NMR Measurement of T1-T2 Spectra with Partial Measurements using Compressive Sensing

Alex Cloninger

Norbert Wiener Center
Department of Mathematics
University of Maryland, College Park
http://www.norbertwiener.umd.edu

Norbert Wiener Center for Harmonic Analysis and Applications

Outline

Outline

Basics of Nuclear Magnetic Resonance

- We will consider multidimensional correlations as measured by NMR
- Specifically look at T1-T2 relaxation properties
  - Relaxation measures how fast magnetic spins ”forget” orientation and return to equalibrium
  - T1 is decay constant for z-component of nuclear spin
  - T2 is decay constant for xy-component perpendicular to magnetic field
- Analysis of multidimensional correlations requires multidimensional inverse Laplace transform
- Due to ill-conditioning of 2D ILT, require a large number of measurements
- Problem is getting T1-T2 map is incredibly slow

Spin-Echo Pulse

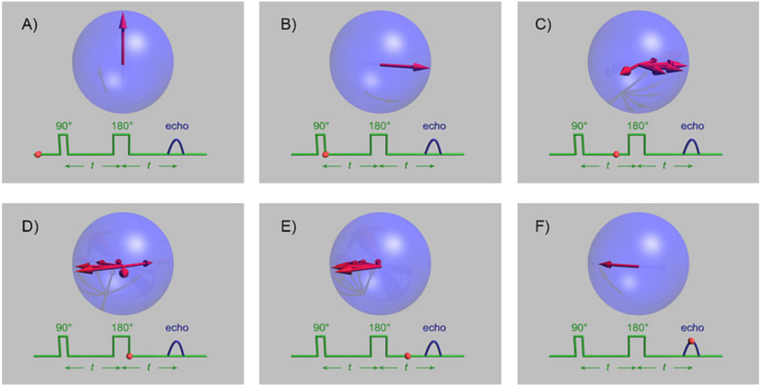
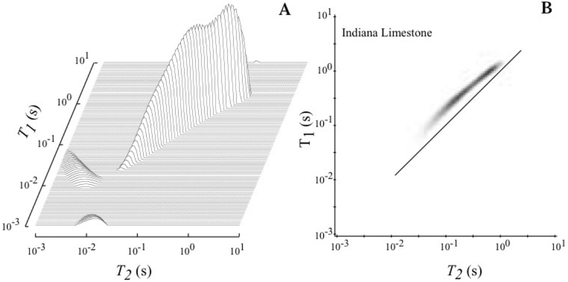


Figure: Animation of Spin-Echo Pulse

T1-T2 Correlations



Math Behind NMR

- Echo measurements are related to T1-T2 correlations via Laplace Transform

$$M(\tau_1, \tau_2) = \int \int (1 - 2e^{-\tau_1/T_1}) e^{-\tau_2/T_2} \mathcal{F}(T_1, T_2) dT_1 dT_2 + E(\tau_1, \tau_2)$$

- We'll consider more general 2D Fredholm Integral

$$M(\tau_1, \tau_2) = \int \int k_1(\tau_1, T_1) k_2(\tau_2, T_2) \mathcal{F}(T_1, T_2) dT_1 dT_2 + E(\tau_1, \tau_2)$$

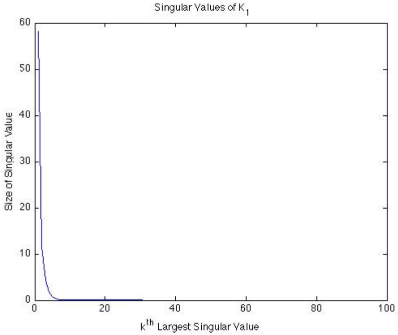
where $E(\tau_1, \tau_2) \sim \mathcal{N}(0, \epsilon)$

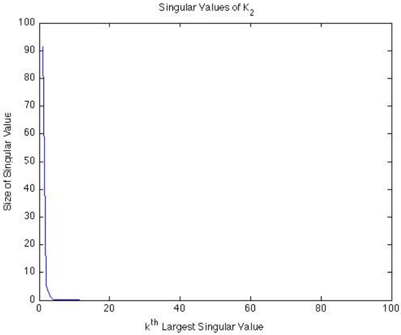
- Discretize to

$$\mathbf{M} = \mathbf{K_1 F K_2'} + \mathbf{E}$$

Problems with Inverse Laplace Transform

- $k_1(\tau 1, T_1)$ and $k_2(\tau_2, T_2)$ are smooth continuous
  - Means $\mathbf{K_1}$ and $\mathbf{K_2}$ are ill-conditioned
  - Makes inversion difficult





Outline

Algorithm for Approximation of $\mathcal{F}(T_1, T_2)$

Definition (Objective Function for Recovery)

Wish to find minimizer to

$$\min_{F \geq 0} \|M - K_1 F K_2^*\|_2 + \alpha \|F\|_2$$

- Because inversion unstable, $\alpha\|F\|_2$ term smooths solution
  - Called Tikhonov regularization with parameter α
- Possible to determine choice of α that minimizes bias
- Computationally unwieldy

Norbert Wiener Center for Harmonic Analysis and Applications

Outline of Preexisting Algorithm

1. Reduce the dimension of the problem in order to make the minimization manageable
2. Holding alpha fixed, solve the Tikhonov regularization problem
3. Having solved the minimization problem, determine a more optimal alpha to use in the regularization problem
4. Loop between Steps 2 & 3 until convergence upon the optimal alpha

Step 1: Dimension Reduction

- Assume M is $N_1 \times N_2$ data matrix
- To reduce the dimension of the problem, let

$$K_1 = U_1 \Sigma_1 V_1^*, \quad K_2 = U_2 \Sigma_2 V_2^*$$

where $\Sigma_1 \in \mathbb{R}^{s_1 \times s_1}$ and $\Sigma_2 \in \mathbb{R}^{s_2 \times s_2}$.

- Choose s_1, s_2 to account for 99% of the energy
- Because of fast decay of kernels, $s_1 \ll N_1$ and $s_2 \ll N_2$.

Step 1: Dimension Reduction

- Change objective function by

$$\begin{aligned} &\min_{F\geq 0} \|U_1^*(M - K_1 F K_2^*)U_2\|_2 + \alpha\|F\|_2 \\ &= \min_{F\geq 0} \|U_1^* M U_2 - \Sigma_1 V_1^* F V_2 \Sigma_2\|_2 + \alpha\|F\|_2 \\ &= \min_{F\geq 0} \|\widetilde{M} - \widetilde{K}_1 F \widetilde{K}_2^*\|_2 + \alpha\|F\|_2. \end{aligned}$$

  - $\widetilde{K}_i = \Sigma_i V_i^*$
  - $\widetilde{M} = U_1^* M U_2 \in \mathbb{R}^{s_1 \times s_2}$
- Note that now minimization only depends on $\widetilde{M}$

Step 2: Tikhonov Regularization

- Now need to solve minimization with compressed data $\widetilde{M}$

$$\min_{F \geq 0} \|\widetilde{M} - \widetilde{K}_1 F \widetilde{K}_2^* \|_2 + \alpha \|F\|_2.$$

- Reform problem into unconstraned optimization for

$$f = \max(0, (\widetilde{K_1} \otimes \widetilde{K_2})' c), \quad c \in \mathbb{R}^{s_1 \times s_2}.$$

- Solve new unconstrained problem using Newton's method

Step 3: Choosing Optimal Alpha

There are multiple ways to choose the optimal alpha:

1. Choose new alpha to be

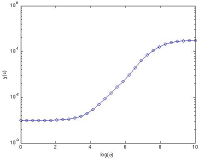
$$\alpha_{new} = \frac{\sqrt{s_1 s_2}}{\|c\|}$$

2. After determining F_α for fixed α, calculate the "fit error"

$$\chi(\alpha) = \frac{\|M - K_1 F_\alpha K_2^*\|_F}{\sqrt{N_1 N_2 - 1}},$$

   which is the standard deviation of the noise in reconstruction.

   - Choose α such that $\frac{d \log \chi}{d \log \alpha} = .1$



Outline

Intuition Behind Compressive Sensing

- Collected $N_1 \times N_2$ data points in M, then "compressed" to $s_1 \times s_2$ matrix $\widetilde{M}$.
    - In practice, $\frac{s_1 s_2}{N_1 N_2} \approx 1\%$

Key Question

Why collect large amounts of data only to throw away over 99% of it?

- At end of day, only $\widetilde{M}$ matters in minimization problem
    - If we had way of measuring $\widetilde{M}$ directly, speedup would be massive

Basic ideas

- Since post-sensing compression is inefficient, why not compress while sensing to boost efficiency?
- Key is that observations in traditional sensing are of form

$$M_{i,j} = \langle M, \delta_{i,j} \rangle.$$

- **Question**: which functions should replace $\delta_{i,j}$ in order to minimize the number of samples needed for reconstruction?
- **Answer**: they should not match image structures; they should mimic the behavior of random noise.

Incoherent Measurements of $\widetilde{M}$

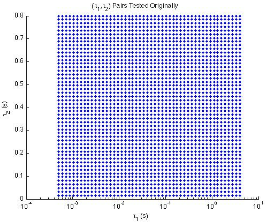
- Remember that $M = U_1 \widetilde{M} U_2'$, so

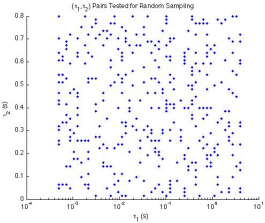
$$\begin{aligned} M_{i,j} &= u_i \widetilde{M} v_j' \\ &= \langle \widetilde{M}, u_i' v_j \rangle. \end{aligned}$$

  where u_i is the i^{th} row of U_1 and v_j is the j^{th} row of U_2
- By structure of U_1 and U_2, rows u_i and v_i are very *incoherent*
  - Incoherent basically means energy is spread out over all elements rather than concentrated
- Because of this, may be possible to observe small number of entries and recover $\widetilde{M}$
  - Even if number of measurements greater than size of $\widetilde{M}$, still could be much less than $N_1 N_2$

Random Sensing

- Consider observing a small number of elements of M randomly
  - We will call set of elements observed Ω

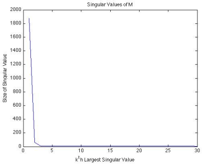




- Fewer measurements directly reduces time spent collecting
- Problem is using these measurements to reconstruct $\widetilde{M}$

Singular Values and Rank of $\widetilde{M}$

- Measurements of $\widetilde{M}$ are underdetermined and noisy.
  - Need a way to bring in more a priori knowledge of $\widetilde{M}$
- Because of ill-conditioned K_1 and K_2, singular values of $\widetilde{M}$ decay quickly
  - Can be closely approximated by only first $r \ll \min(s_1, s_2)$ singular values



Possible to recover M as solution to

$$\begin{array}{ll} \arg\min\limits_{X} & \operatorname{rank}(X) \\ \text{such that} & u_i X v_j' = M_{i,j},\ (i,j) \in \Omega \end{array}$$

Norbert Wiener Center for Harmonic Analysis and Applications

Problems and Reformulation

Two problems:

- Rank minimization is computationally infeasible

Definition (Nuclear Norm)

Let $\sigma_i(M)$ be the i^{th} largest singular value of M. If $\text{rank}(M) = r$, then

$$\|M\|_* := \sum_{i=1}^{r} \sigma_i(M)$$

- Measurements aren't perfect (there is noise)

Definition (Sampling Operator)

For a set of measurements $\{\phi_i\}_{i \in J}$, then for $\Omega \subset J$ the sampling operator is

$$\mathcal{R}_\Omega : \mathbb{C}^{s_1 \times s_2} \to \mathbb{C}^m,$$
$$(\mathcal{R}_\Omega(X))_j = \langle \phi_j, X \rangle, \; j \in \Omega$$

Problems and Reformulation

Recovery Algorithm

Let $y = \mathcal{R}_\Omega(\widetilde{M}) + e$, where $\|e\|_2 < \epsilon$. Wish to recover $\widetilde{M}$ by solving

$$\begin{array}{ll} \arg \min\limits_{X \in \mathbb{R}^{s_1 \times s_2}} & \|X\|_* \\ \text{such that} & \|\mathcal{R}_\Omega(X) - y\|_2 < \epsilon \end{array} \qquad \text{(P*)}$$

- Need to establish that (P*) with these measurements will recover $\widetilde{M}$

Definition

A Parseval tight frame for a d dimensional Hilbert space $\mathcal{H}$ is a collection of elements $\{\phi_j\}_{j\in J} \subset \mathcal{H}$ such that

$$\sum_{j\in J} |\langle f, \phi_j \rangle|^2 = \|f\|^2, \quad \forall f \in H. \qquad (1)$$

Reconstruction and Noise Bounds

- Because u_i and v_j are rows of U_1 and U_2 (who's columns are orthogonal), these measurements form a *tight frame*

Definition

A *Fourier-type Parseval tight frame with incoherence* μ is a Parseval tight frame $\{\phi_j\}_{j\in J}$ on $\mathbb{C}^{d\times d}$ with operator norm satisfying

$$\|\phi_j\|^2 \leq \nu \frac{d}{|J|}, \quad \forall j \in J. \tag{2}$$

- This definition was introduced by David Gross for orthonormal basis
- Can be thought of as singular values bounded in L^∞ norm

Theorem (Cloninger)

Let the measurements in (P) be a Fourier-type Parseval tight frame of size $|J|$, and the true matrix of interest be denoted $\widetilde{M}_0$. If the number of measurements m satisfies*

$$m \geq C\nu nr \log^5 n \log |J|,$$

where ν measures the incoherence of the measurements and $n = \max(s_1, s_2)$, then the result of (P), denoted $\widehat{M}$, will satisfy*

$$\|\widetilde{M}_0 - \widehat{M}\|_F \leq C_1 \frac{\|\widetilde{M}_0 - \widetilde{M}_{0,r}\|_*}{\sqrt{r}} + C_2 \sqrt{\frac{|J|}{m}}\epsilon,$$

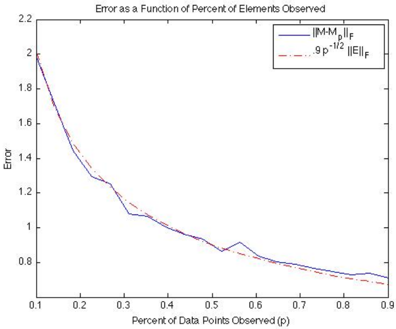
where $\widetilde{M}_{0,r}$ is the best rank r approximation of $\widetilde{M}_0$.

Noise in Practice

- Keep in mind, ϵ is the measure of noise between the ideal matrix $\widetilde{M}_0$ and the compressed data $\widetilde{M}$

$$\|\widetilde{M}_0 - \widetilde{M}\|_F \leq \epsilon$$

- C_2 is a very small constant, and in practice $\sqrt{\frac{N_1 N_2}{m}} \leq \sqrt{5}$



Proof of Theorem

Definition

Let $\mathcal{R}_\Omega : \mathbb{R}^{s_1 \times s_2} \to \mathbb{R}^m$ be a linear operator of measurements. $\mathcal{A}$ satisfies the *restricted isometry property* (RIP) of order r if there exists some small δ_r such that

$$(1-\delta_r)\|X\|_F \leq \|\mathcal{R}_\Omega(X)\|_2 \leq (1+\delta_r)\|X\|_F$$

for all matrices X of rank r.

Theorem (Candès, Fazel)

Let X_0 be an arbitrary matrix in $\mathbb{C}^{m\times n}$ and assume $\delta_{5r} < 1/10$. If the measurements satisfy RIP, then $\widehat{X}$ obtained from solving (P) obeys*

$$\|\widehat{X} - X_0\|_F \leq C_0 \frac{\|X_0 - X_{0,r}\|_*}{\sqrt{r}} + C_1\epsilon,$$

where C_0, C_1 are small constants depending only on the isometry constant.

Proof of Theorem

Lemma

Let $\mathcal{R}_\Omega$ be defined as before. Fix some $0 < \delta < 1$. Let m satisfy

$$m \geq C\nu rn \log^5 n \cdot \log|J|, \tag{3}$$

where C only depends on δ like $C = O\left(1/\delta^2\right)$. Then with high probability, $\sqrt{\frac{|J|}{m}}\mathcal{R}_\Omega$ satisfies the RIP of rank r with isometry constant δ. Furthermore, the probability of failure is exponentially small in $\delta^2 C$.

Since, $\sqrt{\frac{|J|}{m}}\mathcal{R}_\Omega$ satisfies the RIP, can reform (P*) to say

$$\begin{array}{ll} \arg\min & \|X\|_* \\ \text{such that} & \left\|\sqrt{\frac{|J|}{m}}\mathcal{R}_\Omega(X) - \sqrt{\frac{|J|}{m}}y\right\|_2 < \sqrt{\frac{|J|}{m}}\epsilon \end{array}$$

Outline of Proof

- Can restate RIP as

$$\epsilon_r(\mathcal{A}) = \sup_{X \in U_2} |\langle X, (\mathcal{A}^*\mathcal{A} - \mathcal{I})X\rangle| \leq 2\delta - \delta^2$$

where $U_2 = \{X \in \mathbb{C}^{s_1 \times s_2} : \|X\|_F \leq 1, \|X\|_* \leq \sqrt{r}\|X\|_F\}$

- Notation:

Norm: $\|\mathcal{M}\|_{(r)} = \sup_{X \in U_2} |\langle X, \mathcal{M}X\rangle|$; Simplify Terms: $\mathcal{A} = \sqrt{\frac{|J|}{m}}\mathcal{R}_\Omega$

$$\begin{aligned}
\mathbb{E}\epsilon_r(\mathcal{A}) &= \mathbb{E}\,\|\mathcal{A}^*\mathcal{A} - \mathcal{I}\|_{(r)} \\
&\leq \mathbb{E}_\Omega \mathbb{E}_\epsilon \left\| \sum \epsilon_i (\phi_i^* \phi_i - (\phi_i')^* \phi_i') \frac{|J|}{m} \right\|_{(r)} \\
&\leq 2\frac{n}{m} \mathbb{E}_\Omega \mathbb{E}_\epsilon \left\| \sum \epsilon_i \sqrt{\frac{|J|}{n}} \phi_i^* \phi_i \sqrt{\frac{|J|}{n}} \right\|_{(r)}
\end{aligned}$$

Outline of Proof

Lemma

Let $\{V_i\}_{i=1}^m \subset \mathbb{C}^{s_1 \times s_2}$ have uniformly bounded norm, $\|V_i\| \le K$. Let $n = \max(s_1, s_2)$ and let $\{\epsilon_i\}_{i=1}^m$ be iid uniform ± 1 random variables. Then

$$\mathbb{E}_\epsilon \left\| \sum_{i=1}^m \epsilon_i V_i^* V_i \right\|_{(r)} \le C_1 \left\| \sum_{i=1}^m V_i^* V_i \right\|_{(r)}^{1/2}$$

where $C_1 = C_0 \sqrt{r} K \log^{5/2} n \log^{1/2} m$ and C_0 is a universal constant.

For our purposes, $V_i = \sqrt{\frac{|J|}{n}} \phi_i$. Then

$$\begin{aligned}
\mathbb{E}\epsilon_r(\mathcal{A}) &\le 2C_1 \frac{n}{m} \mathbb{E}_\Omega \left\| \sum \sqrt{\frac{|J|}{n}} \phi_i^* \phi_i \sqrt{\frac{|J|}{n}} \right\|_{(r)}^{1/2} \\
&= 2C_1 \sqrt{\frac{n}{m}} \left(\mathbb{E}\|\mathcal{A}^*\mathcal{A}\|\right)^{1/2} \\
&\le 2C_1 \sqrt{\frac{n}{m}} (\mathbb{E}\epsilon_r(\mathcal{A}) + 1)^{1/2}.
\end{aligned}$$

Outline of Proof

Fix some $\lambda \geq 1$ and choose

$$\begin{aligned} m &\geq C\lambda\mu r n \log^5 n \cdot \log |J| \\ &\geq \lambda n (2C_1)^2 \end{aligned}$$

This makes $\mathbb{E}\epsilon_r(\mathcal{A}) \leq \frac{1}{\lambda} + \frac{1}{\sqrt{\lambda}}$.
We can now use result by Talagrand.

Theorem

Let $\{\mathcal{Y}_i\}_{i=1}^m$ be independent symmetric random variables on some Banach space such that $\|\mathcal{Y}_i\| \leq R$. Let $\mathcal{Y} = \sum_{i=1}^m \mathcal{Y}_i$. Then for any integers $l \geq q$ and any $t > 0$

$$\Pr(\|\mathcal{Y}\| \geq 8q\mathbb{E}\|\mathcal{Y}\| + 2Rl + t\mathbb{E}\|\mathcal{Y}\|) \leq (K/q)^l + 2e^{-t^2/256q}, \tag{4}$$

where K is a universal constant.

Use theorem to establish

$$\Pr(\|\epsilon_r(\mathcal{A})\| \geq \epsilon) \leq e^{-C\epsilon^2\lambda}$$

Revised Algorithm

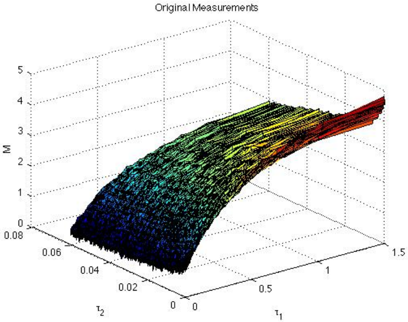
1. Randomly subsample data matrix M at indices Ω
2. Reconstruct $\widetilde{M}$ by solving

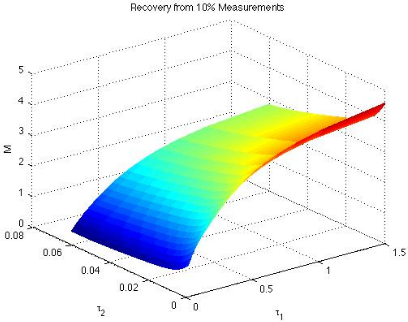
$$\begin{array}{ll} \arg\min & \|X\|_* \\ \text{such that} & \|\mathcal{R}_\Omega(X) - y\|_2 < \epsilon \end{array}$$

3. Holding alpha fixed, solve the Tikhonov regularization problem
4. Having solved the minimization problem, determine a more optimal alpha to use in the regularization problem
5. Loop between Steps 2 & 3 until convergence upon the optimal alpha

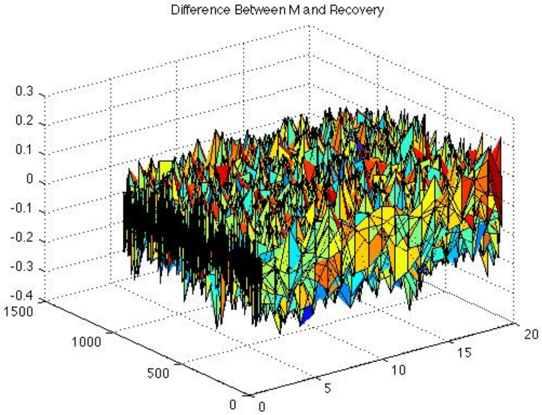
Outline

Recovery from Small Number of Entries

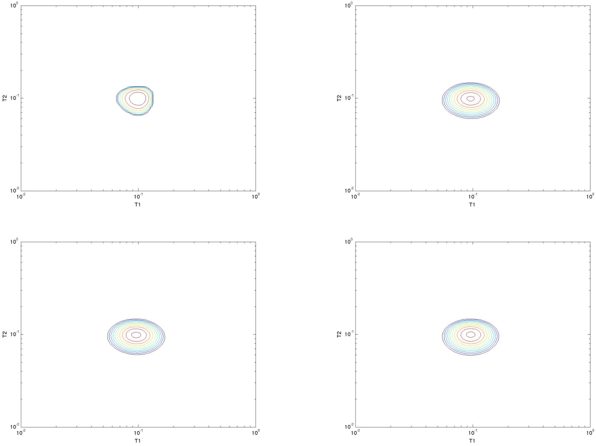




Error Analysis



Simple T1-T2 Map with 30dB SNR



Norbert Wiener Center for Harmonic Analysis and Applications

Figure: (T-L) T1-T2 Map, (T-R) Original Algorithm, (B-L) 30% Measurements, (B-R) 10% Measurements

Simple T1-T2 Map with 15dB SNR

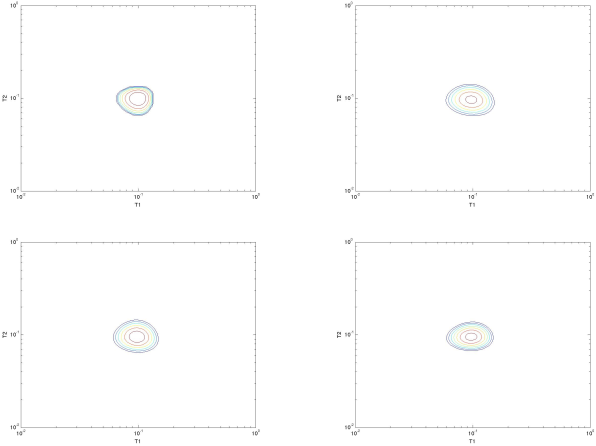


Figure: (T-L) T1-T2 Map, (T-R) Original Algorithm, (B-L) 30% Measurements, (B-R) 10% Measurements

Correlated T1-T2 Map with 30dB SNR

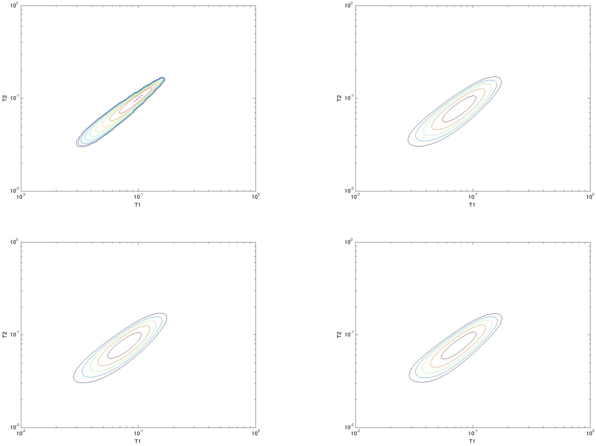


Figure: (T-L) T1-T2 Map, (T-R) Original Algorithm, (B-L) 30% Measurements, (B-R) 10% Measurements

Two Peak T1-T2 Map with 35dB SNR

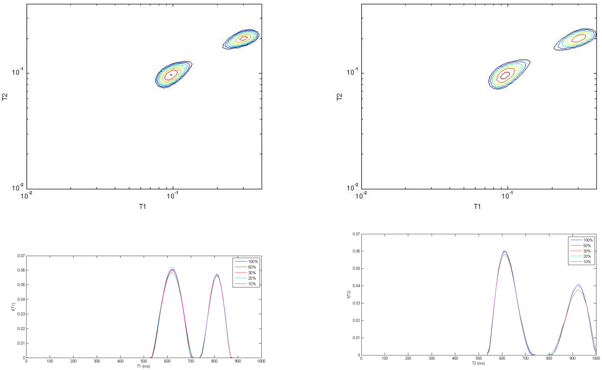


Figure: (T-L) Original Algorithm, (T-R) 20% Measurements, (B-L) 1D ILT of T1, (B-R) 1D ILT of T2

Two Peak T1-T2 Map with 35dB SNR

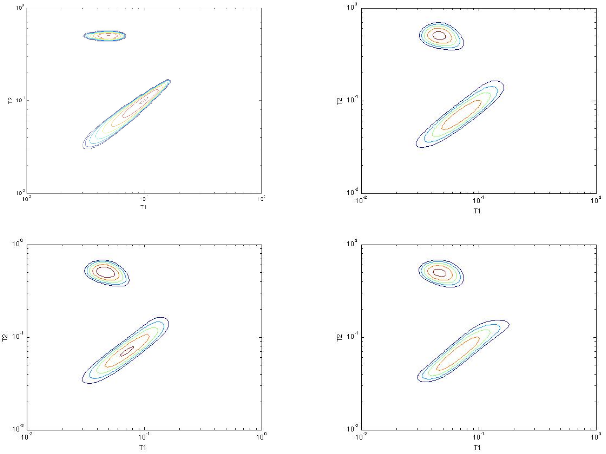
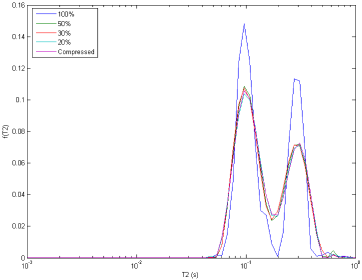


Figure: (T-L) T1-T2 Map, (T-R) Original Algorithm, (B-L) 30% Measurements, (B-R) 20% Measurements

Two Peak Skewing

- Only keeping largest singular values causes “fattening” of peaks
  - Necessary for computational efficiency of any 2D ILT algorithm
- Can be shown by taking 1D ILT of M, $U_1 U_1' M U_2 U_2'$, and reconstructions of M using CS



Preliminary Real Data

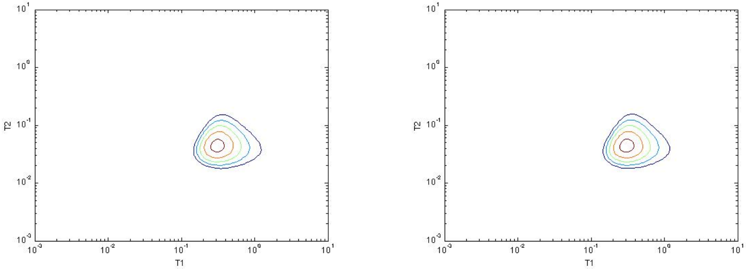


Figure: (L) Original Algorithm, (R) 20% Measurements

Conclusion

- Results recover measurement matrix M almost perfectly
- While small noise is added in compressive sensing reconstruction, it is dwarfed by algorithmic errors
- Next step is to implement on machine and run thorough testing