Kronecker Product Matrices for Compressive Sensing

Marco F. Duarte  Richard Baraniuk
Compressed Sensing: Using signal structure, dimensionality reduction, approximation algorithms to acquire data @ sampling rate below Nyquist
Real-World Signals

- **Compressible** signal: sorted coordinates decay rapidly to zero well-approximated by a $K$-sparse signal (by thresholding or sparse approx.)

- model: weak $\ell_p$-ball: $|x_i| < Si^{-1/p}$

$$\sigma_K(x) := \|x - x_K\|_2 \leq (ps)^{-1/2} SK^{-s}$$

$|x_i|$ power-law decay

$$s = \frac{1}{p} - \frac{1}{2}$$
Real-World Signals

- **Compressible** signal: sorted coordinates decay rapidly to zero
  well-approximated by a $K$-sparse signal
  (by thresholding or sparse approx.)

  - model: weak $\ell_p$-ball:
    \[ |x_i| < Si^{-1/p} \]
    \[
    \sigma_K(x) := \|x - x_K\|_2 \leq (ps)^{-1/2} SK^{-s}
    \]
    \[ s = \frac{1}{p} - \frac{1}{2} \]

Real-World Signals
From Samples to *Measurements*

- Replace **samples** by more general **encoder** based on a few linear projections (inner products)

- Restricted Isometry Property - random matrices

\[ y = \Phi x, \text{ } x \text{ is sparse} \]

\[ M \times 1 \text{ measurements} \]

\[ N \times 1 \text{ sparse signal} \]

\[ M \approx K \ll N \text{ information rate} \]
Mutual Coherence

- Measurements selected from \textit{preset basis} $\Phi$
  (ex: 2D-Fourier measurements in MRI, permuted Walsh measurements in single-pixel camera)

- \textbf{Mutual coherence} between \textit{sparsity basis} $\Psi$ and \textit{measurement basis} $\Phi$:

  $$\mu(\Phi, \Psi) = \max_{\phi, \psi} |\langle \phi, \psi \rangle|$$

- Number of measurements needed for recovery:

  $$M \geq CKN \mu^2(\Phi, \Psi) \log N$$

  $$\mu(\Phi, \Psi) \in \left[\frac{1}{\sqrt{N}}, 1\right]$$  \cite{Candès and Romberg}
Single Pixel Hyperspectral Imaging

Hyperspectral datacube
- *Multidimensional* signals
- *Distributed* measurements

[with D. Takhar, J. Laska, T. Sun, K. Kelly, R. Baraniuk]
Joint Recovery

Distributed
Compressive Sensing (DCS)

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_J \\
\end{bmatrix} =
\begin{bmatrix}
\Phi_1 & 0 & \ldots & 0 \\
0 & \Phi_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \Phi_J \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_J \\
\end{bmatrix}
\]

Joint Recovery

[Joint work with D. Baron, M. Wakin, S. Sarvotham, and R. Baraniuk]
Inspiration: Hyperspectral Imaging

Each band measured separately

Hyperspectral datacube
Inspiration: Hyperspectral Imaging

Hyperspectral datacube

Intra-signal correlations:

spatial sparsity (wavelets)
Inspiration: Hyperspectral Imaging

Inter-signal correlations:

Hyperspectral datacube

spectral sparsity (Fourier)
Idea: Kronecker Products

Hyperspectral datacube

\[ \tilde{\Psi} = \Psi_S \otimes \Psi_F \]
One More Thing...

\[
\Phi = \begin{bmatrix}
\Phi_1 & 0 & \cdots & 0 \\
0 & \Phi_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \Phi_J
\end{bmatrix}
\]
One More Thing...

\[
\tilde{\Phi} = \begin{bmatrix}
\Phi & 0 & \ldots & 0 \\
0 & \Phi & \ldots & 0 \\
0 & 0 & \ldots & \Phi \\
\end{bmatrix}
\]

\[
\tilde{\Phi} = I \otimes \Phi
\]
Kronecker Products for DCS

\[ \tilde{\Phi} = \mathbf{I} \otimes \Phi \quad \tilde{\Psi} = \Psi_1 \otimes \Psi_2 \]
Kronecker Compressive Sensing

\[ \tilde{\Phi} = \Phi_1 \otimes \Phi_2 \]

\[ \tilde{\Psi} = \Psi_1 \otimes \Psi_2 \]

**CS performance metrics** for Kronecker product matrices

Signal classes that are **sparse/compressible** in Kronecker product bases
Kronecker Compressive Sensing

\[ \tilde{\Phi} = \Phi_1 \otimes \Phi_2 \quad \tilde{\Psi} = \Psi_1 \otimes \Psi_2 \]

**CS performance metrics** for Kronecker product matrices

- **Mutual Coherence:**
  For matrices \( \Phi_j, \Psi_j, \ 1 \leq j \leq J, \)

  \[ \mu(\Phi_1 \otimes \ldots \otimes \Phi_J, \Psi_1 \otimes \ldots \otimes \Psi_J) = \prod_{j=1}^{J} \mu(\Phi_j, \Psi_j) \]
Example: Wavelets

\[ \nu : \text{scaling function} \]
\[ \psi : \text{wavelet function} \]
\[ \psi_{i,j} : \text{translated and dilated wavelet} \]

\[ \psi_{i,j}(t) = \frac{1}{2^{i/2}} \psi \left( \frac{t}{2^i} - j \right) \]

\[ g = \nu_0 \nu + \sum_{i \geq 0} \sum_{j=0}^{2^i - 1} w_{i,j} \psi_{i,j}, \]

\[ \nu_0 = \langle g, \nu \rangle \]
\[ w_{i,j} = \langle g, \psi_{i,j} \rangle \]
Higher-Dimensional Wavelets

\[ \psi_{i_2,j_2}(t_2) \]

\[ \psi_{i_1,j_1}(t_1) \]

\[ \psi_{i_1,j_1,i_2,j_2}(t_1, t_2) = \psi_{i_1,j_1}(t_1) \otimes \psi_{i_2,j_2}(t_2) \]
Higher-Dimensional Wavelets

\[ \psi_{i_1,j_1,i_2,j_2}(t_1, t_2) = \psi_{i_1,j_1}(t_1) \otimes \psi_{i_2,j_2}(t_2) \]

Isotropic  \( i_1 = i_2 \)

Anisotropic  \( i_1 a_1 = i_2 a_2 \)

Hyperbolic  \( i_1 \neq i_2 \)
Compressibility in Wavelet Bases

- **Isotropic Besov Space** $B^s_{p,q}$: signals in $L_p$ with $s$ degrees of smoothness (derivatives) in all directions

- Isotropic wavelet characterization:

  $$|g|_{B^s_{p,q}} \approx \left( \sum_i 2^{iqs} \left( \sum_j |w_{i,j}|^p \right)^{q/p} \right)^{1/q}$$

- Compressibility: If the wavelet $\psi$ has more than $s$ vanishing moments, then a signal $g \in B^s_{p,q}$ is $s$-**compressible** in an isotropic wavelet basis
Anisotropic Besov Space $B_{p,q}^{\vec{s}}$, $\vec{s} = (s_1, \ldots, s_D)$: D-dimensional signals in $L_p$ with $s_d$ degrees of smoothness (derivatives) in $d^{th}$ dimension, $1 \leq d \leq D$

Anisotropic wavelet characterization: anisotropy $a_1, \ldots, a_D$

\[
|g|_{B_{p,q}^{\vec{s}}} \lesssim \left( \sum_{n \in \mathbb{N}} 2^{n a_1 s + r(n) (1/D - 1/p)} \left( \sum_{j_1, \ldots, j_D} |w_{r_1(n),j_1,\ldots,j_D}(n,j_D)|^p \right)^{q/p} \right)^{1/q}
\]

\[
r_d(n) = \left\lfloor \frac{a_i n}{a_1} \right\rfloor, \quad 1 \leq d \leq D \\
r(n) = \sum_{d=1}^{D} r_d(n)
\]

\[
s = \frac{D}{\sum_{d=1}^{D} \frac{1}{s_d}}
\]
Compressibility in Wavelet Bases

• If the wavelet $\psi$ is sufficiently smooth*, then a signal $g \in B_{p,q}^s$ is $s$-compressible in an anisotropic wavelet basis with $a_d = s/s_d$, $1 \leq d \leq D$

• If the wavelet $\psi$ is sufficiently smooth*, then a signal $g \in B_{p,q}^s$ is $s'$-compressible in an isotropic wavelet basis with $s' = \min_{1 \leq d \leq D} s_d$

• If the wavelet $\psi$ is sufficiently smooth*, then a signal $g \in B_{p,q}^s$ is $s$-compressible in a hyperbolic wavelet basis $\Psi = \psi \otimes \ldots \otimes \psi$

\[
s = \frac{D}{\sum_{d=1}^{D} \frac{1}{s_d}}
\]
Example: Hyperspectral Data (AVIRIS)

Original data

Spatial Wavelets

Spectral Wavelets

Hyperbolic Wavelets

\( \mathbf{x} \in \mathbb{R}^{128 \times 128 \times 64} \)
Example: Hyperspectral Data (AVIRIS)

Original data

Spectral Wavelets

Hyperbolic Wavelets
Example: Hyperspectral Data (AVIRIS)

\[ x \in \mathbb{R}^{128 \times 128 \times 128} \]

- **Kronecker**
- **Space**
- **Frequency**

Normalized error vs. Number of coefficients \( K \)
Kronecker CS with High-Dimensional Wavelets

- **Theorem:** KCS performs better than independent recovery along dimension $e$ if

$$M < C \mu(\Phi_e, \Psi_e) \frac{\beta_e}{D-1} \left(1 - \frac{\beta}{s_e}\right) \prod_{d \neq e} \mu(\Phi_d, \Psi_d)$$

where

$$\beta = \frac{D}{2 \sum_{d=1}^{D} 1/s_d} - \frac{1}{4},$$

$$\beta_e = \frac{D - 1}{2 \sum_{d \neq e} 1/s_d} - \frac{1}{4}$$

In other words, KCS is **most beneficial** compared to independent CS when applied to dimension $e$ with lowest smoothness $s_e$. 
Compressive Hyperspectral Imaging via Single Pixel Camera

\[ x \in \mathbb{R}^{128 \times 128 \times 64} \]

\( M = 4096 \) measurements per band

450nm-850nm range

[Measured by T. Sun, D. Takhar, K. Kelly]
Compressive Hyperspectral Imaging via Single Pixel Camera

[Measured by T. Sun, D. Takhar, K. Kelly]

$x \in \mathbb{R}^{128 \times 128 \times 64}$

$M = 4096$ measurements per band - \textit{joint recovery}
Experimental Results: Hyperspectral Data (AVIRIS)

SNR, dB

Kronecker Wavelet/Global
Kronecker Fourier/Global
KCS Fourier
KCS Wavelet
Independent Recovery

$\mathbf{x} \in \mathbb{R}^{128 \times 128 \times 16}$
Kronecker Compressive Sensing

- **Succinct** mathematical framework for:
  - distributed measurement techniques
  - multidimensional signal compressibility
- Can be used with **standard recovery algorithms**
- **Same matrix metrics** for CS suitability
- **Same instance-optimality guarantees** for sparse and compressible signals
- **Mutually incoherent** sparsity/measurement bases preferred for each dimension
- Certain bases **transfer compressibility** to higher dimensional ensembles through Kronecker products
- Future work
  - analysis of **suitable sparsity bases** for Kronecker products
  - **extensions** to additional applications (sparsity-based localization) [Cevher, Duarte, Baraniuk]