Gaussian Mixture Prior Models for Imaging of Flow Cross Sections from Sparse Hyperspectral Measurements

Zeeshan Nadir[1]
Michael S. Brown[2]
Mary L. Comer[1]
Charles A. Bouman[1]

3rd IEEE GlobalSIP Conference, Orlando, Florida, USA, Dec 14-16 2015

[1] School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, 47907
[2] Air Force Research Laboratory, Wright-Patterson AFB, OH 45433

Tunable Diode Laser Absorption Tomography (TDLAT)

- What is TDLAT?
  - Use light to measure density and temperature of a gas
  - Measures absorption spectral lines along a small number of paths
  - ~10 paths each with ~4 spectral lines = 40 measurements
- Why is TDLAT useful?
  - Hypersonic flow measurements, and many other applications

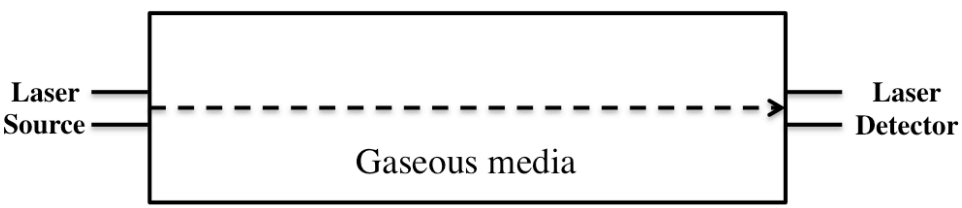
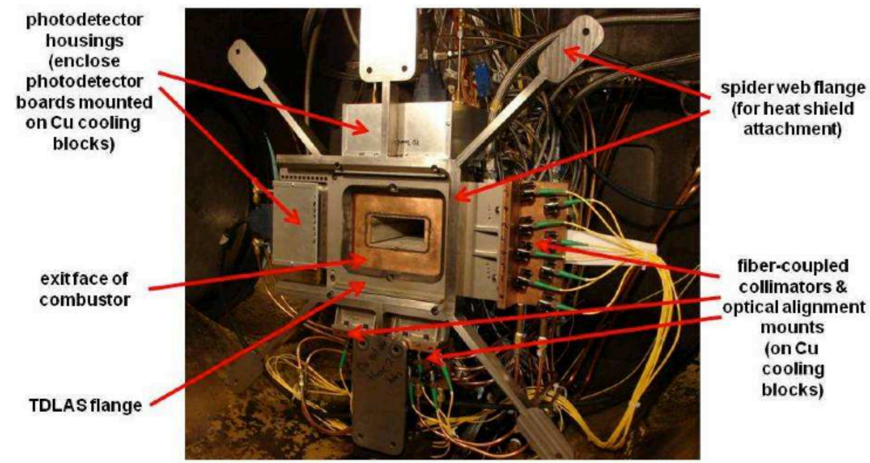
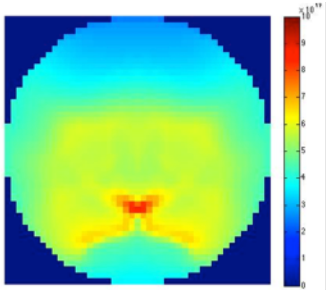


Illustration of TDLAT measurement for a single projection path

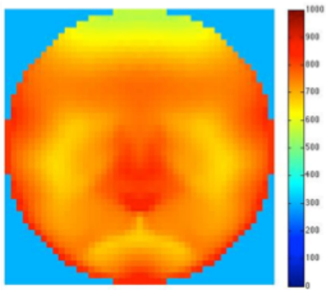


Test fixture

Why TDLAT is Difficult



Concentration
(molecules/cm³)



Temperature
(kelvins)

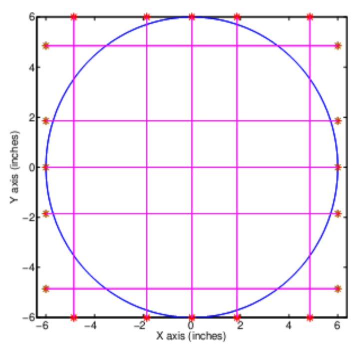
- Why TDLAT is Difficult
  - Nonlinear forward model
  - Highly underdetermined
    - 40 measurements (=10 projections x 4 spectral lines)
    - 3194 unknowns (~= 45×45 grid x 2 unknowns)
- Our solution
  - Use Bayesian inversion (MBIR)
  - Formulate a non-Gaussian prior model
    - Eigenimage decomposition
    - Gaussian mixture distribution
    - Computational fluid dynamics (CFD) training data
  - Multigrid optimization for reconstruction

Reconstruction Framework

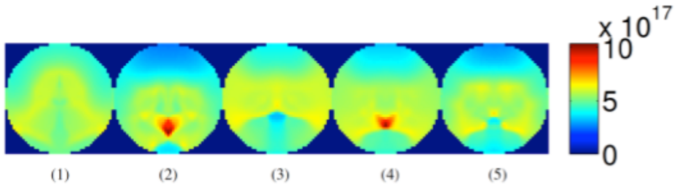
- Model-based iterative reconstruction (MBIR)

$$\hat{x} = \arg\max_{x} \left(\log p\left(y|x\right) + \log p\left(x\right) \right)$$

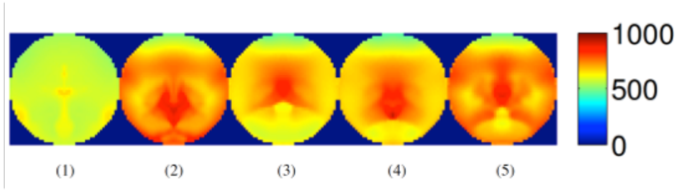
  - y : vector of absorbance data for each path and line spectrum
  - x : vector of unknown molecular concentration and temperature
  - $p\left(y|x\right)$: measurement model; models measurement procedure
  - $p\left(x\right)$: prior model; joint model for molecular concentration and temperature



Schematic of projection paths



Molecular concentration training images (molecules/cm^3)



Temperature training images (kelvins)

Example of CFD images used for training the prior model

Measurement Model: *p(y|x)*

- Nonlinear measurement model given as

$$Y_j = A_j + noise$$

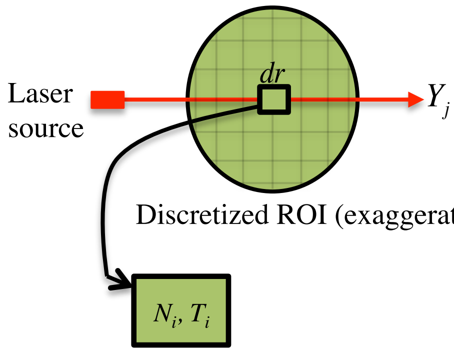
$$A_j = \int_{j^{th}\text{ path}} N(r)S(T(r))\,dr$$

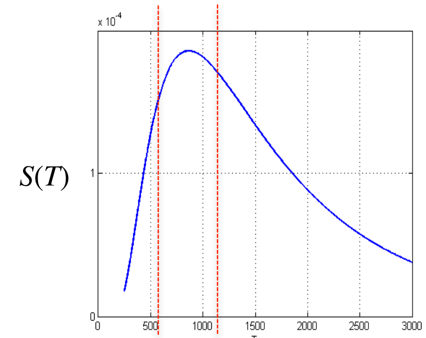
where

$N(r)$: Molar concentration of gas

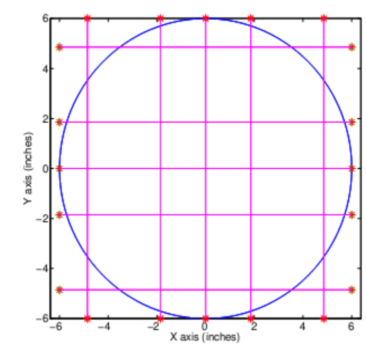
$T(r)$: Temperature of gas

$S(\cdot)$: Nonlinear function





Temperature in Kelvin



Schematic of projection path layout

- Log likelihood of absorbance data *y* given unknown *x* :

$$\log p(y|x) = \frac{-1}{2\sigma^2} \| y - Hf(x) \|_2^2 + \text{constants}$$

where

f : non-linear function defined by light absorption physics

H : forward projector matrix defined by projection layout

σ^2 : noise variance

Non-Gaussian Prior Model based on CFD Training Data

- Train using CFD training data
  - CFD simulations are VERY computationally expensive
  - Very little training data
- Better/accurate prior model trained using sparse training set
  - Non-Gaussian prior model
    - Eigenimage for dimensionality reduction
    - Gaussian mixture model
    - Train using EM algorithm
- MAP estimation
  - Quadratic surrogate => Majorization minimization

Gaussian Mixture Model as Prior

- Gaussian mixture model (GMM): A flexible non-Gaussian distribution
  - Parameter estimation of GMM is difficult
  - Use a **lower dimensional** vector z to express unknown x

$$x = Ez + \mu$$

  - Gaussian mixture distribution of z is given as

$$p(z) = \sum_{m=1}^{M} \frac{\tilde{\pi}_m}{(2\pi)^{\tilde{p}/2} \left|\tilde{R}_m\right|^{1/2}} \exp\left\{-\frac{1}{2}(z-\tilde{\mu}_m)^t \tilde{R}_m^{-1}(z-\tilde{\mu}_m)\right\}$$

where

$\tilde{\pi}_m$ - prior probability of m^{th} mixture component

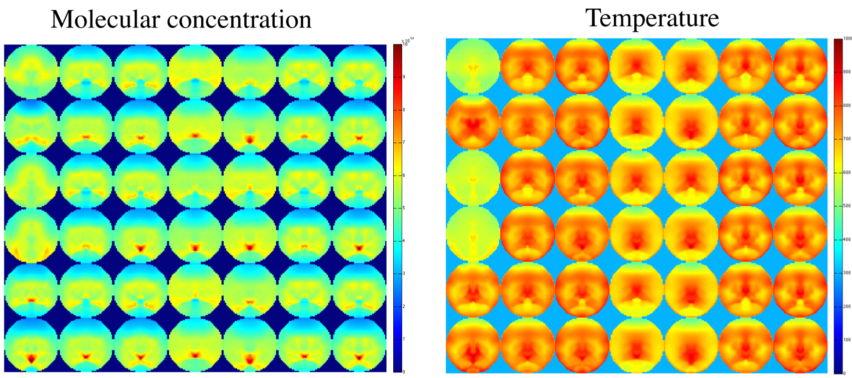
$\tilde{\mu}_m$ - mean of m^{th} mixture component

$\tilde{R}_m$ - covariance of m^{th} mixture component

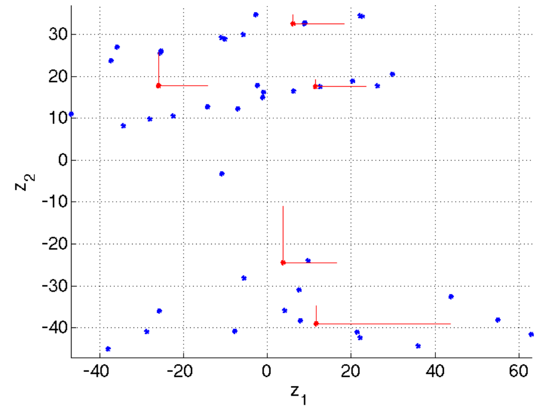
- Model mixture covariance matrices $\tilde{R}_m$ as diagonal matrices

Gaussian Mixture Model Parameter Estimation

- Use EM Algorithm to estimate the parameters

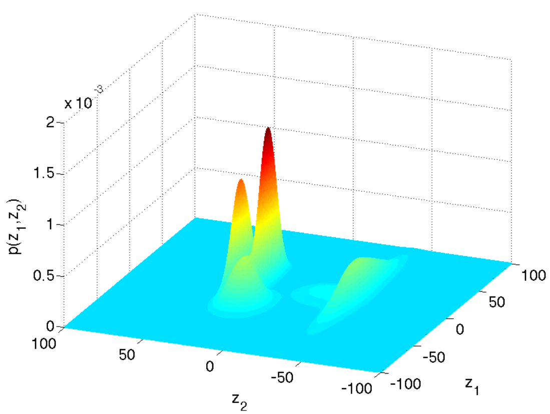


CFD training phantoms

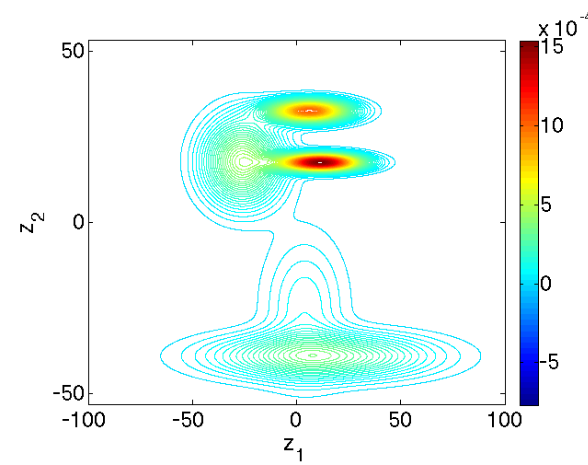


Scatter plot of training data

- The trained model captures non-Gaussian characteristics



Surface plot of Gaussian mixture distribution



Contour plot of Gaussian mixture distribution

Computing the MAP Estimate

- Minimize MAP cost function

$$c(z) = \frac{1}{2\sigma^2} \| y - Hf\left(Ez + \mu\right) \|_2^2 - \log\left(\sum_{k=1}^{M} \frac{\tilde{\pi}_k}{\left(2\pi\right)^{\tilde{p}/2} \left|\tilde{R}_k\right|^{1/2}} \exp\left\{ -\frac{1}{2} \left(z - \tilde{\mu}_k\right)^t \tilde{R}_k^{-1} \left(z - \tilde{\mu}_k\right) \right\} \right)$$

Problem: very complicated to minimize

- Solution: Use majorization minimization with quadratic surrogate function.

$$\tilde{c}\left(z\right) = \frac{1}{2\sigma^2} \| y - Hf\left(Ez + \mu\right) \|_2^2 + \frac{1}{2} \| z - \tilde{\mu} \|_{\tilde{B}}^2$$

Easy to minimize!

Question: How do we find a quadratic surrogate?

Lemma: Surrogate Cost Formulation

- Surrogate MAP cost obtained by using a quadratic approximation for prior

$$c(z;z') = \frac{1}{2\sigma^2} \| y - Hf(Ez+\mu) \|_2^2 + \frac{1}{2} \| z - \tilde{\mu} \|_{\tilde{B}}^2$$

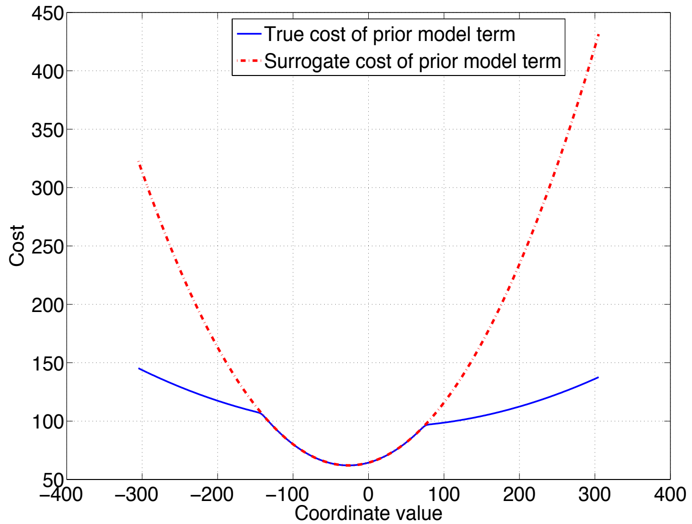
where

z' is the current value of unknown

$$\tilde{B} = \sum_{m=1}^{M} d_m \tilde{B}_m \ ,$$

$$\tilde{\mu} = \tilde{B}^{-1} \left(\sum_{m=1}^{M} d_m \tilde{B}_m \ \tilde{\mu}_m \right),$$

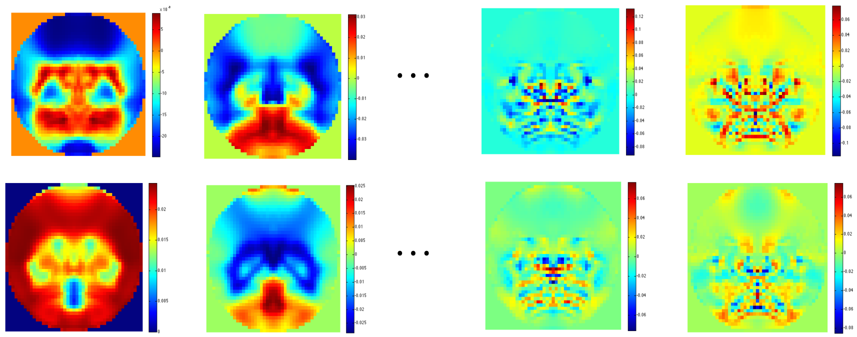
$$d_m = \frac{\tilde{\pi}_m \left| \tilde{B}_m \right|^{1/2} \exp\left\{ -\frac{1}{2} \| z' - \tilde{\mu}_m \|_{\tilde{B}_m}^2 \right\}}{\sum_{j=1}^{M} \tilde{\pi}_j \left| \tilde{B}_j \right|^{1/2} \exp\left\{ -\frac{1}{2} \| z' - \tilde{\mu}_j \|_{\tilde{B}_j}^2 \right\}}$$



Example of surrogate cost for prior model term

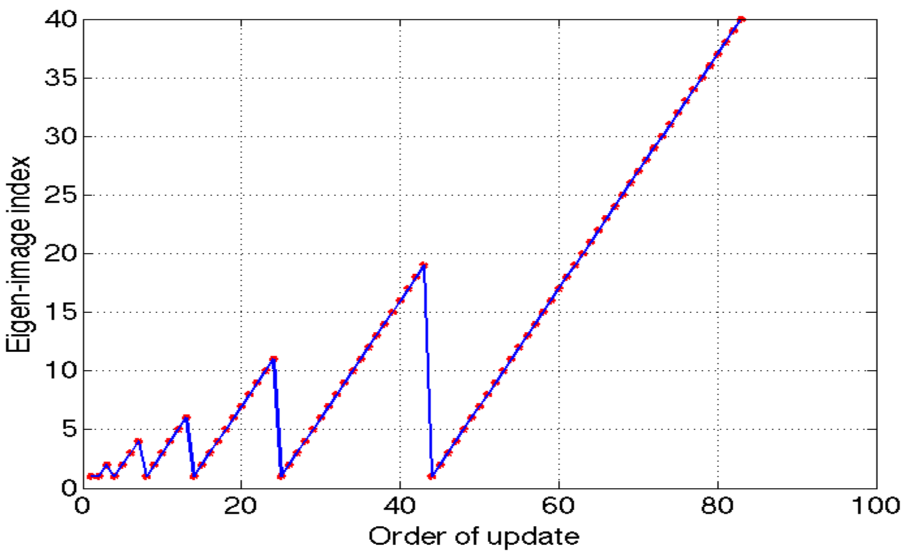
Multigrid Optimization

- Why multigrid?
  - Robust to local minimum in non-convex optimization
  - Faster convergence
- How does it work?
  - Based on eigenimages
  - Goes from largest eigen-values to smallest



eigenimage 1 eigenimage 2 eigenimage 40 eigenimage 41

Illustration showing eigenimages



Coordinate update pattern in multigrid algorithm; $\rho = 1.8$

EXPERIMENTAL RESULTS

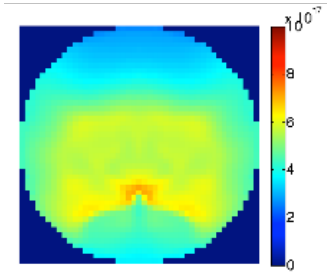
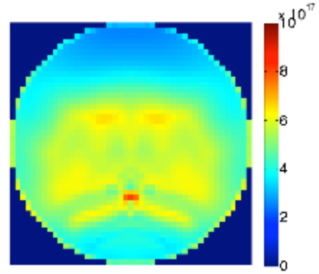
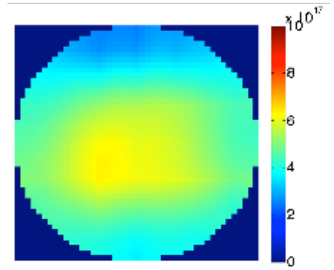
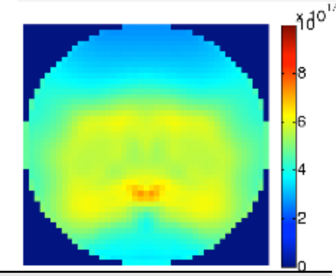
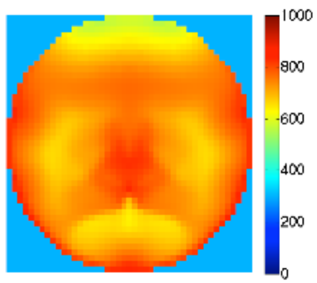
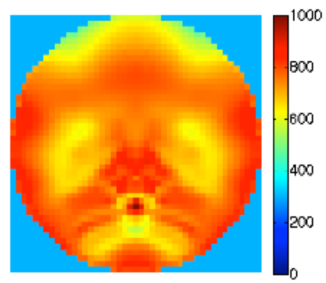
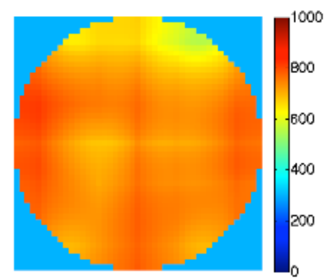
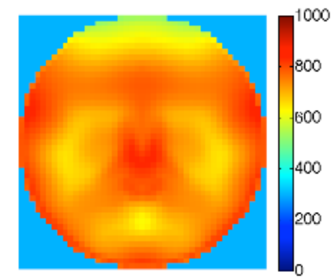
Reconstruction Experiments

- Prior models compared
  - Proper Orthogonal Decomposition (POD)*
  - Gaussian markov random field prior
  - Gaussian mixture model (GMM) prior <= our proposed method
- All results use 42 round-robin cross-validation
- Simulated data with average SNR = 30dB
- Normalized RMSE error:

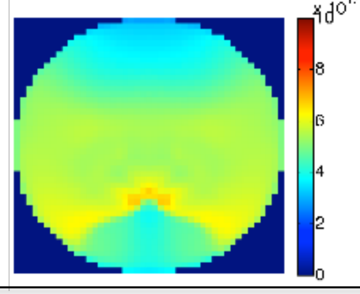
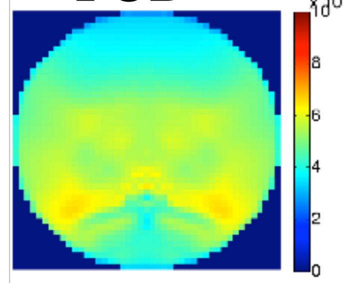
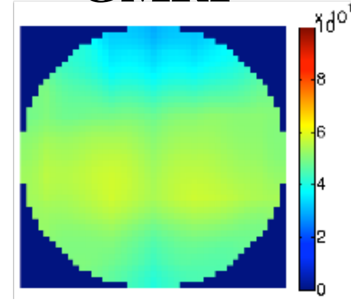
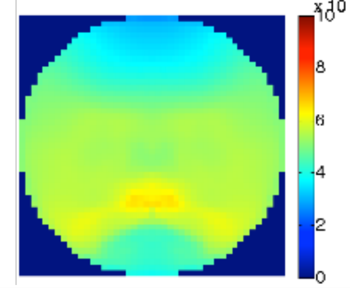
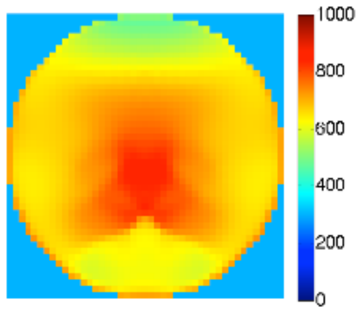
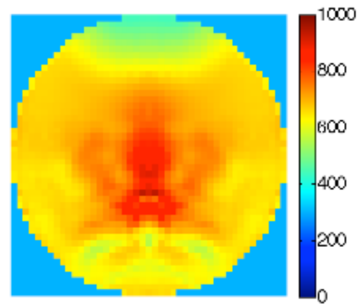
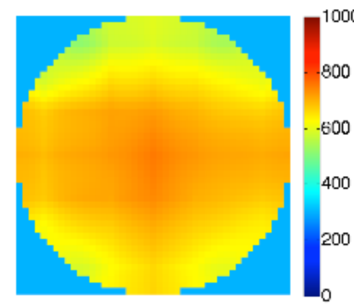
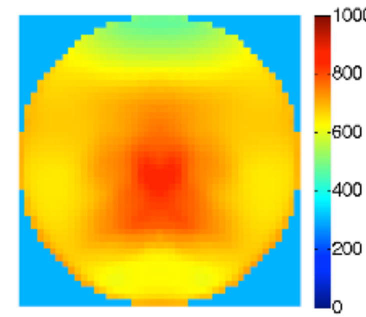
$$\mathrm{NRMSE}(X,Y)=\sqrt{\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i-Y_i)^2}{\max_i(Y_i)-\min_i(X_i)}}$$

* "Hyperspectral tomography based on proper orthogonal decomposition as motivated by imaging diagnostics of unsteady reactive flows" by W. Cai and Lin Ma.

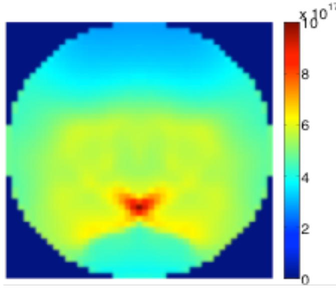
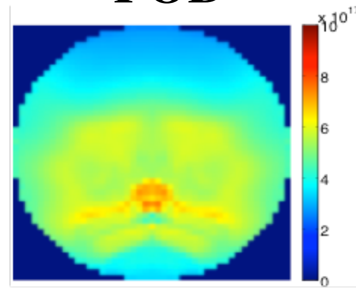
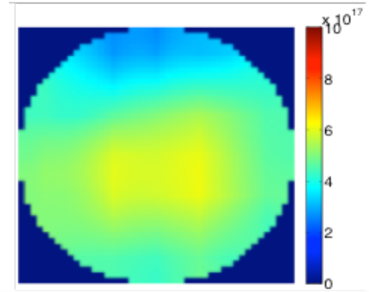
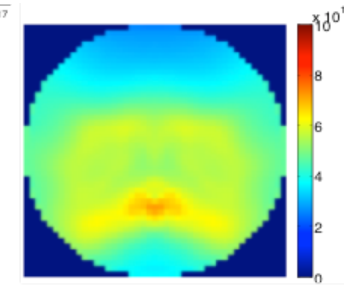
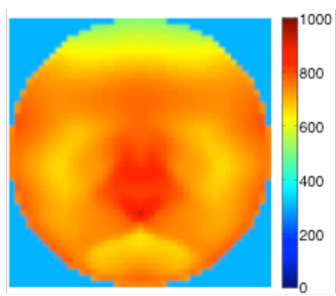
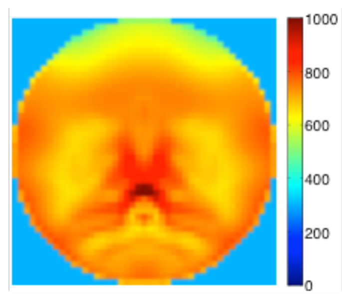
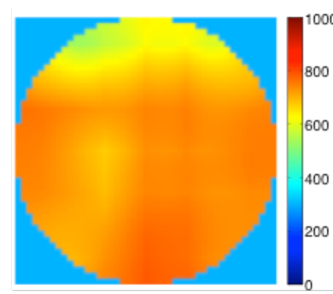
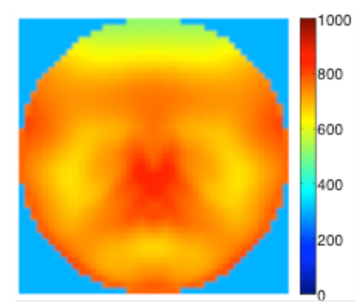
Reconstruction -1

	Ground truth	POD	GMRF Prior	GMM Prior
Molecular Concentration				
NRMSE		10.02%	9.85%	**3.83** %
Temperature				
NRMSE		15.20%	16.22%	**5.44** %

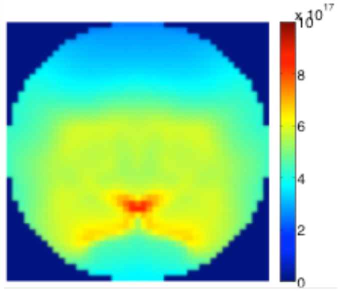
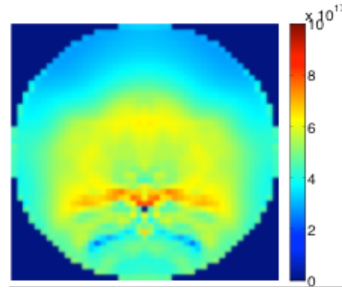
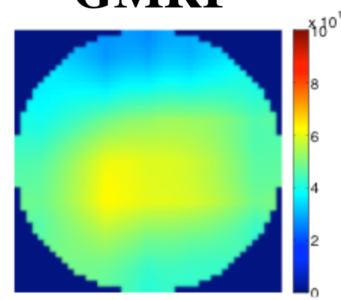
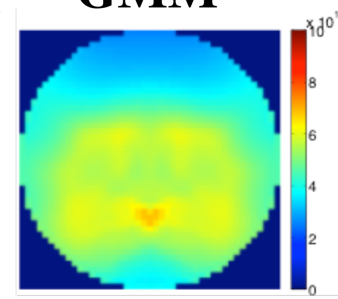
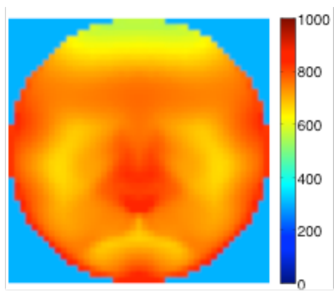
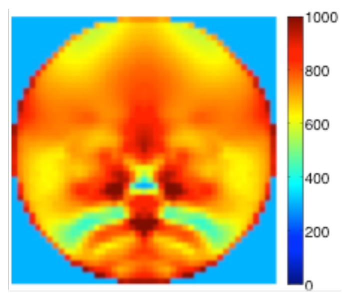
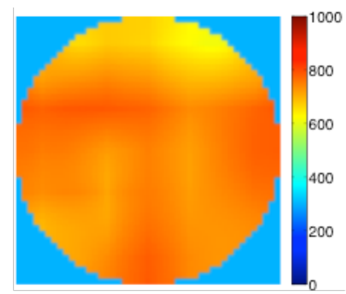
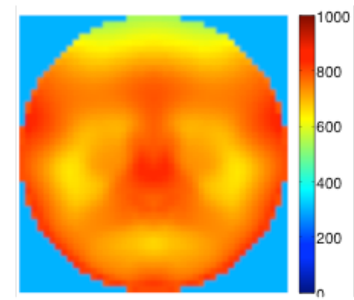
Reconstruction -2

	Ground truth	POD	GMRF	GMM
Molecular Concentration				
NRMSE		8.39%	9.81%	**6.24%**
Temperature				
NRMSE		6.92%	11.17%	**3.82 %**

Reconstruction -3

	Ground truth	POD	GMRF	GMM
Molecular Concentration				
NRMSE		5.30%	7.16%	**3.59%**
Temperature				
NRMSE		6.94%	10.69%	**4.39%**

Reconstruction -4

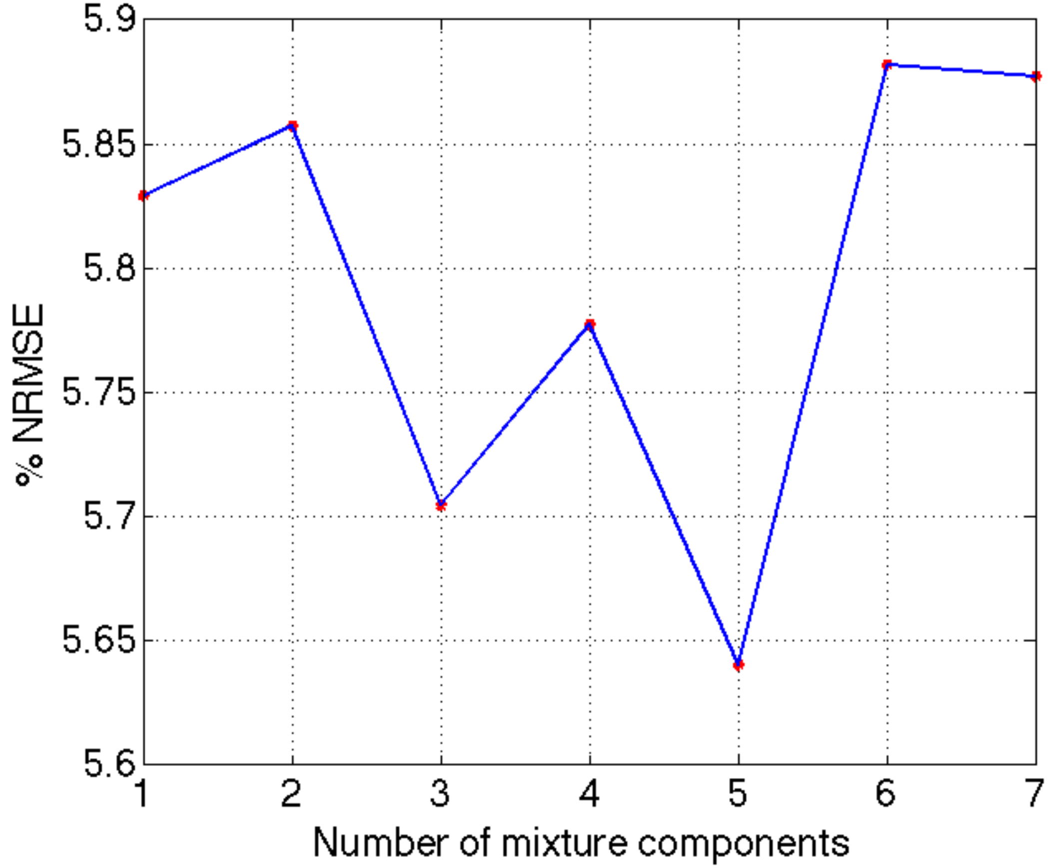
	Ground truth	POD	GMRF	GMM
Molecular Concentration				
NRMSE		5.61%	8.83%	**3.40%**
Temperature				
NRMSE		12.14%	14.78%	**4.81%**

Average Results of All Reconstruction Experiments

	% NRMSE (N)	% NRMSE (T)	% Average NRMSE
POD	9.89	12.13	11.01
GMRF	10.00	13.50	11.75
GMM	**6.14**	**5.14**	**5.64**

Average NRMSE for all 42 reconstruction experiments

Plot of NRMSE vs. Number of Mixture Components



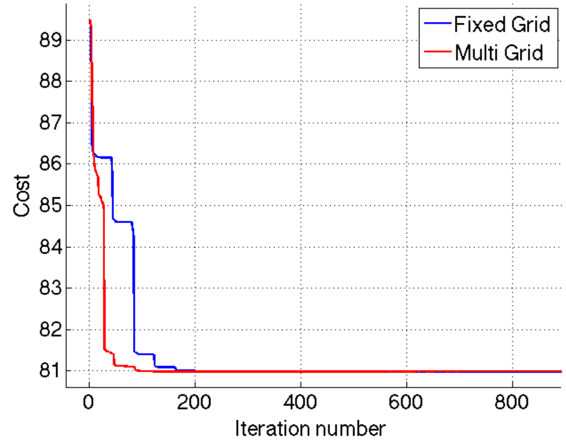
NRMSE plotted against number of mixture components

Convergence Experiments

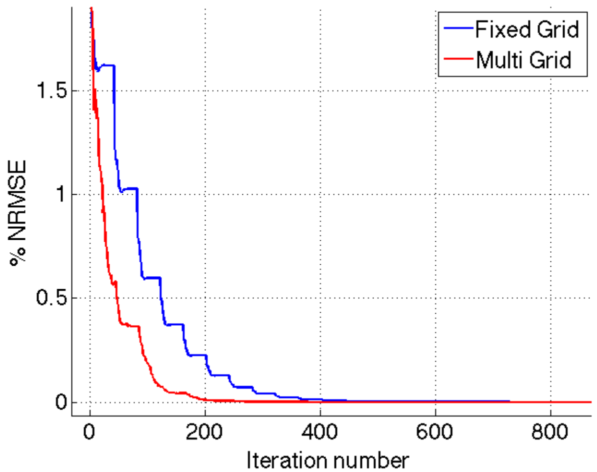
- First run algorithm to achieve "fully converged result"
- Run reconstructions again; compute
  - MAP cost and
  - NRMSE between current and converged result
- Total 42 experiments; for each reconstruction
  - The prior model in the reconstructions is Gaussian mixture model
  - $\rho = 1.8$

Comparison of Convergence

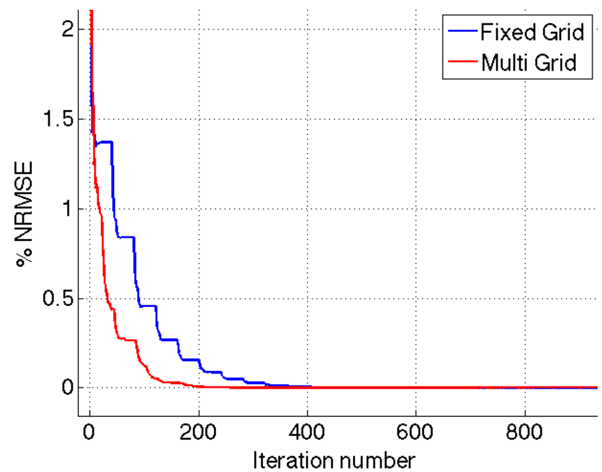
- Plots averaged over 42 experiments; also representative of typical case



MAP Cost



NRMSE between current converged N and current N



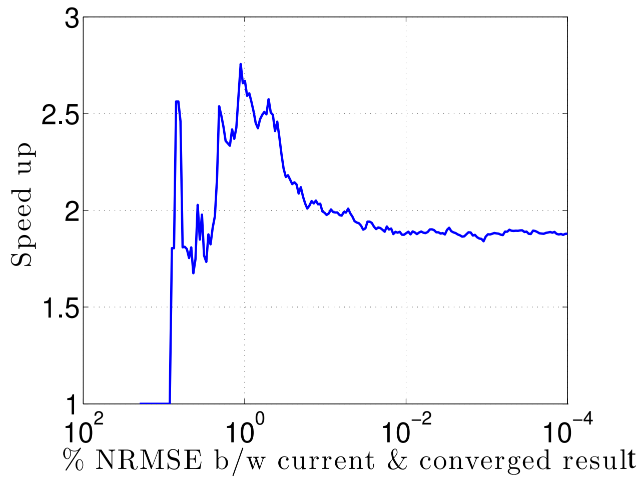
NRMSE between current converged T and current T

CPU Time and Speed up

NRMSE	1%	0.5%	0.1%	0.01%
Average CPU time (sec) (Fixed-grid algorithm)	4.66	8.66	17.48	31.07
Average CPU time (sec) (Multigrid algorithm)	1.72	3.39	8.87	16.57

Time take to achieve specified NRMSE between current and converged result

(Experiments done on Intel core i7 with 32GB of memory using MATLAB)



Average relative benefit of multigrid algorithm in computational time

$$\text{speed up (r)} = \frac{\text{Fixed-grid iterations to achieve } r \text{ \% NRMSE}}{\text{Multigrid iterations to achieve } r \text{ \% NRMSE}}$$

Summary

- Proposed a novel MBIR framework for TDLAT
  - Nonlinear forward model
  - Based on physics of line-spectrum light absorption
- Proposed non-Gaussian GMM prior to model
  - Non-homogeneous characteristics of images
  - Non-Gaussian characteristics of images
- MAP estimation
  - Majorization using surrogate function
  - Multigrid optimization
- Results
  - Reduced NRMSE
  - Fast convergence/reduced computation

Acknowledgements

- We acknowledge

  1. Dr. Mark Hagenmaier of AFRL for the use of CFD simulations

  2. Mr. Ed Tucker of AFRL for his financial support