Quantised Random Embeddings for Dimensionality Reduction

V. Cambareri (joint work w. Prof. L. Jacques), ISP Group, ICTEAM/ELEN

Quantised Random Embeddings

$$\boldsymbol{y} = \mathsf{A}(\boldsymbol{x}) := \mathcal{Q}_\delta(\boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{\xi})$$

where:

- Dataset: $\boldsymbol{x} \in \mathcal{K} \subset \mathbb{R}^n \longrightarrow$ Signatures: $\boldsymbol{y} \in \mathsf{A}(\mathcal{K}) \subset \delta\mathbb{Z}^m$
- Uniform Scalar Quantiser: $\mathcal{Q}_\delta := \delta\lfloor \frac{\cdot}{\delta} \rceil$
- Gaussian (not necessary!) Random Projections: $\boldsymbol{\Phi} \sim \mathcal{N}^{m\times n}(0,1)$
- Uniform Dithering: $\boldsymbol{\xi} \sim \mathcal{U}^{m\times n}([-\frac{\delta}{2}, \frac{\delta}{2}])$

Quasi-Isometry Property

If $\boldsymbol{\Phi}$ is a ε-stable embedding (ℓ_2/ℓ_1 sense) w.r.t. $\mathcal{K}$,

$$(1-\varepsilon)\|\boldsymbol{x}'-\boldsymbol{x}''\|_2 - \Delta \leq \frac{1}{m}\|\boldsymbol{y}'-\boldsymbol{y}''\|_1 \leq (1+\varepsilon)\|\boldsymbol{x}'-\boldsymbol{x}''\|_2 + \Delta$$

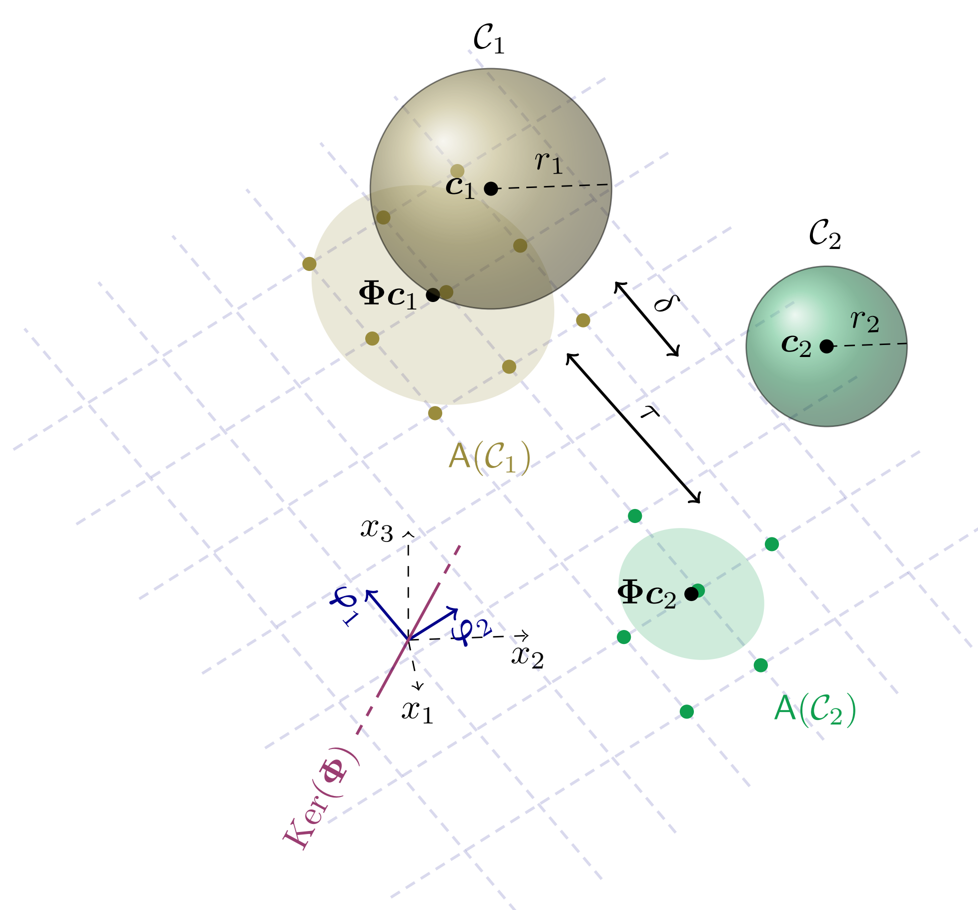
w.h.p. for $\Delta, \varepsilon > 0$ that decay with $m(\uparrow)$, Complexity($\mathcal{K}$) ($\downarrow$), δ ($\downarrow$).

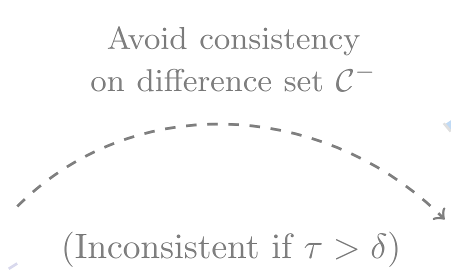
"Rare Eclipses" and Class Separability

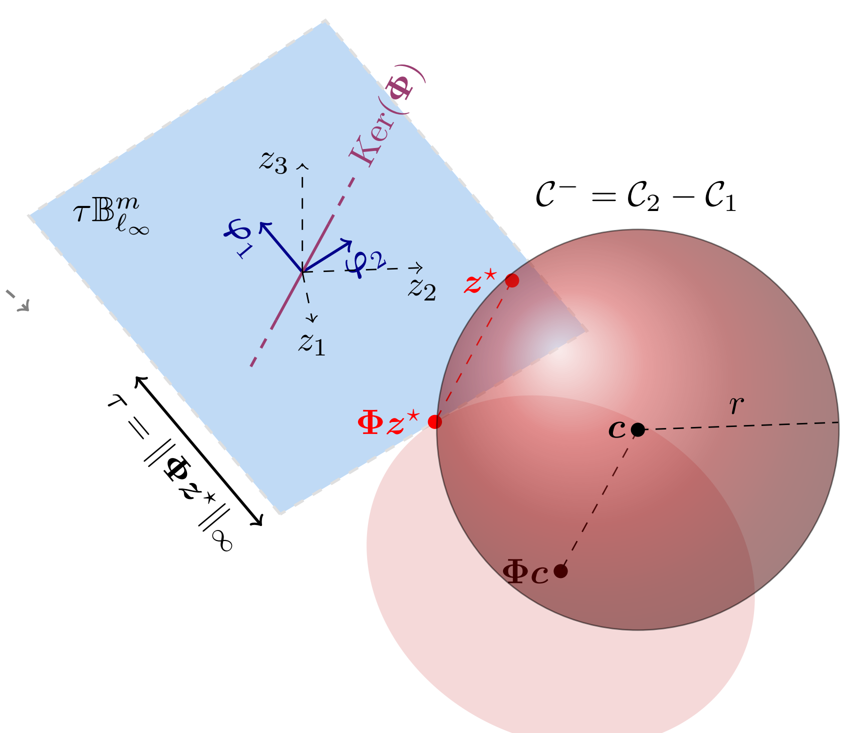
If $\boldsymbol{x}' \in \mathcal{C}' \subset \mathcal{K}, \boldsymbol{x}'' \in \mathcal{C}''\mathcal{K}, \mathcal{C}' \cap \mathcal{C}'' = \emptyset$ disjoint convex sets, then

$$p_\delta := \mathbb{P}[\mathsf{A}(\mathcal{C}') \cap \mathsf{A}(\mathcal{C}') = \emptyset] \geq 1 - \eta$$

with η that decays with m ($\uparrow$), Complexity($\mathcal{K}$) ($\downarrow$), δ ($\downarrow$).



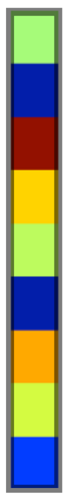




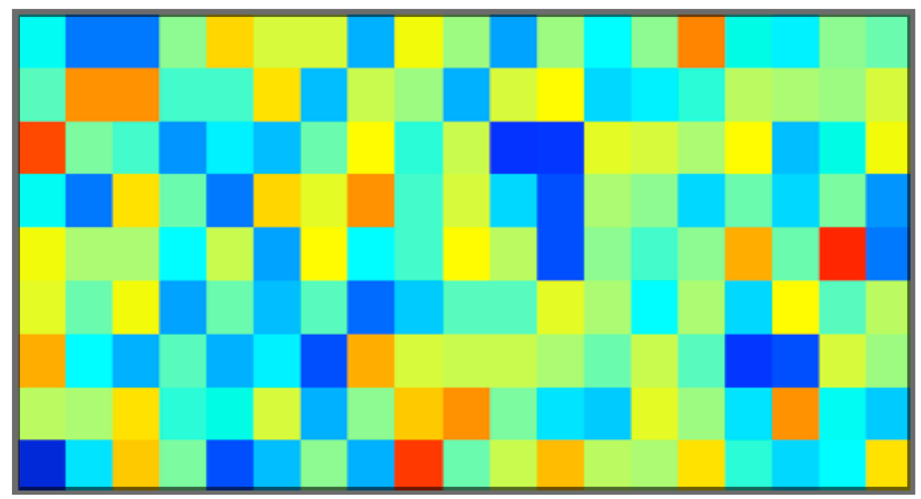
Code here

Linear Compressed Sensing

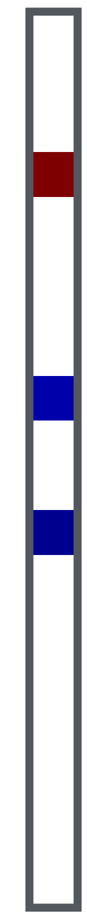
m "queries" | Sensing Operator | Signal



$\simeq$
noise,
model errors



$\cdot$



$\boldsymbol{y}$

$\boldsymbol{A} \in \mathbb{R}^{m \times n}$ (*e.g.*, (sub-)Gaussian)

$\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^n$

Linear Sensing Model:

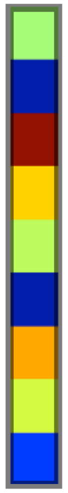
$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \{\boldsymbol{a}_i^* \boldsymbol{x}\}_{i \in [m]}$$

"Low-complexity" *priors* (*e.g.*, sparse, low-rank)

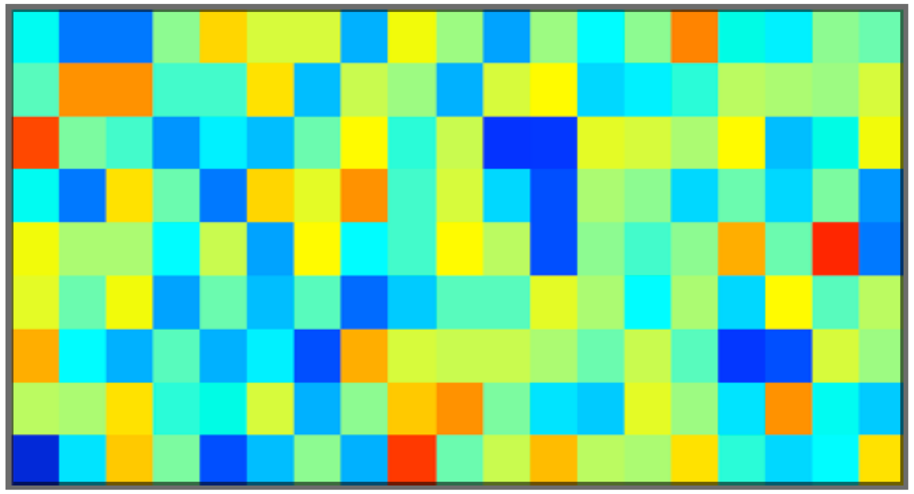
$\mathcal{X} = \Sigma_k^n := \{\boldsymbol{x} \in \mathbb{R}^n : |\mathrm{supp}(\boldsymbol{x})| \leq k\}$

Linear Compressed Sensing

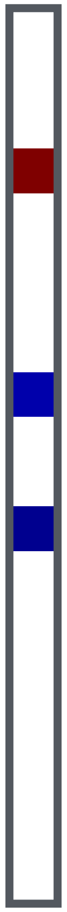
m "queries" Sensing Operator Signal



$\simeq$ noise, model errors



$\cdot$



$\boldsymbol{y}$

$\boldsymbol{A} \in \mathbb{R}^{m \times n}$ (*e.g.*, (sub-)Gaussian)

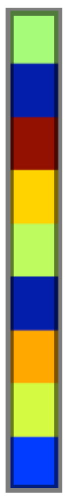
Signal Recovery Algorithm:

$$\hat{\boldsymbol{x}} := \operatorname*{argmin}_{\boldsymbol{\xi} \in \mathbb{R}^n} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\xi}\|_2^2 \text{ s.t. } \boldsymbol{\xi} \in \mathcal{X}$$

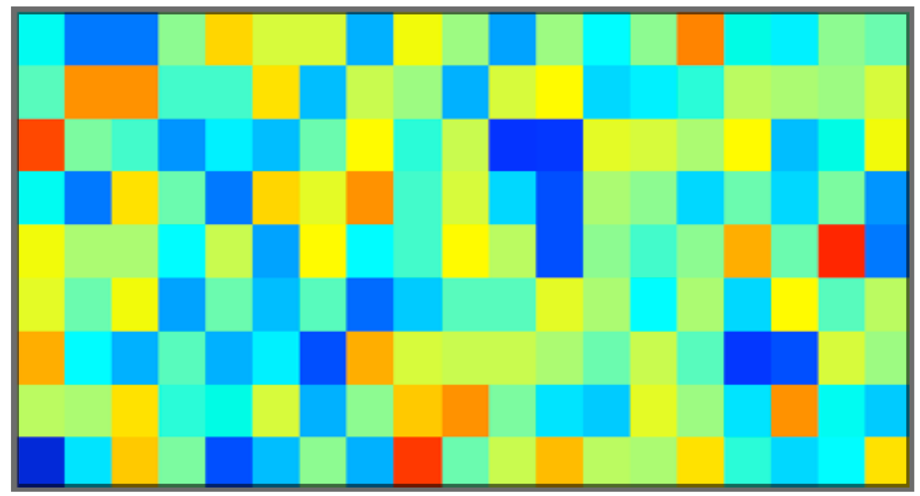
$\simeq$ $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^n$

Linear Compressed Sensing

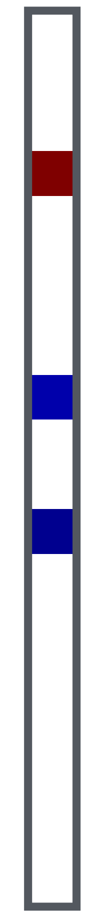
m "queries" | Sensing Operator | Signal



$\simeq$
noise,
model errors



$\cdot$



$\boldsymbol{y}$

$\boldsymbol{A} \in \mathbb{R}^{m \times n}$ (*e.g.*, (sub-)Gaussian)

Signal Recovery Algorithm:

$$\hat{\boldsymbol{x}} := \operatorname*{argmin}_{\boldsymbol{\xi} \in \mathbb{R}^n} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\xi}\|_2^2 \text{ s.t. } \boldsymbol{\xi} \in \mathcal{X} \quad \simeq \quad \boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^n$$

or, in fact, a convexification of $\mathcal{X} \Rightarrow$ Convex problem!