Analysis vs synthesis in weighted sparse recovery

Brock Hargreaves and Özgür Yilmaz, Dec 3rd 2012

SLIM
University of British Columbia

Overview

- When using a redundant transform, such as curvelets, there is choice in how to solve the sparse recovery problem, namely synthesis and analysis formulations which I'll introduce shortly.

- Support information can be incorporated in both formulations and a novel weighting scheme is introduced

- Analysis is a viable candidate and can outperform synthesis (but there is no free lunch, yet)

Examples of redundant transforms

- Curvelet frames (Primary transform of interest for us)
- Gabor frames
- Harmonic frames
- Wavelet frames
- Concatenation of bases/frames

Synthesis formulation

Assume our signal f admits a sparse representation with respect to S^H where S is a redundant transform

$$f = S^H x \quad , \quad \text{x is sparse}$$

<u>Major issue</u>: S redundant $\Longrightarrow$ Uniqueness of above equation lost. Possibly many sparse solutions.

$$\hat{x} = \arg\min_x ||x||_1 \quad \text{s.t.} \quad b = \underbrace{RS^H}x$$

$\leftarrow$ <u>Compressive sensing</u>
Rauhut et al 2008, Randall 2009

Error and reconstruction

$$||\hat{x} - x||_2 \quad \leq \quad C_1 \frac{||x - x_s||_1}{\sqrt{s}}$$

x_s : Best s-term approximation of x

$$\hat{f} = S^H \hat{x}$$

$$\Longrightarrow \quad ||\hat{f} - f||_2 \quad \leq \quad C_1 \frac{||x - x_s||_1}{\sqrt{s}}$$

C_1 : Constant

Remark:

RS^H will rarely satisfy the theoretical conditions required for these bounds to hold when S^H is redundant, however these conditions are only sufficient.

Synthesis weighting

$$\hat{x} = \arg\min_x ||x||_1 \quad \text{s.t.} \quad b = RS^H x$$

We solve instead,

$$\hat{x} = \arg\min_x ||Wx||_1 \quad \text{s.t.} \quad b = RS^H x$$ (Hassan's talk)

W is a diagonal matrix, whose entries selectively penalize different coefficients in the solution:

$$W_{i,i} = \begin{cases} \omega & i \in \text{estimated support of x} \\ 1, & i \notin \text{estimated support of x} \end{cases}$$

Analysis formulation

$$\hat{f} = \arg\min_{f} ||(S^H)^{\dagger} f||_1 \quad \text{s.t.} \quad b = Rf$$

Let S be the curvelet transform, then $S^H S = I$ because curvelets are tight frames:

$$\hat{f} = \arg\min_{f} ||Sf||_1 \ \text{ s.t. } \ b = Rf$$

S is called the analysis operator, equivalently the forward transform

Error (Candès, Eldar, Needell, Randall 2010 and Shidong Li 2012):

$$||\hat{f} - f||_2 \quad \leq \quad C \frac{||Sf - (Sf)_s||_1}{\sqrt{s}} \quad ,$$

$(Sf)_s$: Best s-term approximation of Sf

C : Constant

Weighting: Strategy 1

$$\hat{f} = \arg\min_f ||W(Sf)||_1 \quad \text{s.t.} \quad b = Rf$$

$$W_{i,i} = \begin{cases} \omega & i \in \text{estimated support of x} \\ 1, & i \notin \text{estimated support of x} \end{cases}$$

This weighting is interpreted the exact same way as in synthesis

See: Candès, Walkin, Boyd 2008 and Candès, Needell 2010

Weighting: Strategy 2

Recall that non-simplified version:

$$\hat{f} = \arg\min_f ||(S^H)^\dagger f||_1 \quad \text{s.t.} \quad b = Rf$$

New weighting strategy (Hargreaves-Yilmaz, November 2012):

$$\hat{f} = \arg\min_f ||(S^H \tilde{W})^\dagger f||_1 \quad \text{s.t.} \quad b = Rf$$

$$\tilde{W}_{i,i} = \begin{cases} 1 & i \in \text{estimated support of x} \\ \omega, & i \notin \text{estimated support of x} \end{cases}$$

Possible interpretation: $\tilde{W}$ acting on the columns of S^T influences the subspace in which our solution vector lives (Cosparsity? See Tim Lin's talk.)

Why are the weights inverted?

A simple motivation:

Suppose B is an orthonormal basis, and x sparse

$$f = Bx$$

If we have a perfect estimate, our weight matrix preserves this:

$$f = BWx$$

So in the imperfect setting, we attempt the same:

$$f = BWx \qquad \Longrightarrow \quad x = (BW)^{-1} f = W^{-1} B^{-1} f$$

On support: small

Off support: big $\Longrightarrow$ W should be the opposite

Review: Analysis vs synthesis

Assumption:	x is sparse	Sf is sparse
Problem:	$\min_x \|\|x\|\|_1$	$\min_f \|\|Sf\|\|_1$
Constraint:	$y = RS^H x$	$y = Rf$
$\|\|\hat{f} - f\|\|_2$:	$C_1 \frac{\|\|x - x_s\|\|_1}{\sqrt{s}}$	$C \frac{\|\|Sf - (Sf)_s\|\|_1}{\sqrt{s}}$

Review: Analysis vs synthesis

- Analysis optimizes over signal space, synthesis optimizes over coefficient space
- Not equivalent when S is not a basis and there is an observed performance gap (Elad et al, Nam et al)
- Academic community has not agreed on whether one is better than the other, unless referring to a specific application(Our study: seismic trace interpolation)

Seismic Results

Disclaimer: New weighting strategy is brand new (Nov 2012)

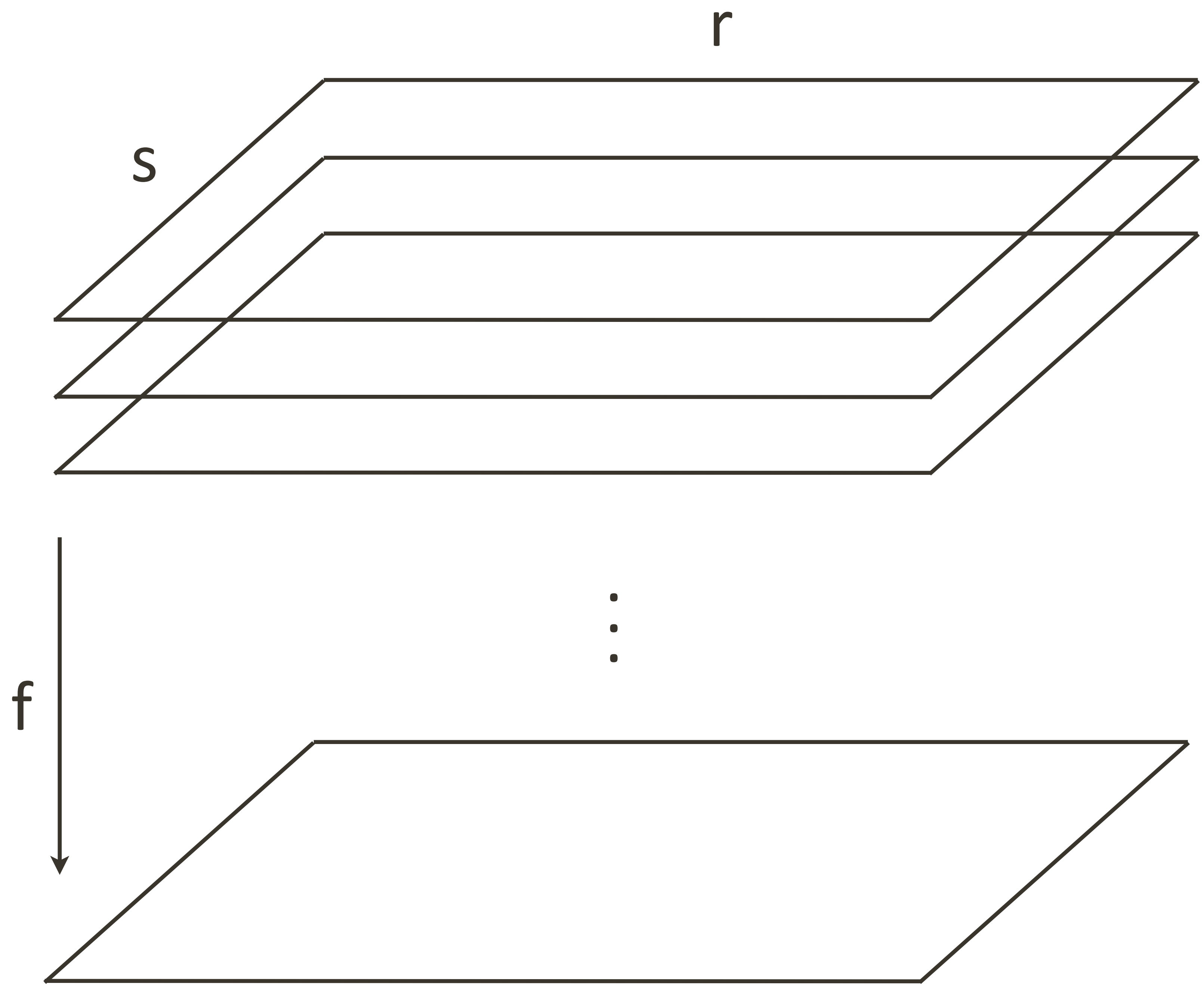
- Gulf of Suez data set: D = 1024 x 256 x 256
  - D = (1:2:300,1:2:100,1:2:100)
  - But we believe we will be able to scale this up
- NESTA: A Fast and Accurate First-order Method for Sparse Recovery (Bobin, Becker, Candès, 2011)
  - One of few publicly available solvers than can solve both analysis and synthesis
  - Mainly driven by a single smoothing parameter, which dictates the tradeoff between speed and accuracy

Frequency content as a support estimate

Recover first frequency slice
└→ Estimate support

Recover second frequency slice using previous support estimate as a weight

⋮

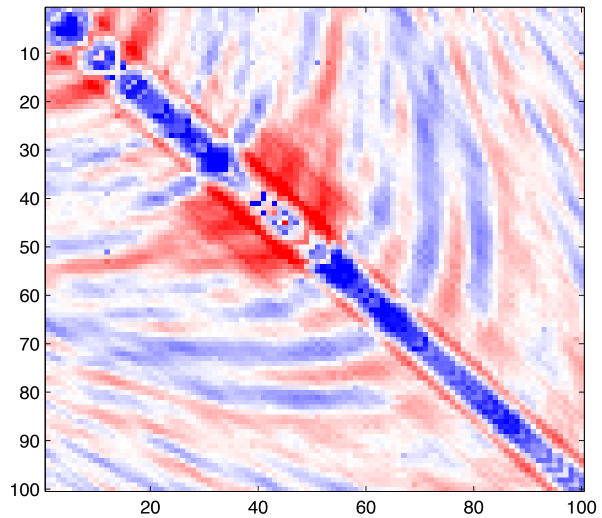


Frequency slice recovery

100 receivers, 100 sources, 2nd frequency slice

Actual

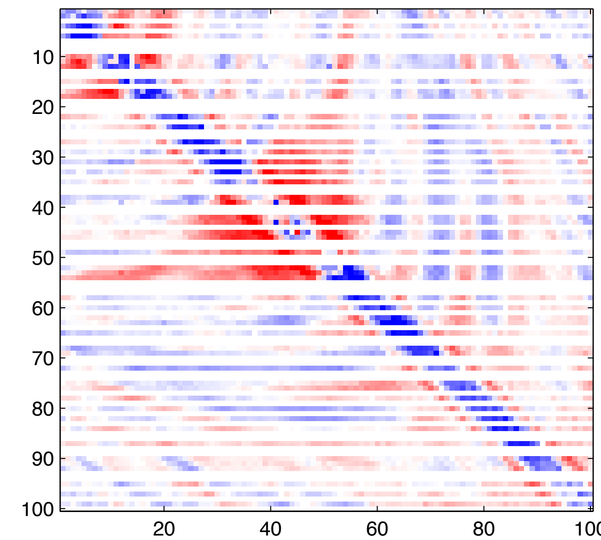
Receiver



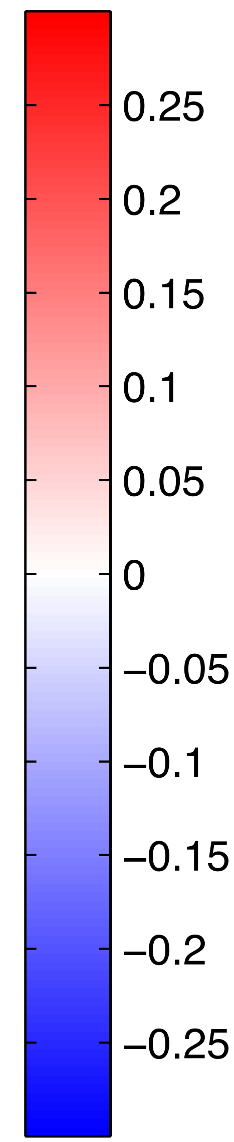
Source

Masked(50%)

Receiver

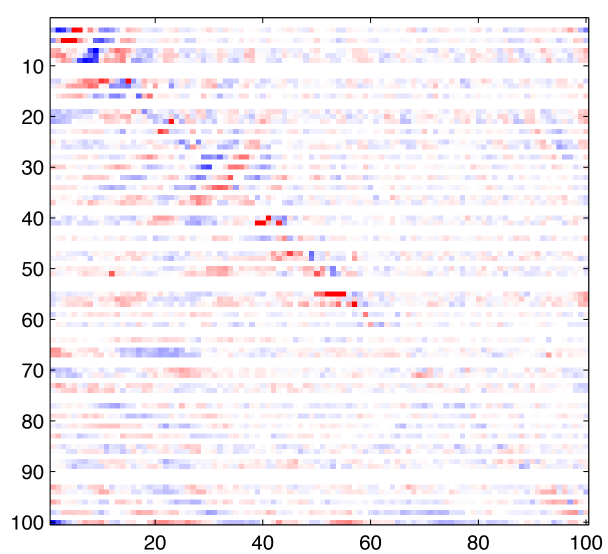


Source



Difference: Weighted Analysis

Receiver

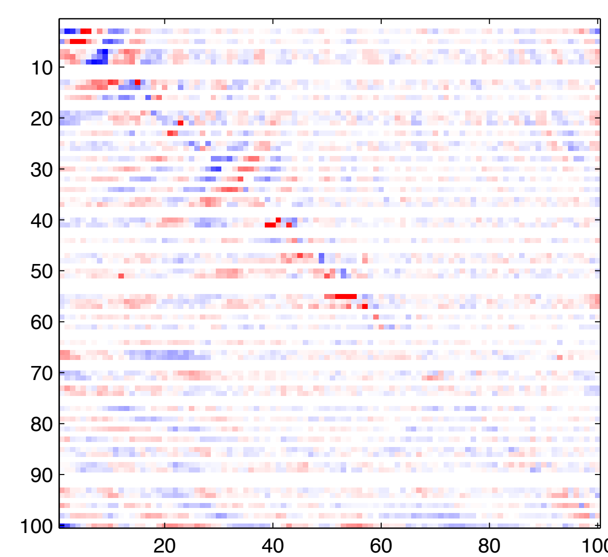


Source

SNR: 10.28

Difference: Weighted Synthesis

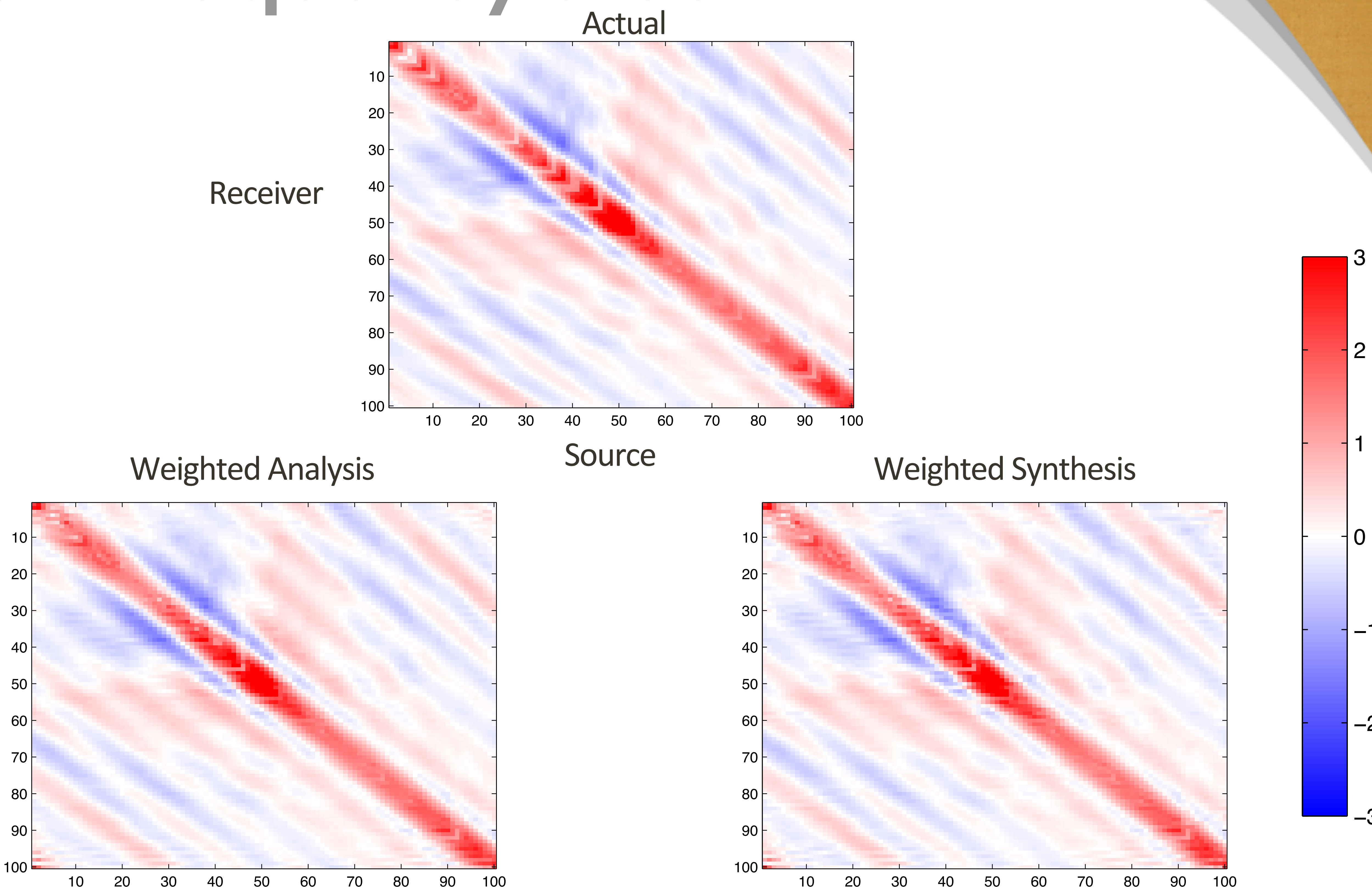
Receiver



Source

SNR: 9.97

15th Frequency slice



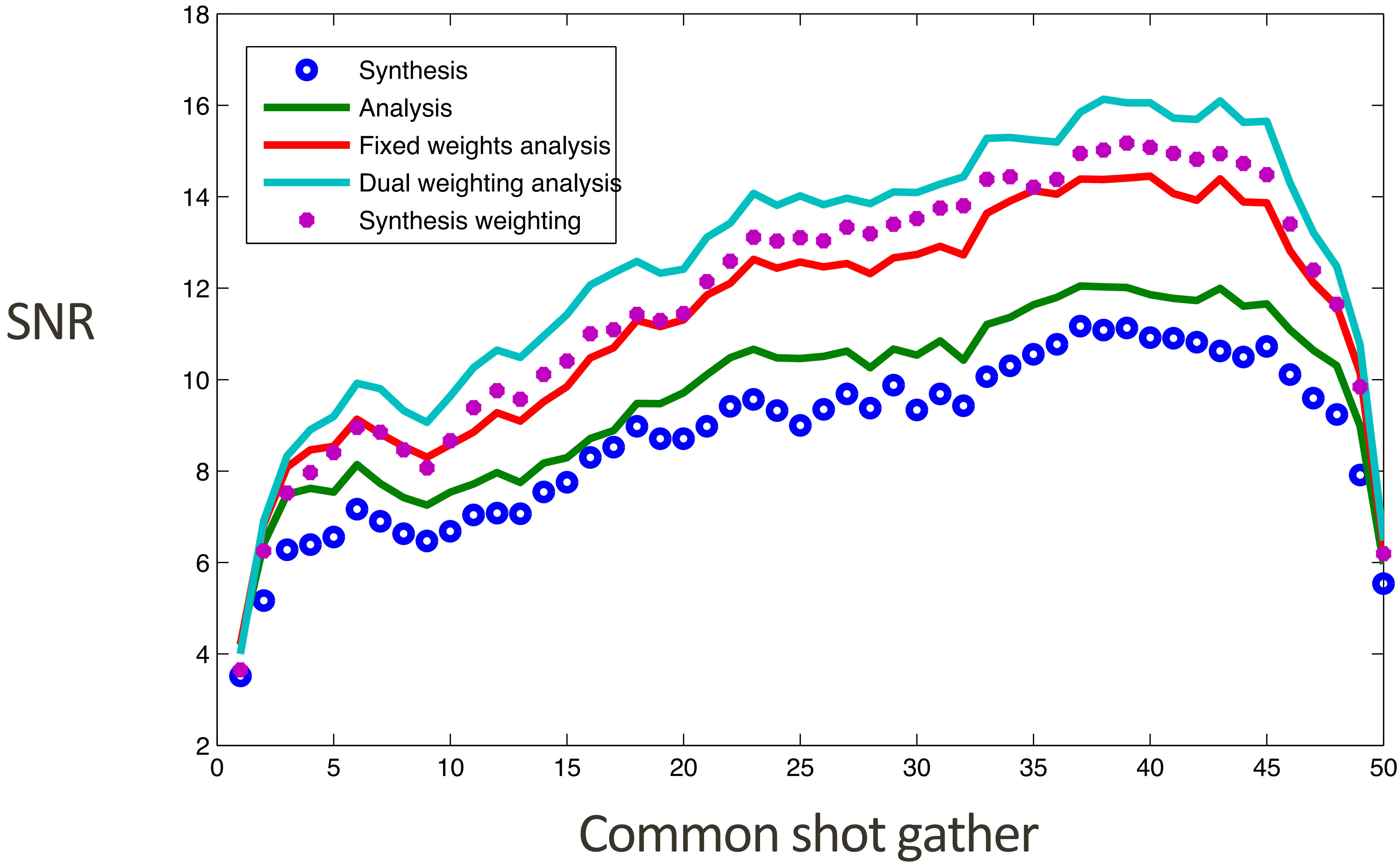
SNR: 20.66

SNR: 19.16

Comparison of algorithms

SLIM

Recovery in source-receiver domain(50 sources, 50 receivers)



But the story isn't over: Choice of dual and optimality

There are infinitely many choices of the analysis operator, why do we choose the canonical one? Optimize over all possible duals as well:

$$\min_{f, \tilde{D}^* D = I} ||\tilde{D}^* f||_1 \quad \text{s.t.} \quad y = Af$$

Paper: Compressed sensing with general frames via Optimal-Dual-Based l1 analysis

Authors: Liu, Mi, Li

Remark: We can also incorporate our weighting strategy here

Remarks and conclusion

- In the new weighting scheme anytime we need to see how the pseudoinverse acts on a vector we need to call LSQR. This is the main bottleneck.

- Conclusion:
  - ***When using NESTA with same parameters***:
    The novel weighting technique introduced performs the best with the assumption that the pseudoinverse is approximated well enough

Future work

- Improve computation time so we can handle large problem sizes and test in various domains such as midpoint-offset
- Weighted analysis with $\ell_p,\ p < 1$ (Ghadermarzy's talk)
- Analysis operator learning? (M. Elad et al 2012/2013)
- Weighted version of optimal analysis on seismic
- Prove theoretical results

References

[1] Mansour, Herrmann, Yilmaz. “Improved wavefield reconstruction via weighted one-norm minimization”, 2012

[2] Liu, Mi, Li. “Compressed sensing with General Frames via Optimal-Dual-Based l1-analysis”, 2012

[3] Nam, Davies, Elad, Gribonval. “The Cosparse Analysis Model and Algorithms”, 2012

[4] Friedlander, Mansour, Saab, Yilmaz. “Recovering compressively sampled signals using partial support information”, 2011

[5] Candès, Eldar, Needell, Randall. “Compressed sensing with coherent and redundant dictionaries”, 2010

[6] Randall. “Sparse Recovery via Convex Optimization”, 2009

[7] Becker, Bobin, Candès. “NESTA: A fast and accurate first-order method for sparse recovery”, 2009

[8] Rauhut, Schnass, Vandergheynst. “Compressed Sensing and Redundant Dictionaries”, 2008

[9] M.Elad, P. Milanfar, R. Rubinstein. “Analysis vs synthesis in signal priors”, 2007

Acknowledgements

- Ozgur Yilmaz
- Curt Da Silva
- Haneet Wason and Rajiv Kumar
- Henryk, Thomas, Ian, and Miranda
- The whole SLIM family

SINBAD

NSERC CRSNG

This work was in part financially supported by the the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, BP, Chevron, ConocoPhillips, Petrobras, PGS, Total SA, and WesternGeco.