Compressed Imaging Algorithm

Tim Tai-Yi Lin
in collaboration with Felix Herrmann
EOS - University of British Columbia

Built on previous work by
Joris Grimbergen, Frank Dessing, E. Candes et. al.
with special thanks to authors of StOMP
D. L. Donoho et. al.

SLIM Seismic Laboratory for Imaging and Modeling

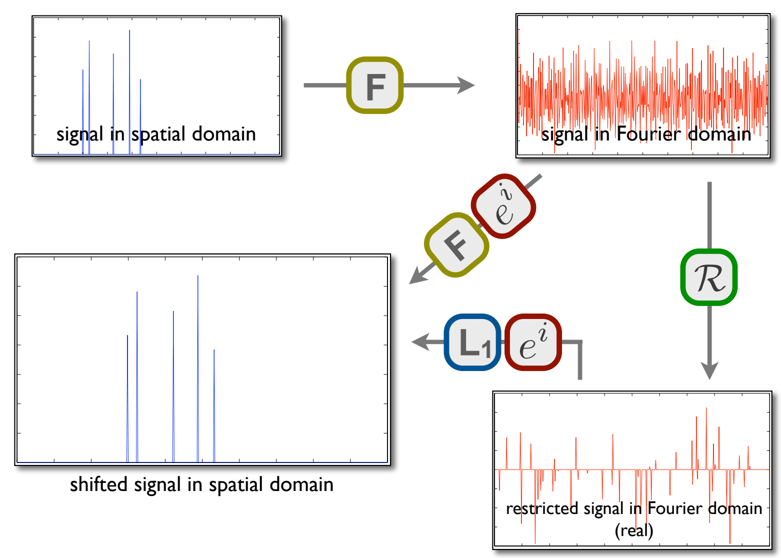
August 29, 2006

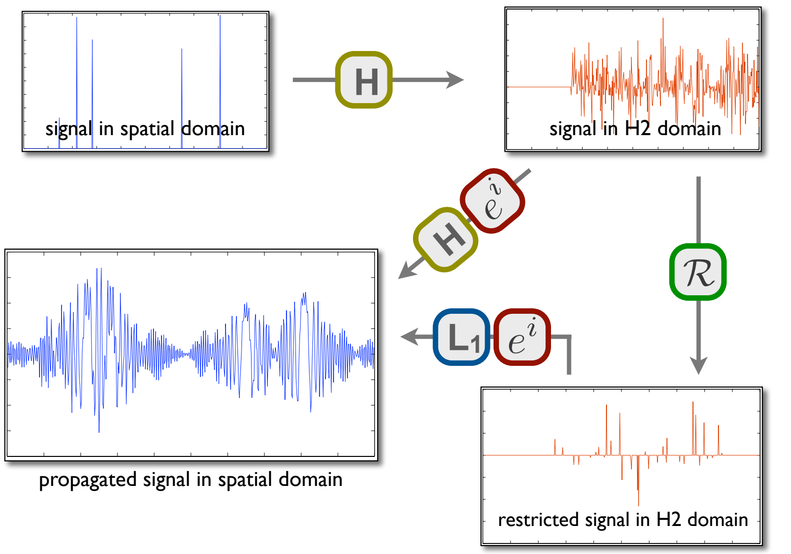
Problem Description

- Very large-scale imaging problems
- Usage of wave propagator that is desirable but expensive to compute
- Imaging is necessarily a recovery problem due to diminished evanescent waves

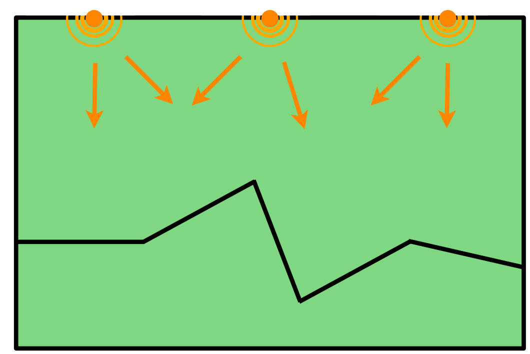
Project Goals

- Employ ideas of compressed sensing
- Deliberately limit signal sampling to reduce computational cost
- L1-minimization recovery reduces blurring due to missing evanescent modes

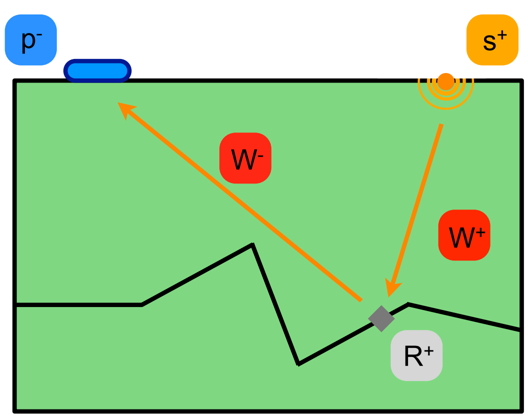




About this Propagator...



Stable Wave Propagator



Stable Wave Propagator

- Property of $W^{\pm}$ is crucial for computation

$$p^- = \int_{x_3 > 0} W^- R^+ W^+ s^+ \, dx_3$$

$$p^- = \sum_{x_3 > 0} W^- R^+ W^+ s^+ \, \Delta x_3$$

Stable Wave Propagator

- No Dip Limitation
- Handles Lateral Variations
- Unconditional Stability
- Low Computational Cost

$W^{\pm}$

One-Way Wave Operator

- Physical behavior of wavefield modeled by coupled differential equation of depth

(Claerbout, 1971; Wapenaar and Berkhout, 1989)

$$\partial_3 \boldsymbol{\mathcal{Q}} + j\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{Q}} = \boldsymbol{\mathcal{D}}$$

wave vector

source vector

- Solution for $\boldsymbol{\mathcal{Q}}$ at any depth

$$\boldsymbol{\mathcal{Q}}(x_3) = \exp(-j\boldsymbol{\mathcal{A}}x_3)\boldsymbol{\mathcal{Q}}_0$$

- Unfortunately this expression is meaningless!

One-Way Wave Operator

- Structure of $\boldsymbol{\mathcal{A}}$ confounds the meaning of its exponentiation, due to it being an operator

$$\boldsymbol{\mathcal{A}} = \begin{pmatrix} 0 & \omega\rho \\ \frac{1}{\omega\rho^{1/2}}(\mathcal{H}_2\rho^{-1/2}) & 0 \end{pmatrix}$$

Two-way Wave Operator

$$\mathcal{H}_2 = k^2(\boldsymbol{x}) + \partial_\mu\partial_\mu$$

- H2 contains information about medium velocity

One-Way Wave Operator

- Decomposition of $\boldsymbol{\mathcal{A}}$ proposed to rectify its usefulness in computation

(Claerbout, 1971; Wapenaar and Berkhout, 1989; de Hoop, 1992)

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{L}}^{-1}$$

Two-way Wave Operator

$$\boldsymbol{\mathcal{H}} = \begin{pmatrix} \mathcal{H}_1 & 0 \\ 0 & -\mathcal{H}_1 \end{pmatrix}$$

$$\mathcal{H}_2 = \mathcal{H}_1\mathcal{H}_1$$

One-Way Wave Operator

- Substitution of $\mathcal{A}$ by its decomposition is performed, and its composition operators is allowed to act on the signal vectors

$$\partial_3 \mathcal{Q} + j\mathcal{A}\mathcal{Q} = \mathcal{D}$$

$$\mathcal{A} = \mathcal{L}\mathcal{H}\mathcal{L}^{-1}$$

$$\partial_3 \boldsymbol{P} + j\boldsymbol{\mathcal{H}}\boldsymbol{P} = \boldsymbol{S} + \boldsymbol{\Theta}\boldsymbol{P}$$

One-way Wave Equation

One-Way Wave Operator

- Solution to one-way wave equation now has the one-way wave operator defined as

$$\mathcal{W}^{\pm}(x_3; x_3') = \exp(\mp j(x_3 - x_3')\mathcal{H}_1)$$

- The new definition is consistent with the standard "migration" model

$$P^- = \sum_{x_3 > 0} W^- \, R^+ \, W^+ \, s^+ \, \Delta x_3$$

Modal Decomposition

- We still need to compute the actual W± Operator
  - this requires structure of $\mathcal{H}_1$

$$\mathcal{H}_2 = \mathcal{H}_1\mathcal{H}_1$$

  - with $\mathcal{H}_2$ defined as

$$\mathcal{H}_2 = k^2(\boldsymbol{x}) + \partial_\mu \partial_\mu$$

  - or, written as a numerical linear operator

$$\underset{\sim}{\mathbf{H}}_2 = \underset{\sim}{\mathbf{C}} + \underset{\sim}{\mathbf{D}}_2.$$

Modal Decomposition

$$\underset{\sim}{\mathbf{H}}_2 = \underset{\sim}{\mathbf{C}} + \underset{\sim}{\mathbf{D}}_2.$$

- Symmetric
- Hermitian
- Self-adjoint

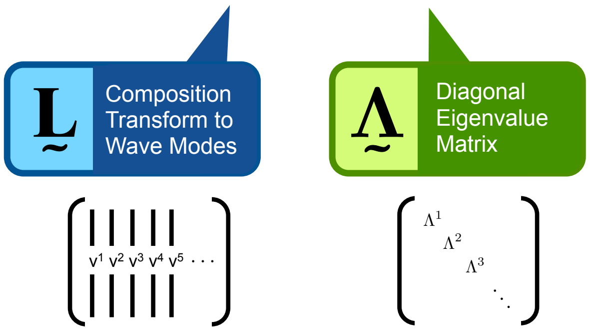
$$\underset{\sim}{\mathbf{C}} = \begin{pmatrix} \left(\frac{\omega}{c_1'}\right)^2 & 0 & \cdots & 0 \\ 0 & \left(\frac{\omega}{c_2'}\right)^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\frac{\omega}{c_M'}\right)^2 \end{pmatrix},$$

$$\underset{\sim}{\mathbf{D}}_2 = \frac{1}{\Delta x_1^2} = \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -2 & 1 \\ 0 & 0 & 0 & \cdots & 1 & -2 \end{pmatrix}$$

Modal Decomposition

- Guaranteed existence of similarity transform decomposition

$$\underset{\sim}{\mathbf{H}}_2 = \underset{\sim}{\mathbf{L}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\mathbf{L}}^{-1} = \underset{\sim}{\mathbf{L}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\mathbf{L}}^{H}$$



Modal Decomposition

- From the structure of $\underset{\sim}{\mathbf{H}}_2$ it is simple to deduce that it's "square root" can be computed as

$$\underset{\sim}{\mathbf{H}}_1 = \underset{\sim}{\mathbf{L}}\underset{\sim}{\boldsymbol{\Lambda}}^{\frac{1}{2}}\underset{\sim}{\mathbf{L}}^{-1} = \underset{\sim}{\mathbf{L}}\underset{\sim}{\boldsymbol{\Lambda}}^{\frac{1}{2}}\underset{\sim}{\mathbf{L}}^{H}$$

- Linear algebra thus allows the propagator to be written in the form:

$$\underset{\sim}{\mathbf{W}}^{\pm}(x_3, x_3') = \underset{\sim}{\mathbf{L}}(x_3') \exp\left\{\mp j(x_3 - x_3')\underset{\sim}{\boldsymbol{\Lambda}}^{\frac{1}{2}}\right\}\underset{\sim}{\mathbf{L}}^{H}(x_3')$$

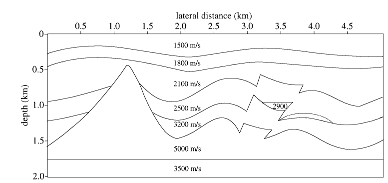
(Grimbergin et. al., 1998)

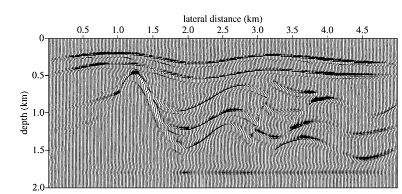
Modal Decomposition

- **Implicit Wavefield Propagation Algorithm**
  - bring signal into frequency domain
  - for each frequency & layer:
    - construct $\underset{\sim}{\mathbf{H}}_2$ operator matrix
    - obtain eigenvalue decomposition of $\underset{\sim}{\mathbf{H}}_2$
    - transform monochromatic signal to eigenvector basis
    - apply phase rotation $\exp\left\{\mp j(x_3 - x_3')\underset{\sim}{\boldsymbol{\Lambda}}^{\frac{1}{2}}\right\}$
    - backward transform signal to space basis
  - combine monochromatic signal & transform back to time domain

Modal Decomposition

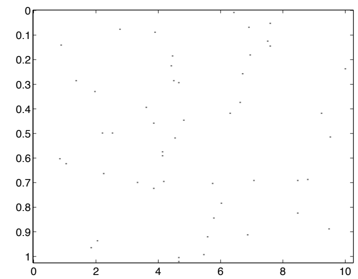
- For propagation examples, refer to Grimbergen et. al. 1998
- Shown to effortlessly handle lateral medium variations without tweaking

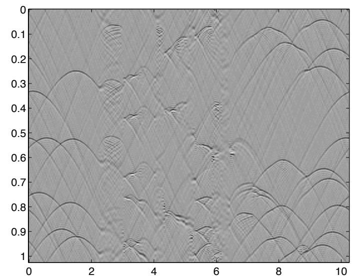




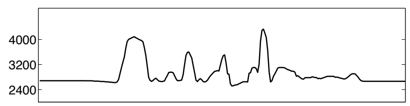
Modal Decomposition

- simple 1-D space/time propagation example





propagated 1.5km down

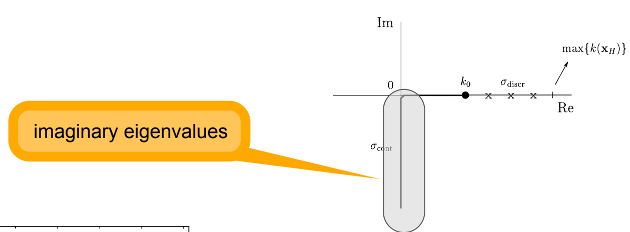


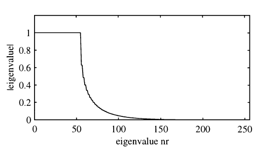
Motivation: Ideal Propagator?

- ✓ No Dip Limitation
- ✓ Handles Lateral Variations
- ✓ Unconditional Stability
- Computational Speed

Motivation: Ideal Propagator?

- spectrum of $\underset{\sim}{\Lambda}$ dictates existence of evanescent wave modes belonging to imaginary eigenvalues

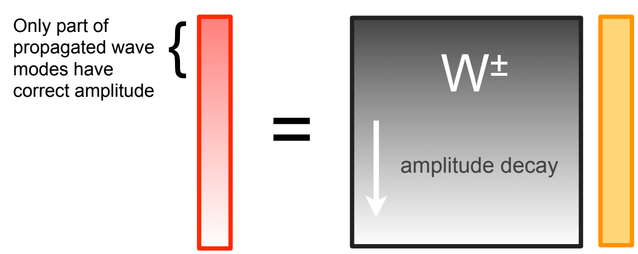




Amplitude of these wave modes decay exponentially as a result

Motivation: Ideal Propagator?

- Our W operator will inevitably be "pseudo-restricted" with a part of the operator having diminished amplitude

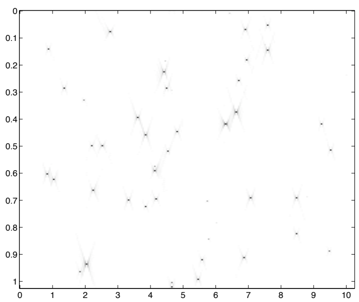


Motivation: Ideal Propagator?

- This causes problems with inverse propagation, defined as Hermitian adjoint

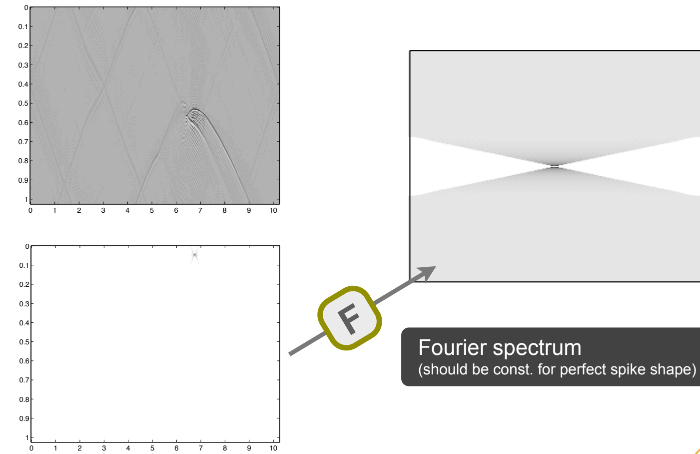
  $$W^T W x$$

  - Evanescent wave modes are not accounted
  - Results in frequency-limited artifact



Inverse propagated wave signal
(should resemble source's perfect spike shape)

Motivation: Ideal Propagator?



Motivation: Ideal Propagator?

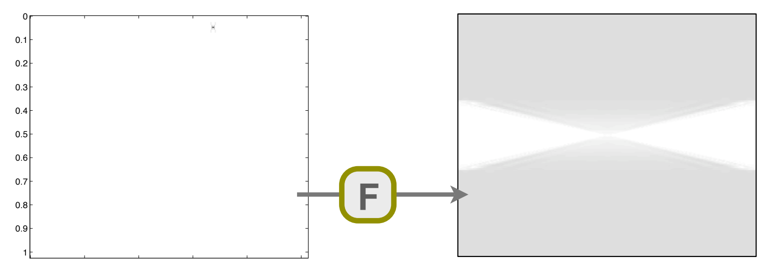
- Inverse propagation can instead be treated as a *least squares* problem to reduce artifact

$$\hat{x} = (A^T A + \epsilon I)^{-1} A^T y$$

- However this must be solved iteratively since the Hessian $(A^T A)$ is ill-conditioned
- This adds a factor of **10x~20x** to the computing cost of each propagation

Motivation: Ideal Propagator?

- However, least-squares do not completely solve the problem of inverse propagation



Inverse propagation by *lsqr()*

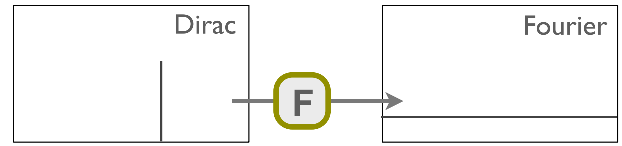
Still missing some frequencies!

Motivation: Ideal Propagator?

- Furthermore, the modal decomposition method is inherently *costly*
  - **synthesis cost:**
    requires solving a full eigenvalue problem with the H2 matrix, which could be $\mathbf{O(n^3)}$ with n being the number of detectors
  - **operation cost:**
    requires a FFT in addition to vector-matrix multiplications which is $\mathbf{O(n^2)}$, with *lsqr* contributing a factor of 10~20 to this cost

Propagation via L1-recovery

- **Fourier** basis is known to be a good measurement basis for sparse recovery due to strong incoherence with Dirac basis (E. Candes, D.L. Donoho)

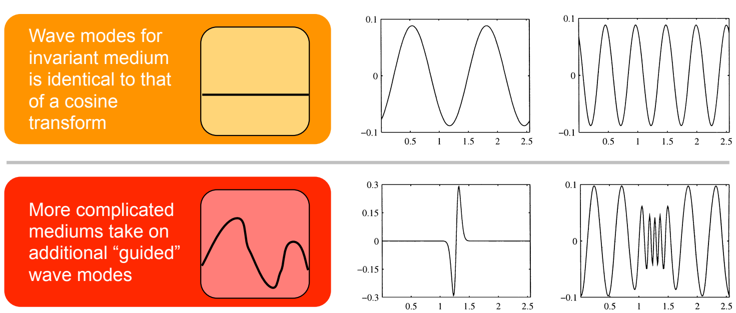


- From UUP we know that it takes only ~5 Fourier wave modes to recover one point spike (disregarding log-like factors)

Inspiration: Wave Modes of H2

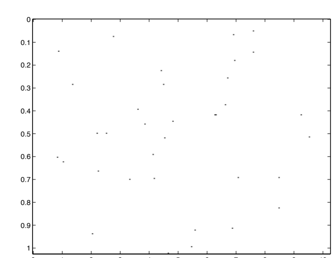
- The wave modes of $\underset{\sim}{\mathbf{H}}_2$ very much resembles a Fourier transform operator's wave modes!

(Grimbergin et. al., 1998)



Propagation via L1-recovery

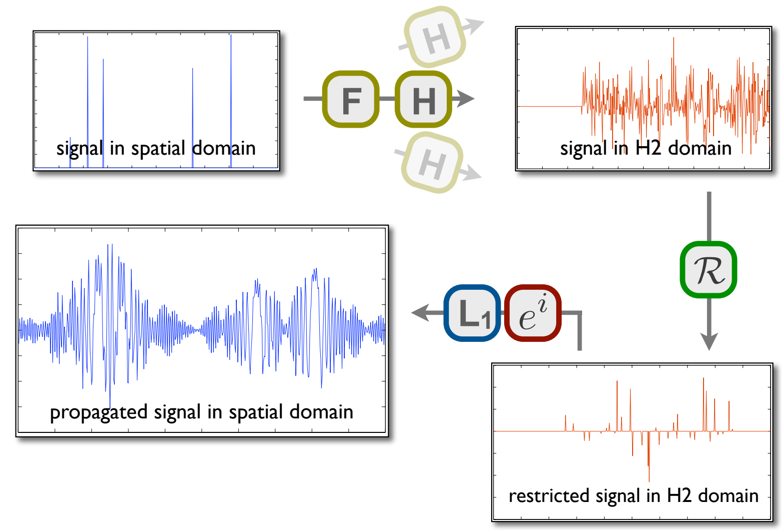
- We can actually directly ignore evanescent wave modes and call it "conveniently restricted out"
- Result is clean spikes without artifacts caused by "incomplete" propagation



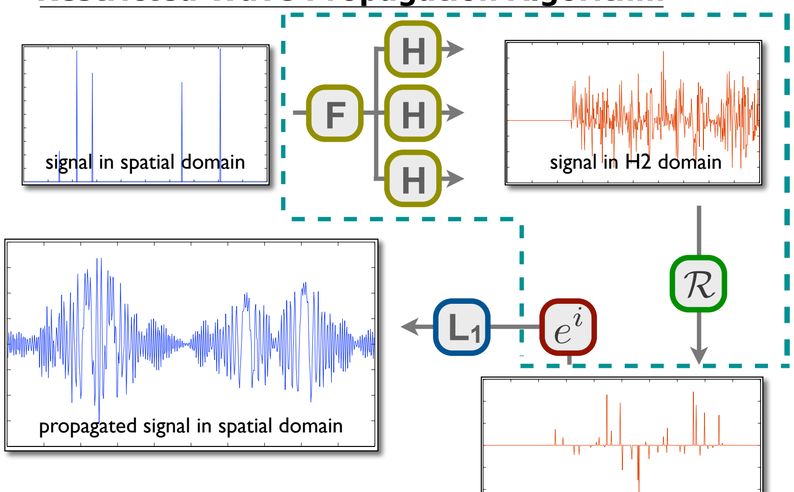
Propagation via L1-recovery

- **Restricted Wave Propagation Algorithm**
  - Decompose signal into freq & H2 wave mode
  - Delete (restrict) most of the signal, for practical cases usually ~90%
  - Construct a much smaller Implicit Wavefield Propagation Algorithm and apply it to restricted signal
  - Use a fast L1-solver to recover the full propagated signal in space/time domain

- **Restricted Wave Propagation Algorithm**



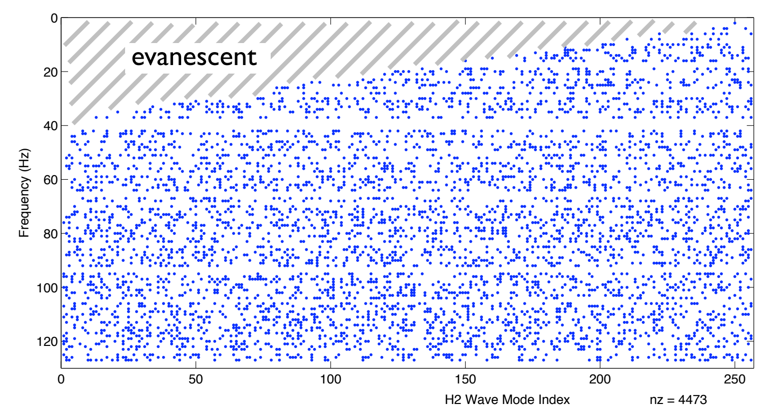
- **Restricted Wave Propagation Algorithm**



Expanding operator size to “force” sparsity

Propagation via L1-recovery

- Restriction index keeps track of restricted signal



Computational Savings

- Reduction in **synthesis cost**
  - a fully restricted frequency eliminates one full eigenvalue problem
  - partially restricted frequencies gain a reduction in the size of the eigenvalue (10% of original size)
- Reduction in **computation cost**
  - Applying the operator now is only O(n), with a factor that is proportional to the fraction of signal surviving restriction

What about the L1-recovery?

- L1-recovery isn't free, which is why we need a fast solver
- **StOMP** can be utilized as a fast approximate L1 solver
- But in reality, any L1 solver can be used as long as it is *fast*

StOMP Computational Costs

- StOMP is approximately equal to 2~5x of an iterative *lsqr* problem. **But**:

Cost of Reduced Operator ≈ Restriction Ratio Cost of Full Operator

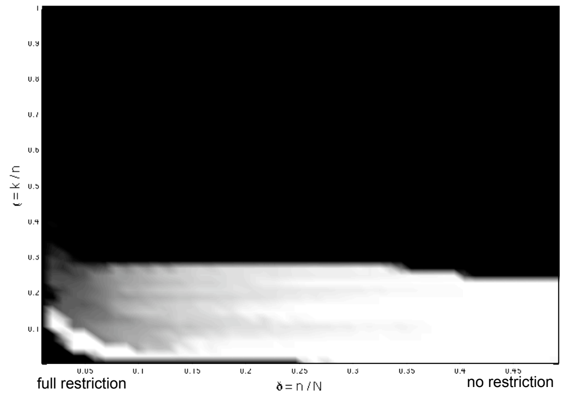
  - i.e., Operating on a signal 10% of the original size will take about 10% of the time taken by a full operator
- StOMP will usually be faster than *lsqr* provided that we restrict more than 80% of the signal

Experimental Results

- Recoverability phase diagrams
- Eigenvalue problem cost reduction

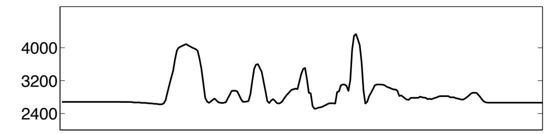
Recoverability Phase Diagrams

- Invariant Medium, 1km down



Hard Problems

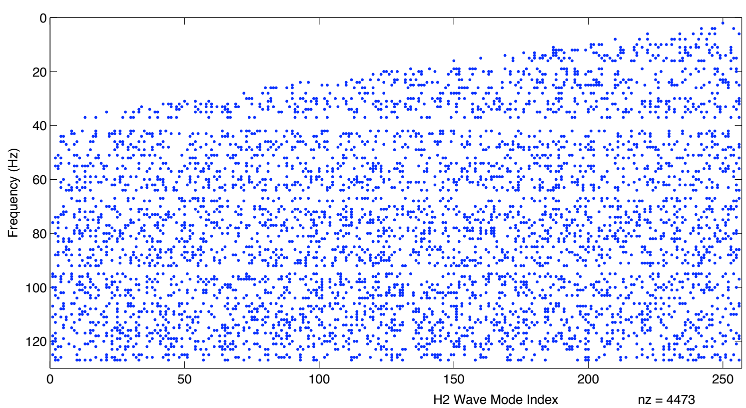
- What do we expect when we inverse-propagate in a “hard” medium?



- Guided wave modes will probably affect recoverability, but hard to predict
- See separate effects of frequency vs. H2 wave mode restriction

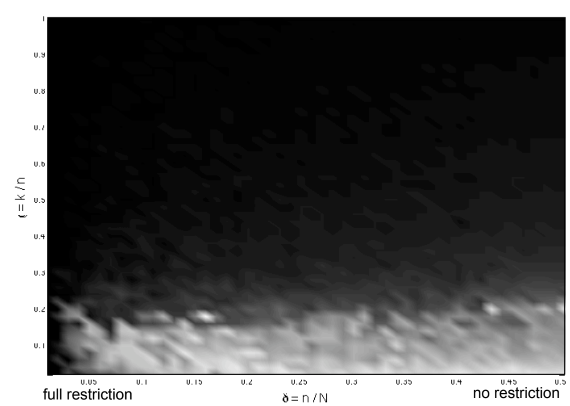
Choosing Restrictions

- Choice of restrictions in frequency and H2 modes



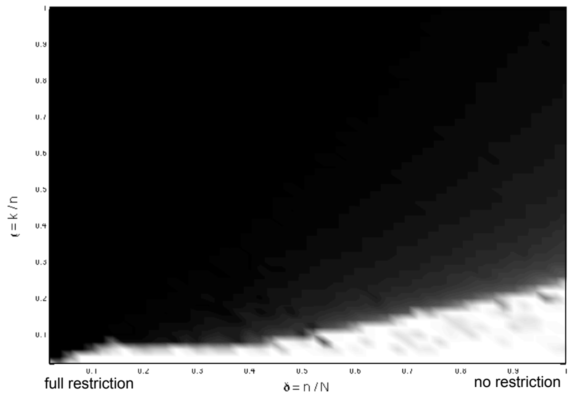
Recoverability Phase Diagrams

- Rapidly Varying Medium, 1km down, freq restriction



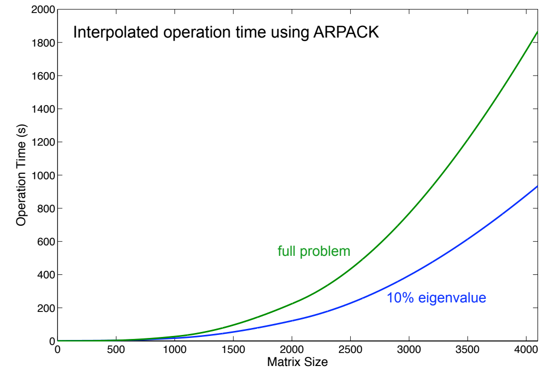
Recoverability Phase Diagrams

- Rapidly Varying Medium, 1km down, H2 restriction



Savings on Eigenvalue Problem

- We additionally save time by computing only a small percentage of eigenvalues



Future Directions

- Optimal Restriction
- Multi-layer Propagation
- Working in curvelet sparsity

Optimal Restriction

- *Restricting whole frequencies* eliminate entire eigenvalue problems, but give less predictable results
- *Pure random restriction* gives predictable results but still require solving eigenvalues

An optimal restriction scheme is proposed to exist

Multi-Layer Propagation

- Multi-Layer propagation is the only way to deal with vertical velocity variations

$$P^- = \sum_{x_3 > 0} W^- W^- W^- R^+ W^+ W^+ W^+ s^+ \, \Delta x_3$$

- Decaying evanescent waves make deep propagations through many layers difficult

> Possible non-linear inv. propagation using L1 Solvers

Curvelet Sparsity

- Stop working in broadband and start working in Curvelet sparsity
- Utilizing Curvelet sparsity is possible by incorporating FDCT into the operator

> Maintain signal sparsity with new sparsity basis

Conclusions

- Reformat inverse propagation and therefore imaging as a sparse recovery problem
- Remove problem with evanescent wave modes
- Faster (or at least as fast) to compute as *lsqr*
- Loosened memory requirements
- Improves with future fast L1 solver

Thanks for your time!