Minimax Emission Computed Tomography using High Resolution Anatomical Side Information and B-Spline Models

Alfred O. Hero, Robinson Piramuthu, Jeffrey A. Fessler, and Steven R. Titus

Abstract

In this paper a minimax methodology is presented for combining information from two imaging modalities having different intrinsic spatial resolutions. The focus application is emission computed tomography (ECT), a low resolution modality for reconstruction of radionuclide tracer density, when supplemented by high resolution anatomical boundary information extracted from a magnetic resonance image (MRI) of the same imaging volume. The MRI boundary within the 2D slice of interest is parameterized by a closed planar curve. The Cramèr-Rao (CR) lower bound is used to analyze estimation errors for different boundary shapes. Under a spatially inhomogeneous Gibbs field model for the tracer density a representation for the minimax MRI-enhanced tracer density estimator is obtained. It is shown that the estimator is asymptotically equivalent to a penalized maximum likelihood (PML) estimator with resolution-selective Gibbs penalty. Quantitative comparisons are presented using the iterative Space Alternating Generalized Expectation Maximization (SAGE-EM) algorithm to implement the PML estimator with and without minimax weight averaging.

Keywords: Multiresolution imaging, planar curves, Fisher information, Expectation Maximization (EM) algorithm, penalized maximum likelihood, Cramèr-Rao (CR) bound, spatially variant Gibbs field model, asymptotic marginalization.

[1]A. Hero (corresponding author), R. Piramuthu, and J. Fessler are with the Dept. of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109-2122. S. R. Titus is with Lattice Semiconductor, OR. This work was supported in part by National Institutes of Health grant RO1-CA-54362.

I. Introduction

This paper is concerned with the following question arising in multiresolution imaging. How should information from two imaging sensors with different intrinsic spatial resolutions and noise statistics be combined in order to better estimate some common feature of the images? It is intuitively reasonable that a feature in a noisy high resolution image should be forced onto the lower resolution image only if the feature can be estimated with sufficiently small variance. This paper provides a theoretical basis for this intuition in the context of minimax optimal fusion of high and low resolution image information for the multi-modality medical imaging application described below. The minimax procedure can be viewed as a simple two step process: 1) apply a spatially variant blur function to the high resolution image, where the amount of blurring is inversely proportional to the observed Fisher information at that spatial location; 2) use an iterative penalized maximum likelihood method to fuse the regions in the images having high Fisher information.

Emission computed tomography (ECT) is a functional imaging modality which reconstructs the radionuclide spatial distribution from gamma rays emitted from a patient after injection of a radioactive tracer. The tracer uptake over a region-of-interest (ROI) is defined as the integral over the ROI of the radionuclide distribution. Tracer uptake estimation is an essential tool in medicine and biological sciences for evaluating metabolic function of living systems. Magnetic resonance imaging (MRI) is an anatomical imaging modality which images tissue structures based on proton spin transverse magnetization differences. ECT and MRI are complementary modalities. MRI has very high spatial resolution and produces exquisitely detailed images of anatomical structures, such as organ boundaries. ECT has far lower spatial resolution than MRI but can track a large number of different biochemical compounds as they are metabolized by organs within the body. Recently, it has been recognized that when functional and anatomical organ boundaries are spatially correlated, MRI and ECT data can be combined to improve the accuracy of uptake estimates [1], [2], [3], [4], [5], [6].

The simplest method for merging the high resolution MRI data, which we call side information, into the low resolution ECT data is post-reconstruction integration: simply integrate the reconstructed ECT image within a boundary extracted from the MRI image. This is the approach taken in [7], [8] for quantification of radiotracer uptake in functional brain imaging. However, resolution mismatch can produce severe bias due to blurring of an organ's tracer intensity across the organ boundary. To avoid such bias it is better to incorporate the MRI side information as an integral part of the ECT image reconstruction process. One of the principal algorithms proposed for this purpose is iterative maximum a posteriori (MAP), equivalently penalized maximum likelihood (PML), tracer reconstruction implemented with a spatially variant Gibbs prior, or penalty, to enforce smoothness and incorporate boundary information. Most penalty functions operate at pixel scale and fall into one of two categories: penalties using inter-pixel line site models [9], [1]; and penalties using tissue-label models at each pixel [10], [6].

MAP/PML methods can also produce bias in uptake estimates when noisy or poorly registered side information leads to boundary mismatch between the ECT and MRI images [11]. Various methods for reducing these mismatch errors have been proposed including: line site blurring [12]; line site weighting according to correlation with functional boundaries [13]; joint estimation of tracer intensity and line sites, [14], [15]; joint estimation of tracer intensity and anatomical tissue labels [16]; and robustification via sub-quadratic potential functions [17]. While more accurate than simpler methods like [12], joint estimation methods [15], [16] have the disadvantage of high computational complexity due in part to the presence of local extrema in the likelihood function. This paper develops a minimax approach to uptake estimation with organ boundary side information that provides protection against boundary mismatch.

There are several important elements that distinguish the approach described in this paper from previous approaches. First, a smooth continuously parameterizable boundary model is used which permits sub-pixel boundary resolution. Second, the boundary Fisher information matrix (FIM) [18] is used to specify *a priori* boundary estimator accuracy. The FIM specifies a lower bound on the covariance of the boundary estimate and also plays a key role in simplifying the implementation of the minimax uptake estimator. Third, a minimum probability of error criterion is adopted to robustify radiotracer estimation errors to boundary mismatch. We show that the minimax radiotracer estimate is equivalent to a PML reconstruction with non-quadratic averaged Gibbs penalty, where the averaging is performed over a "least-favorable" distribution of the boundary. Using the smooth parametric boundary model we show that the averaged Gibbs penalty is asymptotically equivalent to a standard quadratic Gibbs penalty implemented with blurred penalty weights where the blurring kernel is inversely related to the FIM. This approach is illustrated for a particularly simple boundary parameterization using periodic B-splines which, like all polar parametric boundary models, apply only to star shaped organs, i.e., organs whose boundary can be described by a radius function which is a scalar function of angle. This class includes a wide range of complex non-convex shapes but would not include, for example, kidney shaped regions.

In Section II the problem statement along with principal simplifying assumptions are given. In Sections III-A and III-B the standard additive Gaussian statistical model for high resolution MRI measurements is given and a periodic B-spline representation for star-shaped regions is presented. Two boundary estimators are then presented, one implemented by the maximum likelihood principle and the other a simpler edge filtering (EF) implementation. In Section III-C an expression for the FIM is given for the B-spline parameters and the matrix CR bound is used to explore the estimator covariance. An interesting conclusion is that for sufficiently high MRI spatial resolution the trace of the matrix CR bound is minimized for disk shapes. This agrees with intuition that, for fixed MRI spatial resolution, smooth rounded shapes should be the easiest to estimate. Simulations are also given showing that for high MRI contrast-to-noise ratio (CNR), both the ML and the EF boundary estimators have virtually identical mean square error (MSE) performance, both nearly achieving the CR bound. However, as the CNR decreases below a certain threshold the EF estimator error increases more quickly than that of the ML estimator.

In Section IV the focus is turned to the problem of radiotracer reconstruction for the case of perfect side information. Perfect side information occurs when the MRI-derived boundary estimator resolution is uniformly higher than ECT spatial resolution. In ECT the measurements are noisy projections of the object's radiotracer density onto a series of detectors surrounding the object. The ECT problem differs from MRI in two important ways: 1) the ECT measurement noise is Poisson distributed while for MRI it is Gaussian; 2) even without measurement noise the ECT reconstruction problem is ill-conditioned unless a smoothness constraint is imposed on the radiotracer density. Direct optimization of the Poisson likelihood function is intractable and requires an iterative estimation approach. The ill-conditioning is handled in Section IV-A by a set of weights in a Gibbs penalty which enforces smoothness everywhere in the object except across the organ boundary. This leads to an iterative penalized maximum likelihood (PML) reconstruction algorithm, equivalently a MAP algorithm with Gibbs prior, implemented using the space alternating generalized expectation maximization (SAGE-EM) version of the EM algorithm [19].

In Section IV-C the PML algorithm with perfect side information is generalized to the case of noisy boundary estimates. Here the MRI-derived boundary estimator resolution may be lower than ECT spatial resolution in some regions of the image. To deal with imperfect side information a minimax reconstruction algorithm is defined which minimizes the worst case probability of reconstruction error. A solution to this minimax problem is given in Section IV-C which takes the form of a modified PML estimator with non-quadratic penalty. Unfortunately this estimator is not suited to iterative implementation since this would require repeated multi-dimensional integration of a non-linear function of the Gibbs penalty over a least favorable boundary distribution. Using a Bernstein Von-Mises asymptotic theorem on the least favorable densities, in Section IV-D it is shown that the minimax reconstruction can be approximated by a PML algorithm with a resolution-selective Gibbs penalty function. This is a quadratic Gibbs penalty implemented with a blurred or smoothed weight map. The blurring is performed via space-variant convolution of the weight map with a multi-dimensional Gaussian density having local covariance equal to the inverse observed Fisher information matrix [20].

Finally in Section V numerical results are presented. Representative reconstructed concentration density images are shown which indicate significant qualitative improvements using the PML algorithm with resolution-selective Gibbs penalty. These improvements are quantitatively supported with bias-variance tradeoff curves.

II. MRI-aided ECT Uptake Estimation

Consider a square slice $\mathcal{S} = [-W, W] \times [-W, W]$ of an object volume which contains an organ or tumor that selectively absorbs a radionuclide. The MRI boundary estimation problem consists of generating an estimate of the organ boundary from Y_M, a noisy MRI image of the slice. The ECT tracer uptake estimation problem with boundary side information consists of estimating the total radionuclide concentration $\int_{\mathrm{ROI}} \lambda \ dx\, dy$ within the region-of-interest (ROI) contained inside the boundary, where $\lambda = \{\lambda(x, y) : (x, y) \in \mathcal{S}\}$ is the radionuclide concentration density. A block diagram of the general tracer uptake problem with and without side information is shown in Fig. 1 for a generic parameterization $\boldsymbol{\theta}$ of the boundary.

To maintain focus on the principal features of the approach, the following simplifying assumptions are made. First, out-of-plane septal penetration and bleed-through effects are neglected so that all detected radionuclide emissions can be assumed to originate from the slice $\mathcal{S}$. Second, the image is assumed to consist of a single organ, which might be a hot or a cold tumor, in a uniform background. Third, the anatomical boundary is assumed to be star shaped with respect to some point within the ROI, i.e. with respect to that point any directed ray intersects the boundary exactly once. In Section V the ellipsoidal cold spot phantom shown in Fig. 2 will be used to demonstrate and compare performance. Finally, it is assumed that the MRI and ECT images are perfectly registered, i.e. they correspond to the same slice of image data without relative shift, scale or rotation.

III. Anatomical Boundary Extraction

MRI produces an image of tissue which is essentially a noisy, discretized estimate of proton spin transverse magnetization within a 2D slice. The magnetization depends on relaxation parameters which differ for different tissue types and thus MRI can give high contrast and high resolution images of anatomical tumor or organ boundaries.

Many non-parametric approaches to boundary estimation have been proposed, e.g., morphological transforms such as medial axis, skeleton, thinning algorithms, active contours, and Laplacian edge extraction operators [21]. An example of the latter approach is to perform edge detection by thresholding the gradient image followed by an algorithm which connects and smooths the detected edges to form a closed contour. An origin is selected interior to the boundary, e.g. by computing the centroid of the image plane, and the MRI slice is transformed to polar coordinates (ϕ, r) with respect to this origin, $\phi \in [-\pi, \pi)$, $r > 0$. The row index of the image matrix then corresponds to angular position and the column index to radial position of a pixel. The Canny [22], or Marr-Hildreth [23], edge detection gradient operator is then applied to each row of the polar image to extract a coarse estimate of the boundary radius as a function of angle. The coarse estimate is smoothed using a nonlinear median filter to eliminate small scale variations in the radial estimate.The median filter is applied to the angular periodic extension of the image to preserve continuity and guarantee that the smoothed estimate corresponds to a closed contour, i.e., $(\phi, r) = (\phi + 2\pi, r)$, in Cartesian coordinates.

A. B-Spline Boundary Model

Another approach to boundary extraction is to impose a parametric boundary model followed by formulation of a parametric estimation problem. Some of the well-known models used to represent closed boundaries are periodic planar curve models such as Fourier descriptors, fitting of line segments, cubics, Bezier curves, Beta-splines and B-splines [24]. In this paper we adopt the B-spline boundary model. A B-spline consists of a set of K fixed positions, called knots, and piecewise smooth curves, called basis functions, connecting each of the knot positions. For an m-th order B-spline these curves are specified by polynomial functions of degree m. To ensure smoothness at each knot the curve is constrained to have continuous derivatives up to order $m-1$. B-splines can track local shape deformations using a small number of parameters, unlike Fourier descriptors which require many parameters and can have spurious oscillations. This localization property of B-splines is due to their compact support.

For an origin fixed inside the region of interest, the boundary is represented by a radial function $r(\phi)$ continuously indexed by polar angle $\phi \in (-\pi, \pi]$. The m-th order B-spline model specifies the boundary $r(\phi)$ as a linear combination of basis functions $B_i(\phi)$

$$r(\phi) = r_{\boldsymbol{\theta}}(\phi) = \sum_{i=1}^{K} \theta_i B_i(\phi), \qquad \phi \in (-\pi, \pi] \tag{1}$$

where θ_i is a spline coefficient, $\{B_i\}_{i=1}^{K}$ are basis functions,

$$B_i(\phi) = (\phi_{i+m+1} - \phi_i) \sum_{j=1}^{m+1} \frac{[\max(0, \phi_{i+j} - \phi)]^m}{\prod_{\substack{l=0 \\ l \neq j}}^{m+1} (\phi_{i+j} - \phi_{i+l})}, \tag{2}$$

and $i = 1, \ldots, K - m$ and $\{\phi_j\}$ are K angular knot locations. B-spline basis functions (2) are polynomials functions of ϕ and fast recursive numerical computation algorithms are available [25].

It will be convenient to represent the spline representation (1) in vector form

$$r_{\boldsymbol{\theta}}(\phi) = \mathbf{B}^T(\phi)\, \boldsymbol{\theta} \tag{3}$$

where

$$\mathbf{B}(\phi) \quad = \quad [B_1(\phi), \ldots, B_K(\phi)]^T .$$

B. MRI Measurement Model

The noiseless and unblurred MRI image is assumed to be a two-level function in the plane:

$$\mathrm{I}_{\boldsymbol{\theta}}(x, y) = C_{ROI} \cdot \mathrm{I}_{R_{\boldsymbol{\theta}}}(x, y) + C_{BG} \cdot (\mathrm{I}_{R_f}(x, y) - \mathrm{I}_{R_{\boldsymbol{\theta}}}(x, y)), \tag{4}$$

where I_{R_f} and $\mathrm{I}_{R_{\boldsymbol{\theta}}}$ denote indicator functions for the field-of-view R_f and the ROI $R_{\boldsymbol{\theta}}$, respectively (see Fig. 3). The contrast of the MRI image is given by $C_{ROI} - C_{BG}$, where $C_{ROI}, C_{BG} \geq 0$.

Under the assumptions that the MRI scanner has space-invariant Gaussian point spread function $H(x, y)$ and that the measurement noise is additive zero mean white Gaussian [26] we have the following representation for the MRI measurements

$$Y_{\mathrm{M}}(x, y) \quad = \quad \mathrm{I}^s_{\boldsymbol{\theta}}(x, y) + \varepsilon(x, y), \qquad \forall (x, y) \in R_f \tag{5}$$

where

$$\begin{aligned} \mathrm{I}^s_{\boldsymbol{\theta}}(x, y) \quad &= \quad (\mathrm{I}_{\boldsymbol{\theta}} ** H)(x, y) && (6) \\ H(x, y) \quad &= \quad \frac{1}{2\pi\sigma_s^2} \exp\left[-\frac{x^2 + y^2}{2\sigma_s^2}\right] && (7) \end{aligned}$$

and in (6) the 2-dimensional convolution is denoted by $**$. Note that the full-width-half-max (FWHM) spatial resolution of the MRI scanner is proportional to the width σ_s of the symmetric 2D Gaussian-shaped point spread function $H(x, y)$. Note also that for practical systems $\sigma_s \ll 2W$ which is the width of the field of view (FOV).

Under a Gaussian noise assumption the ML estimator of $\boldsymbol{\theta}$ takes the form of a non-linear least squares estimator

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\mathrm{ML}} &= \arg\max_{\boldsymbol{\theta}} \ \ln f_{\boldsymbol{\theta}}(\mathbf{Y}_{\mathrm{M}}) \\ &= \arg\min_{\boldsymbol{\theta}} \ \|\mathbf{Y}_{\mathrm{M}} - \mathrm{I}_{\boldsymbol{\theta}} ** H\|^2. \end{aligned} \tag{8}$$

A conjugate gradient least squares algorithm can be used to minimize the squared norm with respect to the spline parameters $\boldsymbol{\theta}$ and the intensity levels C_{ROI} and C_{BG}. Partial volume effects can be dealt with by performing sub-pixel interpolation. Details of such a ML estimator implementation can be found in [27].

C. Cramèr-Rao Bound for Boundary Estimation

It is well known that under broad conditions [28] the covariance matrix $\mathrm{cov}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) = E_{\boldsymbol{\theta}}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T]$ of any unbiased estimator $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\mathbf{Y}_{\mathrm{M}})$ of a nonrandom parameter vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^T$ satisfies the matrix CR bound

$$\mathrm{cov}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) \geq \mathbf{F}_{\boldsymbol{\theta}}^{-1},$$

which is shorthand notation for: $\mathrm{cov}(\hat{\boldsymbol{\theta}}) - \mathbf{F}_{\boldsymbol{\theta}}^{-1}$ is a non-negative definite matrix. $\mathbf{F}_{\boldsymbol{\theta}}$ is the Fisher information matrix [18], [28]

$$\begin{aligned}\mathbf{F}_{\boldsymbol{\theta}} &= E_{\boldsymbol{\theta}}\left[\nabla_{\boldsymbol{\theta}}^T \ln f(\mathbf{Y}_{\mathrm{M}}; \boldsymbol{\theta}) \ \nabla_{\boldsymbol{\theta}} \ln f(\mathbf{Y}_{\mathrm{M}}; \boldsymbol{\theta})\right] \\ &= E_{\boldsymbol{\theta}}\left[-\nabla_{\boldsymbol{\theta}}^2 \ln f(\mathbf{Y}_{\mathrm{M}}; \boldsymbol{\theta})\right],\end{aligned}$$

where $\nabla_{\boldsymbol{\theta}}$ is a row gradient vector and $\nabla_{\boldsymbol{\theta}}^2 = \nabla_{\boldsymbol{\theta}}^T \nabla_{\boldsymbol{\theta}}$ is a Hessian matrix. Under the MRI measurement model (5) and the B-spline model (3) we derive the following expression for $\mathbf{F}_{\boldsymbol{\theta}}$ in Appendix A

$$\mathbf{F}_{\boldsymbol{\theta}} = C_{\mathrm{CN}} \iint_{-\pi}^{\pi} \exp\left[-\frac{\|\vec{r}_{\boldsymbol{\theta}}(\phi) - \vec{r}_{\boldsymbol{\theta}}(\gamma)\|^2}{4\sigma_s^2}\right] r_{\boldsymbol{\theta}}(\phi)\, r_{\boldsymbol{\theta}}(\gamma)\, \mathbf{B}(\phi)\, \mathbf{B}^T(\gamma)\, d\phi\, d\gamma \tag{9}$$

where

$$C_{\mathrm{CN}} = \frac{(C_{ROI} - C_{BG})^2}{4\pi\sigma_n^2\sigma_s^2}$$

is the normalized contrast. In (9) $\vec{r}_{\boldsymbol{\theta}}(\phi) = [r(\phi)\cos\phi,\ r(\phi)\sin\phi]^T$ denotes the vector in the plane describing the Cartesian coordinates of the boundary at angle ϕ.

While the bound (9) gives a useful performance limit on unbiased estimators of the spline parameters $\boldsymbol{\theta}$, the performance of unbiased estimators $\hat{\mathbf{r}}_{\boldsymbol{\theta}}$ of the boundary $\mathbf{r}_{\boldsymbol{\theta}} = \{r_{\boldsymbol{\theta}}(\phi) : \phi \in [-\pi, \pi)\}$ will be of greater interest in the sequel. Under the B-spline model (3) the boundary function $r_{\boldsymbol{\theta}}(\phi)$, $\phi \in [-\pi, \pi)$, is constrained to be in the span of the spline basis $B_1, \ldots, B_K$. Using a generalization of the constrained CR bound [29] to function estimation, the covariance function $\mathrm{cov}_{\mathbf{r},\boldsymbol{\theta}}(\hat{r}(\phi),\ \hat{r}(\gamma))$ of any unbiased boundary estimator $\hat{r}(\cdot)$ can be shown to satisfy the following bound stated in terms of a non-negative definiteness (n.n.d.) condition (see Appendix B)

$$\mathrm{cov}_{\mathbf{r},\boldsymbol{\theta}}\left(\hat{r}(\phi),\ \hat{r}(\gamma)\right) \ - \ \mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma), \qquad \text{is n.n.d.} \tag{10}$$

where $\mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma)$ is the CR lower bound

$$\mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma) = \mathbf{B}^T(\phi)\ \mathbf{F}_{\boldsymbol{\theta}}^{-1}\ \mathbf{B}(\gamma). \tag{11}$$

In particular, the bound (10) implies that for any integrable scalar function $g(\cdot)$

$$\iint_{-\pi}^{\pi} g(\phi)\left[\mathrm{cov}_{\mathbf{r},\boldsymbol{\theta}}\left(\hat{r}(\phi),\ \hat{r}(\gamma)\right)\right] g(\gamma)\, d\phi d\gamma \ \geq \ \iint_{-\pi}^{\pi} g(\phi)\mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma) g(\gamma)\, d\phi d\gamma. \tag{12}$$

Define the constant α as

$$\alpha \ = \ \sqrt{2}\sigma_s / \min_{\phi} \sqrt{[r_{\boldsymbol{\theta}}(\phi)]^2 + [r_{\boldsymbol{\theta}}'(\phi)]^2}$$

where $r_{\boldsymbol{\theta}}'(\phi) = \partial r_{\boldsymbol{\theta}}(\phi)/\partial\phi$. The parameter α decreases to zero as the MRI scanner spatial resolution increases (small σ_s) or as the boundary shape is magnified ($\mathbf{r}_{\boldsymbol{\theta}} \longrightarrow a\mathbf{r}_{\boldsymbol{\theta}}$, $a > 1$). Roughly speaking, α is small when the resolution of the scanner is high enough to resolve the smallest details in the boundary function.

It is shown in Appendix C that the constrained CR bound (11) has the small α representation:

$$\mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma) = \frac{1}{2\sqrt{\pi}\sigma_s C_{\mathrm{CN}}}\ \mathbf{B}^T(\phi)\left[\int_{-\pi}^{\pi} \frac{r_{\boldsymbol{\theta}}^2(\psi)}{\sqrt{1+\left(\frac{\partial}{\partial\psi}\ln r_{\boldsymbol{\theta}}(\psi)\right)^2}}\,\mathbf{B}(\psi)\mathbf{B}^T(\psi)d\psi\right]^{-1}\mathbf{B}(\gamma) + o(\alpha). \tag{13}$$

D. CR Bound Properties

The following comments pertain to the small α approximation (13) to the CR bound

1. The CR bound decreases as the boundary radius function $r_{\boldsymbol{\theta}}(\phi)$ is scaled up (magnification). It also decreases as the normalized contrast C_{CN} increases. This contrast increases when either the point spread resolution width σ_s or the power level of the noise σ_n^2 decrease. Thus the CR bound predicts that among all possible objects, large and high contrast objects that can be estimated most accurately.

2. Since the CR bound depends on the actual contrast $C_{ROI} - C_{BG}$ only through its square via C_{CN}, the bound predicts that optimal estimator performance is independent of whether the ROI is a hot spot (positive contrast) or a cold spot (negative contrast).
3. An interesting question is: what is the shape of the boundary curve that minimizes the CR bound? Consider the following

$$\begin{aligned}
&\iint_{-\pi}^{\pi} g(\phi)\mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma)g(\gamma)\,d\phi d\gamma \\
&= \frac{1}{2\sqrt{\pi}\sigma_s C_{\text{CN}}} \iint_{-\pi}^{\pi} g(\phi)\mathbf{B}^T(\phi)\left[\int_{-\pi}^{\pi} \frac{r_{\boldsymbol{\theta}}^2(\psi)}{\sqrt{1+\left(\frac{\partial}{\partial\psi}\ln r_{\boldsymbol{\theta}}(\psi)\right)^2}}\mathbf{B}(\psi)\mathbf{B}^T(\psi)d\psi\right]^{-1} \mathbf{B}(\gamma)g(\gamma)\,d\phi d\gamma \\
&\geq \frac{1}{2\sqrt{\pi}\sigma_s C_{\text{CN}}} \min_{\psi} \frac{\sqrt{1+\left(\frac{\partial}{\partial\psi}\ln r_{\boldsymbol{\theta}}(\psi)\right)^2}}{r_{\boldsymbol{\theta}}^2(\psi)} \iint_{-\pi}^{\pi} g(\phi)\mathbf{B}^T(\phi)\left[\int_{-\pi}^{\pi}\mathbf{B}(\psi)\mathbf{B}^T(\psi)d\psi\right]^{-1}\mathbf{B}(\gamma)g(\gamma)\,d\phi d\gamma.
\end{aligned}$$

   where the inequality follows from the non-negative definiteness of the matrix $\int_{-\pi}^{\pi}\mathbf{B}(\psi)\mathbf{B}^T(\psi)d\psi$ (see Appendix B). Note that equality is attained if and only if $\frac{1}{r_{\boldsymbol{\theta}}^2(\psi)}\sqrt{1+\left(\frac{\partial}{\partial\psi}\ln r_{\boldsymbol{\theta}}(\psi)\right)^2}$ is independent of ψ. Since $r_{\boldsymbol{\theta}}(\psi)$ describes a closed curve, the only way for this to occur is if in fact $\frac{\partial}{\partial\psi}\ln r_{\boldsymbol{\theta}}(\psi) = 0$, i.e. $r_{\boldsymbol{\theta}}(\psi)$ is constant [30]. Thus disk shaped objects minimize the CR bound.
4. The worst case boundary shape which maximizes the constrained CR bound (13) is more difficult to specify. A basic property of such shapes is that the radius function $r_{\boldsymbol{\theta}}(\psi)$ have rapid variations as a function of ϕ.

E. Boundary Estimator Performance Comparison

Simulations of the ML and EF boundary algorithms discussed above were performed for a range of contrast-to-noise ratios C_{CN} and compared to the constrained CR bound (13) for the blurred B-spline phantom in Fig. 2. Quadratic B-splines with 16 uniformly spaced knots were used for ML estimation. The sample bias and sample standard deviation of the perimeters of the ML and EF boundary estimators were normalized by the true perimeter and were plotted against the normalized noise standard deviation $\sigma_n/(C_{ROI} - C_{BG})^2$ in Fig. 4. Fig. 4(a) is the plot of bias of the estimators. The ML estimates are virtually unbiased with absolute bias less than 0.2% of ground truth for all but the noisiest data. The bias of EF estimator is smaller than 2% for noise standard deviation less than 10% but increases to a maximum bias of about 9% above this noise threshold. In Figure 4(b), the variance of these estimators are compared to the constrained CR bound. The ML estimates come quite close to achieving the CR bound for noise standard deviation less than 10%. The

EF estimator has variance which uniformly exceeds that of the ML estimator, especially above the threshold of 10%.

IV. Estimation of Radionuclide Concentration

In emission computed tomography (ECT) a radionuclide is injected into a patient and gamma-ray photon emissions are counted by an array of collimated detectors oriented around the patient. The objective is to use these detected counts to estimate or reconstruct the radionuclide spatial concentration density in a square slice $\mathcal{S}$. The slice is discretized into P pixels at locations $\{(x_l, y_m)\}_{l,m}$ and the pixels are lexicographically ordered from 1 to P. The discretized spatial density is denoted by the vector $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_P]^T$ of integrated intensities within each pixel. The measured emission data is a vector of counts $\mathbf{Y}_{\mathrm{E}}$, each element of which corresponds to the number of counts falling into a detector element. These counts are assumed to be statistically independent Poisson random variables with mean vector

$$E[\mathbf{Y}_{\mathrm{E}}] = \mathbf{A}\,\boldsymbol{\lambda} + \boldsymbol{b}$$

where $\mathbf{A}$ is the ECT system matrix and $\boldsymbol{b}$ represents spurious detected counts originating from ambient background radioactivity, septal penetration effects, cosmic rays, scatter, or, in positron emission tomography (PET), random coincidences.

A. Penalized Maximum Likelihood Reconstruction

The most common method for reconstructing ECT images from projections is the non-iterative filtered back projection (FBP) algorithm [31], [32] which does not account for Poisson statistics, scattering, or attenuation. An iterative maximum likelihood (ML) image reconstruction algorithm was implemented with expectation-maximization in [33] which accounts for Poisson statistics and is easily modified to account for attenuation and scattering. However, the ML algorithm suffers from slow convergence and is not well suited to incorporation of image smoothness or anatomical side information. Penalized ML (PML) image reconstruction can enforce image smoothness by introducing roughness penalty functions into the loglikelihood objective function [19], [34]. The PML reconstruction is the maximizer over $\boldsymbol{\lambda}$ of

$$\boldsymbol{\Phi}_{\mathrm{PML}}(\boldsymbol{\lambda}, \boldsymbol{w}) = \ln f(\mathbf{Y}_{\mathrm{E}}; \boldsymbol{\lambda}) - \beta R(\boldsymbol{\lambda}, \boldsymbol{w}) \tag{14}$$

Here, $\ln f(\mathbf{Y}_{\mathrm{E}};\boldsymbol{\lambda})$ is the loglikelihood function of $\boldsymbol{\lambda}$ given the ECT measurements $\mathbf{Y}_{\mathrm{E}}$, $\boldsymbol{w}$ is a vector of weights, $R(\boldsymbol{\lambda},\boldsymbol{w})$ is a quadratic penalty

$$R(\boldsymbol{\lambda},\boldsymbol{w}) = \sum_{j=1}^{P}\sum_{k\in N_j} w_{jk}\cdot(\lambda_j-\lambda_k)^2, \tag{15}$$

and β is a smoothness parameter — a large value of β strongly emphasizes the penalty, and hence encourages smoothness in $\boldsymbol{\lambda}$, while a small β only weakly emphasizes the penalty.

In (15) N_j is the pixel neighborhood of the j-th pixel defining the spatial extent of the dependency structure. $w_{jk}(\hat{\boldsymbol{\theta}})$ is a weighting function which can be used to enforce or relax the roughness penalty by selectively penalizing differences $\lambda_j - \lambda_k$ over a neighborhood.

Fig. 5 shows the second order neighborhood structure that is used for this paper. The pixel j of interest is the grey pixel in the center. The neighborhood N_j is the set of 8 pixels, called second order neighbors, labeled by black dots. The weights w_{jk} are chosen to be symmetric ($w_{jk} = w_{kj}$) and take on a positive value between pixel j and its neighbors and are zero otherwise. By making the weights w_{jk} also depend on a boundary estimate $\hat{\boldsymbol{\theta}}$ it is possible to eliminate or reduce smoothing across the estimated boundary.

B. Weight Assignments

Here several weight assignments are briefly described. These will be compared in Section V.

B.1 Ideal Weights

When exact errorless estimates $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_o$ ($\boldsymbol{\theta}_o$=the true parameter) of the MRI boundary spline parameters are available, the boundary can be forced into the weights by setting $w_{jk}(\boldsymbol{\theta}) = 1$ if pixels j and k are both within the boundary specified by $\boldsymbol{\theta}$ or if they are completely outside of the boundary. Otherwise, w_{jk} is set to zero when j and k are on opposite sides of the boundary. This assignment completely decouples pixels on either side of the boundary yet still encourages smoothness within and exterior to the boundary. A mathematical definition of this assignment is

$$w_{jk}(\boldsymbol{\theta}_o) = \begin{cases} 1, & \mathrm{I}_{R_{\boldsymbol{\theta}_o}}(j) = \mathrm{I}_{R_{\boldsymbol{\theta}_o}}(k) \\ 0, & o.w. \end{cases} \tag{16}$$

where, as in (4), $\mathrm{I}_{R_{\boldsymbol{\theta}_o}}(j) = \mathrm{I}_{R_{\boldsymbol{\theta}_o}}(x_j,y_j)$ is as before the indicator function of the interior of the boundary at the j-th lexicographically ordered pixel located at (x_j,y_j).

B.2 Plug-in Weights

When only an imperfect estimate $\hat{\boldsymbol{\theta}}$ is available the simplest approach is to use it in place of $\boldsymbol{\theta}_o$ in the weight assignment (16). This estimate-and-plug technique, which we call the plug-in weight mapping, gives weights $w_{jk}(\hat{\boldsymbol{\theta}})$ which encourage smoothing without accounting for the distribution of errors in the boundary estimate. We will later refer to these as unsmoothed weights.

B.3 Variance-corrected Weights

A refinement of the estimate-and-plug technique is to average the weights with respect to an empirically determined multivariate density of the estimate $\hat{\boldsymbol{\theta}}$. This produces the "variance-corrected weight mapping"

$$\tilde{w}_{jk}(\hat{\boldsymbol{\theta}}) = \int_{\mathbf{R}^K} w_{jk}(\boldsymbol{\theta})\; f(\boldsymbol{\theta};\hat{\boldsymbol{\theta}})d\boldsymbol{\theta} \tag{17}$$

where $f(\boldsymbol{\theta};\hat{\boldsymbol{\theta}})$ is a multidimensional density function in the argument $\boldsymbol{\theta}$ which is indexed by the estimate $\hat{\boldsymbol{\theta}}$ and w_{jk} is as given in (16). We will later refer to these as smoothed weights. In the next section this method is theoretically justified from a minimax perspective.

C. Gibbs Field Model and a minimax Estimator

The PML estimator (14) is identical in form to a *maximum a posteriori* (MAP) estimator of $\boldsymbol{\lambda}$ when $\boldsymbol{\lambda}$ is taken to be a random vector with a Gaussian prior [35], [3], [36]

$$f_{\boldsymbol{w}}(\boldsymbol{\lambda}) = \rho(\boldsymbol{w}) \cdot \exp\left(-\beta R(\boldsymbol{\lambda}, \boldsymbol{w})\right), \tag{18}$$

where $\rho(\boldsymbol{w})$ is a normalization constant depending only on the weight map $\boldsymbol{w}$ and $\beta > 0$. The prior is known as a Gibbs prior and a simulation of $\boldsymbol{\lambda}$ from the prior is called a Gibbs random field with potential function $R(\boldsymbol{\lambda}, \boldsymbol{w})$.

Let $\epsilon > 0$ be a specified estimation error tolerance. Under the Gibbs prior, the average probability of error associated with any estimator $\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\mathbf{Y})$ based on the ECT/MRI measurements $\mathbf{Y} = \{\mathbf{Y}_{\mathrm{E}}, \mathbf{Y}_{\mathrm{M}}\}$ is

$$\begin{aligned} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}) &= P_{\boldsymbol{w}}\left(\left\|\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\right\| > \epsilon\right) \\ &= \iint_{\left\{\mathbf{Y},\boldsymbol{\lambda}:\left\|\hat{\boldsymbol{\lambda}}(\mathbf{Y})-\boldsymbol{\lambda}\right\|>\epsilon\right\}} d\mathbf{Y}d\boldsymbol{\lambda}\; f(\mathbf{Y}|\boldsymbol{\lambda})f_{\boldsymbol{w}}(\boldsymbol{\lambda}), \end{aligned} \tag{19}$$

where $f(\mathbf{Y}|\boldsymbol{\lambda})$ denotes the conditional density given the Gibbs field $\boldsymbol{\lambda}$.

It is well known, e.g. [37, Ch. 2] or [38, Ch. 4], that for small ϵ the minimum probability of error estimator $\hat{\boldsymbol{\lambda}}^*$ is equivalent to the MAP estimator.

$$\hat{\boldsymbol{\lambda}}^* = \hat{\boldsymbol{\lambda}}^*(\boldsymbol{w}) = \text{argmax}_{\lambda}\{f(\mathbf{Y}|\boldsymbol{\lambda})f_{\boldsymbol{w}}(\boldsymbol{\lambda})\}.$$

When the true boundary $\boldsymbol{\theta} = \boldsymbol{\theta}_o$ is known, one can incorporate it into the weights $\boldsymbol{w} = \boldsymbol{w}(\boldsymbol{\theta}_o)$ of the PML and be ensured of minimum probability of error performance. Assume that $\mathbf{Y}_{\mathrm{E}}$ and $\mathbf{Y}_{\mathrm{M}}$ are conditionally independent given $\boldsymbol{\lambda}$. Then the following factorization holds

$$f(\mathbf{Y}|\boldsymbol{\lambda})f_{\boldsymbol{w}(\boldsymbol{\theta}_o)}(\boldsymbol{\lambda}) = f(\mathbf{Y}_{\mathrm{E}}|\boldsymbol{\lambda})f_{\boldsymbol{w}(\boldsymbol{\theta}_o)}(\boldsymbol{\lambda})f_{\boldsymbol{\theta}_o}(\mathbf{Y}_{\mathrm{M}}). \tag{20}$$

As $f_{\boldsymbol{\theta}_o}(\mathbf{Y}_{\mathrm{M}})$ does not depend on $\boldsymbol{\lambda}$ the MAP estimator maximizing the log of (20) is equivalent to the PML estimator given by (14) with ideal weights (16).

Of course, for an unknown boundary the ideal weights are not implementable. The alternative explored here is to replace the estimation criterion (19) with the worst case probability of estimation error

$$\max_{\boldsymbol{\theta}} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})). \tag{21}$$

The estimator $\hat{\boldsymbol{\lambda}}^*$ which minimizes the worst case probability of error is a minimax estimator which is robust to uncertainty in the boundary parameters $\boldsymbol{\theta}$. An implicit form of this minimax estimator can be specified under compactness and continuity assumptions [39, Thm 2.9.2]. The following is proven in Appendix D.

Proposition 1: For specified weight mapping $\boldsymbol{w}$ and parameter $\boldsymbol{\theta}$ let $P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta}))$ be the risk function for an estimator $\hat{\boldsymbol{\lambda}}$. Assume that the conditions of [39, Thm. 2.9.2] are satisfied. Assume in addition that the function $f(\mathbf{Y}_{\mathrm{E}}|\boldsymbol{\lambda})f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda})$ is continuous and has a strict global maximum over $\boldsymbol{\lambda}$. Then

$$\min_{\hat{\boldsymbol{\lambda}}} \max_{\boldsymbol{\theta}} P_{\boldsymbol{w}(\boldsymbol{\theta})}\left(\left\|\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\right\| > \epsilon\right) = \max_{dP} \int_{\mathbb{R}^K} P_{\boldsymbol{w}(\boldsymbol{\theta})}\left(\left\|\hat{\boldsymbol{\lambda}}^*(dP) - \boldsymbol{\lambda}\right\| > \epsilon\right)\, dP(\boldsymbol{\theta}), \tag{22}$$

where the maximum on the right side of (22) is over all probability measures dP on the Borel subsets of $\mathbb{R}^K$, and $\hat{\boldsymbol{\lambda}}^*(dP)$ has the representation

$$\hat{\boldsymbol{\lambda}}^*(dP) = \text{argmax}_{\boldsymbol{\lambda}}\left\{\int_{\mathbb{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda})f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda})dP(\boldsymbol{\theta})\right\} + O(\epsilon). \tag{23}$$

The notation $dP(\boldsymbol{\theta})$ in the proposition denotes the differential of a probability measure P, and when P is absolutely continuous with respect to Lebesgue measure: $dP(\boldsymbol{\theta}) = f(\boldsymbol{\theta})d\boldsymbol{\theta}$ where $f(\boldsymbol{\theta})$ is a probability density. In the proposition $\hat{\boldsymbol{\lambda}}^*(dP)$ is a Bayes estimator which minimizes the average probability of error; averaged with respect to the probability measure dP. By [39, Thm. 2.10.2], the minimax estimator is extended-Bayes, i.e., given arbitrary $\delta > 0$ there exists a probability measure dP^* for which

$\int_{\mathbf{R}^K} P_{\boldsymbol{w}(\boldsymbol{\theta})}\left(\left\|\hat{\boldsymbol{\lambda}}^*(dP^*) - \boldsymbol{\lambda}\right\| > \epsilon\right)\ dP^*(\boldsymbol{\theta})$ comes to within δ of the maximum value on the right hand side of (22). The measure dP^* is called the *least favorable distribution.* The minimax estimator thus has an average probability of error which is essentially attainable by the Bayes estimator $\hat{\boldsymbol{\lambda}}^*(dP^*)$. Furthermore, by (23),

$$\hat{\boldsymbol{\lambda}}^* = \text{argmax}_{\boldsymbol{\lambda}} \int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP^*(\boldsymbol{\theta}) \tag{24}$$

is, to order ϵ, a minimax estimator of $\boldsymbol{\lambda}$.

Observe that the least favorable distribution dP^* is a function of the weight mapping scheme $\boldsymbol{w}(\cdot)$ but is independent of $\boldsymbol{\lambda}$. Note that since $\boldsymbol{\theta}$ is non-random dP^* cannot be strictly interpreted as a probability measure of a random variable. Rather dP^* is a probability measure which is induced by the minimax criterion. Determination of the dP^* necessary for implementation of the estimator (24) is generally difficult but not infeasible, at least in principle. A sufficient condition for a distribution dP to be least favorable is that it equalize the risk of the PML estimator over $\boldsymbol{\theta}$. For some simple examples see [39], [40] and [41], [42]. Fortunately, for the results of this paper the abstract representation (24) will be sufficient without explicit determination of dP^*.

The estimator (24) can be related to a modified PML estimator with Gibbs-like penalty as follows. Under the conditional independence of MRI and ECT data the factorization (20) holds and the integrand of the right side of (24) has the representation

$$\begin{aligned} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP^*(\boldsymbol{\theta}) &= f(\mathbf{Y}_{\mathrm{E}}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) f_{\boldsymbol{\theta}}(\mathbf{Y}_{\mathrm{M}}) dP^*(\boldsymbol{\theta}) \\ &= f(\mathbf{Y}_{\mathrm{E}}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP^*(\boldsymbol{\theta}|\mathbf{Y}_{\mathrm{M}}) \tilde{f}^*(\mathbf{Y}_{\mathrm{M}}) \end{aligned} \tag{25}$$

where $\tilde{f}^*$ is the marginalized MRI density

$$\tilde{f}^*(\mathbf{Y}_{\mathrm{M}}) = \int f_{\boldsymbol{\theta}}(\mathbf{Y}_{\mathrm{M}}) dP^*(\boldsymbol{\theta}).$$

and $dP^*(\boldsymbol{\theta}|\mathbf{Y}_{\mathrm{M}})$ is the induced least favorable posterior

$$dP^*(\boldsymbol{\theta}|\mathbf{Y}_{\mathrm{M}}) = \frac{f_{\boldsymbol{\theta}}(\mathbf{Y}_{\mathrm{M}})\ dP^*(\boldsymbol{\theta})}{\tilde{f}^*(\mathbf{Y}_{\mathrm{M}})}. \tag{26}$$

Integrating both sides of (26) with respect to $\boldsymbol{\theta}$ and taking the logarithm, we obtain from (24) that the estimator $\hat{\boldsymbol{\lambda}}^*$ is the vector $\boldsymbol{\lambda}$ which maximizes

$$\ln\left\{\int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP^*(\boldsymbol{\theta})\right\} \;=\; \ln f(\mathbf{Y}_{\mathrm{E}}|\boldsymbol{\lambda}) \;+\; \ln \tilde{f}^*(\mathbf{Y}_{\mathrm{M}}) \;+\; \ln\left\{\int_{\mathbf{R}^K} f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP^*(\boldsymbol{\theta}|\mathbf{Y}_{\mathrm{M}})\right\} \tag{27}$$

Using the Gibbs prior distribution (18) and the relation (27), and neglecting constants independent of $\boldsymbol{\lambda}$, the estimator $\hat{\boldsymbol{\lambda}}^*$ of (24) is seen to maximize the following PML objective

$$\boldsymbol{\Phi}_{\mathrm{minimax}}(\boldsymbol{\lambda}) \;=\; \ln f(\mathbf{Y}_{\mathrm{E}}|\boldsymbol{\lambda}) \;+\; \ln \int_{\mathbf{R}^K} \rho(\boldsymbol{w}(\boldsymbol{\theta})) \, \exp\left(-\beta R(\boldsymbol{\lambda}; \boldsymbol{w}(\boldsymbol{\theta}))\right) \, dP^*(\boldsymbol{\theta}|\mathbf{Y}_{\mathrm{M}}). \tag{28}$$

The second additive term in (28) is a penalty which is a nonquadratic function of $\boldsymbol{\lambda}$. Its evaluation involves a complicated integration over all values of $\boldsymbol{\theta} \in \mathbb{R}^K$ to which $dP^*(\boldsymbol{\theta}|\mathbf{Y}_{\mathrm{M}})$ assigns positive measure. This makes (28) ill-suited for iterative numerical maximization, e.g., by Fisher scoring, coordinate ascent, or the EM algorithm, since the integral would have to be evaluated at every iteration. A practical algorithm is given by the approximation in the next section.

The intuition behind the approximation described in the next section is as follows. When the MRI measurements $\mathbf{Y}_{\mathrm{M}}$ are of high quality, e.g. due to merging multiple scans, the posterior $dP^*(\boldsymbol{\theta}|\mathbf{Y}_{\mathrm{M}})$ concentrates as a function of $\boldsymbol{\theta}$ at or near the true parameter $\boldsymbol{\theta}_o$. Thus as the data quality improves the posterior concentrates in the vicinity of $\boldsymbol{\theta}_o$ and two things happen: (*i*) the influence of the induced least favorable prior function on the posterior dissapears; (*ii*) in (28) the integral $\int \exp(-[\beta R - \ln \alpha]) \, dP$ can be approximated by $\exp(-\int [\beta R - \ln \alpha] \, dP)$ making the penalty term in (28) quadratic in $\boldsymbol{\lambda}$. Furthermore, in a manner analogous to the mechanism behind the central limit theorem, the posterior approaches a Gaussian distribution whose mean and covariance are related to the ML estimator of $\boldsymbol{\theta}$ and its Fisher information matrix, respectively.

D. Asymptotic minimax Approximation

Assume that the experiment giving measurements $\mathbf{Y}_{\mathrm{M}}$ is repeated n times yielding the i.i.d. observation set $\underline{\mathbf{Y}_{\mathrm{M}}}^n = \{\mathbf{Y}_{\mathrm{M}}^1, \ldots, \mathbf{Y}_{\mathrm{M}}^n\}$. Let $\hat{\boldsymbol{\theta}}^n = \mathrm{argmax}_{\boldsymbol{\theta}} \, f_{\boldsymbol{\theta}}(\underline{\mathbf{Y}_{\mathrm{M}}}^n)$ be the maximum likelihood estimate of $\boldsymbol{\theta}$ based on $\mathbf{Y}_{\mathrm{M}}^n$. Under general conditions it is well known that smooth posterior densities converge to Gaussian densities when the maximum likelihood estimator is consistent. This property was used by Laplace [43] and is known as the Von-Mises Bernstein Theorem and is commonly stated in terms of convergence of the posterior probability measure $dP(\boldsymbol{\theta}|\underline{\mathbf{Y}_{\mathrm{M}}}^n)$ to a Gaussian measure $dG(\boldsymbol{\theta}|\boldsymbol{T}_n, \boldsymbol{\Gamma}_n)$ where

$$dG(\boldsymbol{\theta}|\boldsymbol{T}_n, \boldsymbol{\Gamma}_n) \;=\; \frac{n}{(2\pi)^{K/2}|\boldsymbol{\Gamma}_n|^{\frac{1}{2}}} \, \exp\left\{-\frac{n}{2}(\boldsymbol{\theta} - \boldsymbol{T}_n)^T \boldsymbol{\Gamma}_n^{-1} (\boldsymbol{\theta} - \boldsymbol{T}_n)\right\} d\boldsymbol{\theta} \tag{29}$$

where $\boldsymbol{T}_n$ and $\boldsymbol{\Gamma}_n$ are a vector and positive definite matrix, respectively, possibly depending on $\underline{\mathbf{Y}_{\mathrm{M}}}^n$ but not on $\boldsymbol{\theta}$. We state the following general version of this convergence result which is discussed in [44, Secs 12.4 and 17.7].

Theorem 1: Let $\boldsymbol{\theta}_o$ be the true value of the K-dimensional real parameter vector $\boldsymbol{\theta}$ and let $dP(\boldsymbol{\theta})$ be a prior distribution. Assume that the likelihood $f_{\boldsymbol{\theta}}(\mathbf{Y}_{\mathrm{M}}^1)$ and $dP(\boldsymbol{\theta})$ are smooth in $\boldsymbol{\theta}$ in the sense of satisfying the regularity conditions (B1)-(B7) of [44, Sec. 17.7]. Then the posterior probability distribution $dP(\boldsymbol{\theta}|\underline{\mathbf{Y}_{\mathrm{M}}}^n)$ satisfies

1. For any open ball $B = B(\boldsymbol{\theta}_o, \epsilon) = \{\boldsymbol{\theta} : \|\boldsymbol{\theta} - \boldsymbol{\theta}_o\| < \epsilon\}$ centered at $\boldsymbol{\theta}_o$ of radius $\epsilon > 0$, $\int_B dP(\boldsymbol{\theta}|\underline{\mathbf{Y}_{\mathrm{M}}}^n) \rightarrow 1$ (i.p.).

2. For any Borel subset V of $\mathbb{R}^K$: $\left|\int_V dP(\boldsymbol{\theta}|\underline{\mathbf{Y}_{\mathrm{M}}}^n) - \int_V dG(\boldsymbol{\theta}||\boldsymbol{T}_n, \boldsymbol{\Gamma}_n)\right| \rightarrow 0$ (i.p.) when, in (29), $\boldsymbol{\Gamma}_n$ is a matrix which converges in probability to the Fisher information $\mathbf{F}_{\boldsymbol{\theta}_o} = E_{\boldsymbol{\theta}_o}\left[-\nabla^2_{\boldsymbol{\theta}_o} \ln f_{\boldsymbol{\theta}_o}(\mathbf{Y}_{\mathrm{M}}^1)\right]$, and $\boldsymbol{T}_n$ is an estimator which converges in probability to $\boldsymbol{\theta}_o$.

The practical interpretation of Assertion 2 of the Theorem is that a multivariate Gaussian distribution with mean vector $\boldsymbol{T}_n$ and covariance matrix $\boldsymbol{\Gamma}_n$ is an asymptotically accurate approximation to the posterior distribution. Note that the Theorem applies to a large class of sequences $\boldsymbol{T}_n$ and $\boldsymbol{\Gamma}_n$. Note also that the asymptotic form of the posterior is *independent* of the explicit form of the prior $dP(\boldsymbol{\theta})$. This is the key to simplification of the modified PML criterion (28).

An important special case of Theorem 1 is obtained when $\boldsymbol{T}_n$ is taken as the maximum likelihood estimate $\hat{\boldsymbol{\theta}} = \mathrm{argmax}_{\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\underline{\mathbf{Y}_{\mathrm{M}}}^n)$ and $\boldsymbol{\Gamma}_n$ is taken as the *observed Fisher information* evaluated at $\hat{\boldsymbol{\theta}}$:

$$\hat{\mathbf{F}}_{\hat{\boldsymbol{\theta}}} = -\frac{1}{n}\nabla^2_{\hat{\boldsymbol{\theta}}} \ln f(\underline{\mathbf{Y}_{\mathrm{M}}}^n|\hat{\boldsymbol{\theta}}). \tag{30}$$

Use of these assignments of $\boldsymbol{T}_n$ and $\boldsymbol{\Gamma}_n$ in Theorem 1 give an asymptotic approximation to the posterior which is identical to the profile posterior approximation proposed in [45]. Another common proposal is to take $\boldsymbol{\Gamma}_n$ in Theorem 1 to be the *expected Fisher information*

$$\mathbf{F}_{\hat{\boldsymbol{\theta}}} = E_{\boldsymbol{\theta}}\left[-\nabla^2_{\boldsymbol{\theta}} \ln f(\mathbf{Y}_{\mathrm{M}}^1|\boldsymbol{\theta})\right]|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{31}$$

which is simply the Fisher information formula (9) evaluated at $\hat{\boldsymbol{\theta}}$. In terms of the asymptotic theorem, the use of either of the information matrices (30) or (31) in the Gaussian approximation (29) is equivalent. However, there have been studies showing that the observed information has better finite sample approximation properties in certain situations [20]. The observed Fisher information matrix for the B-spline boundary model is derived in Appendix F.

Theorem 1 allows us to obtain the following key result, proven in Appendix E.

Proposition 2:

Let the assumptions underlying Theorem 1 be satisfied and define the smoothed weights

$$\tilde{w}_{jk}(\hat{\boldsymbol{\theta}}) = \int_{\mathbf{R}^K} w_{jk}(\boldsymbol{\theta})\, dG(\boldsymbol{\theta}|\boldsymbol{T}_n, \boldsymbol{\Gamma}_n) \tag{32}$$

where $dG(\boldsymbol{\theta}; \boldsymbol{T}_n, \boldsymbol{\Gamma}_n)$ is a multivariate Gaussian distribution with mean vector $\boldsymbol{T}_n$ and covariance matrix $\boldsymbol{\Gamma}_n$ specified in Assertion 2 of Theorem 1. Then, within a $\boldsymbol{\lambda}$-independent additive constant C, the minimax objective function $\boldsymbol{\Phi}_{\text{minimax}}(\boldsymbol{\lambda}; \boldsymbol{w})$ defined in (28) converges from above to the PML objective function $\boldsymbol{\Phi}_{\text{PML}}(\boldsymbol{\lambda}; \tilde{\boldsymbol{w}})$ defined in (14) in the sense that

$$\boldsymbol{\Phi}_{\text{PML}}(\boldsymbol{\lambda}; \tilde{\boldsymbol{w}}) \;\le\; \boldsymbol{\Phi}_{\text{minimax}}(\boldsymbol{\lambda}; \boldsymbol{w}) - C \;\le\; \boldsymbol{\Phi}_{\text{PML}}(\boldsymbol{\lambda}; \tilde{\boldsymbol{w}}) + \gamma_n \tag{33}$$

where the sequence $\gamma_n > 0$ converges to zero in probability.

It follows from the above that, if the ML estimator $\hat{\boldsymbol{\theta}}$ is consistent, the minimax estimator of $\boldsymbol{\lambda}$ is asymptotically equivalent to a PML estimator $\hat{\boldsymbol{\lambda}}$ implemented via Gibbs penalty with smoothed weights

$$\hat{\boldsymbol{\lambda}}^*_{PML} = \text{argmax}_{\boldsymbol{\lambda}} \left\{ \ln f(\mathbf{Y}_{\text{E}}|\boldsymbol{\lambda}) - \beta \sum_{j=1}^{P} \sum_{k \in N_j} \tilde{w}_{jk}(\hat{\boldsymbol{\theta}})(\lambda_j - \lambda_k)^2 \right\}, \tag{34}$$

where $\tilde{w}_{jk}(\hat{\boldsymbol{\theta}})$ is the K-dimensional convolution

$$\tilde{w}_{jk}(\hat{\boldsymbol{\theta}}) = (w_{jk} \star g)(\hat{\boldsymbol{\theta}}), \tag{35}$$

and

$$g(\boldsymbol{\theta}) = \frac{\left|\mathbf{F}_{\hat{\boldsymbol{\theta}}}\right|^{1/2}}{\left(\sqrt{2\pi}\right)^K} \exp\left\{ -\frac{1}{2} \boldsymbol{\theta}^T \mathbf{F}_{\hat{\boldsymbol{\theta}}}^{-1} \boldsymbol{\theta} \right\}.$$

Unlike the non-asymptotic minimax estimator $\hat{\boldsymbol{\lambda}}^*$ in (28) which would require averaging the $\boldsymbol{\lambda}$-dependent Gibbs prior density, the asymptotic minimax estimator $\hat{\boldsymbol{\lambda}}^*_{PML}$ in (34) only requires averaging the weights in the Gibbs prior. Since the weights do not depend on $\boldsymbol{\lambda}$ this computation can be computed prior to computing $\hat{\boldsymbol{\lambda}}^*_{PML}$. Furthermore, by discretization of the convolution, (35) can be efficiently computed via the multi-dimensional FFT [30].

V. Numerical Results

The ellipsoidal 64×64 pixel gray level phantom shown in Fig. 2 was smoothed using a Gaussian kernel with width $\sigma_s = 0.75$ pixels, which is 15% of the average radius of the inscribed ellipsoidal boundary (the ROI), and additive Gaussian noise was added to form the MRI image. Parallel projections of the phantom typical of a positron emission tomography (PET) geometry were computed with 64 radial bins and 60 angles uniformly spaced over 180^o. Poisson noise was then added with rates equal to the average intensities in each of the projection bins plus a 15% uniform background. Attenuation was not included. The scan time for PET was adjusted to obtain 1 million detected counts not including background events. PML reconstructions of the phantom were implemented using 60 iterations of SAGE3 (Space Alternating Generalized EM [19]) with a quadratic penalty and second order Gibbs weights. The reconstructions were performed on 400 independent realizations of emission data to obtain empirical estimates of bias and variance.

We first illustrate a spatially variant resolution property of the minimax averaged Gibbs weighting scheme. Fig. 6 shows a typical realization of the weight maps for ideal weights extracted directly from the noiseless phantom, unsmoothed (plug-in) weights extracted from the noisy phantom using the ML estimation technique described in Section III, and the corresponding smoothed (variance-corrected) weights. An outer elliptical boundary guard region and a hard-limit were added to the weight averaging in order to prevent the reconstruction of $\boldsymbol{\lambda}$ from leaking mass accross the estimated boundary. The hard-limit boundary is obtained by assigning zero weights to pairs of pixels that have the smallest magnitude averaged Gibbs weights. In each panel in Fig. 6 a location where the pixel is black indicates that the weight is zero, and a location where the pixels is white denotes a weight value of one. For example, in the upper left panel of Fig. 6(a) no smoothing of $\boldsymbol{\lambda}$ is enforced in the "left" direction at pixel locations which are labeled in black. The ideal weight maps in Fig. 6(a) capture the boundary of the ellipsoidal phantom without error. The unsmoothed weight maps in Fig. 6(b) show significant errors in the estimated boundary, especially on the left side. The smoothed weight maps in Fig. 6(c) resemble blurred versions of the ideal weight maps, the blurring being highest where the error of the unsmoothed weights is predicted by the observed Fisher information to be highest.

Fig. 7 shows representative reconstructions obtained by applying filtered back projection (FBP), PML without side information, PML with ideal side information, PML with unsmoothed weights (estimate-and-plug), and PML with minimax smoothed weights. Observe that the boundary of ROI is severely blurred for both FBP and PML without side information. This is because the spatial smoothing associated with these two algorithms is spatially invariant. One thus would expect the tracer uptake estimates for these reconstructions to have high negative bias. The PML with ideal side information enforces the high resolution ideal weight map in Fig. 6(a) into the ECT reconstruction which virtually eliminates smoothing of radiotracer across the boundary. The PML with unsmoothed weights does better than the FBP and PML without side information but considerably worse than the PML with ideal weights due to the enforcement of the erroneous high resolution weight map into ECT. The PML with smoothed weights appears to have recovered the ROI almost as well as the PML with ideal weights.

We next turn to the bias and standard deviation of total uptake estimates derived from the ECT image reconstructions. Total uptake was estimated for each of the ideal, unsmoothed, and smoothed PML weighting algorithms by simply summing the reconstructed pixel values within each of the estimated ROI's determined by the respective weight maps. For FBP and PML without side information, the ROI's were estimated by thresholding the reconstructed image. We compared the minimax smoothing technique using true Fisher information, expected Fisher information, and observed Fisher information.

In Fig. 8 the average RMS error of the uptake estimates is shown as a function of the smoothing parameter β for the ECT reconstructions: FBP, PML without side information, PML with EF extracted side information (unsmoothed weights), and PML with ML extracted side information (unsmoothed weights). These figures clearly indicate the benefit of side information. Note also that, as expected from the MRI benchmark studies presented in Section III-E, use of ML boundary estimates in the weights gives better uptake estimates than using suboptimal EF boundary estimates.

In Fig. 9 the % standard deviation is plotted against % bias of total uptake as β ranges from small (upper part of curves) to large (lower part of curves). The FBP and PML without side information algorithms performed off scale and are not shown. The improvement in bias-variance tradeoff due to minimax smoothing with ideal, expected, or observed Fisher information is obvious from the figure. The bias-variance curves using three different types of Fisher information are quite close to each other and to the ideal curve. On the other hand, as contrasted with the results reported in [20], Fig. 10 gives clear indication that smoothing with expected Fisher information significantly outperforms smoothing with observed Fisher information smoothing as the MRI noise variance σ_n^2 increases.

VI. Conclusion and Future Work

We have presented an asymptotically minimax tracer uptake estimation algorithm for merging high resolution MRI boundary information into low resolution ECT data. The algorithm merges the MRI boundary information into the ECT reconstruction by a penalized maximum likelihood method with a resolution selective Gibbs penalty implemented with a set of smoothed weights. Here the amount of smoothing is spatially variant and inversely proportional to the Fisher information. The effect of this spatially variant blurring is that the Gibbs penalty forces a segment of the estimated boundary into the ECT reconstruction only if the boundary estimate has high resolution, i.e., low variance as measured by the inverse Fisher information. In this way the penalty accomplishes resolution matching between the ECT and MRI data.

Several restrictive assumptions were made in this paper to maintain focus on the fundamentals. The PML structure of the asymptotic minimax approximation, given by Proposition 2, holds for any continuous boundary parameterization. As a consequence, while the implementation details will differ, the general method is applicable to any continuously parameterizable non-star shaped organ in 2D or 3D. 3D parameterizations under investigation include: thin plate splines, spherical harmonics, and 3D prolate spheroidal functions. With the addition of continuously variable translation, scale, and rotation variables into the unknown parameter vector $\boldsymbol{\theta}$ the method can also be extended to the case of imperfectly registered images. In principle, the approach could also be extended to multiple organs, although this would require conditioning on their number and multiple hypothesis testing.

Acknowledgement

The authors thank the three reviewers of this paper for comments and suggestions which led to improvements in the presentation. They also gratefully acknowledge insightful comments during the formative stages of this work by Dr. Leslie Rogers of the Division of Nuclear Medicine, University of Michigan.

I. APPENDIX

A. Fisher Information Matrix for B-Spline Coefficients

Under the linear Gaussian model (5) the loglikelihood of the noisy MRI data is given by Girsanov's formula [46, Ch 13]

$$\ln f(\mathbf{Y}_{\mathrm{M}};\boldsymbol{\theta}) = \left(-\frac{1}{2\sigma_n^2}\right)\int_{R_f}\left[\mathbf{Y}_{\mathrm{M}}(x,y) - (\mathrm{I}_{\boldsymbol{\theta}} ** H)(x,y)\right]^2 \; dx\; dy + C \tag{36}$$

where C is a term independent of $\boldsymbol{\theta}$, and $R_f = [-W, W] \times [-W, W]$ is the square domain of the MRI image. The Hessian matrix of the loglikelihood is

$$\begin{aligned} \nabla^2_{\boldsymbol{\theta}} \ln f(\mathbf{Y}_{\mathrm{M}};\boldsymbol{\theta}) \quad = \quad & -\frac{1}{\sigma_n^2}\iint\limits_{R_f} \nabla^T_{\boldsymbol{\theta}}\mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\cdot\nabla_{\boldsymbol{\theta}}\mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\; dx\; dy \\ & +\frac{1}{\sigma_n^2}\iint\limits_{R_f}\left[\mathbf{Y}_{\mathrm{M}}(x,y) - \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\right]\nabla^2_{\boldsymbol{\theta}}\mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\; dx\; dy \end{aligned} \tag{37}$$

Since, for θ equal to the true parameter, $\mathbf{Y}_{\mathrm{M}}(x,y) - \mathrm{I}^s_{\boldsymbol{\theta}}(x,y) = \varepsilon(x,y)$ is zero mean the expectation of the second term on the right hand side is zero and thus

$$\mathbf{F}_{\boldsymbol{\theta}} = \frac{1}{\sigma_n^2} \iint\limits_{-W}^{W} \nabla_{\boldsymbol{\theta}}^T \mathrm{I}_{\boldsymbol{\theta}}^s(x,y) \cdot \nabla_{\boldsymbol{\theta}} \mathrm{I}_{\boldsymbol{\theta}}^s(x,y)\, dx\ dy. \tag{38}$$

Now, using equations (4) and (7),

$$\nabla_{\boldsymbol{\theta}} \mathrm{I}_{\boldsymbol{\theta}}^s(x,y) = \frac{(C_{ROI} - C_{BG})}{2\pi\sigma_s^2}\, \nabla_{\boldsymbol{\theta}} \iint\limits_{R_{\boldsymbol{\theta}}} \exp\left[-\frac{(x-\tau_1)^2 + (y-\tau_2)^2}{2\sigma_s^2}\right] d\tau_1 d\tau_2.$$

A change of variable in the above integral from rectangular to polar coordinates, and the definition

$$C_s = (C_{ROI} - C_{BG})/(2\pi\sigma_s^2)$$

give

$$\begin{aligned}
\nabla_{\boldsymbol{\theta}} \mathrm{I}_{\boldsymbol{\theta}}^s(x,y) &= C_s \cdot \nabla_{\boldsymbol{\theta}} \left\{ \iint\limits_{R_{\boldsymbol{\theta}}} \exp\left[-\frac{(x-\tau_1)^2 + (y-\tau_2)^2}{2\sigma_s^2}\right]\ d\tau_1\ d\tau_2 \right\} \\
&= C_s \cdot \nabla_{\boldsymbol{\theta}} \left\{ \int_{\phi=0}^{2\pi} \int_{\zeta=0}^{r_{\boldsymbol{\theta}}(\phi)} \exp\left[-\frac{(x-\zeta\cos\phi)^2 + (y-\zeta\sin\phi)^2}{2\sigma_s^2}\right]\ \zeta\ d\zeta\ d\phi \right\} \\
&= C_s \cdot \int_{\phi=0}^{2\pi} \exp\left[-\frac{(x-r_{\boldsymbol{\theta}}(\phi)\cos\phi)^2 + (y-r_{\boldsymbol{\theta}}(\phi)\sin\phi)^2}{2\sigma_s^2}\right]\ r_{\boldsymbol{\theta}}(\phi)\, \nabla_{\boldsymbol{\theta}} r_{\boldsymbol{\theta}}(\phi)\ \ d\phi \\
&\qquad \text{(By Leibnitz's rule)}
\end{aligned}$$

Now use $\nabla_{\boldsymbol{\theta}} r_{\boldsymbol{\theta}}(\phi) = \mathbf{B}(\phi)$, substitute the above into (38), complete the square in the exponent of the integrand, and identify

$$\|\vec{r}_{\boldsymbol{\theta}}(\phi) - \vec{r}_{\boldsymbol{\theta}}(\gamma)\|^2 = r_{\boldsymbol{\theta}}^2(\phi) + r_{\boldsymbol{\theta}}^2(\gamma) - 2r_{\boldsymbol{\theta}}(\phi) r_{\boldsymbol{\theta}}(\gamma)\cos(\phi-\gamma)$$

to obtain

$$\mathbf{F}_{\boldsymbol{\theta}} = \frac{C_s^2}{\sigma_n^2} \int_{-W}^{W} dx \int_{-W}^{W} dy \int_0^{2\pi} d\phi \int_0^{2\pi} d\gamma\ p(x,y) q(\boldsymbol{\theta},\phi,\gamma) \mathbf{G}(\boldsymbol{\theta},\phi,\gamma) \tag{39}$$

where

$$\begin{aligned}
q(\boldsymbol{\theta},\phi,\gamma) &= \exp\left(-\frac{\|\vec{r}_{\boldsymbol{\theta}}(\phi) - \vec{r}_{\boldsymbol{\theta}}(\gamma)\|^2}{4\sigma_s^2}\right) \\
\mathbf{G}(\boldsymbol{\theta},\phi,\gamma) &= r_{\boldsymbol{\theta}}(\phi) r_{\boldsymbol{\theta}}(\gamma) \mathbf{B}(\phi) \mathbf{B}^T(\gamma) \\
p(x,y) &= \exp\left(-\frac{1}{\sigma_s^2}\left[\left(x-\frac{a}{2}\right)^2 + \left(y-\frac{b}{2}\right)^2\right]\right)
\end{aligned}$$

and

$$\begin{aligned} a &= r_{\boldsymbol{\theta}}(\phi)\cos\phi + r_{\boldsymbol{\theta}}(\gamma)\cos\gamma \\ b &= r_{\boldsymbol{\theta}}(\phi)\sin\phi + r_{\boldsymbol{\theta}}(\gamma)\sin\gamma. \end{aligned}$$

Under the assumption that $\sigma_s \ll W$ the integral over (x, y) can be performed by integration of the unnormalized bivariate Gaussian probability density $p(x, y)$ resulting in the expression

$$\mathbf{F}_{\boldsymbol{\theta}} = \frac{C_s^2 \pi \sigma_s^2}{\sigma_n^2} \int_0^{2\pi} d\phi \int_0^{2\pi} d\gamma \; q(\boldsymbol{\theta}, \phi, \gamma)\mathbf{G}(\boldsymbol{\theta}, \phi, \gamma), \tag{40}$$

which is readily verified to be identical to the expression (9).

B. Cramèr-Rao Bound for Boundary Estimation

The CR bound for finite dimensional parameters subject to constraints has been studied in [29], [47] and [48]. While proof of the general extension theorem is outside of the scope of this paper, it can be shown that the theory in [29] extends to continuously indexed parameters, e.g. functions such as $\mathbf{r} = \{r(\phi) : \phi \in [-\pi, \pi)\}$. In particular, the CR bound for unbiased estimates of $\mathbf{r}$ subject to the linear constraint (3) can be stated as: $\mathrm{cov}_{\mathbf{r}}\,(r(\phi), r(\gamma)) - \mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma)$ is a symmetric non-negative definite (n.n.d.) kernel where, analogously to [29, Lemma 2 Eq. (19)] or [48, Thm. 1 Eq. (6)], $\mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma)$ is the symmetric positive definite kernel

$$\begin{aligned} \mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma) &= \mathbf{B}^T(\phi)[\mathbf{B}(\cdot) \odot \mathbf{F}_{\mathbf{r}}(\cdot,\star) \odot \mathbf{B}^T(\star)]^{-1}\mathbf{B}(\gamma) \\ &= \mathbf{B}^T(\phi)\left[\iint_{-\pi}^{\pi} \mathbf{B}(\psi)\mathbf{F}_{\mathbf{r}}(\psi,\eta)\mathbf{B}^T(\eta)\, d\psi d\eta\right]^{-1} \mathbf{B}(\gamma). \end{aligned} \tag{41}$$

In the above equation the integral inner product operator $g(\cdot) \odot f(\cdot) = \int_{-\pi}^{\pi} g(u)f(u)du$ has been introduced to simplify notation. The symmetric function $\mathbf{F}_{\mathbf{r}}(\psi,\eta)$ is the (unconstrained) Fisher information corresponding to unconstrained estimation of $\mathbf{r}$

$$\mathbf{F}_{\mathbf{r}}(\psi,\eta) = -E_{\mathbf{r}}\left[\frac{\partial^2}{\partial r(\psi)\partial r(\eta)} \ln f(\mathbf{Y}_{\mathrm{M}};\mathbf{r})\right] \tag{42}$$

$$= C_{\mathrm{CN}}\, r(\psi) r(\eta) \exp\left(-\frac{\|\vec{r}(\psi) - \vec{r}(\eta)\|^2}{4\sigma_s^2}\right). \tag{43}$$

In view of the form of the FIM for $\boldsymbol{\theta}$ derived in the previous appendix, it is evident that $F_{\boldsymbol{\theta}} = \mathbf{B}(\cdot) \odot \mathbf{F}_{\mathbf{r}}(\cdot,\star) \odot \mathbf{B}^T(\star)$. Thus we have

Lemma 1: Let $\mathbf{B}(\phi)$ be a vector function $\mathbf{B} : \mathbb{R} \to \mathbb{R}^K$ and let $\mathbf{r}(\phi)$ be a scalar function on $\mathbb{R}$ which is constrained to have the form $\mathbf{r}(\phi) = \mathbf{B}(\phi)\boldsymbol{\theta}$ for some $\boldsymbol{\theta} \in \mathbb{R}^k$. Assume that the likelihood function $f_{\boldsymbol{\theta}}(\mathbf{Y}_\mathrm{M})$ satisfies conditions (15)-(19) of [29]. Then any unbiased estimator $\hat{\mathbf{r}}$ of the function $\mathbf{r}$ satisfies the following *constrained CR bound*

$$\mathrm{cov}_{\mathbf{r}}\left(r(\phi), r(\gamma)\right) - \mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma), \qquad \text{is n.n.d.,} \tag{44}$$

where

$$\mathrm{CRB}_{\mathbf{r},\boldsymbol{\theta}}(\phi,\gamma) = \mathbf{B}^T(\phi) F_{\boldsymbol{\theta}}^{-1} \mathbf{B}(\gamma).$$

Proof of Lemma 1

Define the $K \times K$ matrix

$$\mathbf{B}(\cdot) \odot \mathbf{B}^T(\cdot) = \int_{-\pi}^{\pi} \mathbf{B}(\psi)\mathbf{B}^T(\psi) d\psi.$$

This matrix is non-singular since $\mathbf{B} = [B_1, \ldots, B_K]^T$ is a column vector of linearly independent functions. Indeed, there exists a non-singular lower triangular Gramm-Schmidt orthogonalization matrix $\mathbf{L}$ such that $\tilde{\mathbf{B}}(\phi) = \mathbf{L}\mathbf{B}(\phi)$ is a vector of orthonormal functions and therefore $\mathbf{B}(\cdot) \odot \mathbf{B}^T(\cdot) = \mathbf{L}[\tilde{\mathbf{B}}(\cdot) \odot \tilde{\mathbf{B}}^T(\cdot)]\mathbf{L}^T = \mathbf{L}\mathbf{L}^T$ which is obviously non-singular. Thus we can define the projection operator

$$P(\phi,\gamma) = \mathbf{B}^T(\phi)[\mathbf{B}(\cdot) \odot \mathbf{B}^T(\cdot)]^{-1}\mathbf{B}(\gamma)$$

which projects square integrable functions on $[-\pi,\pi)$ onto the linear span of the functions $B_1, \ldots, B_K$. In particular, for any such function $g(u)$ lying in the span of these functions

$$P(\phi,\cdot) \odot g(\cdot) = \int_{-\pi}^{\pi} P(\phi,u) g(u) du = g(\phi)$$

and therefore

$$[\delta(\phi,\cdot) - P(\phi,\cdot)] \odot g(\cdot) = 0$$

for all $\phi \in [-\pi,\pi)$, where $\delta(\phi,\gamma) = \delta(\phi-\gamma)$ is a dirac delta function.

Now let $\hat{r}(\phi)$ be an arbitrary unbiased estimator of the B-spline boundary $r(\phi)$, where $r(\phi)$ lies in the span of $B_1, \ldots, B_K$. Then $\hat{r}_P(\phi) = P(\phi,\cdot) \odot \hat{r}(\cdot)$ is also an unbiased estimator since

$$E_{\mathbf{r}}[\hat{r}_P(\phi)] = P(\phi,\cdot) \odot E[\hat{r}(\cdot)] = P(\phi,\cdot) \odot r(\cdot) = r(\phi).$$

Furthermore, the covariance function of $\hat{r}_P(\phi)$ cannot exceed that of $\hat{r}(\phi)$

$$\mathrm{cov}_{\mathbf{r}}\left(\hat{r}(\phi), \hat{r}(\gamma)\right) - \mathrm{cov}_{\mathbf{r}}\left(\hat{r}_P(\phi), \hat{r}_P(\gamma)\right) \quad \text{is n.n.d.} \tag{45}$$

This is easily shown by substitution of the orthogonal decomposition

$$\hat{r}(\phi) - r(\phi) = P(\phi, \cdot) \odot [\hat{r}(\cdot) - r(\cdot)] + [\delta(\phi, \cdot) - P(\phi, \cdot)] \odot [\hat{r}(\cdot) - r(\cdot)].$$

into $\mathrm{cov}_{\mathbf{r}} (\hat{r}(\phi), \hat{r}(\gamma)) = E_{\mathbf{r}}[(\hat{r}(\phi) - r(\phi))(\hat{r}(\gamma) - r(\gamma))]$.

Now as the estimator $\hat{r}_P(\phi)$ lies in the span of $B_1, \dots B_K$ there exists a vector $\hat{\boldsymbol{\theta}}$ such that

$$\hat{r}_P(\phi) = \mathbf{B}^T(\phi)\hat{\boldsymbol{\theta}}.$$

Thus the covariance function of $\hat{r}_P(\phi)$ is

$$\mathrm{cov}_{\mathbf{r}} (\hat{r}_P(\phi), \hat{r}_P(\gamma)) = \mathbf{B}^T(\phi)\mathrm{cov}_{\boldsymbol{\theta}} \left(\hat{\boldsymbol{\theta}}\right) \mathbf{B}(\gamma) \tag{46}$$

where $\mathrm{cov}_{\boldsymbol{\theta}} \left(\hat{\boldsymbol{\theta}}\right)$ is the $K \times K$ covariance matrix of $\hat{\boldsymbol{\theta}}$. Furthermore, as $\hat{r}_P(\phi)$ is an unbiased estimator of the true $r(\phi)$, $\hat{\boldsymbol{\theta}}$ is an unbiased estimator of the true $\boldsymbol{\theta}$. This can be seen from the sequence of equalities

$$\mathbf{B}^T(\phi)\boldsymbol{\theta} = r(\phi) = E_{\boldsymbol{\theta}}[\mathbf{B}^T(\phi)\hat{\boldsymbol{\theta}}] = \mathbf{B}^T(\phi)E_{\boldsymbol{\theta}}[\hat{\boldsymbol{\theta}}]$$

so that $\mathbf{B}^T(\phi)[\boldsymbol{\theta} - E_{\boldsymbol{\theta}}[\hat{\boldsymbol{\theta}}]] = 0$. Hence

$$\left[\mathbf{B}(\cdot) \odot \mathbf{B}^T(\cdot)\right] \; [\boldsymbol{\theta} - E_{\boldsymbol{\theta}}[\hat{\boldsymbol{\theta}}]] = 0.$$

which implies that $E_{\boldsymbol{\theta}}[\hat{\boldsymbol{\theta}}] = \boldsymbol{\theta}$ as $\mathbf{B}(\cdot) \odot \mathbf{B}^T(\cdot)$ is a non-singular matrix.

Finally, as $\hat{\boldsymbol{\theta}}$ is unbiased, application of the CR bound to the covariance of $\hat{\boldsymbol{\theta}}$ gives: $\mathrm{cov}_{\boldsymbol{\theta}} \left(\hat{\boldsymbol{\theta}}\right) \geq \mathbf{F}_{\boldsymbol{\theta}}^{-1}$. Using this in (46) obtain

$$\mathrm{cov}_{\mathbf{r}} (\hat{r}_P(\phi), \hat{r}_P(\gamma)) - \mathbf{B}^T(\phi)\mathbf{F}_{\boldsymbol{\theta}}^{-1}\mathbf{B}(\gamma) \quad \text{is n.n.d..}$$

Therefore, using the above and (45):

$$\mathrm{cov}_{\mathbf{r}} (\hat{r}(\phi), \hat{r}(\gamma)) - \mathbf{B}^T(\phi)\mathbf{F}_{\boldsymbol{\theta}}^{-1}\mathbf{B}(\gamma) \quad \text{is n.n.d.,}$$

which is the form of the CR bound claimed in (44).

□

C. Asymptotic form of CR Bound

In this appendix the small α (high MRI spatial resolution) representation (13) is established for the boundary estimation CR bound (11). We state this result as the following

Lemma 2: Assume that for all ϕ: $r_{\boldsymbol\theta}(\phi) > 0$ and $r_{\boldsymbol\theta}(\phi)$, $r'_{\boldsymbol\theta}(\phi)$, $r''_{\boldsymbol\theta}(\phi)$ are bounded. Then

$$\mathrm{CRB}_{\boldsymbol\theta}(\phi,\gamma) = \frac{1}{2\sqrt{\pi}\sigma_s C_{\mathrm{CN}}}\,\mathbf{B}^T(\phi)\left[\int_{-\pi}^{\pi}\frac{r_{\boldsymbol\theta}^2}{\sqrt{1+\left(\frac{\partial}{\partial\psi}\ln r_{\boldsymbol\theta}(\psi)\right)^2}}\,\mathbf{B}(\psi)\mathbf{B}^T(\psi)d\psi\right]^{-1}\mathbf{B}(\gamma) + o(\alpha). \tag{47}$$

where

$$\alpha = \max_{\phi}\left\{\frac{\sqrt{2}\sigma_s}{\|\vec{r}'_{\boldsymbol\theta}(\phi)\|}\right\} = \frac{\sqrt{2}\sigma_s}{\min_{\phi}\sqrt{[r_{\boldsymbol\theta}(\phi)]^2 + [r'_{\boldsymbol\theta}(\phi)]^2}}. \tag{48}$$

Proof of Lemma 2

First, the following simple identitities are easily established using the definition $\vec{r}_{\boldsymbol\theta}(\phi) = [r_{\boldsymbol\theta}(\phi)\cos\phi, r_{\boldsymbol\theta}(\phi)\sin\phi]^T$

$$\begin{aligned}\|\vec{r}'_{\boldsymbol\theta}(\phi)\|^2 &= [r_{\boldsymbol\theta}(\phi)]^2 + [r'_{\boldsymbol\theta}(\phi)]^2\\ \|\vec{r}''_{\boldsymbol\theta}(\phi)\|^2 &= [r_{\boldsymbol\theta}(\phi)]^2 + [2r'_{\boldsymbol\theta}(\phi)]^2 - 2r_{\boldsymbol\theta}(\phi)r'_{\boldsymbol\theta}(\phi) + [r''_{\boldsymbol\theta}(\phi)]^2\end{aligned}$$

In particular the assumed boundedness of $r_{\boldsymbol\theta}, r'_{\boldsymbol\theta}$ and $r''_{\boldsymbol\theta}$ implies that $\|\vec{r}'_{\boldsymbol\theta}\|^2$ and $\|\vec{r}''_{\boldsymbol\theta}\|^2$ are also bounded.

The next step is to derive the following asymptotic form (9) for the FIM $\mathbf{F}_{\boldsymbol\theta}$.

$$\begin{aligned}\mathbf{F}_{\boldsymbol\theta} &= C_{\mathrm{CN}}\iint_{-\pi}^{\pi}\exp\left[-\frac{\|\vec{r}_{\boldsymbol\theta}(\phi)-\vec{r}_{\boldsymbol\theta}(\gamma)\|^2}{4\sigma_s^2}\right] r_{\boldsymbol\theta}(\phi)\, r_{\boldsymbol\theta}(\gamma)\,\mathbf{B}(\phi)\,\mathbf{B}^T(\gamma)\,d\phi\,d\gamma & (49)\\ &= 2C_{\mathrm{CN}}\sqrt{\pi}\sigma_s\int_{-\pi}^{\pi}\frac{r_{\boldsymbol\theta}^2(\gamma)}{\sqrt{[r_{\boldsymbol\theta}(\gamma)]^2+[r'_{\boldsymbol\theta}(\gamma)]^2}}\,\mathbf{B}(\gamma)\,\mathbf{B}^T(\gamma)\,d\gamma + o(\alpha). & (50)\end{aligned}$$

To show this start with the Taylor series with remainder for the vector $\vec{r}_{\boldsymbol\theta}(\phi) = \vec{r}_{\boldsymbol\theta}(\gamma) + \vec{r}'_{\boldsymbol\theta}(\gamma)(\phi-\gamma) + \frac{1}{2}\vec{r}''_{\boldsymbol\theta}(\xi)(\phi-\gamma)^2$ where ξ is a point in $(0, 2\pi]$, $\vec{r}'_{\boldsymbol\theta}(\gamma) = \partial\vec{r}_{\boldsymbol\theta}(\gamma)/\partial\gamma$ and $\vec{r}''_{\boldsymbol\theta}(\gamma) = \partial^2\vec{r}_{\boldsymbol\theta}(\gamma)/\partial\gamma^2$. This gives

$$\|\vec{r}_{\boldsymbol\theta}(\phi)-\vec{r}_{\boldsymbol\theta}(\gamma)\|^2 = \left\|\vec{r}'_{\boldsymbol\theta}(\gamma)\right\|^2(\phi-\gamma)^2 + \left\langle\vec{r}'_{\boldsymbol\theta}(\gamma),\ \vec{r}''_{\boldsymbol\theta}(\xi)\right\rangle(\phi-\gamma)^3 + \frac{1}{4}\|\vec{r}''_{\boldsymbol\theta}(\gamma)\|^2(\phi-\gamma)^4 \tag{51}$$

Now separate the integral of (50) over ϕ and γ to obtain the representation

$$\mathbf{F}_{\theta} = C_{\mathrm{CN}}\int_{-\pi}^{\pi}A(\gamma)r_{\theta}(\gamma)\,\mathbf{B}^T(\gamma)\,d\gamma, \tag{52}$$

where

$$A(\gamma) = \int_{-\pi}^{\pi}\exp\left[-\frac{\|\vec{r}_{\boldsymbol\theta}(\phi)-\vec{r}_{\boldsymbol\theta}(\gamma)\|^2}{4\sigma_s^2}\right] r_{\boldsymbol\theta}(\phi)\,\mathbf{B}(\phi)\,d\phi.$$

It will next be shown that

$$A(\gamma) = \sqrt{2\pi}\alpha(\gamma)r_{\theta}(\gamma)\,\mathbf{B}(\gamma) + o(\alpha). \tag{53}$$

where $\alpha(\gamma)$ is defined as

$$\alpha(\gamma) = \frac{\sqrt{2}\sigma_s}{\|\vec{r}^{\,\prime}_{\boldsymbol{\theta}}(\gamma)\|}. \tag{54}$$

Substitute (51) into the exponent of the integrand of $A(\gamma)$, and make a change of variable $u = (\phi - \gamma)/\alpha(\gamma)$ to obtain the representation

$$\begin{aligned} A(\gamma) &= \sqrt{2\pi}\alpha(\gamma) \int_{-\pi}^{\pi} h(\phi) \frac{1}{\sqrt{2\pi}\alpha(\gamma)}\, e^{-(\phi-\gamma)^2/(2\alpha^2(\gamma))} d\phi \\ &= \sqrt{2\pi}\alpha(\gamma) \int_{-\pi/\alpha(\gamma)}^{\pi/\alpha(\gamma)} h(\alpha u + \gamma) \frac{1}{\sqrt{2\pi}}\, e^{-u^2/2} du \\ &= \sqrt{2\pi}\alpha(\gamma)\, E\left[h(\alpha U + \gamma) I_A(U)\right] \end{aligned} \tag{55}$$

where U is a standard zero mean unit variance Gaussian random variable and $I_A(u)$ is the indicator function of the interval $A = [-\pi/\alpha(\gamma),\ \pi/\alpha(\gamma)]$.

$$h(\phi) \;=\; \exp\left(- \frac{\left\langle \vec{r}^{\,\prime}_{\boldsymbol{\theta}}(\gamma),\ \vec{r}^{\,\prime\prime}_{\boldsymbol{\theta}}(\xi) \right\rangle (\phi-\gamma)^3 + \frac{1}{4}\|\vec{r}^{\,\prime\prime}_{\boldsymbol{\theta}}(\gamma)\|^2 (\phi-\gamma)^4}{4\sigma_s^2} \right)\, r_{\boldsymbol{\theta}}(\phi)\, \mathbf{B}(\phi)$$

and using the definition (54) of $\alpha(\gamma)$

$$h(\alpha(\gamma)u + \gamma) \;=\; \exp\left(-\frac{1}{2}\left[\frac{\left\langle \vec{r}^{\,\prime}_{\boldsymbol{\theta}}(\gamma),\ \vec{r}^{\,\prime\prime}_{\boldsymbol{\theta}}(\xi) \right\rangle}{\|\vec{r}^{\,\prime}_{\boldsymbol{\theta}}(\gamma)\|^2}\, \alpha(\gamma) u^3 + \frac{\|\vec{r}^{\,\prime\prime}_{\boldsymbol{\theta}}(\gamma)\|^2}{8\|\vec{r}^{\,\prime}_{\boldsymbol{\theta}}(\gamma)\|^2}\, \alpha(\gamma)^2 u^4 \right] \right)\, r_{\boldsymbol{\theta}}(\alpha(\gamma)u+\gamma)\, \mathbf{B}(\alpha(\gamma)u+\gamma)$$

Since the second derivative of the function $h(\cdot)$ is bounded, expansion of the LHS of the previous equation in a Taylor series with remainder about $u = 0$ and substitution into (55) yields

$$\begin{aligned} E\left[h(\alpha(\gamma)U + \gamma) I_A(U)\right] &= h(\alpha(\gamma)u+\gamma)|_{u=0} \;+\; \alpha h^{'}(\alpha(\gamma)u+\gamma)|_{u=0}\, E[U I_A(U)] \\ &\qquad + \frac{\alpha^2(\gamma)}{2} h^{''}(\alpha(\gamma)\eta + \gamma)\, E[U^2 I_A(U)] \\ &= h(\alpha(\gamma)u+\gamma)|_{u=0} \;+\; o(\alpha) \end{aligned}$$

where $\alpha = \max_\gamma \alpha(\gamma)$ as given in (48), and

$$\begin{aligned} |o(\alpha)| &= \left| \frac{\alpha^2(\gamma)}{2} h^{''}(\alpha(\gamma)\eta+\gamma) E[U^2 I_A(U)] \right| \\ &\le \frac{\alpha^2}{2} \max_\phi |h^{''}(\phi)|. \end{aligned}$$

where η is a point in $[-\pi/\alpha(\gamma), \pi/\alpha(\gamma))]$.

Since $h(\alpha(\gamma)u + \gamma)|_{u=0} = r_{\boldsymbol{\theta}}(\gamma)\ \mathbf{B}(\gamma)$ this establishes relation (53). Plugging (53) into (52) one obtains (40). Using the identity $\frac{\partial}{\partial\phi} \ln r(\phi) = r^{'}(\phi)/r(\phi)$ and the continuity of matrix inverses for invertible matrices: $[Q + o(\alpha)]^{-1} = Q^{-1} + o(\alpha)$, finishes the proof.

$\square$

D. Proof of Proposition 1

First, note that for a fixed function $\hat{\boldsymbol{\lambda}}$ the maximization over $\boldsymbol{\theta} \in \mathbb{R}^K$ in (21) is equivalent to maximization over probability distributions dP on $\mathbb{R}^K$

$$\max_{\boldsymbol{\theta}} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})) = \max_{dP} \int_{\mathbf{R}^K} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})) dP(\boldsymbol{\theta})$$

To see this, define $g(\boldsymbol{\theta}) = P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta}))$ and let $d\delta(\boldsymbol{\theta}_1)$ denote the probability distribution which assigns probability one to a single point $\boldsymbol{\theta}_1 \in \mathbb{R}^K$. Now, since the space of probability distributions $\{dP\}$ on $\mathbb{R}^K$ contains the space of single point distributions $\{d\delta\}$ on $\mathbb{R}^K$ we have

$$\max_{\boldsymbol{\theta}} g(\boldsymbol{\theta}) = \max_{d\delta} \int_{\mathbf{R}^K} g(\boldsymbol{\theta}) d\delta(\boldsymbol{\theta}) \leq \max_{dP} \int_{\mathbf{R}^K} g(\boldsymbol{\theta}) dP(\boldsymbol{\theta}).$$

On the other hand, as $\int_{\mathbf{R}^K} dP(\boldsymbol{\theta}) = 1$,

$$\max_{dP} \int_{\mathbf{R}^K} g(\boldsymbol{\theta}) dP(\boldsymbol{\theta}) \leq \max_{\boldsymbol{\theta}} g(\boldsymbol{\theta}) \int dP(\boldsymbol{\theta}) = \max_{\boldsymbol{\theta}} g(\boldsymbol{\theta}).$$

Hence $\max_{\boldsymbol{\theta}} g(\boldsymbol{\theta}) = \max_{dP} \int_{\mathbf{R}^K} g(\boldsymbol{\theta}) dP(\boldsymbol{\theta})$ and (21) is equivalent to

$$\min_{\hat{\boldsymbol{\lambda}}} \max_{\boldsymbol{\theta}} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})) = \min_{\hat{\boldsymbol{\lambda}}} \max_{dP} \int_{\mathbf{R}^K} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})) dP(\boldsymbol{\theta})$$

Second, under the assumptions of [39, Thm. 2.9.2] the solution to the minimax problem occurs at a saddlepoint so that

$$\min_{\hat{\boldsymbol{\lambda}}} \max_{dP} \int_{\mathbf{R}^K} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})) dP(\boldsymbol{\theta}) \quad = \quad \max_{dP} \min_{\hat{\boldsymbol{\lambda}}} \int_{\mathbf{R}^K} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})) dP(\boldsymbol{\theta}).$$

Third, for fixed dP the integral on the right hand side of the above equation is simply the average probability of error, which by Fubini, has the equivalent form

$$\int_{\mathbf{R}^K} P_e(\hat{\boldsymbol{\lambda}}, \boldsymbol{w}(\boldsymbol{\theta})) dP(\boldsymbol{\theta}) = \int d\mathbf{Y} \int_{\|\hat{\boldsymbol{\lambda}}(\mathbf{Y}) - \boldsymbol{\lambda}\| > \epsilon} d\boldsymbol{\lambda} \int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP(\boldsymbol{\theta}).$$

As $\int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP(\boldsymbol{\theta}) \geq 0$ the above expression is minimized when $\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\mathbf{Y})$ is selected as the midpoint of the sphere of radius ϵ which maximizes the integral

$$\int_{B(\hat{\boldsymbol{\lambda}}, \epsilon)} d\boldsymbol{\lambda} \int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP(\boldsymbol{\theta})$$

[40, Corollary 4.1.1]. From the assumed continuity of $f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda})$ it follows that this integral is continuous and therefore by the mean value theorem

$$\max_{\hat{\boldsymbol{\lambda}}} \int_{B(\hat{\boldsymbol{\lambda}}, \epsilon)} d\boldsymbol{\lambda} \int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP(\boldsymbol{\theta}) = \max_{\hat{\lambda}} \left\{ \int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP(\boldsymbol{\theta}) \right\} V_K \, \epsilon^K + o(\epsilon^K) \qquad (56)$$

where V_K denotes the volume of the unit sphere in $\mathbb{R}^K$. Since the expression in brackets $\{\ \}$ has a strict maximum and is continuous in $\boldsymbol{\lambda}$ the solution $\hat{\boldsymbol{\lambda}}$ to (56) is

$$\text{argmax}_{\hat{\boldsymbol{\lambda}}} \left\{ \int_{\mathbf{R}^K} f(\mathbf{Y}|\boldsymbol{\lambda}) f_{\boldsymbol{w}(\boldsymbol{\theta})}(\boldsymbol{\lambda}) dP(\boldsymbol{\theta}) \right\} + O(\epsilon)$$

as claimed. This establishes the proposition.

$\square$

E. Proof of Proposition 2

Only a sketch of the salient parts of the proof is given. We eliminate the index n in the following for simplicity. Define the conditional expectation operator $E_{\boldsymbol{\theta}|\mathbf{Y}_\text{M}}[Z] = \int_{\mathbf{R}^K} Z\ dP^*(\boldsymbol{\theta}|\mathbf{Y}_\text{M})$. Then we have from (28)

$$\boldsymbol{\Phi}_{\text{minimax}}(\boldsymbol{\lambda}, \boldsymbol{w}) = \ln f(\mathbf{Y}_\text{E}|\boldsymbol{\lambda}) + \ln E_{\boldsymbol{\theta}|\mathbf{Y}_\text{M}} \left[\exp\left\{-\beta R(\boldsymbol{\lambda}; \boldsymbol{\theta}) + \ln \rho(\boldsymbol{w}(\boldsymbol{\theta}))\right\}\right].$$

Since the exponential function is convex (equivalently the log function is concave) Jensen's inequality applied to the RHS of the above equation gives

$$\begin{aligned} \boldsymbol{\Phi}_{\text{minimax}}(\boldsymbol{\lambda}, \boldsymbol{w}) &\geq \ln f(\mathbf{Y}_\text{E}|\boldsymbol{\lambda}) - \beta\ E_{\boldsymbol{\theta}|\mathbf{Y}_\text{M}} \left[R(\boldsymbol{\lambda}; \boldsymbol{\theta})\right] + C \\ &= \ln f(\mathbf{Y}_\text{E}|\boldsymbol{\lambda}) - \beta \sum_{j=1}^{P} \sum_{k \in N_j} \tilde{w}_{jk}(\hat{\boldsymbol{\theta}})(\lambda_j - \lambda_k)^2 + C, \end{aligned} \tag{57}$$

where

$$C = E_{\boldsymbol{\theta}|\mathbf{Y}_\text{M}} \left[\ln \rho(\boldsymbol{w}(\boldsymbol{\theta}))\right]. \tag{58}$$

Hence $\boldsymbol{\Phi}_{\text{minimax}}(\boldsymbol{\lambda}; \boldsymbol{w}) \geq \boldsymbol{\Phi}_{\text{PML}}(\boldsymbol{\lambda}; \tilde{\boldsymbol{w}}) + C$, which establishes the lower inequality in (33).

We next show the upper inequality in (33). Under the assumption that $dP(\boldsymbol{\theta}|\mathbf{Y}_\text{M})$ is smooth, we can apply Theorem 1. Assertion 2 of Theorem 1 implies that for integrable $h(\boldsymbol{\theta})$ and any Borel subset V of $\mathbb{R}^K$ the Lebesgue-Steiltjes integral $\int_V h(\boldsymbol{\theta})\ dP(\boldsymbol{\theta}|\mathbf{Y}_\text{M})$ converges in probability to $\int_V h(\boldsymbol{\theta})\ dG(\boldsymbol{\theta}|\boldsymbol{T}, \boldsymbol{\Gamma})$, where dG is the Gaussian probability measure (29) cited in Theorem 1. On the other hand, we have by Assertion 1 of Theorem 1 that for any open ball B of radius ϵ centered at the true parameter $\boldsymbol{\theta}_o$,

$$\left| \int_{\mathbf{R}^K} h(\boldsymbol{\theta})\ dP(\boldsymbol{\theta}|Ym) - \int_B h(\boldsymbol{\theta})\ dP(\boldsymbol{\theta}|\mathbf{Y}_\text{M}) \right| \leq \Delta \tag{59}$$

where Δ converges to zero in probability. Now specializing to $h(\boldsymbol{\theta}) = \rho(\boldsymbol{w}(\boldsymbol{\theta})) \cdot \exp(-\beta R(\boldsymbol{\lambda}; \boldsymbol{\theta}))$, since ϵ is arbitrary and $\exp(\cdot)$ is a continuous function, we have

$$\left| \int_B h(\boldsymbol{\theta})\ dP(\boldsymbol{\theta}|\mathbf{Y}_\text{M}) - \exp\left(-\beta \int_B R(\boldsymbol{\lambda}; \boldsymbol{\theta})\ dP(\boldsymbol{\theta}|\mathbf{Y}_\text{M})\ +\ \int_B \ln \rho(\boldsymbol{w}(\boldsymbol{\theta}))\ dP(\boldsymbol{\theta}|\mathbf{Y}_\text{M})\right) \right| \leq \delta, \tag{60}$$

where δ goes to zero as ϵ goes to zero. Combining (59), (60) and the definition (28) of $\mathbf{\Phi}_{\text{minimax}}(\boldsymbol{\lambda}, \boldsymbol{w})$

$$\mathbf{\Phi}_{\text{minimax}}(\boldsymbol{\lambda}, \boldsymbol{w}) \quad \leq \quad \ln f(\mathbf{Y}_{\text{E}}|\boldsymbol{\lambda}) - \beta \int_{\mathbf{R}^K} R(\boldsymbol{\lambda}; \boldsymbol{\theta})\, dP(\boldsymbol{\theta}|\mathbf{Y}_{\text{M}}) + C + \gamma$$

where C is given in (58) and $\gamma > 0$ converges to zero in probability. From (14) and (15) the first two terms on the right hand side of the inequality can be recognized as $\mathbf{\Phi}_{\text{PML}}(\boldsymbol{\lambda}, \tilde{\boldsymbol{w}})$, which finishes the proof.

□

F. Derivation of Observed Fisher Information

The *observed* Fisher information [20] $\hat{\mathbf{F}}_{\hat{\theta}}$ is given by $- \left.\nabla^2_{\boldsymbol{\theta}} \{\ln f(\mathbf{Y}_{\text{M}}; \boldsymbol{\theta})\}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$. From the Hessian of the loglikelihood as in (37)

$$\begin{aligned}\hat{\mathbf{F}}_{\hat{\boldsymbol{\theta}}} \quad &= \quad \frac{1}{\sigma_n^2} \iint\limits_{R_f} \nabla^T_{\boldsymbol{\theta}} \mathrm{I}^s_{\boldsymbol{\theta}}(x,y) \cdot \nabla_{\boldsymbol{\theta}} \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\, dx\, dy \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ &\qquad - \frac{1}{\sigma_n^2} \iint\limits_{R_f} [\mathbf{Y}_{\text{M}}(x,y) - \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)]\, \nabla^2_{\boldsymbol{\theta}} \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\, dx\, dy \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}\end{aligned} \tag{61}$$

The first term on the right hand side of the above equation can be identified from (38) as the *true* Fisher information evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. We refer to this as the *expected* Fisher information. Therefore

$$\hat{\mathbf{F}}_{\hat{\boldsymbol{\theta}}} = \mathbf{F}_{\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} - \frac{1}{\sigma_n^2} \iint\limits_{R_f} [\mathbf{Y}_{\text{M}}(x,y) - \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)]\ \nabla^2_{\boldsymbol{\theta}} \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\, dx\, dy \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{62}$$

Now define $\hat{\varepsilon}(x, y, \hat{\boldsymbol{\theta}})$ to be an "estimate" of the noise term as follows

$$\hat{\varepsilon}(x,y,\hat{\boldsymbol{\theta}}) = \mathbf{Y}_{\text{M}}(x,y) - \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

then the observed Fisher information can be written in terms of expected Fisher information as

$$\hat{\mathbf{F}}_{\hat{\boldsymbol{\theta}}} = \mathbf{F}_{\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} - \frac{1}{\sigma_n^2} \iint\limits_{R_f} \hat{\varepsilon}(x,y,\hat{\boldsymbol{\theta}})\ \nabla^2_{\boldsymbol{\theta}} \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)\, dx\, dy \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \tag{63}$$

Since $r_{\boldsymbol{\theta}}(\phi) = \mathbf{B}^T(\phi)\boldsymbol{\theta}$ is linear in $\boldsymbol{\theta}$ its second derivative is equal to zero. Thus

$$\nabla^2_{\boldsymbol{\theta}} \mathrm{I}^s_{\boldsymbol{\theta}}(x,y) = C_s \int_{-\pi}^{\pi} \mathcal{G}(x,y,\theta,\phi)\, d\phi$$

where

$$\begin{aligned}[\mathcal{G}]_{i,j}(x,y,\theta,\phi) \quad &= \quad B_i(\phi)B_j(\phi) \left[1 + \frac{r_{\boldsymbol{\theta}}(\phi)(x\cos\phi + y\sin\phi) - r^2_{\boldsymbol{\theta}}(\phi)}{\sigma_s^2}\right] \\ &\qquad \cdot \exp\left[-\frac{(x - r_{\boldsymbol{\theta}}(\phi)\cos\phi)^2 + (y - r_{\boldsymbol{\theta}}(\phi)\sin\phi)^2}{2\sigma_s^2}\right], \qquad (i,j = 1,\ldots,K).\end{aligned}$$

Substituting the above in (63),

$$\left[\hat{\mathbf{F}}_{\hat{\boldsymbol{\theta}}}\right]_{i,j} = \left[\mathbf{F}_{\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}\right]_{i,j} - \frac{C_s}{\sigma_n^2}\int_{-\pi}^{\pi} B_i(\phi)B_j(\phi) \iint\limits_{R_f} \hat{\varepsilon}(x,y,\boldsymbol{\theta}) \cdot \exp\left[-\frac{\|\boldsymbol{x}-\vec{r}_{\boldsymbol{\theta}}\|^2}{2\sigma_s^2}\right] \cdot \left[1 + \frac{r_{\boldsymbol{\theta}}(\phi)(x\cos\phi + y\sin\phi) - r_{\boldsymbol{\theta}}^2(\phi)}{\sigma_s^2}\right] \, dx\, dy\, d\phi\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{64}$$

where

$$\|\boldsymbol{x}-\vec{r}_{\boldsymbol{\theta}}\|^2 = (x - r_{\boldsymbol{\theta}}(\phi)\cos\phi)^2 + (y - r_{\boldsymbol{\theta}}(\phi)\sin\phi)^2.$$

We can then write this as

$$\left[\hat{\mathbf{F}}_{\hat{\boldsymbol{\theta}}}\right]_{i,j} = \left[\mathbf{F}_{\hat{\boldsymbol{\theta}}}\right]_{i,j} - \Delta\mathbf{F}_{ij}$$

where

$$\Delta\mathbf{F}_{ij} = \frac{C_s}{\sigma_n^2}\int_{-\pi}^{\pi} B_i(\phi)B_j(\phi) \iint\limits_{R_f} (\mathbf{Y}_{\mathrm{M}}(x,y) - \mathrm{I}^s_{\boldsymbol{\theta}}(x,y)) \cdot \exp\left[-\frac{\|\boldsymbol{x}-\vec{r}_{\boldsymbol{\theta}}\|^2}{2\sigma_s^2}\right] \cdot \left[1 + \frac{r_{\boldsymbol{\theta}}(\phi)(x\cos\phi + y\sin\phi) - r_{\boldsymbol{\theta}}^2(\phi)}{\sigma_s^2}\right] \, dx\, dy\, d\phi\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}.$$

References

[1] G. Gindi, M. Lee, A. Rangarajan, and G. Zubal, "Bayesian reconstruction of functional images using anatomical information as priors," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 670–680, December 1993.

[2] K. M. Hanson, "Bayesian reconstruction based on flexible priors," *J. Opt. Soc. Am.*, vol. 10, no. 5, pp. 997–1004, May 1993.

[3] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Transactions on Medical Imaging*, vol. 8, no. 2, pp. 194–203, 1989.

[4] V. E. Johnson, "A model for segmentation and analysis of noisy images," *Journal of the American Statistical Association*, vol. 89, no. 425, pp. 230–241, mar 1994.

[5] S. J. Lee, G. R. Gindi, I. G. Zubal, and A. Rangarajan, "Using ground-truth data to design priors in Bayesian SPECT reconstruction," in *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot,

and R. D. Paola, editors, Kluwer, 1995.

[6] Y. Zhang, J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Joint estimation for incorporating MRI anatomic images into SPECT reconstruction," in *Proc. of IEEE Nuclear Science Symposium*, volume 3, pp. 1256–1260, 1994.

[7] H. W. Müller-Gärtner, J. M. Links, J. L. Prince, R. N. Bryan, E. McVeigh, J. P. Leal, C. Davatzikos, and J. J. Frost, "Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects," *J. Cerebral Blood Flow and Metabolism*, vol. 12, pp. 571–83, 1992.

[8] O. Rousset, Y. Ma, M. Kamber, and A. C. Evans, "3D simulations of radiotracer uptake in deep nuclei of human brain," *Computerized Medical Imaging and Graphics*, vol. 17, no. 4-5, pp. 373–379, July-October 1993.

[9] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, November 1984.

[10] V. E. Johnson, J. Bowsher, R. Jaszak, and T. Turkington, "Analysis and reconstruction of medical images using prior information," Technical report, Institute of statistics and decision sciences, Duke University, 1993.

[11] B. Lipinski, H. Herzog, E. Rota Kops, W. Oberschelp, and H. W. Müller-Gärtner, "MR guided PET reconstruction and problems with anatomical misinformation," *NeuroImage*, vol. 2, no. 2, pp. S31, June 1995.

[12] J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Regularized emission image reconstruction using imperfect side information," *IEEE Transactions on Nuclear Science*, vol. 39, no. 5, pp. 1464–1471, Oct. 1992.

[13] X. Ouyang, W. H. Wong, V. E. Johnson, X. Hu, and C. T. Chen, "Incorporation of correlated structural images in PET image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp.

627–640, December 1994.

[14] R. Leahy and X. H. Yan, "Incorporation of anatomical MR data for improved functional imaging with PET," in *Information Processing in Medical Imaging*, A. C. F. Colchester and D. J. Hawkes, editors, pp. 105–120, Springer-Verlag, New York, 1991.

[15] X. H. Yan and R. Leahy, "Map image reconstruction using intensity and line processes for emission tomography data," in *Proceedings of SPIE*, pp. 158–169, San Jose, CA, 1991.

[16] Y. Zhang, J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Incorporating MRI region information into SPECT reconstruction using joint estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, volume 4, pp. 2307–2310, 1995.

[17] B. Lipinski, H. Herzog, E. Rota Kops, W. Oberschelp, and H. W. Müller-Gärtner, "Expectaion maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information," *IEEE Transactions on Medical Imaging*, vol. MI-16, no. 2, pp. 129–136, April 1997.

[18] R. A. Fisher, "Theory of statistical estimation," *Proc. Cambridge Philosophical Society*, vol. 22, pp. 700–725, 1925.

[19] J. A. Fessler and A. Hero, "Penalized maximum likelihood image reconstruction using space alternating generalized EM algorithms," *IEEE Transactions on Image Processing*, vol. 4, no. 10, pp. 1417–1429, October 1995.

[20] B. Efron and D. V. Hinkley, "Assessing the accuracy of the maximum likelihood estimator: observed versus expected Fisher information," *Biometrika*, vol. 65, no. 3, pp. 457–487, 1978.

[21] E. R. Dougherty, editor, *Digital image processing methods*, Marcel Dekker, Inc., 1994.

[22] J. Canny, "A computational approach to edge detection," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. COM-26, pp. 297–307, 1977.

[23] D. Marr and E. C. Hildreth, "Theory of edge detection," *Proc. Royal Soc. Lond.*, vol. 207, pp. 187–217,

1980.

[24] M. Bret, *Image Synthesis*, Kluwer Academic Publishers, 1992.

[25] C. Chui, *Multivariate splines*, SIAM, CBMS-NSF series in applied mathematics, no. 54, 1988.

[26] W. A. Edelstein, G. H. Glover, C. J. Hardy, and R. W. Redington, "Intrinsic signal to noise ratio in NMR imaging," *Magnetic Resonance in Medicine*, vol. 3, pp. 606–618, 1986.

[27] S. R. Titus, *Improved penalized likelihood reconstruction of anatomically correlated emission computed tomography data*, PhD thesis, The University of Michigan, Ann Arbor, December 1996.

[28] I. A. Ibragimov and R. Z. Has'minskii, *Statistical estimation: Asymptotic theory*, Springer-Verlag, New York, 1981.

[29] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Transactions on Information Theory*, vol. IT-36, pp. 1285–1301, Nov. 1990.

[30] R. Piramuthu and A. O. Hero, "Theory and implementation of minmax ECT image reconstruction with MRI side information," Technical report, Comm. and Sig. Proc. Lab. (CSPL), Dept. EECS, University of Michigan, Ann Arbor, August 1998.

[31] T. F. Budinger and G. T. Gullberg, "Three dimensional reconstruction in nuclear medicine emission imaging," *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, pp. 2–20, 1974.

[32] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*, IEEE press, 1988.

[33] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. MI-1, No. 2, pp. 113–122, Oct. 1982.

[34] K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing," *IEEE Transactions on Medical Imaging*, vol. 9, no. 4, pp. 439–446, Dec. 1990.

[35] E. Levitan and G. T. Herman, "Maximum a posteriori probability expectation maximization algorithm for image reconstruction emission tomography," *IEEE Transactions on Medical Imaging*, vol. MI-6, no. 3, pp. 185–192, 1987.

[36] P. J. Green, "Bayesian reconstructions from emission tomography using a modified EM algorithm," *IEEE Transactions on Medical Imaging*, vol. 11, no. 1, pp. 81–90, March 1990.

[37] H. L. Van-Trees, *Detection, Estimation, and Modulation Theory: Part I*, Wiley, New York, 1968.

[38] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1988.

[39] T. S. Ferguson, *Mathematical Statistics - A Decision Theoretic Approach*, Academic Press, Orlando FL, 1967.

[40] E. L. Lehmann, *Theory of Point Estimation*, Wiley, New York, 1983.

[41] B. Baygun and A. O. Hero, "Optimal simultaneous detection and estimation under a false alarm constraint," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 688–703, 1995.

[42] B. Baygun and A. O. Hero, "An iterative solution to the min-max simultaneous detection and estimation problem," in *Proc. of the IEEE Workshop on Statistical Signal and Array Processing*, pp. 8–11, Corfu, Greece, June 1996.

[43] P. S. Laplace, *Théorie analytique des probabilités*, Paris, Imprimerie royale, 1847.

[44] L. LeCam, *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York, 1986.

[45] L. Tierney and J. B. Kadane, "Accurate approximations for posterior moments and marginal densities," *J. Am. Statist. Assoc.*, vol. 81, pp. 82–86, 1986.

[46] R. Elliot, *Stochastic Calculus and Applications*, Springer-Verlag, New York N.Y., 1982.

[47] T. L. Marzetta, "A simple derivation of the constrained multiple parameter Cramèr-Rao bound," *Signal*

Processing, vol. SP-6, pp. 2247–2249, June 1993.

[48] P. Stoica and B. C. Ng, "On the Cramèr-Rao lower bound under parametric constraints," *Signal Processing Letters*, vol. SPL-5, pp. 177–179, July 1998.

Biography of the authors

Alfred O. Hero III

Alfred O. Hero III, was born in Boston, MA. in 1955. He received the B.S. in Electrical Engineering (summa cum laude) from Boston University (1980) and the Ph.D from Princeton University (1984), both in Electrical Engineering. While at Princeton he held the G.V.N. Lothrop Fellowship in Engineering. Since 1984 he has been with the of Dept. of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, where he is currently Professor and Director of the Communications and Signal Processing Laboratory. He has held positions of Visiting Scientist at M.I.T. Lincoln Laboratory (1987 - 1989), Visiting Professor at Ecole Nationale de Techniques Avancees (ENSTA) and Ecole Superieure d'Electricite, Paris, France (1990), and William Clay Ford Fellow at Ford Motor Company (1993). He has been a consultant to a wide spectrum of industries including Echlin International Corp., EG&G, MIT Lincoln Laboratory, Irwin Magnetics, and Ford Motor Co. His research has been supported by NIH, NSF, AFOSR, NSA, ARO and ONR in the area of estimation and detection, statistical communications, signal processing and image processing. Alfred Hero is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), a member of Tau Beta Pi, the American Statistical Association (ASA) , and Commission C of the International Union of Radio Science (URSI) . He has served as Associate Editor for the IEEE Transactions on Information Theory and Chairman of the Statistical Signal and Array Processing (SSAP) Technical Committee. He is currently treasurer of the Conference Board of the IEEE Signal Processing Society . He was Chairman for Publicity for the 1986 IEEE International Symposium on Information Theory. He was General Chairman for the 1995 IEEE International Conference on Acoustics, Speech, and Signal Processing, held in Detroit, Michigan. He is co-chair for the 1999 IEEE Information Theory Workshop and the 1999 IEEE Workshop on Higher Order Statistics .

Robinson Piramuthu

Robinson Piramuthu received his B.E. degree in Electrical Engineering from Regional Engineering College – Trichy, India in 1994. He received his M.S. degree in 1996 from the University of Florida – Gainesville with a major in Control Systems. He is currently pursuing his Ph.D. in the University of Michigan – Ann Arbor with a major in Signal Processing. He is a member of the signal processing society of the IEEE.

Jeffrey A. Fessler

Jeff Fessler received the BSEE degree from Purdue University in 1985, the MSEE degree from Stanford University in 1986, and the M.S. degree in Statistics from Stanford University in 1989. From 1985 to 1988 he was a National Science Foundation Graduate Fellow at Stanford, where he earned a Ph.D. in electrical engineering in 1990. He has worked at the University of Michigan since then. From 1991 to 1992 he was a Department of Energy Alexander Hollaender Post-Doctoral Fellow in the Division of Nuclear Medicine. From 1993 to 1995 he was an Assistant Professor in Nuclear Medicine and the Bioengineering Program. Since 1995 he has been with the Department of Electrical Engineering and Computer Science, where he is an Associate Professor. He received the Francois Erbsmann award for his IPMI93 presentation. He serves as an associate editor for IEEE Transactions on Medical Imaging and IEEE Transactions on Image Processing, and was co-chair of the 1997 SPIE conference on Image Reconstruction and Restoration. His research interests are in statistical aspects of imaging problems.

Steven R. Titus

To be submitted later.

Captions

Fig. 1. Radionuclide tracer concentration estimation with and without anatomical side information. Here $\boldsymbol{\theta}$ denotes an arbitrary parameterization of the anatomical boundary.

Fig. 2. Ellipsoidal cold spot phantom with a contrast of 66.67%.

Fig. 3. Representation of a two-level MRI image.

Fig. 4. (a) Bias of ML and EF boundary estimates vs. the MRI noise level. (b) Performance of the ML and EF boundary extraction methods relative to constrained CR bound vs MRI noise level. Vertical error bars shown cover $\pm$ one standard deviation.

Fig. 5. Second order neighborhood for a 2-D discrete image. The neighbors (black) of the center pixel (grey) are labeled according to their relation to the center, with D for down, U for up, L for left and R for right.

Fig. 6. For each of these three types of weights namely, (a) ideal, (b) unsmoothed and (c) smoothed weights, four images are given corresponding to $w_{j,k}$ for $j \equiv (x_j, y_j)$ and the four cases: $k = (x_{j-1}, y_j)$ (left), $k = (x_j, y_{j+1})$ (up), $k = (x_{j+1}, y_{j+1})$ (up-right), and $k = (x_{j-1}, y_{j+1})$ (up-left), respectively. In each figure a location where the pixel is black indicates that the weight is zero, and a location where the pixel is white denotes a weight value of one. The MRI additive noise parameter is $\sigma_n = 0.5$ in Figs. 6.b and 6.c.

Fig. 7. Reconstructed images using (a) FBP with Hamming window, (b) uniform weights, (c) ideal weights, (d) unsmoothed weights and (e) smoothed weights, for the same realization of MRI and PET data with MRI noise $\sigma_n = 0.36$ (6% of contrast). The PML-SAGE3 method with $\beta = 1$ was used for (c-e).

Fig. 8. Average RMS reconstruction error for various emission reconstruction methods.

Fig. 9. % Std Vs % Bias of total uptake estimate for MRI noise $\sigma_n = 0.36$.

Fig. 10. (a) % Bias of total uptake estimate Vs MRI noise σ_n. (b) % Std of total uptake estimate Vs MRI noise σ_n.

Figures

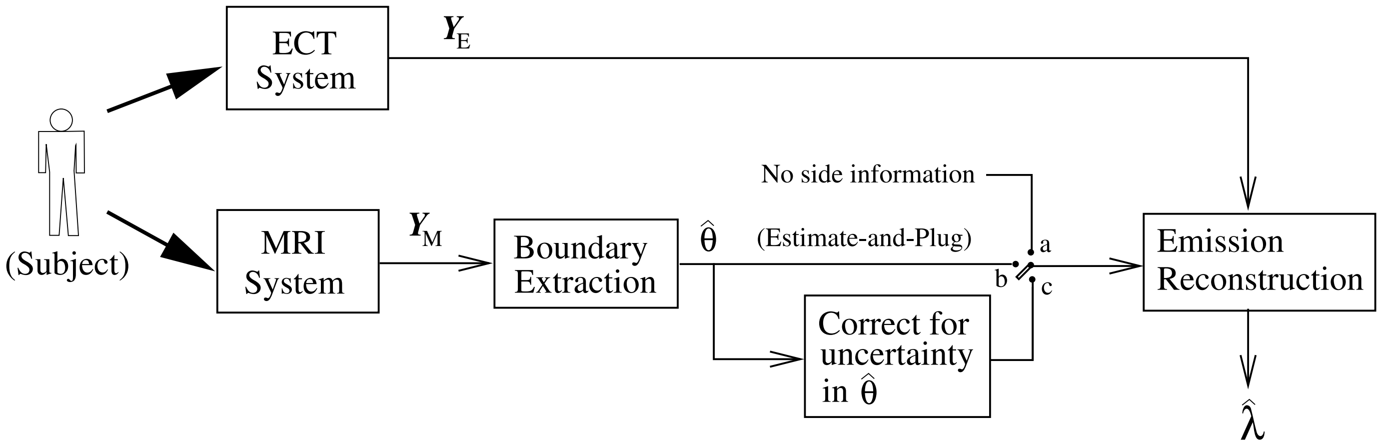


Fig. 1.

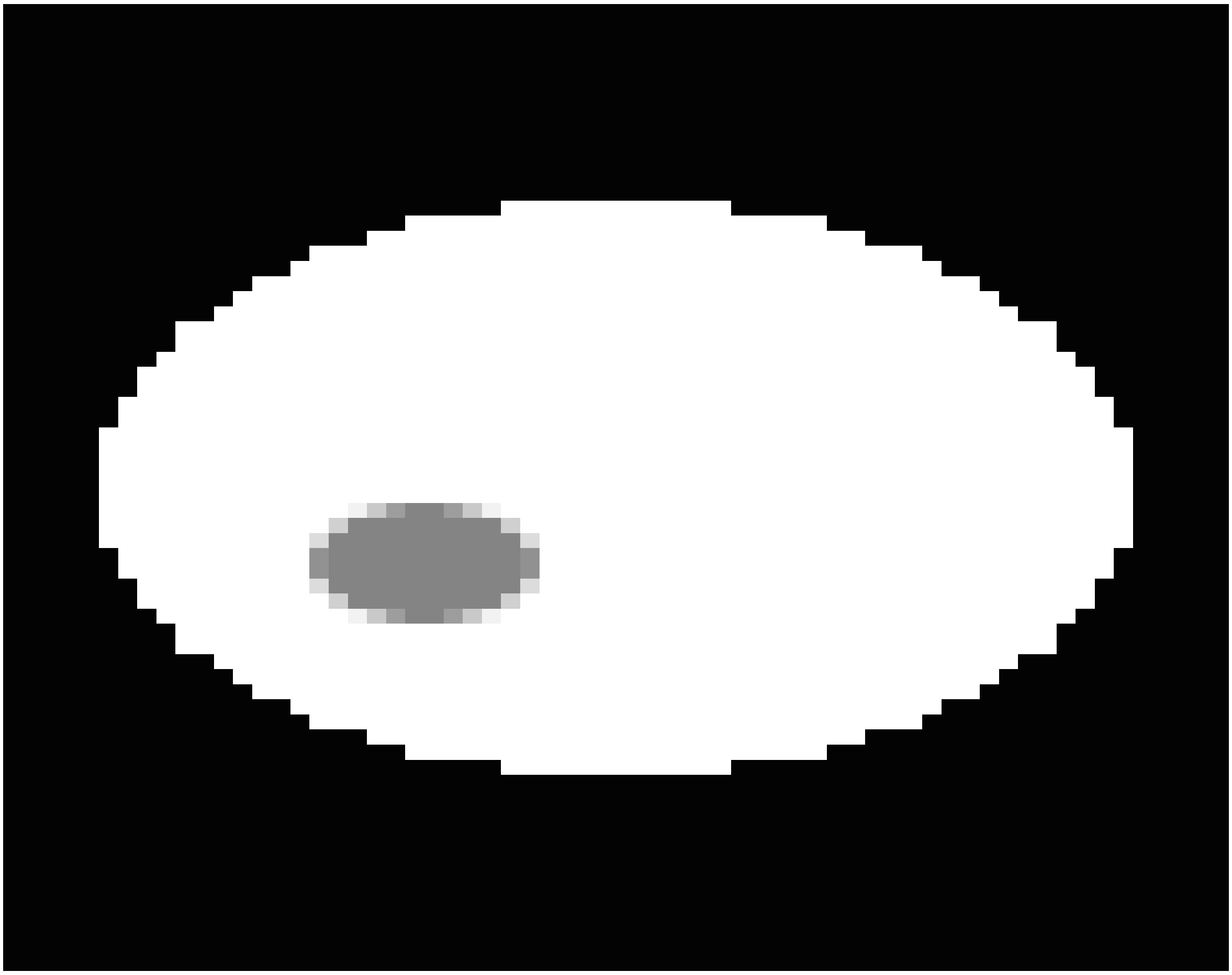
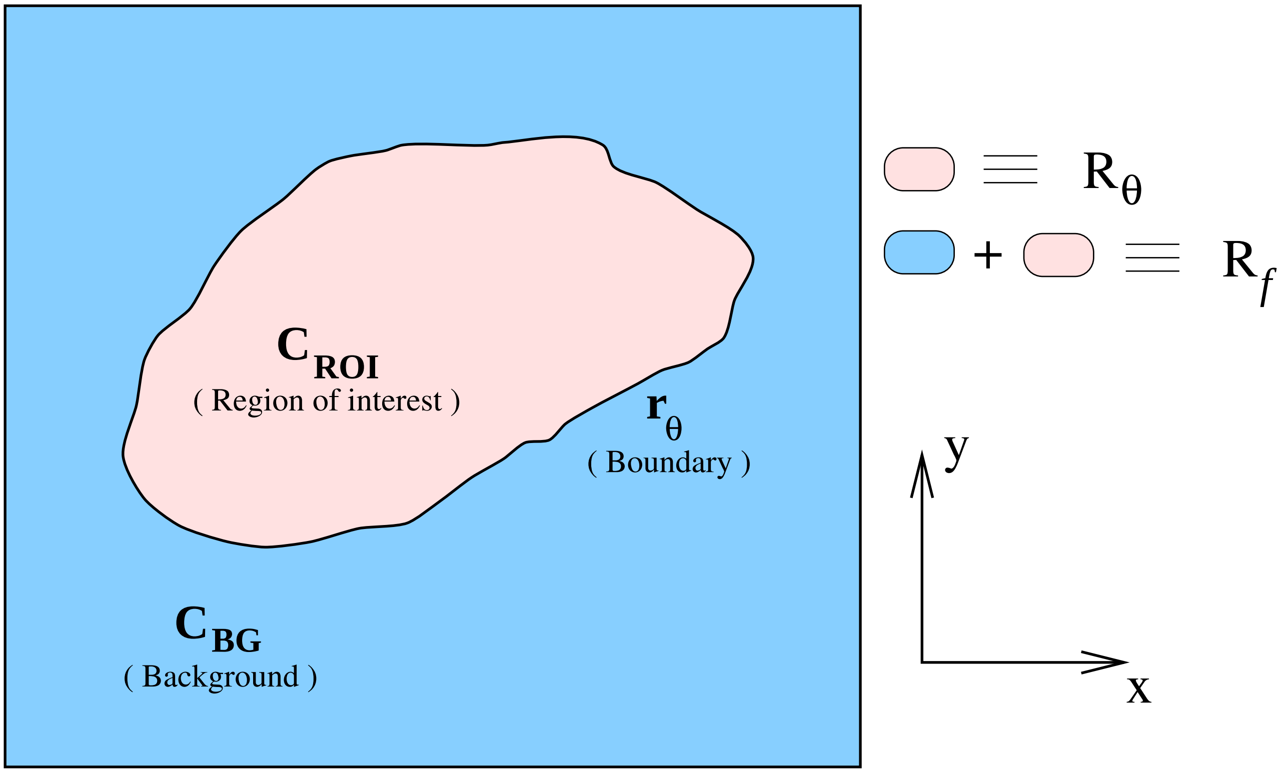


Fig. 2.


C_{ROI}
(Region of interest)
r_θ
(Boundary)
C_{BG}
(Background)
$\equiv$ R_θ
+ $\equiv$ R_f
y
x

.

Fig. 3.

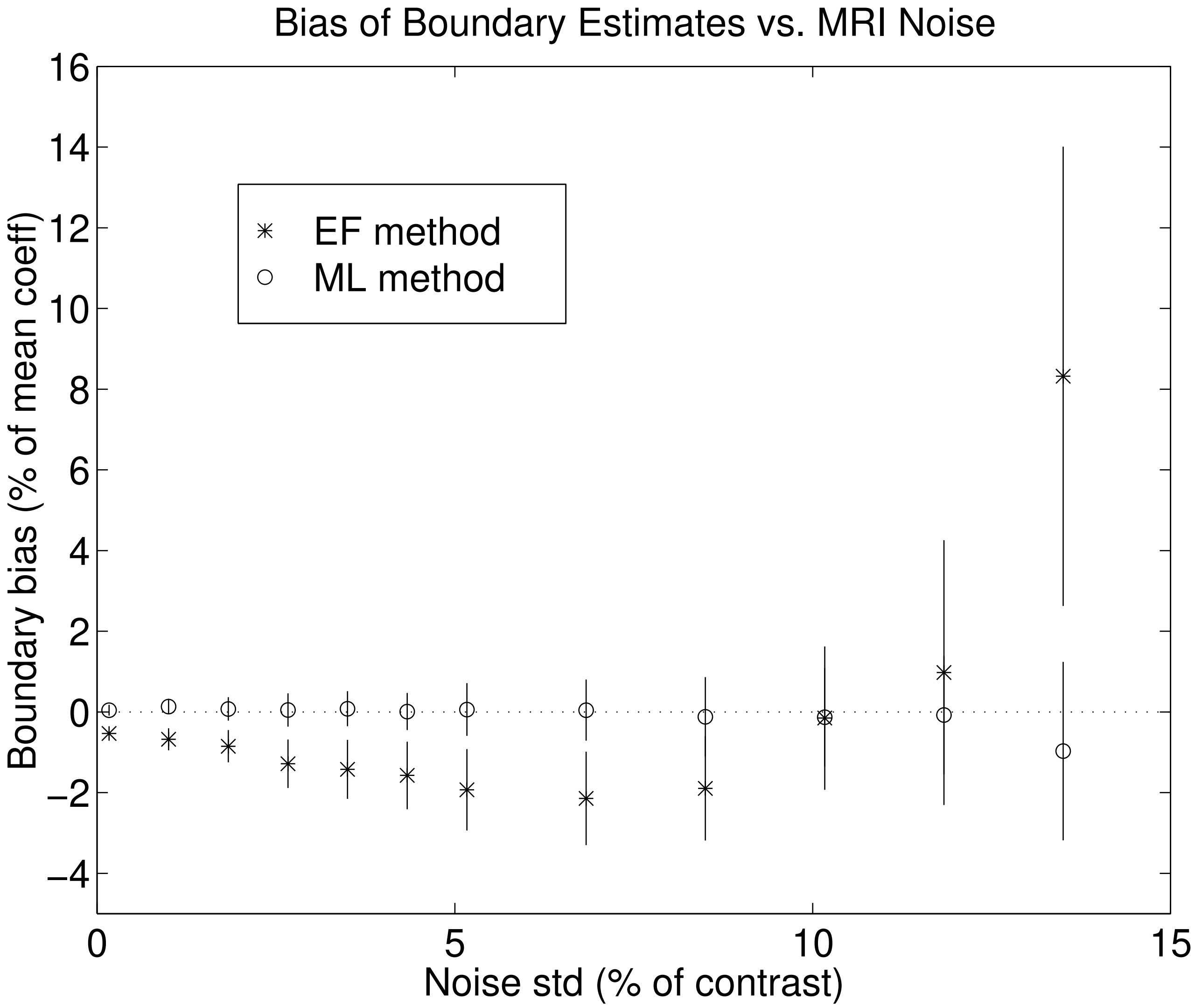
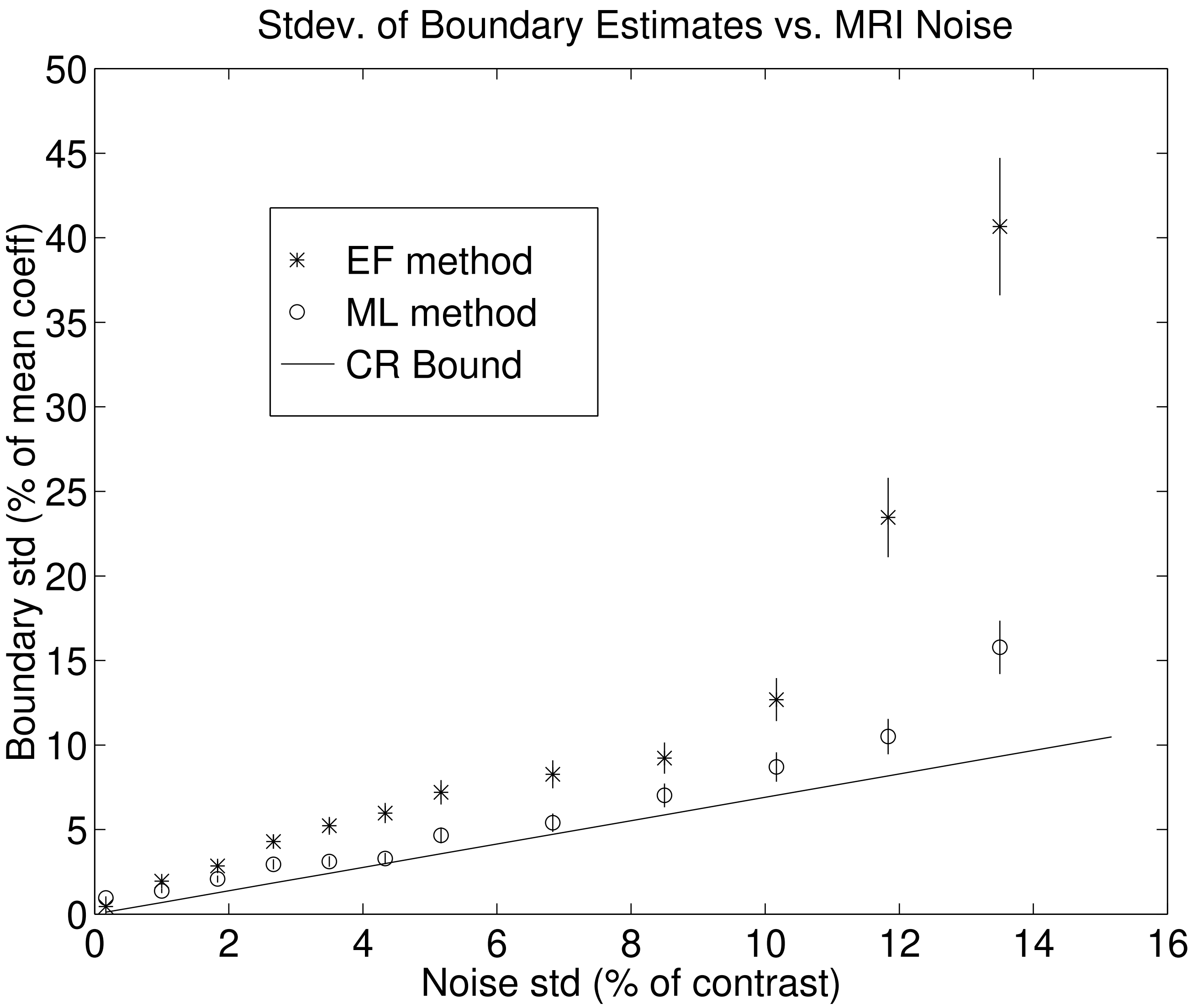

Bias of Boundary Estimates vs. MRI Noise
EF method
ML method
Boundary bias (% of mean coeff)
Noise std (% of contrast)
16
14
12
10
8
6
4
2
0
−2
−4
0
5
10
15

Fig. 4. (a)



Stdev. of Boundary Estimates vs. MRI Noise
Boundary std (% of mean coeff)
Noise std (% of contrast)
EF method
ML method
CR Bound
0
5
10
15
20
25
30
35
40
45
50
0
2
4
6
8
10
12
14
16

Fig. 4. (b)

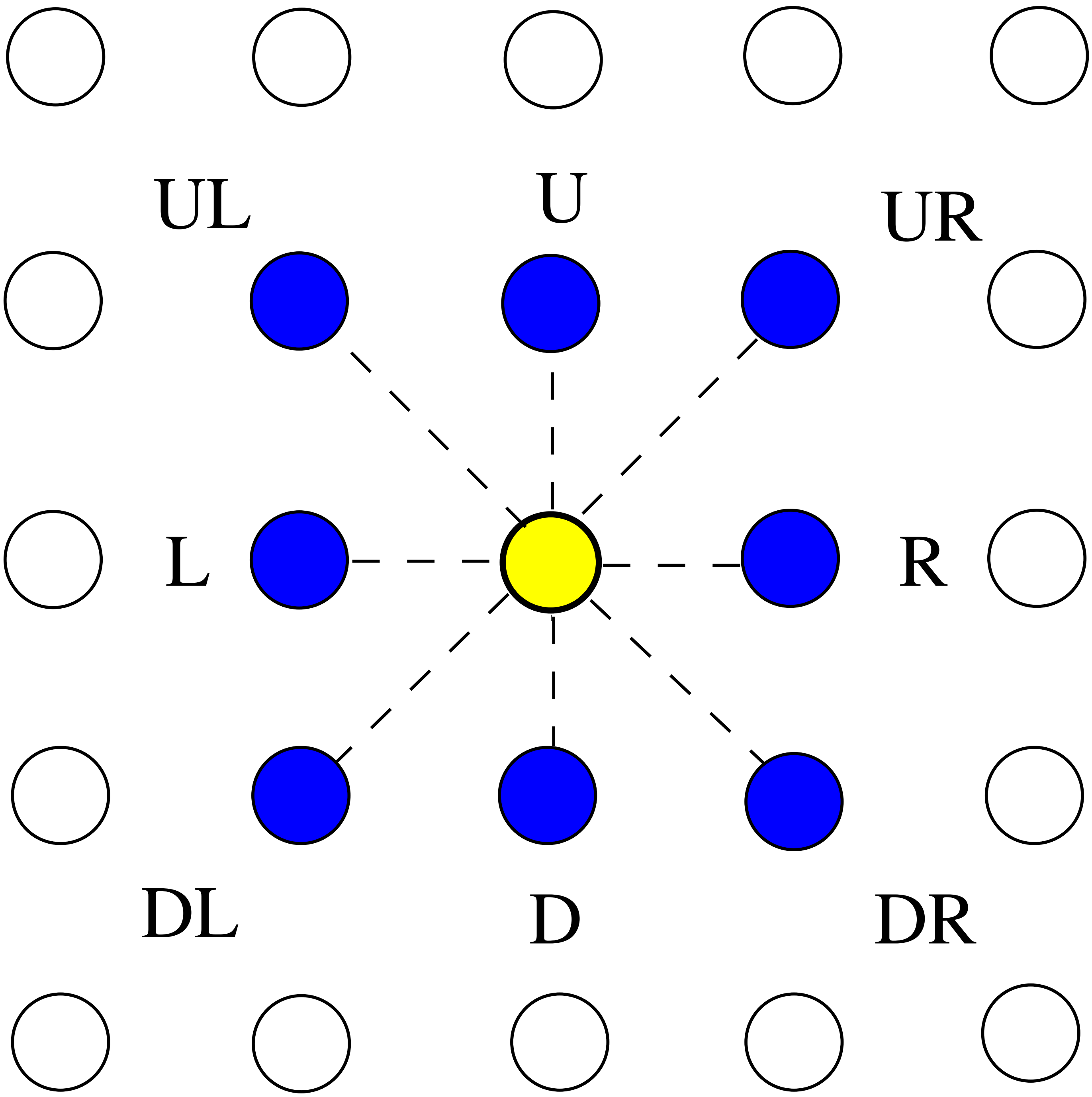


Fig. 5.

[Left, Up; Up−Right, Up−Left] weight matrices

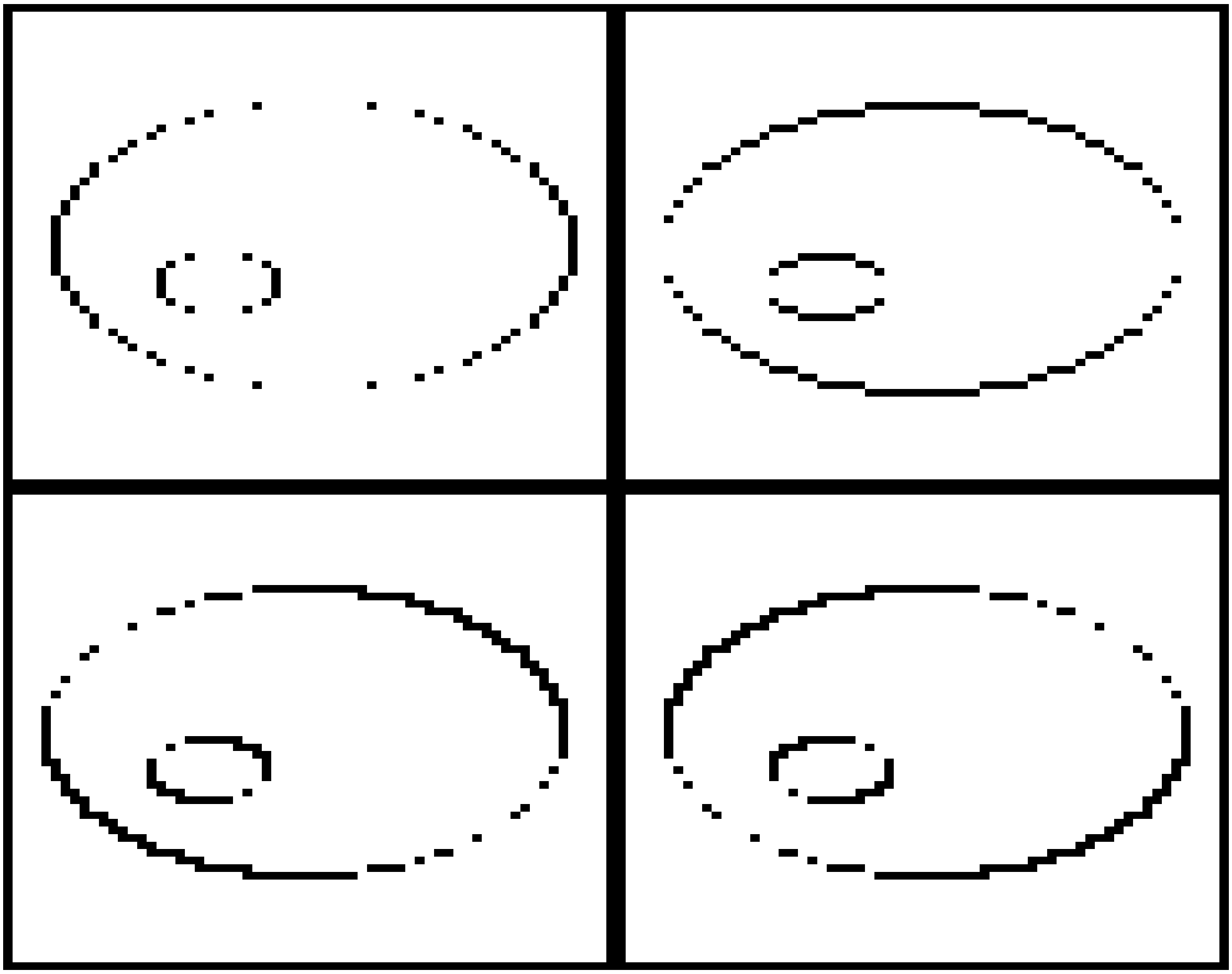


Fig. 6. (a)

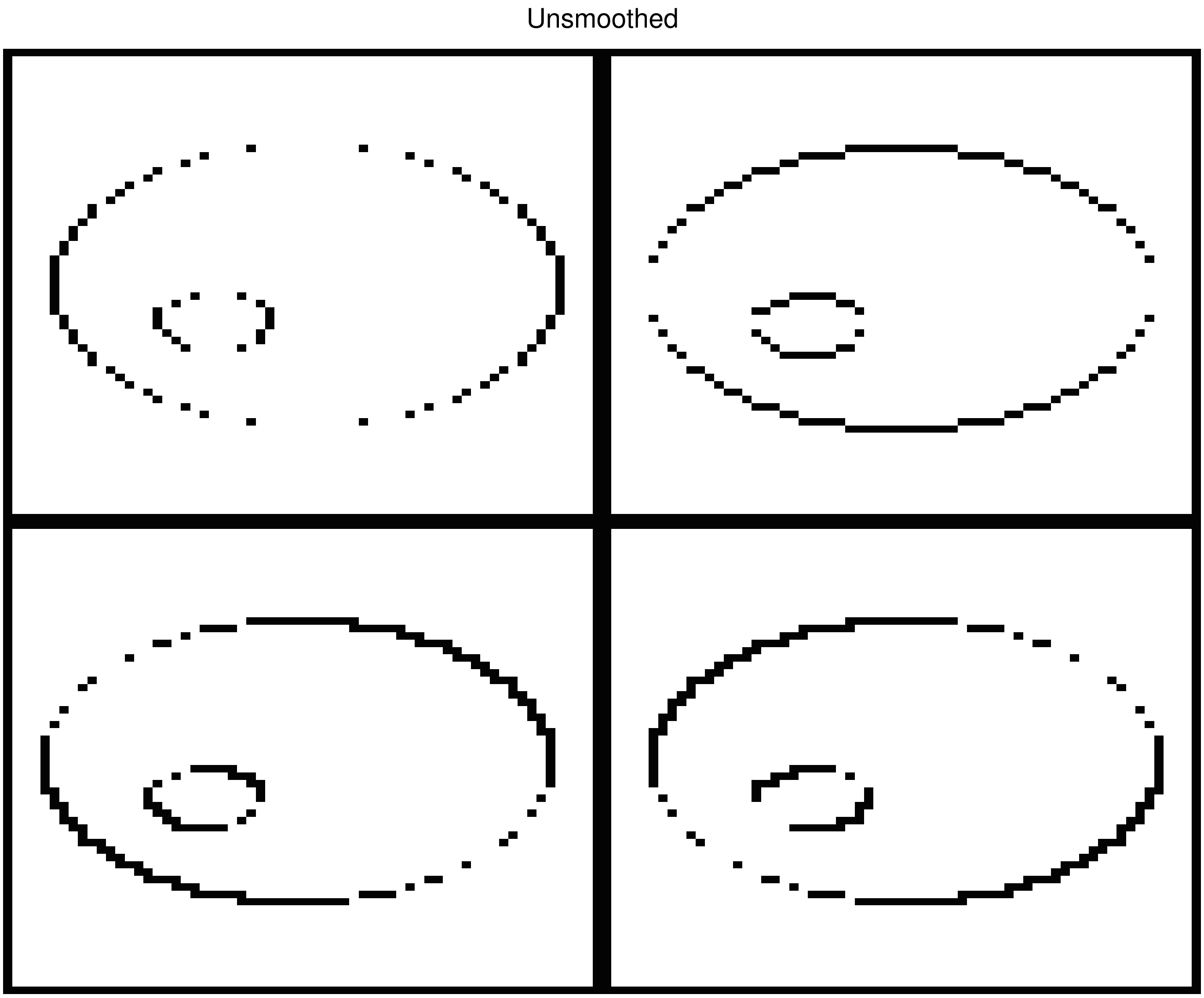


Fig. 6. (b)

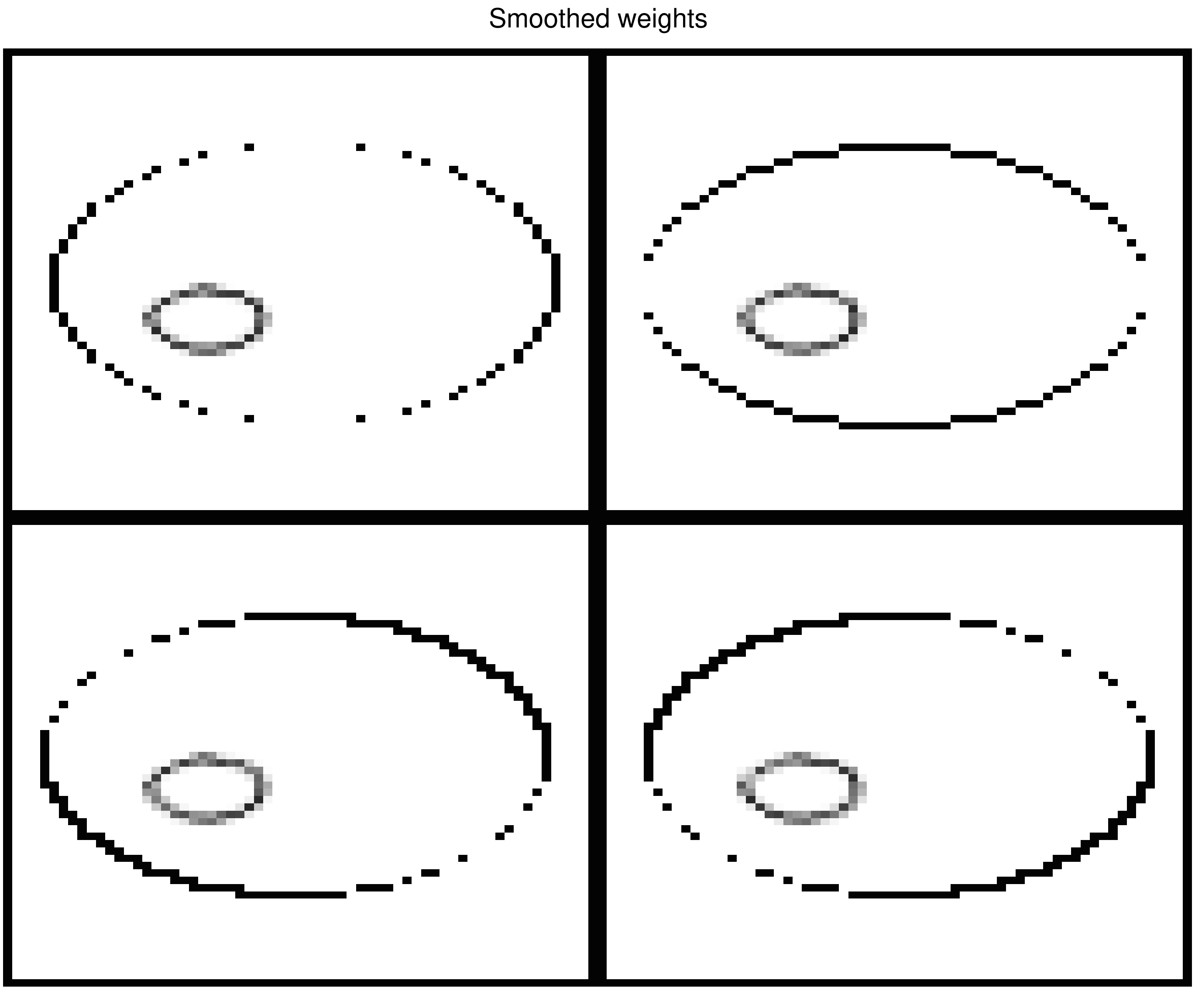


Fig. 6. (c)

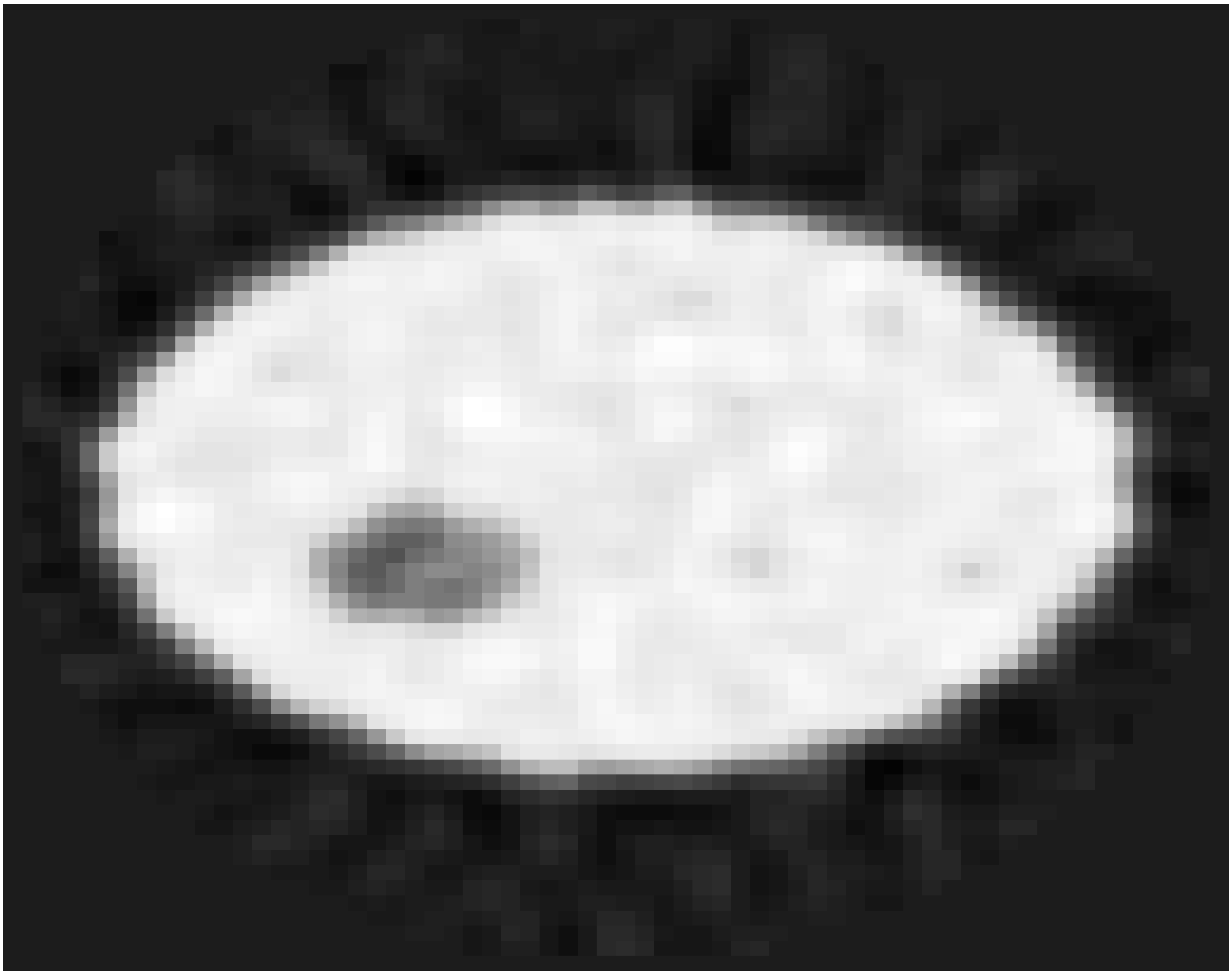


Fig. 7. (a)

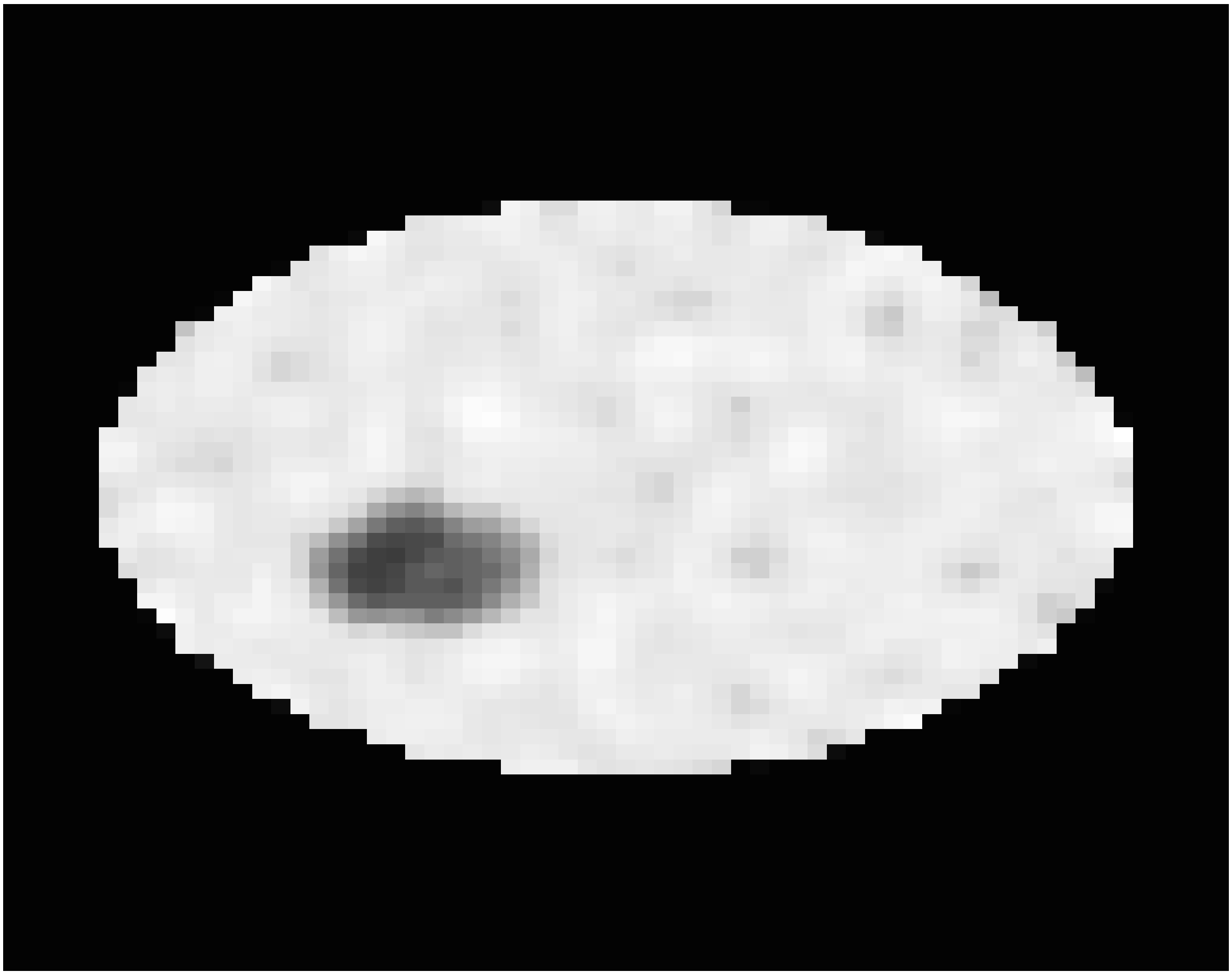


Fig. 7. (b)

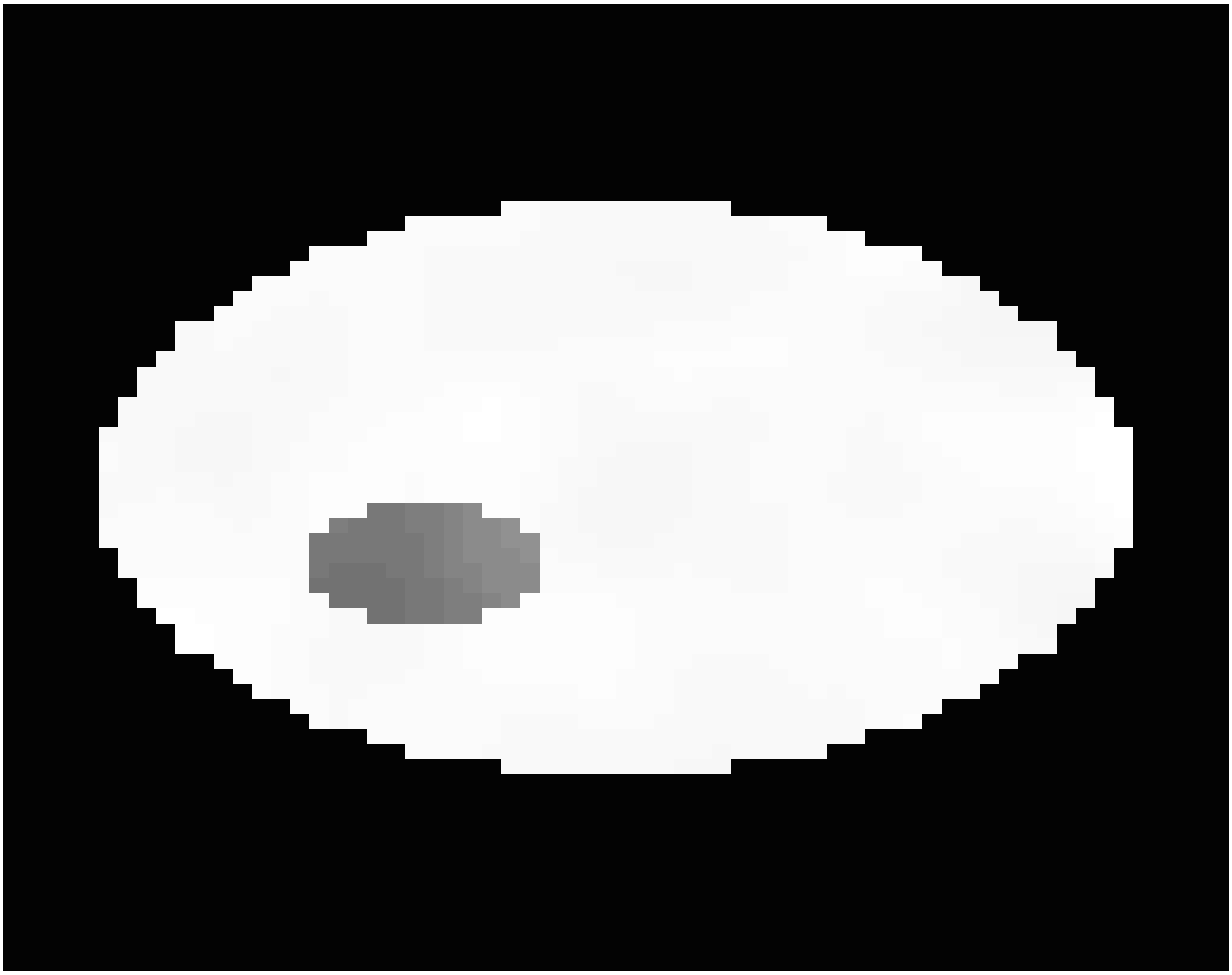


Fig. 7. (c)

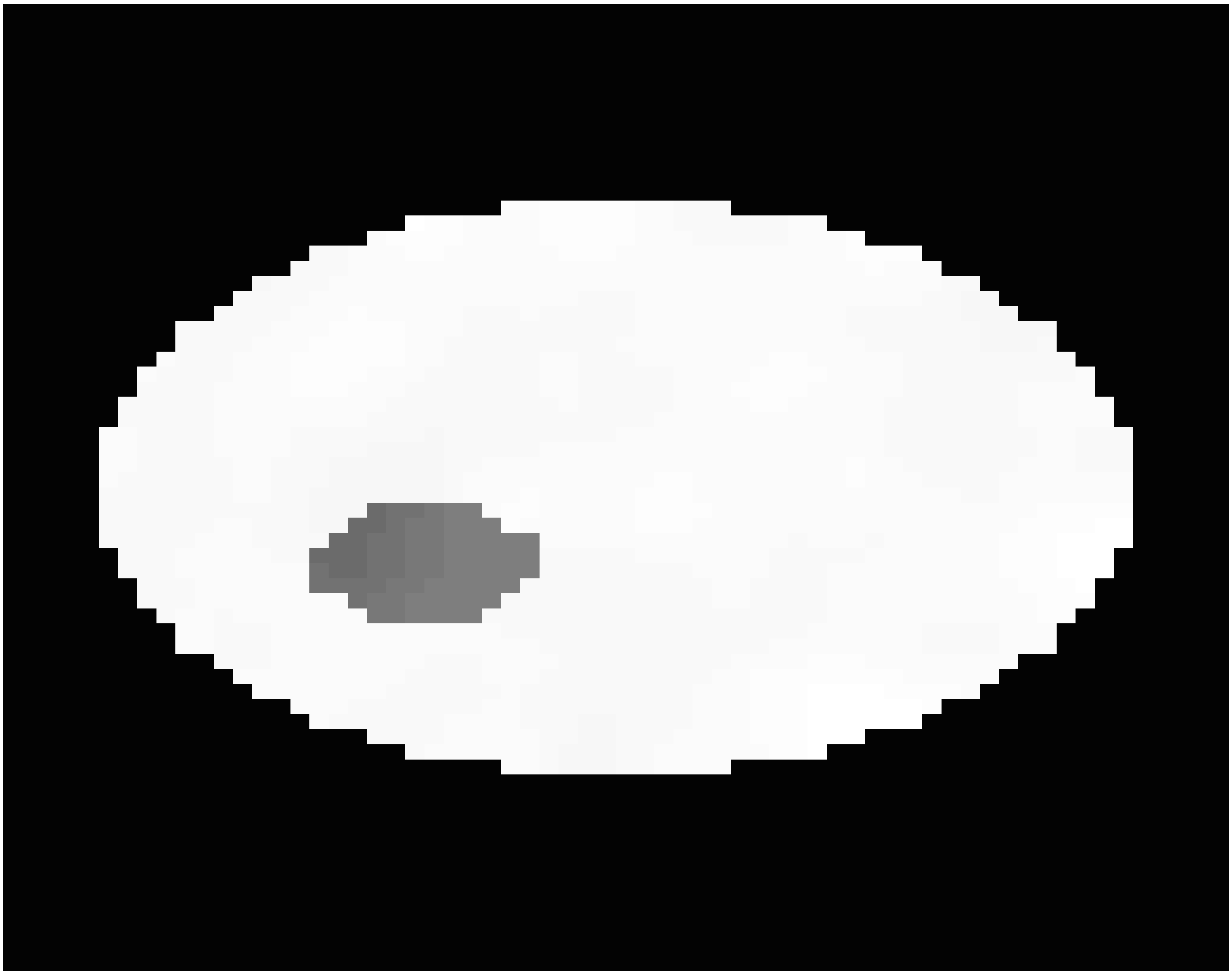


Fig. 7. (d)

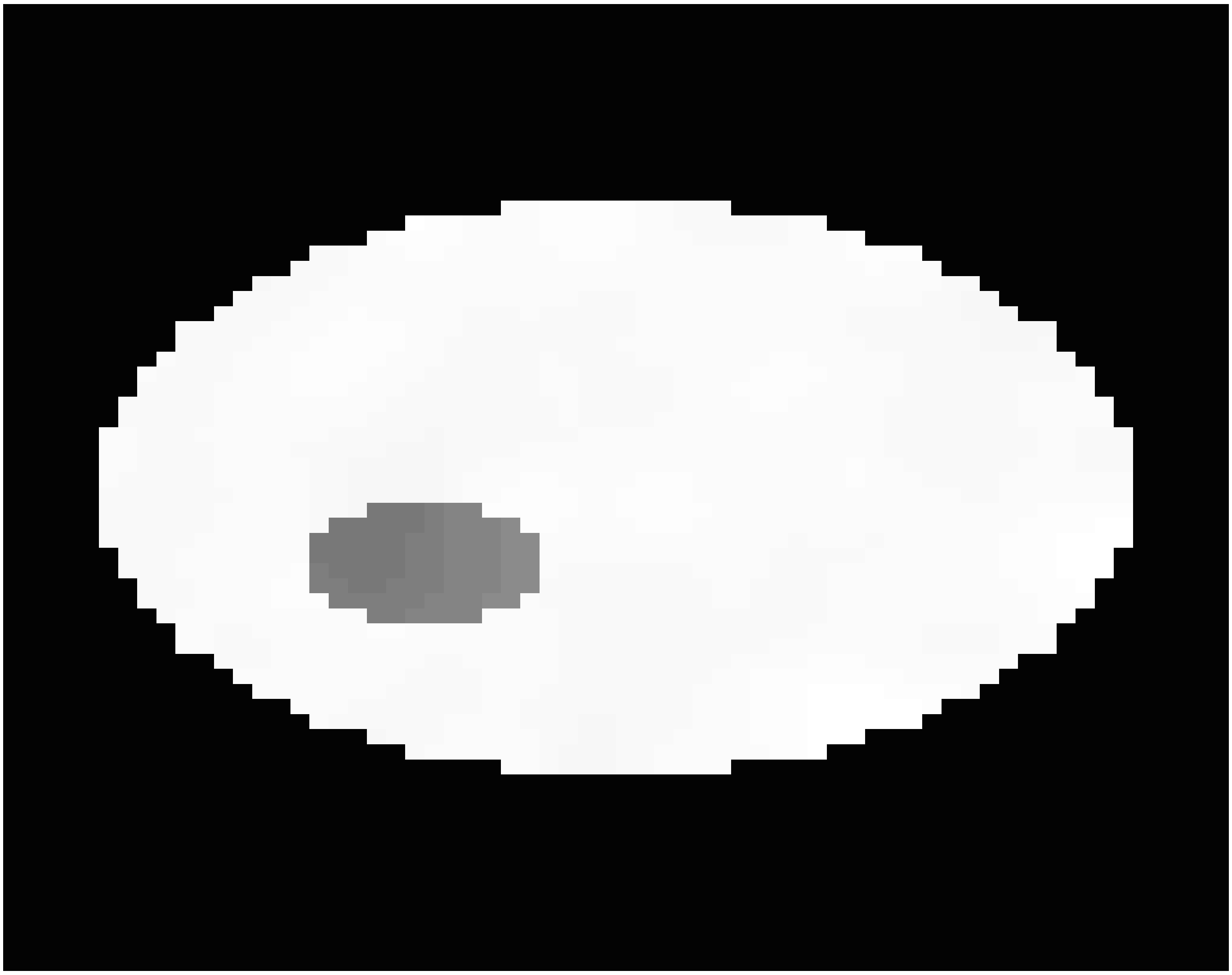


Fig. 7. (e)

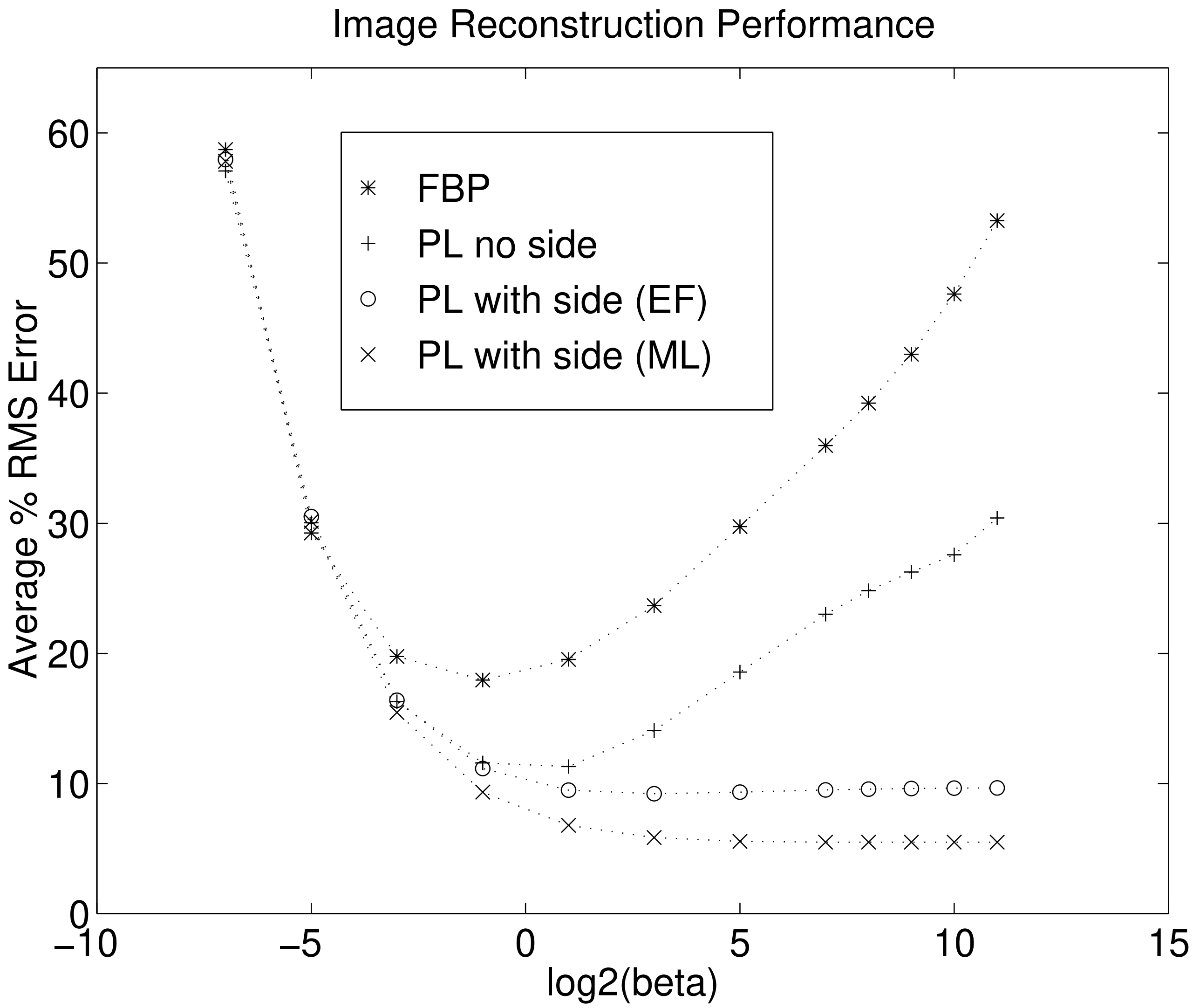


Image Reconstruction Performance
Average % RMS Error
log2(beta)
FBP
PL no side
PL with side (EF)
PL with side (ML)
60
50
40
30
20
10
0
–10
–5
0
5
10
15

Fig. 8.

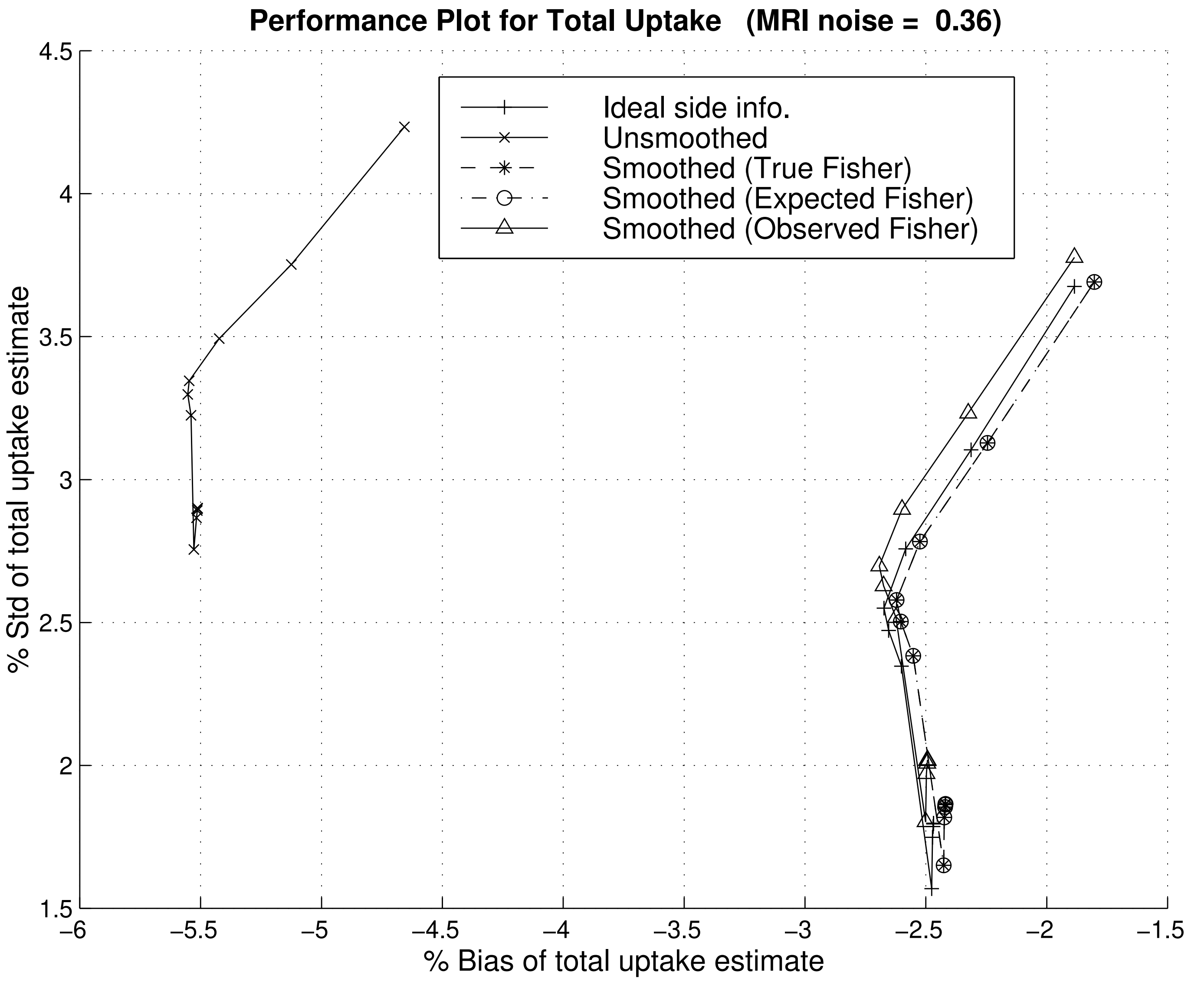

Performance Plot for Total Uptake (MRI noise = 0.36)
Ideal side info.
Unsmoothed
Smoothed (True Fisher)
Smoothed (Expected Fisher)
Smoothed (Observed Fisher)
% Std of total uptake estimate
% Bias of total uptake estimate
4.5
4
3.5
3
2.5
2
1.5
−6
−5.5
−5
−4.5
−4
−3.5
−3
−2.5
−2
−1.5

Fig. 9.

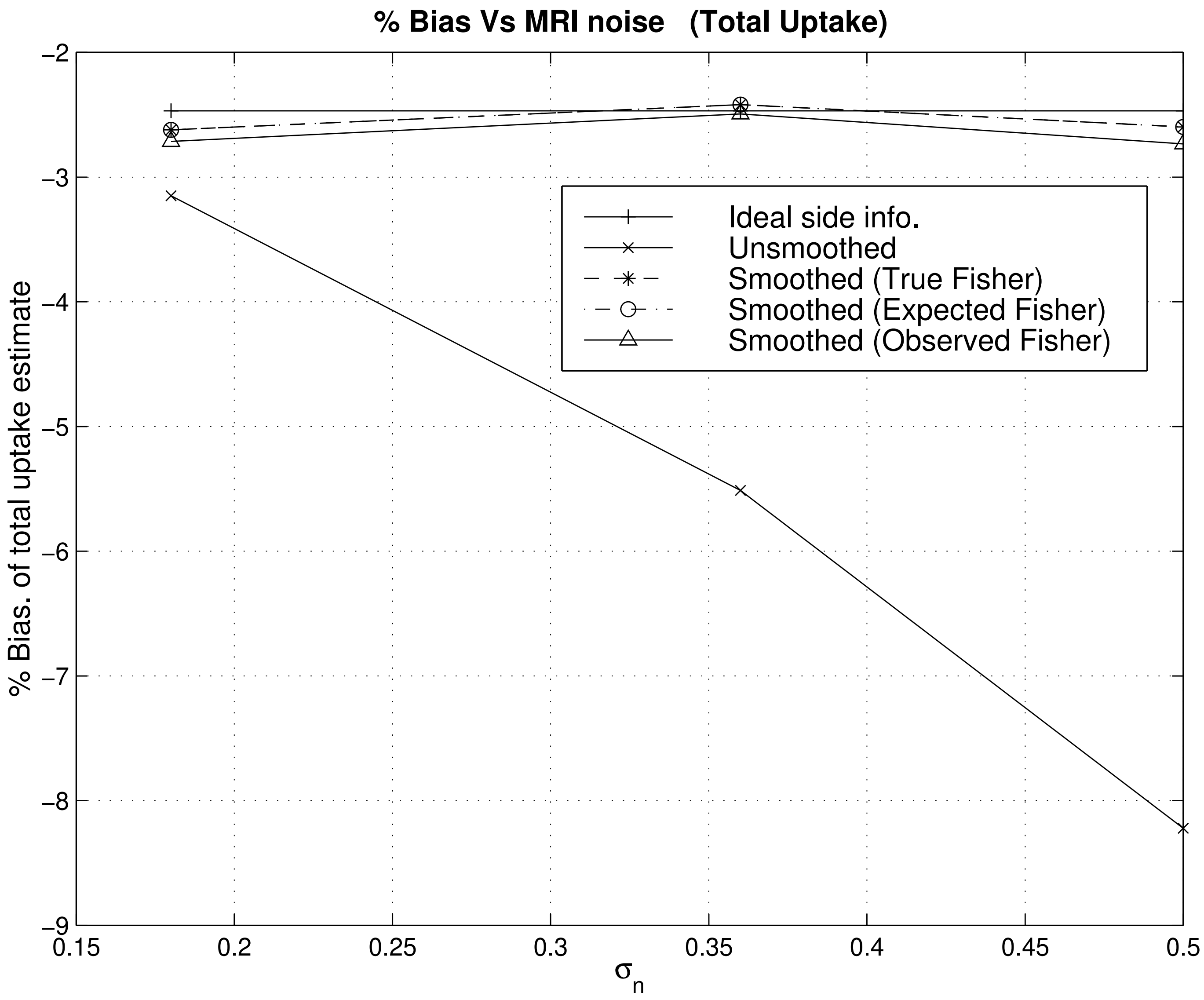

% Bias Vs MRI noise (Total Uptake)
% Bias. of total uptake estimate
σn
−2
−3
−4
−5
−6
−7
−8
−9
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Ideal side info.
Unsmoothed
Smoothed (True Fisher)
Smoothed (Expected Fisher)
Smoothed (Observed Fisher)

Fig. 10. (a)

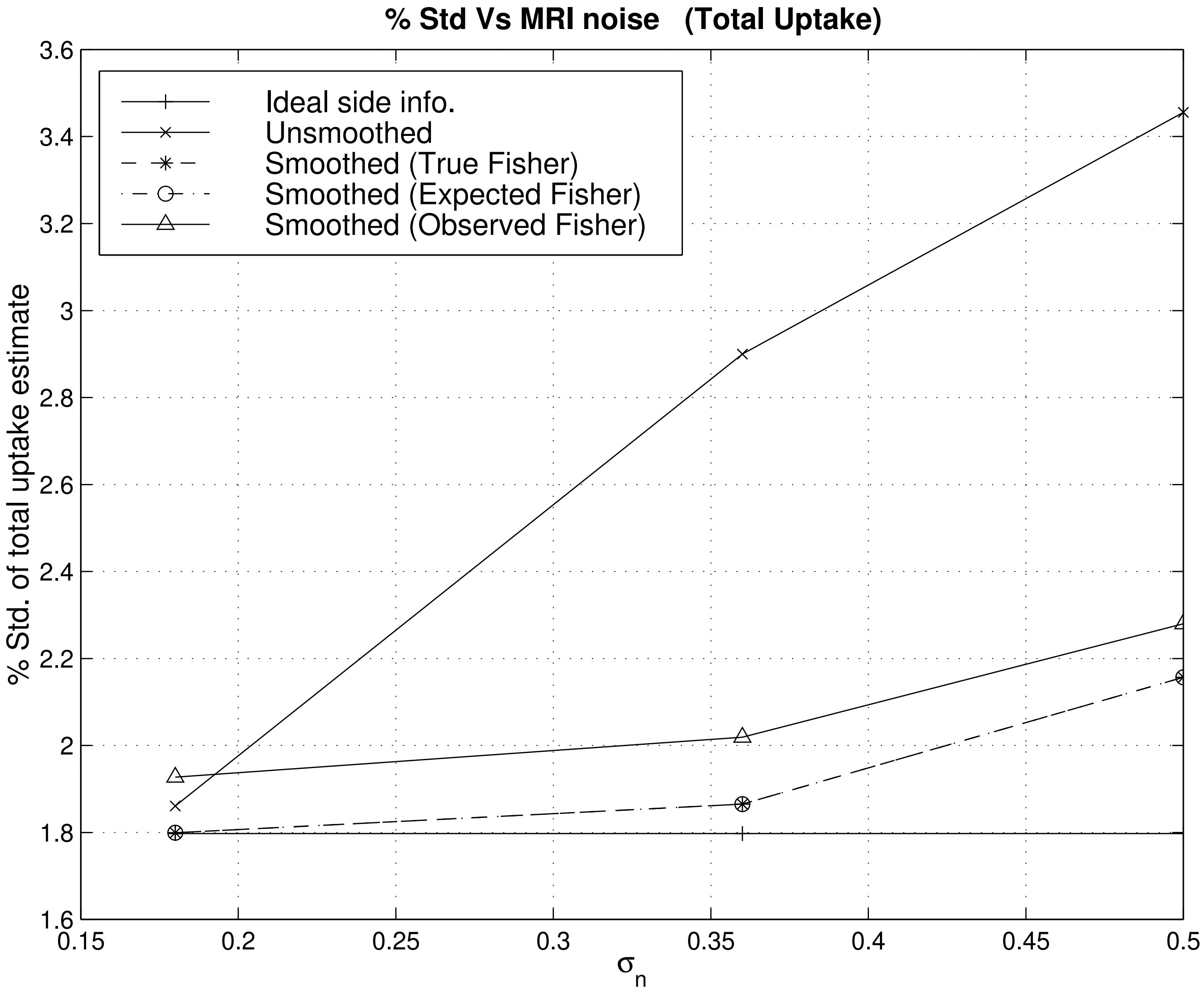

% Std Vs MRI noise (Total Uptake)
Ideal side info.
Unsmoothed
Smoothed (True Fisher)
Smoothed (Expected Fisher)
Smoothed (Observed Fisher)
% Std. of total uptake estimate
3.6
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
σ_n

Fig. 10. (b)