Approximate Variance Images for Penalized-Likelihood Image Reconstruction

Jeffrey A. Fessler
University of Michigan

NSS/MIC Nov. 1997

Introduction

- Statistical image reconstruction methods are nonlinear estimators
  $\Rightarrow$ space-variant pixel variances
- Potential applications of variance maps:
  - reconstruction algorithm evaluation
  - imaging system design
  - medical diagnosis (confidence)
  - choosing simulation parameters
  - ???
- Fast approximate variance maps may be useful (cf simulations)
- Variance maps for FBP images: well-known but little used...

Poisson Statistical Model

$$
\begin{aligned}
Y_i &\sim \text{Poisson}\{\bar{Y}_i(\underline{\lambda}^{\text{true}})\} \\
\bar{Y}_i(\underline{\lambda}) &= \sum_j c_i g_{ij} \lambda_j + r_i
\end{aligned}
$$

- Y_i measured emission counts
- $\bar{Y}_i$ modeled mean of Y_i
- λ_j unknown activity in the jth pixel
- $\boldsymbol{G}$ geometric system response $\boldsymbol{G} = \{g_{ij}\}$
- c_i ray factors *e.g.* attenuation and detector efficiency
- r_i random coincidences and scatter.

Image reconstruction: estimate image $\underline{\lambda}$ from sinogram $\underline{Y}$

Penalized-Likelihood Estimators

Log-likelihood:

$$L(\underline{\lambda}, \underline{Y}) = \sum_{i=1}^{n} y_i \log \bar{Y}_i(\underline{\lambda}) - \bar{Y}_i(\underline{\lambda}) + \text{constant}$$

Estimator:

$$\hat{\underline{\lambda}} = \arg\max_{\underline{\lambda} \geq 0} \Phi(\underline{\lambda}, \underline{Y})$$

Penalized-Likelihood Objective Function:

$$\Phi(\underline{\lambda}, \underline{Y}) = L(\underline{\lambda}, \underline{Y}) - \beta R(\underline{\lambda}),$$

where $R(\underline{\lambda})$ is a roughness penalty function.

Fast converging algorithms available for finding minimizer $\hat{\underline{\lambda}}$ of Φ.

Covariance Approximation

Estimator defined implicitly $\Rightarrow$ no explicit expression for covariance.

Approximation from (Fessler, IEEE Tr. Image Proc., Mar. 1996):

$$\boxed{\mathrm{Cov}\{\hat{\underline{\lambda}}\} \approx [\boldsymbol{F} + \beta \boldsymbol{R}]^{-1} \boldsymbol{F} [\boldsymbol{F} + \beta \boldsymbol{R}]^{-1}}$$

- $\boldsymbol{F} = \boldsymbol{G}' \boldsymbol{D}(u_i) \boldsymbol{G}$ Fisher-information matrix
- $\boldsymbol{D}(u_i)$ Diagonal matrix with $\boldsymbol{D}_{ii} = u_i$
- $u_i = c_i^2 / \bar{Y}_i$ Inverse of measurement variance
- $\boldsymbol{R} = \nabla^2 R$ Hessian of the penalty.

Covariance approximation improves with increasing scan time.

Variance

Variance map: image of the diagonal elements of $\mathrm{Cov}\{\hat{\underline{\lambda}}\}$.

$$\begin{aligned}
\mathrm{Var}\{\hat{\lambda}_j\} &= [\mathrm{Cov}\{\hat{\underline{\lambda}}\}]_{jj} = \underline{e}_j' \mathrm{Cov}\{\hat{\underline{\lambda}}\} \underline{e}_j \\
&\approx \underline{e}_j' [\boldsymbol{F} + \beta \boldsymbol{R}]^{-1} \boldsymbol{F} [\boldsymbol{F} + \beta \boldsymbol{R}]^{-1} \underline{e}_j \\
&= \underline{x}' \boldsymbol{G}' \boldsymbol{D}(u_i) \boldsymbol{G} \underline{x} \\
&= \sum_{i=1}^{n} u_i [\boldsymbol{G} \underline{x}]_i^2
\end{aligned}$$

where $\underline{e}_j$ is the jth standard unit vector and

$$[\boldsymbol{G}' \boldsymbol{D}(u_i) \boldsymbol{G} + \beta \boldsymbol{R}] \underline{x} = \underline{e}_j.$$

One would have to solve this system of equations once for each pixel.
Too expensive (simulations would be cheaper!): $\therefore$ approximate further.

Fisher Information Approximation

From (Fessler and Rogers, IEEE Tr. Image Proc., Sep. 1996):

$$\boldsymbol{F} = \boldsymbol{G}'\boldsymbol{D}(u_i)\boldsymbol{G} \approx \boldsymbol{D}(\kappa_j)\boldsymbol{G}'\boldsymbol{G}\,\boldsymbol{D}(\kappa_j)$$

where $\kappa_j = \sqrt{\frac{\sum_{i=1}^n g_{ij}^2 u_i}{\sum_{i=1}^n g_{ij}^2}}$ is the "effective certainty" of the jth pixel.
(Normalized backprojection of inverse ray variances.)
For homoscedastic Gaussian noise, the κ_j's would all be equal.

New Covariance Approximation

$$\mathrm{Cov}\{\hat{\underline{\lambda}}\} \approx \boldsymbol{D}(\kappa_j^{-1})[\boldsymbol{G}'\boldsymbol{G} + \beta\boldsymbol{R}_2]^{-1}\boldsymbol{G}'\boldsymbol{G}[\boldsymbol{G}'\boldsymbol{G} + \beta\boldsymbol{R}_2]^{-1}\boldsymbol{D}(\kappa_j^{-1})$$

where $\boldsymbol{R}_2 = \boldsymbol{D}(\kappa_j^{-1})\boldsymbol{R}\boldsymbol{D}(\kappa_j^{-1})$

Proposed Variance Approximation

$$\mathrm{Var}\{\hat{\lambda}_j\} \approx \frac{\sigma_j^2(\beta/\kappa_j^2)}{\kappa_j^2}$$

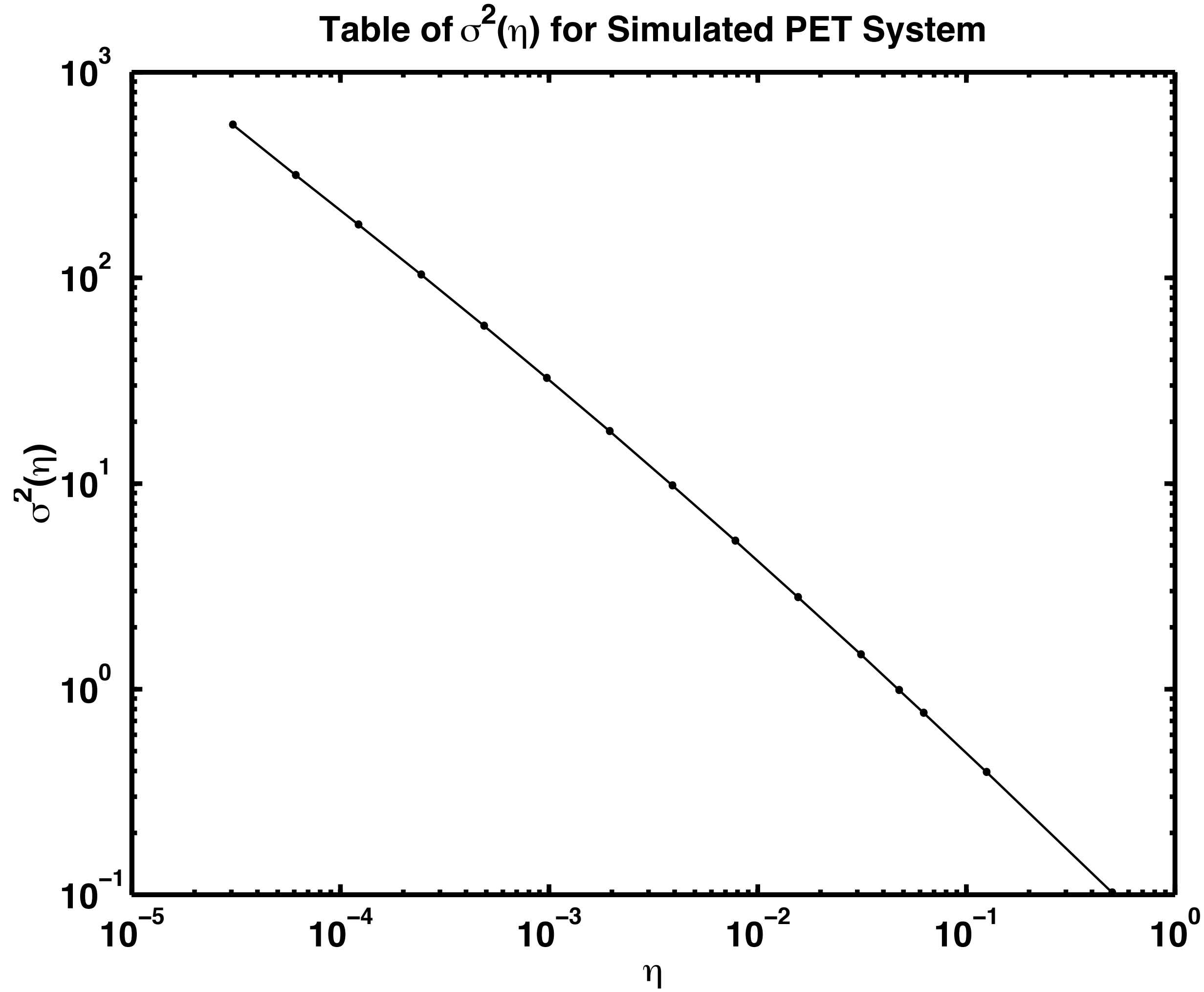
where $\sigma_j^2(\eta) \triangleq \underline{e}_j'[\boldsymbol{G}'\boldsymbol{G} + \eta \boldsymbol{R}]^{-1}\boldsymbol{G}'\boldsymbol{G}[\boldsymbol{G}'\boldsymbol{G} + \eta \boldsymbol{R}]^{-1}\underline{e}_j$

- In PET the σ_j^2 function(s) depend only on the system geometry and the penalty function, $\Rightarrow$ precompute / tabulate once.
- All object-dependent factors are contained in the κ_j's.
- $\sigma_j^2(\eta)$ is the variance of $\hat{\lambda}_j$ under homoscedastic Gaussian noise and reconstruction with regularization parameter η.

For shift-invariant systems:

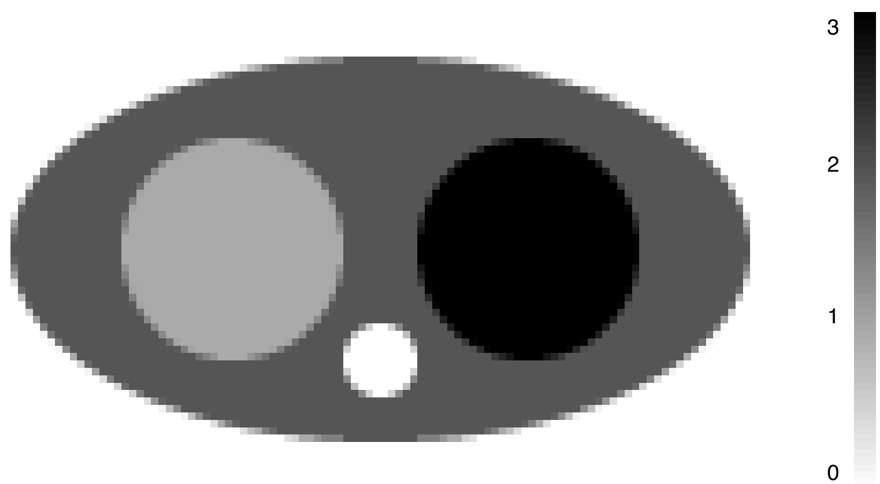
- The σ_j^2 functions are all identical
- $\sigma^2(\eta)$ easily computed using FFTs

Example "Table"

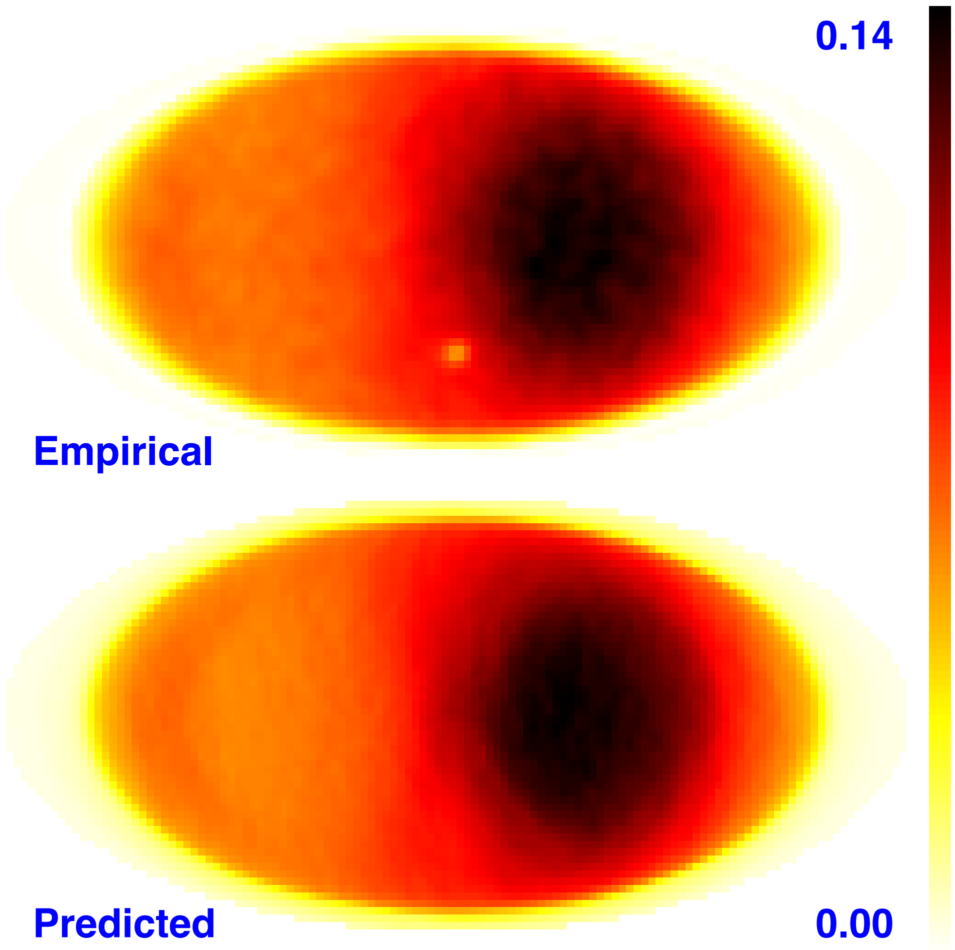


Simulation

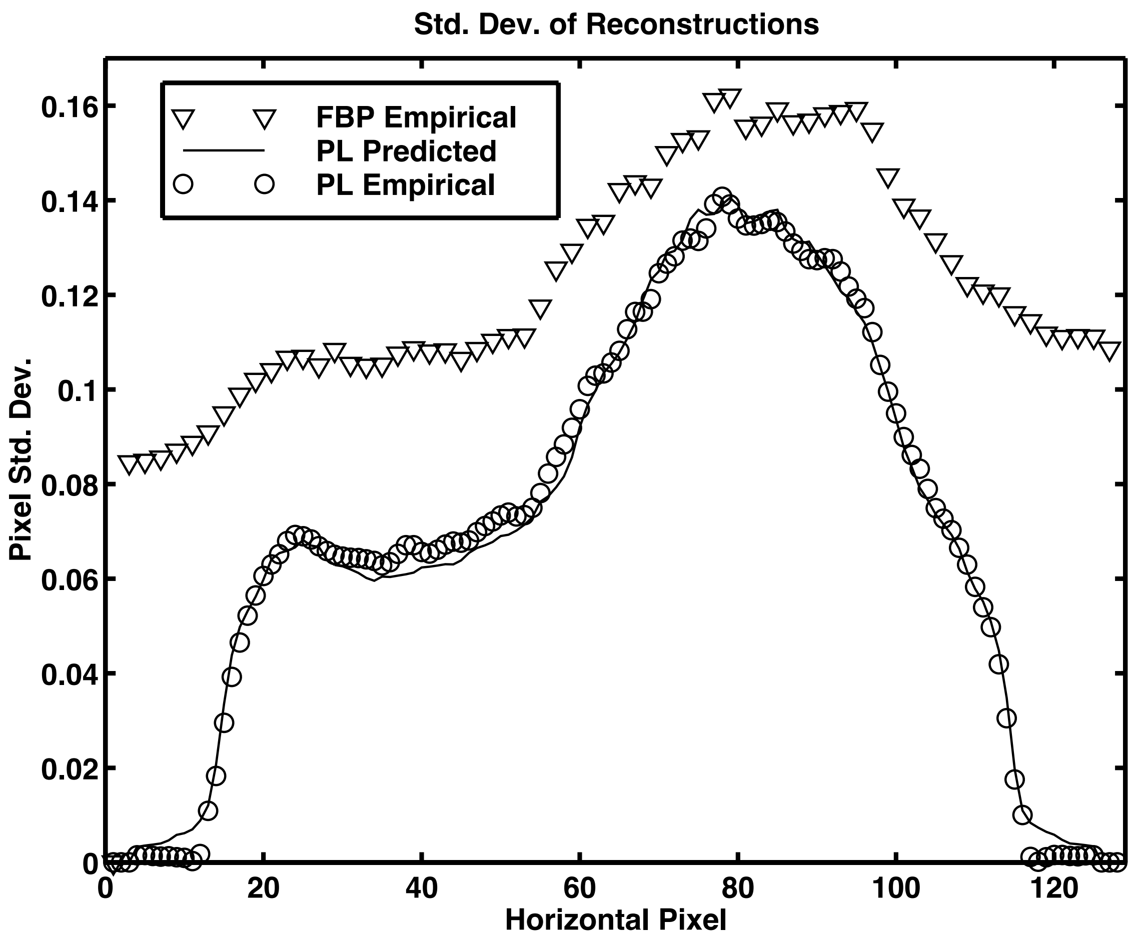
- 2000 realizations
- PET digital emission phantom / nonuniform attenuation
- Modified quadratic penalty
- 10 iterations of PML-SAGE-3
- Nonnegativity enforced



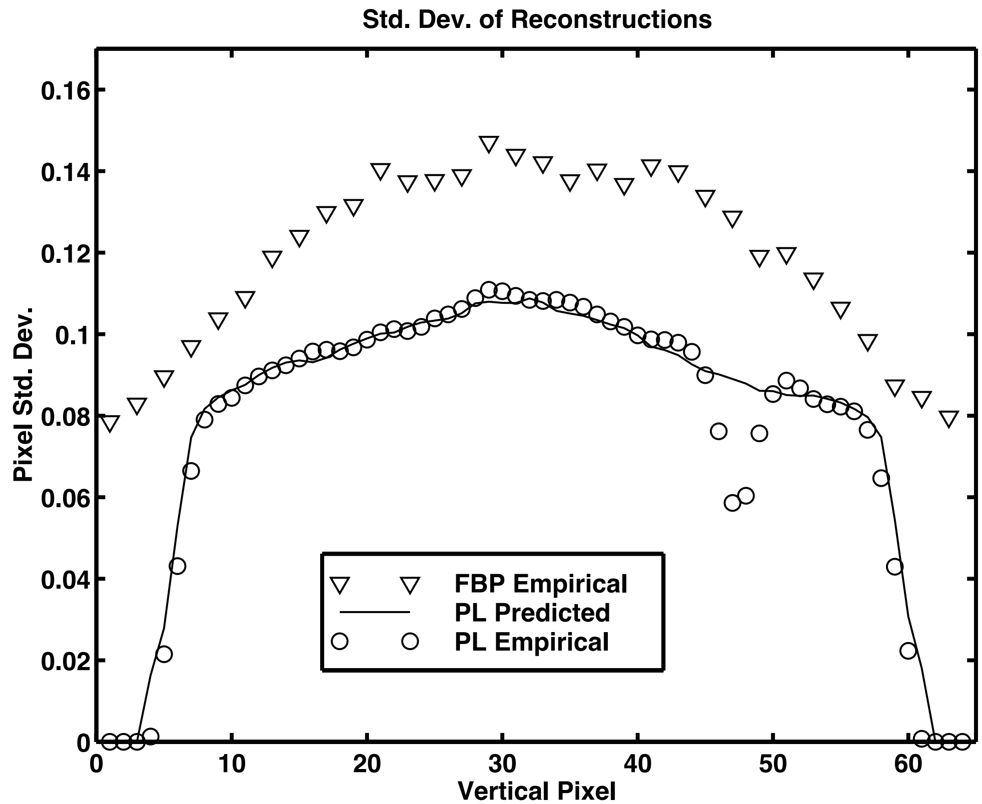
Standard Deviation Maps



Center Horizontal Profile

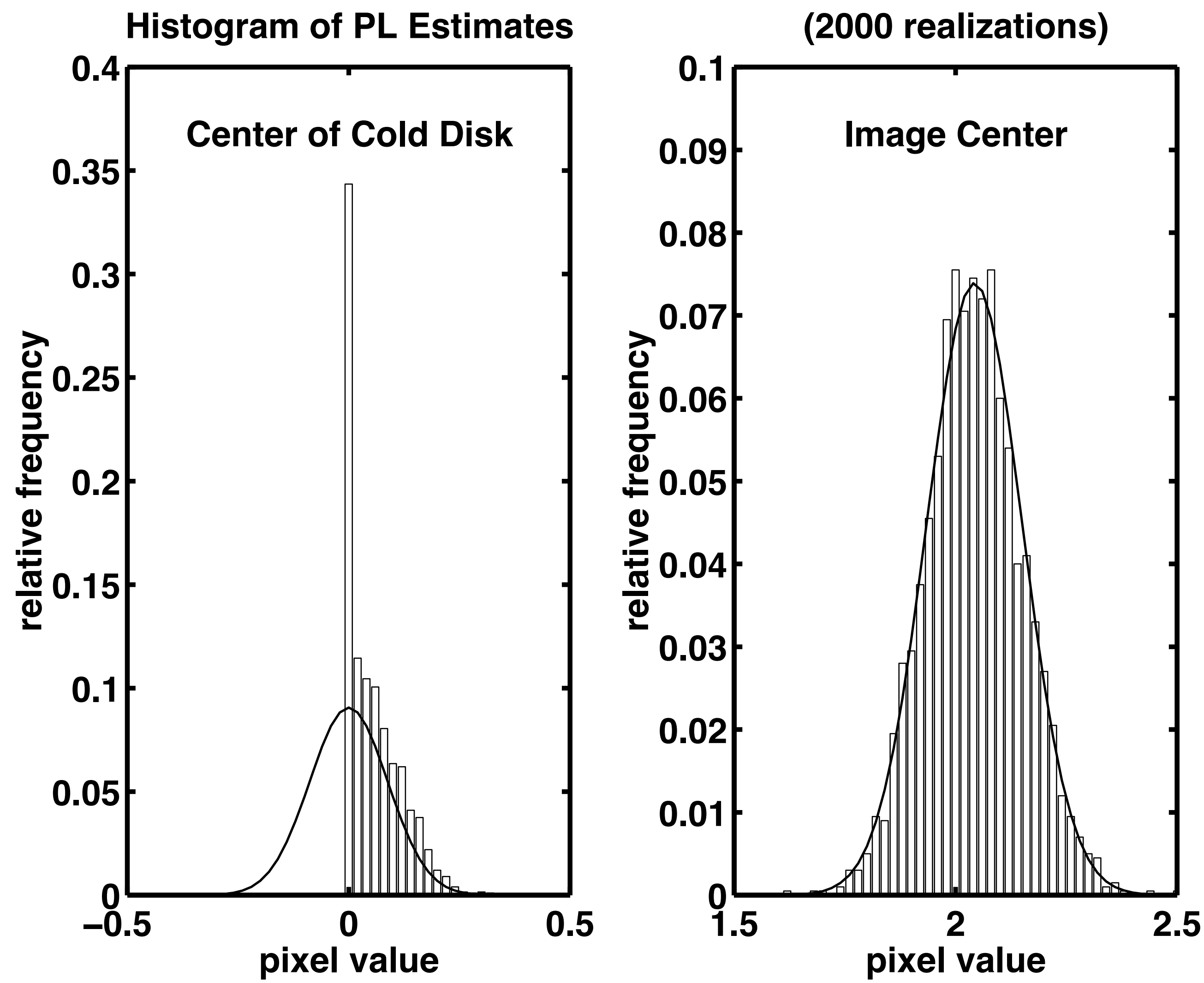


Center Vertical Profile

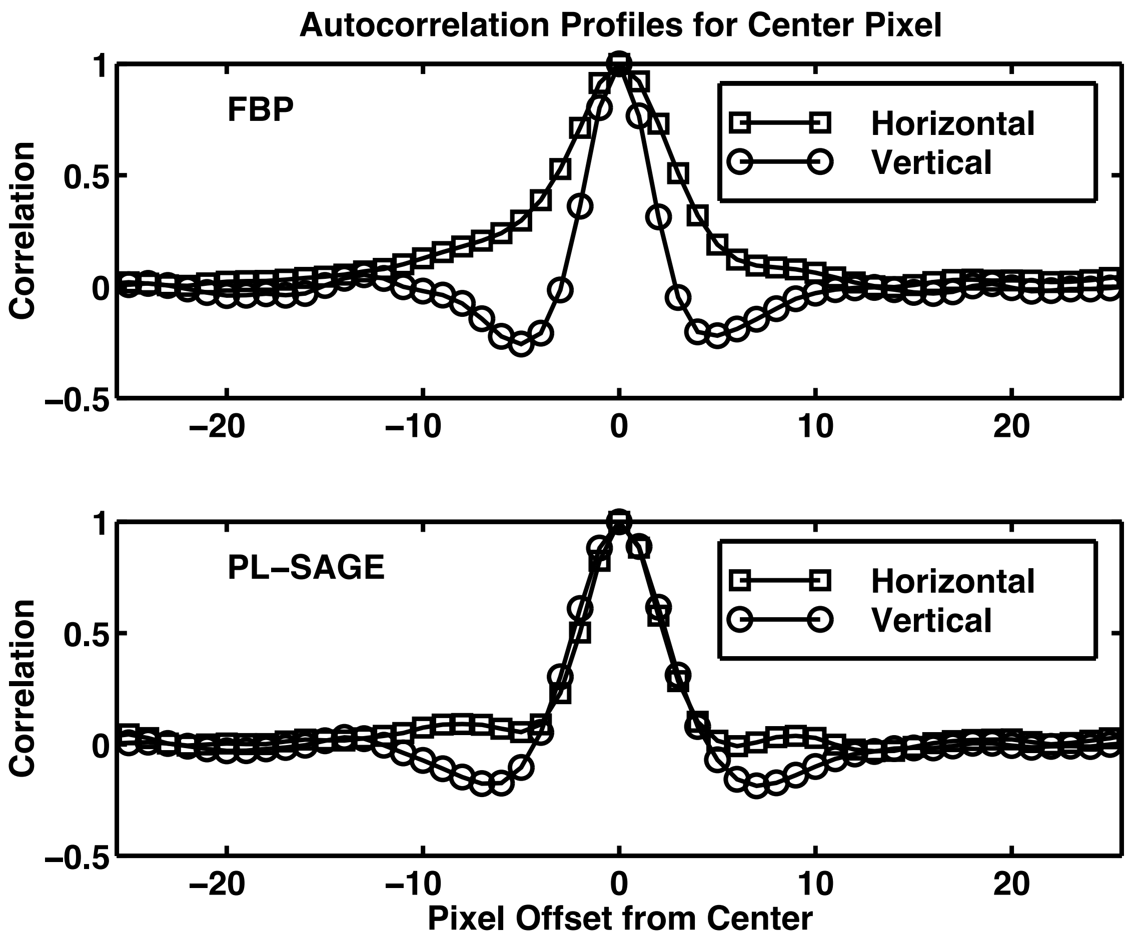


Mismatch in cold spot where nonnegativity constraint is very active.

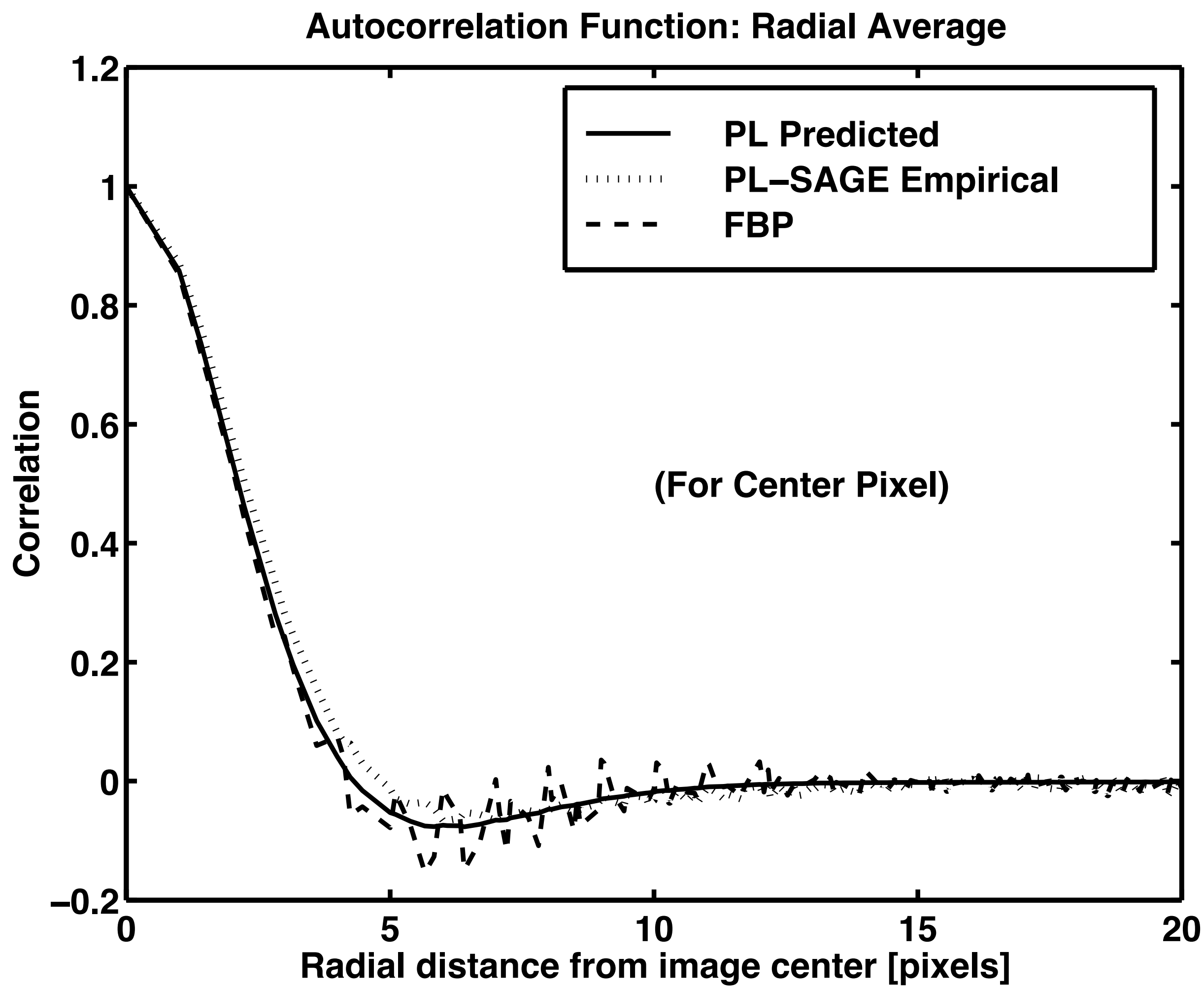
Histograms


Histogram of PL Estimates
(2000 realizations)
Center of Cold Disk
Image Center
relative frequency
pixel value
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
–0.5
0.5
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
1.5
2
2.5

Autocorrelation Functions



Autocorrelation Function: Radial Average



Summary and Future Work

- Fast approximation for pixel variances in penalized-likelihood or penalized weighted least-squares image reconstruction methods
- Very fast for shift-invariant systems
- Over-estimates variance in low-count regions
- Refinement needed for asymmetric autocorrelation functions
- Extend to 3D and shift-variant systems
- When is it useful?

Preprints: http://www.eecs.umich.edu/~fessler/