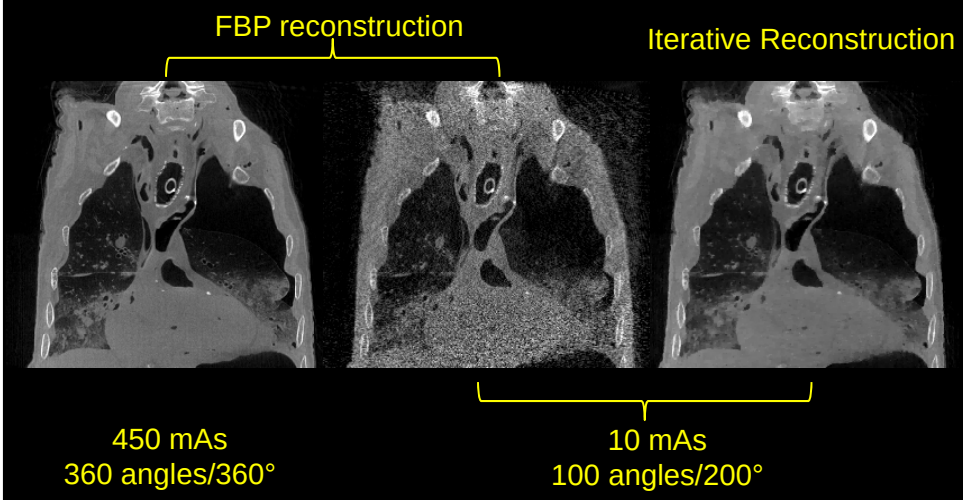
Iterative Reconstruction Methods in Computed Tomography

J. Webster Stayman

Dept. of Biomedical Engineering, Johns Hopkins University

Power of Iterative Reconstruction

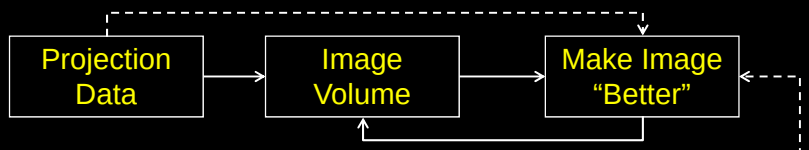


Learning Objectives

- Fundamentals of iterative methods
  - Approach to forming an iterative algorithm
  - Identify particular classes of methods
- Advantages of iterative approaches
  - Intuition behind what these algorithms do
  - Flexibility of these techniques
- Images produced by iterative methods
  - Differences from traditional reconstruction
  - Image properties associated with iterative reconstruction

Iterative Reconstruction

- What is iterative reconstruction?



- How can we make the image better?
  - Get a better match to the data
    - Requires a data model
  - Enforce desirable image properties
    - Encourage smoothness, edges, etc.
  - Need a measure of "better"

Building an Iterative Technique

- Define the objective
  - Find the volume that best fits the data and desired image quality

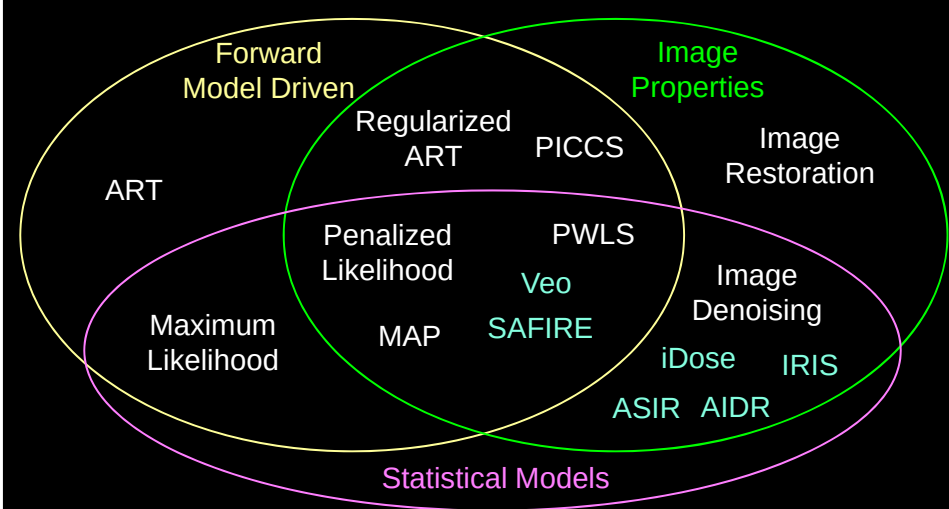
$$\text{volume} = \arg\max\left\{\left\|\text{data, model}\right\| \,\&\, \left\|\text{image properties}\right\|\right\}$$

$$\hat{\mu} = \arg\max\left\{\left\|\text{y}, \bar{\text{y}}(\mu)\right\| \,\&\, \left\| f(\mu)\right\|\right\}$$

- Devise an algorithm that solves the objective
  - Iteratively solve for $\hat{\mu}$
  - Decide when to stop iterating (and how to start)

Objective Function Optimization Algorithm

Reconstruction Choices



*Disclaimer: The exact details of commercially available reconstruction methods are not known by the author.

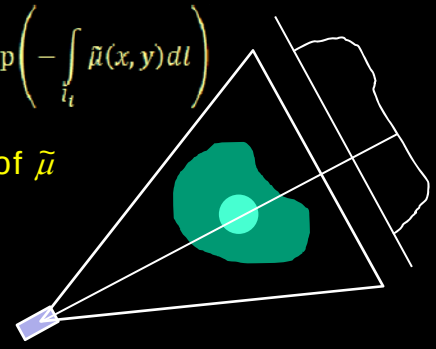
Model-based Approaches

- Transmission Tomography Forward Model
  - Projection Physics, Beer's Law

$$P(\text{photon survives}) = \exp\left(-\int_{l_i} \tilde{\mu}(x,y)\,dl\right)$$

$$\bar{y}_i = E[\text{number of photons}] = I_0 \exp\left(-\int_{l_i} \tilde{\mu}(x,y)\,dl\right)$$

- Need a Parameterization of $\tilde{\mu}$

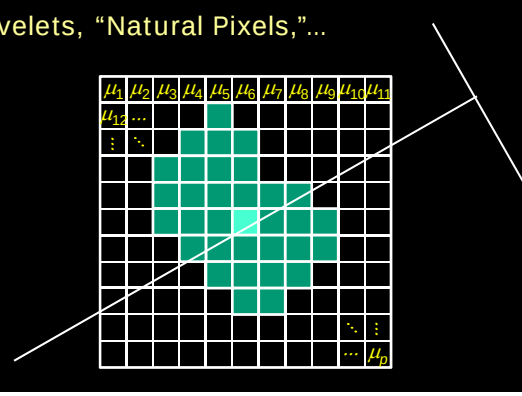


Parameterization of the Object

- Continuous-domain object
- Want finite number of parameters
- Choices of basis functions:
  - Point Samples - Bandlimited
  - Contours – Piecewise constant
  - Blobs, Wavelets, "Natural Pixels,"...
- Voxel Basis

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

$$\tilde{\mu}(x,y) \cong \mathbf{B}\boldsymbol{\mu}$$



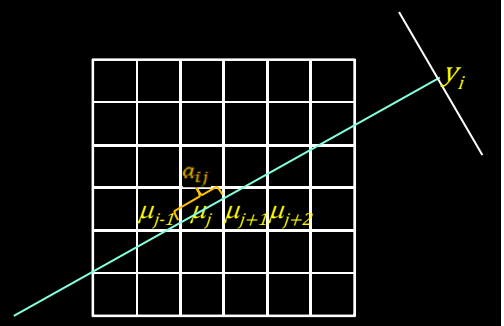
Projection

- Linear operation $\bar{y}_i = I_0 \exp\left(-\int_{l_i} \tilde{\mu}(x,y)\,dl\right)$
- Discrete-Discrete for parameterized problem

$$y_i = I_0 \exp\left(-\sum_{j=1}^{p} a_{ij}\mu_j\right)$$

- System matrix

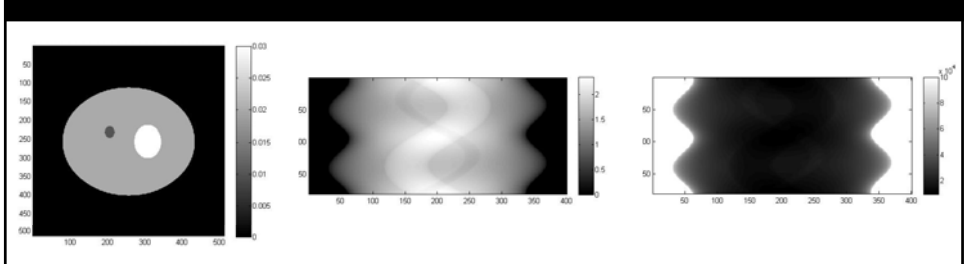
$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{ip} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{np} \end{bmatrix}$$



Forward Model

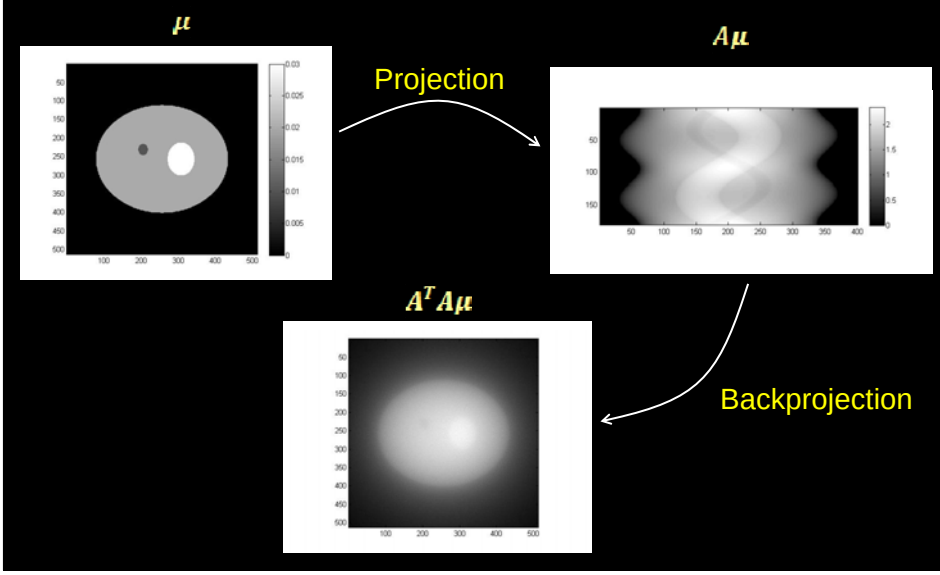
- Mean measurements as a function of parameters

$$\bar{y}(\mu) = I_0 \exp(-A\mu)$$



μ $\quad$ $A\mu$ $\quad$ $I_0\exp(-A\mu)$

Aside: Backprojection



A Simple Model-Driven Approach

- Forward Model $\bar{y}(\mu) = I_0 \exp(-\mathbf{A}\mu)$
- Objective and Estimator

$$\hat{\mu} = \arg\min \left\| y - \bar{y}(\mu) \right\| \quad \hat{\mu} \cong \arg\min \left\| -\log\left(\frac{y}{I_0}\right) + \mathbf{A}\mu \right\|$$

$$\hat{\mu} = \arg\min \left\| \mathbf{A}\mu - l \right\|^2 = \left[\mathbf{A}^T \mathbf{A} \right]^{-1} \mathbf{A}^T l$$

Projection-Backprojection Backprojection

- For a squared-distance metric
  - Many possible algorithms
  - Can be equivalent to ART, which seeks $\mathbf{A}\mu = l$

Flexibility of Iterative Methods

$$\hat{\mu} = \arg\min \left\| \mathbf{A}\mu - l \right\|^2 = \left[\mathbf{A}^T \mathbf{A} \right]^{-1} \mathbf{A}^T l$$

- For a well-sampled, parallel beam case

$$\mathbf{A}^T \mathbf{A} x \approx \frac{1}{r} ** x \quad \rightarrow \quad \left[\mathbf{A}^T \mathbf{A} \right]^{-1} x \approx F^{-1} \left\{ \left| \rho_{2D} \right| \right\} ** x$$

- For other cases

$$\left[\mathbf{A}^T \mathbf{A} \right]^{-1} \mathbf{A}^T l$$

- Iterative methods implicitly handle the geometry
  - Find the correct inversion for the specific geometry
  - Cannot overcome data nullspaces (needs complete data)

More Complete Forward Models

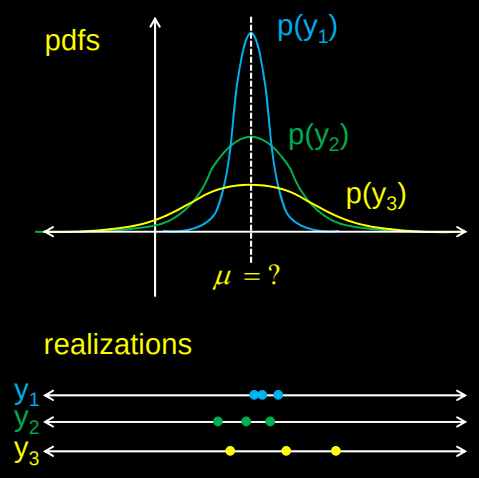
- More physics
  - Polyenergetic beam, energy-dependent attenuation
  - Detector effects, finite size elements, blur
  - Source effects, finite size element
  - Scattered radiation
- Noise
  - Data statistics
  - Quantum noise – x-ray photons
  - Detector Noise

A Simple Estimation Problem

3 Random Variables

Different std dev ($\sigma_1, \sigma_2, \sigma_3$)

Best way to estimate μ?



Maximum Likelihood Estimation

- Find the parameter values most likely to be responsible for the observed measurements.

- Likelihood Function

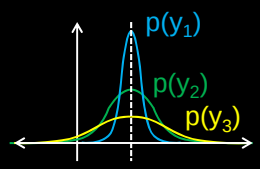
$$L(y;\mu) = p(y_1, y_2, \ldots, y_N \,|\, \mu_1, \mu_2, \ldots, \mu_M)$$

- Maximum Likelihood Objective Function

$$\hat{\mu} = \arg\max L(y;\mu)$$

ML for the Simple Example

Likelihood function: $L(y,\mu) = p(y_i \mid \mu) = \prod_{i=1}^{3} p(y_i \mid \mu)$



$$p(y_i \mid \mu) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2\sigma_i^2}(y_i - \mu)^2\right]$$

Maximize over μ: $\hat{\mu} = \arg\max L(y;\mu)$

$$\frac{\partial}{\partial \mu} L(y,\mu) = 0 \Longrightarrow \boxed{\hat{\mu} = \frac{\sum_{i=1}^{3} \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^{3} \frac{1}{\sigma_i^2}}}$$

ML for Tomography

- Need a noise model
- Depends on the statistics of the measurements
- Depends on the detection process
- Common choices
  - Poisson – x-ray photon statistics

    $y_i \sim \text{Poisson}\{\bar{y}_i(\mu)\}$ $\quad P_{y_i} = \exp[-\bar{y}_i(\mu)]\frac{[y_i(\mu)]^{y_i}}{y_i!}$
  - Poisson-Gaussian mixtures – photons + readout noise

    $y_i \sim \text{Poisson}\{y_i(\mu)\} + \text{N}\{0, \sigma_{ro}^2\}$
  - Gaussian (nonuniform variances) – approx. many effects

    $y_i \sim \text{N}\{\bar{y}_i(\mu), \sigma_i^2\}$ $\quad P_{y_i} = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2\sigma_i^2}(y_i - \bar{y}_i(\mu))^2\right)$

Gaussian Likelihood Function

- Marginal Likelihoods

$$l_i = -\log\left(\frac{y_i}{I_0}\right) \qquad p(l_i \mid \mu) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2\sigma_i^2}\left(l_i - [\mathbf{A}\mu]_i\right)^2\right]$$

- Likelihood

$$L(l \mid \mu) = p(l \mid \mu) = \prod_{i=1}^{N} p(l_i \mid \mu) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2\sigma_i^2}\left(l_i - [\mathbf{A}\mu]_i\right)^2\right]$$

- Log-Likelihood

$$\log L(l \mid \mu) \cong \sum_{i=1}^{N} -\frac{1}{2\sigma_i^2}\left(l_i - [\mathbf{A}\mu]_i\right)^2$$

- Objective Function and Estimator

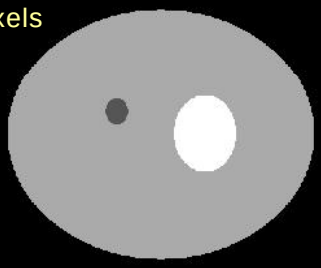
$$\hat{\mu} = \arg\max \log L(l \mid \mu) = \left[\mathbf{A}^T \mathbf{D}\left[\frac{1}{\sigma_i^2}\right]\mathbf{A}\right]^{-1} \mathbf{A}^T \mathbf{D}\left[\frac{1}{\sigma_i^2}\right] l$$

Iterative Algorithms

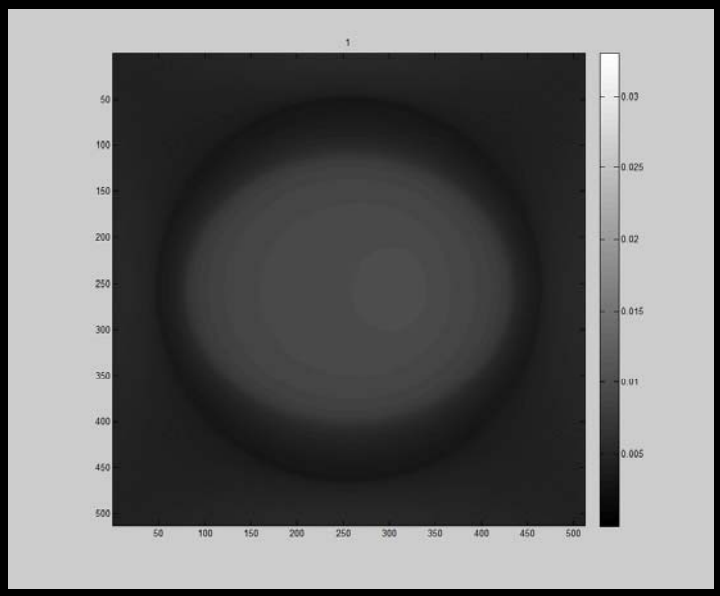
- Plethora of iterative approaches
  - Expectation-Maximization – General Purpose Methodology
  - Gradient-based Methods – Coordinate Ascent/Descent
  - Optimization Transfer – Paraboloidal Surrogates
  - Ordered-Subsets methods
- Properties of iterative algorithms
  - Monotonicity
  - True convergence
  - Speed
  - Complexity

Statistical Reconstruction Example

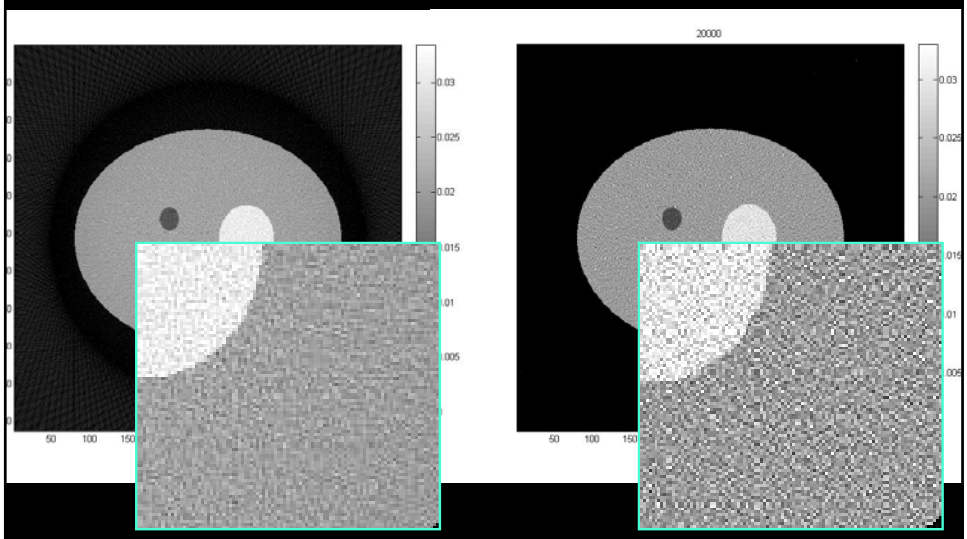
- Test Case
  - Single slice x-ray transmission problem
  - 512 x 512 0.3 mm volume
  - 400 detector bins over 180 angles (360 degrees)
  - Poisson noise: 1e5 counts per 0.5 mm detector element
  - SO: 380 mm, DO: 220 mm
- Reconstruction
  - Voxel basis: 512 x 512 0.3 mm voxels
  - Maximum-likelihood objective
  - EM-type algorithm
  - Initial image – constant value
  - Lots of iterations



ML-EM Iterations



FBP vs ML-EM Comparison



Enforcing Desirable Properties

- FBP
  - Filter designs – cutoff frequencies
- Iterative methods
  - Modify the objective to penalize "bad" images
  - Discourage noise
  - Preserve desirable image features
  - Other prior knowledge
- Image-domain Denoising

$$\hat{\mu} = \arg\max F(\mu) - \beta R(\mu)$$

- Model-based Reconstruction

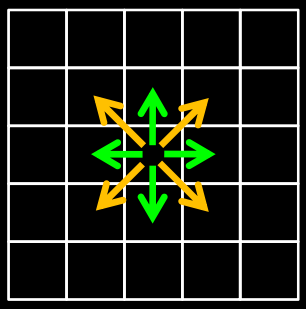
$$\hat{\mu} = \arg\max F(y, \mu) - \beta R(\mu)$$

Local Control of Image Properties

- Pairwise Penalty
  - Penalize the difference between neighboring voxels

$$R(\mu) = \sum_j \sum_{k \in N} w_{jk} \psi\left(\mu_j - \mu_k\right)$$

$$k_{1st} = \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix} \quad k_{2nd} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & -1 & \frac{-1}{\sqrt{2}} \\ -1 & 4 + \frac{4}{\sqrt{2}} & -1 \\ \frac{-1}{\sqrt{2}} & -1 & \frac{-1}{\sqrt{2}} \end{bmatrix}$$

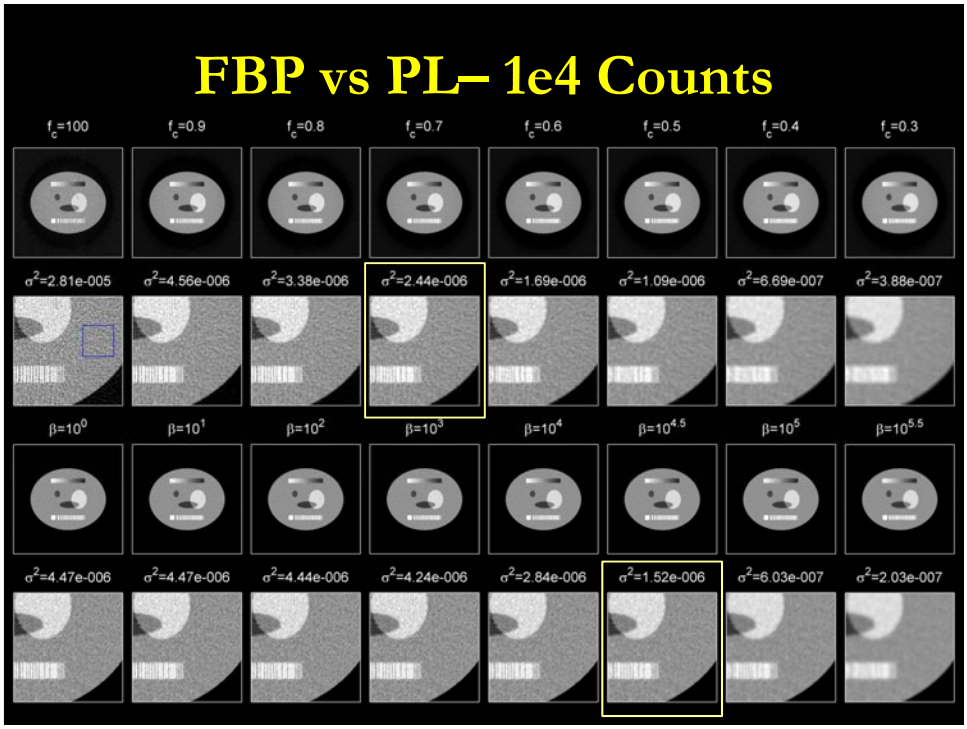


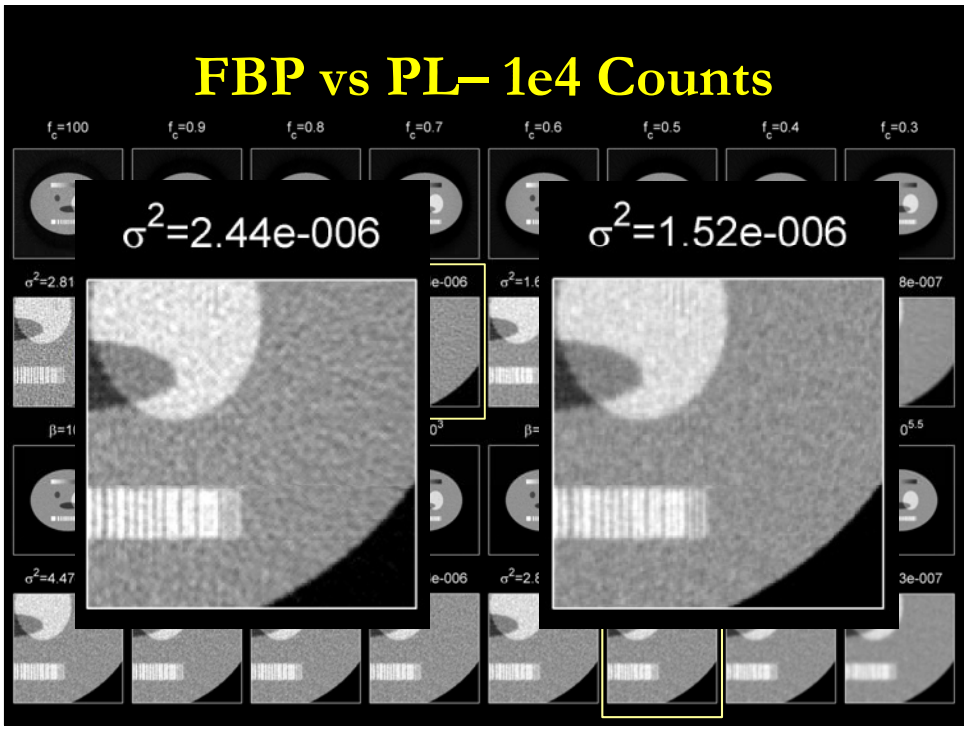
- Quadratic Penalty

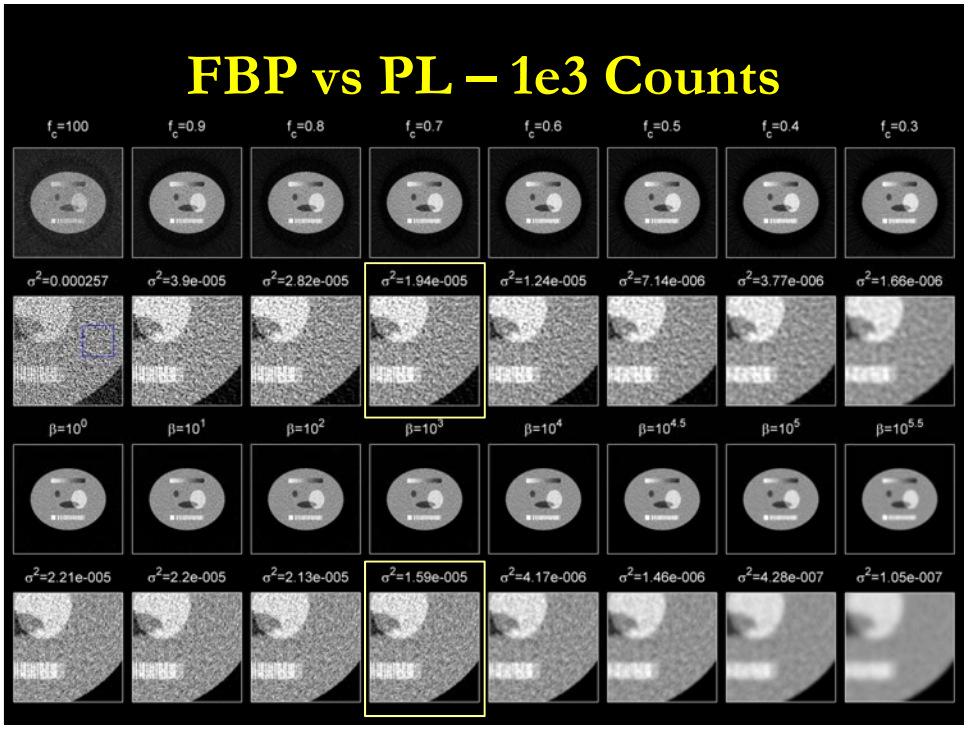
$$R(\mu) = \sum_j \sum_{k \in N} w_{jk} \left(\mu_j - \mu_k\right)^2$$

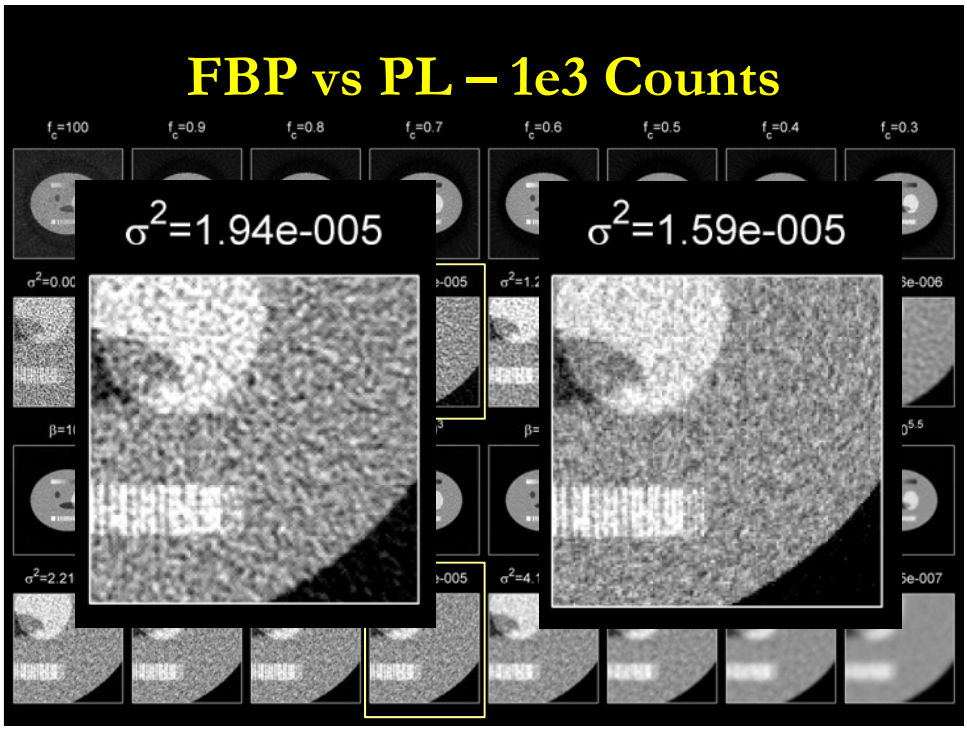
Penalized-Likelihood Example

- Forward Model
  - Fan-beam Transmission Tomography
  - 512 x 512 0.3 mm volume
  - 400 detector bins over 180 angles (360 degrees)
  - Poisson: {1e4,1e3} counts per 0.5 mm detector element
  - SO: 380 mm, DO: 220 mm
- Objective Function
  - Poisson Likelihood
  - Quadratic, first-order Penalty
- Algorithm
  - Separable Paraboloidal Surrogates
  - 400 iterations – well-converged


FBP vs PL– 1e4 Counts
f_c=100
f_c=0.9
f_c=0.8
f_c=0.7
f_c=0.6
f_c=0.5
f_c=0.4
f_c=0.3
σ^2=2.81e-005
σ^2=4.56e-006
σ^2=3.38e-006
σ^2=2.44e-006
σ^2=1.69e-006
σ^2=1.09e-006
σ^2=6.69e-007
σ^2=3.88e-007
$\beta=10^0$
$\beta=10^1$
$\beta=10^2$
$\beta=10^3$
$\beta=10^4$
$\beta=10^{4.5}$
$\beta=10^5$
$\beta=10^{5.5}$
σ^2=4.47e-006
σ^2=4.47e-006
σ^2=4.44e-006
σ^2=4.24e-006
σ^2=2.84e-006
σ^2=1.52e-006
σ^2=6.03e-007
σ^2=2.03e-007


FBP vs PL– 1e4 Counts
f_c=100
f_c=0.9
f_c=0.8
f_c=0.7
f_c=0.6
f_c=0.5
f_c=0.4
f_c=0.3
σ^2=2.44e-006
σ^2=1.52e-006


FBP vs PL – 1e3 Counts
f_c=100
f_c=0.9
f_c=0.8
f_c=0.7
f_c=0.6
f_c=0.5
f_c=0.4
f_c=0.3
σ²=0.000257
σ²=3.9e-005
σ²=2.82e-005
σ²=1.94e-005
σ²=1.24e-005
σ²=7.14e-006
σ²=3.77e-006
σ²=1.66e-006
β=10⁰
β=10¹
β=10²
β=10³
β=10⁴
β=10^4.5
β=10⁵
β=10^5.5
σ²=2.21e-005
σ²=2.2e-005
σ²=2.13e-005
σ²=1.59e-005
σ²=4.17e-006
σ²=1.46e-006
σ²=4.28e-007
σ²=1.05e-007


FBP vs PL – 1e3 Counts
f_c=100
f_c=0.9
f_c=0.8
f_c=0.7
f_c=0.6
f_c=0.5
f_c=0.4
f_c=0.3
σ²=1.94e-005
σ²=1.59e-005

Other Penalties/Energy Functions

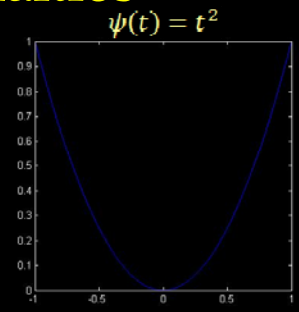
$$R(\mu) = \sum_j \sum_{k \in N} w_{jk} \psi(\mu_j - \mu_k)$$

- Quadratic penalty
  - Tends to enforce smoothness throughout image
  - Increasingly penalizes larger pixel differences
- Non-quadratic penalties
  - Attempt to preserve edges in the image
  - Once pixel differences become large allow for decreased penalty (perhaps a relative decrease)
- Flexibility: Even more penalties
  - Wavelets and other bases, non-local means, etc.

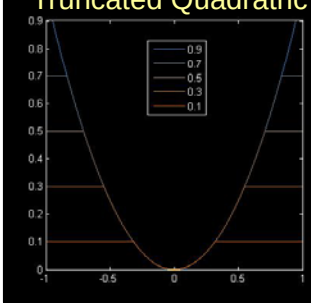
Nonquadratic Penalties

- Choices
  - Truncated Quadratic
    $\psi(t;\delta) = [t^2]_\delta$
  - Lange Penalty
    $\psi(t;\delta) = \delta \left|\frac{t}{\delta}\right| - \log\left[1 + \left|\frac{t}{\delta}\right|\right]$
  - P-Norm
    $\psi(t;p) = \frac{1}{p}|t|^p$

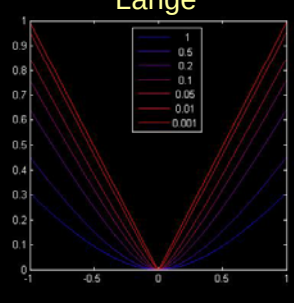
$\psi(t) = t^2$



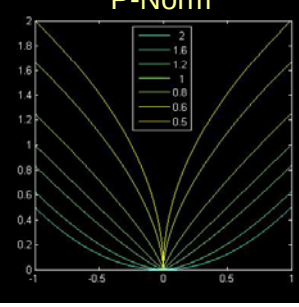
Truncated Quadratric

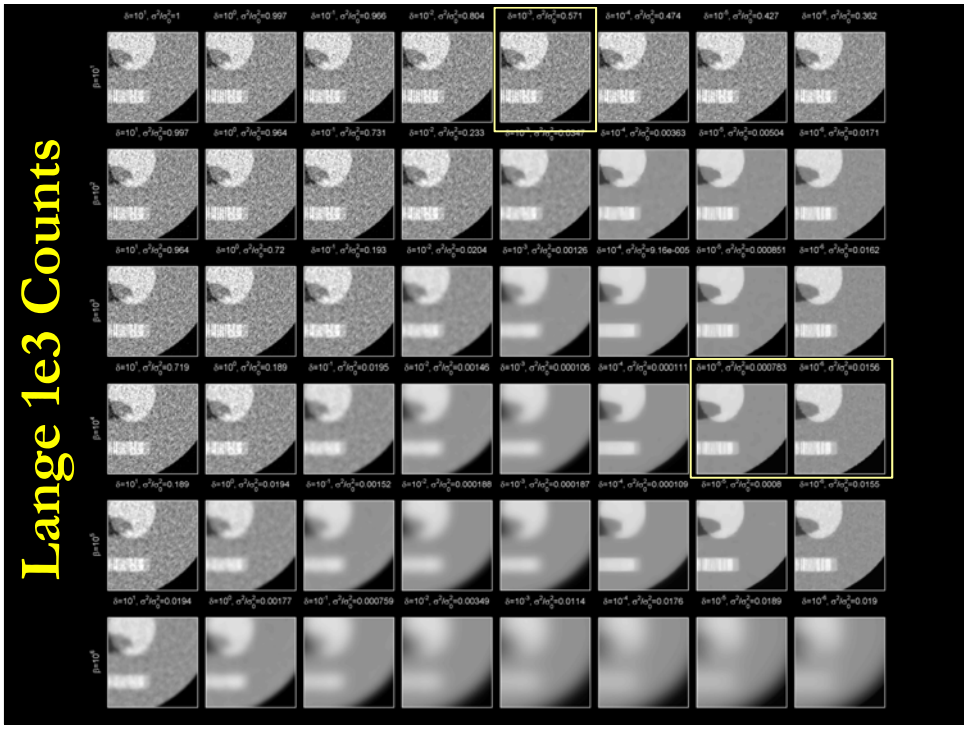


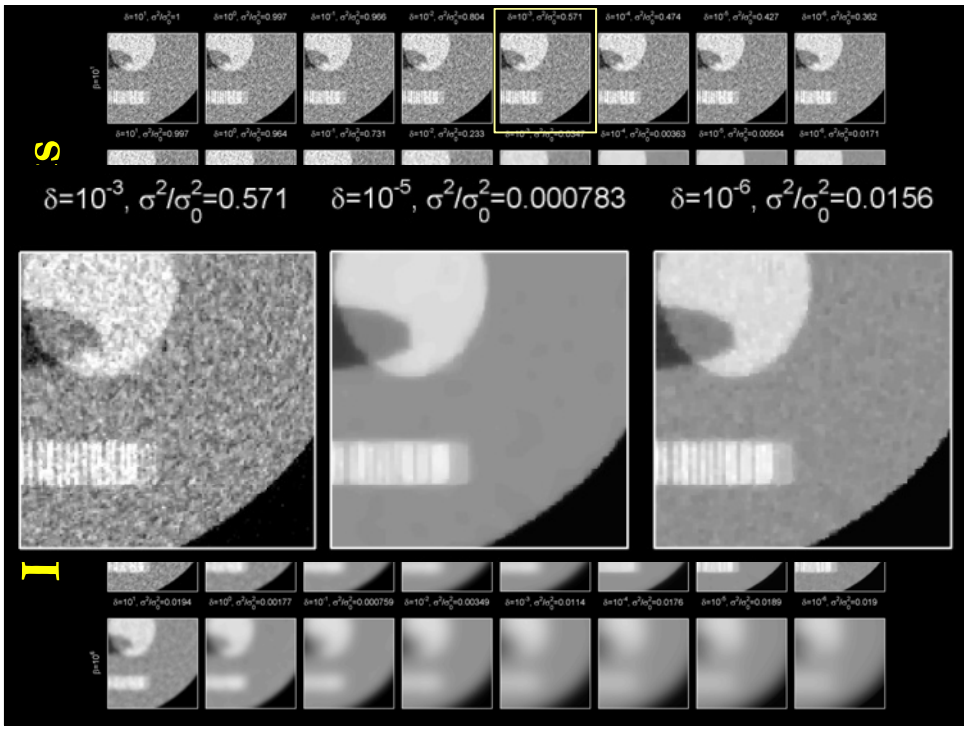
Lange



P-Norm

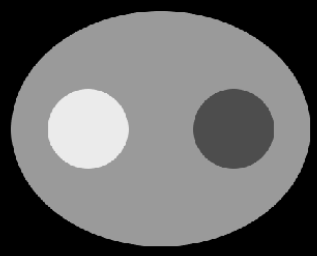



Lange 1e3 Counts


δ=10^-3, σ^2/σ_0^2=0.571
δ=10^-5, σ^2/σ_0^2=0.000783
δ=10^-6, σ^2/σ_0^2=0.0156

Closer Look at Image Properties

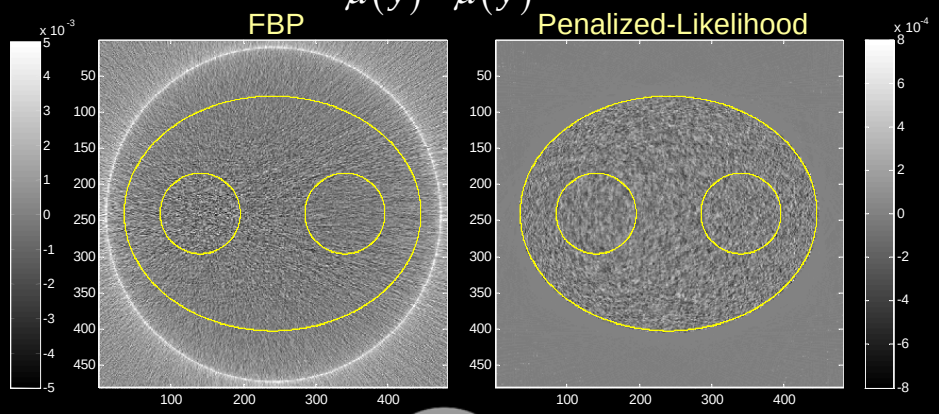
- Test Case
  - Single slice x-ray transmission problem
  - 480 x 480 0.8 mm volume
  - 1000 detector bins over 360 angles (360 degrees)
  - Poisson noise: 1e5 counts / 0.76 mm detector element
  - SO: 600 mm, DO: 600 mm



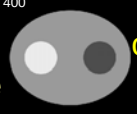
- Reconstruction
  - Penalized-likelihood objective
  - Shift-Invariant Quadratic Penalty
  - Separable paraboloidal surrogates
  - 200 iterations
  - Voxel basis: 480 x 480 0.8 mm voxels

FBP vs PL, Noise Properties

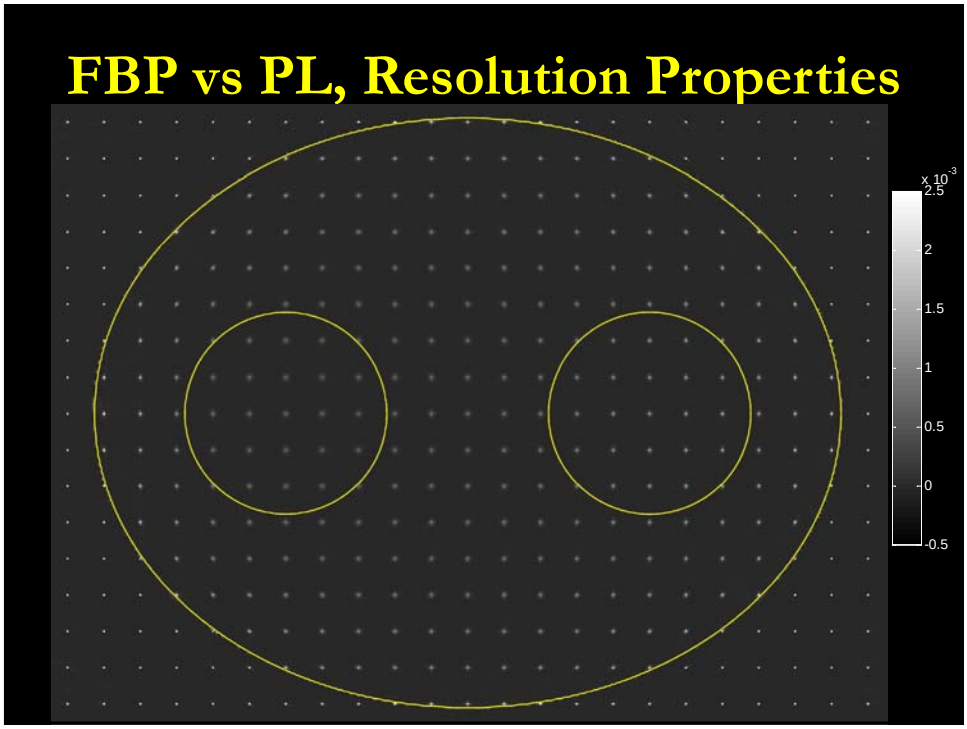
$$\hat{\mu}(y) - \hat{\mu}(\bar{y})$$

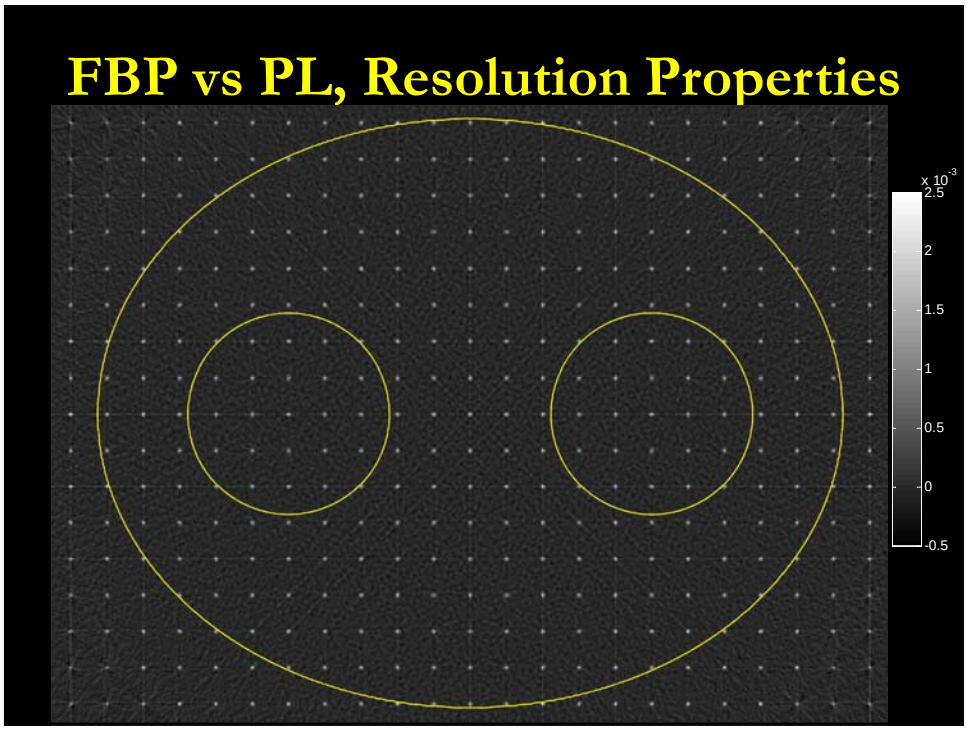


- Noise in FBP
  - Shift-variant variance
  - Shift-variant covariance



- Noise in Quadratic PL
  - Relatively shift-invariant variance (in object)
  - Shift-variant covariance


FBP vs PL, Resolution Properties
x 10^-3
2.5
2
1.5
1
0.5
0
-0.5


FBP vs PL, Resolution Properties
x 10^-3
2.5
2
1.5
1
0.5
0
-0.5

PL Image Properties – Thorax

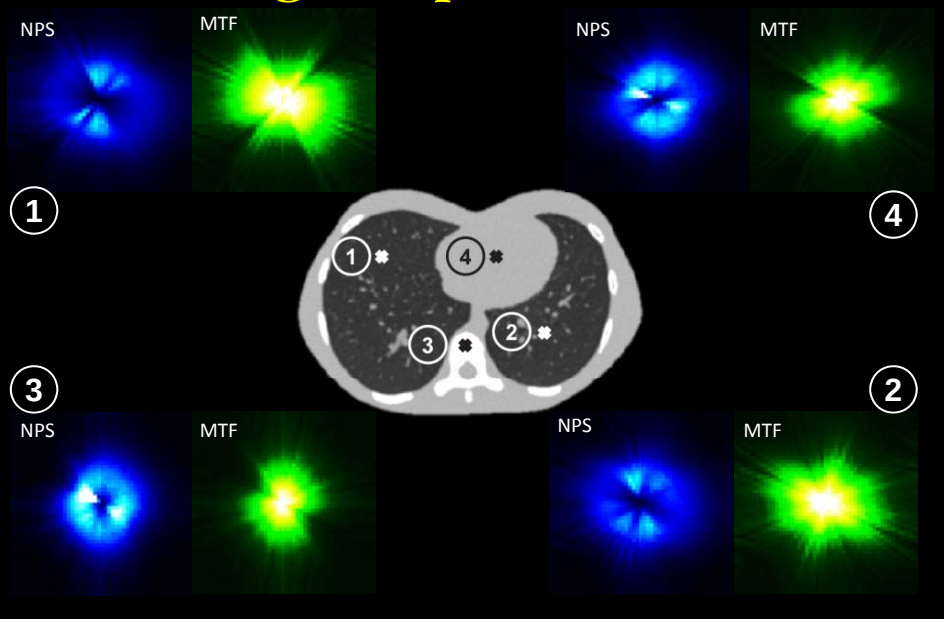


Image Properties

- Filtered-Backprojection
  - Largely shift-invariant spatial resolution
  - Shift-variant, object-dependent noise
- Uniform Quadratic Penalized Likelihood
  - Shift-variant, object-dependent spatial resolution
  - Shift-variant, object-dependent noise
- Edge-preserving Penalty Methods
  - Shift-variant, object-dependent spatial resolution
  - Shift-variant, object-dependent noise
  - Noise-resolution properties may not even be locally smooth

Learning Objectives I

- Fundamentals of iterative methods
  - Approach to forming an iterative algorithm
    - Forward Model
    - Objective Function
    - Optimization Algorithm
  - Identify particular classes of methods
    - Model-based vs. Image Denoising approaches
    - Statistical vs. Nonstatistical approaches
    - Kinds of regularization

Learning Objectives II

- Advantages of iterative approaches
  - Intuition behind what these algorithms do
    - Fitting reconstruction to observations
    - Data weighting by information content
    - Importance of regularization
  - Flexibility of these techniques
    - Arbitrary geometries
    - Sophisticated modeling of physics
    - General incorporation of desired image properties through regularization

Learning Objectives III

- Images produced by iterative methods
  - Differences from traditional reconstruction
    - Regularization is key
    - Image properties are tied to statistical weighting
    - Can depend on algorithm when iterative solution has not yet converged
  - Image properties associated with iterative reconstruction
    - Highly dependent on regularization
    - Data- and object dependent
    - Shift-variant; different noise, texture than FBP

Acknowledgements

I-STAR Laboratory

Imaging for Surgery, Therapy, and Radiology
www.jhu.edu/istar

Faculty and Scientists
Jeff Siewerdsen
Yoshito Otake
Sebastian Schafer
Adam Wang
Wojtek Zbijewski

BME Students
Hao Dang
Grace Gang
Sajendra Nithiananthan
Steven Tilly II
Jennifer Xu

CS Students
Sureerat Reaungamornrat
Ali Uneri

Hopkins Collaborators

School of Medicine
John A Carrino (Radiology)
Gary Gallia (Neurosurgery)
A Jay Khanna (Orthopaedic Surgery)
Martin Radvany (Interventional Neuroradiology)
Doug Reh (Oto H&N Surgery)
Mahadevappa Mahesh (Radiology)
Marc Sussman (Thoracic Surgery)

School of Engineering
Greg Hager (CS)
Junghoon Lee (CS)
Jerry Prince (ECE)
Russell Taylor (CS)

Support

National Institutes of Health
Carestream Health
Varian Medical Systems