Fast Compressive Sampling Using Structurally Random Matrices

Presented by:
Thong Do (thongdo@jhu.edu)
The Johns Hopkins University

A joint work with
Prof. Trac Tran, The Johns Hopkins University
Dr. Lu Gan, Brunel University, UK
Compressive Sampling Framework

- **Main assumption:** $K$-sparse representation of an input signal

$$x_{N \times 1} = \Psi_{N \times N} \alpha_{N \times 1}$$

$\Psi$: sparsifying transform; $\alpha$: transform coefficients;

- **Compressive sampling:**

$$y_{M \times 1} = \Phi_{M \times N} x_{N \times 1}$$

- $\Phi_{M \times N}$: random matrices, random row subset of an orthogonal matrix (partial Fourier), etc.

- **Reconstruction:** $L1$-minimization (Basis Pursuit)

$$\hat{\alpha} = \arg \min \| \alpha \|_1 \quad \text{s.t.} \quad y = \Phi \Psi \alpha$$

$$\hat{x} = \Psi \hat{\alpha}$$

$\alpha_{N \times 1}$ has $K$ nonzero entries
A Wish-list of the Sampling Operator

• **Optimal Performance:**
  – require the minimal number of compressed measurements

• **Universality:**
  – incoherent with various families of signals

• **Practicality:**
  – fast computation
  – memory efficiency
  – hardware friendly
  – streaming capability
Current Sensing Matrices

• Random matrices [Candes, Tao, Donoho]
  ✓ Optimal performance
  ✗ Huge memory and computational complexity
  ✗ Not appropriate in large scale applications

• Partial Fourier [Candès et. al.]
  ✓ Fast computation
  ✗ Non-universality
    • Only incoherent with signals sparse in time
    • Not incoherent with smooth signals such as natural images.

• Other methods (Scrambled FFT, Random Filters, …)
  ✗ Either lack of universality or no theoretical guarantee
Motivation

• Significant performance improvement of scrambled FFT over partial Fourier

✓ A well-known fact [Baraniuk, Candès]

✗ But no theoretical justification

Original 512x512 Lena image

Reconstruction from 25% of measurements: 16.5 dB
Motivation

• Significant performance improvement of scrambled FFT over partial Fourier

✅ A well-known fact [Baraniuk, Candès]

❌ But no theoretical justification
Our Contributions

• Propose the concept of **Structurally random ensembles**
  – Extension of Scrambled Fourier Ensemble
• Provide **theoretical guarantee** for this novel sensing framework
• Design sensing ensembles with **practical features**
  – Fast computable, memory efficient, hardware friendly, streaming capability etc.
Proposed CS System

• Pre-randomizer:
  • *Global randomizer*: random permutation of sample indices
  • *Local randomizer*: random sign reversal of sample values
Proposed CS System

• Compressive Sampling
  \[ y = D(T(P(x))) = A(x) \]
  • Pre-randomize an input signal (P)
  • Apply fast transform (T)
  • Pick up a random subset of transform coefficients (D)

• Reconstruction
  • Basis Pursuit with sensing operator and its adjoint:
  \[ A = D(T(P(\Psi(\bullet)))) \quad A^* = \Psi^*(P^*(T^*(D^*(\bullet)))) \]

Compressed measurements → D (Random downsampler) → T (FFT, WHT, DCT,...) → P (Pre-randomizer) → Reconstruction Basis Pursuit → signal recovery

Input signal \( x = \Psi \alpha \)
Proposed CS System

- Compressive Sampling
  \[ y = D(T(P(x))) = A(x) \]
  - Pre-randomize an input signal (P)
  - Apply fast transform (T)
  - Pick up a random subset of transform coefficients (D)
- Reconstruction
  - Basis Pursuit with sensing operator and its adjoint:
    \[ A = D(T(P(\Psi(\bullet)))) \quad A^* = \Psi^*(P^*(T^*(D^*(\bullet)))) \]
Structurally Random Matrices

- Structurally random matrices with **local randomizer**: a product of 3 matrices

**Random downsampler**
\[
\begin{align*}
\Pr(d_{ii} = 0) &= 1 - \frac{M}{N} \\
\Pr(d_{ii} = 1) &= \frac{M}{N}
\end{align*}
\]

**Fast transform**
*FFT, WHT, DCT, ...*

**Local randomizer**
\[
\Pr(d'_{ii} = \pm 1) = \frac{1}{2}
\]
Structurally Random Matrices

- Structurally random matrices with **global randomizer**: a product of 3 matrices

\[
\begin{align*}
\Pr(d_{ii} = 0) &= 1 - \frac{M}{N} \\
\Pr(d_{ii} = 1) &= \frac{M}{N}
\end{align*}
\]

Random downsampler

Fast transform
\(FFT, WHT, DCT, \ldots\)

Global randomizer
**Uniformly random permutation matrix**

Partial Fourier
Sparse Structurally Random Matrices

- With local randomizer:
  - Fast computation
  - Memory efficiency
  - Hardware friendly
  - Streaming capability

Random downsampler
Pr($d_{ii} = 0$) = $1 - M / N$
Pr($d_{ii} = 1$) = $M / N$

Block-diagonal WHT, DCT, FFT, etc.

Local randomizer
Pr($d'_{ii} = \pm 1$) = $1 / 2$
Sparse Structurally Random Matrices

- With global randomizer
  - Fast computation
  - Memory efficiency
  - Hardware friendly
  - Nearly streaming capability

Random downsampler
Pr($d_{ii} = 0$) = 1 - $M/N$
Pr($d_{ii} = 1$) = $M/N$

Block-diagonal WHT, DCT, FFT, etc.

Global randomizer
Uniformly random permutation matrix
Theoretical Analysis

• **Theorem 1**: Assume that the maximum absolute entries of a structurally random matrix $\Phi_{M \times N}$ and an orthonormal matrix $\Psi_{N \times N}$ is not larger than $1/\sqrt{\log N}$ with high probability, **coherence** of $\Phi_{M \times N}$ and $\Psi_{N \times N}$ is not larger than $O(\sqrt{\log N / s})$
  
  – $s$: the average number of nonzero entries per row of $\Phi_{M \times N}$

• **Proof**:
  
  – Bernstein’s concentration inequality of sum of independent random variables

• **The optimal coherence** (Gaussian/Bernoulli random matrices): $O(\sqrt{\log N / N})$
Theoretical Analysis

- **Theorem 2**: With the previous assumption, sampling a signal using a structurally random matrix guarantees exact reconstruction (by Basis Pursuit) with high probability, provided that

\[ M \sim (KN / s) \log^2 N \]

- \( s \): the average number of nonzero entries per row of the sampling matrix
- \( N \): length of the signal,
- \( K \): sparsity of the signal

- **Proof**:
  - follow the proof framework of [Candès2007] and previous theorem of coherence of the structurally random matrices

- **The optimal number of measurements required by Gaussian/Bernoulli dense random matrices**: \( K \log N \)

---

Simulation Results: Sparse 1D Signals

- Input signal sparse in DCT domain
  - $N=256$, $K=30$

- Reconstruction:
  - Orthogonal Matching Pursuit (OMP)

- **WHT256 + global randomizer**
  - Random permutation of samples indices
    + Walsh-Hadamard

- **WHT256 + local randomizer**
  - Random sign reversal of sample values
    + Walsh-Hadamard

- **WHT8 + global randomizer**
  - Random permutation of samples indices
    + block $8 \times 8$ diagonal Walsh-Hadamard

✓ The fraction of nonzero entries: $1/32$
✓ 32 times sparser than Scrambled FFT, i.i.d Gaussian ensemble,…
Simulation Results: Compressible 2D Signals

• Experiment set-up:
  – Test images: $512 \times 512$ Lena and Boat;
  – Sparsifying transform $\psi$: Daubechies 9/7 wavelet transform;
  – L1-minimization solver: GPSR [Mario, Robert, Stephen]
  – Structurally random sensing matrices $\Phi$:
    • WHT512 & local randomizer
      – Random sign reversal of sample values + $512 \times 512$ block diagonal Hadamard transform;
      – Full streaming capability
    • WHT32 & global randomizer
      – Random permutation of sample indices + $32 \times 32$ block diagonal Hadamard transform;
      – Highly sparse: The fraction of nonzero entries is only $1/2^{13}$
Rate-Distortion Performance: *Lena*

Partial FFT in wavelet domain:
- transform the image into wavelet coeffs
- sense these coeffs (rather than sense directly image pixels) using partial FFT

Full streaming capability

8000 times sparser than the Scrambled FFT

Partial FFT
- Partial FFT in wavelet domain
- Scrambled FFT
- WHT512 + local randomizer
- WHT32 + global randomizer

R-D performance of 512x512 Lena
Reconstructed Images: *Lena*

- Reconstruction from 25% of measurements using GPSR

Original 512x512 Lena image  Partial FFT: **16** dB  Partial FFT in wavelet domain: **30.1** dB

Scrambled FFT: **29.3** dB  WHT512 + local randomizer: **28.4** dB  WHT32 + global randomizer: **29** dB
Rate-Distortion Performance: *Boat*

R-D performance of 512x512 Boat image
Future Research

• Develop theoretical analysis of structurally random matrices with greedy, iterative reconstruction algorithms such as OMP

• Application of structurally random matrices to high dimensionality reduction
  – Fast Johnson-Lindenstrauss transform using structurally random matrices

• Closer to deterministic compressive sampling
  – Replace a random downsampler by a deterministic downsampler or a deterministic lattice of measurements*
  – Develop theoretical analysis for this nearly deterministic framework

Conclusions

- Structurally random matrices
  - Fast computable;
  - Memory efficient:
    - Highly sparse solution: random permutation $\Rightarrow$ fast block diagonal transform $\Rightarrow$ random sampling;
  - Streaming capability:
    - Solution with full streaming capability: random sign flipping $\Rightarrow$ fast block diagonal transform $\Rightarrow$ random sampling;
- Hardware friendly;
- Performance:
  - Nearly optimal theoretical bounds;
  - Numerical simulations: comparable with completely random matrices
References


• thanglong.ece.jhu.edu/CS/fast_cs_SRMS.rar

Thank you for your attention!