On Compressive Sensing of Sparse Covariance Matrices Using Deterministic Sensing Matrices

Alihan Kaplan\textsuperscript{1} Volker Pohl\textsuperscript{1} Dae Gwan Lee\textsuperscript{2}

\textsuperscript{1}Institute of Theoretical Inform. Technology
Technical University of Munich
80333 München, Germany

\textsuperscript{2}Faculty of Mathematics and Geography
Catholic University of Eichstätt-Ingolstadt
85071 Eichstätt, Germany
Overview

1. Motivation: Covariance matrix sketching

2. Some Simulations

3. Statistical Restricted Isometry property (StRIP)

4. Probabilistic Recovery Guarantee of StRIP

5. Construction Example
Motivation: Covariance matrix sketching

Given:

- $x \in \mathbb{C}^N$, a vector of $N$ independent zero-mean random variables
- covariance matrix $X = \mathbb{E}[xx^*]$, sparse in most applications
- $m$ linear measurements $y = Ax$, with measurement matrix $A \in \mathbb{C}^{m \times N}$

Determine $X$ from $Y$

$$Y = \mathbb{E}[yy^*] = A\mathbb{E}[xx^*]A^* = AXA^*$$

using vectorization...

$$\tilde{y} = (\bar{A} \otimes A) \tilde{x}$$

with $\tilde{y} = \text{vec}\{Y\}$ and $\tilde{x} = \text{vec}\{X\}$

$\Rightarrow$ Compressive Sensing setting!
Recap of relevant results in Compressive Sensing

Problem setting: \( y = Ax \) with \( A \in \mathbb{C}^{m \times N} \), where \( m \ll N \)

**Definition (Restricted Isometry Property)**

\( A \in \mathbb{C}^{m \times N} \) is said to fulfill the \( k \)-th restricted isometry property (abbrv. RIP) with the restricted isometry constant \( \delta_k \) (abbrv. RIC) if

\[
(1 - \delta_k) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k) \|x\|_2^2
\]

holds for all \( k \)-sparse \( x \in \mathbb{C}^N \).

**Theorem**

*If \( A \) fulfills the \( 2k \)-th RIP with RIC

\[
\delta_{2k} < \frac{1}{3}
\]

then every \( k \)-sparse \( x \) can be recovered uniquely by the \( \ell_1 \)-minimization (convex).*
Covariance matrix sketching & Compressive Sensing

Covariance matrix sketching as compressive sensing problem:

\[
\min \|\tilde{x}\|_0 \quad \text{subject to} \quad \tilde{y} = (\bar{A} \otimes A) \tilde{x}
\]

with \( \tilde{y} = \text{vec}\ \{Y\} \) and \( \tilde{x} = \text{vec}\ \{X\} \)

**Question:** Are there ”good” (deterministic) matrices for compressive sensing with Kronecker structure?

→ convex relaxation: \( \ell_1 \)-minimization instead of \( \ell_0 \)-minimization

- Result on RIC by Duarte and Baraniuk: \( \delta_k (\bar{A} \otimes A) \geq \delta_k (A) \)

⇒ For fixed sparsity, RIC of the Kronecker structured matrix is lower bounded by the RIC of the non-Kronecker matrix despite having quadratically more measurements.
Simulations: non-Kronecker structured matrices

**Random Gaussian:** Each entry of the sensing matrix is a Gaussian random variable.

**Random Partial Fourier:** Rows of the DFT matrix are chosen at random to form the sensing matrix.

**EHF (Equiangular harmonic frames):** The sensing matrix is a "carefully" chosen minor of the DFT matrix.
Simulation: Kronecker structured Gaussian matrices seem to be "bad" for CS.

From previous slide: \( \delta_k(A) \leq \delta_k(A \otimes A) \leq 2\delta_k(A) + \delta_k(A)^2 \)
Simulation: Kronecker structured partial Fourier

In contrast to Gaussian: random Fourier & Kronecker structured random Fourier matrices perform similarly.
Simulation: Kronecker structured EHF

→ An attempt of explanation based on *Statistical Isometry Property* (Def. by Calderbank, Howard, Jafarpour, 2010).
Standard CS versus StRIP Approach

Standard CS

- random sensing matrices
- deterministic vectors \( x \)
- recovery guarantee for all \( k \)-sparse vectors \( x \)
- recovery guarantee with high probability for random choice of \( A \)

⇒ randomness in the choice of the sensing matrix \( A \)

StRIP

- deterministic sensing matrix
- stochastic data vectors \( x \)
- recovery guarantee with high probability for random choice of \( x \)
- recovery guarantee for deterministic choice of \( A \)

⇒ randomness in the choice of the data vector \( x \)
Statistical Restricted Isometry Property (StRIP) - 1

Following definition of StRIP is by Calderbank, Howard and Jafarpour (2010).

Definition (StRIP & UStRIP)

- **A** = $\frac{1}{\sqrt{m}} \Phi \in \mathbb{C}^{m \times N}$ has $(k, \delta, \epsilon)$ - StRIP if

\[
(1 - \delta) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta) \|x\|_2^2
\]

holds with probability $1 - \epsilon$ for a random $k$-sparse vectors $x$ (uniformly distributed over all $k$-sparse vectors).

- **A** is $(k, \delta, \epsilon)$-uniqueness-guaranteed StRIP (UStRIP) if

\[
Ax = Az \iff z = x, \quad \forall k\text{-sparse } z
\]

satisfied with probability $1 - \epsilon$. 
Statistical Restricted Isometry Property (StRIP) - 2

Definition ($\eta$-StRIP)

$A = \frac{1}{\sqrt{m}} \Phi \in \mathbb{C}^{m \times N}$ with all entries of $\Phi$ having absolute value 1, is $\eta$-StRIP if St1 - St3 holds.

**St1:**
- rows of $\Phi$ are orthogonal
- sum of all elements in a row is equal to zero

**St2:**
- columns of $\Phi$ form a multiplicative group under pointwise multiplication

**St3:**
- $\exists \eta > 0$ s.t. $\left| \sum_{l=1}^{m} \phi_j[l] \right|^2 \leq m^{2-\eta}$ $\forall l$ apart from the identity column
Probabilistic recovery guarantee of StRIP

Theorem (Calderbank, Howard, Jafarpour)

Let \( \mathbf{A} = \frac{1}{\sqrt{m}} \Phi \in \mathbb{C}^{m \times N} \) be an \( \eta \)-StRIP matrix with \( \eta > 1/2 \). If \( k < 1 + (N - 1)\delta \) and

\[
m \geq \frac{(k \log N)}{\delta^2}
\]

for some constant \( c > 0 \) then \( \mathbf{A} \) is \((k, \delta, 2\epsilon)\)-UStRIP with

\[
\epsilon = 2 \exp \left( - \left( \delta - \frac{k-1}{N-1} \right)^2 \frac{m^\eta}{8k} \right).
\]

- Theorem connects structure of a deterministic CS matrix with probabilistic recovery guarantee.
- Easy applicability of StRIP on deterministic matrices (checking RIP is NP-Hard).
- Linear scaling of the number of measurements \( m \) with the sparsity \( k \).
- **Main idea:** use this theorem for Kronecker structured matrices.
\( \eta \)-StRIP for Kronecker structured matrices

**Theorem**
Assume \( A \in \mathbb{C}^{n \times N} \) is \( \eta_A \)-StRIP and \( B \in \mathbb{C}^{m \times M} \) is \( \eta_B \)-StRIP, then the following holds.

(a) \( \overline{A} \) is \( \eta_{\overline{A}} \)-StRIP with \( \eta_{\overline{A}} = \eta_A \).

(b) The matrix \( C = A \otimes B \in \mathbb{C}^{nm \times NM} \) is \( \eta_C \)-StRIP with

\[
\eta_C = \begin{cases} 
\eta_A \frac{\ln(n)}{\ln(nm)} & \text{if } n \eta_A \leq m \eta_B \\
\eta_B \frac{\ln(m)}{\ln(nm)} & \text{if } n \eta_A > m \eta_B
\end{cases}
\]

**Corollary**
If \( A \in \mathbb{C}^{m \times N} \) is \( \eta \)-StRIP, then the matrix \( \overline{A} \otimes A \in \mathbb{C}^{m^2 \times N^2} \) is \( (\eta/2) \)-StRIP.

- linear scaling of the number of measurements \( m^2 \) with the sparsity
  \[ m^2 \geq ck \log N \]
- search for deterministic matrices \( A \) s.t. \( \eta_A > 1 \).
\textbf{η-StRIP matrices with } \eta > 1

- Coherence \( \mu \) of a matrix \( A \) is defined by

\[
\mu(A) = \max_{i,j \text{ s.t. } i \neq j} |\langle a_i, a_j \rangle| \]

- using the Welch bound, for a matrix \( A \) fulfilling the \( \eta \)-StRIP definition follows:

\[
\sqrt{\frac{N - m}{m (N - 1)}} \leq \mu(A) \leq \frac{1}{\sqrt{m \eta}}
\]

\( \Rightarrow \) upper bound on \( \eta \):

\[
\eta \leq 1 + \ln \left( \frac{N - 1}{N - m} \right) \frac{1}{\ln (m)}
\]

\( \Rightarrow \) any \( \eta \)-StRIP matrix coming very close to the Welch bound
Example (Equiangular Harmonic Frames)

EHFs are partial Fourier matrices with \( \mu = \sqrt{\frac{N-m}{m(N-1)}} \)

construction is based on difference sets

\( \{0, 1, 3\} \) forms a \((7, 3, 1)\) diff. set in \( \mathbb{Z}_7 \) (integers modulo 7)

\[
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & \omega^1 & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 \\
1 & \omega^2 & \omega^4 & \omega^6 & \omega^1 & \omega^3 & \omega^5 \\
1 & \omega^3 & \omega^6 & \omega^2 & \omega^5 & \omega^1 & \omega^4 \\
1 & \omega^4 & \omega^1 & \omega^5 & \omega^2 & \omega^6 & \omega^3 \\
1 & \omega^5 & \omega^3 & \omega^1 & \omega^6 & \omega^4 & \omega^2 \\
1 & \omega^6 & \omega^5 & \omega^4 & \omega^3 & \omega^2 & \omega^1 \\
\end{bmatrix}
\]

\[
DFT_7 = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & \omega^1 & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 \\
1 & \omega^2 & \omega^4 & \omega^6 & \omega^1 & \omega^3 & \omega^5 \\
1 & \omega^3 & \omega^6 & \omega^2 & \omega^5 & \omega^1 & \omega^4 \\
1 & \omega^4 & \omega^1 & \omega^5 & \omega^2 & \omega^6 & \omega^3 \\
1 & \omega^5 & \omega^3 & \omega^1 & \omega^6 & \omega^4 & \omega^2 \\
1 & \omega^6 & \omega^5 & \omega^4 & \omega^3 & \omega^2 & \omega^1 \\
\end{bmatrix}
\]

\[
EHF_3 = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & \omega^1 & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 \\
1 & \omega^3 & \omega^6 & \omega^2 & \omega^5 & \omega^1 & \omega^4 \\
\end{bmatrix}
\]
Summary

- Investigation of Kronecker structured sensing matrices for compressive sensing
- Used the StRIP approach for Kronecker structured sensing matrices
- Proved statistical recovery guarantees where the number of measurements scales linearly with the sparsity

**Deterministic Matrices**

**Random Partial Fourier**

StRIP approach

no explanation yet