Lecture 10: Combining Multiple Representations

David Donoho
Statistics Department
Stanford University

Collaborator

Figure 1: Xiaoming Huo

Contributor

Figure 2: Jean-Luc Starck

Background Reading

- Wavelet, "Adaptive Wavelet Analysis"
- Mallat, "A Wavelet Tour of Signal Processing"

Books

http://www.stat.stanford.edu/~jstarck

X. Huo Thesis: Combined Image Representation, 1999

Chen, Donoho, Saunders, Basis Pursuit, SIAM J Comp 1999

Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

• Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

• Books:
- Mallat, "A Wavelet Tour of Signal Processing"
- Wickerhauser, "Adapted Wavelet Analysis"

• Background Reading

- Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

- Mallat, "A Wavelet Tour of Signal Processing"

- Wavelet, "Adaptive Wavelet Analysis"

Figure 2: Jean-Luc Starck

Collaborator

Figure 1: Xiaoming Huo

Contributor

Background Reading

- Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

- Mallat, "A Wavelet Tour of Signal Processing"

- Wavelet, "Adaptive Wavelet Analysis"

Books

http://www.stat.stanford.edu/~jstarck

X. Huo Thesis: Combined Image Representation, 1999

Chen, Donoho, Saunders, Basis Pursuit, SIAM J Comp 1999

Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

• Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

• Books:
- Mallat, "A Wavelet Tour of Signal Processing"
- Wickerhauser, "Adapted Wavelet Analysis"

- Background Reading

- Donoho and Huo, "Uncertainty Principles and Ideal Atomic Decomposition"

- Mallat, "A Wavelet Tour of Signal Processing"

- Wavelet, "Adaptive Wavelet Analysis"
Many Available Transforms

• Sinusoids
• Cosine Packets
• Wavelets
• Wavelet Packets
• Anisotropic Wavelet Bases
• Ridgelets
• Curvelets
• New Developments
• Cosine Packets
• Wavelets
• Sinusoids

Many Available Transforms


Lack of Universality

• There is no universal basis

• Different Bases good for Different Purposes

• Here good means Sparsity

• Wavelets for point discontinuities

• Ridgelets for discontinuities along lines

Example 1.

• Sinusoids for High-frequency oscillatory phenomena

• Wavelets for Impulsive phenomena

Mallat-Zhang (1993) Several Bases of Interest

Dictionary

\( \{ \phi \in \mathcal{V} \} = \mathcal{F} \)

\( \mathcal{F} \phi = \mathcal{V} \)

Combined Representation

\( \mathcal{V} = \bigcup \mathcal{F} \phi \cap \cdots \cap \mathcal{F} \phi \)

Combined Bases into Dictionaries

Example 2.

• Wavelets for line discontinuities

• Ridgelets for point discontinuities

Example 1.
If \( S \) is the superposition of several ‘simple’ phenomena, \( S_1, S_2, \ldots, S_d \), such that each \( S_i \) is sparsely represented in \( \Phi_i \), we dream of a representation of the superposed object which is the superposition of the sparse decompositions.

Problem – generally completely dense.

\[
\min_{\alpha} \| \alpha \|_0 \\
\text{s.t.} \quad S = \sum_{\gamma} \alpha_{\gamma} \phi_{\gamma},
\]

\[\| \alpha \|_0 = \# \{ \gamma : \alpha_{\gamma} \neq 0 \} \]

Search through \( \Phi \) looking for a sparse subset providing exact decomposition.

Each \( S_i \) can be represented by any single one of the dictionaries.

Recompression is non-unique.

\( \Phi \) is overcomplete (several bases).

\( \Phi \) is overcomplete.

Recompression of \( S_i \) by wrong dictionary \( \Phi_i \) is highly non-sparse.

Minimax approach

Three Approaches

Method of Frames

- \( P_2 \) : \( \min \| \alpha \|_2 \\
\text{s.t.} \quad S = \sum_{\gamma} \alpha_{\gamma} \phi_{\gamma} \)

Solvable by linear algebra

\[
\alpha = (\Phi^T \Phi)^{-1} \Phi^T S.
\]

Problem – generally completely dense.

\[
\Phi_1 \cdots \Phi_d = \Phi
\]

\[
\min_{\alpha} \| \alpha \|_2 \\
\text{s.t.} \quad S = \sum_{\gamma} \alpha_{\gamma} \phi_{\gamma},
\]

Search through \( \Phi \) looking for a sparse subset providing exact decomposition.
If have two ortho bases, can use Alternating Soft Thresholding

\[
\alpha_1 = 0 \\
\alpha_2 = \eta \circ \Phi_2 (y - \Phi_1 \alpha_1) \\
\alpha_1 = \eta \circ \Phi_1 (y - \Phi_2 \alpha_1) \\
0 = 1 \\
\]

\[l_1 \Phi - \Phi_1 \circ \alpha = 1\]
\[l_2 \Phi - \Phi_2 \circ \alpha = 1\]
\[0 = 1\]

Bruce, et. al (1997?)

If Soft Thresholding, fixed point is solution of BP objective.

Example of Sparse Decomposition via BP – 2

Example of Sparse Decomposition via BP – 1

BP via Alternating Thresholding

If have two ortho bases, can use Alternating Soft Thresholding

Modern interior-point optimization. Solvable by linear programming.

\[
\|a\|_q^p = \left\{ \begin{array}{cl} a^q & : q \neq 1 \\
\|a\|_1 & : (p, q) \\
\end{array} \right.
\]

Basis Pursuit
Examples of BP-like Processes

- Xiaoming Huo – Stanford 1999 Ph.D.
- J.L. Starck – Stanford 2000 Visitor

Combined Wavelet/Beamlet Rep. – X. Huo

Combined Curvelet/Wavelet Rep. – J.L. Starck
Combined Ridgelet/Wavelet Rep. – Starck

Experimental Results

Empirically, solution of BP is frequently quite sparse;

Many examples: Chen, Donoho, Saunders (1999)

Surprising Phenomenon

Empirically, solution of BP is frequently quite sparse;

Bp solution may perfectly recover the specific atoms and specific

coefficient used in the synthesis.

Separation at Moderate Dynamic Range

True across a wide range of amplitude ratios between the sinusoid

and spike components.

• BP recovered exactly the indices and coefficients of the terms

and spikes.

• Dictionary a combined time/frequency dictionary of sinusoids

S a sum of 2 sinusoids and 2 spikes.

S. Chen (Stanford Thesis, 1996)

Example

\[ \begin{align*}
\text{(a) Dynamic} & 1 \text{ on log scale} \\
\text{(b) MP: DCT Coefs on log scale} \\
\text{(c) MP: DIRAC Coefs on log scale} \\
\text{(d) BP: DCT Coefs on log scale} \\
\text{(e) BP: DIRAC Coefs on log scale}
\end{align*} \]
Separation at High Dynamic Range

Comparison

Atomic Decomposition

Ideal Atomic Decomposition

Atomic Decomposition

Goal for Today

• In certain dictionaries $\Phi$
  • Signal sufficiently sparse combination of terms from dictionary
  • Unique Solution to $P_1$
  • Unique Solution to $P_0$
  • Solutions are the same
  • We refer to solution $P_0$.

Refinements

• How much sparsity is required?
• What about dictionaries makes this possible?
• What about $\ell^1$ makes this possible?

Matching Pursuit (Greedy) recovery of indices and coefficients was only approximate and became very inaccurate when the sinusoidal and spike components were at very different amplitudes.
We claim that

- Under certain sparsity conditions,
- In certain dictionaries

Minimum $\ell_1$-norm decomposition is an ideal atomic decomposition.

**Theorem 0.1**

Let $S = \sum_{\gamma \in T} \alpha_{\gamma} \phi_{\gamma} + \sum_{\gamma \in W} \alpha_{\gamma} \phi_{\gamma}$ where $T$ is a subset of the "time domain" $\{(1, \tau)| \tau = 0, 1, \ldots, N-1\}$ and $W$ is a subset of the "frequency domain" $\{(2, w)| w = 0, 1, \ldots, N-1\}$. If $|T| + |W| < \sqrt{N}$, then $(P_0)$ has a unique solution. Meanwhile, there exist $T, W$ so that $|T| + |W| = \sqrt{N}$ and $(P_0)$ has a non-unique solution.

**Theorem 0.2**

Let $S = \sum_{\gamma \in T \cup W} \alpha_{\gamma} \phi_{\gamma}$ with $T, W$ as in Theorem 0.1. If $|T| + |W| < \frac{1}{2} \sqrt{N}$, then $(P_1)$ has a unique solution, which is also the unique solution of $(P_0)$. Meanwhile, there exist $T, W$ so that $|T| + |W| = \sqrt{N}$ and $(P_1)$ has a non-unique solution.

If the signal $S$ truly has a very sparse decomposition in the time/frequency dictionary, the solution is unique, and basis pursuit (\(\ell_1\) decomposition) will find it.

Applying the terminology, we have:

- $N^\wedge = |M| + |L|$
- $N^\wedge > |M| + |L|$
- $\{t \in \mathbb{Z}\} = S \cap \mathbb{Z}$
- $\mathcal{L} = M \mathcal{E} + \mathcal{F}$
- $\mathcal{F}$ is a set of functions
- $\mathcal{E}$ is a set of elements
- $\mathcal{L}$ is a set of elements

We claim that

- The minimum $\ell_1$-norm decomposition is an ideal atomic
- In certain dictionaries
- Under certain sparsity conditions

Combined dictionary: $\Phi = \Phi_1 \cup \Phi_2$

- Spike basis $\phi_1(t) = \{t = \tau\}$
- Fourier basis $\phi_2(w) = \frac{1}{\sqrt{N}} \exp\left(\frac{2\pi i tw}{N}\right)$ for $w = 0, 1, \ldots, N-1$
Digression 1: Uncertainty Principle

Underlying Theorem: An Uncertainty Principle

The combined analysis of a signal into both time and frequency domains cannot yield a transform pair which is simultaneously sparse in both domains. A signal cannot be sparsely represented from both the time and frequency sides simultaneously. Otherwise, the atomic decomposition would be nonunique, so no unique interpretation is possible. The nonlinearity of the \( \ell_1 \) norm is responsible for our phenomenon.

Byproduct

Ideal atomic decomposition depends on very particular properties of the \( \ell_1 \) norm. In effect, \( (\ell_1, P_1) \) asks to find the member of a linear subspace closest to the origin in the \( \ell_1 \) norm. This closest point problem (which would be a linear problem in the \( \ell_2 \) norm) is highly nonlinear in the \( \ell_1 \) norm, and the nonlinearity is responsible for our phenomenon.

Digression 2: Nonlinearity of \( \ell_1 \)

Ideal atomic decomposition depends on very particular properties of the \( \ell_1 \) norm. New uncertainty principle and generalization to other pairs than time-frequency.
Precedent: Logan's Phenomenon

• Decompose signal \( S(t) = B(t) + N(t) \) – Ω-bandlimited function and impulsive noise.
• Find \( \tilde{B} \) Ω-bandlimited \( L_1 \) closest to \( S \).
• Suppose product \( \| \Omega \| \text{supp}(N) \| \leq c \).

Perfect separation: \( \tilde{B} = B \).

Phenomenon highly nonlinear in sense that perfect reconstruction holds at all signal/noise ratios.

Logan (1965), here bandlimiting holds at all signal/noise ratios.


Perfect separation \( B = \tilde{B} \).
Suppose product \( \| \Omega \| \text{supp}(N) \| \leq c \).
Find \( \tilde{B} \) bandlimited \( L_1 \) closest to \( S \).

Theorem 0.3 Let \( f(\theta) \) function on the circle \([0, 2\pi] \), a superposition of sinusoids and wavelets.

Wavelet/Sinusoid Dictionaries

\[
\psi_\lambda(\theta) = \sum_{n=0}^{\infty} c_n e^{in\theta},
\]

\( c_n \) is the Meyer-Lemarié wavelets, and \( n_0 = 2^j \).

There is a constant \( C \) with the following property. Let \( N_j(\text{Wavelets}) \) be the number of Meyer wavelets at resolution level \( j \) and let \( N_j(\text{Sinusoids}) \) be the number of sinusoids at resolution level \( j \).

There is a constant \( C \) with the following property. Let \( N_j \) be the number of Meyer wavelets and \( n_0 = 2^j \).

There is a constant \( C \) with the following property. Let \( N_j \) be the number of Meyer wavelets and \( n_0 = 2^j \).

There is a constant \( C \) with the following property. Let \( N_j \) be the number of Meyer wavelets and \( n_0 = 2^j \).

\[
N_{\text{Wavelets}}(\Omega) + N_{\text{Sinusoids}}(\Omega) \leq C \cdot 2^{j/2}, j = j_0 + 1, \ldots, \infty
\]
Wavelet/Sinusoid Dictionaries [Cont.]

Consider the overcomplete dictionary \( \Phi \) consisting of Meyer-Lemarié wavelets and of sinusoids at frequencies \( n \geq 2^{j_0} + 1 \).

There is at most one way of decomposing a function \( f \) in the form

\[
\sum_{\lambda} \alpha_{\lambda} \psi_{\lambda} + \sum_{n \geq n_0} c_n \phi_n = f
\]

This is the unique solution to the minimum \( \ell_1 \)-optimization problem

\[
\min \sum |\alpha_{\lambda}| + \sum_{n \geq n_0} |c_n|
\]

Interpretation

No assumption about the sparsity or non-sparsity of the representation of \( f \).

A phenomenon near scale \( 2^{-j_0} \) and frequency \( 2^{j_0} \) cannot

be detected at scale \( 2^{-j_0} \).

Sinusoid/Wavelet Uncertainty Principle

Intuitively

A phenomenon near scale \( 2^{-j_0} \) and frequency \( 2^{j_0} \) cannot

be detected at scale \( 2^{-j_0} \).

Formally

If a function \( f \) has at most \( C \cdot 2^{j_0} \) nonzero wavelet coefficients and

nonzero sinusoid coefficients at each level \( j \), then the function is zero.

Wavelet/Ridgelet Dictionary

No matter what the decomposition, two systems become highly disjoint.

Sparse sums of sinusoids lead to sparse sums of wavelets at high resolution.

Theoretically more nonzero terms at high resolution than we can at low resolution.

Notion of sparsity level-dependent.

If sparsity holds, gives ideal atomic decomposition.

No assumption about the sparsity of the representation of \( f \).

Consider the overcomplete dictionary consisting of wavelets and sinusoids.

Wavelet/Sinusoid Dictionaries [Cont.]
Theorem [Conclusion]

There is $C > 0$ so that if $N_j(\text{Wavelets}) + N_j(\text{Ridgelets}) \leq C \cdot 2^j/2$, then the function is vanishing.

Formally, $f$ cannot have a sparse representation in both the wavelets and the ridgelets bases.

A phenomenon occurring near scale $2^{-j}$ and frequency $2^j$.

Initiatively, level-dependent sparsity.

Wavelet/Ridgelet Uncertainty Principle

Intuitively, a phenomenon occurring near scale $2^{-j}$ and frequency $2^j$ cannot have a sparse representation in both the wavelets and the ridgelets bases.

Formally, if an $L^2$ function $f$ has at most $C \cdot 2^{j/2}$ nonzero wavelet coefficients, then $f$ is vanishing.

Agenda

- Time/Frequency: $\ell_0$
- Time/Frequency: $\ell_1$
- Appl: Bandlimited Approximation
- From Complex to Real Sinusoids
- General Basis Pairs
- Multi-scale Bases
- Add: Bounded Approximation
Theorem 0.4
Suppose \((x_t)_{N^{-1}t=0}^N\) has \(N_t\) nonzero elements and that its Fourier transform \((\hat{x}_w)_{N^{-1}w=0}^N\) has \(N_w\) nonzero elements. Then

\[ N_t N_w \geq N \]

and so

\[ N_t + N_w \geq 2 \sqrt{N} \]

(5)

Donoho and Stark (1989)
Proof of Theorem 1

Suppose $S$ had two decompositions:

$$S = \Phi \alpha_1, S = \Phi \alpha_2$$

0 = $\Phi (\alpha_1 - \alpha_2)$.

so $\alpha_1 - \alpha_2 \in \mathbb{N}$

Suppose both decompositions sparse:

$$\|\alpha_i\|_0 < \sqrt{N}.$$ 

Then there must be a transform pair $(x, -\hat{x})$ with

$$N t + N w < 2 \sqrt{N}.$$ 

This would violate the $\ell_0$ uncertainty principle!

Therefore $\ell_0$ optimization is unique if $S = \Phi \alpha$ and $\|\alpha\|_0 < \sqrt{N}$.

2. Uniqueness of $\ell_1$ Optimization

For $\alpha$ unique solution, need

$$\|\tilde{\alpha}\|_1 > \|\alpha\|_1.$$ 

$$\Phi \tilde{\alpha} = \Phi \alpha$$

Equivalently, for every $\delta \in \mathbb{N}$ we must have

$$\|\alpha + \delta\|_1 - \|\alpha\|_1 > 0,$$

unless $\delta = 0$.

For a unique solution, need

$$N^\wedge > 0 \|\nu\|.$$ 

optimization is unique if $S \nu = 0$ and $\nu \Phi = S \nu$.

Now $\nu = 0$.

Suppose both decompositions sparse:

$$N \in \mathbb{N}, 1 \|\nu\| < 1 \|\delta\|$$

2. Uniqueness of $\ell_1$ Optimization

$\nu \Phi = S \nu = \nu \Phi$.

Suppose $S$ had two decompositions:

$$N^\wedge > 0 \|\nu\|.$$ 

This would violate the uncertainty principle. Therefore $N^\wedge > a^N + 1 \|x\| - 1 \|\hat{x}\|$, with $N \in \mathbb{N}$.

Then there must be a transform pair $(x, -\hat{x})$ such that

$$N^\wedge > a^N + 1 \|x\| - 1 \|\hat{x}\|.$$ 

Suppose both decompositions sparse:

$$N \in \mathbb{N}, 1 \|\nu\| < 1 \|\delta\|$$

2. Uniqueness of $\ell_1$ Optimization

$\nu \Phi = S \nu = \nu \Phi$.

Suppose $S$ had two decompositions:
Theorem 0.5

Let $T$ be a subset of the time domain and $W$ be a subset of the frequency domain. Then

$$\mu(T, W) \leq |T| + |W| + \sqrt{N} + 1.$$  \hfill (10)

In particular, if $|T| + |W| \leq \frac{1}{2} \sqrt{N}$, then $\mu(T, W) < \frac{1}{2}$, and the optimization problem $(P_1)$ has a unique solution.

Proof

Lemma 0.6

Consider the capacity defined by the optimization problem

$$\inf_{x \in \mathcal{L}} \|x\|_1 + \frac{1}{2} \|\hat{x}\|_1.$$ \hfill (11)

Theorem 0.2

Let $\mathcal{L}$ be a subset of the time domain and let $\mathcal{W}$ be a subset of the frequency domain. Then

$$\frac{1}{2} + \frac{\mu}{(\mathcal{L}, \mathcal{W})}.$$ \hfill (12)

Lemma 0.7

Consider the capacity defined by the optimization problem

$$\min_{x} \|x\|_1 + \frac{1}{2} \|\hat{x}\|_1,$$ \hfill (13)

subject to

$$\hat{x} \in \mathcal{W},$$ \hfill (14)

and the frequency-side capacity defined by the optimization problem

$$\min_{\hat{x}} \|x\|_1 + \frac{1}{2} \|\hat{x}\|_1,$$ \hfill (15)

subject to

$$\hat{x} \in \mathcal{L}.$$ \hfill (16)

Then

$$\mu(x, \hat{x}) = \frac{1}{2} \left( \|x\|_1 + \|\hat{x}\|_1 \right) + \frac{1}{2} \left( \|x\|_1 + \|\hat{x}\|_1 \right).$$ \hfill (17)

Real Sinusoids

\( \phi(w) \)

orthobasis for \( l^2, N \).

\[ \tilde{x}_w = \langle x, \phi(w) \rangle \]

Fourier-Bessel coefficients.

\( T \) and \( W \) subsets of \( t \)- and \( w \)-index space, respectively.

\[ \tilde{\mu}(T, W; \phi) = \sup \sum_{T} |x_t| + \sum_{W} |\tilde{x}_w| \parallel x \parallel_1 + \parallel \hat{x} \parallel_1 \]

Earlier \( \mu(T, W) \) special case with \( \phi(w) = \frac{1}{\sqrt{N}} e^{i 2\pi wt/N} \).

\[ \tilde{M} = \max_w \max_t |\phi(w)(t)| \]

Then

\[ \max_{|n|} \max_{w} |(i^n \phi)|^2 = J^N \]

Put

\[ \frac{N}{c} = J^N \]

we have

\[ |(1)-(\frac{N}{c})^\frac{1}{c}| = (i)^{1-N} \]

\[ 1 - \frac{2}{N} \cdots 1 = \eta \]

\[ \cos(n/\eta) \frac{N}{c} = (i)^{n/\eta} \]

\[ \sin(n/\eta) \frac{N}{c} = (i)^{1-1/\eta} \]

\[ \frac{N}{c} / 1 = (i)^{0/\eta} \]

Domain \( t = 0, 1, \cdots, N - 1 \), with \( N \) even.

Real Fourier Basis

\( \text{Real } \phi \)

\( \text{Orthobasis } \phi \)

\( \text{for } l^2 \)

\( \tilde{\phi} \)

\( \text{Real } Bessel \)

\( \text{Orthobasis } \phi \)

\( \text{for } l^2 \)
Theorem 0.8

Let $\Phi_1$ be the basis of spikes and let $\Phi_2$ be the basis of real sinusoids. If $S$ is a superposition of atoms from sets $T$ and $W$ and $|T| + |W| \leq \sqrt{N}/2$, then

1. the solution to $(P_0)$ is unique;
2. the solutions of $(P_1)$ is unique;
3. the two solutions are identical.

General Pairs of Bases

Theorem 0.9

$\Phi_1$ and $\Phi_2$ orthonormal bases for $\mathbb{R}^N$; $M(\Phi_1, \Phi_2) = \sup \{ |\langle \phi_1, \phi_2 \rangle| : \phi_1 \in \Phi_1, \phi_2 \in \Phi_2 \}$.

Suppose $S = \Phi\alpha$, where $\alpha$ obeys $\|\alpha\|_0 < 1 + (1 + M - 1)$, then $\alpha$ is the unique solution to $(P_1)$ and also the unique solution to $(P_0)$.

Small $M$ offers possibility of Ideal Atomic Decomposition.

General Basis Pairs Uncertainty Principle

Theorem 0.10

Let $\Phi_1$ and $\Phi_2$ be orthonormal bases for $\mathbb{R}^N$. Let $T$ index the collection of nonzero coefficients for $x$ in basis 1, and $W$ index the collection of nonzero coefficients for $x$ in basis 2. Then, $|T| + |W| \geq (1 + M - 1)$.

Suppose $S = \Phi\alpha$, where $\alpha$ obeys $\|\alpha\|_0 < 1 + (1 + M - 1)$, then $\alpha$ is the unique solution to $(P_1)$ and also the unique solution to $(P_0)$.

Small $M$ offers possibility of Ideal Atomic Decomposition.

Then

$$\frac{\sqrt{N}}{2} \geq |T| + |W|$$

and $M$ is a superposition of atoms from sets $T$ and $W$. If $S$ is a superposition of atoms from sets $T$ and $W$, then $S$ is an ideal atomic decomposition.

Lemma 0.11

For any pair of orthonormal bases $\Phi_1$ and $\Phi_2$, $M(\Phi_1, \Phi_2) \geq 1/\sqrt{N}$. Small $M$ bases are mutually incoherent.
Random Orthogonal Bases \( \Phi_1, \Phi_2 \) random orthogonal matrices, uniform on \( O(N) \)

\[
M(\Phi_1, \Phi_2) \approx 2 \sqrt{\log(N)/N}
\]

**Table 1:** Table of the medians of the maximum amplitude in a real \( N \times N \) random orthogonal matrix, out of 100 generations.

<table>
<thead>
<tr>
<th>Size</th>
<th>( \sqrt{N} )</th>
<th>( \sqrt{N} )</th>
<th>( \sqrt{N} )</th>
<th>( \sqrt{N} )</th>
<th>( \sqrt{N} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>0.658</td>
<td>0.510</td>
<td>0.389</td>
<td>0.294</td>
<td>0.221</td>
</tr>
<tr>
<td>64</td>
<td>0.864</td>
<td>0.884</td>
<td>0.910</td>
<td>0.919</td>
<td>0.930</td>
</tr>
<tr>
<td>128</td>
<td>0.934</td>
<td>0.952</td>
<td>0.958</td>
<td>0.973</td>
<td>0.977</td>
</tr>
<tr>
<td>256</td>
<td>0.941</td>
<td>0.960</td>
<td>0.966</td>
<td>0.981</td>
<td>0.985</td>
</tr>
<tr>
<td>512</td>
<td>0.947</td>
<td>0.966</td>
<td>0.970</td>
<td>0.990</td>
<td>0.989</td>
</tr>
<tr>
<td>1024</td>
<td>0.953</td>
<td>0.972</td>
<td>0.976</td>
<td>0.994</td>
<td>0.993</td>
</tr>
</tbody>
</table>

**Impulse Resistant Scrambling**

- Form random orthogonal matrix \( U \), known only to sender and intended recipient.
- \( M = \max_{i,j} |U_{ij}| \).
- Place \( K < M - 1 \) nonzero elements in \( N \)-vector \( S \).
- Scrambling: \( E = US \).
- Transmit \( E \) in the clear.
- Intended recipient: \( S = U^T E \).
- Scrambling-resistant to arbitrary changes in fewer than \( K \) components of \( E \).

**Corollary**

- Completely resistant to arbitrary changes in fewer than \( K \) components of \( E \).
- In general: Transmit \( O(\sqrt{N}/\log(N)) \) real numbers in a vector of length \( N \).
- Immune to \( O(\sqrt{N}/\log(N)) \) gross errors.

---

**Application: Speech Scrambling**

A. D. Wyner (1979), Sloane (1983)

- Form random orthogonal matrix \( U \), known only to sender and intended recipient.
- Scrambling: \( E = US \).
- Transmit \( E \) in the clear.
- Intended recipient: \( S = U^T E \).

**V. D. Wyner (1979), Sloane (1983)**
MultiScale Bases

Key Feature of $(\text{Spike,Sinusoid})$ pair: $M$ small for large $N$: $M = O \left( \frac{N - 1}{2} \right)$

More generally, $M$ roughly 1.

Orthobases $\Phi_1$ and $\Phi_2$; consider capacity defined by optimization problem

$$\min_{\| \Phi_1^T x \|_1 + \| \Phi_2^T x \|_1} \langle x, \phi_\gamma \rangle = 1$$

Previous analysis $\mathcal{V}(\gamma)$ does not depend on $\gamma$ or at most weakly.

$$1 = \langle \phi, x \rangle, \quad \text{subject to} \quad 1 \| x_2 \Phi_1 \| + 1 \| x_2 \Phi_2 \| = O$$

More generally, $\mathcal{V}_1$ roughly 1.

$$(\frac{1}{2} - N) O = N$$

Key Feature of $(\text{Spike,Sinusoid})$ pair:

MultiScale Bases

Idea for Wavelets and Sinusoids

Wavelets and Sinusoids are based on common dyadic subbands

Block Diagonal Structure

Definition 0.12 Joint block diagonal structure

Orthogonal direct sum decomposition of $\mathbb{R}^N$ as $\mathbb{R}^N = X_0 \oplus X_1 \oplus \ldots \oplus X_J$.

Grouping of indices $\Gamma_1$, $j$ for basis 1 so that

span $(\phi_\gamma: \gamma \in \Gamma_1, j) = X_j$

and grouping of indices $\Gamma_2$, $j$ for basis 2 so that

span $(\phi_\gamma: \gamma \in \Gamma_2, j) = X_j$.
Lemma 0.13

\[ M_j = M(\{\phi_\gamma: \gamma \in \Gamma_1, j\}, \{\phi_\gamma: \gamma \in \Gamma_2, j\}) \]

\[ \text{blockwise mutual incoherence.} \]

Then if \( S \) is superposition of \( N_1, j \) terms from \( \Gamma_1, j \) and \( N_2, j \) terms from \( \Gamma_2, j \), and \( N_1 + N_2 < \frac{1}{2} M_j - 1 \), the solutions of each \( (P_0, j) \) and each \( (P_1, j) \) are unique and are the same.

Real bi-sinusoids

\[ e^{i \omega}(t) = b_j(\omega) \cos\left(\frac{2 \pi \omega t}{N}\right) - b_j(\omega') \cos\left(\frac{2 \pi \omega' t}{N}\right) \]

\( w < 2^j \cdot \frac{4}{3} \), \( \sigma = 1 \);

\[ e^{i \omega}(t) = b_j(\omega) \cos\left(\frac{2 \pi \omega t}{N}\right) + b_j(\omega') \cos\left(\frac{2 \pi \omega' t}{N}\right) \]

\( w \geq 2^j \cdot \frac{4}{3} \), \( \sigma = 1 \);

\[ e^{i \omega}(t) = b_j(\omega) \sin\left(\frac{2 \pi \omega t}{N}\right) - b_j(\omega') \sin\left(\frac{2 \pi \omega' t}{N}\right) \]

\( w < 2^j \cdot \frac{4}{3} \), \( \sigma = 2 \);

\[ e^{i \omega}(t) = b_j(\omega) \sin\left(\frac{2 \pi \omega t}{N}\right) + b_j(\omega') \sin\left(\frac{2 \pi \omega' t}{N}\right) \]

\( w \geq 2^j \cdot \frac{4}{3} \), \( \sigma = 2 \).

Here \( \omega \) is the “twin” of \( \omega \), and obeys

\[ 2^j - \omega' = \omega - 2^j \cdot \frac{4}{3} \]

\[ 2^j + 1 - \omega' = \omega' - 2^j + 1 \cdot \frac{4}{3} \]

\[ \omega < 2^j \cdot \frac{4}{3} \; \text{and} \; \omega > 2^j \cdot \frac{4}{3}. \]

\[ \left(1 + \frac{\omega'}{\omega}, \frac{\omega'}{\omega}\right) \in \mathbb{N} \]

\[ N/\omega = \epsilon(m)^q + \epsilon(n)^q \]

\[ f \leq f_1 \]

The solutions of each \( \Phi_1 \) and each \( \Phi_2 \) are unique and are the same.

The dictionary \( \Phi \) is joint block diagonal.

\[ M_j = 2^{-j/2} \]

\[ \text{Application: Wavelets and bi-sinusoids} \]

\[ \text{Real bi-sinusoids} \]

\[ \text{Meyer-Lemarié Wavelets} \]
To do ...

Summary

Contact

Problem of Multiple Representations

Theoretical explanation

Empirical Successes

• More about Multiscale Bases

• More about Non-orthogonal Bases

• It is unnecessarily pessimistic

• M is unnecessarily pessimistic

• Necessary

Figure 3: George Dantzig

Figure 3: George Dantzig