Multilevel Optimization for Multi-Modal X-ray Data Analysis

Zichao (Wendy) Di

Mathematics & Computer Science Division
Argonne National Laboratory

August 8, 2017

Outline

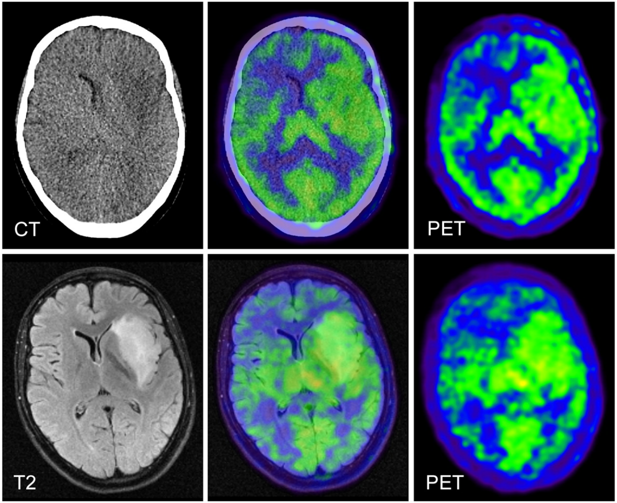
Outline

Multi-Modality Imaging
- Example: X-ray Tomography
- Forward Model
- Optimization Algorithms
- Numerical Results

Multilevel-Based Acceleration
- Other Applications of Multilevel Methods
- Introduction of MG/OPT
- Application of Multilevel in Tomographic Reconstruction
- Numerical Results

Summary and Outlook

More is Better



CT:hard-tissue; T2(MRI):soft tissue; PET: functional characteristics [1]

[1] Boss, A. and et al. (2010). Hybrid PET/MRI of Intracranial Masses: Initial Experiences and Comparison to PET/CT, JNM

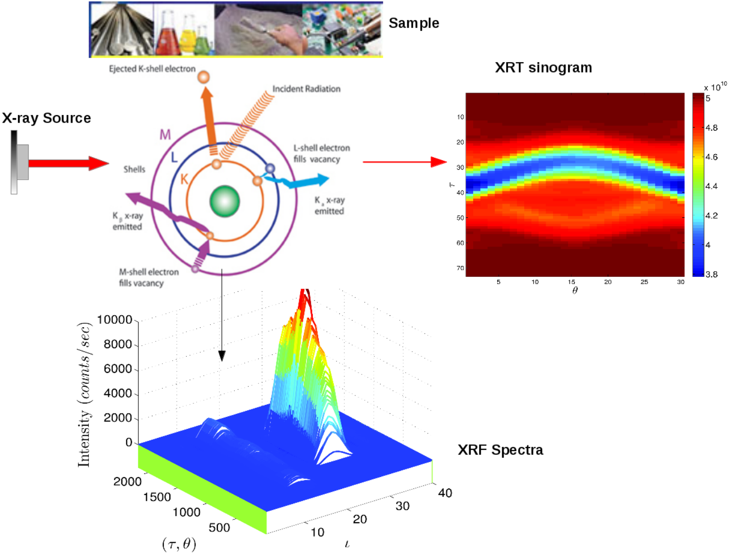
Outline

- Multi-Modality Imaging
  - Example: X-ray Tomography
  - Forward Model
  - Optimization Algorithms
  - Numerical Results
- Multilevel-Based Acceleration
- Summary and Outlook

Tomographic Image Reconstruction

- Non-invasive imaging technique to visualize internal structures of object.
- *Tomography applications:* physics, chemistry, astronomy, geophysics, medicine, etc.
- *Tomographic imaging modalities:* X-ray transmission, ultrasound, magnetic resonance, X-ray fluorescence, etc.
- *Task:* estimate distribution of physical quantities in sample from measurements.
- Limited angle tomography reconstruction is naturally ill-conditioned.

Schematic Experimental Setup with Two Modalities



Reconstruction Approaches

- Traditional: filtered backprojection
  - Restricted to simple tomographic model
  - Requires a large number of projections
- Alternatively, iterative reconstruction from single data modality
  - Requires much less data acquisition, results in higher accuracy

Our Goal:

Formulate a joint inversion integrating XRF and XRT data to improve the reconstruciton quality of elemental map.

Outline

X-ray Transmission (XRT)

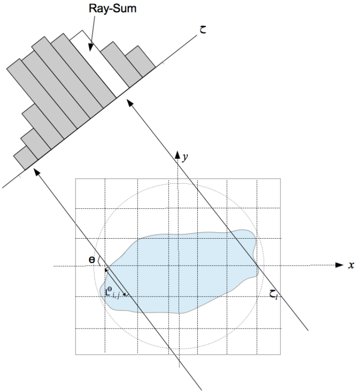
- Traditionally, the XRT projection of the object from beam line (θ, τ) is modeled as

$$\mathbf{F}^{T}_{\theta,\tau}(\tilde{\boldsymbol{\mu}}^{\boldsymbol{E}}) = I_0 \exp\left\{ -\sum_{v} L_v \tilde{\mu}^{E}_{v} \right\}.$$

to directly solve the linear attenuation coefficient $\tilde{\mu}^{E}_{v}$ for each voxel v,

- In our approach,notice $\tilde{\mu}^{E}_{v} = \sum_{e} \mathcal{W}_{v,e} \mu^{E}_{e}$,

$$\mathbf{F}^{T}_{\theta,\tau}(\boldsymbol{\mathcal{W}}) = I_0 \exp\left\{ -\sum_{v,e} L_v \mu^{E}_{e} \mathcal{W}_{v,e} \right\}$$



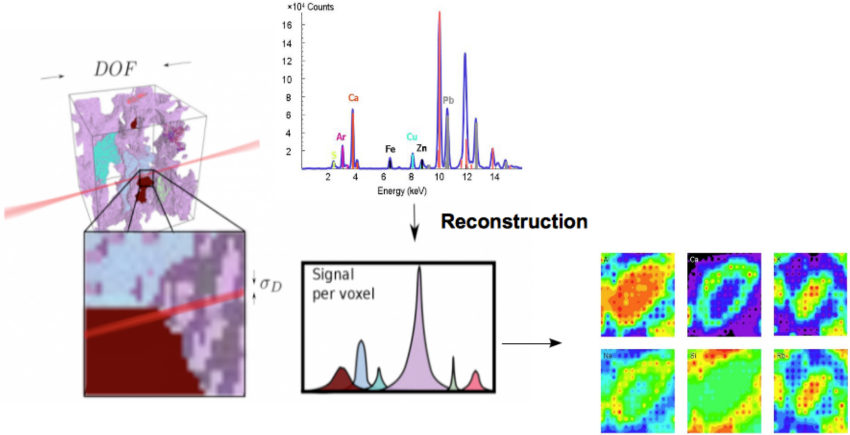
I_0: incident photon flux

μ^{E}_{e}: mass attenuation coefficient of element $e \in \mathcal{E}$ at beam incident energy E

$\boldsymbol{\mathcal{W}} = \mathcal{W}_{v,e}$: tensor denoting how much of element e is in voxel v

$\mathbf{L} = [L_v]$: tensor of intersection length of beam line (θ, τ) with the voxel v

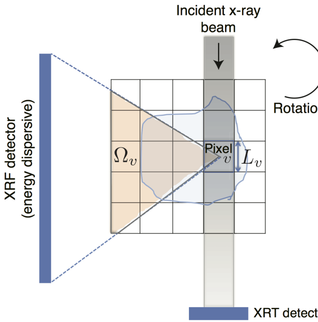
X-ray Fluorescence (XRF)



Mathematical Model of XRF

- First, we obtain the unit fluorescence spectrum

$$\mathbf{M}_e = F^{-1}\left(F(I_e) * F\left(\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{\frac{-\mathbf{x}^2}{2\sigma^2}\right\}\right)\right)$$



- Then, the XRF spectrum, $\mathbf{F}^R_{\theta,\tau}$

$$= \sum_{v,e,d} \frac{L_v \mathcal{W}_{v,e} \mathbf{M}_e}{n_d} \exp\left\{-\sum_{v',e'} \mathcal{W}_{v',e'}\left(\mu^E_{e'} L_{v'} \mathbb{I}_{v' \in \mathcal{U}^v} + \mu^{E_e}_{e'} P_{v,v',d}\right)\right\}$$

$\mathbf{P} = [P_{v,v',d}]$: tensor of intersection length of fluorescence detectorlet path d with the voxel v'

Outline

Multi-Modality Imaging

- Example: X-ray Tomography
- Forward Model
- Optimization Algorithms
- Numerical Results

Multilevel-Based Acceleration

Summary and Outlook

Resulting Optimization: Joint Reconstruction (JRT)

Goal:

Find $\mathcal{W}$ so that $\mathbf{F}^R_{\theta,\tau}(\mathcal{W}) = \mathbf{D}^R_{\theta,\tau}$ and $\mathbf{F}^T_{\theta,\tau}(\mathcal{W}) = D^T_{\theta,\tau}$

$$\min_{\mathcal{W}\geq 0} \phi(\mathcal{W}) = \sum_{\theta,\tau} \left(\frac{1}{2} \left\| \mathbf{F}^R_{\theta,\tau}(\mathcal{W}) - \mathbf{D}^R_{\theta,\tau} \right\|^2 + \frac{\beta_1}{2} \left\| \mathbf{F}^T_{\theta,\tau}(\mathcal{W}) - \beta_2 D^T_{\theta,\tau} \right\|^2 \right)$$

where

- $\mathbf{D}^R_{\theta,\tau} \in \mathbb{R}^{n_E}$: measurement data of XRF signal detected at angle θ from light beam τ
- $D^T_{\theta,\tau} \in \mathbb{R}$: the measurement data of XRT signal detected at angle θ from light beam τ
- $\beta_1, \beta_2 > 0$ are scaling parameters to balance the ability of each modality to fit the data, and detect the relative variability between the data sources.

Note: Optimization differs on how $\mathbf{F}^R_{\theta,\tau}$ and $\mathbf{F}^T_{\theta,\tau}$ are combined

How Multimodality Can Help

Consider an **overdetermined** (more equations than the number of unknowns) but **rank deficient** system which has an **infinitude of solutions**:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0.5 \end{bmatrix}$$

Another such rank-deficient (not full-rank) system:

$$\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}$$

However, combine these two can form a **full rank** and **consistent** system with a unique solution $\mathbf{x}$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0.5 & 1 \\ 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0.5 \\ 1.5 \\ 3 \end{bmatrix}$$

Optimization Solver

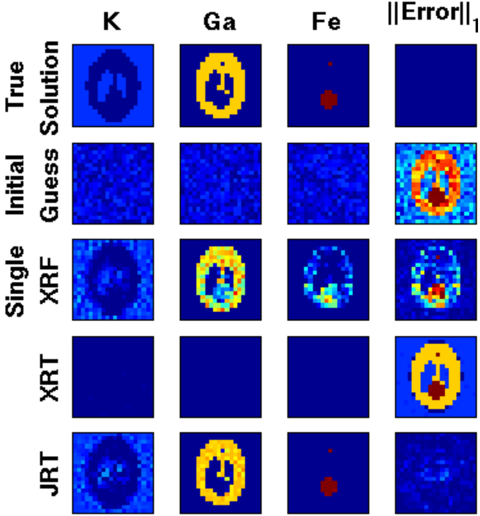
In experiments, we use truncated-Newton (TN) method with preconditioned conjugate gradient (PCG) to provide a search direction:

- To satisfy the bound constraints, the projected PCG is applied to the reduced Newton system
- One TN iteration typically requires $\kappa(O(10))$ PCG iterations

Main expense of each outer iteration is $\kappa + 2$ function-gradient evaluations.

Outline

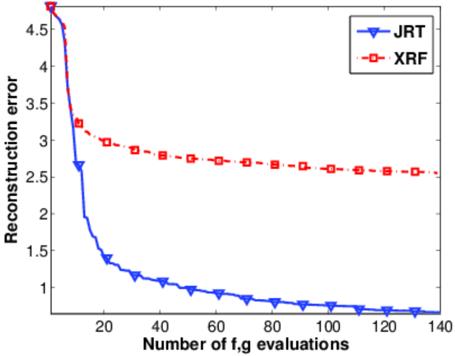
JRT versus Single XRF Reconstruction (Synthetic)[1]

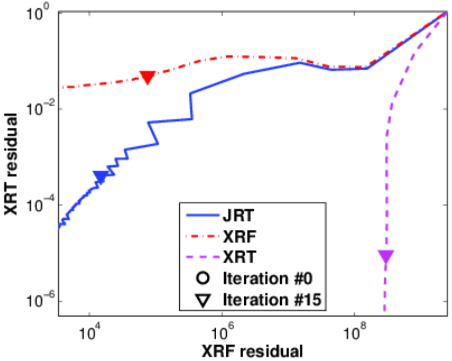


Element Unit: $g/\mu m^2$

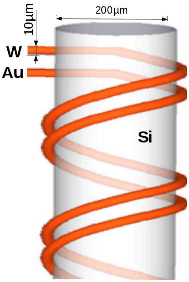
[1] Di, Leyffer, and Wild (2016). An Optimization-Based Approach for Tomographic Inversion from Multiple Data Modalities, SIAMIS

Convergence

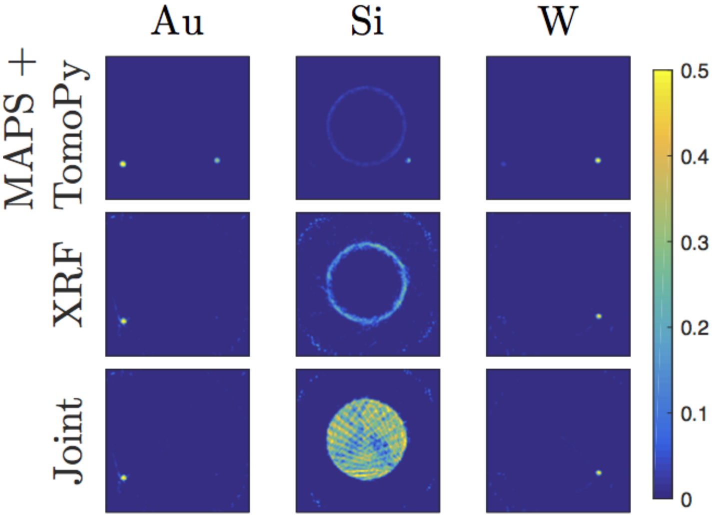




Addressing Self-Absorption Effect [1]



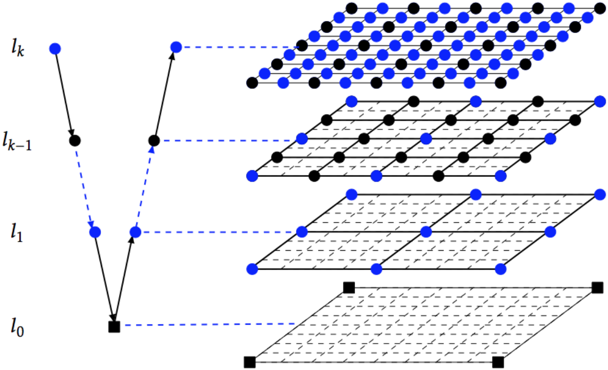
Sample: Solid glassrod (Silicon) with 2 wires (Tungsten, Gold).



[1] Di and et al. (2017). Joint reconstruction of x-ray fluorescence and transmission tomography, Optics Express

Outline

Motivation



$N = 9^2$, $flops = 2\times 10^7$

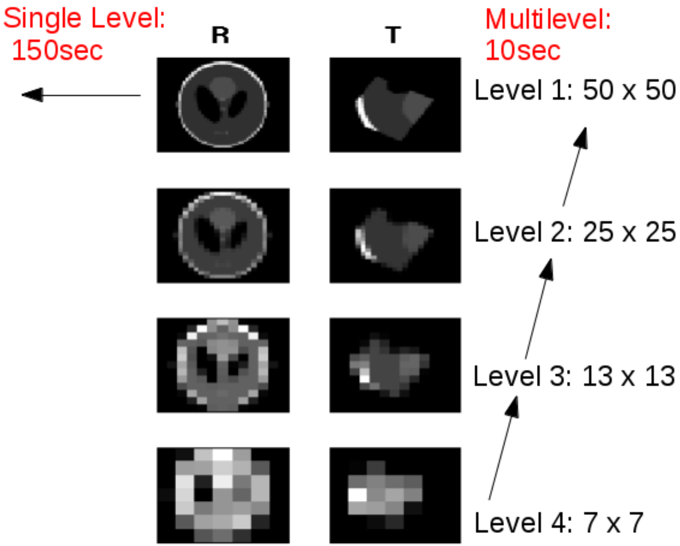
$N = 5^2$, $flops = 6\times 10^6$

$N = 3^2$, $flops = 3\times 10^6$

$N = 2^2$, $flops = 8\times 10^5$

Outline

Multilevel Acceleration of Image Registration



Outline

Introduction of MG/OPT

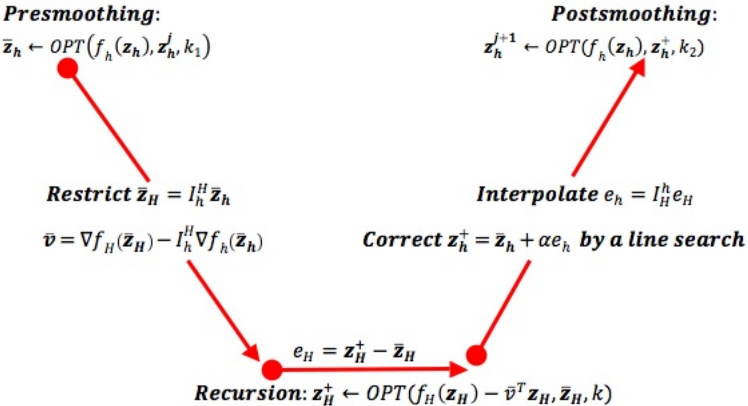
- Multigrid optimization algorithm (MG/OPT) is a general framework to accelerate a traditional optimization algorithm[1].
- Recursively use coarse problems to generate search directions for fine problems.
- MG/OPT can deal with more general problems in an optimization perspective, in particular, it is able to handle inequality constraints in a natural way.
- Multiple options to design the hierarchy of the problem:
  - through image space.
  - through data space.

[1] Nash, S. G. (2000). A Multigrid Approach to Discretized Optimization Problems, OMS

MG/OPT

Given:

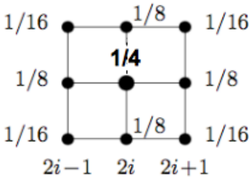
- High-resolution model $f_h(z_h)$, easier-to-solve low-resolution model $f_H(z_H)$
- $\mathbf{z}^+ \leftarrow \mathrm{OPT}(f(\mathbf{z}), \bar{\mathbf{z}}, k)$
- A restriction operator I_h^H and an interpolation operator I_H^h
- An initial estimate $\mathbf{z}_h^0$ of the solution $\mathbf{z}_h^*$ on the fine level
- Integers k_1 and k_2 satisfying $k_1 + k_2 > 0$



Outline

Interpolation/Restriction Operators

- Restriction on 2D parameter $\mathcal{W}_h \in \mathbb{R}^2$ to produce $\bar{\mathcal{W}}_H$ using full weighted matrix:



- Restriction on gradient $\hat{I}_h^H = C\, I_h^H$ where C balances the order difference between $\phi_h(\bar{\mathcal{W}}_h)$ and $\phi_H(I_h^H \bar{\mathcal{W}}_h)$

- Interpolation operator: $I_H^h = 4(I_h^H)^T$.

Coarse Grid Surrogate Model

- Shift experimental data for coarse level:

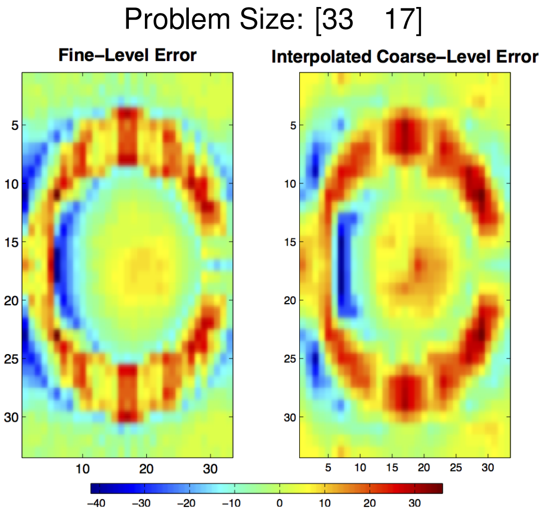
$$\tilde{\mathbf{D}}_{\theta,\tau} = \mathbf{D}_{\theta,\tau} - (\mathbf{F}^{,h}_{\theta,\tau}(\bar{\mathcal{W}}_h) - \mathbf{F}^{,H}_{\theta,\tau}(I_h^H \bar{\mathcal{W}}_h))$$

The surrogate model:

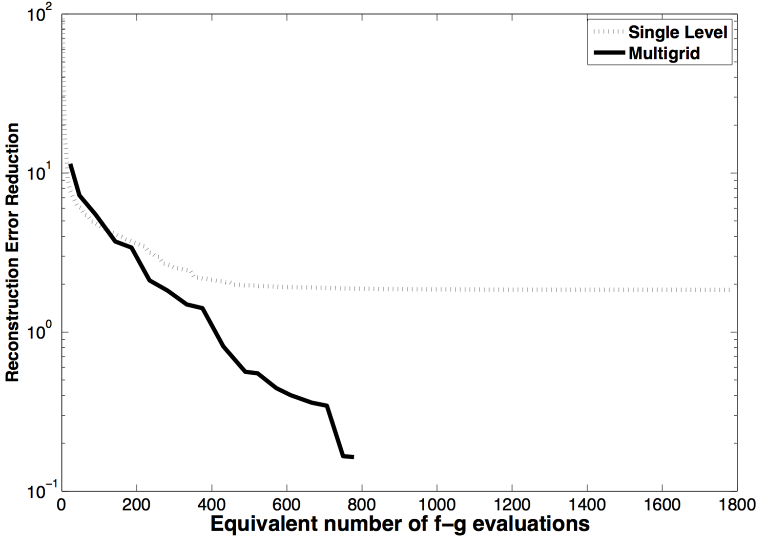
$$\tilde{\phi}_H(\mathcal{W}_H) = \sum_{\theta,\tau} \left(\frac{1}{2} \left\| \mathbf{F}^{,H}_{\theta,\tau}(\mathcal{W}_H) - \tilde{\mathbf{D}}_{\theta,\tau} \right\|^2 \right) - \left(\nabla \phi_H(I_h^H \bar{\mathcal{W}}_h) - \hat{I}_h^H \nabla \phi_h(\bar{\mathcal{W}}_h) \right)^T \mathcal{W}_H$$

Outline

MG/OPT Search Directions versus Error



2-level MG/OPT Reconstruction



Outline

Summary

- Established a link between X-ray transmission and X-ray fluorescence datasets by reformulating their corresponding physical models.
- Developed a simultaneous optimization approach for the joint inversion, and achieved a dramatic improvement of reconstruction quality with no extra computational cost.
- Proposed a multigrid-based optimization framework to further reduce the computational cost of the reconstruction problem.
- Preliminary results show that coarsening in voxel space improves accuracy/speed.

Making More of More

- Extend our multimodal analysis tool to data from different instruments, involving varying spatial resolution and contrast mechanisms.
- Guided by the hierarchical nature of our multilevel algorithm, we will investigate new data acquisition strategies and allow for flexible and adaptive sampling approaches.
- Enable true real-time feedback.

Acknowledgement

This work was funded under the project ”The Tao of Fusion: Pathways for Big-data Analysis of Energy Materials at Work” and ”Base optimization (DOE-ASCR DE-AC02-06CH11357)”.

Thanks to
Sven Leyffer, Stefan Wild, and
collaborators at APS:
Francesco De Carlo, Si Chen, Doga Gursoy, Young Pyo Hong, Chris Jacobsen, Stefan Vogt