Statistical methods for X-ray CT image reconstruction

Jeffrey A. Fessler

EECS Department
BME Department
Department of Radiology
The University of Michigan

GEMS CT

June 13, 2001

Outline

- Group/Lab
- Statistical image reconstruction

  Choices / tradeoffs / considerations:
  - 1. Object parameterization
  - 2. System physical modeling
  - 3. Statistical modeling of measurements
  - 4. Objective functions and regularization
  - 5. Iterative algorithms

  Short course lecture notes:

  `http://www.eecs.umich.edu/~fessler/talk`
- Statistical reconstruction for X-ray CT with beam hardening
- Incomplete data tomography
- Future goals

Students

• El Bakri, Idris	X-ray CT image reconstruction
• Ferrise, Gianni	Signal processing for direct brain interface
• Jacobson, Matt	PET image reconstruction
• Kim, Jeongtae	Image registration/reconstruction for radiotherapy
• Naik, Vipul	Bioluminescence imaging
• Stayman, Web	Regularization methods for tomographic reconstruction
• Sotthivirat, Saowapak	Optical image restoration
• Sutton, Brad	MRI image reconstruction
• Yendiki, Anastasia	Regularization methods for image reconstruction

Image computing laboratory, Department of Electrical Engineering and Computer Science

Collaborations with colleagues in Biomedical Engineering, EECS, Nuclear Engineering, Nuclear Medicine, Radiology, Radiation Oncology, Physical Medicine, Anatomy and Cell Biology, Biostatistics

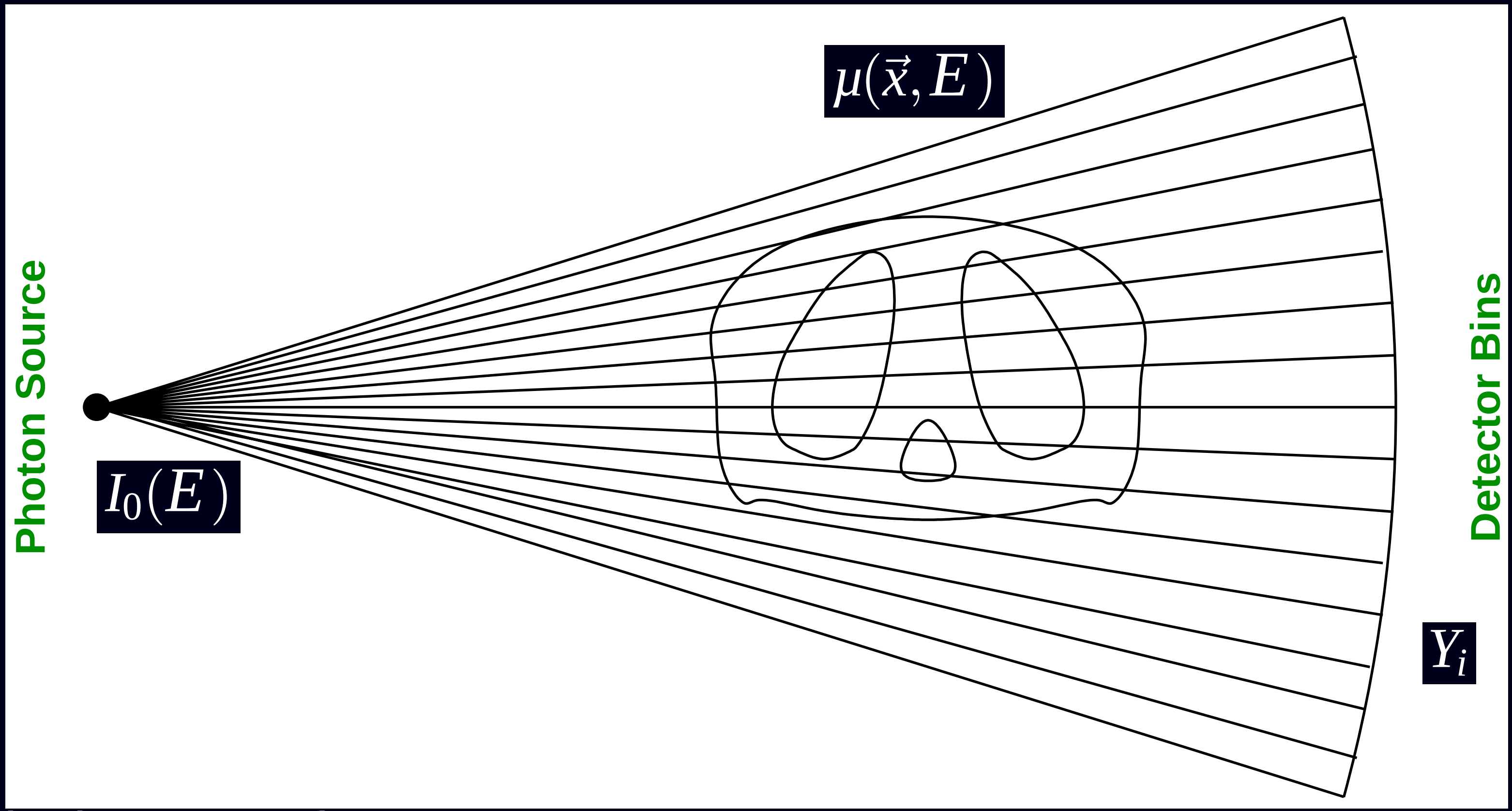
Research Goals

- Develop methods for making “better” images
  (modeling of imaging system physics and measurement statistics)
- Faster algorithms for computing/processing images
- Analysis of the properties of image formation methods
- Design of imaging systems based on performance bounds

Impact

- ASPIRE (A sparse iterative reconstruction environment) software
  (about 40 registered sites worldwide)
- PWLS reconstruction used routinely for cardiac SPECT at UM,
  following 1996 ROC study. (several thousand patients scanned)
- Pittsburgh PET/CT “side information” scans reconstructed using ASPIRE
- Consulted for GEMS/PET for 2D and 3D OSEM implementation.

X-ray CT Data Collection



Transmission scanning geometry.

X-ray CT Reconstruction Problem - Illustration

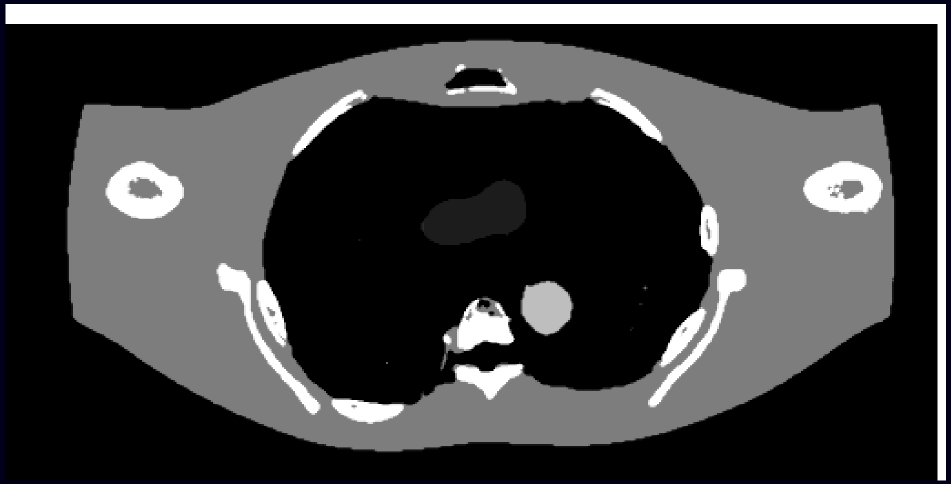
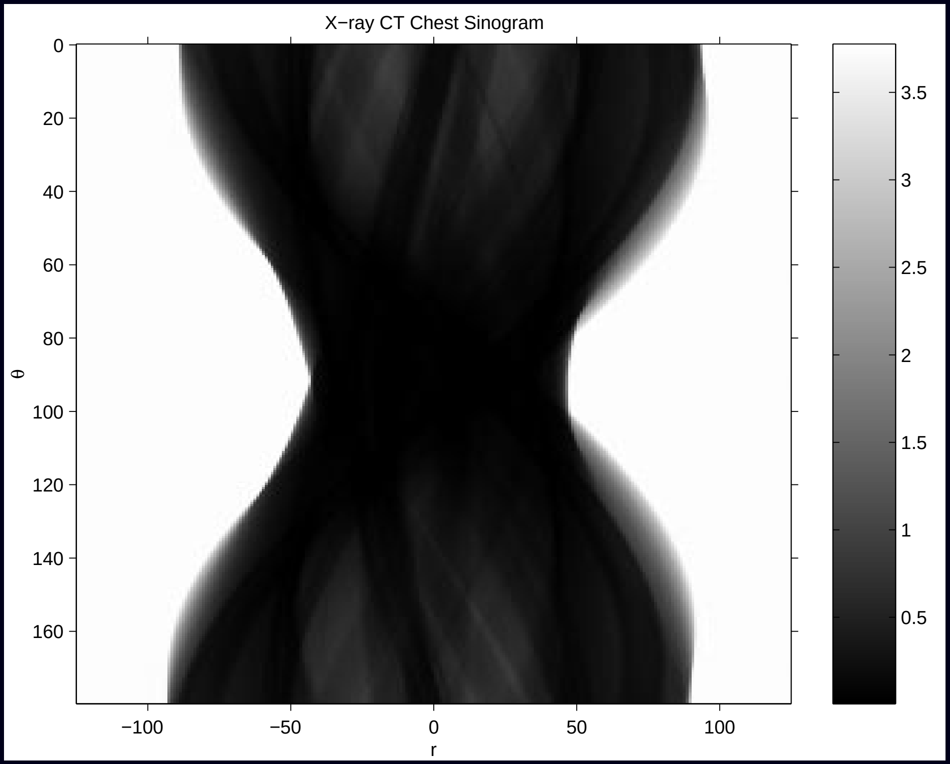


Image
$\mu(\vec{x}, E_{\text{eff}})$



Sinogram
$\{Y_i\}$

Why Statistical/Iterative Methods?

Physics of imaging

- Reduced *artifacts*, increased quantitative *accuracy*
  source spectrum, beam hardening, scatter, ...
- System detector response models (*possibly* improved spatial resolution)

Statistics

- Appropriate statistical models (reduced image noise and/or *dose*)
- FBP treats all rays equally

Geometry

- Non-Radon geometries (helical, cone-beam)
- “Missing” data, *e.g.*, truncation (long object)
- Gated cardiac helical CT

Prior knowledge

- Object constraints (*e.g.*, nonnegativity)
- Material properties

Tradeoffs...

- Computation time
  (solution: ordered-subset algorithms and multiprocessor computing)
- Model complexity
- Software / algorithm complexity
- Complexity of analysis of nonlinear methods

Reconstruction Methods

(Simplified View)

Analytical (FBP)	**Iterative** (OSEM?)

Reconstruction Methods

- **ANALYTICAL**
  - FBP
  - BPF
  - Gridding
  - ...
- **ITERATIVE**
  - **Algebraic** (y = Ax)
    - ART
    - MART
    - SMART
    - ...
  - **Statistical**
    - **Least Squares**
      - CG
      - CD
      - ISRA
      - ...
    - **Poisson Likelihood**
      - EM (etc.)
      - OSEM
      - SAGE
      - CG
      - Int. Point
      - GCA
      - PSCD
      - FSCD
      - ...

Five Categories of Choices

1. Object parameterization: $\lambda(\vec{x})$ or $\mu(\vec{x}, E)$ vs $\boldsymbol{f} = (f_1, \ldots, f_{n_p}) \in \mathbb{R}^{n_p}$
2. Model for system physics
3. Measurement statistical model $Y_i \sim \boxed{?}$
4. Objective function: data-fit / regularization
5. Algorithm / initialization

No perfect choices - one can critique all approaches!

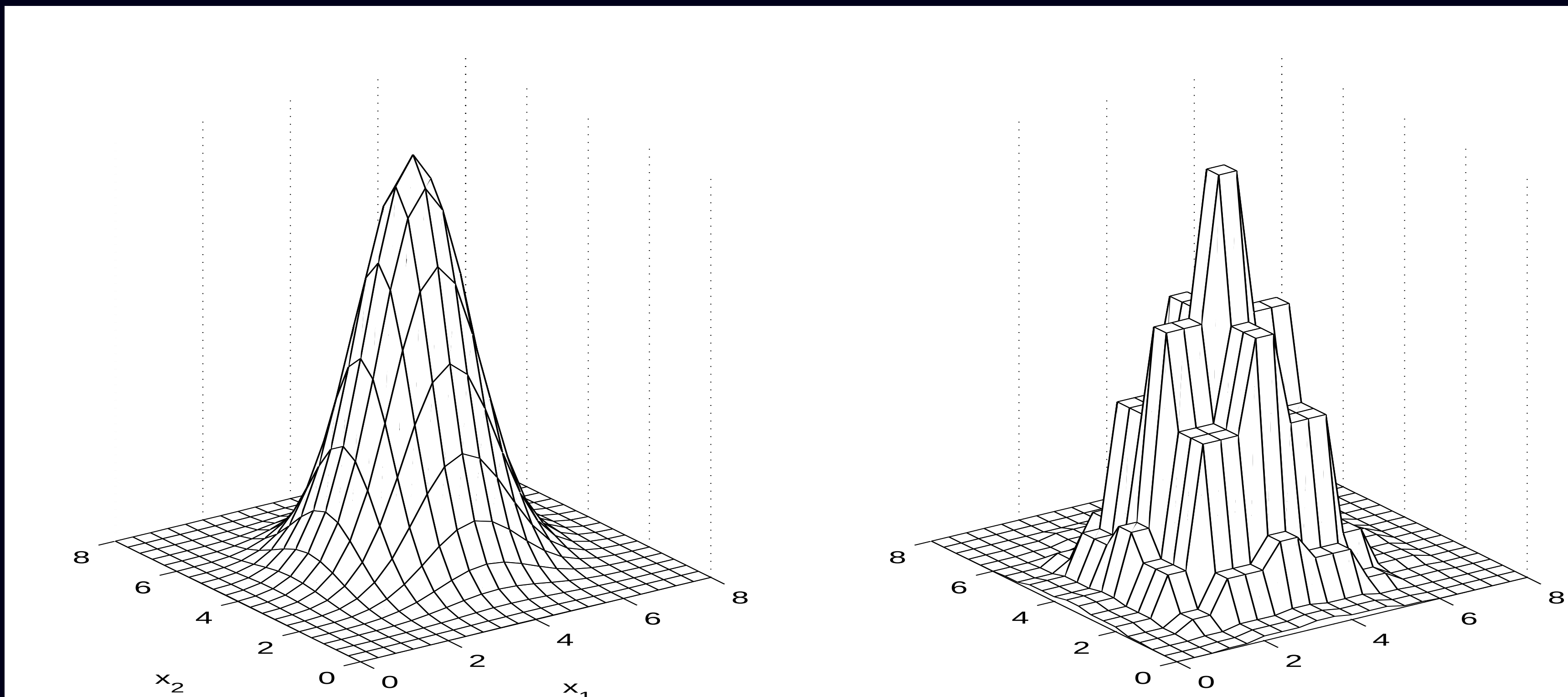
Choices impact:

- Image spatial resolution
- Image noise
- Quantitative accuracy
- Computation time
- Memory
- Algorithm complexity

Choice 1. Object Parameterization (Transmission)

Series expansion

$$\text{attenuation map} \;\rightarrow\; \mu(\vec{x}) \approx f(\vec{x}) = \sum_{j=1}^{n_p} f_j \, b_j(\vec{x}) \;\leftarrow\; \text{“basis functions”}$$



Object $\mu(\vec{x})$ Pixelized approximation $f(\vec{x})$

Basis Functions

Choices

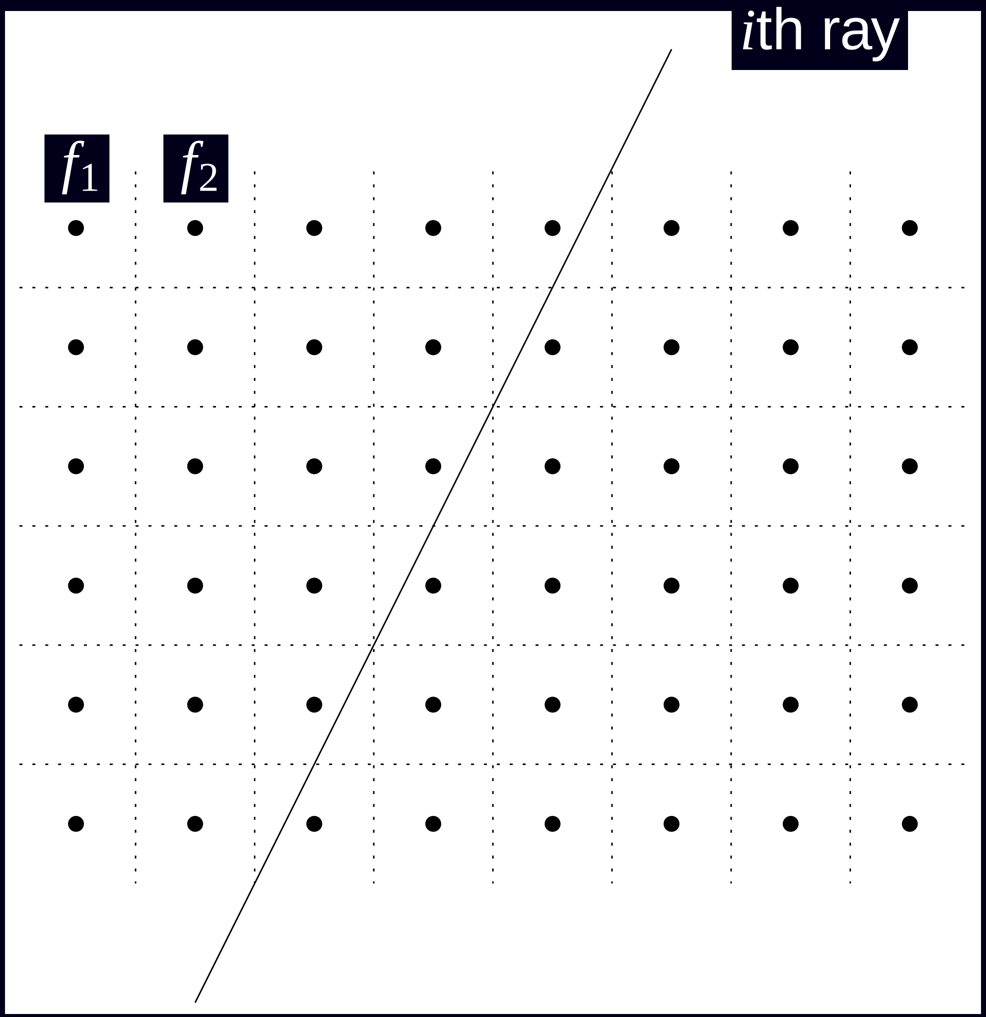
- Fourier series
- Circular harmonics
- Wavelets
- Kaiser-Bessel windows
- Overlapping disks
- B-splines (pyramids)
- Polar grids
- Logarithmic polar grids
- "Natural pixels"
- Point masses (Dirac δ's)
- pixels / voxels
- ...

Considerations

- Represent object "well" with moderate n_p
- system matrix elements $\{a_{ij}\}$ "easy" to compute
- The $n_d \times n_p$ system matrix: $\boldsymbol{A} = \{a_{ij}\}$, should be sparse (mostly zeros).
- Easy to represent nonnegative functions

  e.g., if $f_j \geq 0$, then $f(\vec{x}) \geq 0$, *i.e.*, $b_j(\vec{x}) \geq 0$.

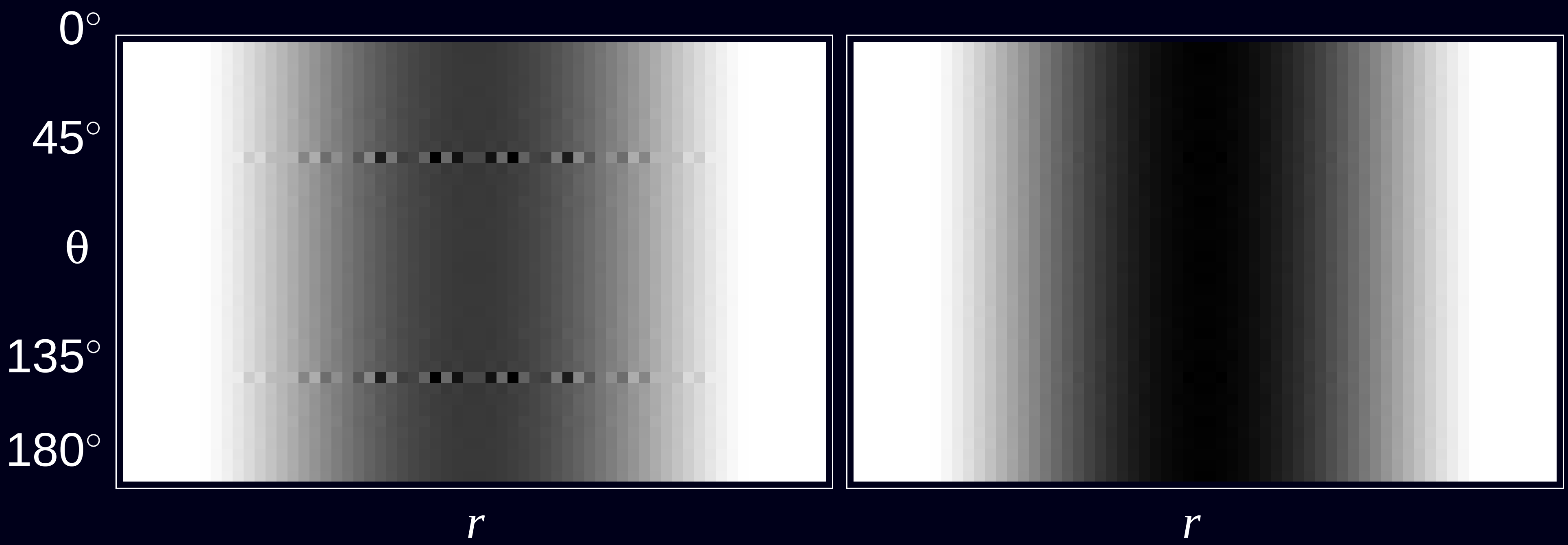
Point-Lattice Projector/Backprojector



Implicit in conventional pixel-driven backprojection used for FBP.
System matrix elements (a_{ij}'s) determined by linear interpolation.

Point-Lattice Artifacts

Forward projections (sinograms) of a uniform disk object:



Basis:	Point Lattice	Square pixels
System:	Linear Interpolation	Strip Integrals

Choice 2. System Model (Transmission)

System matrix $A = \{a_{ij}\}$

- scanner **geometry**
- detector width
- source size
- detector response
- collimation
- ...

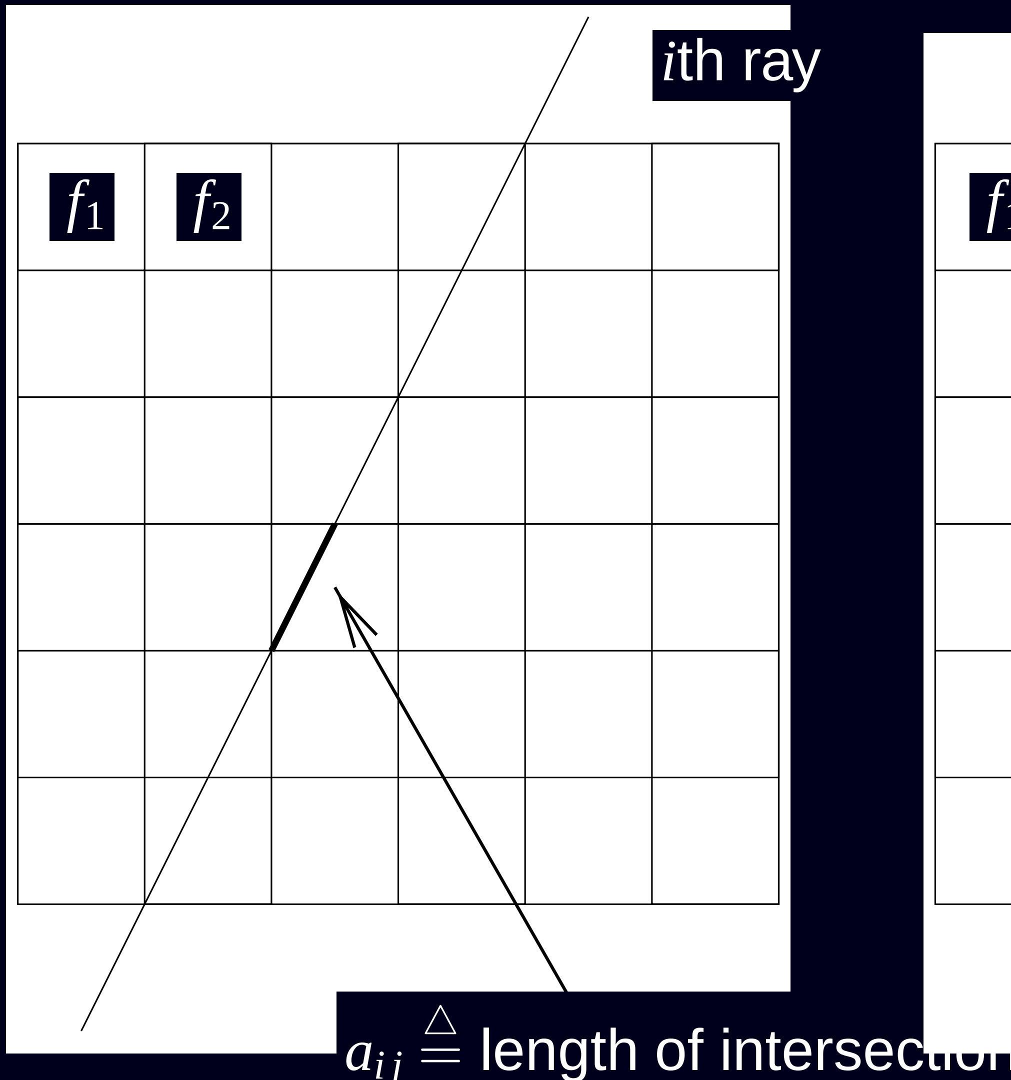
Other physical effects

- source spectrum
- detector efficiency
- Beer's law
- scatter
- detector after-glow?
- ...

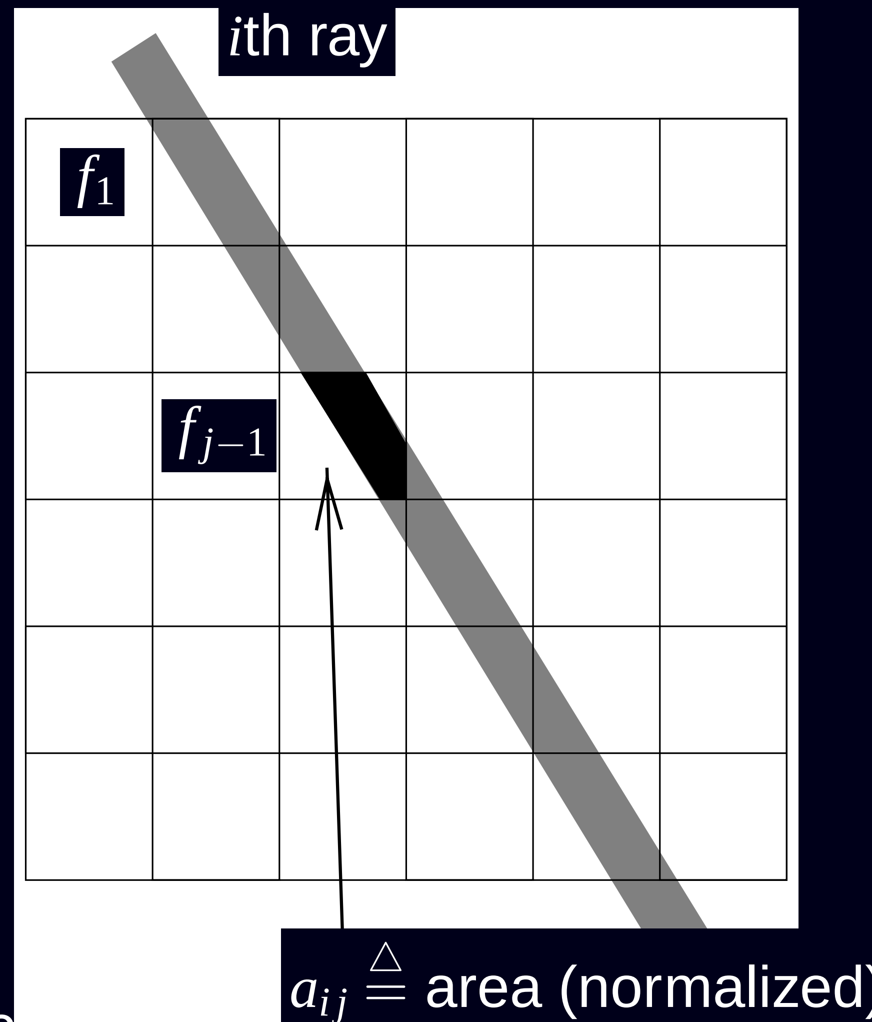
Considerations

- Accuracy vs computation
- Store a_{ij}'s or on-the-fly computing of forward and backprojection?
- Model uncertainties
  (*e.g.*, drifts in source spectrum with tube heating)
- Artifacts due to over-simplifications

“Line Length” System Model



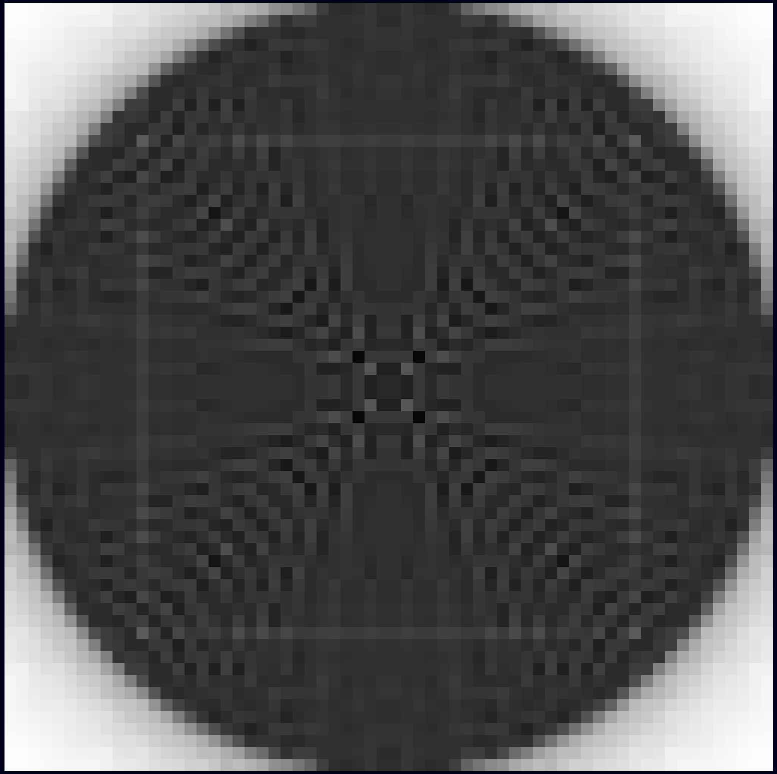
“Strip Area” System Model



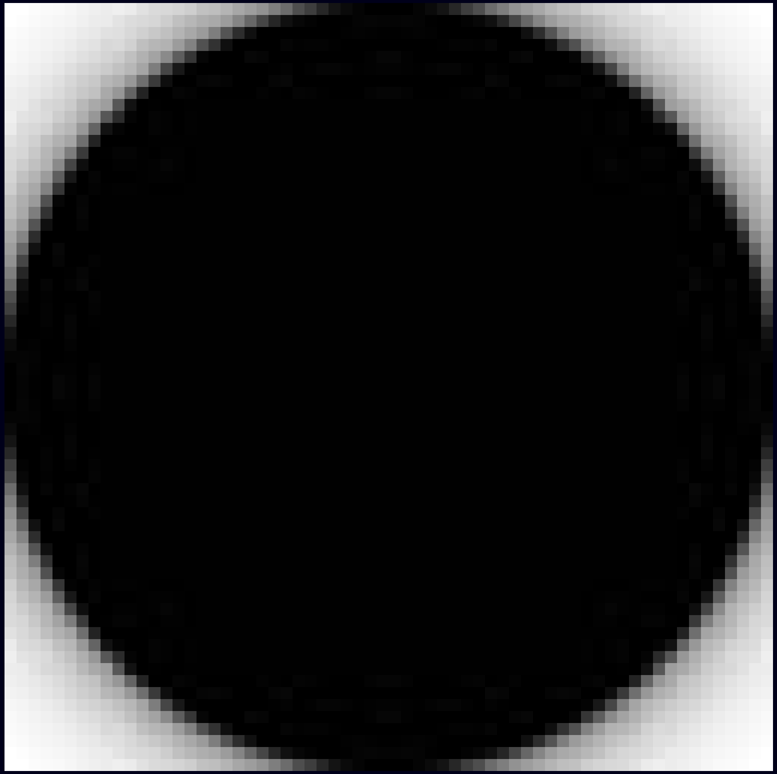
Sensitivity Patterns (Emission)

$$\sum_{i=1}^{n_d} a_{ij} \approx s(\vec{x}_j) = \sum_{i=1}^{n_d} s_i(\vec{x}_j)$$

Line Length



Strip Area



Forward- / Back-projector "Pairs"

Forward projection (image domain to projection domain):

$$\boldsymbol{A}\boldsymbol{f} = \left\{ \sum_{j=1}^{n_p} a_{ij} f_j \right\}_{i=1}^{n_d}.$$

Backprojection (projection domain to image domain):

$$\boldsymbol{A}'\boldsymbol{y} = \left\{ \sum_{i=1}^{n_d} a_{ij} y_i \right\}_{j=1}^{n_p}.$$

Too often $\boldsymbol{A}'\boldsymbol{y}$ is implemented as $\boldsymbol{B}\boldsymbol{y}$ for some "backprojector" $\boldsymbol{B} \neq \boldsymbol{A}'$.

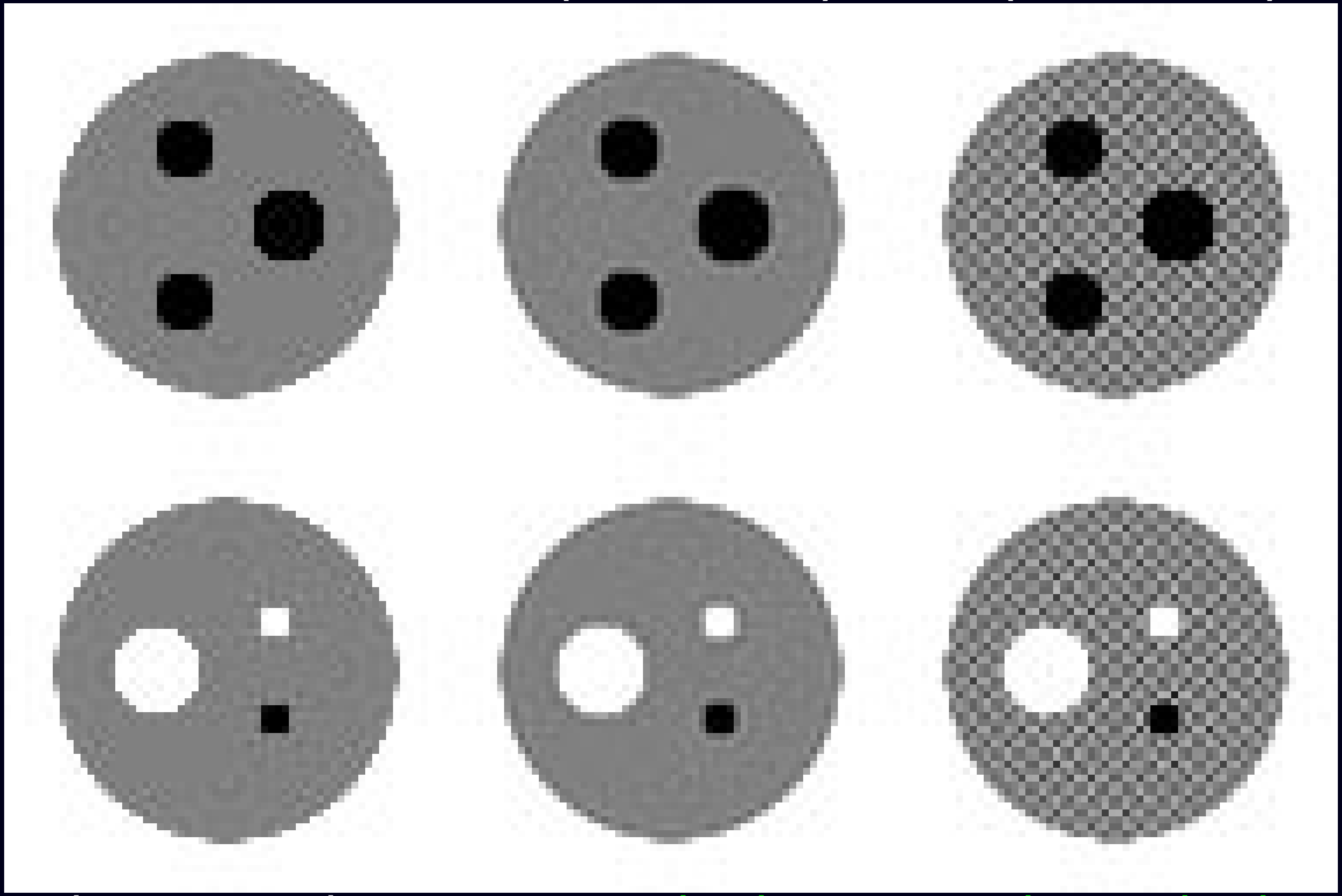
Least-squares solutions (for example):

$$\hat{\boldsymbol{f}} = [\boldsymbol{A}'\boldsymbol{A}]^{-1}\boldsymbol{A}'\boldsymbol{y} \neq [\boldsymbol{B}\boldsymbol{A}]^{-1}\boldsymbol{B}\boldsymbol{y}.$$

Mismatches accumulate with iterations!

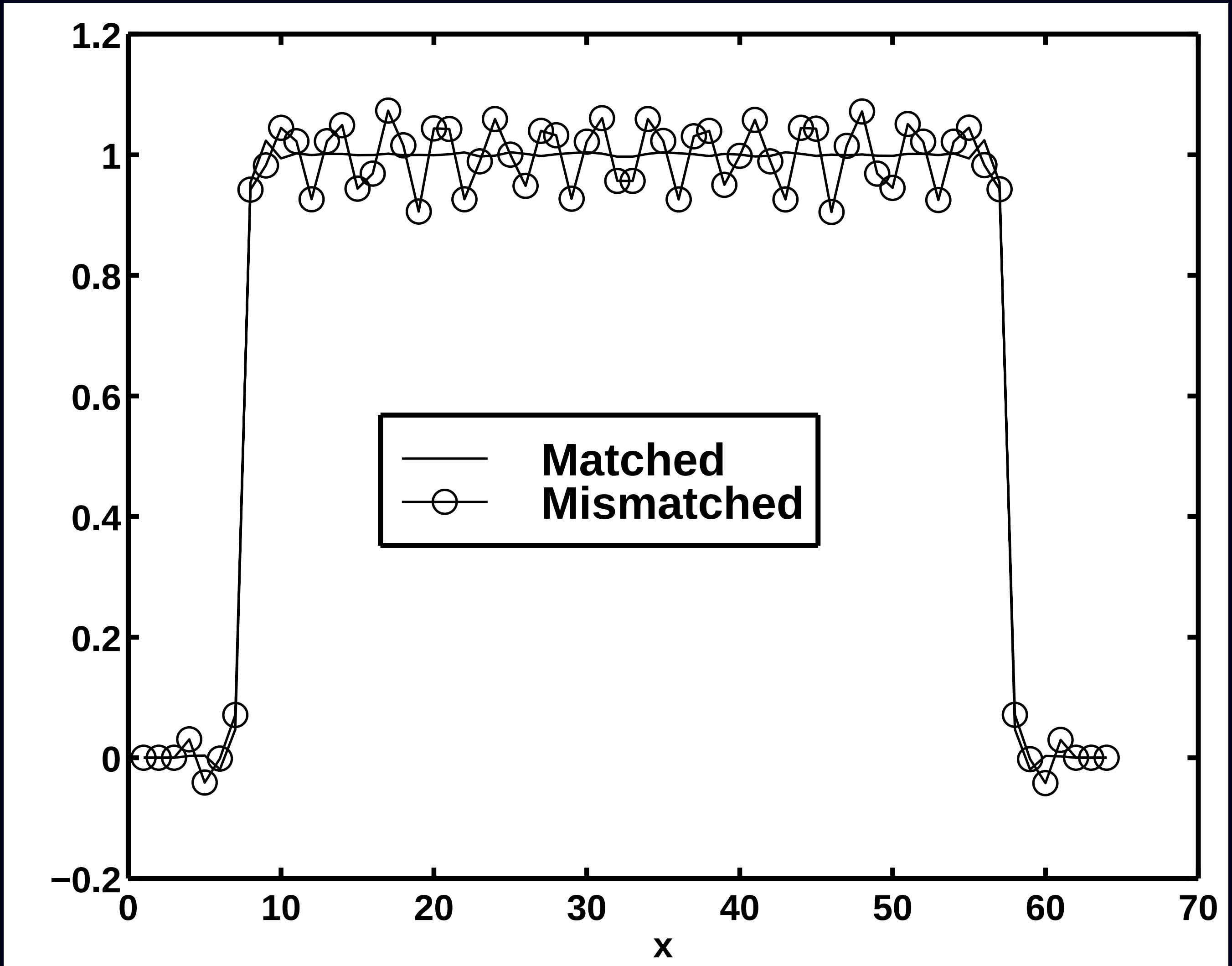
Mismatched Backprojector $B \neq A'$ (3D PET)

$\boldsymbol{\lambda}$ | $\hat{\boldsymbol{\lambda}}$ (PWLS-CG) | $\hat{\boldsymbol{\lambda}}$ (PWLS-CG)



$(64 \times 64 \times 4)$ | Matched | Mismatched

Horizontal Profiles



Choice 3. Statistical Models (Transmission)

After modeling the system physics, we have a deterministic "model:"

$$Y_i \approx E[Y_i] = \bar{y}_i(\boldsymbol{f}) + r_i = \int I_0(E) \exp\left(-\sum_{j=1}^{n_p} a_{ij} f_j(E) \right) dE + r_i.$$

Statistical modeling is concerned with the " $\approx$ " aspect.

- $I_0(E)$: source spectrum
- r_i: scatter background, etc.

Random Phenomena

- Number of photons
- Photon energy
- Photon absorption
- Compton scatter
- Detection probability
- Readout noise
- ...

Statistical Model Considerations

- More accurate models:
  - can lead to lower variance images,
  - can reduce bias
  - may incur additional computation,
  - may involve additional algorithm complexity

  (*e.g.*, transmission Poisson model can have nonconcave log-likelihood)
- Statistical model errors (*e.g.*, deadtime)
- Incorrect models (*e.g.*, log-processed transmission data)

Statistical Model Choices (Transmission)

- "None." Assume $Y_i - r_i = \bar{y}_i(\boldsymbol{f})$. "Solve algebraically" to find $\boldsymbol{f}$.
- White Gaussian noise. Ordinary least squares: minimize $\|(\boldsymbol{Y}-\boldsymbol{r})-\bar{\boldsymbol{y}}(\boldsymbol{f})\|^2$
- Non-White Gaussian noise. Weighted least squares: minimize

$$\|(\boldsymbol{Y}-\boldsymbol{r})-\bar{\boldsymbol{y}}(\boldsymbol{f})\|_W^2 = \sum_{i=1}^{n_d} w_i\,(y_i - r_i - \bar{y}_i(\boldsymbol{f}))^2.$$

- Ordinary Poisson model (ignoring or precorrecting for background)

$$Y_i \sim \text{Poisson}\{\bar{y}_i(\boldsymbol{f})\}.$$

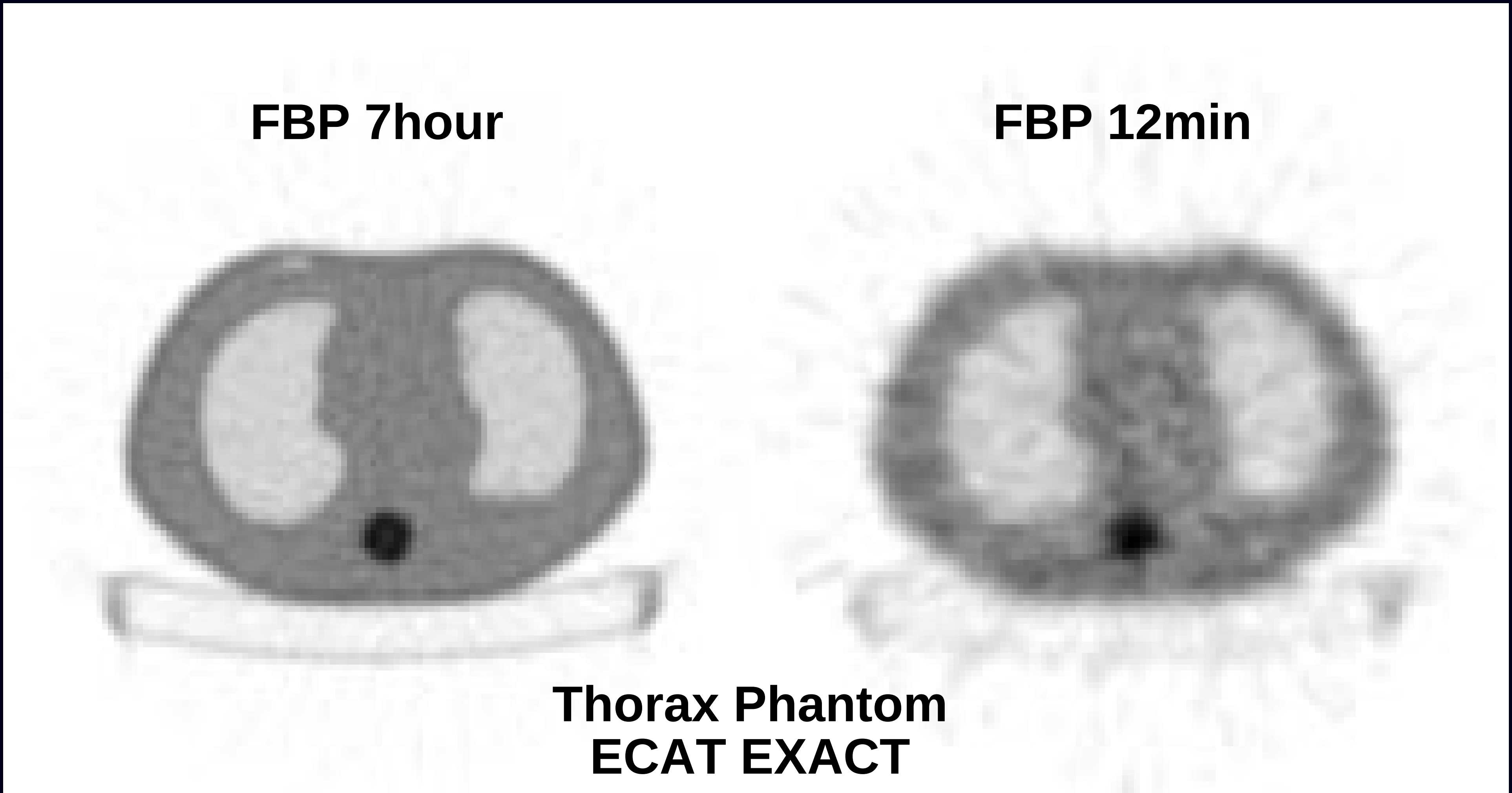
- Poisson model

$$Y_i \sim \text{Poisson}\{\bar{y}_i(\boldsymbol{f}) + r_i\}.$$

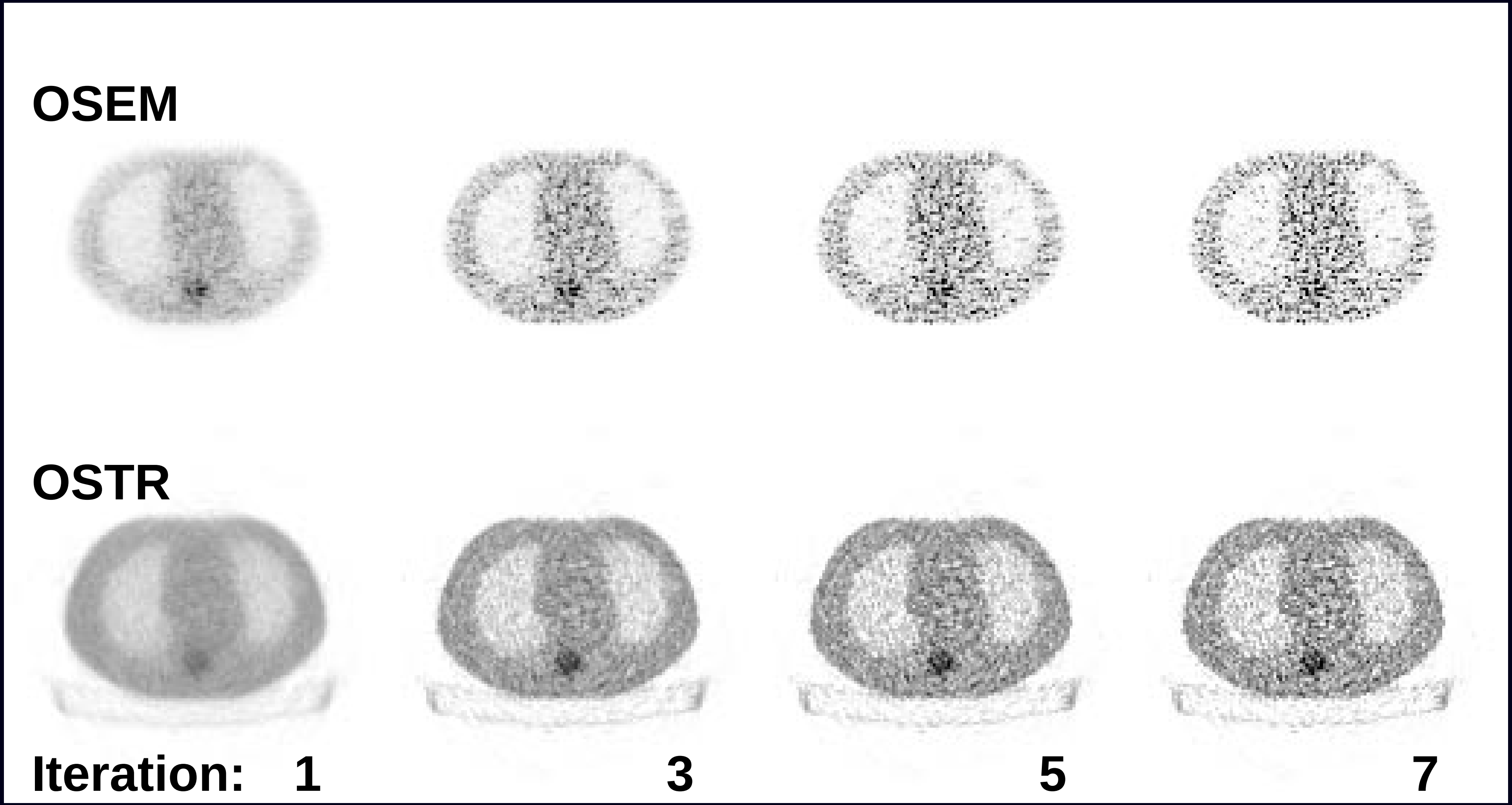
- Shifted Poisson model (for randoms precorrected PET)

$$Y_i = Y_i^{\text{prompt}} - Y_i^{\text{delay}} \sim \text{Poisson}\{\bar{y}_i(\boldsymbol{f}) + 2r_i\} - 2r_i.$$

PET Transmission Phantom



Effect of Statistical Model (PET Transmission Scan)



Choice 4. Objective Functions

Components:

- *Data-fit* term
- *Regularization* term (and regularization parameter β)
- Constraints (*e.g.*, nonnegativity)

$$\Phi(\boldsymbol{f}) = \text{DataFit}(\boldsymbol{Y}, \boldsymbol{A}\boldsymbol{f}) - \beta \cdot \text{Roughness}(\boldsymbol{f})$$

$$\boxed{\hat{\boldsymbol{f}} \overset{\triangle}{=} \arg\max_{\boldsymbol{f} \geq 0} \Phi(\boldsymbol{f})}$$

"Find the image that 'best fits' the measurements"

Actually *three* choices to make for Choice 4 ...

Distinguishes "statistical methods" from "algebraic methods" for "$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{f}$."

Why Objective Functions?

(vs "procedure" *e.g.*, adaptive neural net with wavelet denoising)

Theoretical reasons

ML is based on maximizing an objective function: the log-likelihood

- ML is asymptotically consistent
- ML is asymptotically unbiased
- ML is asymptotically efficient (under true statistical model...)
- Penalized-likelihood achieves uniform CR bound asymptotically

Practical reasons

- Stability of estimates (if Φ and algorithm chosen properly)
- Predictability of properties (despite nonlinearities)
- Empirical evidence (?)

Choice 4.1: Data-Fit Term

- Least squares, weighted least squares (quadratic data-fit terms)
- Reweighted least-squares
- Model-weighted least-squares
- Norms robust to outliers
- Log-likelihood of statistical model. Poisson emission case:

$$L(\boldsymbol{\lambda}; \boldsymbol{Y}) = \log P[\boldsymbol{Y} = \boldsymbol{y}; \boldsymbol{\lambda}] = \sum_{i=1}^{n_d} y_i \log([\boldsymbol{A}\boldsymbol{\lambda}]_i + r_i) - ([\boldsymbol{A}\boldsymbol{\lambda}]_i + r_i) - \log y_i!$$

Poisson probability mass function (PMF):

$$P[\boldsymbol{Y} = \boldsymbol{y}; \boldsymbol{\lambda}] = \prod_{i=1}^{n_d} e^{-\bar{y}_i} \bar{y}_i^{y_i} / y_i! \quad \text{where} \quad \bar{\boldsymbol{y}} \triangleq \boldsymbol{A}\boldsymbol{\lambda} + \boldsymbol{r}$$

Considerations

- Faithfulness to statistical model vs computation.
- Effect of statistical modeling errors.

Choice 4.2: Regularization

Forcing too much “data fit” gives noisy images.
Ill-conditioned problems: small data noise causes large image noise.

Solutions:

- **Noise-reduction methods**
  - Modify the *data* (prefilter or extrapolate sinogram data)
  - Modify an *algorithm* derived for an ill-conditioned problem (stop before converging, post-filter)
- **True regularization methods**

  Redefine the *problem* to eliminate ill-conditioning
  - Use bigger pixels (fewer basis functions)
  - Method of sieves (constrain image roughness)
  - Change objective function by adding a roughness penalty / prior

$$R(\boldsymbol{f}) = \sum_{j=1}^{n_p} \sum_{k \in N_j} \psi(f_j - f_k)$$

Noise-Reduction vs True Regularization

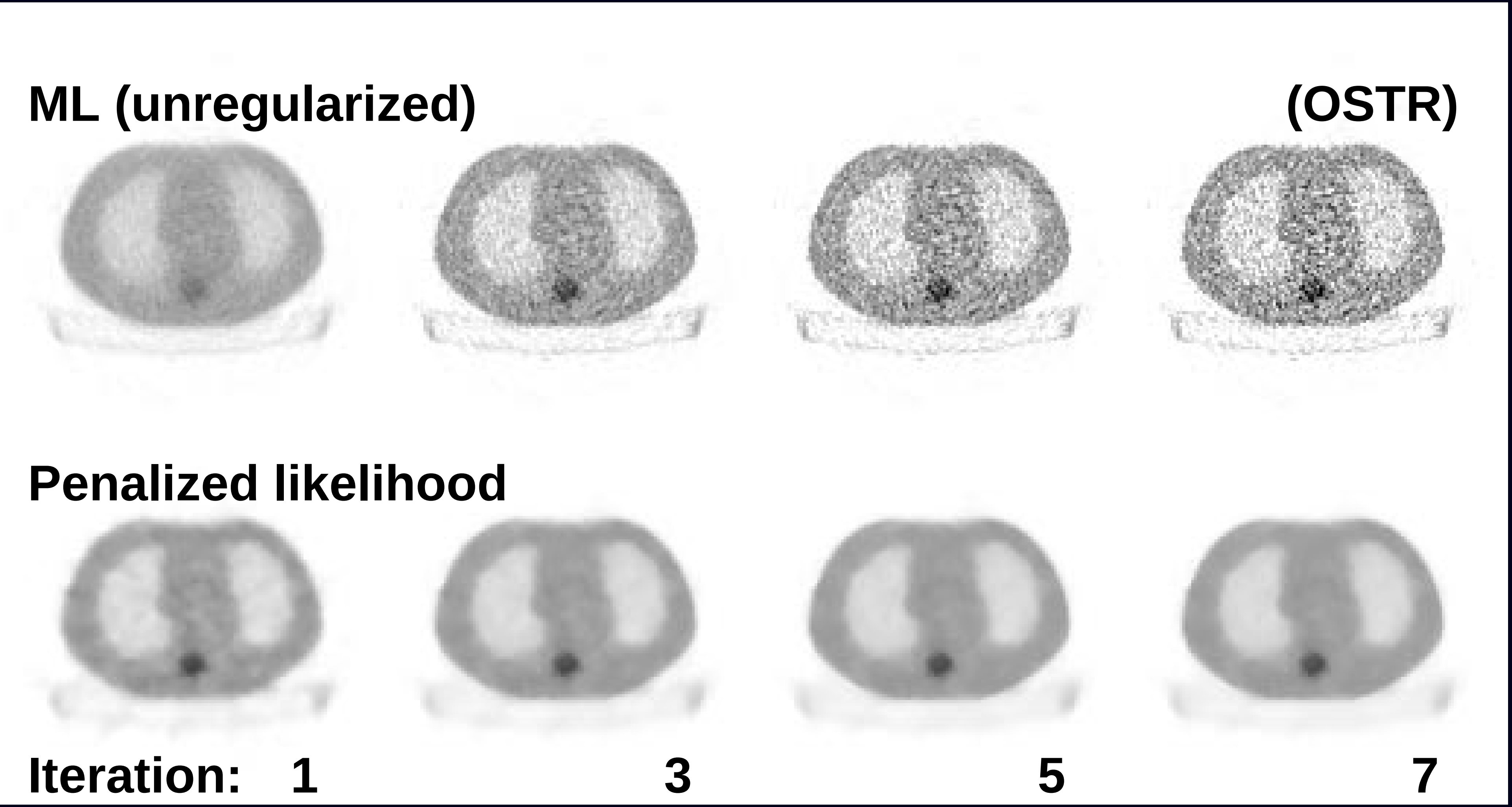
Advantages of **“noise-reduction”** methods

- Simplicity (?)
- Familiarity
- Appear less subjective than using penalty functions or priors
- Only fiddle factors are # of iterations, amount of smoothing
- Resolution/noise tradeoff usually varies with iteration
  (stop when image looks good - in principle)

Advantages of **true regularization** methods

- Stability
- Predictability
- Resolution can be made object independent
- Controlled resolution (*e.g.* spatially uniform, edge preserving)
- Start with (*e.g.*) FBP image $\Rightarrow$ reach solution faster.

Unregularized vs Regularized Reconstruction



Roughness Penalty Function Considerations

$$R(\boldsymbol{f}) = \sum_{j=1}^{n_p} \sum_{k \in N_j} \psi(f_j - f_k)$$

- Computation
- Algorithm complexity
- Uniqueness of maximum of Φ
- Resolution properties (edge preserving?)
- # of adjustable parameters
- Predictability of properties (resolution and noise)

Choices

- separable vs nonseparable
- quadratic vs nonquadratic
- convex vs nonconvex

This topic is actively debated!

Nonseparable Penalty Function Example

Example

f_1	f_2	f_3
f_4	f_5	

$$R(\boldsymbol{f}) = (f_2 - f_1)^2 + (f_3 - f_2)^2 + (f_5 - f_4)^2 + (f_4 - f_1)^2 + (f_5 - f_2)^2$$

2	2	2
2	1	

$R(\boldsymbol{f}) = 2$

3	3	1
2	2	

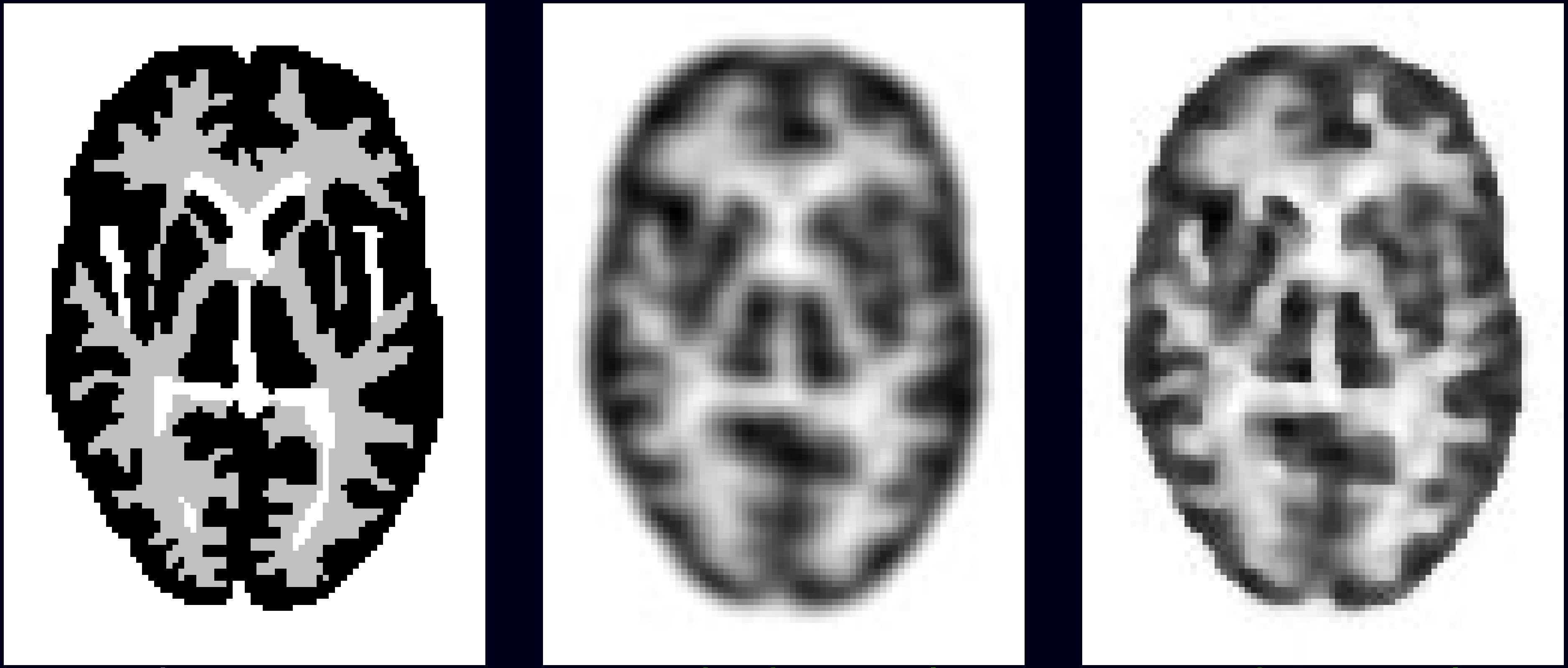
$R(\boldsymbol{f}) = 6$

1	3	1
2	2	

$R(\boldsymbol{f}) = 10$

Rougher images $\Rightarrow$ greater $R(\boldsymbol{f})$

Penalty Functions: Quadratic vs Nonquadratic



Phantom | Quadratic Penalty | Huber Penalty

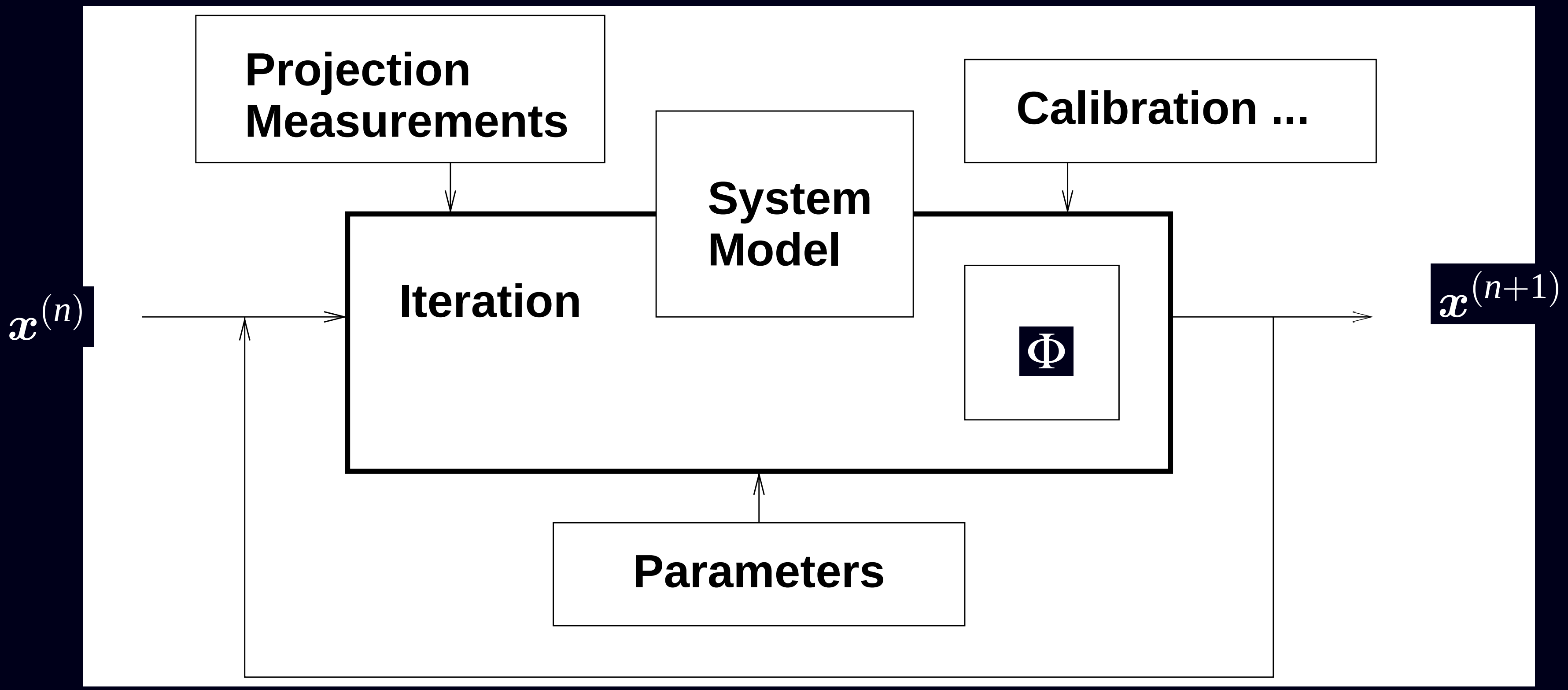
Summary of Modeling Choices

1. Object parameterization: $\lambda(\vec{x})$ or $\mu(\vec{x})$ vs $\boldsymbol{f}$
2. Model of system physics
3. Measurement statistical model $Y_i \sim \boxed{?}$
4. Objective function: data-fit / regularization / constraints

Reconstruction Method = Objective Function + Algorithm

5. Iterative algorithm
   ML-EM, MAP-OSL, PL-SAGE, PWLS+SOR, PWLS-CG, . . .

Choice 5. Algorithms



Deterministic iterative mapping: $x^{(n+1)} = M\left(x^{(n)}\right)$

All algorithms are imperfect. No single best solution.

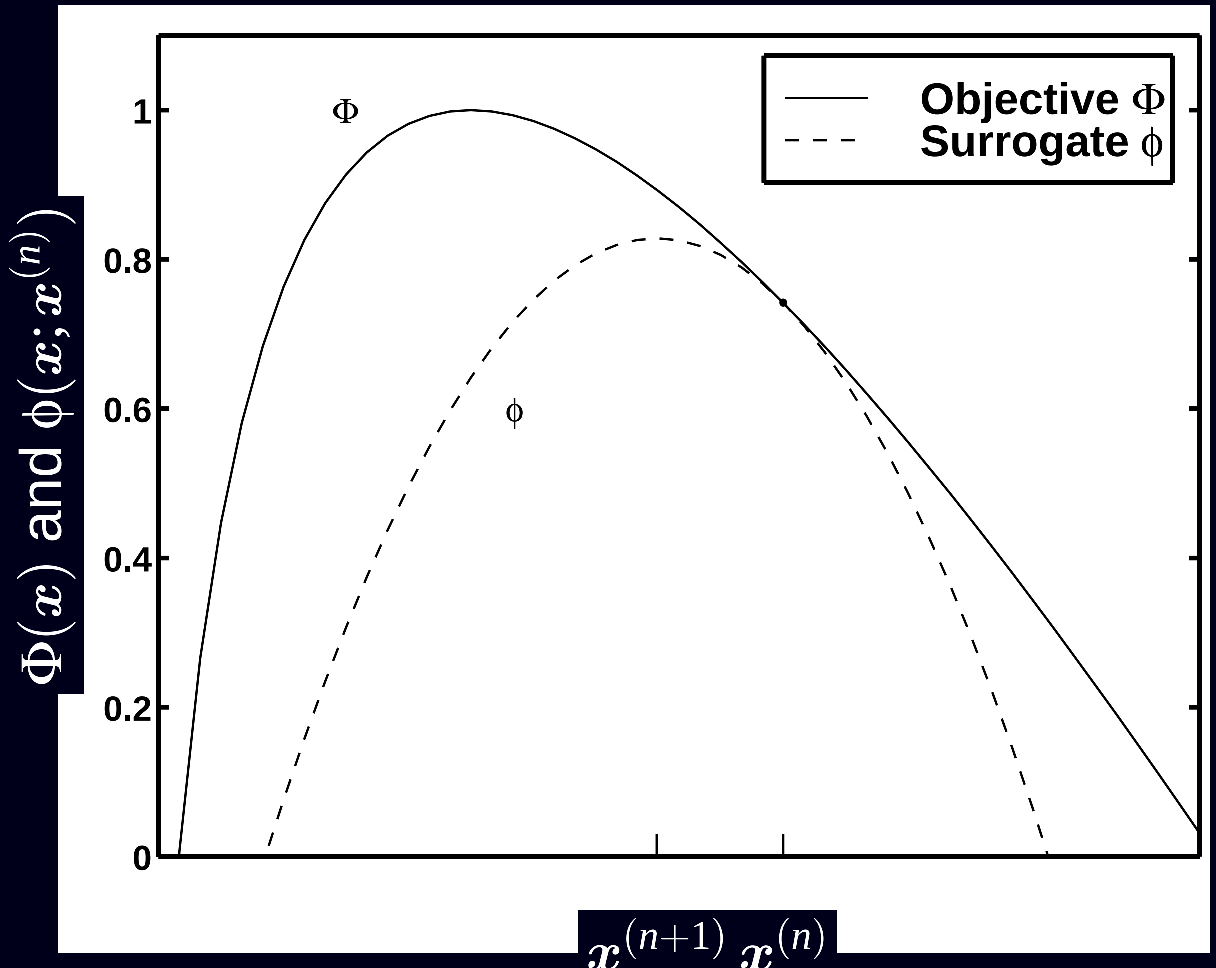
Ideal Algorithm

$$\boldsymbol{x}^\star \triangleq \arg\max_{\boldsymbol{x} \geq \boldsymbol{0}} \Phi(\boldsymbol{x})$$ (global maximum)

stable and convergent	$\{\boldsymbol{x}^{(n)}\}$ converges to $\boldsymbol{x}^\star$ if run indefinitely
converges quickly	$\{\boldsymbol{x}^{(n)}\}$ gets "close" to $\boldsymbol{x}^\star$ in just a few iterations
globally convergent	$\lim_n \boldsymbol{x}^{(n)}$ independent of starting image
fast	requires minimal computation per iteration
robust	insensitive to finite numerical precision
user friendly	nothing to adjust (*e.g.*, acceleration factors)
monotonic	$\Phi(\boldsymbol{x}^{(n)})$ increases every iteration
parallelizable	(when necessary)
simple	easy to program and debug
flexible	accommodates any type of system model (matrix stored by row or column or projector/backprojector)

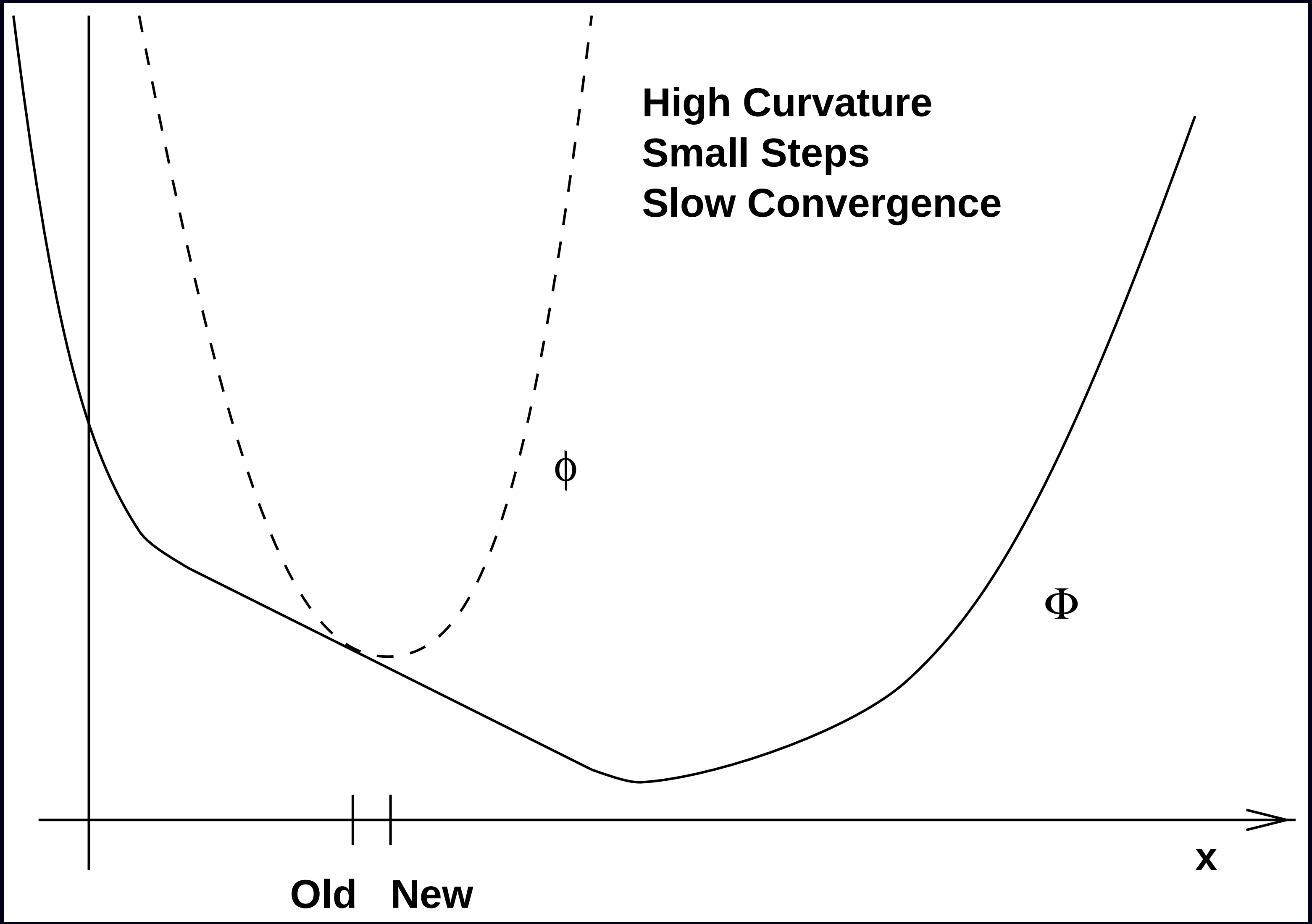
Choices: forgo one or more of the above

Optimization Transfer Illustrated

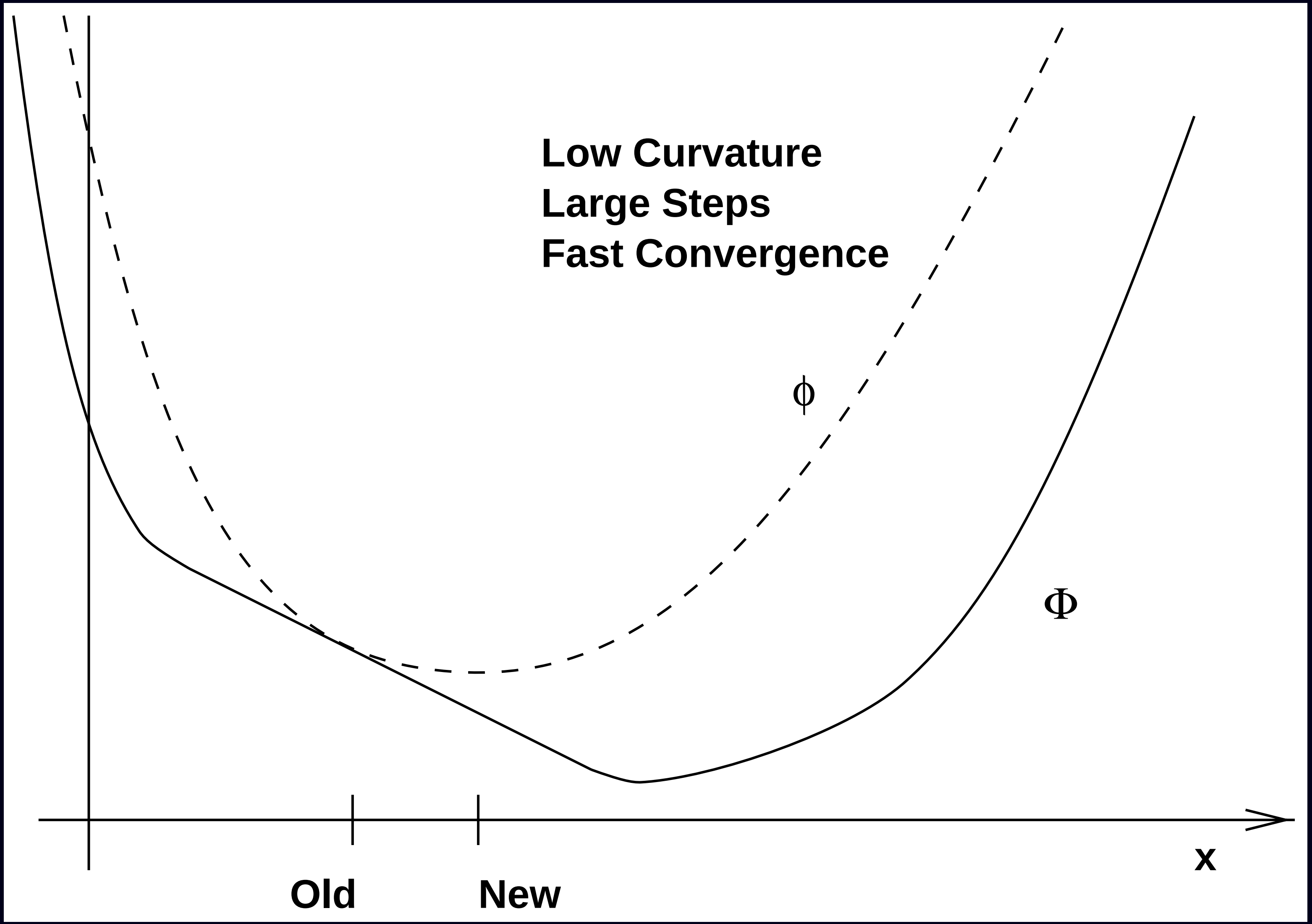


$$x^{(n+1)} = \arg\max_{x \geq 0} \phi(x; x^{(n)})$$

Convergence Rate: Slow



Convergence Rate: Fast



Summary

- General principles of statistical image reconstruction
- Optimization transfer
- Principles apply to both emission and transmission reconstruction.
- Predictability of resolution / noise and controlling spatial resolution argues for regularized objective-function
- Still work to be done...

An Open Problem

Still no algorithm with all of the following properties:

- Nonnegativity easy
- Fast converging
- Intrinsically monotone global convergence
- Accepts any type of system matrix
- Parallelizable

Statistical image reconstruction for polyenergetic X-ray CT

Idris A. Elbakri and Jeffrey A. Fessler

EECS Department
The University of Michigan

(Based on SPIE ’01)

GEMS CT

June 13, 2001

Outline

- Introduction
- Statistical model
- Algorithm
- Ordered-subsets
- Results

Introduction

- Beam-hardening can cause severe artifacts if ignored
- Previous correction methods are only approximation (*e.g.*, Joseph and Spital, JCAT, 1978)
- Previous correction methods are not statistical
- Previous statistical reconstruction methods have ignored beam hardening

Object Model

$$\mu(\vec{x}, E) = \sum_{k=1}^{K} m_k(E)\rho_k(\vec{x})\alpha_k(\vec{x})$$

Arguments:

- $\vec{x}$ spatial position
- E photon energy

Unknown functions:

- μ linear attenuation coefficient
- ρ_k density of kth material

Known quantities:

- K number of materials (*e.g.*, $K = 2$ for bone / soft tissue)
- m_k mass attenuation coefficient of kth material
- α_k fraction of kth material at location $\vec{x}$
  (Currently: 0 or 1 from segmenting JS FBP image)

Statistical model

$$Y_i \sim \text{Poisson}\left\{ \int I_0(E) \exp\left(-\sum_{k=1}^{K} m_k(E) \sum_{j=1}^{n_p} a_{ij}^k \rho_j \right) dE + r_i \right\}$$

"Known" quantities:

- Y_i ith element of measured sinogram
- a_{ij}^k system matrix $a_{ij}^k = a_{ij} \alpha_k(\vec{x}_j)$
- I_0 source spectrum and detector sensitivity
- r_i scatter and/or detector readout bias

Unknown quantities:

- ρ_j density of jth voxel

Goal: Reconstruct density vector $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{n_p})$
from measurement vector $\boldsymbol{Y} = (Y_1, \ldots, Y_{n_d})$.

Approach

- Penalized-likelihood objective function
- Edge-preserving regularization function
- Optimization transfer to derive "scaled gradient ascent" algorithm

$$\rho_j^{(n+1)} = \left[\rho_j^{(n)} - \frac{1}{d_j} \frac{d}{d\rho_j} \Phi(\boldsymbol{\rho}) \Big|_{\boldsymbol{\rho}=\boldsymbol{\rho}^{(n)}} \right]_+ .$$

- The $[\cdot]_+$ enforces nonnegativity constraint.
- For suitable step sizes $\{d_j\}$, algorithm monotonically increases Φ.
- Table lookup to compute

$$F_k(t_{\text{water}}, t_{\text{bone}}) = \int I_0(E) m_k(E) e^{-[m_{\text{water}}(E) t_{\text{water}} + m_{\text{bone}}(E) t_{\text{bone}}]} \, dE .$$

Ordered-Subsets Acceleration

The gradient involves forward and backprojections:

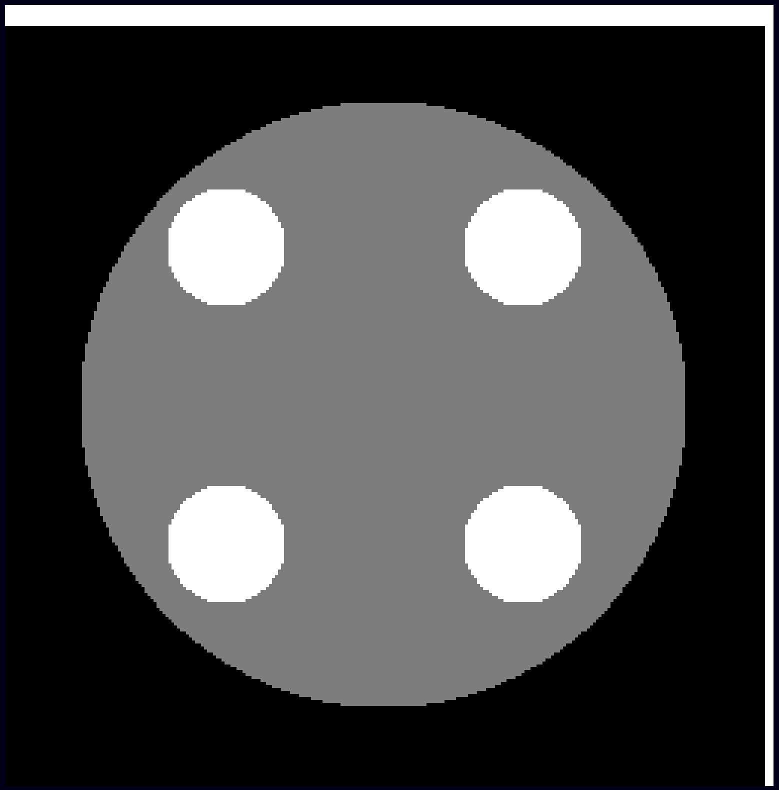
$$\frac{d}{d\rho_j}\Phi(\boldsymbol{\rho}) = \sum_{i=1}^{n_d} a_{ij}\left(1 - Y_i/\bar{y}_i^{(n)}\right)\sum_{k=1}^{K}\alpha_j^k F_k([\boldsymbol{A}_1\boldsymbol{\rho}^{(n)}]_i, [\boldsymbol{A}_2\boldsymbol{\rho}^{(n)}]_i) + \cdots$$

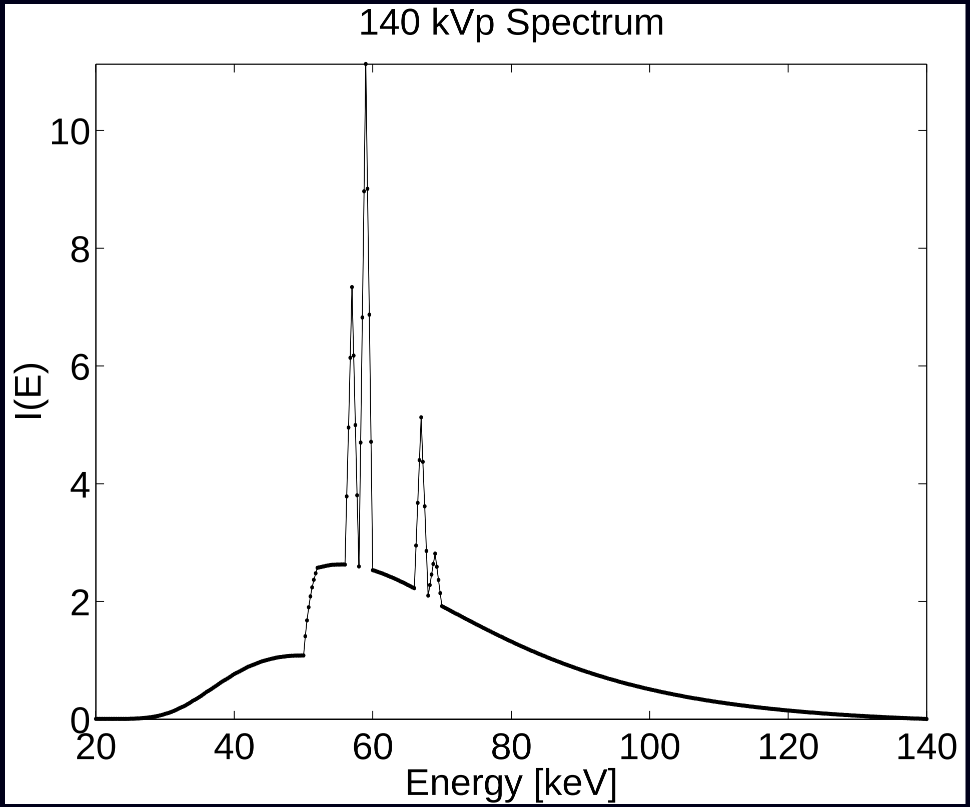
Ordered-subsets concept (Hudson and Larkin, TMI, 1994):

- Replace full backprojection with “downsampled” backprojection (cf FBP with angular downsampling)
- Only forward project ($[\boldsymbol{A}\boldsymbol{\rho}]_i$) for the needed projection views of each subset
- Cycle through all subsets of projection angles
- Accelerates “convergence” $\approx$ by number of subsets in early iterations
- Original formulation does not converge (GE, Siemens, etc., all sell it for PET and SPECT nevertheless...)
- Ahn and Fessler have developed truly convergent formulation

Disk Phantom CT Simulation

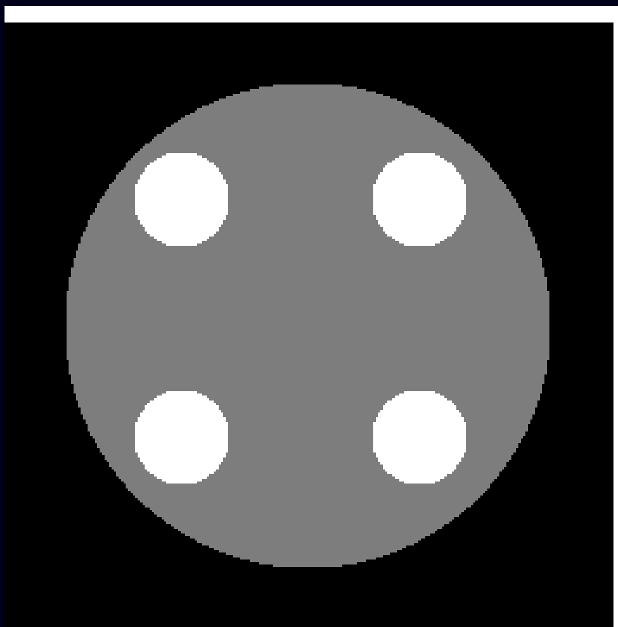
Phantom



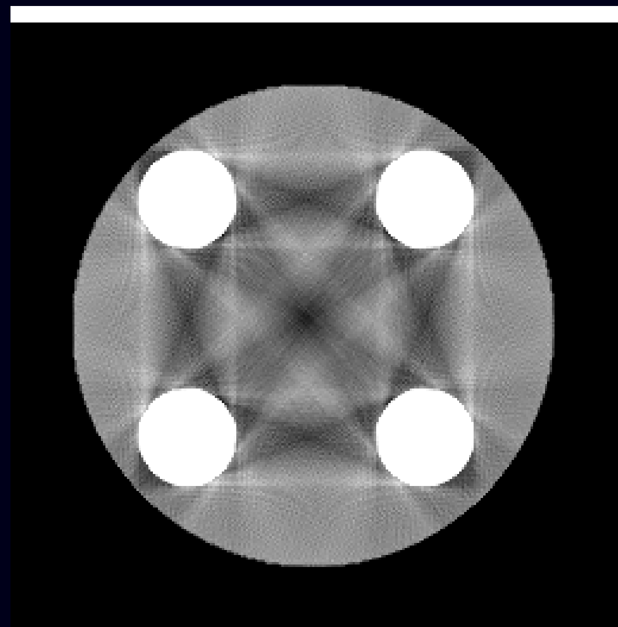


256×256 image, 1.6mm pixels,, water = 1.0g/cm^3, bone = 2.0g/cm^3.
600 rays by 500 angles, 1.3mm spacing, $2.2 \cdot 10^7$ incident photons/ray.
Images windowed from 0.8 to 1.2 g/cm^3.

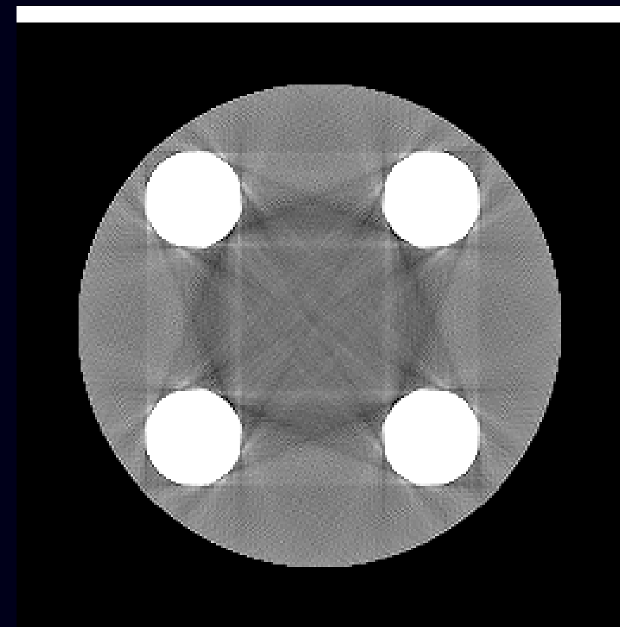
Disk Phantom Results: Conventional



Phantom

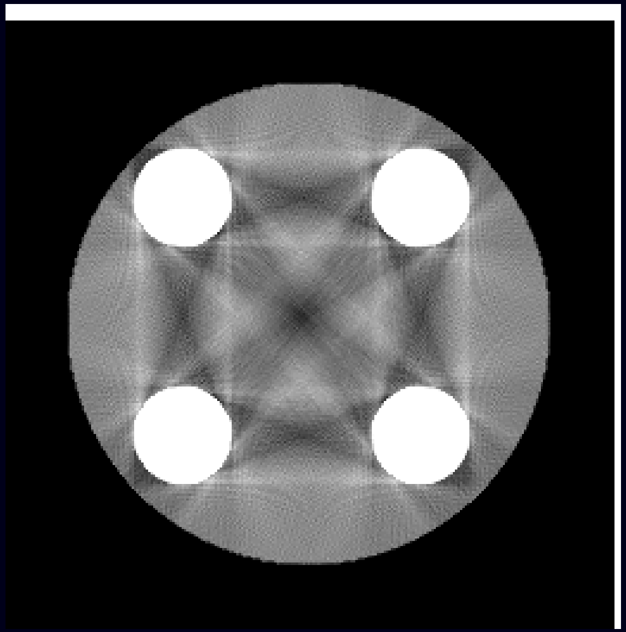


FBP
Uncorrected

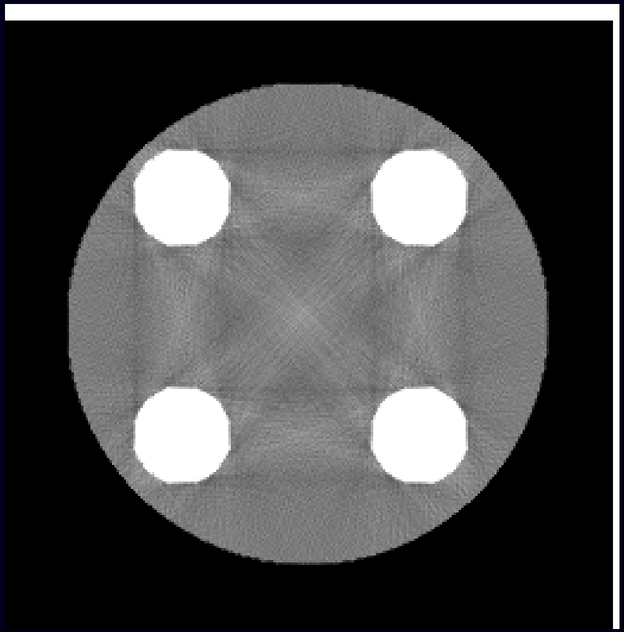


Iterative
Monoenergetic Model

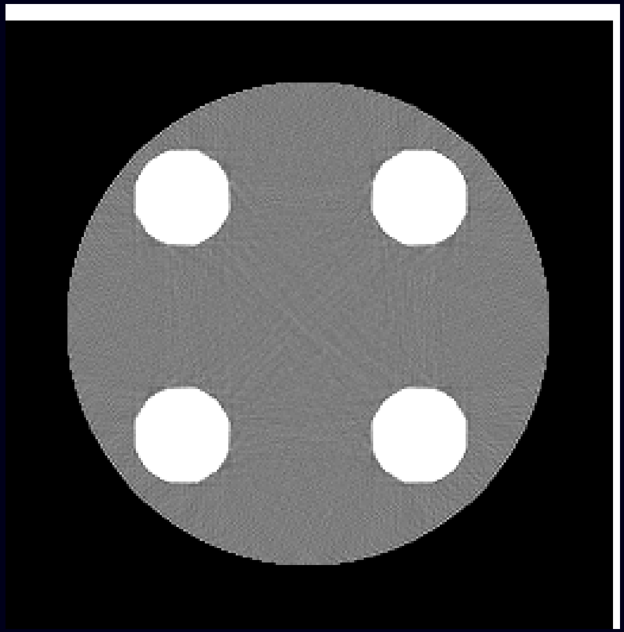
Disk Phantom Results: Corrected



FBP
Soft Tissue Correction



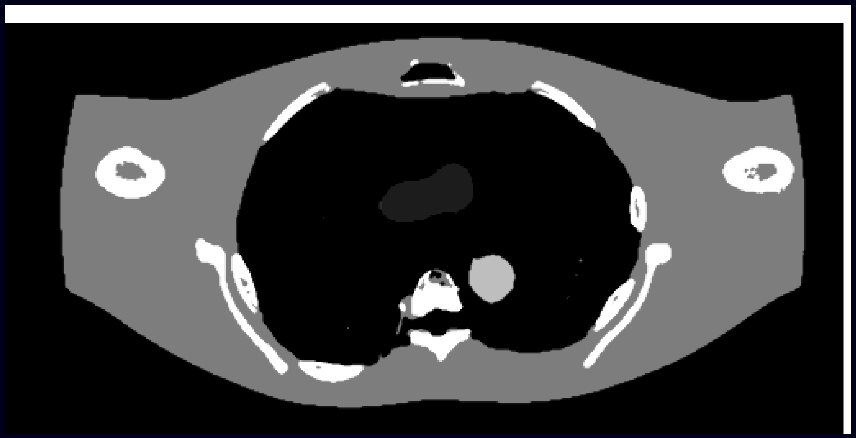
FBP
Joseph and Spital

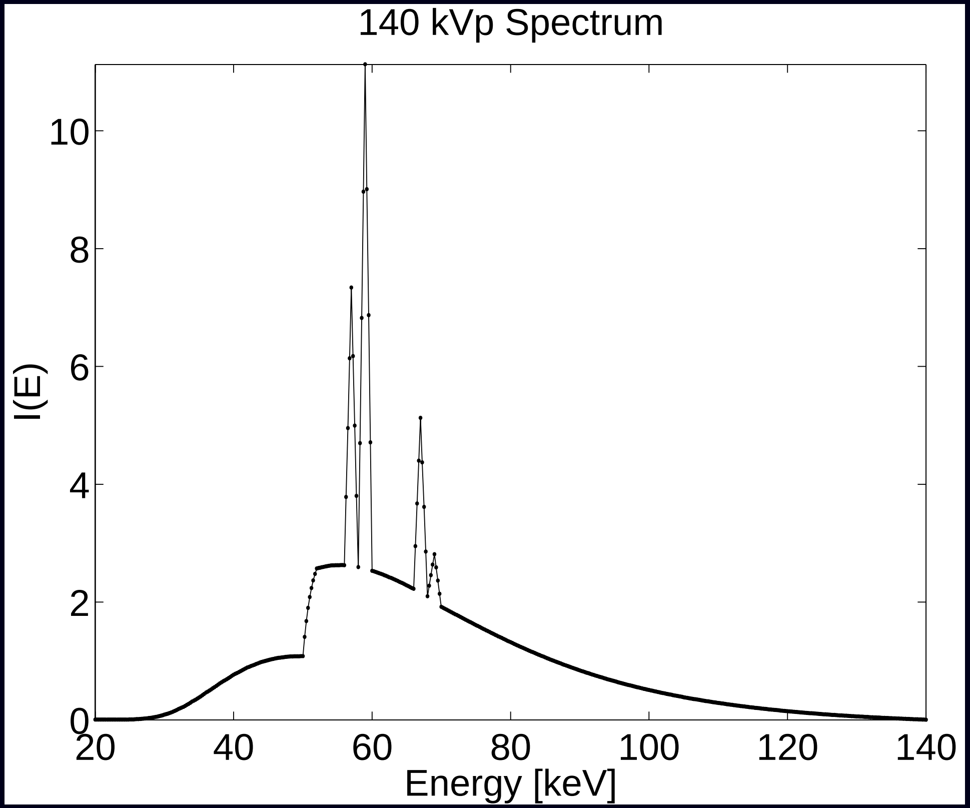


Statistical
Polyenergetic Model
20 iterations, 20 subsets

Chest Phantom CT Simulation

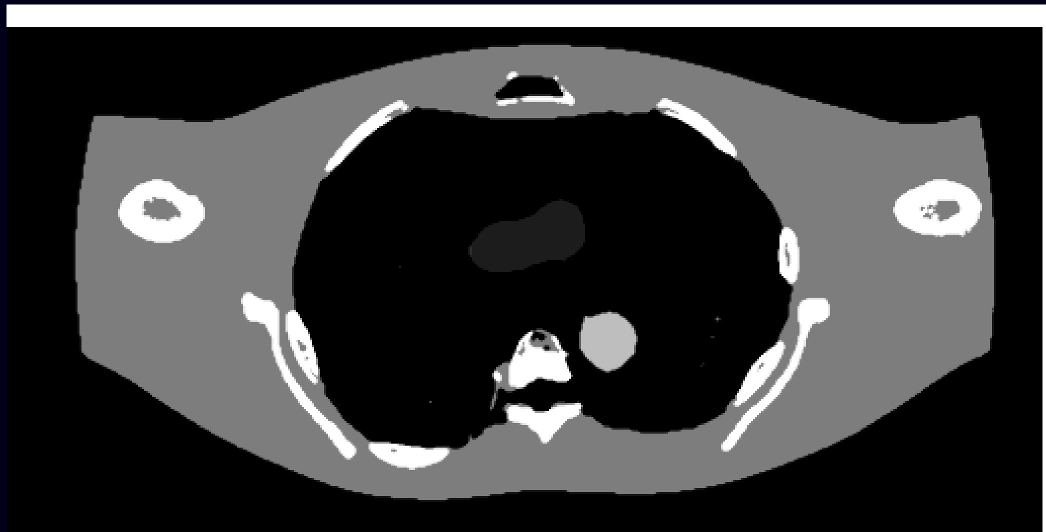
Phantom





512×256 image, 0.8mm pixels.
700 rays by 600 angles, 1.3mm spacing, $1.77 \cdot 10^5$ incident photons/ray.
Images windowed from 0.8 to 1.2 g/cm^3.

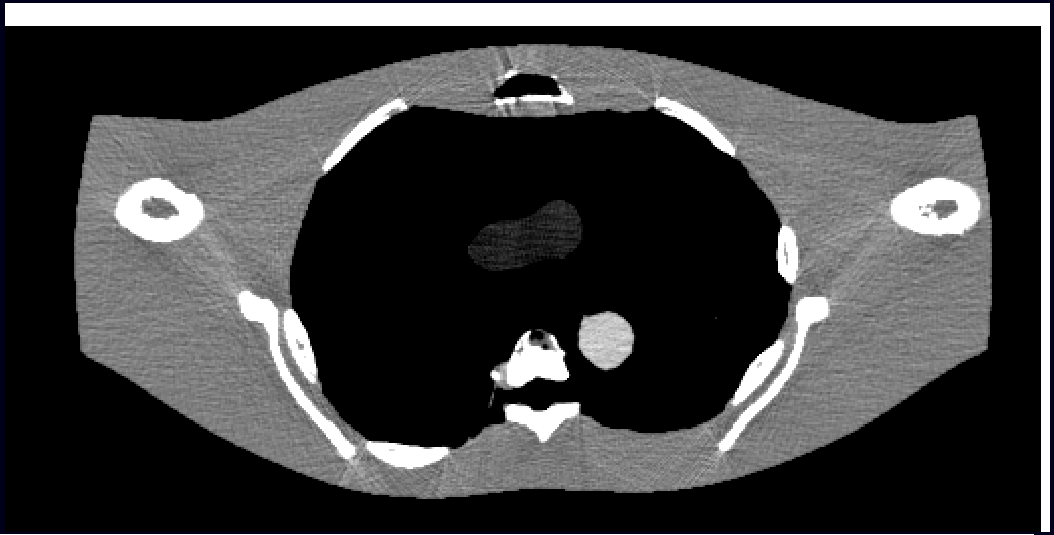
Chest Phantom Results 1



Phantom

FBP
Soft tissue correction

Chest Phantom Results 2



Joseph and Spital

Statistical
Polyenergetic model
10 iterations, 40 subsets

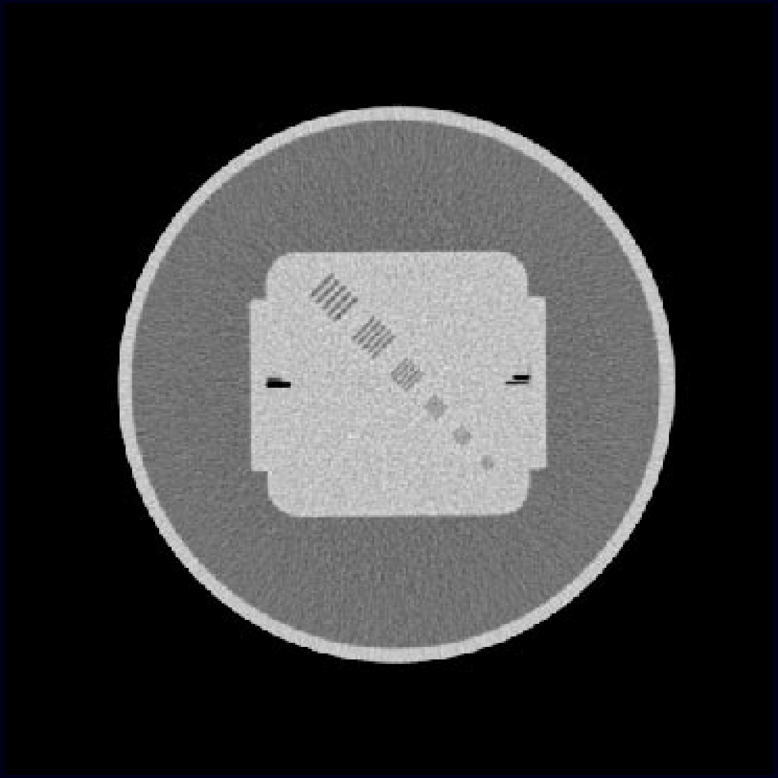
Summary (Beam Hardening)

- First statistical reconstruction method with full polyenergetic model.
- Beam-hardening artifacts nearly eliminated.
- Ordered-subsets approach helps contain computation time.

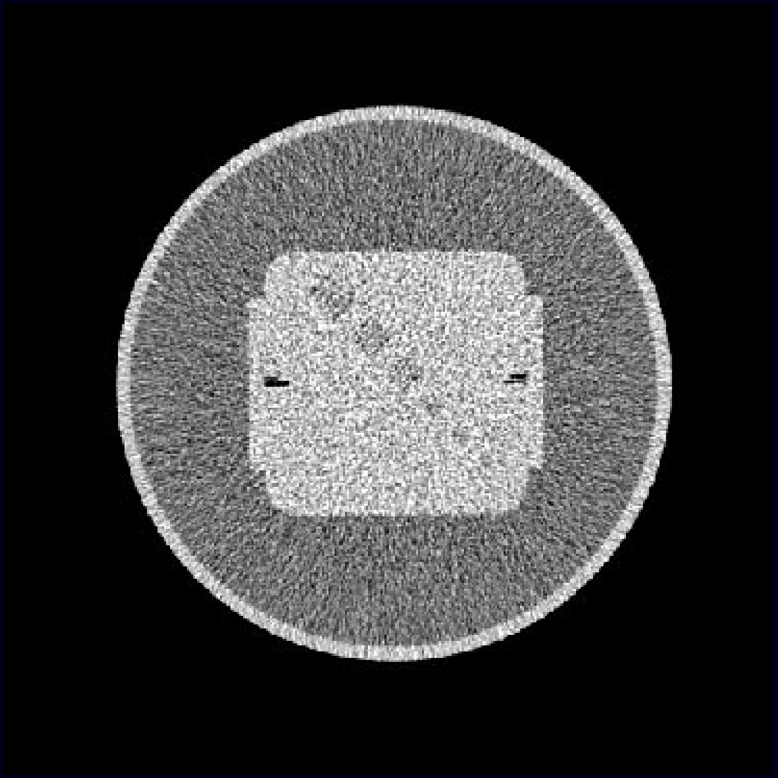
Future Work

- Segmentation of FBP/JS image is adequate starting place.
  Explore re-segmenting statistical reconstruction.
- Ideal approach: simultaneous (iterative) segmentation / reconstruction.
- Extension to contrast agent case $(K = 3)$ conceptually straightforward.
- Extension to metal implants conceptually straightforward.

GEMS Data: Preliminary Results



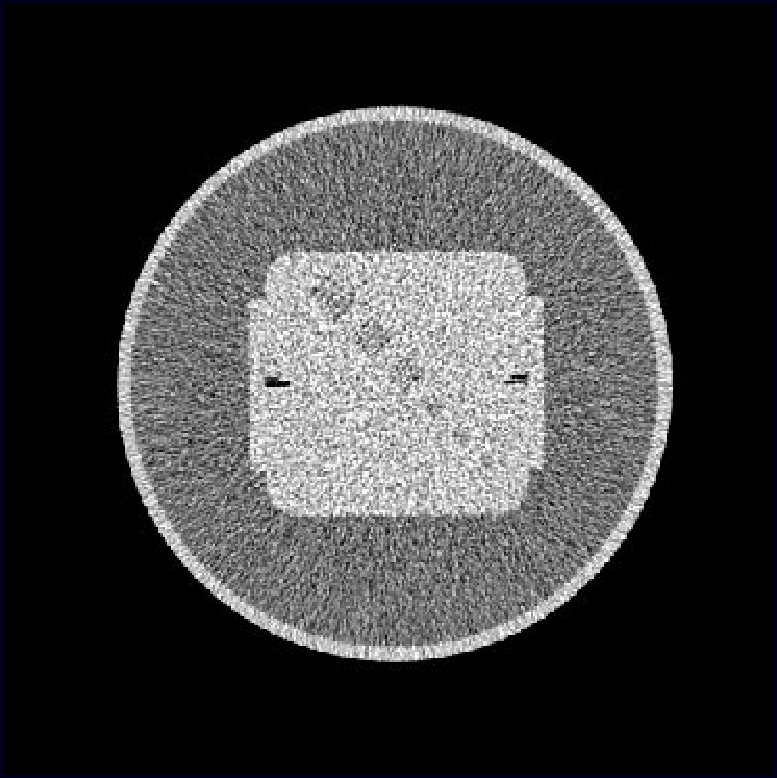
GE FBP, 190mAs



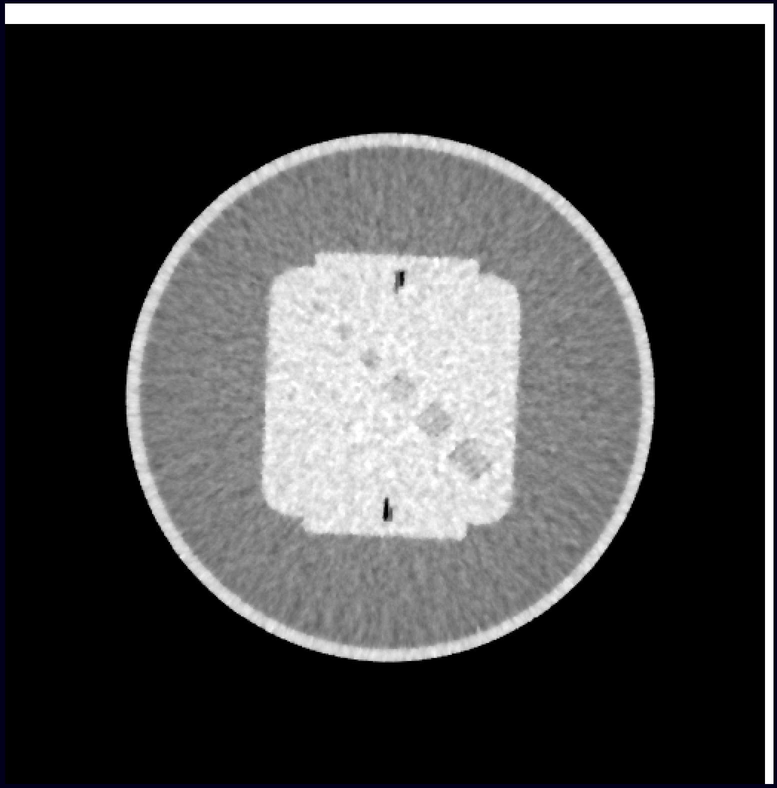
GE FBP, 10mAs

888 channels by 984 views over 360°. bin spacing is 1.0239mm.
source-to-iso = 541mm, iso-to-detector = 408.075mm.
Sinograms precorrected for *everything*.

GE 10mAs Preliminary Results



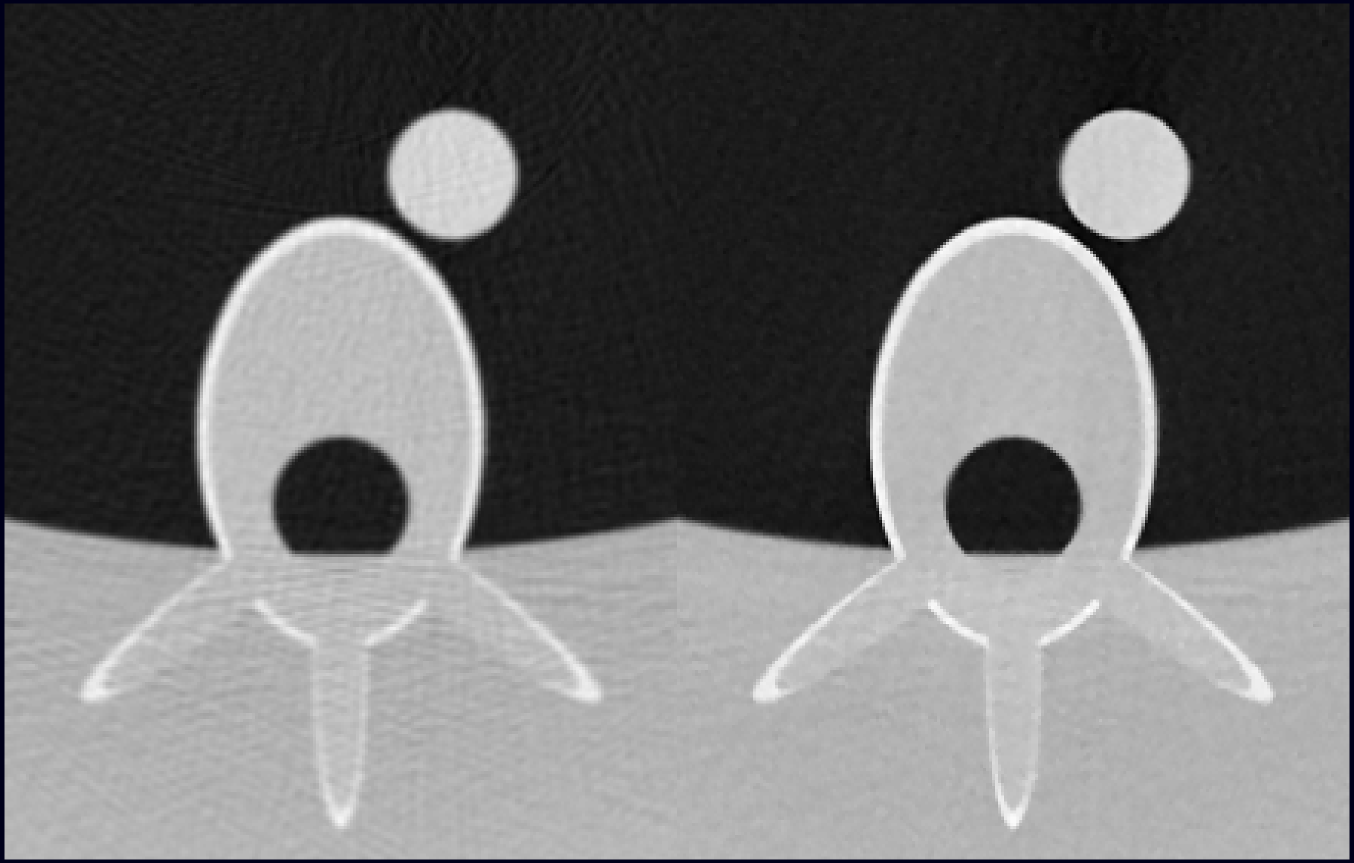
GE FBP



“Iterative”

$\log_2 \beta = 10$, $\delta = 10$. 5 iterations, 41 subsets.
Unweighted regularized least squares.

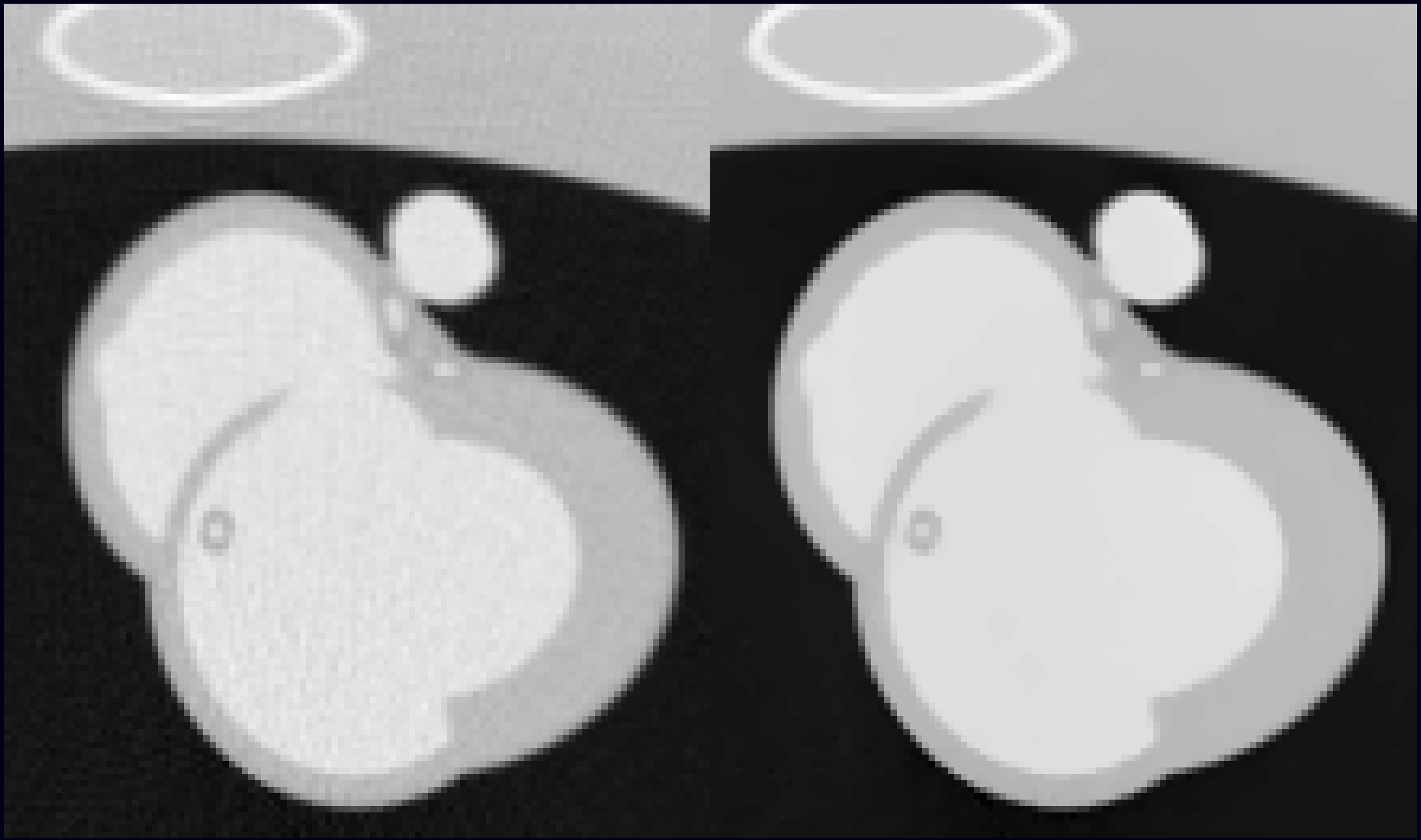
GE CRD Simulation: Spine



GE FBP　　　　　　　　　　　　Iterative

1000^2 flat panel fan-beam monoenergetic (?) projections

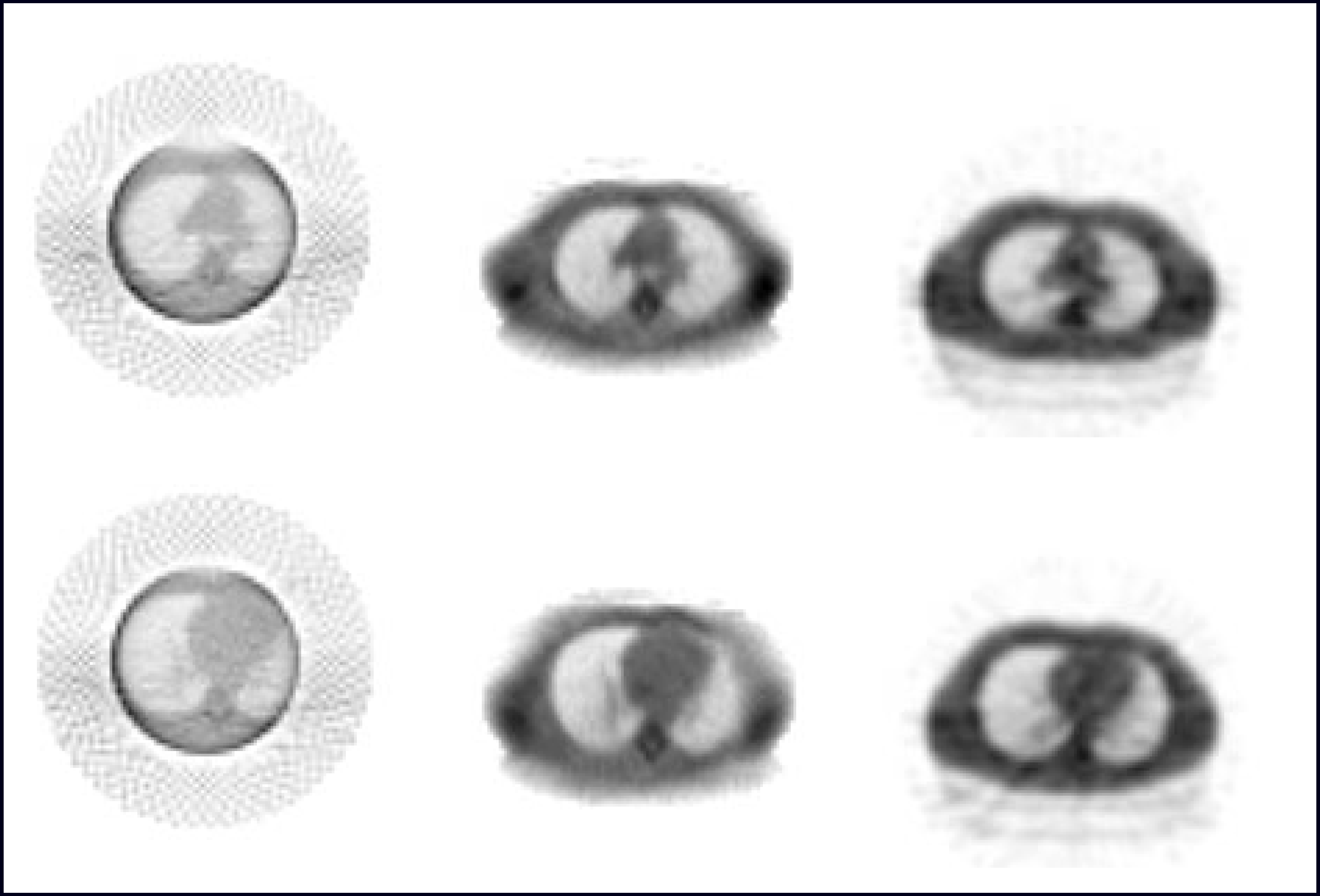
GE CRD Simulation: Cardiac



GE FBP

Iterative

Truncated Fan-Beam SPECT Transmission



Truncated FBP

Truncated PWLS

Untruncated FBP

Overall Summary

Physics

- Modeling source spectrum reduces beam hardening effects

Statistics

- Reduced noise using statistical methods

Geometry

- Truncated fan-beam data

Prior knowledge

- Nonnegativity
- Mass attenuation of water and bone

Future Work

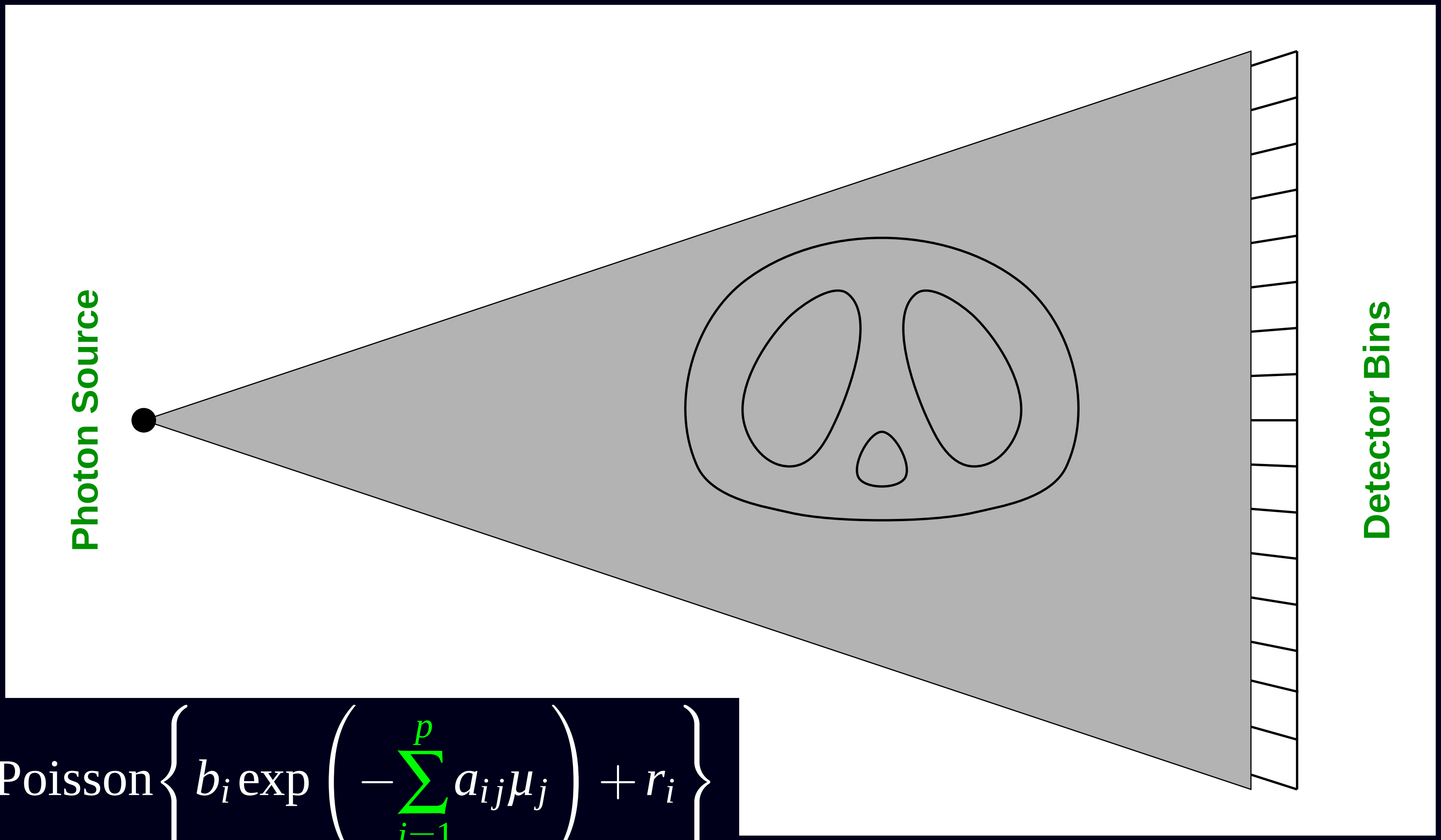
- Real X-ray CT data with appropriate physics and statistics!
- Refinements of beam hardening algorithm
- Helical and cone-beam geometries
- Gated cardiac scans?
- Compton scatter (for larger cone angles)?

Fast Maximum Likelihood Transmission Reconstruction using Ordered Subsets

Jeffrey A. Fessler, Hakan Erdoğan

EECS Department, BME Department, and
Nuclear Medicine Division of Dept. of Internal Medicine
The University of Michigan

Transmission Scans



$$Y_i \sim \text{Poisson}\left\{ b_i \exp\left(-\sum_{j=1}^{p} a_{ij}\mu_j \right) + r_i \right\}$$

Each measurement Y_i is related to a single "line integral" through the object.

Transmission Scan Statistical Model

$$Y_i \sim \text{Poisson}\left\{ b_i \exp\left(-\sum_{j=1}^{p} a_{ij}\mu_j \right) + r_i \right\}, \quad i = 1, \ldots, N$$

- N number of detector elements
- Y_i recorded counts by ith detector element
- b_i blank scan value for ith detector element
- a_{ij} length of intersection of ith ray with jth pixel
- μ_j linear attenuation coefficient of jth pixel
- r_i contribution of room background, scatter, and emission crosstalk

(Monoenergetic case, can be generalized for dual-energy CT)
(Can be generalized for additive Gaussian detector noise)

Maximum-Likelihood Reconstruction

$$\hat{\mu} = \arg\max_{\mu \geq 0} L(\mu) \quad \text{(Log-likelihood)}$$

$$L(\mu) = \sum_{i=1}^{N} Y_i \log\left[b_i \exp\left(-\sum_{j=1}^{p} a_{ij}\mu_j\right) + r_i\right] - \left[b_i \exp\left(-\sum_{j=1}^{p} a_{ij}\mu_j\right) + r_i\right]$$

Transmission ML Reconstruction Algorithms

- Conjugate gradient

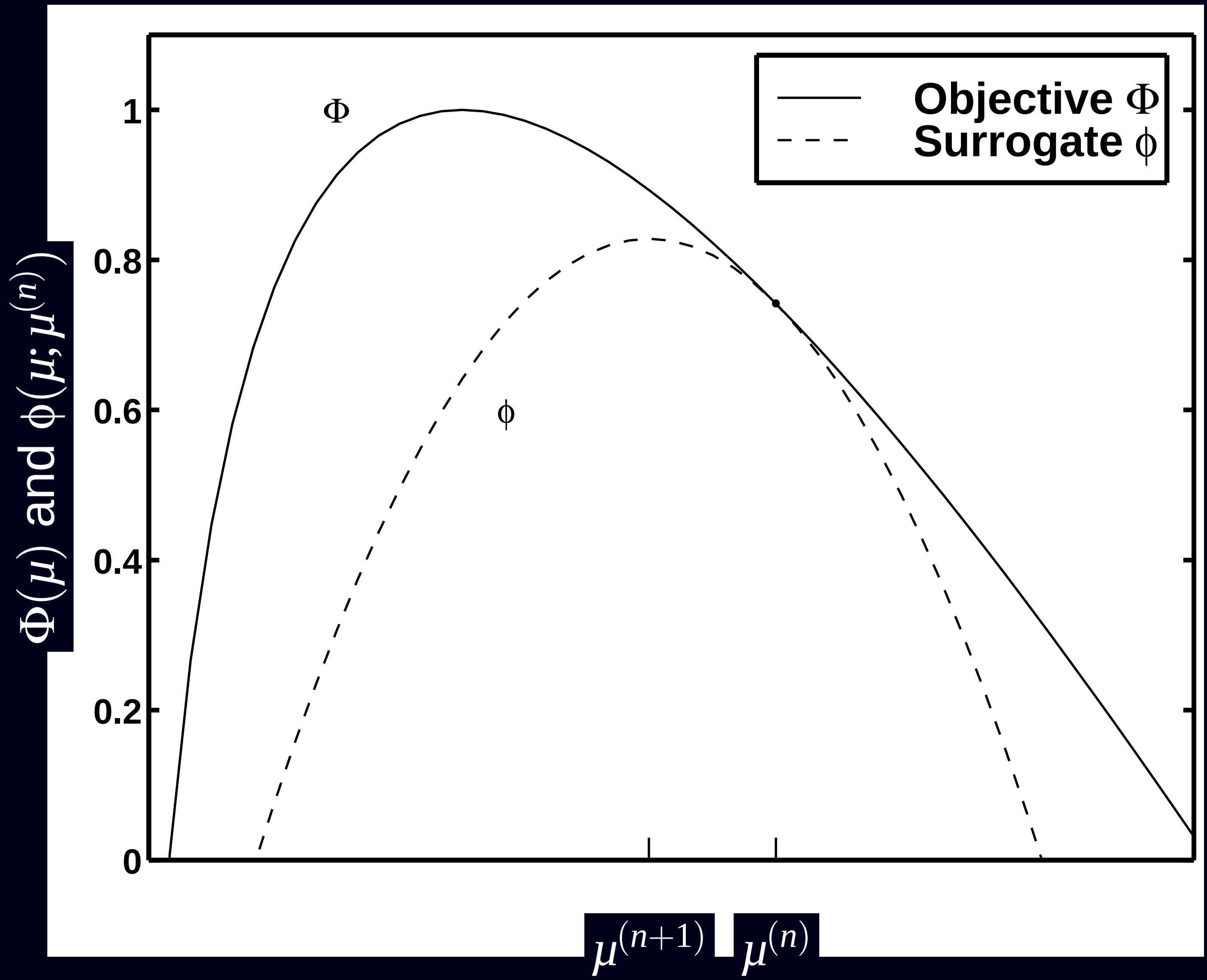
  Mumcuoğlu *et al.*, T-MI, Dec. 1994
- Paraboloidal surrogates coordinate ascent (PSCA)

  Erdoğan and Fessler, T-MI, 1999
- Ordered subsets separable paraboloidal surrogates

  Erdoğan *et al.*, PMB, Nov. 1999
- ~~Transmission expectation maximization (EM) algorithm~~

  Lange and Carson, JCAT, Apr. 1984

Optimization Transfer Illustrated



Parabola Surrogate Function

- $h(l) = y\log(be^{-l} + r) - (be^{-l} + r)$ has a parabola surrogate: $q_{im}^{(n)}$
- Optimum curvature of parabola derived by Erdoğan (T-MI, 1999)
- Replace likelihood with paraboloidal surrogate

$$L(\mu^{(n)}) = \sum_{i=1}^{N} h_i\left(\sum_{j=1}^{p} a_{ij}\mu_j\right) \geq Q_1(\mu;\mu^{(n)}) = \sum_{i=1}^{N} q_{im}^{(n)}\left(\sum_{j=1}^{p} a_{ij}\mu_j\right)$$

- $q_{im}^{(n)}$ is a simple quadratic function
- Iterative algorithm:

$$\boxed{\mu^{(n+1)} = \arg\max_{\mu \geq 0} Q_1(\mu;\mu^{(n)})}$$

- Maximizing $Q_1(\mu;\mu^{(n)})$ over μ is equivalent to (reweighted) least-squares.
- Natural algorithms
  - Conjugate gradient
  - Coordinate ascent

Separable Paraboloid Surrogate Function

- Parabolas are convex functions
- Apply De Pierro's "additive" convexity trick (T-MI, Mar. 1995)

$$\sum_{j=1}^{p} a_{ij}\mu_j = \sum_{j=1}^{p} \frac{a_{ij}}{a_i}\left[a_i(\mu_j - \mu_j^{(n)})\right] + \left[\boldsymbol{A}\mu^{(n)}\right]_i \quad \text{where} \quad a_i \triangleq \sum_{j=1}^{p} a_{ij}$$

- Move summation over pixels outside quadratic

$$Q_1(\mu;\mu^{(n)}) = \sum_{i=1}^{N} q_{im}^{(n)}\left(\sum_{j=1}^{p} a_{ij}\mu_j\right)$$

$$\geq \; Q_2(\mu;\mu^{(n)}) = \sum_{i=1}^{N}\sum_{j=1}^{p}\frac{a_{ij}}{a_i} q_{im}^{(n)}\left(a_i(\mu_j - \mu_j^{(n)}) + \left[\boldsymbol{A}\mu^{(n)}\right]_i\right)$$

$$= \sum_{j=1}^{p} Q_{2j}^{(n)}(\mu_j), \quad \text{where} \quad Q_{2j}^{(n)}(x) \triangleq \sum_{i=1}^{N}\frac{a_{ij}}{a_i} q_{im}^{(n)}\left(a_i(x - \mu_j^{(n)}) + \left[\boldsymbol{A}\mu^{(n)}\right]_i\right)$$

- Separable paraboloidal surrogate function $\Rightarrow$ trivial to maximize (cf EM)

Iterative algorithm:

$$\mu_j^{(n+1)} = \arg\max_{\mu_j \geq 0} Q_{2j}^{(n)}(\mu_j) = \left[\mu_j^{(n)} + \frac{\frac{\partial}{\partial \mu_j} Q_{2j}^{(n)}(\mu^{(n)})}{-\frac{\partial^2}{\partial \mu_j^2} Q_{2j}^{(n)}(\mu^{(n)})} \right]_+$$

$$= \left[\mu_j^{(n)} + \frac{1}{-\frac{\partial^2}{\partial \mu_j^2} Q_{2j}^{(n)}(\mu^{(n)})} \frac{\partial}{\partial \mu_j} L(\mu^{(n)}) \right]_+$$

$$= \left[\mu_j^{(n)} + \frac{\sum_{i=1}^N (y_i / \bar{y}_i^{(n)} - 1) b_i \exp\left(-\left[\boldsymbol{A}\mu^{(n)}\right]_i\right)}{\sum_{i=1}^N a_{ij}^2 a_i c_i^{(n)}} \right]_+, \quad j = 1, \ldots, p$$

- $c_i^{(n)}$'s related to parabola curvatures
- Parallelizable (ideal for multiprocessor workstations)
- Monotonically increases the likelihood each iteration
- Intrinsically enforces the nonnegativity constraint
- Guaranteed to converge if unique maximizer
- Natural starting point for forming ordered-subsets variation

Ordered Subsets Algorithm

- Each $\sum_{i=1}^{N}$ is a backprojection
- Replace “full” backprojections with partial backprojections
- Partial backprojection based on angular subsampling
- Cycle through subsets of projection angles

Pros

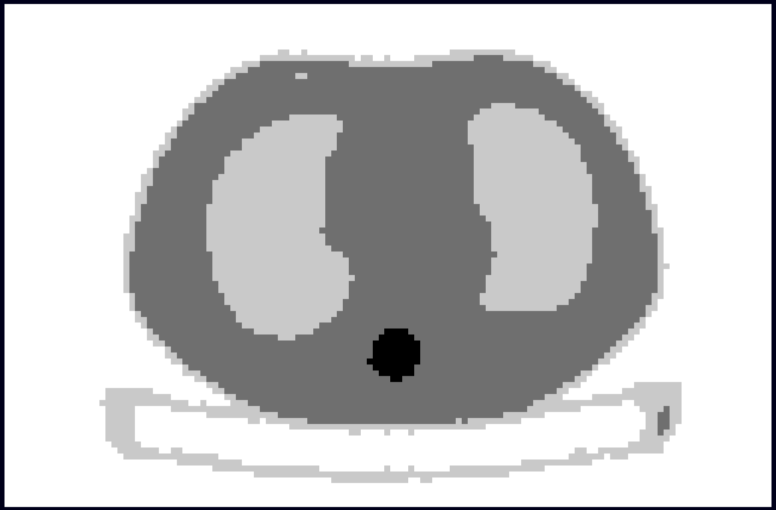
- Accelerates “convergence”
- Very simple to implement
- Reasonable images in just 1 or 2 iterations
- Regularization easily incorporated

Cons:

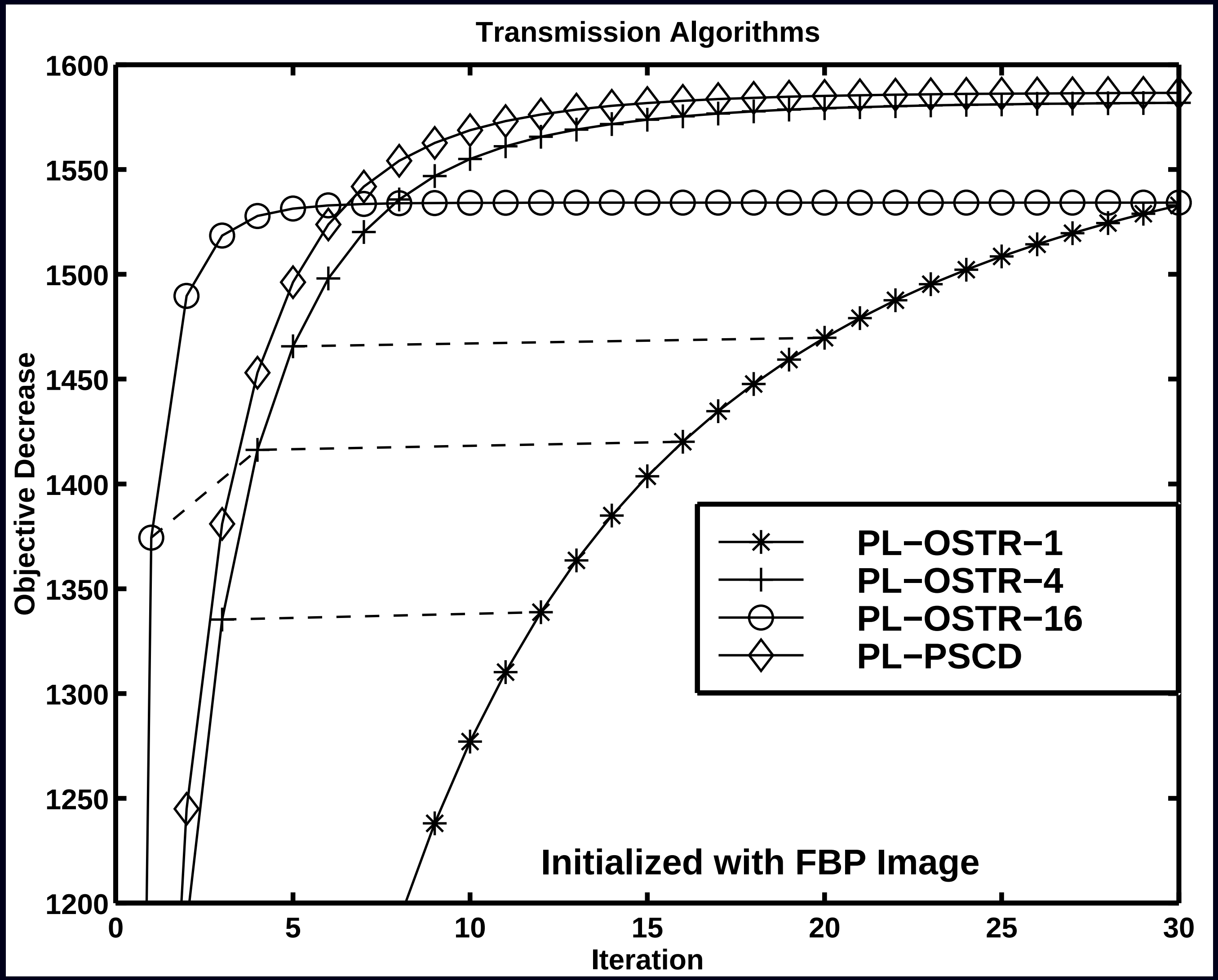
- Does not converge to true maximizer
- Makes analysis of properties difficult

Phantom Study

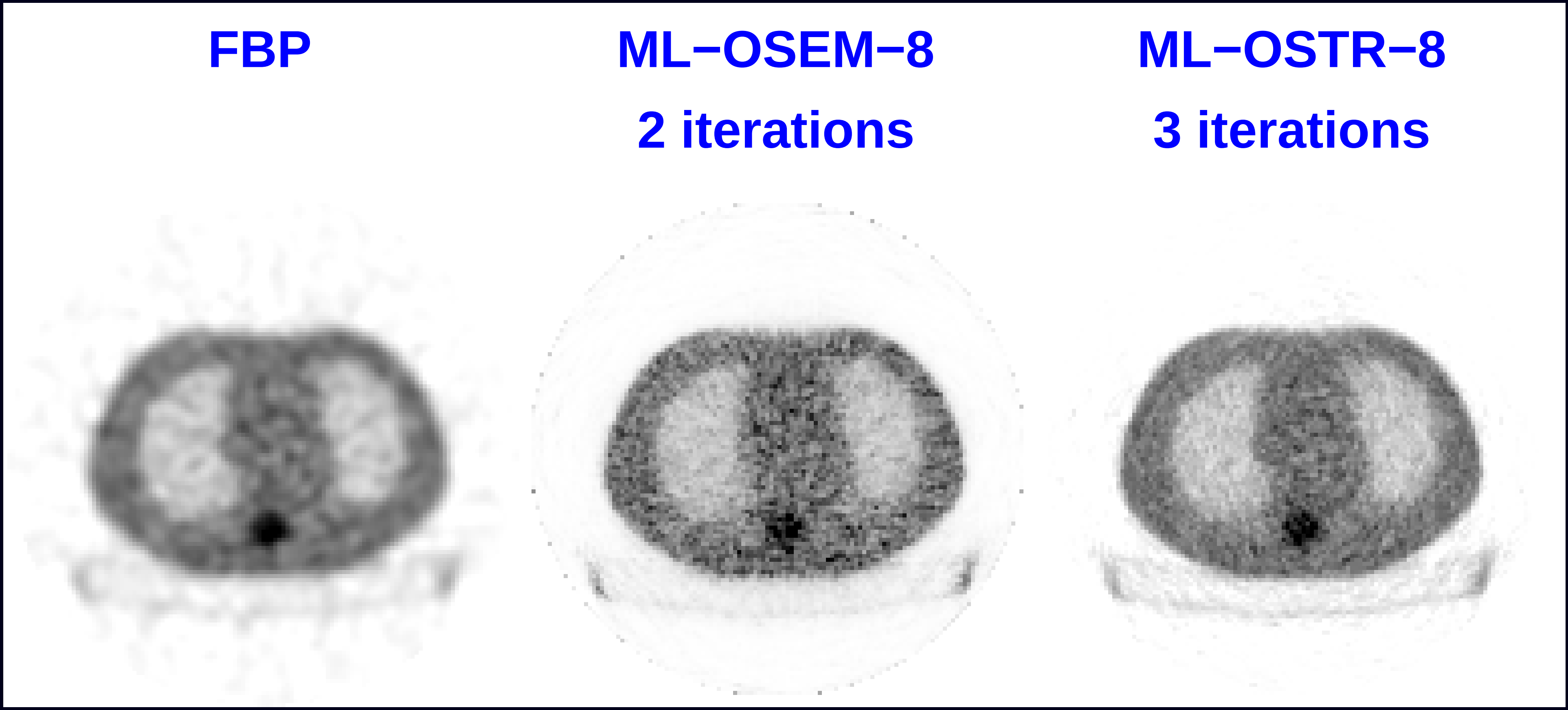
- 12-minute PET transmission scan
- Anthropomorphic thorax phantom (Data Spectrum, Chapel Hill, NC)
- Sinogram: 160 3.375mm bins by 192 angles over 180°
- Image: 128 by 128 4.2mm pixels
- Ground truth determined from 15-hour scan, FBP reconstruction / segmentation



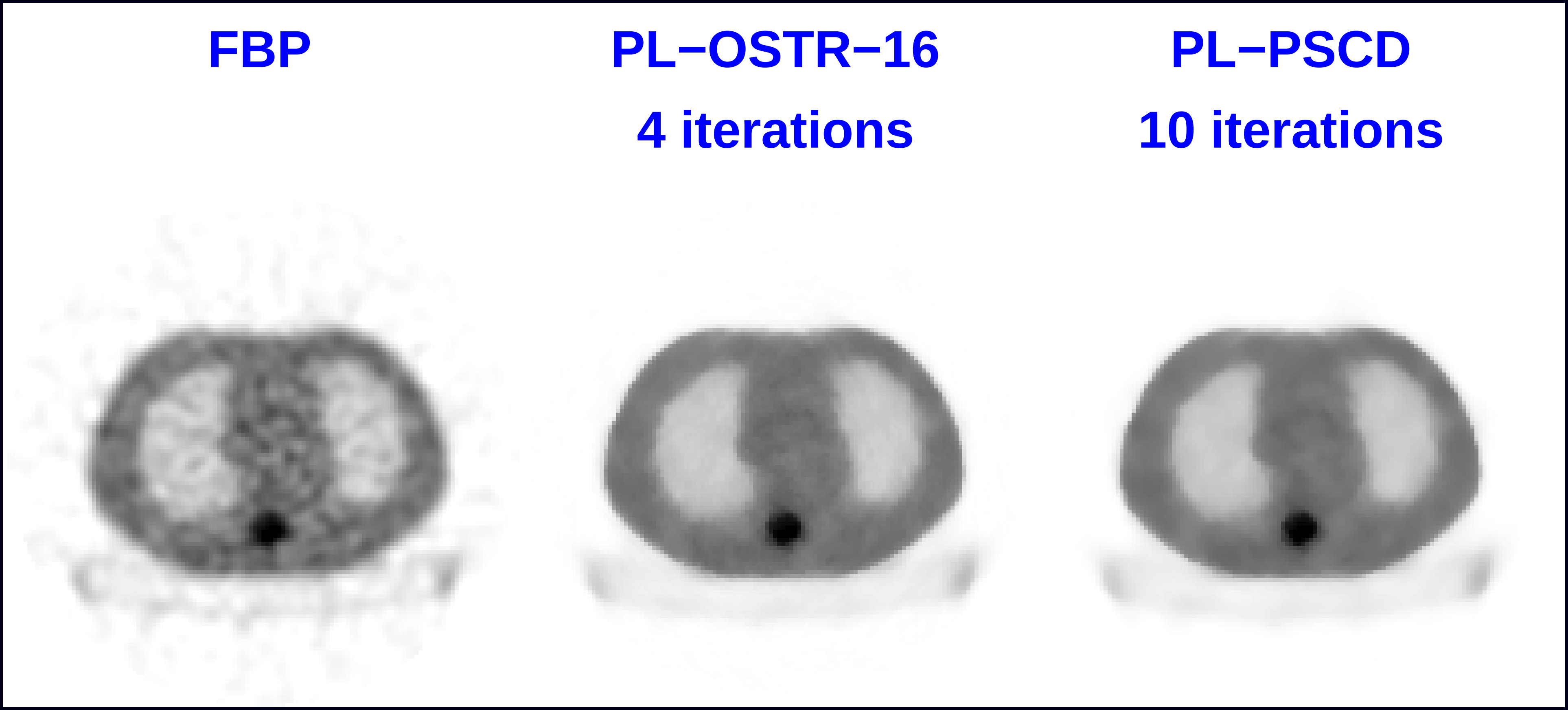
Algorithm Convergence



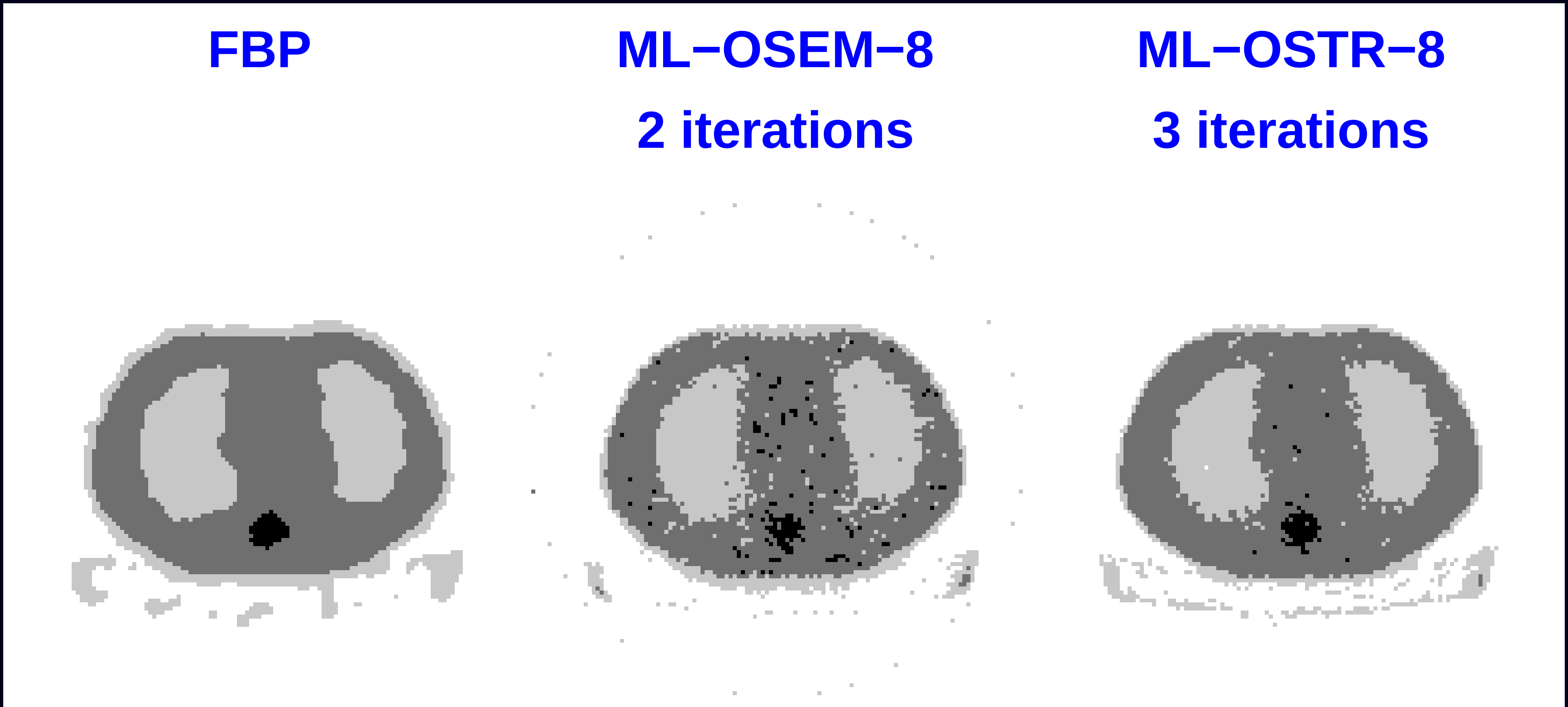
Reconstructed Images



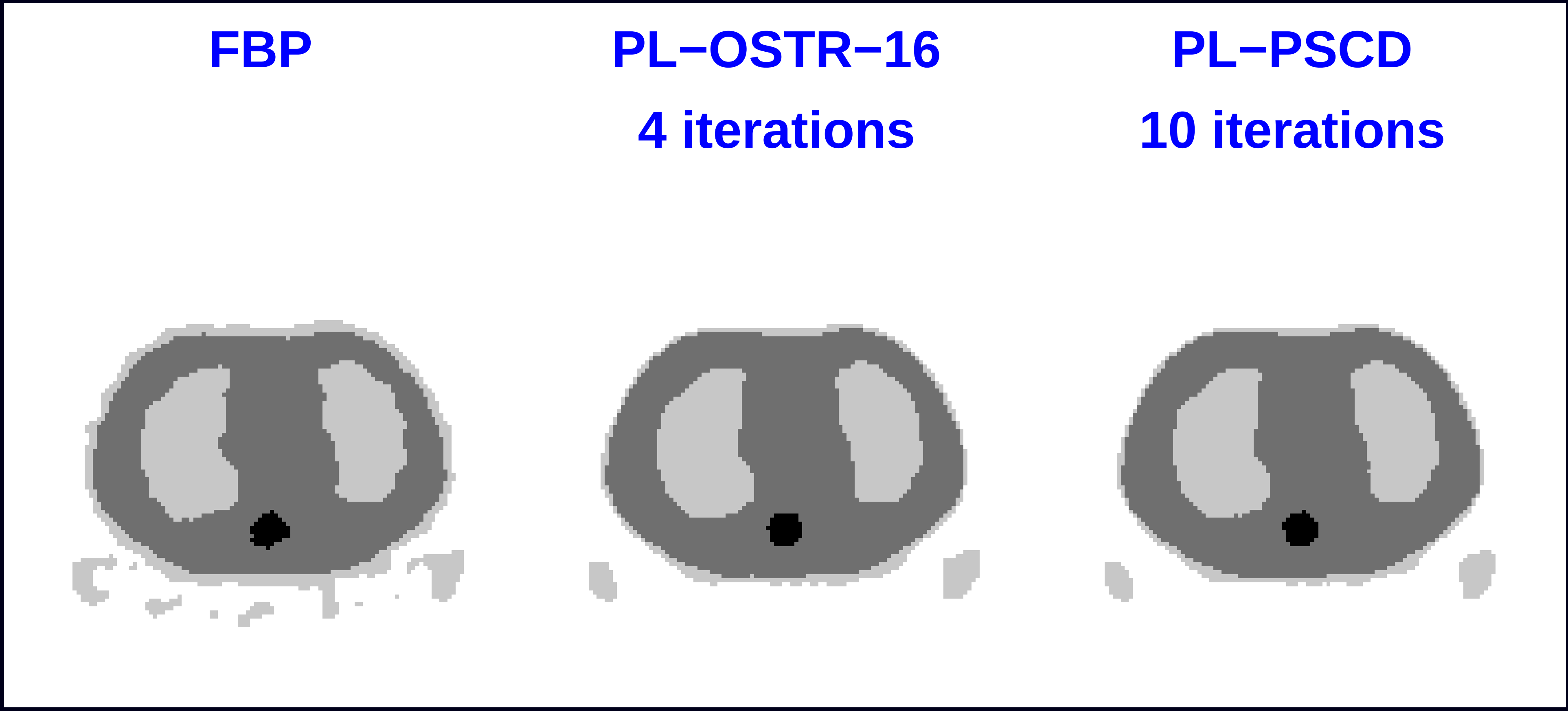
Reconstructed Images

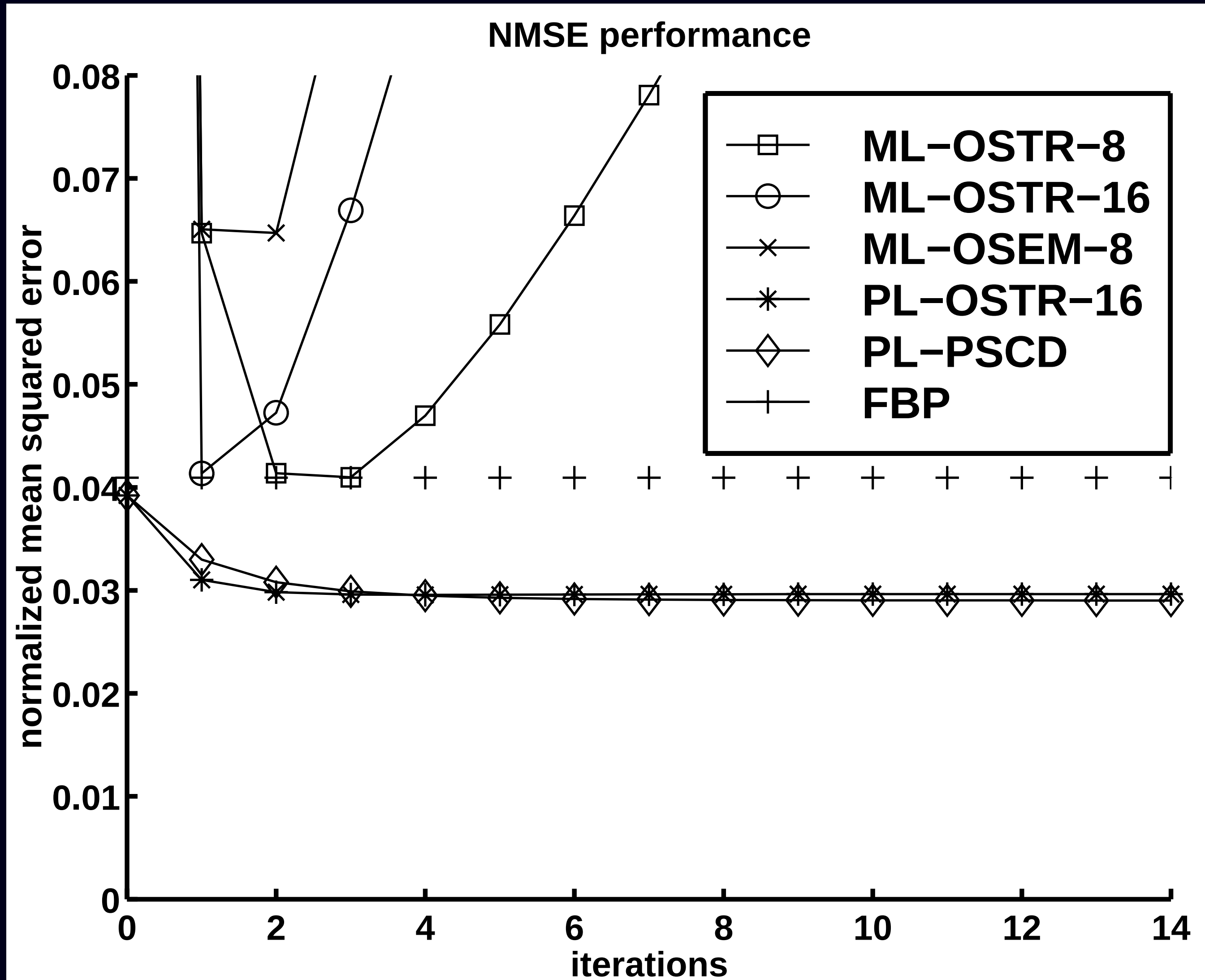


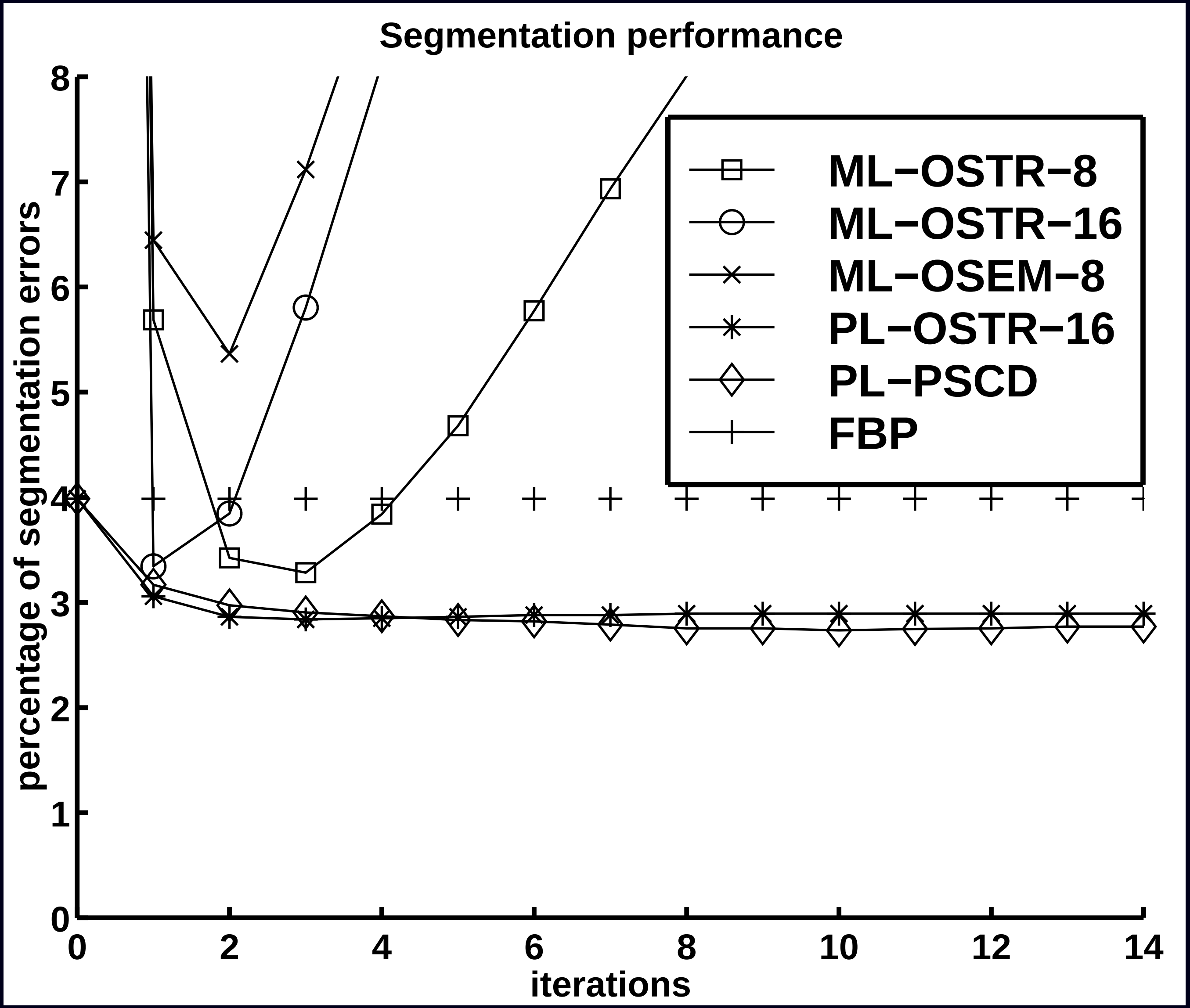
Segmented Images



Segmented Images

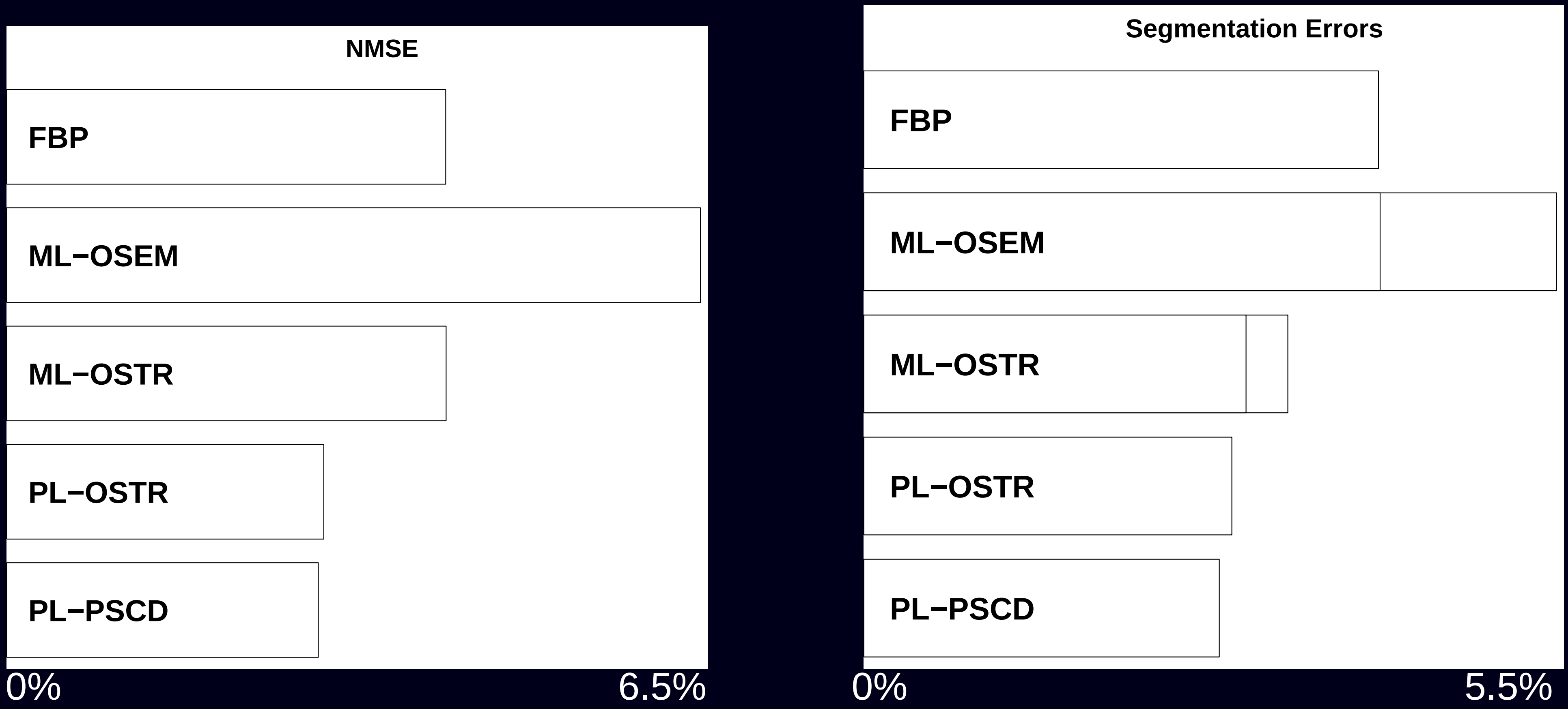



NMSE performance
normalized mean squared error
iterations
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0
2
4
6
8
10
12
14
ML−OSTR−8
ML−OSTR−16
ML−OSEM−8
PL−OSTR−16
PL−PSCD
FBP

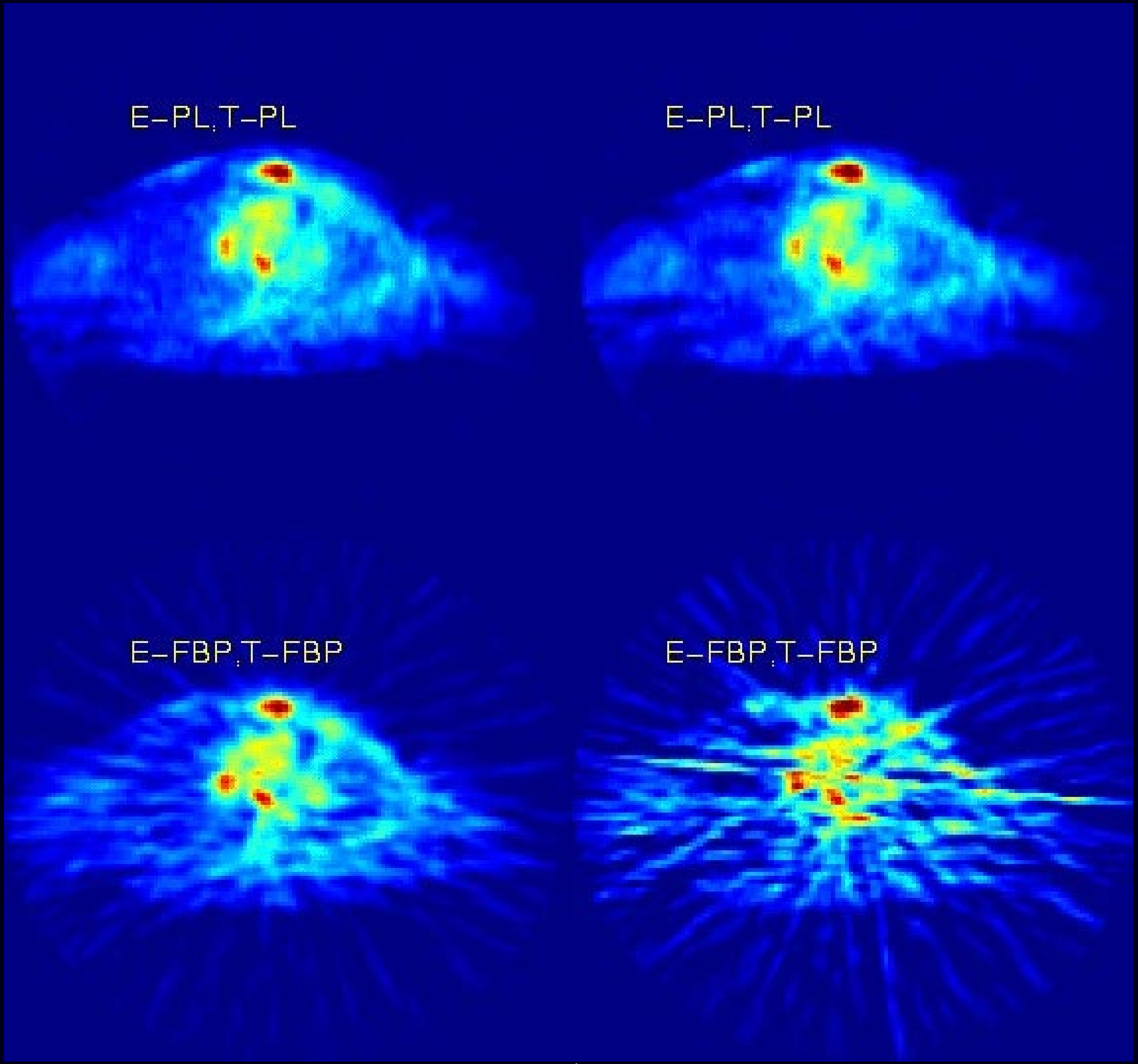


Segmentation performance
percentage of segmentation errors
iterations
0
1
2
3
4
5
6
7
8
0
2
4
6
8
10
12
14
ML−OSTR−8
ML−OSTR−16
ML−OSEM−8
PL−OSTR−16
PL−PSCD
FBP

Quantitative Results



FDG PET Patient Data, PL-OSTR vs FBP



(15-minute transmission scan | 2-minute transmission scan)