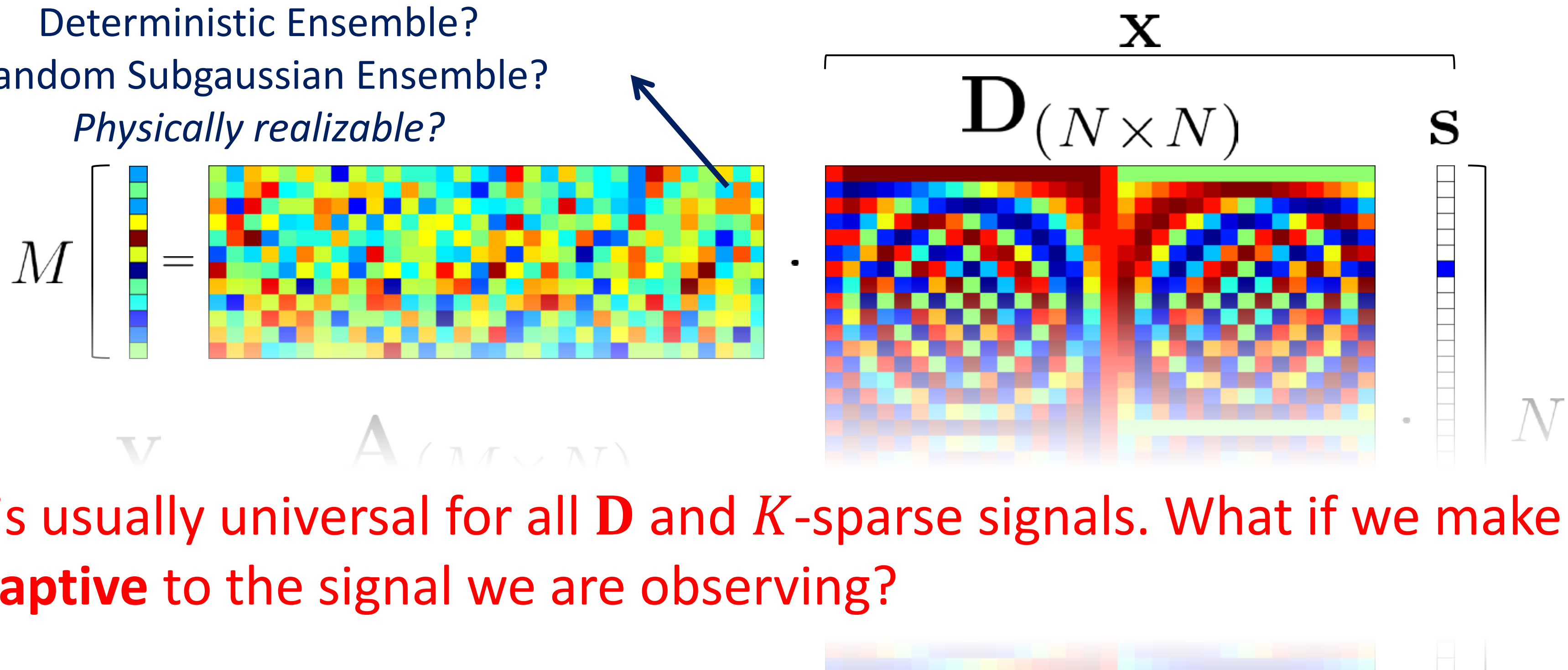
$$\hat{\mathbf{s}} = \arg\min_{\boldsymbol{\xi}\in\mathbb{R}^n} \|\boldsymbol{\xi}\|_1 \;\; s.t. \;\; \|\mathbf{y} - \mathbf{W}\boldsymbol{\xi}\|_2 \leq \|\boldsymbol{\nu}\|_2$$

verifies $\|\hat{\mathbf{s}} - \mathbf{s}\|_2 \leq C_0 \frac{\|\mathbf{s}-\mathbf{s}_K\|_1}{\sqrt{K}} + C_1\|\boldsymbol{\nu}\|_2$ [CRT,2006].

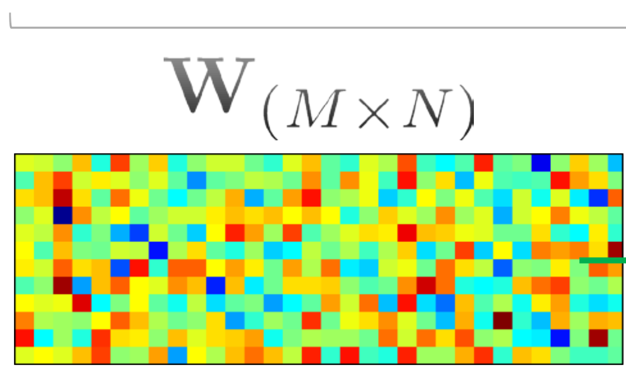
THE SENSING MATRIX DESIGN PROBLEM



Deterministic Ensemble?
Random Subgaussian Ensemble?
Physically realizable?



A is usually universal for all **D** and K-sparse signals. What if we make it **adaptive** to the signal we are observing?



Information-preserving guarantees
w.r.t. K-sparse signals?
(Near-isometric embedding?)

- Let $\mathbf{x}, \mathbf{s}$: random vectors (RV) whose spectral (energy) distribution is **localized**.
- Let $\mathbf{K}_{\mathbf{x}}$: correlation matrix of the input RV,

$$\mathbf{K}_{\mathbf{x}} = \mathbb{E}[\mathbf{x}\mathbf{x}^{\dagger}] = \mathbf{Q}\Lambda_{\mathbf{x}}\mathbf{Q}^{\dagger}$$

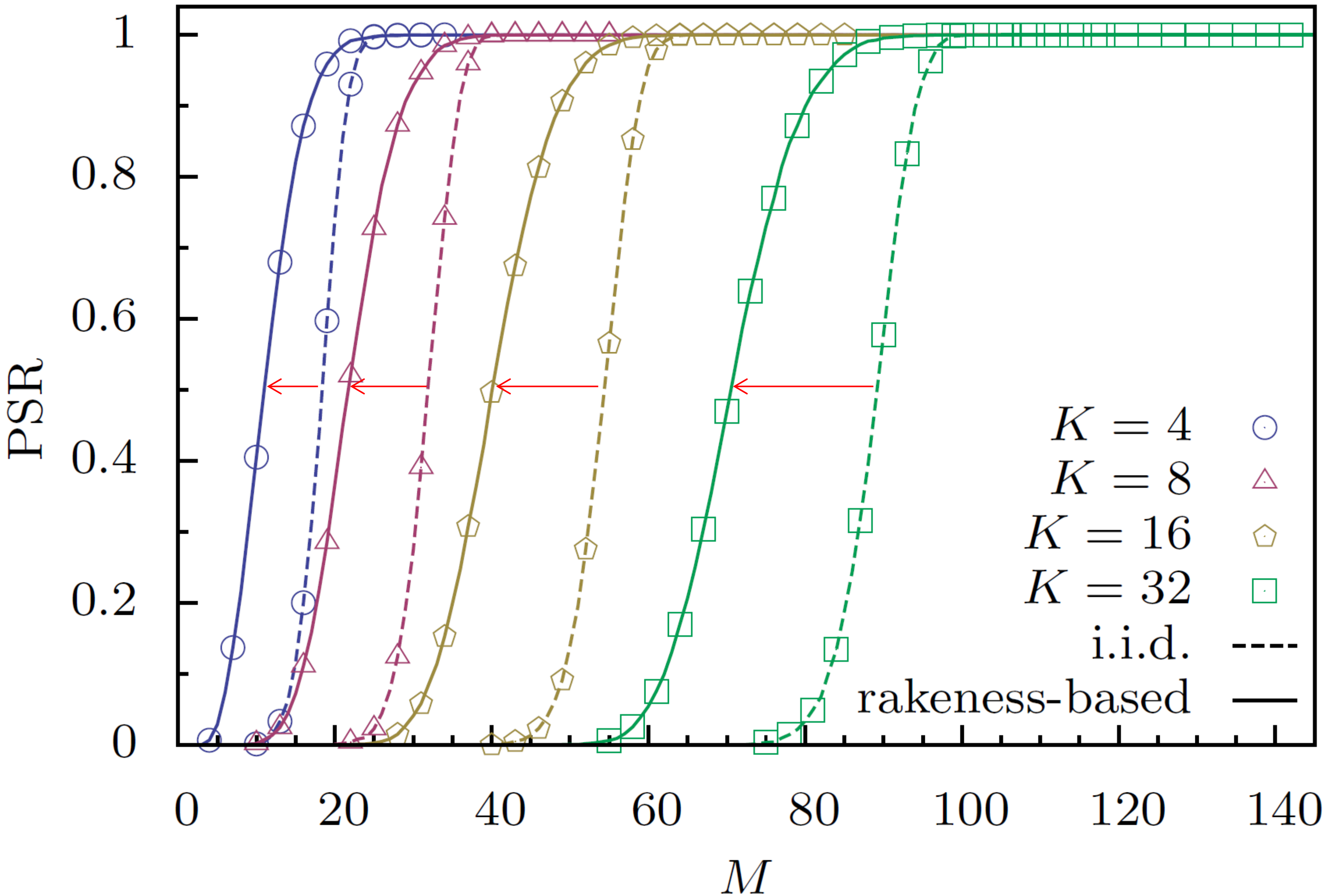
with $\mathbf{Q}$ an orthonormal basis of eigenvectors.

- **Localization** as a deviation from the same-energy white case:

$$\mathcal{L}(\Lambda_{\mathbf{x}}) = \sum_{j=0}^{n-1} \left(\frac{\lambda_{\mathbf{x},j}}{e_{\mathbf{x}}} - \frac{1}{n} \right)^2 \in [0, 1 - 1/n]$$

- A simple measure of **anisotropy** in the partition of $e_{\mathbf{x}} = \mathrm{tr}(\Lambda_{\mathbf{x}})$ along the eigenvectors $\mathbf{q}_j$ / eigenvalues $\lambda_{\mathbf{x},j}$.

- Conjecture:
  - **"Optimal" random projections** for **white** RV (universal, worst-case)
    → *isotropic* subgaussian RV (i.i.d. Bernoulli or Gaussian)
  - **"Optimal" random projections** for **localized** RV (non-universal)
    →*non-isotropic* subgaussian RV that **maximize** the **raked energy**
- Empirical evidence:
  - Assume $\mathbf{x}$ has K-sparse realizations
  - Choose *a posteriori* out of 10^4 i.i.d. random Gaussian measurements the $M < N$ with the **highest energy**
  - Observe the **probability of successful recovery** (PSR) of a sparse signal


1
0.8
0.6
0.4
0.2
0
PSR
0
20
40
60
80
100
120
140
M
K = 4
K = 8
K = 16
K = 32
i.i.d.
rakeness-based

- We define **rakeness** as:

$$\rho(\mathbf{A}_j,\mathbf{x}) = \mathbb{E}_{\mathbf{A}_j,\mathbf{x}}\left[\|y_j\|_2^2\right] = \mathbb{E}_{\mathbf{A}_j,\mathbf{x}}\left[\|\mathbf{A}_j\mathbf{x}\|_2^2\right]$$

i.e. the expected "affinity" of random projection vectors $\mathbf{A}_j$ to the task of collecting the **energy** in $\mathbf{x}$.

- The **maximum rakeness** optimization problem will be:

$$\max \rho(\mathbf{A}_j,\mathbf{x}) = \max \operatorname{tr}\left(\mathbf{K}_{\mathbf{A}_j}\mathbf{K}_{\mathbf{x}}\right) = \max \operatorname{tr}\left(\Lambda_{\mathbf{A}_j}\Lambda_{\mathbf{x}}\right)$$

with $\mathbf{K}_{\mathbf{A}_j} = \mathbf{Q}\,\Lambda_{\mathbf{A}_j}\,\mathbf{Q}^\dagger$ (as proved in [MRS,2012]) and under

  - **Average energy constraint:** $e_{\mathbf{A}_j} = \operatorname{tr}\left(\Lambda_{\mathbf{A_j}}\right) = 1$
  - **Localization constraint**: the random projection vectors must be *less localized* than $\mathbf{x}$ $\Rightarrow$ tuning parameter τ to **balance *isotropicity* against *localization***.

- We define **rakeness** as:

$$\rho(\mathbf{A}_j, \mathbf{x}) = \mathbb{E}_{\mathbf{A}_j,\mathbf{x}}\left[\|y_j\|_2^2\right] = \mathbb{E}_{\mathbf{A}_j,\mathbf{x}}\left[\|\mathbf{A}_j\mathbf{x}\|_2^2\right]$$

i.e. the expected "affinity" of random projection vectors $\mathbf{A}_j$ to the task of collecting the **energy** in $\mathbf{x}$.

- We may let $\mathbf{K}_{\mathbf{A}_j} = \mathbf{Q}\,\Lambda_{\mathbf{A}_j}\,\mathbf{Q}^\dagger$. The **maximum rakeness** optimization problem will be:

$$\max_{\Lambda_{\mathbf{A}_j}} \operatorname{tr}\left(\Lambda_{\mathbf{A}_j}\Lambda_{\mathbf{x}}\right)$$

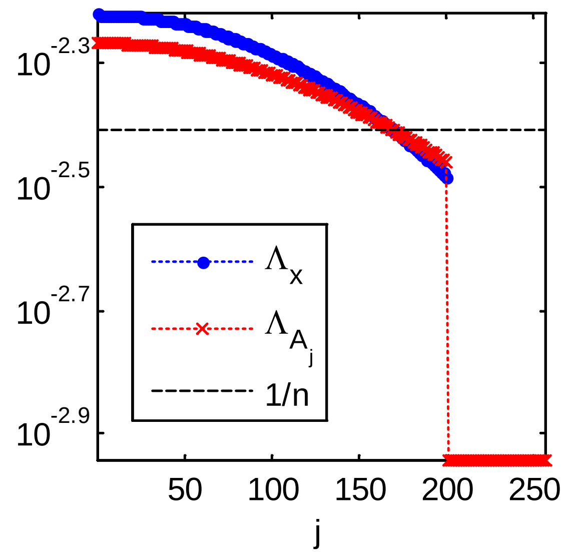
s.t.

$$e_{\mathbf{A}_j} = 1$$

$$\mathcal{L}(\Lambda_{\mathbf{A}_j}) \leq \tau\mathcal{L}(\Lambda_{\mathbf{x}}), \tau \in [0, 1]$$

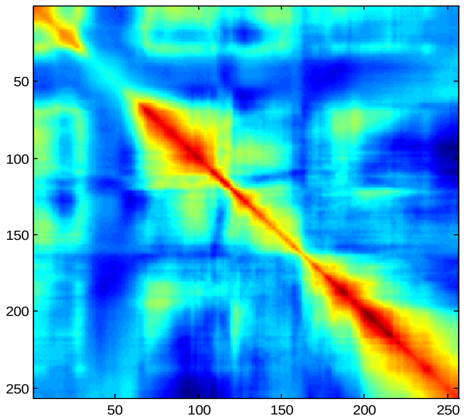
Average Energy Constraint

Localization Constraint



A Rakeness-based Design Flow

$\mathbf{K}_{\mathbf{x}}$



Analysis →

$$\mathbf{K}_{\mathbf{x}} = \mathbb{E}\left(\mathbf{x}\mathbf{x}^{\dagger}\right) = Q\Lambda_{\mathbf{x}}Q^{\dagger}$$

$$\Rightarrow \mathcal{L}(\Lambda_{\mathbf{x}})$$

↓

$$\max_{\mathbf{A}_j} \rho(\mathbf{A}_j, \mathbf{x})$$

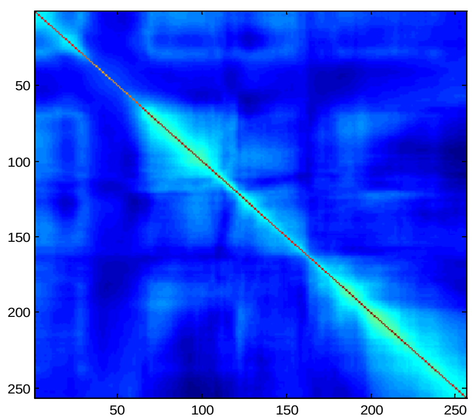
s.t.

$$e_{\mathbf{A}_j} = 1$$

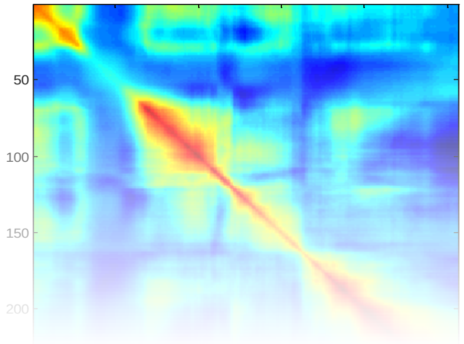
$$\mathcal{L}(\Lambda_{\mathbf{A}_j}) \leq \tau \mathcal{L}(\Lambda_{\mathbf{x}}), \tau \in [0, 1]$$

← Synthesis

$\mathbf{K}_{\mathbf{A}_j}$



$\mathbf{K}_\mathbf{x}$



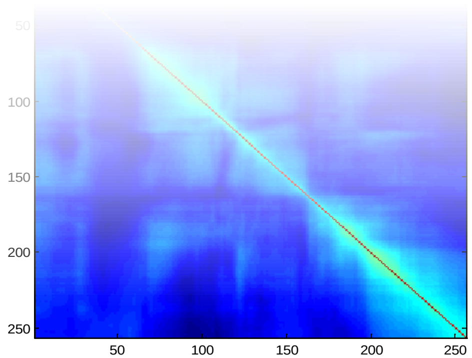
Analysis

$$\mathbf{K}_\mathbf{x} = \mathbb{E}\left(\mathbf{x}\mathbf{x}^\dagger\right) = Q\Lambda_\mathbf{x}Q^\dagger$$

$$\Rightarrow \mathcal{L}(\Lambda_\mathbf{x})$$

For any $\tau \in [0,1]$ so that $\mathbf{K}_{\mathbf{A}_j}$ is **positive definite** the projections will allocate more energy along the **principal components** of $\mathbf{x}$, while allocating a non-null fraction of it along the others.

$\mathbf{K}_{\mathbf{A}_j}$



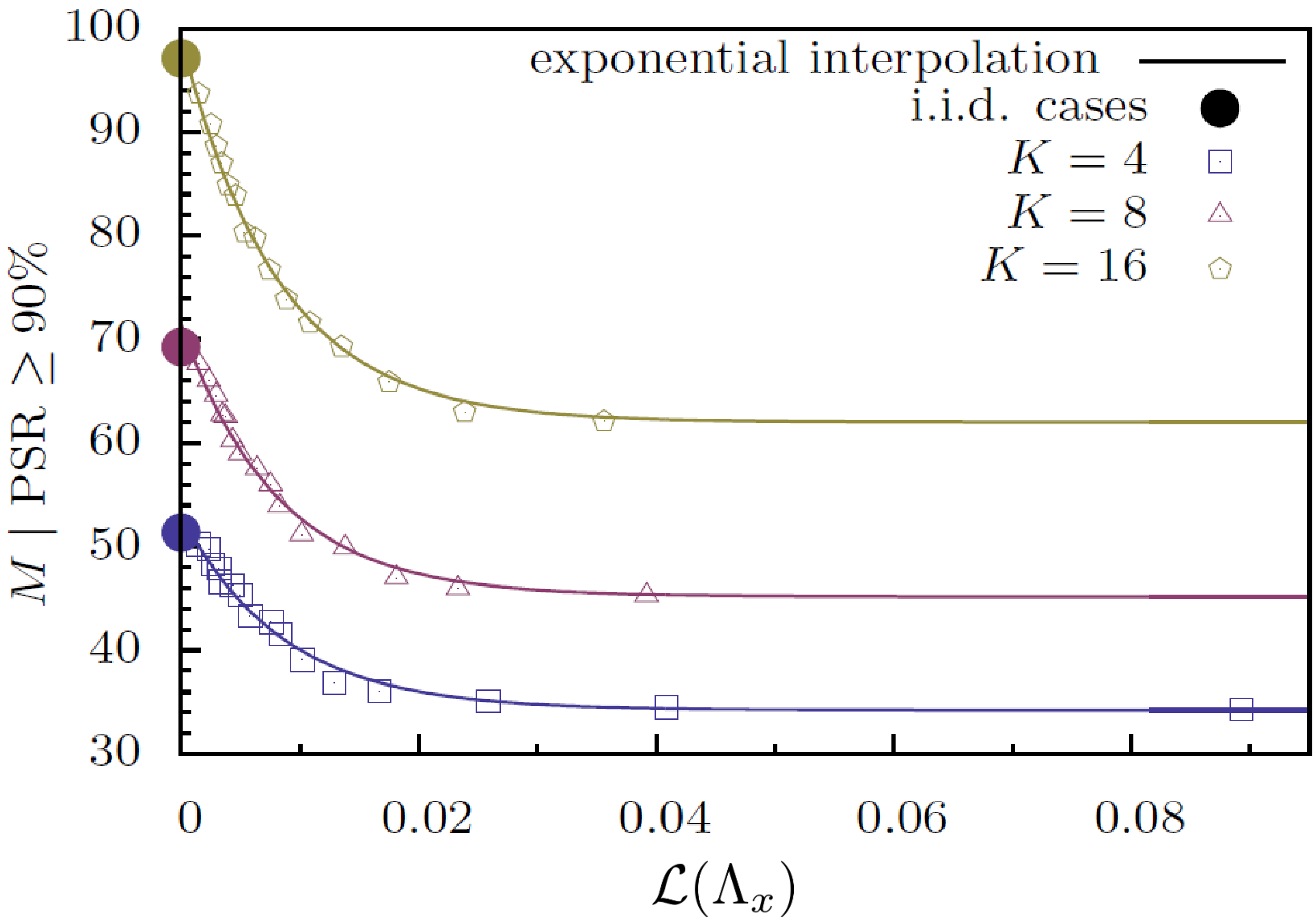
Synthesis

s.t.

$$e_{\mathbf{A}_j} = 1$$

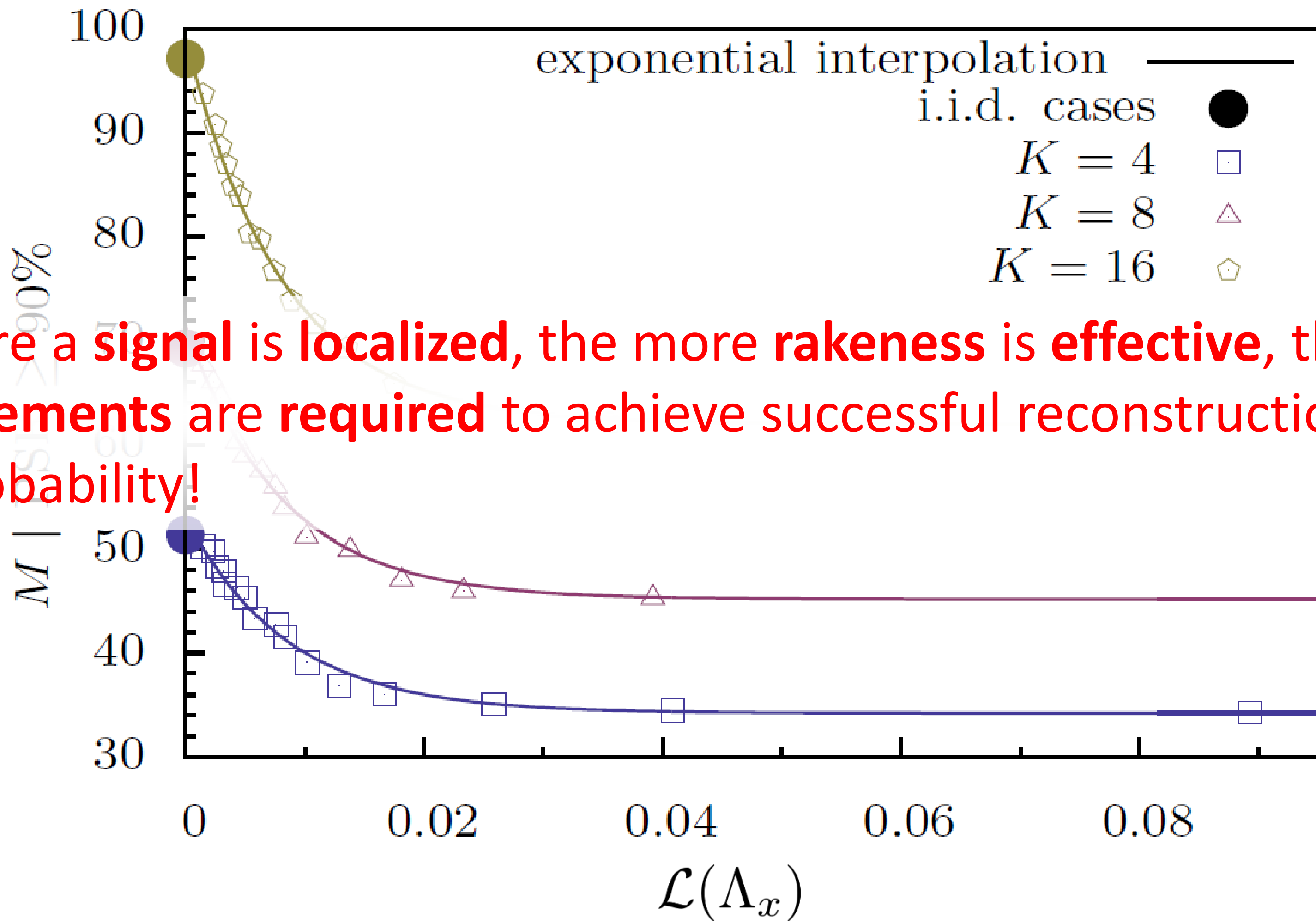
$$\mathcal{L}(\Lambda_{\mathbf{A}_j}) \leq \tau\mathcal{L}(\Lambda_\mathbf{x}), \tau \in [0,1]$$

- $N = 256$, K-sparse signals with $\mathbf{K_x}$ matching a given **localization** and 30 dB superimposed AWGN noise. $\mathbf{K}_{\mathbf{A}_j}$ designed with $\tau = 0.5$.



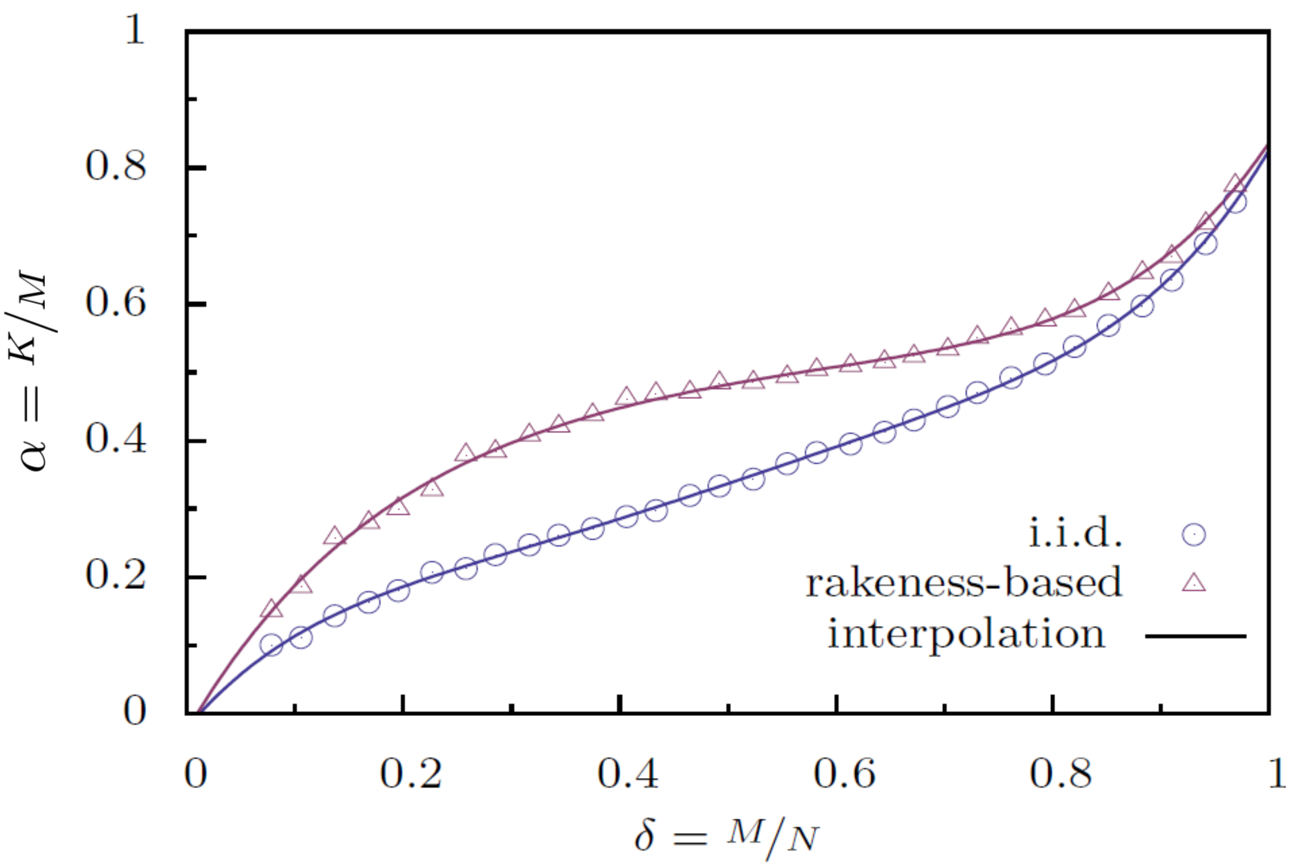
A Rakeness-based Design Flow: Example

- $N = 256$, K-sparse signals with $\mathbf{K_x}$ matching a given **localization** and 30 dB superimposed AWGN noise. $\mathbf{K}_{\mathbf{A}_j}$ designed with $\tau = 0.5$.



The more a **signal** is **localized**, the more **rakeness** is **effective**, the **less measurements** are **required** to achieve successful reconstruction with high probability!

- Rakeness-based CS raises the Donoho – Tanner phase transition curve [DT,2009] w.r.t. ℓ_1 minimization (PSR $>$ 90% as $f(\frac{M}{N},\frac{K}{M})$) **w.r.t. localized signals** (in this example $L(\Lambda_{\mathbf{x}}) = 0.03, \tau = 0.5$).



- Synthesis with **Gaussian** random matrices: *Easy*

$$\mathbf{A}_j^w \sim \mathcal{N}(0, \frac{1}{n}\mathbf{I}_n) \in \mathbb{R}^n$$

$$\mathbf{K}_{\mathbf{A}_j} = \mathbb{E}\left(\mathbf{A}_j^\dagger \mathbf{A}_j\right) = Q\Lambda_{\mathbf{A}_j}Q^\dagger \quad \Rightarrow \quad \mathbf{A}_j \sim \mathcal{N}(0, \mathbf{K}_{\mathbf{A}_j}) \in \mathbb{R}^n$$

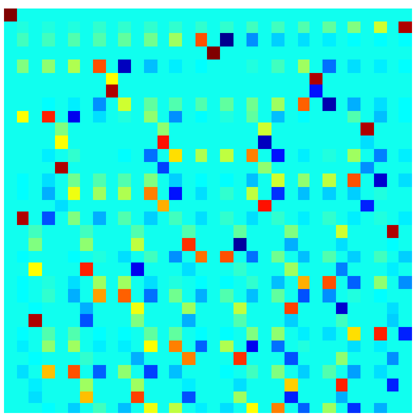
$$\mathbf{A}_j = \mathbf{A}_j^w \sqrt{\mathbf{K}_{\mathbf{A}_j}}$$

- Synthesis with **Bernoulli** random matrices: *Non-trivial*
  - Linear Probability Feedback (stationary case) [MRS,2012]
  - Quadratic Integer Programming (general case, hard problem) [CFLMRS,2014]
  - The *Arcsin Law* (general case, non-general applicability) [VM,1966 and CFLMRS,2014]

SSIGPRO

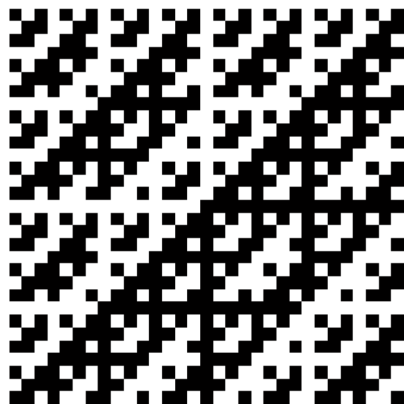
- [CT,2005] E.J. Candes, and T. Tao. "Decoding by linear programming."*Information Theory, IEEE Transactions on* 51.12 (2005): 4203-4215.
- [CRT,2006] E.J. Candes, J. K. Romberg, and T. Tao. "Stable signal recovery from incomplete and inaccurate measurements." *Communications on pure and applied mathematics* 59.8 (2006): 1207-1223.
- [DT,2009] D. Donoho, and J. Tanner. "Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367.1906 (2009): 4273-4293.
- [RRS,2010] Ranieri, J.; Rovatti, R.; Setti, G., "Compressive sensing of localized signals: Application to Analog-to-Information conversion," *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on* , vol., no., pp.3513,3516, May 30 2010-June 2 2010
- [MRS,2012] M. Mangia, R. Rovatti, and G. Setti. "Rakeness in the design of Analog-to-Information Conversion of Sparse and Localized Signals." *Circuits and Systems I: Regular Papers, IEEE Transactions on* 59.5 (2012): 1001-1014. **Guillemin-Cauer Award**.
- [CMPRS, 2013] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A rakeness-based design flow for analog-to-information conversion by compressive sensing." *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE, 2013.
- [CFLMRS, 2014] Caprara, A. ; Furini, F.; Lodi, A.; Mangia, M.; Rovatti, R.; Setti, G., "Generation of Antipodal Random Vectors With Prescribed Non-Stationary 2-nd Order Statistics," *Signal Processing, IEEE Transactions on* , 62.6 (2014): 1603-1612.
- [VM,1966] J.H. Van Vleck, and D. Middleton. "The spectrum of clipped noise." *Proceedings of the IEEE* 54.1 (1966): 2-19.

- In some cases $\mathbf{A}$ is designed from a finite set *(design space)* of *physically realizable* sensing vectors (e.g. an orthonormal basis).
- Examples: partial Hadamard, partial Fourier (e.g. MRI) matrix ensembles
- Problems:
  - $\mathbf{A}$ *coherent* w.r.t. $\mathbf{D}$ (correlated columns in $\mathbf{W}$)
  - Less degrees of freedom to apply rakeness-based designs to localized signals



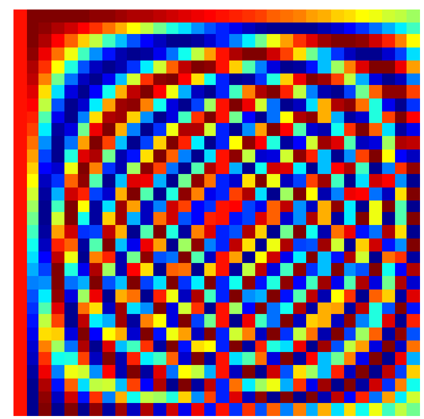
$\mathbf{W}$

$=$



$\mathbf{A}$

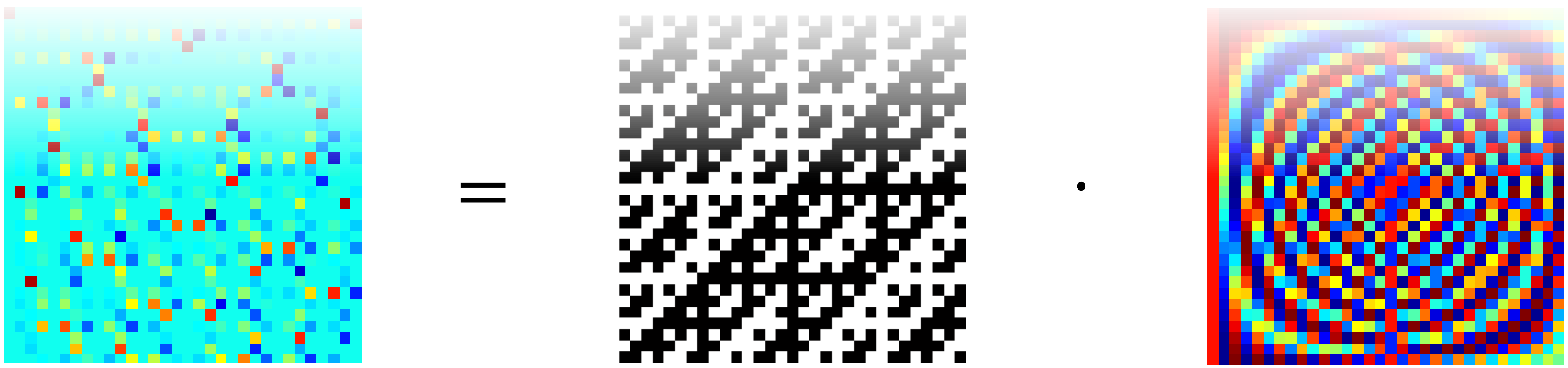
$\cdot$



$\mathbf{D}$

- In some cases $\mathbf{A}$ is designed from a finite set *(design space)* of *physically realizable* sensing vectors (e.g. an orthonormal basis).
- Examples: partial Hadamard, partial Fourier (e.g. MRI) matrix ensembles
- Problems:
  - $\mathbf{A}$ *coherent* w.r.t. $\mathbf{D}$ (correlated columns in $\mathbf{W}$)

In such a constrained design space, is there an **adaptive** method to fine-tune the sensing matrix to **localized signals**?



$\mathbf{W}$ = $\mathbf{A}$ · $\mathbf{D}$

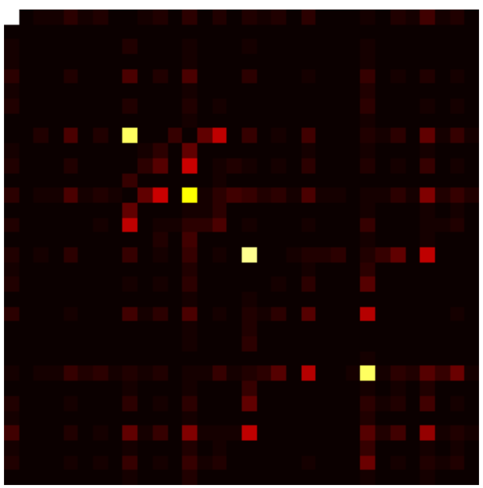
Localized Signals and Correlated Measurements

- Localized signals generally imply correlated measurements

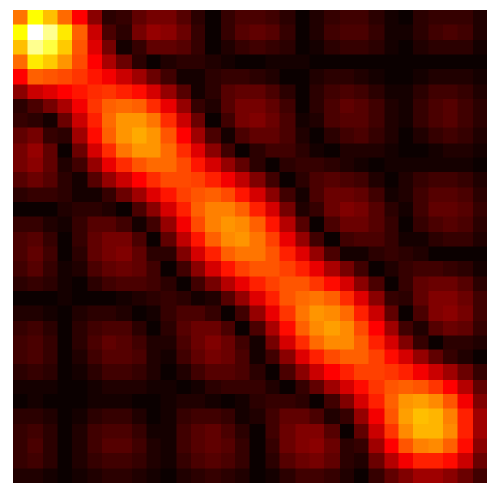
$$\mathbf{K}_{\mathbf{y}} = \mathbf{A}\mathbf{K}_{\mathbf{x}}\mathbf{A}^{\dagger} \qquad \mathbf{K}_{\mathbf{y}_T} = \mathbf{A}_T\mathbf{K}_{\mathbf{x}}\mathbf{A}_T^{\dagger}, T \subset [1, n]$$

- <u>Note</u>: here $\mathbf{A}$ is deterministic (e.g. Hadamard matrix $\mathbf{H}_n, n = 2^q, q \in \mathbb{N}$), T is a randomly chosen subset of basis vectors in $\mathbf{A}_T$
- Which is the subset $T^{\star}$ with cardinality $m < n$ carrying *maximum information* w.r.t. the others?
- This is very close to an *experimental design* problem

$\mathbf{K}_{\mathbf{y}}$



$\mathbf{K}_{\mathbf{x}}$



- Assume for now the measurements are correlated and Gaussian, then the *differential entropy*

$$h(\mathbf{y}_T) \stackrel{\text{def}}{=} -\int_{\theta \in \mathbb{R}^m} f_{\mathbf{y}_T}(\theta) \log f_{\mathbf{y}_T}(\theta) d\theta$$

$$h(\mathbf{y}_T) = \frac{1}{2} \log(2\pi e)^m \det \mathbf{K}_{\mathbf{y}_T} \leq \frac{1}{2} \log \left(2\pi e \frac{\operatorname{tr} \mathbf{K}_{\mathbf{y}_T}}{m} \right)^m$$

E.T. Jaynes, ca. 1982

- With this information measure,

$$T^\star = \underset{T \subset [0, n-1]}{\arg\max} \; h(\mathbf{y}_T) \; s.t. \; |T| = m$$

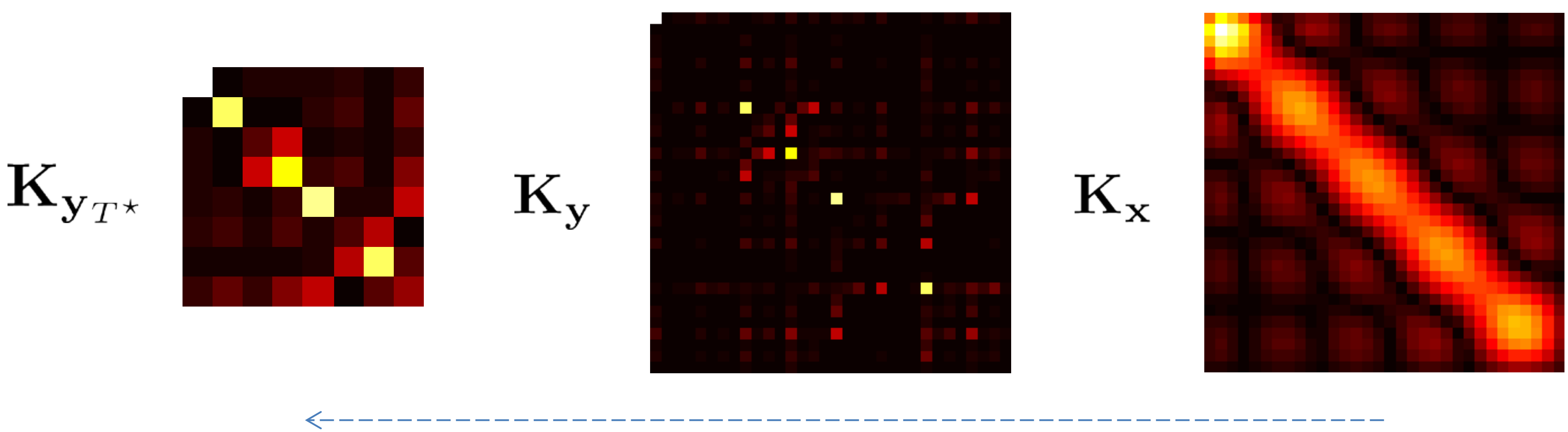
$$= \underset{T \subset [0, n-1]}{\arg\max} \; \log \det \mathbf{K}_{\mathbf{y}_T} \; s.t. \; |T| = m$$

and we form $\mathbf{y}_{T^\star} = \mathbf{A}_{T^\star}\mathbf{x}$ with the rows of $\mathbf{A}$ selected by $T^\star$.

- This is also known as *D-optimal design* or MaxDet w.r.t. $\mathbf{K}_{\mathbf{y}}$.

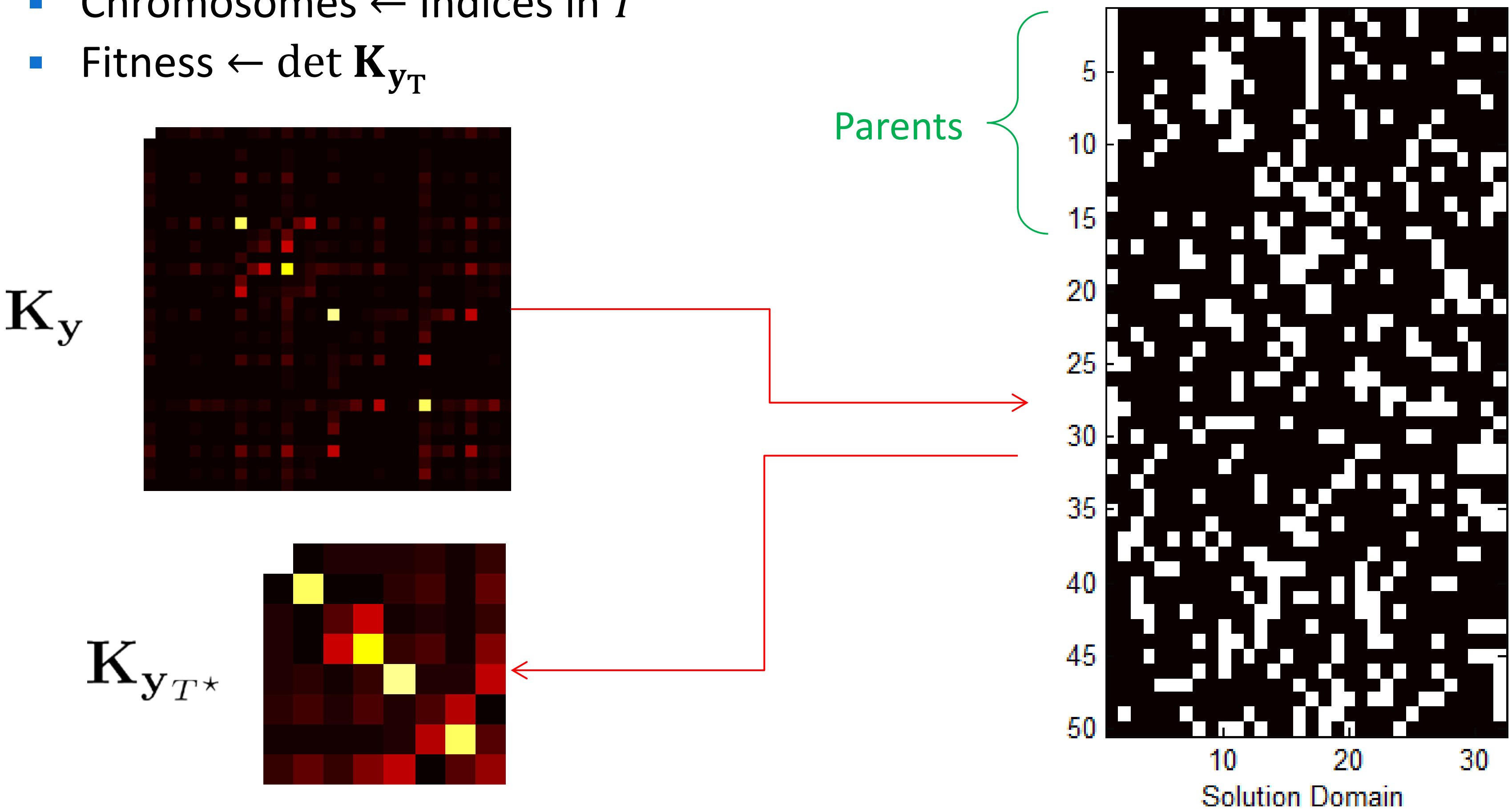
Rationale and Gaussian Approximation

- When $\mathbf{A}$ is deterministic, $\mathbf{y}$ depends on $f_{\mathbf{x}}(x)$.
  - $f_{\mathbf{x}}(x)$ Gaussian $\Rightarrow$ maximum entropy
  - $f_{\mathbf{x}}(x)$ *approximately* Gaussian $\Rightarrow$ near-maximum entropy
  - $f_{\mathbf{x}}(x)$ *non*-Gaussian $\Rightarrow$ $\mathbf{y}_{T^\star}$ is the measurement set with *least linear predictability* (each measurement has maximum prediction error w.r.t. the remaining $m-1$)

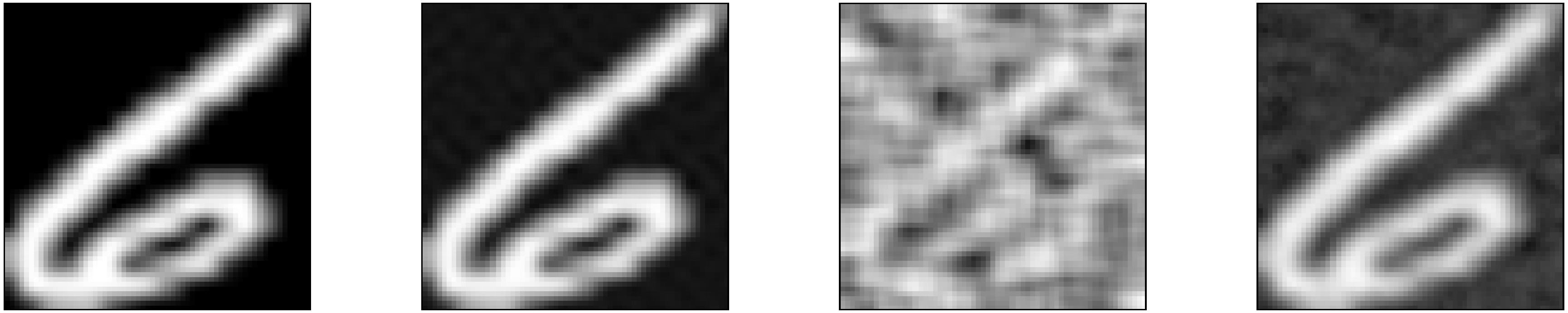


A Maximum Determinant Heuristic

- Problem: Maximum Determinant Principal Submatrix of $\mathbf{K_y}$ (hard)
- Exact solution: Branch-and-Bound (Ko et al., 1995)
- Heuristic (high-entropy) solution by a simple evolutionary algorithm:
  - Chromosomes $\leftarrow$ Indices in T
  - Fitness $\leftarrow \det \mathbf{K}_{\mathbf{y}_T}$



Experimental Results: Handwritten Digits



(b) Original image, $n = 4096$ pixel

(c) $T^\star$ PHE, RSNR = 36.82 dB

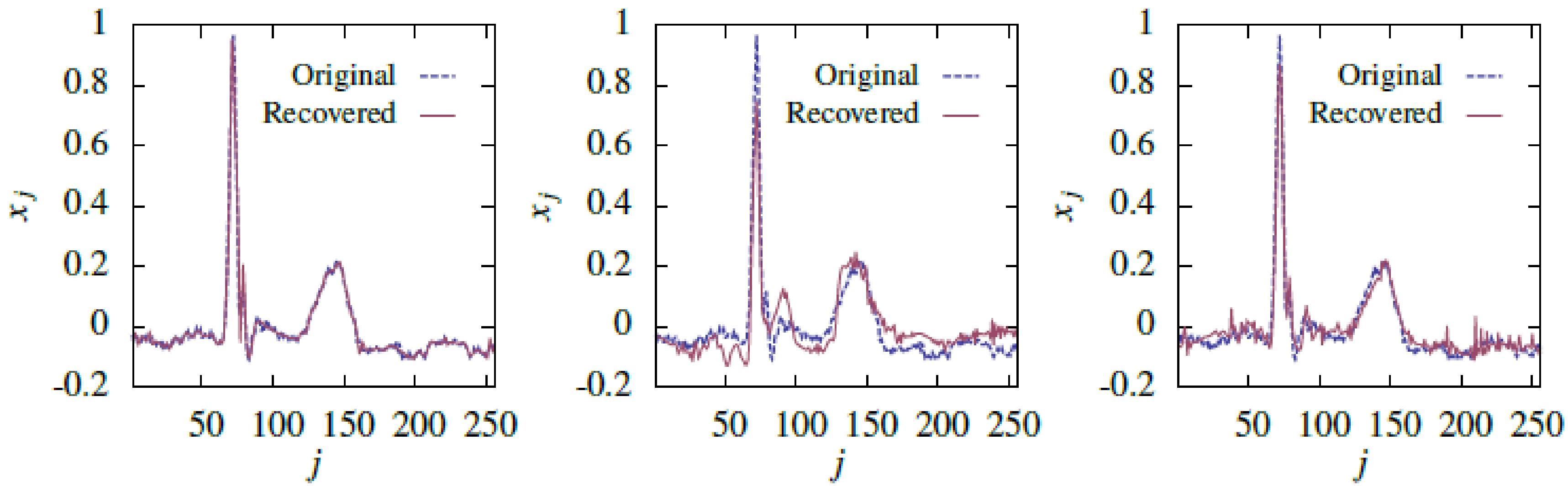
(d) Random PHE, RSNR = 1.88 dB

(e) RBE, RSNR = 19.85 dB

m	MaxDet pool PHE	Random PHE	RBE
1024 ($^n/_4$)	**36.57**	1.51	20.63
1365 ($\lfloor ^n/_3 \rfloor$)	**39.63**	2.89	26.08

- Average RSNR (dB) over 20 sample images, 25 MaxDet pool PHE, 25 Random PHE and 50 RBE sensing matrices
- The dataset is approximately sparse on the Daubechies-4 wavelet basis.

Experimental Results: ECG Tracks



(a) $T^\star$ PHE, RSNR = 15.94 dB

(b) Random PHE, RSNR = 7.40 dB

(c) RBE, RSNR = 12.85 dB

m	MaxDet pool PHE	Random PHE	RBE
64 ($n/4$)	**15.12**	2.68	6.94
85 ($\lfloor n/3 \rfloor$)	**17.20**	3.73	11.11

- Average RSNR (dB) over 50 sample ECG tracks, 25 MaxDet pool PHE, 25 Random PHE and 50 RBE sensing matrices.
- The dataset is approximately sparse on the Coiflet-3 wavelet basis.

- **Non-adaptive sensing strategies** are general, but *underperforming* if more signal-domain priors are available
- **Structured sparsity** priors are commonly used during signal recovery (decoding): optimally designed measurements (encoding) could further improve performances
- **Adaptive sensing strategies** leverage on such priors, although (often) lacking rigorous signal recovery guarantees
  - *Maximum-Energy* Measurements from Correlated Random Matrix Ensembles
  - *Maximum-Entropy* Measurements from Deterministic Matrix Ensembles
  - Many other adaptive designs exist (see bibliography)

- V. Cambareri, R. Rovatti, and G. Setti, "Maximum Entropy Hadamard Sensing of Sparse and Localized Signals" *ICASSP 2014,* Florence, Italy, pp. 2376-2380
- A. Caprara; F. Furini; A. Lodi; M. Mangia; R. Rovatti; G. Setti, "Generation of Antipodal Random Vectors With Prescribed Non-Stationary 2-nd Order Statistics," *Signal Processing, IEEE Transactions on* , 62.6 (2014): 1603-1612.
- V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A rakeness-based design flow for analog-to-information conversion by compressive sensing." *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE, 2013.
- M. Mangia, R. Rovatti, and G. Setti. "Rakeness in the design of Analog-to-Information Conversion of Sparse and Localized Signals." *Circuits and Systems I: Regular Papers, IEEE Transactions on* 59.5 (2012): 1001-1014. **Guillemin-Cauer Award**.
- Ranieri, J.; Rovatti, R.; Setti, G., "Compressive sensing of localized signals: Application to Analog-to-Information conversion," *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on* , vol., no., pp.3513,3516, May 30 2010-June 2 2010

- Wang, Zhongmin, and Gonzalo R. Arce. "Variable density compressed image sampling." *Image Processing, IEEE Transactions on* 19.1 (2010): 264-270.
- Seeger, Matthias W., and Hannes Nickisch. "Compressed sensing and Bayesian experimental design." *Proceedings of the 25th international conference on Machine learning*. ACM, 2008.
- Ji, Shihao, Ya Xue, and Lawrence Carin. "Bayesian compressive sensing."*Signal Processing, IEEE Transactions on* 56.6 (2008): 2346-2356.
- Carson, William R., et al. "Communications-inspired projection design with application to compressive sensing." *SIAM Journal on Imaging Sciences* 5.4 (2012): 1185-1212.
- Duarte-Carvajalino, J.M., et al. "Task-driven adaptive statistical compressive sensing of Gaussian mixture models." *Signal Processing, IEEE Transactions on* 61.3 (2013): 585-600.
- Chen, W., M. R. D. Rodrigues, and I. J. Wassell. "Projections design for statistical compressive sensing: A tight frame based approach." *IEEE Transactions on Signal Processing* 61.4 (2013): 2016-2029.
- Puy, Gilles, Pierre Vandergheynst, and Yves Wiaux. "On variable density compressive sampling." *Signal Processing Letters, IEEE* 18.10 (2011): 595-598.

- Jaynes, Edwin T. "On the rationale of maximum-entropy methods." *Proceedings of the IEEE* 70.9 (1982): 939-952.
- Jaynes, Edwin T. "Information theory and statistical mechanics." *Physical review* 106.4 (1957): 620.
- Jaynes, Edwin T. "Information theory and statistical mechanics. II." *Physical review* 108.2 (1957): 171.
- Ko, Chun-Wa, Jon Lee, and Maurice Queyranne. "An exact algorithm for maximum entropy sampling." *Operations Research* 43.4 (1995): 684-691.
- Donoho, David L., et al. "Maximum entropy and the nearly black object."*Journal of the Royal Statistical Society. Series B (Methodological)* (1992): 41-81.
- Elad, Michael. "Optimized projections for compressed sensing." *Signal Processing, IEEE Transactions on* 55.12 (2007): 5695-5702.

Thank you for your attention.

Questions?