

# Regularization for Uniform Spatial Resolution Properties in Penalized-Likelihood Image Reconstruction

J. Webster Stayman\*, *Student Member, IEEE*, and Jeffrey A. Fessler, *Member, IEEE*

**Abstract**—Traditional space-invariant regularization methods in tomographic image reconstruction using penalized-likelihood estimators produce images with *nonuniform* spatial resolution properties. The local point spread functions that quantify the smoothing properties of such estimators are space-variant, asymmetric, and object-dependent even for space-invariant imaging systems. We propose a new quadratic regularization scheme for tomographic imaging systems that yields increased spatial uniformity and is motivated by the least-squares fitting of a parameterized local impulse response to a desired global response. We have developed computationally efficient methods for PET systems with shift-invariant geometric responses. We demonstrate the increased spatial uniformity of this new method versus conventional quadratic regularization schemes in simulated PET thorax scans.

**Index Terms**—Gauss–Markov prior, PET, tomography.

## I. INTRODUCTION

STATISTICAL image reconstruction methods provide improved noise and resolution properties over conventional nonstatistical methods such as filtered backprojection (FBP). However, methods based purely on the maximum-likelihood estimate produce overly noisy images. This noise may be reduced by stopping the iterative procedure used to find the maximum-likelihood estimate before convergence [1], by iterating until convergence followed by post-smoothing [2], or by including a roughness penalty term in the objective function [3]. It is difficult to control resolution properties with stopping criteria. Post-smoothing methods allow for better resolution control but require iteration until convergence. Since unregularized algorithms converge slowly, penalized-likelihood methods are desirable.

However, there are disadvantages with penalized-likelihood methods that use conventional regularization schemes. Space-invariant penalties lead to object-dependent nonuniform resolution properties [4], [5]. For emission tomography, such estimators tend to smooth the image more in high-count regions than in low-count regions. The local point spread functions [4], [6] that quantify this space-variant smoothing can also be highly asymmetric, indicating anisotropic smoothing. These

asymmetric point spread functions mean that objects within an image are distorted nonuniformly. For example, circular objects will appear elliptical due to anisotropic blurring (see Fig. 15).

These distortions have been noted by clinical colleagues in positron emission tomography (PET) scans. Lymph nodes are often found near the edge of an anatomical slice where the point spread functions are particularly asymmetric. Lymph nodes, which appeared essentially radially symmetric in FBP reconstructions (due to the isotropic smoothing of FBP), appeared elliptical in penalized-likelihood image reconstructions using traditional regularization methods.

Since conventional regularization produces images with nonuniform resolution properties, one also cannot select the regularization parameter intuitively. With FBP the noise-resolution tradeoff is controlled through the cutoff frequency  $f_c$  of the filter. There is a direct relationship between  $f_c$  and the *global* full-width half-maximum (FWHM) resolution of the reconstructed image. Such a direct relation does not exist with penalized-likelihood reconstructions with conventional regularizations.

One attempt to analyze and reduce the resolution nonuniformity was presented in [4]. The shift-variant regularization method proposed in [4], which is based on the aggregate certainty of measurement rays intersecting each pixel, provides increased spatial uniformity over conventional space-invariant regularization. However, the local point spread functions are still highly asymmetric.

In this paper we present a parameterization of the quadratic roughness penalty function, which in turn parameterizes the local impulse response functions. We then propose a novel method for determining the penalty function coefficients motivated by a least-squares fitting of the parameterized local impulse response to a desired shift-invariant response (Section III). We describe a computationally efficient noniterative method for computing the coefficients for an idealized PET system (Section VI). This new method provides increased spatial uniformity compared to the certainty-based method of [4] and to conventional regularization techniques. We demonstrate this increased uniformity through an investigation of the local point spread functions (Section V). In addition, we perform a noise investigation on simulated data as well as a qualitative investigation using digital thorax phantom data (Section VI).

Since the proposed quadratic regularization method provides nearly global resolution uniformity, one can use the direct relationship between the regularization parameter and the global

Manuscript received August 10, 1998; revised March 8, 2000. This work was supported in part by the Whitaker Foundation and NIH Grants CA-60711 and CA-54362. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was R. Leahy. *Asterisk indicates the corresponding author.*

\*J. W. Stayman and J. A. Fessler are with the EECS Department, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: stayman@eeecs.umich.edu).

Publisher Item Identifier S 0278-0062(00)05463-X.

FWHM resolution to specify a desired resolution for reconstruction. Therefore, the proposed regularization possesses the intuition of FBP with respect to resolution and performs better than FBP in terms of variance.

Whether uniform spatial resolution is essential is an open question. Uniform resolution properties may, in fact, be undesirable for certain tasks. One could use statistical criteria to choose the regularization parameter [7] and accept the anisotropic smoothing properties of the estimator. Alternatively, one may desire specific nonuniform resolution properties through modification of the penalty (e.g., incorporation of anatomical side and boundary information [8], [9]). Qi and Leahy have investigated a shift-variant regularization method that optimizes the local contrast to noise ratio in an attempt to improve lesion detectability [10], [11].

However, for high-resolution PET images, the geometric distortions due to conventional regularizations may be undesirable for tasks requiring shape preservation. Therefore, resolution uniformity would be important. For cross-patient studies or single-patient studies taken over a period of time, one would presumably desire the same resolution properties across images for comparison. Similarly, for full-body PET scans with multiple table positions and cross-modality image registration, these space-variant resolution properties and geometric distortions can contribute to registration errors. As mentioned above, in some cases one may desire nonuniform resolution properties. The methods described in this paper can also be applied to user-specified nonuniform resolution criteria (e.g., regions of isotropic smoothing with sharp boundaries) using space-varying regularization methods as in [12] and [13].

In this paper we focus on the resolution properties of penalized-likelihood estimators that are iterated until convergence. Other studies have investigated resolution properties of unregularized maximum-likelihood expectation-maximization algorithms as a function of iteration [6], [14].

Real imaging systems usually have intrinsically nonuniform resolution properties. Single-photon-emission computed tomography (SPECT) systems have depth-dependent resolution [15], and PET systems often have resolution nonuniformity due to crystal penetration effects [16]. While the analysis presented here applies generally, the resulting design can be computationally expensive. We have developed a computationally fast practical method for an idealized PET imaging system with a shift-invariant geometric response, but including ray-dependent attenuation and detector effects. The central region in the field of view of many PET systems tends to be nearly shift-invariant and can be accurately modeled in this way.

## II. BACKGROUND

We focus on emission tomography, although the method applies generally. Let  $\underline{\lambda} = [\lambda_1, \dots, \lambda_p]'$  represent the non-negative emission rates for an object discretized into  $p$  pixels, where  $'$  denotes the Hermitian transpose. Detectors surrounding the object count photons (SPECT) or photon pairs (PET) that are emitted from the object. Measurements are denoted by the

random vector  $\underline{Y} = [Y_1, \dots, Y_N]'$ . These measurements are Poisson with means given by

$$\bar{Y}_i(\underline{\lambda}) = \sum_{j=1}^p a_{ij}\lambda_j + r_i^E, \quad i = 1, \dots, N$$

where the  $a_{ij}$ 's represent nonnegative constants that characterize the tomographic system, and the  $r_i^E$ 's are nonnegative constants that specify the contribution due to background events (background radiation, random coincidences, scatter, etc.). Given measurements  $\underline{Y}$ , we would like to reconstruct  $\underline{\lambda}$ , assuming the  $a_{ij}$ 's and  $r_i^E$ 's are known.

We will focus on penalized-likelihood estimators (PLE's) of the form

$$\hat{\underline{\lambda}}(\underline{Y}) = \arg \max_{\underline{\lambda} \in \underline{\Lambda}} L(\underline{\lambda}, \underline{Y}) - R(\underline{\lambda})$$

where

$\underline{\Lambda}$	set of feasible images;
$L(\underline{\lambda}, \underline{Y})$	log-likelihood;
$R(\underline{\lambda})$	roughness penalty.

For the Poisson model, the log-likelihood is

$$L(\underline{\lambda}, \underline{Y}) = \sum_{i=1}^N Y_i \log \bar{Y}_i(\underline{\lambda}) - \bar{Y}_i(\underline{\lambda}) - \log Y_i!$$

We focus on pairwise roughness penalties of the following form

$$R(\underline{\lambda}) = \sum_{j=1}^p \frac{1}{2} \sum_{k=1}^p w_{jk} \psi(\lambda_j - \lambda_k) \quad (1)$$

where  $\psi$  is a symmetric convex function.

In the case of a quadratic penalty,  $\psi(x) = x^2/2$  and the roughness penalty may be written in matrix form:  $R(\underline{\lambda}) = (1/2)\underline{\lambda}'\mathbf{R}\underline{\lambda}$ , where the  $p \times p$  matrix  $\mathbf{R}$  has elements defined by

$$\mathbf{R}_{jk} = \begin{cases} \sum_{l=1}^p \frac{1}{2} (w_{lj} + w_{jl}), & k = j \\ -w_{jk}, & k \neq j. \end{cases} \quad (2)$$

For a space-invariant penalty using a first-order neighborhood, the conventional choice is  $w_{jk} = \beta$  for the horizontal and vertical neighbors, and zero otherwise. The regularization parameter  $\beta$  controls the noise-resolution tradeoff. Large  $\beta$  values induce smoother reconstructions, hence lower noise. For a second-order penalty, one often includes  $w_{jk} = \beta/\sqrt{2}$  for the diagonal neighbors in addition to the first-order neighbors.

The mean of an estimator is given by

$$\underline{\mu}(\underline{\lambda}) = E_{\underline{\lambda}}[\hat{\underline{\lambda}}(\underline{Y})] = \int \hat{\underline{\lambda}}(\underline{Y}) f(\underline{Y}; \underline{\lambda}) d\underline{Y}$$

where  $f$  is the Poisson measurement distribution. The local impulse response [4] at the  $j$ th pixel is defined as

$$\underline{l}^j \triangleq \lim_{\delta \rightarrow 0} \frac{\underline{\mu}(\underline{\lambda} + \delta \underline{e}^j) - \underline{\mu}(\underline{\lambda})}{\delta} = \frac{\partial}{\partial \lambda_j} \underline{\mu}(\underline{\lambda})$$

where  $\underline{e}^j$  represents the  $j$ th unit vector.<sup>1</sup> The local impulse response depends on the estimator  $\hat{\underline{\lambda}}$ , the object  $\underline{\lambda}$ , and the pixel

<sup>1</sup>Throughout the paper, the superscript  $j$  represents an index, not an exponent.

position  $j$ . From [4], for PLE's with quadratic penalties,<sup>2</sup> the local impulse response may be well approximated by

$$\underline{l}^j \approx \underline{l}^j(\mathbf{R}) \triangleq [\mathbf{A}'\mathbf{D}\mathbf{A} + \mathbf{R}^{\text{sym}}]^{-1} \mathbf{A}'\mathbf{D}\mathbf{A}\underline{c}^j \quad (3)$$

where  $\mathbf{A}$  is a  $N \times p$  matrix of the  $\{a_{ij}\}$  elements, for emission tomography,<sup>3</sup>  $\mathbf{D} = D[1/\overline{Y}_i(\lambda)]$  is a diagonal matrix with elements  $1/\overline{Y}_i(\lambda)$ , and  $\mathbf{R}^{\text{sym}} \triangleq (1/2)(\mathbf{R} + \mathbf{R}')$  is the symmetric component of  $\mathbf{R}$ . When  $\overline{Y}_i(\lambda)$  is unknown, one can estimate the local impulse response  $\underline{l}^j$  by using a simple plug-in technique where the observed measurements  $y_i$  replace  $\overline{Y}_i(\lambda)$ . Since  $D[1/y_i]$  is sandwiched between the projection and back-projection operators, there is an implicit smoothing and even noisy  $y_i$  tend to produce relatively accurate estimates of  $\mathbf{A}'\mathbf{D}\mathbf{A}$ .

The approximation (3) for the local impulse response is the tool we use below for the design and evaluation of different quadratic regularization methods.

### III. PENALTY DESIGN METHODS

Our goal is to find a penalty function  $R(\lambda)$  that yields reconstructed images with some arbitrary desired space-invariant response. For example, we may desire penalty functions that produce a global impulse response with a Gaussian shape and some specified FWHM resolution. If we restrict ourselves to quadratic penalty functions, we can formulate such problems in terms of the design of the penalty matrix  $\mathbf{R}$ . Equivalently, we may design  $\mathbf{R}^{\text{sym}}$ , since only the symmetric component of  $\mathbf{R}$  affects the objective function for quadratic penalties. We restrict  $\mathbf{R}$  to be nonnegative definite to maintain the concavity of the penalized-likelihood objective function.

Therefore, we would like to find a nonnegative definite  $\mathbf{R}$  according to an optimization criterion such as the following:<sup>4</sup>

$$\hat{\mathbf{R}} = \arg \min_{\mathbf{R} \geq 0} \sum_{j=1}^p d(\underline{l}^j(\mathbf{R}), \underline{l}_0^j) \quad (4)$$

where  $d(\underline{l}^j, \underline{l}_0^j)$  is some measure of disparity between the local impulse response,  $\underline{l}^j$  and a desired space-invariant response,<sup>5</sup>  $\underline{l}_0^j$ . Solving (4) by plugging in (3) appears to be computationally intractable.

Practical penalties use only a small neighborhood of pixels for the penalty support (e.g., first- and second-order neighborhoods). Therefore we reformulate the penalty design problem in

<sup>2</sup>In [4], an approximation for the local impulse response was derived for symmetric  $\mathbf{R}$ . For an asymmetric  $\mathbf{R}$ , the scalar  $\lambda' \mathbf{R} \lambda = (\lambda' \mathbf{R} \lambda)' = \lambda' \mathbf{R}' \lambda$ . Therefore,  $R(\lambda) = (1/2)[(1/2)\lambda' \mathbf{R} \lambda + (1/2)\lambda' \mathbf{R}' \lambda] = (1/2)\lambda' [(1/2)(\mathbf{R} + \mathbf{R}')] \lambda = (1/2)\lambda' \mathbf{R}^{\text{sym}} \lambda$ . If an asymmetric  $\mathbf{R}$  matrix were used, only the symmetric component of  $\mathbf{R}$  would influence the objective function.

<sup>3</sup>The formulation given in (3) also holds for transmission tomography with  $\mathbf{D} = D[\overline{Y}_i(\lambda) - r_i]^2 / \overline{Y}_i(\lambda)$ .

<sup>4</sup>The notation  $\mathbf{R} \geq 0$  indicates that this minimization is over nonnegative definite  $\mathbf{R}$ .

<sup>5</sup>One might choose a space-variant  $\underline{l}_0^j$  for user-specified nonuniform resolution properties. A space-invariant  $\underline{l}_0^j$  is required only for the practical implementation discussed in Section IV. For a desired space-invariant response  $\underline{l}_0^j$  is a function of the pixel position  $j$  only in that the desired response must be centered at pixel  $j$ . That is, since the local impulse response at pixel  $j$  is centered at pixel  $j$ , we must shift the desired response to that location for comparison using  $d(\cdot, \cdot)$ .

terms of these small support neighborhoods by parameterizing the penalty matrix.

#### A. Penalty Matrix Parameterization

For a shift-invariant quadratic penalty, one can treat the penalty matrix  $\mathbf{R}$  as a space-invariant filtering operator. Therefore multiplying  $\mathbf{R}$  by the image  $\underline{\lambda}$  is equivalent to convolving the image with a kernel,<sup>6</sup>

$$\mathbf{R}\underline{\lambda} \equiv k_R(m, n) * * \lambda(m, n).$$

For example, the conventional first-order penalty described below (2) has the following kernel

$$k_{R_0} = \beta \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \quad (5)$$

The design of a space-invariant  $\mathbf{R}$  is like a filter design problem with constraints on the kernel  $k_R$ . Since  $\mathbf{R}$  should yield a zero penalty for uniform regions, the filter represented by  $k_R$  should have zero DC gain. (The kernel elements must sum to zero.) Since only the symmetric portion of  $\mathbf{R}$  influences the penalized-likelihood objective function and the local impulse response in (3), we need only to consider symmetric kernels<sup>7</sup>  $k_R$  for representing the action of a space-invariant  $\mathbf{R}^{\text{sym}}$ . Last, we choose to require that  $\mathbf{R}^{\text{sym}}$  be nonnegative definite to guarantee concavity of the penalized-likelihood objective function. Therefore, for the space-invariant penalty, we restrict kernels to those whose Fourier transform is nonnegative, so that the eigenvalues of  $\mathbf{R}^{\text{sym}}$  are nonnegative.

To achieve these goals, we parameterize the kernel  $k_R$  in terms of a small number of bases such as those having the following form

$$b_{(k,l)}(m, n) = 2\delta(m, n) - \delta(m-k, n-l) - \delta(m+k, n+l)$$

where  $\delta(\cdot)$  represents a 2-D discrete impulse function, and  $m$  and  $n$  represent spatial coordinates. A collection of such  $b_{(k,l)}(m, n)$  functions for various  $(k, l)$  pairs forms a basis for valid kernels of space-invariant  $\mathbf{R}^{\text{sym}}$  matrices. For example, for a first-order neighborhood

$$b_{(1,0)} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad b_{(0,1)} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

is a basis for valid kernels of  $\mathbf{R}^{\text{sym}}$ . For a second-order neighborhood,  $\{b_{(1,0)}, b_{(0,1)}, b_{(1,1)}, b_{(1,-1)}\}$  forms a valid basis set, where

$$b_{(1,-1)} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad b_{(1,1)} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 0 \end{bmatrix}.$$

<sup>6</sup>We use  $\equiv$  since the left-hand side (LHS) is a vector, but the right-hand side (RHS) is a 2-D image. The two sides are equivalent in that the vector is a lexicographic reordering of the 2-D image.

<sup>7</sup>Consider a horizontal penalty and two neighboring pixels; one on the left and one on the right. For a symmetric  $\mathbf{R}^{\text{sym}}$ , the penalty applied on the right pixel from the left pixel is the same as the penalty applied on the left from the right. Therefore, the left and right sides of the kernel must be the same for the same penalty to be applied in both directions.

In general, any valid kernel for a space-invariant penalty matrix  $\mathbf{R}^{\text{sym}}$  may be specified by a linear combination of such basis functions:

$$k_R = \sum_{k,l} r_{kl} b_{(k,l)} \equiv \mathbf{B}\underline{r} \quad (6)$$

where  $r_{kl}$  represent the basis coefficients. Let  $\underline{r}$  represent the vector of all  $r_{kl}$  for a given neighborhood of support. Define  $n_w$  to be the number of  $(k, l)$  pairs (the number of basis functions) and define the  $p \times n_w$  matrix  $\mathbf{B}$  with column vectors of lexicographically ordered basis functions,  $b_{(k,l)}$ .

For uniform resolution properties, we *require* a space-variant regularization matrix  $\mathbf{R}$ . Therefore, we extend the kernel representation (6) and let  $\mathbf{R}$  be parameterized by a *space-variant* set of coefficients  $\underline{r}^j$ , where  $j$  represents the  $j$ th pixel.

Let  $(n_j, m_j)$  denote the spatial coordinates of the  $j$ th pixel and define  $\mathbf{B}^j$  to be a  $p \times n_w$  matrix of shifted basis functions, with each column having elements defined by the lexicographically ordered bases,  $b_{(k,l)}(m - m_j, n - n_j)$ . In the case of a space-invariant  $\mathbf{R}$  matrix,  $\mathbf{R}\underline{e}^j = \mathbf{B}^j \underline{r}$ . (With the choice of symmetric bases described previously,  $\mathbf{R} = \mathbf{R}^{\text{sym}}$ .) To parameterize space-variant  $\mathbf{R}$ , we define the  $j$ th column of  $\mathbf{R}$  by

$$\mathbf{R}\underline{e}^j = \mathbf{B}^j \underline{r}^j, \quad j = 1, \dots, p. \quad (7)$$

In this case,  $\mathbf{R}$  no longer equals  $\mathbf{R}^{\text{sym}}$  in general. However, we may form  $\mathbf{R}^{\text{sym}}$  by  $(1/2)(\mathbf{R} + \mathbf{R}')$ , as stated previously in footnote 2. The parameterization (7) allows for the specification of valid shift-variant  $\mathbf{R}^{\text{sym}}$  by the set of coefficients  $\{\underline{r}^j\}_{j=1}^p$ . To guarantee the nonnegative definiteness of  $\mathbf{R}^{\text{sym}}$  it is sufficient to restrict  $\underline{r}^j$  to be nonnegative. A nonnegative definite  $\mathbf{R}$  means that the penalty  $R(\underline{\lambda}) = \underline{\lambda}' \mathbf{R} \underline{\lambda}$  is always nonnegative. If we restrict  $\underline{r}^j$  to be nonnegative, the local penalty at any pixel will be nonnegative due to our selection of bases. Therefore, the penalty on the entire image will be nonnegative as well.

Using the parameterization described in (7), the problem of determining the  $p \times p$  matrix  $\mathbf{R}$  in (4) is simplified to the problem of determining the  $n_w \cdot p$  coefficients  $\{\underline{r}^j\}_{j=1}^p$ . The penalty design problem thus reduces to the following optimization problem

$$\{\underline{r}^k\}_{k=1}^p = \arg \min_{\{\underline{r}^k\}_{k=1}^p \geq 0} \sum_{j=1}^p d(\underline{l}^j(\{\underline{r}^k\}_{k=1}^p), \underline{l}_0^j). \quad (8)$$

Although this minimization requires less computation than (4), it still appears to be impractical since all  $p$  of the  $\underline{r}^j$  vectors would need to be found simultaneously.

### B. Circulant Simplifications

When the operator  $\mathbf{A}'\mathbf{A}$  is approximately shift-invariant, we may approximate  $\mathbf{A}'\mathbf{A}$  by  $\mathbf{Q}'\mathbf{\Omega}\mathbf{Q}$ , where  $\mathbf{Q}$  is a 2-D discrete Fourier matrix operator and  $\mathbf{\Omega}$  is a diagonal matrix representing a frequency domain filtering operator.<sup>8</sup> However, even if  $\mathbf{A}'\mathbf{A}$

<sup>8</sup>For an ideal tomographic system, the diagonal elements of  $\mathbf{\Omega}$  are approximately the well-known  $1/\rho$  frequency response of the back-projected projection operator. When  $\mathbf{A}'\mathbf{A}$  is nearly shift-invariant, we may compute the elements of  $\mathbf{\Omega}$  by taking the 2-D discrete Fourier transform of  $\mathbf{A}'\mathbf{A}\underline{e}^{j_0}$ , where  $j_0$  is a fixed pixel in the image (usually the center pixel, in practice). When  $j_0$  is not the center pixel we must include appropriate complex exponentials to account for the shifting property of Fourier transforms.

is a shift-invariant operator,  $\mathbf{A}'\mathbf{DA}$  will be shift-variant because of the nonuniform diagonal weighting. Although  $\mathbf{A}'\mathbf{DA}$  is not globally shift-invariant, it is approximately *locally* shift-invariant and we make the following approximation [17] to (3)

$$\begin{aligned} \underline{l}^j(\mathbf{R}) &\approx [\mathbf{Q}'\mathbf{\Omega}^j\mathbf{Q} + \mathbf{Q}'\mathbf{\Psi}^j\mathbf{Q}]^{-1}\mathbf{Q}'\mathbf{\Omega}^j\mathbf{Q}\underline{e}^j \\ &= \mathbf{Q}' \left[ \frac{\mathbf{\Omega}^j}{\mathbf{\Omega}^j + \mathbf{\Psi}^j} \right] \mathbf{Q}\underline{e}^j \end{aligned} \quad (9)$$

where the division is an element-by-element division,  $\mathbf{\Omega}^j \triangleq D[\mathbf{F}\{\mathbf{A}'\mathbf{DA}\underline{e}^j\}]$ , and  $\mathbf{\Psi}^j \triangleq D[\mathbf{F}\{\mathbf{R}^{\text{sym}}\underline{e}^j\}]$ . ( $\mathbf{F}\{\cdot\}$  represents the  $p \times p$  discrete 2-D Fourier transform operator.)

Since local impulse response functions usually vary smoothly with position, we expect that the  $\underline{r}^j$  coefficients of our penalty design will also be smoothly varying. This is also implied by the above locally shift-invariant approximation. For this reason we use the approximation  $\mathbf{R}^{\text{sym}}\underline{e}^j \approx \mathbf{B}^j \underline{r}^j$ . To illustrate this approximation, consider a simple 1-dimensional example with a single  $[-1 \ 2 \ -1]$  basis. For a single basis function there is a single coefficient  $r^j$  for each position  $j$ . In terms of (2), this means  $w_{j, j+1} = -r^j$  and  $w_{j+1, j} = -r^{j+1}$ . If  $r^j$  is smoothly varying (i.e.,  $r^j \approx r^{j+1}$ ), then  $w_{j, j+1} \approx w_{j+1, j}$  and  $\mathbf{R}$  is nearly symmetric. Substituting  $\mathbf{R}^{\text{sym}}\underline{e}^j \approx \mathbf{B}^j \underline{r}^j$  into (9) yields

$$\begin{aligned} \underline{l}^j(\mathbf{R}) &\approx \underline{l}_F^j(\underline{r}^j) \\ \underline{l}_F^j(\underline{r}^j) &\triangleq \mathbf{F}^{-1} \left\{ \frac{\mathbf{F}\{\mathbf{A}'\mathbf{DA}\underline{e}^j\}}{\mathbf{F}\{\mathbf{A}'\mathbf{DA}\underline{e}^j\} + \mathbf{F}\{\mathbf{B}^j \underline{r}^j\}} \right\}. \end{aligned} \quad (10)$$

Combining (8) and (10) yields a separable minimization problem, i.e.,  $\underline{l}^j$  depends only on  $\underline{r}^j$  and not  $\underline{r}^k$  for  $k \neq j$ . Therefore we may determine each  $\underline{r}^j$  separately by

$$\hat{\underline{r}}^j = \arg \min_{\underline{r}^j \geq 0} d(\underline{l}_F^j(\underline{r}^j), \underline{l}_0^j), \quad j = 1, \dots, p. \quad (11)$$

If  $d(u, v) = \|u - v\|^2$ , then (11) is a set of  $p$  constrained nonlinear least-squares (CNLLS) problems, since the dependence on  $\underline{r}^j$  is in the denominator of (10). We have implemented this method using a BFGS quasi-Newton method, but it is still computationally expensive. Thus, we further simplify this nonlinear optimization problem into a linear least-squares problem. Working in the frequency domain simplifies the design problem, as described next.

### C. Linearized Penalty Design

Define  $\underline{L}^j(\underline{r}^j) \triangleq \mathbf{F}\{\underline{l}_F^j(\underline{r}^j)\}$  to be the local frequency response and let  $\underline{L}_0^j \triangleq \mathbf{F}\{\underline{l}_0^j\}$  be the desired frequency response. To solve (11), we want to choose  $\underline{r}^j$  so that  $\underline{L}^j(\underline{r}^j) \approx \underline{L}_0^j$ , i.e., from (10)

$$\underline{L}^j(\underline{r}^j) = \frac{\mathbf{F}\{\mathbf{A}'\mathbf{DA}\underline{e}^j\}}{\mathbf{F}\{\mathbf{A}'\mathbf{DA}\underline{e}^j\} + \mathbf{F}\{\mathbf{B}^j \underline{r}^j\}} \approx \underline{L}_0^j. \quad (12)$$

Rearranging (12) by cross multiplying and simplifying yields

$$\mathbf{F}\{\mathbf{A}'\mathbf{DA}\underline{e}^j\} \odot (1 - \underline{L}_0^j) \approx \underline{L}_0^j \odot \mathbf{F}\{\mathbf{B}^j \underline{r}^j\} \quad (13)$$

where  $\odot$  represents element-by-element multiplication. We can now design the penalty coefficients as a weighted least-squares solution to (13). Specifically, we choose  $\underline{r}^j$  such that

$$\hat{\underline{r}}^j = \arg \min_{\underline{r}^j \geq 0} \|\Phi^j \underline{r}^j - \underline{d}^j\|^2 \quad (14)$$

with

$$\Phi^j \triangleq V^j D[\underline{L}_0^j] \mathbf{F}\{\mathbf{B}^j\} \quad (15)$$

$$\underline{d}^j \triangleq V^j D[1 - \underline{L}_0^j] \mathbf{F}\{\mathbf{A}' \mathbf{D} \mathbf{A} \underline{c}^j\}. \quad (16)$$

(For matrices,  $\mathbf{F}\{\cdot\}$  operates on each column.) The matrix  $V^j$  is a (possibly shift-variant) least-squares weighting,<sup>9</sup> where  $[V^j]' V^j$  is a symmetric positive definite  $p \times p$  matrix.

After one chooses a desired frequency response  $\underline{L}_0$ , or equivalently a desired impulse response  $\underline{l}_0$ , one can use the NNLS (nonnegative least-squares) algorithm in [18] to perform the minimization (14) for each pixel position  $j$  to obtain the coefficients  $\{\hat{\underline{r}}^j\}_{j=1}^p$ . This provides nearly uniform resolution properties matching a specified response.

One possible practical inconvenience of the proposed method (14) is that  $\mathbf{R}$  must be recalculated for every desired response  $\underline{l}_0$ . We describe a method in the Appendix that yields a convenient class of penalty matrices  $\mathbf{R}$  that span a range of spatial resolutions for a specific class of desired impulse responses.

#### IV. PRACTICAL IMPLEMENTATION

While the penalty design method (14) gives a simple form for the calculation of  $\hat{\underline{r}}^j$ , in the form described above it still requires more computation than we would like for routine use. In this section, we outline a computationally efficient method for closely approximating the parameterized penalty with coefficients given by (14).

Consider each of the terms in (15) and (16). Determination of  $\mathbf{F}\{\mathbf{B}^j\}$  requires a single calculation of the 2-D-FFT (fast Fourier transform) of each of the  $n_w$  2-D basis functions. (Different  $j$  only shift the bases. One could incorporate these shifts with relatively little computational overhead by multiplying  $\mathbf{F}\{\mathbf{B}^{j_0}\}$  by appropriate complex exponentials.) For a shift-invariant  $\underline{l}_0$ , the remaining portion of (15) may also be computed once with simple matrix multiplications. Therefore, we can precompute  $\Phi^{j_0}$ , and determine  $\Phi^j$  from  $\Phi^{j_0}$  by complex phase shifts. (This step is eliminated below.) For  $\underline{d}^j$ , the matrix multiplications including  $V^j$  and  $D[1 - \underline{L}_0]$  may be precalculated as well. However, for a direct implementation of (14), one would have to compute the 2-D-FFT of  $\mathbf{A}' \mathbf{D} \mathbf{A} \underline{c}^j$  for every pixel  $j$ , which would be computationally expensive.

Therefore, the key to a practical implementation of (14) is the efficient calculation of  $\mathbf{F}\{\mathbf{A}' \mathbf{D} \mathbf{A} \underline{c}^j\}$  for  $j = 1, \dots, p$ . In general this term would need to be calculated explicitly, which would be quite time consuming because of the size of  $\mathbf{A}$  and the number of pixels  $p$ . However, consider a system matrix that

has the factorable form  $\mathbf{A} = D[c_i] \mathbf{G}$ , where  $\mathbf{G}' \mathbf{G}$  is an approximately shift-invariant operator and represents the geometric system response. The diagonal matrix  $D[c_i]$  contains known ray-dependent effects such as detector efficiency and attenuation factors, where  $c_i$  is a multiplicative factor for the  $i$ th measurement,  $y_i$ . In this case, we may write

$$\mathbf{A}' \mathbf{D} \mathbf{A} = \mathbf{G}' \mathbf{W} \mathbf{G} \quad (17)$$

where for emission tomography  $\mathbf{W} = D[q_i] = D[c_i^2 / \bar{Y}_i]$ . Such a system model is appropriate for modeling an idealized PET system, where the geometric response is shift-invariant. This is a relatively good approximation for real PET systems near the center of the field of view.

One can show that in an idealized continuous system, if the continuous equivalent of  $\mathbf{W}$  is a *radially constant*<sup>10</sup> sinogram scaling operator, then the continuous equivalent of  $\mathbf{G}' \mathbf{W} \mathbf{G}$  can be expressed as a position-independent blurring operation [19]. This property should be approximately true in the discrete case. If  $\mathbf{W}$  were radially constant, we would only need one computation of  $\mathbf{F}\{\mathbf{G}' \mathbf{W} \mathbf{G} \underline{c}^j\}$ .

In practice, the elements of  $\mathbf{W}$  are not globally radially constant. However, since the projection of a single pixel forms a relatively narrow trace in sinogram space (only a few radial bins in width), we can approximate  $\mathbf{W}$  *locally* by a position-dependent radially constant matrix  $\mathbf{W}^j$ . This property is illustrated in Fig. 1.

Consider a single pixel in the image represented by the unit vector  $\underline{c}^j$ . The operation of  $\mathbf{G}$  on  $\underline{c}^j$  forms a relatively narrow sinusoidal trace in sinogram space. Such a projection is shown in Fig. 1(a). The backprojection of this sinusoidal trace produces the familiar  $1/r$  response centered at the given pixel. This image and an enlarged region about the pixel of interest are shown in Figs. 1(b) and (c). Recall from (17) that the effective sinogram weighting  $\mathbf{W}$  is  $c_i^2 / \bar{Y}_i$ . In reconstructions where  $\bar{Y}_i$  is unknown, we choose  $\mathbf{W} = D[q_i]$ , where  $q_i \triangleq \hat{c}_i^2 / \max\{y_i, t_c\}$ . The  $\hat{c}_i$  terms are estimates of the detector efficiencies and attenuation correction factors made from a normalization scan and transmission scan, and  $y_i$  are the measurements. The  $\max\{y_i, t_c\}$  is included to avoid inordinate ray weighting for low count  $y_i$ , by choosing  $t_c > 0$ .

The vector  $\underline{q} = [q_1, \dots, q_N]'$  represents a lexicographically reordered 2-D array of scaling values that is  $n_a$  angles by  $n_b$  radial bins in size, where  $N = n_a \cdot n_b$ . A typical  $\underline{q}$  is presented in Fig. 1(d). Fig. 1(e) shows the weighted sinogram  $\mathbf{W} \mathbf{G} \underline{c}^j$  for the single pixel's projection using this particular weighting.

Instead of using  $\underline{q}$ , we would like to approximate the weighting with a local radially constant version  $\tilde{q}^j$ . The associated diagonal weighting matrix is  $\mathbf{W}^j = D[\tilde{q}^j]$ . To choose  $\tilde{q}^j$ , consider the following. Let  $\mathbf{U} = \mathbf{G}' \mathbf{W} \mathbf{G}$ , where  $\mathbf{W} = D[\underline{q}]$ . Decompose the system matrix  $\mathbf{G}$  by rows into separate submatrices for each projection angle so that  $\mathbf{G} = [\mathbf{G}'_1, \dots, \mathbf{G}'_{n_a}]'$  with  $\mathbf{G}'_n \in \mathbb{R}^{n_b \times p}$ . Similarly, decompose the weighting vector into  $\underline{q} = [[q^1]', \dots, [q^{n_a}]']'$  with  $q^n \in \mathbb{R}^{n_b}$ .

<sup>9</sup>An equivalent weighted least-squares problem may be stated using a weighted norm  $\|x\|_{V^j} = \|V^j x\|$  in (14) and eliminating  $V^j$  from (15) and (16).

<sup>10</sup>A "radially constant"  $\mathbf{W}$  scales all of a sinogram's radial elements for a particular projection angle by the same scalar value. Such a  $\mathbf{W}$  would have the block scaled-identity form  $\mathbf{W} = D[w^{(1)} I, w^{(2)} I, \dots, w^{(n_a)} I]$ , where  $n_a$  is the number of angles.

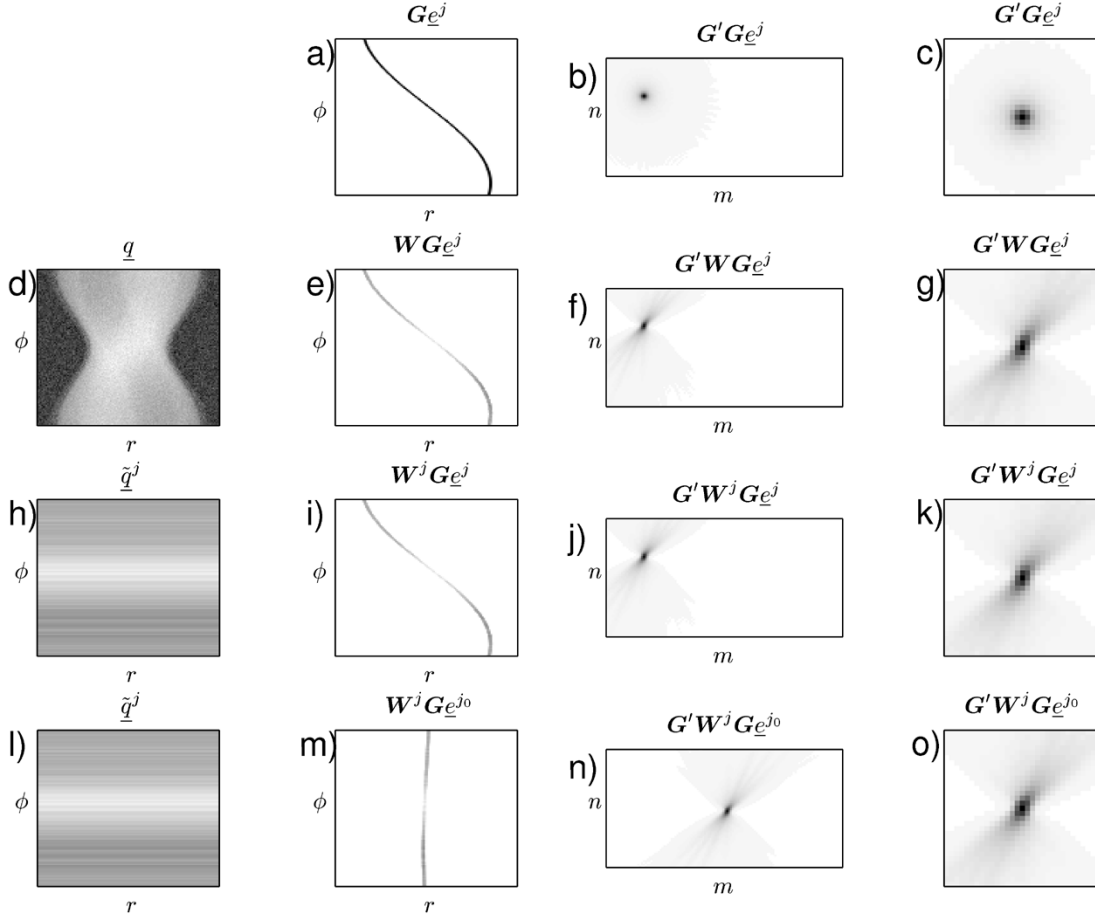


Fig. 1. Approximation of  $\mathbf{G}'\mathbf{W}\mathbf{G}$  with local radially constant weightings (see text). (a) shows the unweighted projection of a single pixel in the sinogram domain. (b) and (c) show the backprojection of this unweighted sinogram where (c) is an enlarged portion of (b). Variables  $n$  and  $m$  index image coordinates, and variables  $\phi$  and  $r$  index the sinogram measurements. (d)–(g) show the weighted projection (e) and backprojection (f) and (g) for a typical weighting (d). (h)–(k) show the weighted projection-backprojection using an approximate radially constant weighting (h). (l)–(o) show the weighted projection-backprojection for a fixed central pixel  $j_0$ . Note that (g), (k), and (o) are nearly identical. (d), (e), (h), (i), (l), and (m) have a logarithmic color scale.

Due to the  $1/r$  response in tomography,  $\mathbf{U}$  is concentrated about its diagonal. Therefore,

$$\mathbf{U} = \sum_{n=1}^{n_a} \mathbf{G}'_n D[\underline{q}^n] \mathbf{G}_n \approx \sum_{n=1}^{n_a} D[\underline{k}^n] \mathbf{G}'_n \mathbf{G}_n D[\underline{k}^n] \triangleq \tilde{\mathbf{U}} \quad (18)$$

where the  $j$ th element of  $\underline{k}^n$  is

$$[\underline{k}^n]_j \triangleq \sqrt{\frac{\sum_{i=1}^{n_b} [\mathbf{G}_n]_{ij}^2 [\underline{q}^n]_i}{\sum_{i=1}^{n_b} [\mathbf{G}_n]_{ij}^2}}, \quad n = 1, \dots, n_a. \quad (19)$$

This is the unique choice of  $\underline{k}^n$  that makes  $\mathbf{U} \approx \tilde{\mathbf{U}}$  with equality along diagonals for each of the  $n_a$  terms in the summations in (18) (i.e., the diagonal elements of  $\mathbf{G}'_n D[\underline{q}^n] \mathbf{G}_n$  and  $D[\underline{k}^n] \mathbf{G}'_n \mathbf{G}_n D[\underline{k}^n]$  are identical for each  $n = 1, \dots, n_a$ ). The approximation would be exact if the  $q_i$ 's were all equal. However, since the local impulse response at pixel  $j$  relies predominantly on the  $q_i$ 's that intersect pixel  $j$ ,  $\mathbf{U}$  and  $\tilde{\mathbf{U}}$  will be nearly equal. This approximation is reasonable even for very nonuniform  $\underline{q}$  since  $[\underline{k}^n]_j$  tend to vary slowly as a function of

$j$  because of the implicit smoothing in (19). Similarly, since  $\mathbf{G}'\mathbf{G}\underline{e}^j$  concentrates around  $\underline{e}^j$  (cf. [4]),

$$\mathbf{U}\underline{e}^j \approx \sum_{n=1}^{n_a} [\underline{k}^n]_j^2 \mathbf{G}'_n \mathbf{G}_n \underline{e}^j.$$

We choose  $\tilde{\underline{q}}^j = [[\underline{k}^1]_j^2 \mathbf{1}', \dots, [\underline{k}^{n_a}]_j^2 \mathbf{1}']'$  to form a radially constant, position-dependent weighting  $\tilde{\underline{q}}^j$ , where  $\mathbf{1}$  is a column vector of ones of length  $n_b$ .

Fig. 1(h) shows the radially constant weighting  $\tilde{\underline{q}}^j$  using this technique on the weights in Fig. 1(d) for pixel  $j$ . When applied to the projection of  $\underline{e}^j$ , the result is very close to the weighting using  $\underline{q}$ . The close agreement between the  $\mathbf{W}$  and  $\mathbf{W}^j$  weightings can be seen by comparing weighted sinograms in Fig. 1(e) and (i). Similarly, the agreement is very close for the backprojected weighted sinograms,  $\mathbf{G}'\mathbf{W}\mathbf{G}\underline{e}^j$  and  $\mathbf{G}'\mathbf{W}^j\mathbf{G}\underline{e}^j$ , shown in Fig. 1(f) and (j), respectively. Zoomed versions of the backprojected weighted sinogram for regular and local radially constant weightings are shown in Fig. 1(g) and (k). Note the close agreement in the image domain as well as the sinogram domain.

Since  $\mathbf{G}'\mathbf{W}\mathbf{G}$  is an approximately shift-invariant operator for radially constant  $\mathbf{W}$ ,  $\mathbf{G}'\mathbf{W}\mathbf{G}\underline{e}^j$  approximately equals a shifted  $\mathbf{G}'\mathbf{W}^j\mathbf{G}\underline{e}^{j_0}$  for an arbitrary fixed pixel  $j_0$  (e.g., the center pixel

in the image). Plots of  $\mathbf{G}'\mathbf{W}^j\mathbf{G}\underline{e}^{j_0}$  for  $j_0$  equal to the center pixel are shown in Fig. 1(n) and (o). Note the close agreement between  $\mathbf{G}'\mathbf{W}^j\mathbf{G}\underline{e}^{j_0}$  and  $\mathbf{G}'\mathbf{W}\mathbf{G}\underline{e}^j$ . Therefore, we need only to calculate  $\mathbf{G}\underline{e}^{j_0}$  once, rather than for every  $j$ .

In terms of  $\Phi^j$  in (14), under the radially constant  $\mathbf{W}$  approximation and a shift-invariant weighting  $V^{j_0}$ , we need only  $\Phi^{j_0}$  and may replace (14) with

$$\begin{aligned} \hat{r}^j &= \arg \min_{\underline{r}^j \geq 0} \|\Phi^{j_0} \underline{r}^j - \tilde{d}^j\|^2 \\ \tilde{d}^j &\triangleq V^{j_0} D[1 - \underline{L}_0] \mathbf{F}\{\mathbf{G}'\mathbf{W}^j\mathbf{G}\underline{e}^{j_0}\} \end{aligned} \quad (20)$$

where  $\tilde{d}^j$  is a ‘‘centered’’ version of  $d^j$ . This step eliminates the need for complex phase shifts. Nevertheless, direct implementation of (20) would still require  $p$  2-D-FFT’s.

To simplify further, define  $\underline{w}^j = [[\kappa_j^1]^2, \dots, [\kappa_j^{n_a}]^2]'$ . The vector  $\underline{w}^j$  contains all of the distinct angular weighting values in the radially constant  $\mathbf{W}^j$ . Since  $\tilde{d}^j$  is linear in the elements of  $\mathbf{W}^j$ , we may write  $\tilde{d}^j(\underline{w}^j) \equiv \mathbf{M}\underline{w}^j$ , where  $\mathbf{M}$  is a  $p \times n_a$  matrix. We find  $\mathbf{M}$  by superposition as follows.

Define  $\mathbf{W}^{\delta_n}$  to be the weighting matrix with radially constant values having unit values at angle  $n$ , and zero otherwise. (i.e.,  $\mathbf{W}^{\delta_n} = D[[\underline{0}', \dots, \underline{0}', \underline{1}', \underline{0}', \dots, \underline{0}']]$ , where the  $\underline{1}$  vector appears in the  $n$ th block and  $\underline{1}$  and  $\underline{0}$  are column vectors of length  $n_b$ .) For each angle  $n$  define

$$\underline{m}_n \triangleq V^{j_0} D[1 - \underline{L}_0] \mathbf{F}\{\mathbf{G}'\mathbf{W}^{\delta_n}\mathbf{G}\underline{e}^{j_0}\}, \quad n = 1, \dots, n_a. \quad (21)$$

Then by superposition  $\mathbf{M} = [\underline{m}_1, \dots, \underline{m}_{n_a}]$ .

For the unconstrained case, (20) has the closed form linear solution  $\hat{r}^j = [[\Phi^{j_0}]' \Phi^{j_0}]^{-1} [\Phi^{j_0}]' \tilde{d}^j$ . Let

$$\mathbf{T} = [[\Phi^{j_0}]' \Phi^{j_0}]^{-1} [\Phi^{j_0}]' \mathbf{M}$$

be the combined  $n_w \times n_a$  matrix operator. Therefore we may determine unconstrained solution of (20) by  $\hat{r}^j = \mathbf{T}\underline{w}^j$ .

However, for the matrix  $\mathbf{R}$  to be nonnegative definite, we need to solve the constrained optimization problem (20). It is straightforward to modify the NNLS algorithm of [18] using  $\Phi^{j_0}$  and  $\mathbf{M}$  to provide the constrained solution.

For simplicity in our implementation, we have used the sub-optimal greedy approach presented in Table I, which yields non-negative  $\hat{r}^j$  and nearly the same results as NNLS but with a slight computational speedup and simpler implementation. This procedure takes one step for each negative element in  $\hat{r}^j$  and will complete in at most  $n_w$  steps. For small  $n_w$ , one could precompute the  $2^{n_w} - 1$  possible reduced  $\mathbf{H}^{-1}$  matrices for further speedup.

As described in the beginning of this section, direct implementation of the design given by (14)–(16) requires  $p$  backprojections,  $n_w + p$  2-D FFT’s, and  $p$  applications of the NNLS algorithm. Using the simplifications described in this section, we perform the one-time precomputation of  $\Phi^{j_0}$  and  $\mathbf{M}$  for a given system geometry using  $n_a$  backprojections and  $n_w + n_a$  2-D FFT’s. The coefficients  $\hat{r}^j$  may be determined with  $p$  calculations of (19), which is on the order of one backprojection, and  $p$  applications of the algorithm in Table I (or the NNLS algorithm).

TABLE I  
ROUTINE USED TO CONSTRAIN KERNEL COEFFICIENTS  $\hat{r}^j$

---

Set $\mathbf{H} = [\Phi^{j_0}]' \Phi^{j_0}$ and $\mathbf{C} = [\Phi^{j_0}]' \mathbf{M}$ .
Let $\hat{r}^j = \mathbf{T}\underline{w}^j$ .
while $\hat{r}^j$ contains negative values,
Let $i$ equal the index of the minimum value of $\hat{r}^j$ .
Remove the $i$ th row and $i$ th column of $\mathbf{H}$ .
Remove the $i$ th row from $\mathbf{C}$ .
Set the $i$ th element of $\hat{r}^j$ to zero.
Find the remaining elements of $\hat{r}^j$ by $\mathbf{H}^{-1} \mathbf{C}\underline{w}^j$ .
end

---

Computing (19) requires  $4n_G + 2p$  floating point operations (flops), where  $n_G = n_a n_b p \alpha$  and  $\alpha$  is the fraction of nonzero elements of  $\mathbf{G}$  (or  $\mathbf{A}$ ). Assuming the  $\mathbf{H}^{-1} \mathbf{C}$  are precomputed, the algorithm in Table I requires at most  $2n_w^2 n_a$  flops per pixel. Therefore, calculation of all  $\hat{r}^j$  coefficients using the method summarized in Table I requires at most  $4n_G + 2p + pn_a n_w^2$  flops. Since the precalculation of  $\mathbf{M}$  using (21) requires approximately  $n_G + O(n_a p \log p)$  flops and  $O(n_w p \log p)$  flops for  $\Phi^{j_0}$ , the entire precalculation is  $n_G + O((n_a + n_w) p \log p)$ . This precalculation need only be performed once for a specific system geometry and choice of  $\underline{l}_0$ .

In contrast, computation of the design given by (14)–(16) is dominated by the calculation of  $d^j$  in (16). This term requires  $2n_G n_a n_b \alpha + O(p \log p)$  flops for a single pixel location  $j$ . Therefore, even without calculating (14) and (15) we require at least  $2n_G^2 + O(p^2 \log p)$  flops to evaluate  $\hat{r}^j$  for all  $j$ . Clearly, much of the computational advantage of the proposed method is due to the order reduction of  $n_G^2$  to  $n_G$ .

For 2-D reconstructions performed in the following section, 30 iterations of the SAGE algorithm [20] on a 266 MHz Pentium II processor took 18.5 s for the conventional space-invariant first-order penalty given by the kernel in (5), and 20.1 s for the proposed penalty with precomputed  $\Phi^{j_0}$  and  $\mathbf{M}$ . The precalculation of  $\Phi^{j_0}$  and  $\mathbf{M}$  took 23.1 s.<sup>11</sup> Thus, the method is very practical. (We performed the reconstructions using the ASPIRE iterative reconstruction libraries[21].)

## V. SIMULATION RESULTS

### A. Resolution Uniformity

This section provides simulation results comparing the relative resolution uniformity of different regularization schemes. Fig. 2 shows the  $128 \times 64$  emission image  $\underline{\lambda}$  (with 3-mm square pixels) used for the investigation as in [4]. The image has a warm background ellipse, a cold left disk, and hot right disk with relative emission intensities of 2, 1, and 3, and attenuation coefficients of 0.0096, 0.003, and 0.013/mm, respectively. The PET system model included projection data with 128 radial bins and 110 angles uniformly spread over  $180^\circ$  with 6-mm-wide strip integrals (3-mm center-to-center spacing), and detector efficiencies with a pseudorandom log-normal variance with  $\sigma = 0.3$  to model detector efficiency effects.

<sup>11</sup>The constrained nonlinear least-squares penalty given in (11) took about 2 h.

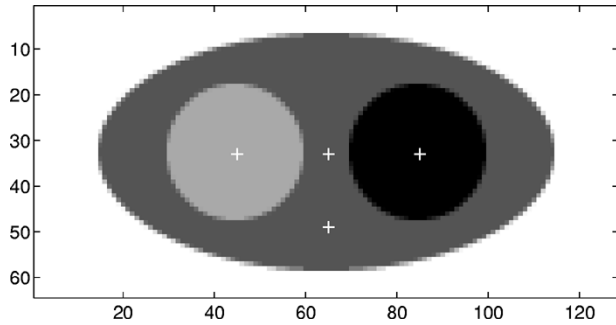


Fig. 2. Digital phantom used for investigation of resolution properties of different regularizations, with four pixels of interest marked.

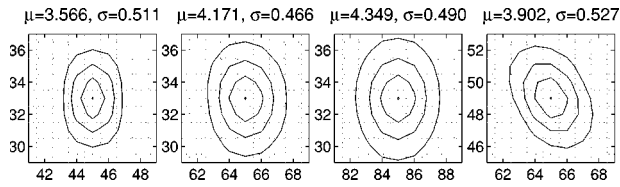


Fig. 3. Local PSF's for space-invariant penalty.

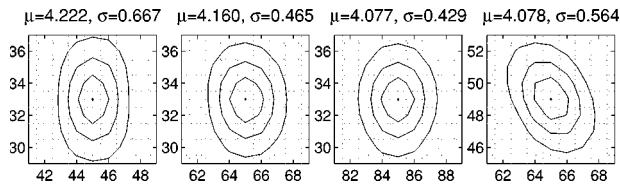


Fig. 4. Local PSF's for certainty-based penalty.

To compare the relative spatial uniformity of these regularization methods we used (3) to calculate local point spread functions (PSF's). We approximated the solution of (3) using 100 iterations of a coordinate ascent algorithm initialized with a Fourier approximation of the target response given in (23). Since the responses are space-variant, we investigated four different locations in the object, represented by the white "+" marks in Fig. 2. We systematically examined numerous additional spatial locations which yielded similar conclusions (i.e., these are representative results).

Results of this impulse response survey are presented in Figs. 3–8. For each penalty, PSF contours at 25%, 50%, 75%, and 99% of peak value are shown. These contours were generated using the `contour` command in Matlab 5.3. The pixel boundaries are represented by the dotted grid in each plot. Above each set of contours are estimates of the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of the FWHM resolution in pixels, which quantify the mean resolution and radial variation at that location.

All reconstruction methods and penalties were designed with a target resolution of 4.0 pixels (1.2 cm) FWHM resolution. (The relationship between global FWHM resolution and  $\beta$ , and how to calculate  $\beta$  is discussed in [5].) For the estimation of (3) for penalized-likelihood methods, where  $q_i$  in (17) must be computed, we used the noiseless measurements  $\bar{Y}_i$  with a mean of 1 million counts and we chose  $q_i = c_i^2 / \max\{\bar{Y}_i, t_c\}$  with  $t_c = 10$ .

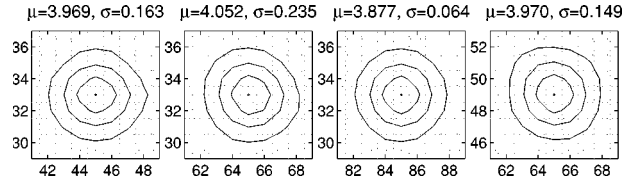


Fig. 5. Local PSF's for CNLLS penalty.

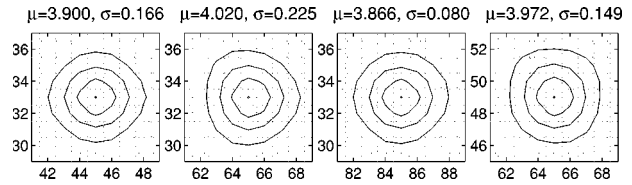


Fig. 6. Local PSF's for proposed penalty.

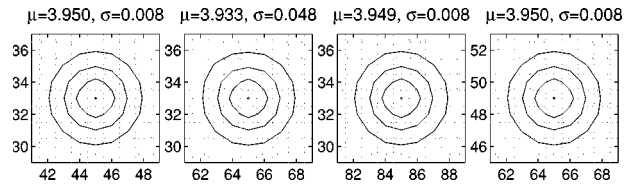


Fig. 7. Local PSF's for FBP.

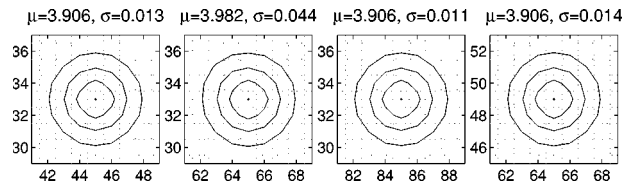


Fig. 8. Local PSF's for penalized unweighted least squares.

PSF's for penalized-likelihood with the conventional space-invariant penalty are shown in Fig. 3. This penalty has the first-order kernel given in (5). These local PSF's are highly asymmetric and space-variant, blurring more in high-count regions (85, 33) than in low-count regions. The certainty-based penalty of [4] shown in Fig. 4 provides some improvement by making the mean FWHM close to 4.0 pixels. However, the responses are still quite asymmetric.

PSF's for the constrained nonlinear least-squares (CNLLS) penalty given by (11) are shown in Fig. 5. For this design, four basis functions and second-order neighborhood were used. We used a target response  $\bar{l}_0$  equal to the response of a penalized unweighted least-squares (PULS) estimator with a conventional first-order penalty [see (23)]. The contours for these PSF's are nearly radially symmetric and near the 4.0 pixel FWHM target resolution. The PSF's of the proposed regularization method (20) are presented in Fig. 6. We used the same basis set and target response as the CNLLS penalty design. This penalty was designed using the computational simplifications in Section IV and the simplifications presented in the Appendix. The PSF contours are also quite symmetric and the average FWHM resolution is within 5% of the target resolution of 4.0 pixels.

In addition to the penalized-likelihood methods, we present results for filtered backprojection (FBP) and a PULS estimator with a conventional first-order shift-invariant penalty. Both



TABLE II  
SUMMARY OF MEAN ABSOLUTE RADIAL DEVIATION RESULTS

Estimator	Pixel Set			
	A	B	C	D
PLE: Conventional	0.29	0.26	0.27	0.23
PLE: Certainty-Based	0.33	0.26	0.31	0.20
PLE: CNLLS Penalty	0.19	0.10	0.06	0.08
PLE: Proposed Penalty	0.24	0.12	0.07	0.09
Filtered Backprojection	0.02	0.02	0.02	0.02
PULS	0.05	0.05	0.05	0.05

(A) All Phantom Pixels (B) Interior Pixels  
(C) Cold Disc Pixels (D) Hot Disc Pixels

of these methods should produce shift-invariant and isotropic smoothing properties since we are considering a tomographic systems whose intrinsic response is shift-invariant. [The impulse response for a PULS estimator is given by (23) in the Appendix.] In a real PET (SPECT) system, FBP will have nonuniform resolution due to the detector response. We chose a constrained least-squares filter for FBP, which essentially matches the smoothing properties of FBP and PULS (see [5]).

The local PSF's for FBP are shown in Fig. 7, and for PULS in Fig. 8. These responses are nearly perfectly symmetric. Recall, our proposed penalty is designed with a target PSF given by the PULS response (23). Note the similarity between these method's responses as shown in Figs. 6 and 8.

As a quantitative assessment of the resolution uniformity, we calculated the mean absolute radial deviation of the 50% contour from the 2.0 pixel half-maximum target radius. Then we calculated the average value of this deviation over a set of sample locations within the phantom. We performed these calculations over four pixel sets: Set A consists of all pixels within the digital phantom object; Set B contains roughly 80% of the interior pixels of the phantom excluding the outer edge pixels; Set C contains all pixels in the cold disc; and Set D contains all pixels in the hot disc. These results are summarized in Table II. All values are in pixels. The certainty-based penalty and the conventional penalty have the greatest deviation, while the CNLLS penalty and the proposed penalty are more uniform. The improvement in uniformity with these penalties is more dramatic for the interior pixels (Sets B, C, and D), indicating that these penalties provide less uniform resolution at the edges of the phantom. FBP and PULS have the lowest deviations with no variation between sets.

The calculated  $\hat{r}^j$  coefficients for the CNLLS penalty and the proposed method are presented in Fig. 9. The coefficient values are presented as four images (since we used four basis functions) for both methods, separated by dotted lines. Each image pixel corresponds to the coefficient of a given basis function at that pixel location. The scale is logarithmic, except for the value zero, which is represented in white. The largest discrepancies between the coefficients appear at the edge or outside the object in the digital phantom. Additionally, we see that the nonnegativity constraints are fairly active (as represented by the white regions). Future designs may be able to obtain increased resolution uniformity by relaxing the nonnegativity constraints on  $\hat{r}^j$ .

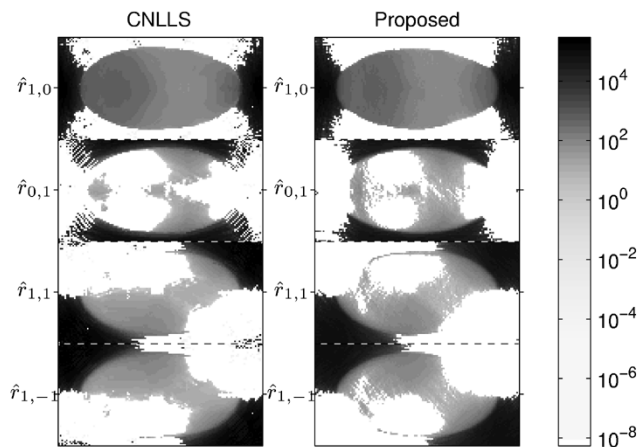


Fig. 9. Comparison of calculated  $\hat{r}^j$  values for the CNLLS penalty and the proposed penalty. Note the logarithmic color scale. White regions indicate a value of zero.

The CNLLS penalty and the proposed penalty yield similar  $\hat{r}^j$  coefficients and produce similar local impulse responses. Hence, we conclude that the computational simplifications proposed in Section IV do not change the calculated values of  $\hat{r}^j$  significantly and that such simplifications are appropriate for providing a computationally efficient algorithm for calculating our proposed penalty for uniform resolution properties.

### B. Noise Properties

The results presented above describe the resolution properties of the estimators. As in [4], we also investigated the noise properties. To form sample standard deviation images, we simulated 400 noisy measurement realizations for the digital phantom in Fig. 2. The PET model included 10% random coincidences and averaged 1 million counts per realization.

We reconstructed each of these 400 realizations using 30 iterations of the SAGE algorithm [20] with the same regularization methods used above in the resolution properties investigation. For all of the statistical methods except the CNLLS penalty, we use the measurements  $y_i$  for calculation of  $\mathbf{R}$ . Because of the extensive computation time associated with calculation of the CNLLS penalty, the noiseless,  $\bar{Y}_i$  were used, i.e., the same penalty based on the noiseless measurements was used for all realizations.

The results of this noise investigation are presented in Fig. 10. The sample standard deviation images are shown on the LHS of the figure. Horizontal and vertical profiles of these images are shown in the remaining plots. The horizontal profile is taken through the image center and the vertical profile is taken through the center of the cold disk. These profiles are represented by dotted lines in the images. Pixel standard deviations in these plots are expressed in terms of a percentage of the background ellipse intensity. If one included error bars on these plots, the error bars would be smaller than the plot markers. Therefore we have eliminated the error bars for clarity. For conventional regularization, the standard deviation estimate is nearly uniform. FBP and PULS generally have the highest standard

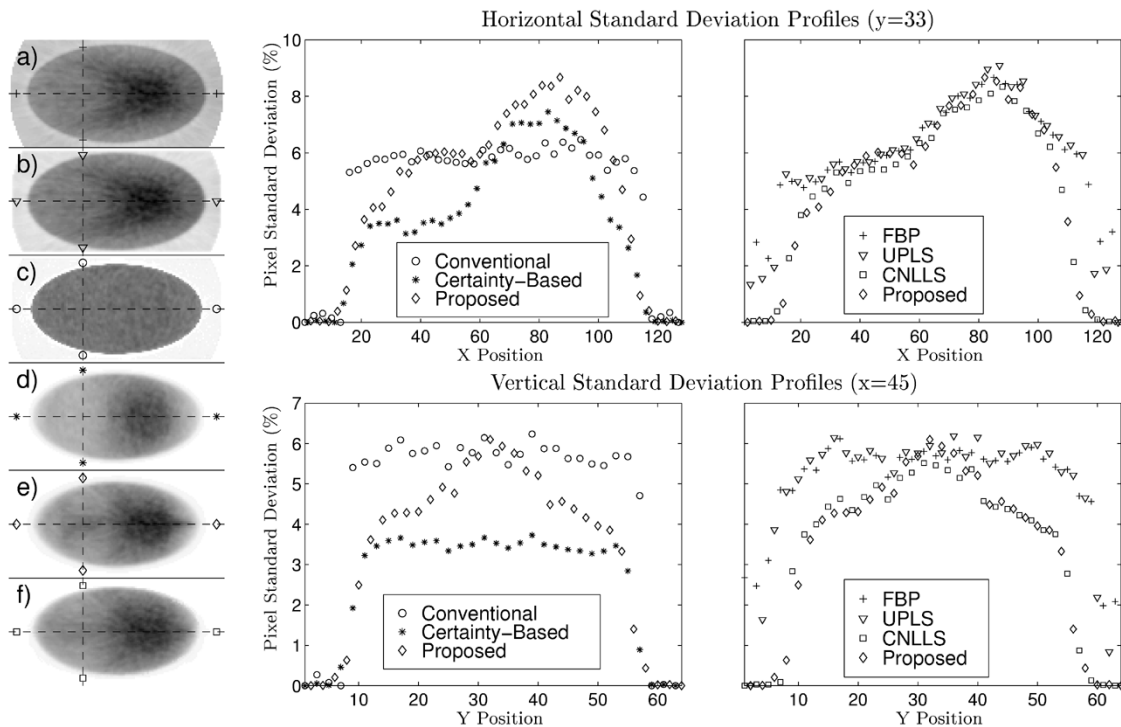


Fig. 10. Sample standard deviation images and profiles. (a) Filtered backprojection (+). (b) Penalized unweighted least-squares PULS ( $\nabla$ ). (c) PLE with conventional regularization (o). (d) PLE with certainty-based penalty (\*). (e) PLE with proposed penalty ( $\diamond$ ). (f) PLE with CNLLS penalty ( $\square$ ).

deviation and the certainty-based penalty have the lowest standard deviation. Not only do FBP and PULS share similar resolution properties, but also similar noise properties. The close agreement in standard deviation between the proposed method and the CNLLS penalty further justifies our computationally efficient design technique.

At first glance, it appears that uniform resolution properties come at the price of a variance increase as compared with the certainty-based penalty. However, the certainty-based penalty and the proposed penalty have *different* resolution properties. The certainty-based reconstruction often has a greater maximum diameter of the local PSF's (compare Figs. 4 and 6). This causes increased smoothing, yielding a reconstruction with lower variance.

We would like to produce a resolution-noise curve comparing the relative performance of these two methods over a range of target resolutions, but this is difficult because they have different resolution properties. Using the angularly averaged FWHM as a resolution metric (cf. [4]) unfairly handicaps estimators with isotropic resolution properties. Estimators with anisotropic responses can reduce noise by smoothing "optimally" in each direction while maintaining the same average FWHM as an estimator with isotropic responses. Rather than creating resolution-noise curves where each point on the curve corresponds to a single resolution value and a single standard deviation, we created banded "curves" as follows. For the ordinate, we used the sample standard deviations of pixel values in images reconstructed from 400 noisy sinogram realizations, for each of several target spatial resolutions. For each target resolution we also computed the local PSF and found the smallest and largest diameters of its half-maximum contour. We specified the abscissae in

the banded plot as the interval between the minimum and maximum diameters. For each pixel location and target resolution, these plots describe the (single) pixel standard deviation value as well as the *range* of spatial resolutions spanned by the local PSF. A method with isotropic resolution properties would appear as a single line in such plots, whereas a method with a highly anisotropic response appears as a thick band. These tradeoff curves were calculated for the four pixels positions shown in Fig. 2.

Curves for the conventional and proposed penalties are shown in Fig. 11. The lighter band with "+" symbols on the border represents the resolution/noise tradeoff curve for the proposed regularization, while the darker band with "o" symbols on the border is the curve for reconstruction with conventional regularization. (The light band partially obscures the dark band, however the borders are marked by symbols and lines so that the degree of overlap is visible.)

We also produced a banded resolution/noise tradeoff plot using the certainty-based regularization of [4]. Since the certainty-based technique produced a curve nearly identical to the conventional regularization, we have omitted the plot. Similar behavior was observed in [4] using a mean FWHM resolution criterion. Essentially this means each pixel simply moves up or down its resolution/noise curve to the specified resolution. This is another indication that the certainty-based method does not yield isotropic resolution properties. While the average FWHM resolution may be improved, the PSF's are still anisotropic yielding a wide resolution band in our banded resolution/noise tradeoff curves.

In Fig. 11 the banded curves for the proposed penalty span a small resolution range (i.e., the curve is thin horizontally), in-

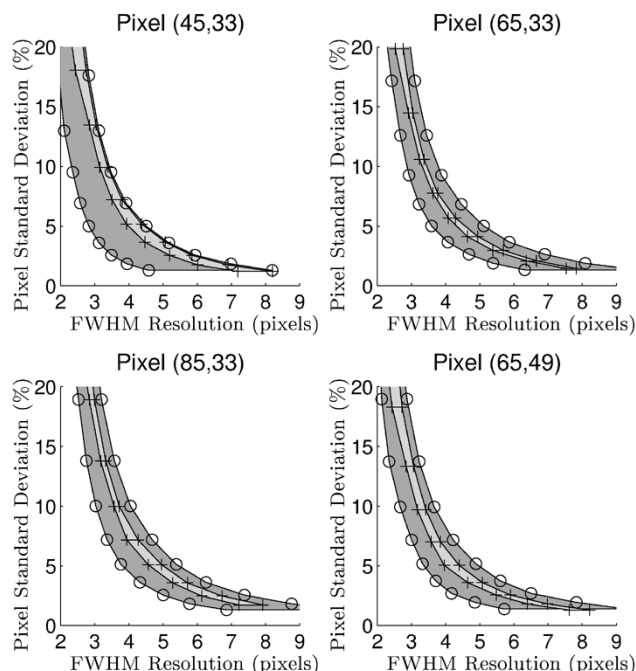


Fig. 11. Resolution/noise tradeoff for penalized-likelihood emission image reconstruction with conventional (o/dark) and proposed penalties (+/light).

dicating isotropic smoothing properties relative to the conventional penalty. If our design were ideal, minimum and maximum FWHM resolution would be identical and we would have a line instead of a band. Note that the proposed penalty band lies inside the conventional penalty band. If the proposed penalty band laid above the conventional penalty band over the same resolution interval, then the proposed penalty would arguably have worse noise properties. The proposed penalty band generally lies in the center of the conventional penalty band. However, this is not the case for the pixel (45, 33) in the cold disk. Note that the PSF for the conventional penalty at this pixel is especially asymmetric (see Fig. 3) having the largest difference between the min and max FWHM resolutions. If this PSF yields an “optimal” kind of smoothing (with its predominantly vertical orientation), it is logical that an isotropic response would decrease the variance little with additional horizontal smoothing (note that max resolution for the conventional PSF is very close to the 4.0 target). Using this rationale the proposed penalty bands for the other pixel locations lie roughly in the middle of the conventional penalty’s band since the PSF’s for these points are less asymmetric (with the max resolution greater than 4.0 and min less than 4.0 pixels). The isotropic response reduces the max resolution and increases the min resolution as compared with the conventional PSF. The “optimal” smoothing of the conventional PSF is arguably not so directionally dependent in this case and an isotropic response can provide roughly the same variance. While these two methods have different resolution properties, it appears that our penalty design has not adversely affected the noise properties of the estimator.

It is difficult to *globally* compare the proposed penalty with the conventional and certainty-based methods for an entire image reconstruction because they possess different resolution

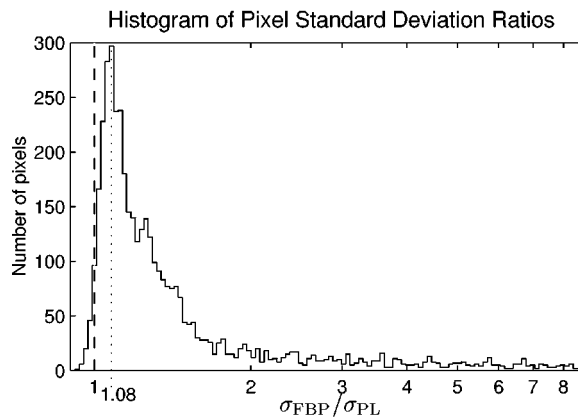


Fig. 12. Histogram showing the distribution of the ratio of the pixel standard deviation using filtered backprojection ( $\sigma_{\text{FBP}}$ ) to the pixel standard deviation using a PLE with the proposed regularization ( $\sigma_{\text{PL}}$ ).

properties for every pixel. On the other hand, FBP and the proposed penalty both yield nearly the same PSF’s, so a comparison seems more appropriate. Since these methods have nearly the same resolution properties, we can fairly identify which provides better global noise properties. Note, particularly in the vertical profile in Fig. 10, that reconstructions based on the proposed penalty have lower variance than FBP.

There are a few points in Fig. 10 where the standard deviation estimate is slightly greater for the proposed penalty. To illustrate the relative global noise properties of FBP and the PLE with the proposed regularization, we generated a histogram of the relative variance. For each pixel in the object, we calculated the ratio of the sample standard deviation at that pixel using filtered backprojection ( $\sigma_{\text{FBP}}$ ) to the sample standard deviation at that pixel using the PLE with the proposed regularization ( $\sigma_{\text{PL}}$ ). For pixels where  $\sigma_{\text{FBP}}/\sigma_{\text{PL}}$  is greater than one, filtered backprojection has higher standard deviation. This histogram is shown in Fig. 12. The vertical dashed line indicates the position where this ratio equals one. For nearly every pixel the PLE with the proposed regularization produces lower variance estimates and, for those pixels that have higher variances the difference is only slight. More than 50% of the pixels have over a 20% reduction in reconstructed pixel standard deviation.

In addition to the variance investigation, we present a correlation investigation. By specifying the desired resolution properties of PULS, do we also inherit the correlation properties? To address this issue we have included a set of typical correlation maps in Fig. 13 for FBP, PULS, and the PLE’s with conventional, certainty-based, and proposed penalties. These maps represent the absolute value of the correlation between each pixel and pixel (65, 49). FBP and PULS have nearly identical correlation maps (particularly inside the object). The PLE’s with conventional and certainty-based penalties have similar correlation maps, but are noticeably different due to the different resolutions. The proposed method is shown in Fig. 13(e). The structure of the correlation immediately surrounding (65, 49) is quite similar to FBP and PULS, having lost the nearly isotropic effect of the other PLE’s. This behavior is somewhat counterintuitive since PLE’s usually have much narrower correlation sidelobes

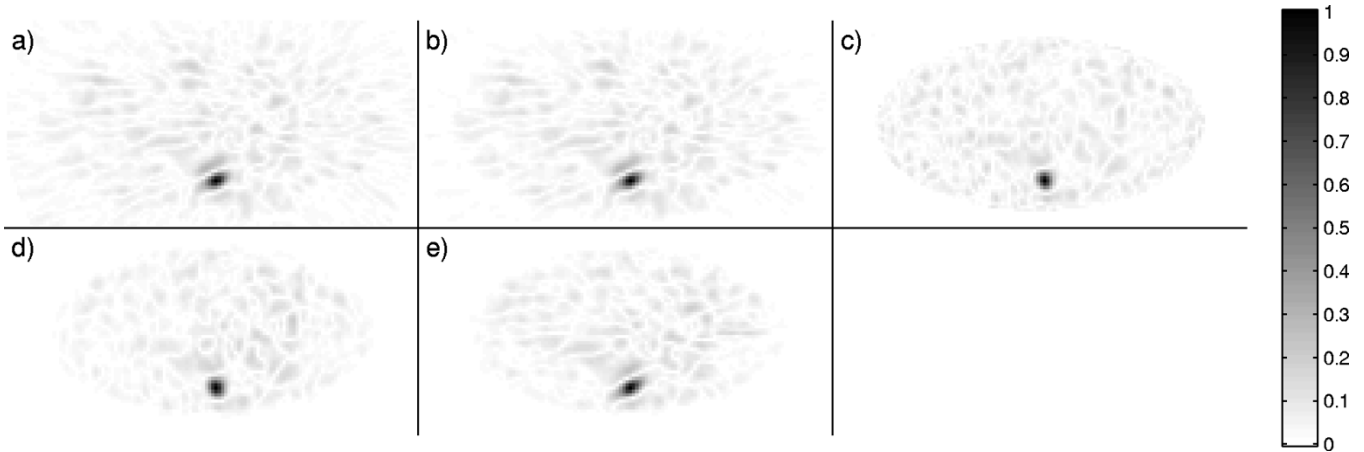


Fig. 13. Sample absolute correlation maps shown for pixel (65, 49). (a) Filtered backprojection (FBP). (b) Penalized unweighted least-squares (PULS). (c) PLE with conventional regularization. (d) PLE with certainty-based penalty. (e) PLE with proposed penalty.

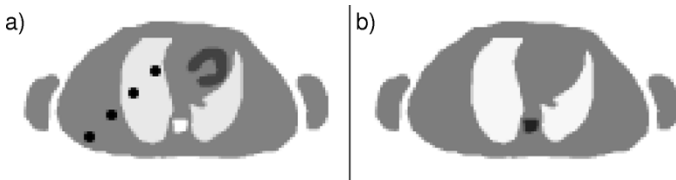


Fig. 14. Digital thorax phantom used for reconstruction of different regularizations: (a) is the emission image and (b) is the transmission image.

than FBP and PULS. Perhaps such correlation properties are inextricably tied to resolution uniformity. Further investigation is required.

## VI. THORAX PHANTOM RESULTS

In the previous section, we investigated our proposed regularization technique using quantitative measures of noise and resolution. In this section, we demonstrate the qualitative improvement using our proposed regularization technique through reconstruction of a noiseless thorax phantom. Both transmission and emission images for the digital phantom are shown in Fig. 14. The phantom is  $128 \times 64$  and has 0.42 cm square pixels. Relative emission intensities for the lungs, spine, and heart are 0.4, 0.0, and 3.0, respectively, with the background soft tissue having a relative intensity of 2. In addition, there are four round tumors with a relative intensity of 4. These simulated tumors are radially symmetric, neglecting discretization effects. In the transmission image, the attenuation coefficient of the lungs is 0.001/mm, the spine is 0.016/mm, and the remaining soft tissue is 0.0096/mm. The PET system model includes 160 radial bins and 192 angles space uniformly over  $180^\circ$ , with 3.375-mm strip integrals and 3.375-mm center-to-center spacing.

We reconstructed the noiseless emission measurements using FBP, PULS, and penalized weighted least-squares (PWLS) estimators with the conventional, certainty-based, and proposed penalties. All statistical methods enforced nonnegativity of the image and negatives in the image reconstructed via FBP were set to zero. All methods used a target FWHM resolution of 3.0 pixels (1.25 cm). For PULS and PWLS with conventional regularization, the penalties were chosen so that  $R_0$  corresponds to the shift-invariant first-order penalty with kernel as in (5).

The proposed penalty uses the  $\beta$ -independent design (26) with second-order bases, and the same target  $R_0$  as PULS.

The reconstructions using these methods are presented in Fig. 15. The FBP reconstruction in Fig. 15(a) has uniform resolution properties. This is evident from the uniformly smooth edges and radially symmetric tumors. Similarly, the PULS reconstruction in Fig. 15(b) shows the expected nearly identical results. (Recall the nearly identical PSF's of FBP and PULS in Section V.) The reconstruction using conventional regularization is shown in Fig. 15(c). There are distortions of the four round tumors (particularly in the lungs) in this reconstruction. The tumors are stretched vertically and appear elliptical. Another indication of resolution nonuniformity is evident at the outer boundaries of the arms. These boundaries are sharper than those in FBP and PULS. The reconstruction with certainty-based penalty in Fig. 15(d) shows some improvement. Most notably, the outer edges of the arms are smoothed in a more uniform fashion. However, the tumors are still smoothed preferentially in the vertical direction. Fig. 15(d) shows the reconstruction with our proposed penalty. The resolution uniformity appears much improved over the other PWLS methods. The tumors appear nearly radially symmetric and the edges appear much more uniformly smoothed.

## VII. DISCUSSION

Conventional space-invariant regularization methods for penalized-likelihood image reconstruction produce images with space-variant resolution properties. Although the certainty-based method of [4] attempts to provide more uniform resolution, as we have seen in our investigations, that method does not provide truly isotropic resolution properties.

We have presented a new method for designing a shift-variant penalty function that attempts to provide uniform resolution properties. The proposed method is motivated by a least-squares fitting of a parameterized local impulse response to a desired response  $l_0$ . We have developed fast methods to calculate this penalty for an idealized PET system whose geometric response is shift-invariant (while including ray-dependent attenuation and detector effects). This method yields nearly space-invariant and nearly symmetric local point spread functions at FWHM

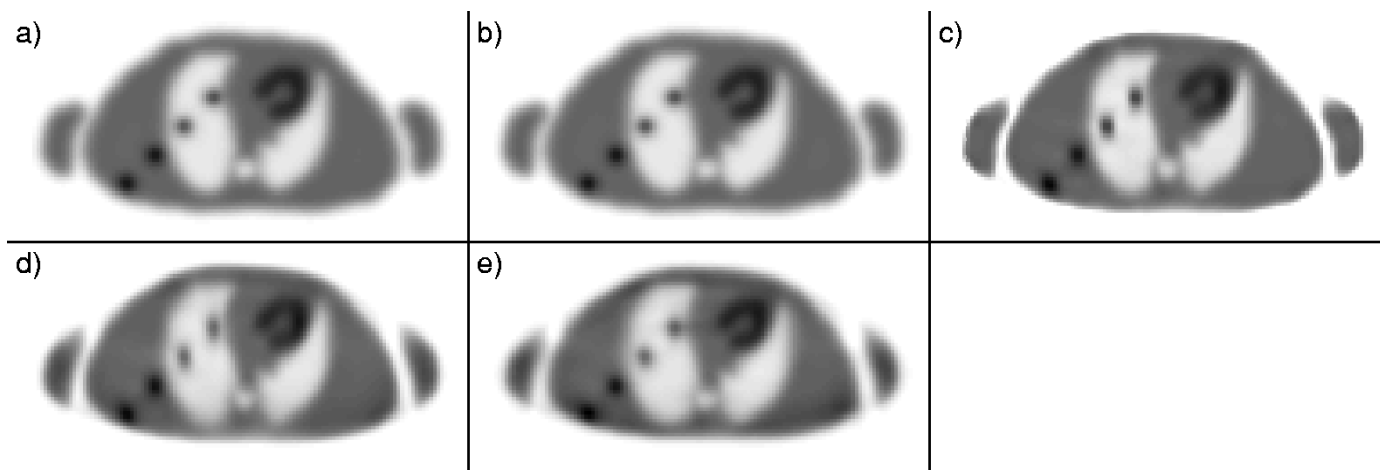


Fig. 15. Reconstruction of thorax phantom data using (a) filtered backprojection; (b) penalized unweighted least-squares; (c) PWLS with conventional regularization; (d) PWLS with certainty-based penalty; and (e) PWLS with the proposed regularization scheme.

resolutions very close to specified target resolutions. Additionally, we applied this novel regularization in the reconstruction of simulated thorax phantom data and demonstrated increased resolution uniformity.

Providing a regularization scheme that yields uniform resolution properties makes the selection of the regularization parameter ( $\beta$ ) more intuitive. One may simply specify the desired global resolution of the reconstructed image in  $\underline{l}_0$ . Additionally, creating nearly the same resolution properties in both the statistical (PLE) and traditional (FBP) reconstruction techniques provides a fair ground for comparing the noise properties of the two methods. As expected, we observed that by using a likelihood-based estimator and taking the noise model into account, one can reduce estimator variance.

While one may arguably desire space-variant resolution properties, one would most likely want to be able to control regional resolution properties, while maintaining radially symmetric responses. Our proposed methods can be modified to provide such control, allowing for predictable and intuitive specification of resolution properties in image reconstruction. Recall from (4), one could choose  $\underline{l}_0^j$  to be a shift-variant set of desired local impulse responses. Generally, such a choice of  $\underline{l}_0^j$ , will increase computation time. (However, one could use the technique presented in the Appendix and simply specify a shift-variant set of  $\beta$  values.)

As demonstrated in Fig. 6, the proposed second-order penalty still yields slight asymmetries in the point spread functions. Similarly, in Fig. 15(e), there are slight resolution nonuniformities evident in the reconstruction. In particular, the edges of the phantom have subtle smoothing differences between PWLS with the proposed penalty, and the FBP and PULS reconstructions in Fig. 15(a) and (b) (especially near the arms). The mean absolute deviation study summarized in Table II also indicates increased nonuniformity at the edges for the proposed design. There are many possible solutions that merit future investigation. The nonnegativity constraint on  $\underline{r}^j$  may be too strong a condition. (Recall this is a sufficient condition for nonnegativity of  $\mathbf{R}$ .) This constraint could be relaxed providing increase design freedom, yet still maintaining a nonnegative definite  $\mathbf{R}$  and a concave objective.

One may also achieve slightly better resolution uniformity at the expense of additional computation by using higher-order neighborhoods. Additionally, choices of  $\underline{l}_0$  other than (24) may require larger neighborhoods to obtain good fits to the desired response.

With additional improvements, the question of noise performance may arise. If the resolution properties are truly identical, does penalized-likelihood still outperform FBP? The variance improvements we have seen with our proposed method over FBP are marginal in some regions. As resolution properties are matched exactly will the advantages disappear? If so, does this hold for shift-variant systems as well?

The correlation images for our proposed penalty appear very similar to FBP and PULS. It appears that uniform resolution may come at the cost of wider correlation sidelobes. Further investigation of the tradeoffs between resolution, noise, and correlation is required for both space-invariant and space-variant systems.

The design given in (14) has the advantage of being able to provide more uniform resolution in systems even when the inherent system response is shift-variant (e.g., SPECT). However, there is no computationally efficient method for computing these roughness penalties yielding uniform resolution properties for systems where  $\mathbf{A}'\mathbf{A}$  is space-variant. Therefore, the ideas used in the practical implementation presented here need to be extended to shift-variant tomographic systems. Additionally, the methods presented here are for 2-D reconstruction. We also plan to investigate 3D penalties for resolution uniformity in volumetric reconstructions [22].

#### APPENDIX SIMPLIFIED PENALTY DESIGN FOR A SPECIFIC CLASS OF DESIRED RESPONSES

In the design given by (14)–(16), for every desired response  $\underline{l}_0$  one must recompute a new penalty matrix  $\mathbf{R}$ . For example, one might want to perform reconstructions with a set of desired responses with different FWHM resolutions. Each resolution requires a separate  $\mathbf{R}$  calculation, much of which may be precomputed as in Section IV. For further simplification, in this Ap-

pendix we present a specific class of desired impulse responses that may be specified which require only a single penalty matrix computation.

As mentioned below (2), for traditional space-invariant penalties, the  $w_{jk}$  terms in (2) include the regularization parameter  $\beta$ , which controls the mean global resolution. For shift-invariant penalties where  $\beta$  is a simple multiplicative factor we may write  $\mathbf{R} = \beta\mathbf{R}_0$ , where  $\mathbf{R}_0$  specifies the relative penalty strength between pixel pairs and  $\beta$  controls the mean global resolution. Therefore it is simple to generate new  $\mathbf{R}$  for different desired resolutions. (One does not have to recompute  $\mathbf{R}_0$ .)

Just as the conventional shift-invariant penalty is a simple function of  $\beta$ , we would like to design the penalty matrix  $\mathbf{R}$  as a product of a user-selected  $\beta$  and a  $\beta$ -independent  $\mathbf{R}_*$ , i.e.,  $\mathbf{R} = \beta\mathbf{R}_*$ , yet still yields uniform resolution properties. In terms of our parameterization of  $\mathbf{R}$ , we would like factorable coefficients such that  $\underline{r}^j = \beta\underline{s}^j$ . Making this substitution into (14) yields

$$\hat{\underline{r}}^j = \beta\hat{\underline{s}}^j, \quad \hat{\underline{s}}^j \triangleq \arg \min_{\underline{s}^j \geq 0} \|\beta\Phi^j \underline{s}^j - \underline{d}^j\|^2. \quad (22)$$

The penalty matrix  $\mathbf{R}_*$  is completely specified by  $\{\underline{s}^j\}_{j=1}^p$ . However, the minimization in (22) depends on  $\beta$ . We eliminate this dependence by an particular choice of the target frequency response  $\underline{L}_0$  and the weighting  $\mathbf{V}^j$  in (15) and (16).

Let us consider the idealized PET system where (17) is an appropriate system matrix factorization. In this case, the local impulse response of an unweighted least-squares estimator with penalty matrix  $\mathbf{R} = \beta\mathbf{R}_0$  and with  $c_i$ 's are all unity, is

$$\underline{l}_0^j = [\mathbf{G}'\mathbf{G} + \beta\mathbf{R}_0]^{-1}\mathbf{G}'\mathbf{G}\underline{e}^j. \quad (23)$$

If  $\mathbf{R}_0$  is chosen to be a space-invariant penalty, the response  $\underline{l}_0^j$  is approximately independent of the choice of  $j$  since  $\mathbf{G}'\mathbf{G}$  is a nearly shift-invariant operator. That is,  $\underline{l}_0^j$  is nearly the same (with appropriate shifts) for all  $j$ . This particular choice of  $\underline{l}_0$  has a form very similar to the local impulse response in (3) and has resolution controlled by the parameter  $\beta$ .

Using the simplifications discussed above (9), we express the frequency response of (23) as

$$\underline{L}_0^j \approx \frac{\mathbf{F}\{\mathbf{G}'\mathbf{G}\underline{e}^j\}}{\mathbf{F}\{\mathbf{G}'\mathbf{G}\underline{e}^j\} + \beta\mathbf{F}\{\mathbf{R}_0\underline{e}^j\}}. \quad (24)$$

Similarly, we may write

$$1 - \underline{L}_0^j \approx \frac{\beta\mathbf{F}\{\mathbf{R}_0\underline{e}^j\}}{\mathbf{F}\{\mathbf{G}'\mathbf{G}\underline{e}^j\} + \beta\mathbf{F}\{\mathbf{R}_0\underline{e}^j\}}. \quad (25)$$

For the particular choice (23) of  $\underline{l}_0$ , the denominators of (24) and (25) are identical. Additionally,  $\beta$  is in the numerator of (25) and not in the numerator of (24). If we choose a least-squares weighting of  $\mathbf{V}^j = D[\mathbf{F}\{(\mathbf{G}'\mathbf{G} + \beta\mathbf{R}_0)\underline{e}^j\}]$  the denominators of (24) and (25) disappear in (15) and (16), and we can rewrite the penalty design as

$$\begin{aligned} \hat{\underline{s}}^j &= \arg \min_{\underline{s}^j \geq 0} \|\Phi^j \underline{s}^j - \underline{d}^j\|^2 \\ \Phi^j &= D[\mathbf{F}\{\mathbf{G}'\mathbf{G}\underline{e}^j\}]\mathbf{F}\{\mathbf{B}^j\} \\ \underline{d}^j &= D[\mathbf{F}\{\mathbf{R}_0\underline{e}^j\}]\mathbf{F}\{\mathbf{G}'\mathbf{W}\mathbf{G}\underline{e}^j\}. \end{aligned} \quad (26)$$

The design (26) is independent of  $\beta$ , as desired.

Once we have calculated the parameters  $\{\hat{\underline{s}}^j\}_{j=1}^p$  using (26), we construct the penalty matrix  $\mathbf{R}_*$  using (7) with  $\underline{r}^j = \beta\hat{\underline{s}}^j$ . Since only the symmetric component of  $\mathbf{R}_*$  affects the penalty, we use  $\mathbf{R}_*^{\text{sym}}$ , which requires less memory. This  $\mathbf{R}_*^{\text{sym}}$  has been designed to provide global isotropic resolution properties and, because of the least-squares weighting leading to (26),  $\mathbf{R}_*^{\text{sym}}$  is independent of the choice of the regularization parameter  $\beta$ . Therefore, once  $\mathbf{R}_*^{\text{sym}}$  is calculated one may specify a desired global resolution through  $\beta$ . The penalty matrix is given by the simple relation  $\mathbf{R}^{\text{sym}} = \beta\mathbf{R}_*^{\text{sym}}$ . (A method relating  $\beta$  to the FWHM resolution is discussed in [5].)

The computational simplifications discussed in Section IV can also be applied to (26).

#### ACKNOWLEDGMENT

The authors wish to thank J. Qi for an insightful discussion regarding the generality of the proposed design.

#### REFERENCES

- [1] E. Veklerov and J. Llacer, "Stopping rule for the MLE algorithm based on statistical hypothesis testing," *IEEE Trans. Med. Imag.*, vol. 6, pp. 313–319, Dec. 1987.
- [2] D. L. Snyder and M. I. Miller, "The use of sieves to stabilize images produced with the EM algorithm for emission tomography," *IEEE Trans. Nuclear Sci.*, vol. 32, pp. 3864–3871, Oct. 1985.
- [3] K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing," *IEEE Trans. Med. Imag.*, vol. 9, pp. 439–446, June 1981. Corrections, *IEEE Trans. Med. Imag.*, vol. 10, p. 288.
- [4] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs," *IEEE Trans. Imag. Process.*, vol. 5, pp. 1346–1358, Sept. 1996.
- [5] J. A. Fessler, "Resolution properties of regularized image reconstruction methods," Comm. and Sig. Process. Lab., Dept. EECS, Univ. Mich., Ann Arbor, MI, Tech. Rep. 297, Aug. 1995.
- [6] D. W. Wilson and B. M. W. Tsui, "Spatial resolution properties of FB and ML-EM reconstruction methods," in *Proc. IEEE Nuclear Sci. Symp. Med. Imag. Conf.*, vol. 2, 1993, pp. 1189–1193.
- [7] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Imag. Process.*, vol. 1, pp. 322–336, July 1992.
- [8] R. Leahy and X. H. Yan, "Statistical models and methods for PET image reconstruction," in *Proc. Stat. Comp. Sect. Amer. Stat. Assoc.*, 1991, pp. 1–10.
- [9] J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Regularized emission image reconstruction using imperfect side information," *IEEE Trans. Nuclear Sci.*, vol. 39, pp. 1464–1471, Oct. 1992.
- [10] J. Qi and R. M. Leahy, "A theoretical study of the contrast recovery and variance of MAP reconstructions with applications to the selection of smoothing parameters," *IEEE Trans. Med. Imag.*, vol. 18, pp. 293–305, Apr. 1999.
- [11] —, "Resolution and noise properties of MAP reconstruction for fully 3D PET," in *Proc. 1999 Int. Mtg. on Fully 3D Imag. Recon. in Rad. Nuc. Med.*, 1999, pp. 35–39.
- [12] S. J. Reeves, "Optimal space-varying regularization in iterative image restoration," *IEEE Trans. Imag. Process.*, vol. 3, pp. 319–323, May 1994.
- [13] J. Nunez and J. Llacer, "Variable resolution Bayesian image reconstruction," in *Proc. IEEE Wkshp. Nonlin. Sig. Imag. Proc.*, 1995.
- [14] J. S. Liow and S. C. Strother, "The convergence of object dependent resolution in maximum likelihood based tomographic image reconstruction," *Phys. Med. Biol.*, vol. 38, pp. 55–70, Jan. 1993.
- [15] E. J. Hoffman, S. C. Huang, D. Plummer, and M. E. Phelps, "Quantitation in positron emission tomography: 6 effect of nonuniform resolution," *J. Comp. Assisted Tomog.*, vol. 6, pp. 987–999, Oct. 1982.
- [16] J. G. Rogers, "A method for correcting the depth-of-interaction blurring in PET cameras," *IEEE Trans. Med. Imag.*, vol. 14, pp. 146–50, Mar. 1995.

- [17] J. A. Fessler and S. D. Booth, "Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction," *IEEE Trans. Imag. Process.*, vol. 8, pp. 688–99, May 1999.
- [18] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
- [19] J. A. Fessler. (1997, Mar.) Spatial resolution properties of penalized weighted least-squares image reconstruction with model mismatch. Tech. Rep. 308, Comm. Sig. Process. Lab., Dept. EECS, Univ. Michigan, Ann Arbor, MI. [Online]. Available: <http://www.eecs.umich.edu/~fessler>
- [20] J. A. Fessler and A. O. Hero, "Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms," *IEEE Trans. Imag. Process.*, vol. 4, pp. 1417–29, Oct. 1995.
- [21] J. A. Fessler. (1995, July) ASPIRE 3.0 user's guide: A sparse iterative reconstruction library. Tech. Rep. 293, Comm. Signal Process. Lab., Dept. EECS, Univ. Michigan, Ann Arbor, MI. [Online]. Available: <http://www.eecs.umich.edu/~fessler>
- [22] J. W. Stayman and J. A. Fessler, "Penalty design for uniform spatial resolution in 3D penalized-likelihood image reconstruction," in *Proc. 1999 Intl. Mtg. Fully 3D Imag. Recon. in Rad. Nuclear Med.*, 1999.