# Regularized Image Reconstruction Algorithms for Dual-Isotope Myocardial Perfusion SPECT (MPS) Imaging Using a Cross-Tracer Prior

Xin He, *Member, IEEE*, Lishui Cheng, *Student Member, IEEE*, Jeffrey A. Fessler, *Fellow, IEEE*, and Eric C. Frey*, *Senior Member, IEEE*

*Abstract*—In simultaneous dual-isotope myocardial perfusion SPECT (MPS) imaging, data are simultaneously acquired to determine the distributions of two radioactive isotopes. The goal of this work was to develop penalized maximum likelihood (PML) algorithms for a novel cross-tracer prior that exploits the fact that the two images reconstructed from simultaneous dual-isotope MPS projection data are perfectly registered in space. We first formulated the simultaneous dual-isotope MPS reconstruction problem as a joint estimation problem. A cross-tracer prior that couples voxel values on both images was then proposed. We developed an iterative algorithm to reconstruct the MPS images that converges to the maximum *a posteriori* solution for this prior based on separable surrogate functions. To accelerate the convergence, we developed a fast algorithm for the cross-tracer prior based on the complete data OS-EM (COSEM) framework. The proposed algorithm was compared qualitatively and quantitatively to a single-tracer version of the prior that did not include the cross-tracer term. Quantitative evaluations included comparisons of mean and standard deviation images as well as assessment of image fidelity using the mean square error. We also evaluated the cross tracer prior using a three-class observer study with respect to the three-class MPS diagnostic task, i.e., classifying patients as having either no defect, reversible defect, or fixed defects. For this study, a comparison with conventional ordered subsets-expectation maximization (OS-EM) reconstruction with postfiltering was performed. The comparisons to the single-tracer prior demonstrated similar resolution for areas of the image with large intensity changes and reduced noise in uniform regions. The cross-tracer prior was also superior to the single-tracer version in terms of restoring image fidelity. Results of the three-class observer study showed that the proposed cross-tracer prior and the convergent algorithms improved the image quality of dual-isotope MPS images compared to OS-EM.

*Index Terms*—Dual isotope imaging, emission computed tomography, joint estimation, maximum a posteriori (MAP) reconstruction.

X. He is with the Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD 21287 USA (e-mail: xhe6@jhmi.edu).

L. Cheng is with the Department of Biomedical Engineering and Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD 21287 USA (e-mail: lcheng14@jhmi.edu).

J. A. Fessler is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA.

*E. C. Frey is with the Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD 21287 USA (e-mail: efrey@jhmi.edu).

## I. INTRODUCTION

PENALIZED maximum likelihood (PML) image reconstruction methods have been used extensively to address the ill-posed nature of emission image reconstruction. In particular, various penalty functions have been developed to penalize the maximum likelihood (ML) solution of the emission reconstruction problem to lower the noise in nuclear medicine images. These penalty functions, or priors, penalize images that exhibit large differences between nearby voxels.

In some nuclear medicine imaging protocols, multiple images are obtained for the same patient characterizing different functional or anatomical properties. For example, in simultaneous dual-isotope myocardial perfusion SPECT (MPS) imaging, two images are acquired simultaneously, representing the distributions of two radioactive isotopes. Due to the simultaneous acquisition, the two images are perfectly registered in time and space. We hypothesize that common information in the registered data can be used to improve the overall image quality. The goal of this study was to develop a cross-tracer prior that uses this common information and penalizes both images simultaneously, to develop converging PML algorithms for the cross-tracer prior, and to compare the new reconstruction method, using task-based performance measures, to a conventional separate reconstruction technique.

### A. Simultaneous Dual-Isotope Myocardial Perfusion SPECT

In a simultaneous dual-isotope MPS imaging protocol, radiopharmaceuticals labeled with two isotopes, $^{99m}$Tc and $^{201}$Tl, are simultaneously present in the patient body. Projection data are acquired and two images are reconstructed, thus simultaneously characterizing the stress and rest states of myocardial perfusion. Since the projection data are acquired simultaneously, the two reconstructed images are thus perfectly registered. The diagnosis of coronary artery disease (CAD) using MPS considers defect presence in both images and classifies patients as being normal, having reversible ischemia, or fixed perfusion defects, as shown in Table I.

One limiting problem in simultaneous dual isotope imaging is crosstalk contamination between the projection data from the two isotopes. We have previously developed methods to model [1], [2] and compensate for [3] this contamination using iterative reconstruction algorithms. We have also shown the efficacy of these methods in improving image quality compared to

TABLE I
THREE CLASSES IN SIMULTANEOUS DUAL-ISOTOPE
MPS IMAGING

|  | Class 1 | Class 3 | Class 2 |
|---|---|---|---|
|  | Normal | Reversible defect | Fixed defect |
| Tc-99m (stress) | No* | Yes** | Yes |
| Tl-201 (rest) | No | No | Yes |

\* No denotes perfusion defect absence
\*\* Yes denotes perfusion defect presence

TABLE II
COMPARISON OF TWO-CLASS AUC VALUES USING
SINGLE AND DUAL ISOTOPE DATA

|  | Normal vs. Reversible | | Fixed vs. Reversible | |  |
|---|---|---|---|---|---|
|  | Separate | Simultaneous | Separate | Simultaneous |  |
| Tc data only | 0.888 | 0.886 | 0.888 | 0.870 | Tl data only |
| Tc+Tl data | 0.946 | 0.937 | 0.917 | 0.890 | Tc+Tl data |

the case of no crosstalk compensation. While crosstalk compensation was effective, task performance was still degraded compared to separate acquisition. We believe that a major reason for this is the increase in noise in the compensated dual isotope image compared to the separate acquisition. Specifically, the compensation method can, at best, estimate only the mean of the crosstalk, leaving behind the noise from the crosstalk. As an example, assume that a uniform background is added to projections, which are thus corrupted by Poisson noise. Subtracting the background would remove the bias, but not the noise from the crosstalk. The goal of this work is to exploit the perfect registration of the simultaneous dual isotope images to reduce the impact of this noise on task performance.

### B. Motivation for the Cross-Tracer Prior

To motivate the hypothesis that the perfectly registered images can be used to improve image quality, we performed an experiment using data from [4]. In this study, we evaluated the image quality of dual-isotope MPS imaging using a large simulated image ensemble, reconstructed with OS-EM algorithms with attenuation, scatter and detector response compensation. The image ensemble is described in more detail in Section III. We used a two-class observer study to assess defect detectability using either the image of the isotope itself, or with the addition of data from the other isotope. We investigated two cases. First we investigated defect detection performance in the $^{99m}$Tc image with and without access to the $^{201}$Tl image. In this case, the $^{201}$Tl image was defect free regardless of the state of the $^{99m}$Tc image. Thus, detecting the defect in the $^{99m}$Tc data distinguished between normal and reversible lesions. In the second case, we investigated defect detection performance in the $^{201}$Tl image with and without access to the $^{99m}$Tc image. For this case, the $^{99m}$Tc image had a defect present in all cases, and detecting the defect in the $^{201}$Tl image thus distinguished between patients with fixed or reversible defects.

Table II shows the AUC values obtained using reconstructed images from both separate and simultaneous acquisitions. The left half of Table II shows the AUC values for classifying normal patients versus patients with reversible defects using the $^{99m}$Tc image alone or using both images. Similarly, the right half of Table II shows the AUC values for classifying patients with fixed versus reversible defects using the $^{201}$Tl image only or both images. In both cases, using both the $^{99m}$Tc and $^{201}$Tl images resulted in a higher AUC value than using only one of the images. These results indicate that, despite the fact that the image from one of the isotopes did not explicitly contain information about whether there was a defect present in the other image, access to

that image improved defect detection performance. This observation suggests that successfully using this information during the reconstruction might improve the overall image quality, i.e., the ability to correctly classify into the three diagnostic classes.

### C. Application of the Penalized Maximum Likelihood (PML) Algorithms

As shown above, incorporating the information provided by the other isotope may improve the image quality, and simultaneous acquisition provides a convenient means to enable this since the acquired data are perfectly registered. The goal of this work was thus to develop fast and convergent PML algorithms that integrate the information on images from both isotopes. Many challenges exist in accomplishing this goal.

The first challenge was development of a prior that couples the information on both stress and rest images. One way of achieving this would be to use 4D maximum *a posteriori* (MAP) reconstruction algorithms. In these algorithms, a time domain clique is employed and a time domain smoothing term is incorporated into the prior to penalize differences in corresponding voxel values on images from different time frames. This is possible because it is assumed that the voxel values are continuous in the time domain. In dual-isotope MPS imaging, however, the voxel values are not necessarily related in the two images. Instead, general anatomical features, such as anatomical boundaries, are common. We used this, as well as constraints on the prior form needed to develop a convergent algorithm, and designed a cross-tracer prior that will be described in detail below. Based on this cross-tracer prior, the dual-isotope MPS reconstruction algorithm was formulated as a joint estimation problem.

A second challenge lies in the development of a convergent PML algorithm for the proposed cross-tracer prior. Converging PML algorithms have been investigated for different types of priors, and many optimization techniques have been used in the literature [5]–[8]. We chose to use optimization transfer methods [9], which replace the original cost function at each step with a surrogate function. The surrogate function is selected so that it is easier to optimize and often a closed-form solution update equation exists. In addition, compared to other approaches, reduced computation time is possible for a carefully chosen surrogate function. Thus, the essence of the second challenge is to carefully choose surrogate functions. De Pierro proposed certain surrogate functions for MAP algorithms based on the concavity of the log-likelihood function and the convexity of certain penalty functions, but derived closed form solutions only for quadratic priors [7]. Based on De Pierro's approach, Chang demonstrated that closed-form solutions are possible for a group of nonquadratic Gibbs priors [10]. In

this work, we adopted Chang's approach to derive a separable surrogate function for the proposed cross-tracer prior.

The third challenge is to accelerate the algorithm using ordered subsets while still maintaining the convergence of the algorithm. A fast converging MAP algorithm, *row-action* maximum-likelihood algorithm (RAMLA), was proposed by Browne *et al.*, based on the row-action Algebraic Reconstruction Technique (ART) algorithm, by introducing a subset-independent diminishing step size [11]. Based on Browne's work, Ahn *et al.* proposed a modified block sequential regularized expectation maximization (BSREM) algorithm, which allows more convenient selection of the relaxation parameters [8]. However, the relaxation scheme selection must be performed carefully, and the convergence rate of the algorithms depends on the relaxation scheme. Gunawardana [12] studied the application of the incremental EM algorithm [13] to PET image reconstruction. A similar method, COSEM, was independently proposed by Hisao *et al.* [5], [14], [15] for emission tomography reconstruction. Both methods perform an alternating descent on the complete data and image estimate. These methods provide an accelerated and convergent framework for emission image reconstruction without requiring relaxation parameters [12], [16], [17]. However, closed form solutions have been derived previously only for quadratic priors [14]. In this work we derived a separable surrogate function for the proposed cross-tracer prior and implemented a COSEM-MAP-like fast and convergent image reconstruction algorithm.

### ALGORITHMS

In the following, we first formulate the dual-isotope reconstruction problem as a joint estimation problem and define the joint estimation objective functions. Second, we introduce the cross-tracer prior and formulate a surrogate function that decouples the cross-tracer prior. Using this surrogate, we develop a PML algorithm that converges to the MAP solution. Finally, we develop accelerated algorithms that use ordered subsets.

### D. *Formulating a Joint Estimation Problem*

As described previously, the goal is to reconstruct both stress and rest images simultaneously, resulting in a joint estimation problem. Erdogan *et al.* formulated the objective function for a joint estimation problem for simultaneous reconstruction of attenuation and emission images from PET scans [18]. Inspired by Erdogan's formula, we propose the following objective function:

$$\Phi(\mathbf{x}, \mathbf{y}) = -L(\mathbf{x}) - L(\mathbf{y}) + \beta \Lambda(\mathbf{x}, \mathbf{y}) \qquad (1)$$

where $\mathbf{x}$ and $\mathbf{y}$ are the image vectors of the two isotopes, $L(\mathbf{x})$ and $L(\mathbf{y})$ are the log likelihood of images $\mathbf{x}$ and $\mathbf{y}$ given the measured data, and $\Lambda(\mathbf{x}, \mathbf{y})$ is the potential function for the prior, which is a function of both images.

If $\Lambda(\mathbf{x}, \mathbf{y})$ can be written as the sum of the energy function of $\mathbf{x}$ and the energy function of $\mathbf{y}$, i.e.,

$$\Lambda(\mathbf{x}, \mathbf{y}) = \Lambda_{\mathbf{x}}(\mathbf{x}) + \Lambda_{\mathbf{y}}(\mathbf{y}) \qquad (2)$$

then the objective function $\Phi(\mathbf{x}, \mathbf{y})$ can be written as

$$\Phi(\mathbf{x}, \mathbf{y}) = -L(\mathbf{x}) - L(\mathbf{y}) + \beta \Lambda_{\mathbf{x}}(\mathbf{x}) + \beta \Lambda_{\mathbf{y}}(\mathbf{y}). \qquad (3)$$

In this case, the minimization can be achieved by minimizing the terms involving $\mathbf{x}$ and $\mathbf{y}$ separately. In other words, the reconstruction algorithm could minimize (3) by reconstructing the images $\mathbf{x}$ and $\mathbf{y}$ separately, i.e., without referencing the information in the other image.

However, incorporating the information from both images into the reconstruction process of each individual images requires a prior that couples the voxel values of both $\mathbf{x}$ and $\mathbf{y}$.

### E. *Definition of Cross-Tracer Prior*

Because the images of the two isotopes, $\mathbf{x}$ and $\mathbf{y}$, are perfectly registered, if neighboring voxels values in one image are similar, suggesting the neighboring voxels belong to the same organ, then the corresponding pixels on the other image probably belong to the same organ too. Thus, the prior should penalize pixel differences in both images simultaneously. We define a cross-tracer prior as

$$\Lambda(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{J} \sum_{k \in N_j} \omega_{jk} \psi(x_j, x_k, y_j, y_k) \qquad (4)$$

where $\omega_{jk}$ is the weighting factor of the clique consisting of pixel $j$ and pixel $k$, and

$$\psi(x_j, x_k, y_j, y_k) = \sqrt{1 + \left(\frac{x_j - x_k}{\delta}\right)^2 + \left(\frac{y_j - y_k}{\eta}\right)^2} - 1. \qquad (5)$$

In (4) and (5), $J$ is the number of voxels in the image of each isotope, $N_J$ represents the neighborhood of voxel $j$, the constants $\omega_{jk}$ are non-negative weights, and $\delta$ and $\eta$ are the two adjustable parameters that define the prior. In principle, $\delta$ and $\eta$ may be different due to differences in the scale of the two images.

We chose the hyperbolic form of the prior so that it is convex and edge preserving. To understand how it meets the other design goal of favoring smoothness in portions of an image where the image for the other radionuclide is smooth, consider the potential function in terms of the difference in pixel values

$$\psi(\Delta x_{jk}, \Delta y_{jk}) = \sqrt{1 + \left(\frac{\Delta x_{jk}}{\delta}\right)^2 + \left(\frac{\Delta y_{jk}}{\eta}\right)^2} - 1 \qquad (6)$$

where $\Delta x_{jk} = (x_j - x_k)$ and $\Delta y_{jk} = (y_j - y_k)$. Consider the partial derivative of the potential function, which is analogous to the force that tries to reduce the difference between neighboring pixels, in the image $\mathbf{x}$

$$\frac{\partial \psi(\Delta x_{jk}, \Delta y_{jk})}{\partial \Delta x_{jk}} = \frac{\delta^{-2} \Delta x_{jk}}{\sqrt{1 + \left(\frac{\Delta x_{jk}}{\delta}\right)^2 + \left(\frac{\Delta y_{jk}}{\eta}\right)^2}}. \qquad (7)$$

For simplicity, consider the case where $\delta$ and $\eta$ are 1. Fig. 1 shows a plot of the derivative of the potential function for several different values of $\Delta y_{jk}$. Note that the "force" increases more rapidly as a function of the difference in the pixel values in the $\mathbf{x}$ image when the difference in pixel values in the $\mathbf{y}$
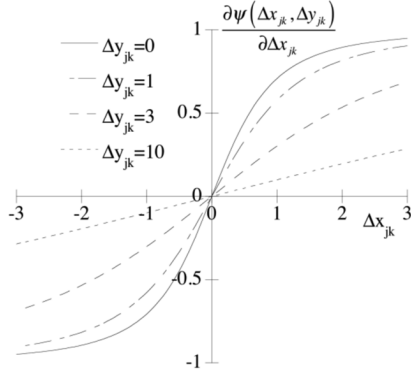
Fig. 1. Graph of partial derivative of cross-tracer prior potential function with respect to pixel differences in the $\mathbf{x}$ image for various values of pixel difference in the $\mathbf{y}$ image.

image is smaller. Thus, small differences in the pixel values of neighboring pixels in the $\mathbf{y}$ image favor small differences in the pixel values in the $\mathbf{x}$ image. Similarly, large differences in neighboring pixel values in the $\mathbf{y}$ image apply less "force" to smooth differences in the pixels in the $\mathbf{x}$ image. In other words, if the differences in both the $\mathbf{y}$ image and $\mathbf{x}$ image are small, then the partial derivative is large, and the prior will provide more smoothing. On the other hand, if the differences in both the $\mathbf{y}$ and $\mathbf{x}$ images are large, then the partial derivative will be small, providing less smoothing. When the difference in one image is large but small in the other image, the smoothing will be somewhere in between the previous two cases.

### F. Convergent PML Algorithm With the Cross-Tracer Prior Separable Surrogate Function of the Prior

To obtain a closed-form solution for the update equation, we first construct a separable surrogate function for the penalty function at a given iteration using the estimate obtained from the previous iteration. The surrogate function is denoted by, $\Lambda^{(n)}(\mathbf{x}, \mathbf{y})$ and should satisfy the following conditions:

$$\text{C1)} \quad \Lambda^{(n)}(\mathbf{x}^{(n)}, \mathbf{y}^{(n)}) = \Lambda(\mathbf{x}^{(n)}, \mathbf{y}^{(n)}) \tag{8}$$

and

$$\text{C2)} \quad \Lambda^{(n)}(\mathbf{x}, \mathbf{y}) \geq \Lambda(\mathbf{x}, \mathbf{y}) \text{ for all } \mathbf{x} \text{ and } \mathbf{y} \tag{9}$$

where $\mathbf{x}^n$ and $\mathbf{y}^{(n)}$ are the estimates of $\mathbf{x}$ and $\mathbf{y}$ after the $n$th iteration.

We now derive the separable surrogates for the hyperbolic prior function in (5). Let

$$\psi(x_j, x_k, y_j, y_k) = \lambda(x_j - x_k, y_j - y_k) \tag{10}$$

and define

$$\lambda'_x(x, y) = \frac{\partial \lambda(x, y)}{\partial x}$$

and

$$\lambda'_y(x, y) = \frac{\partial \lambda(x, y)}{\partial y}. \tag{11}$$

Further, let

$$\gamma_x(x, y) = \frac{\partial \lambda(x, y)}{\partial x} \cdot \frac{1}{x}$$

and

$$\gamma_y(x, y) = \frac{\partial \lambda(x, y)}{\partial y} \cdot \frac{1}{y}. \tag{12}$$

Given the above assumptions, in Appendix I we derived a separable surrogate function, $\Lambda^{(n)}(\mathbf{x}, \mathbf{y})$, given by

$$\Lambda^{(n)}(\mathbf{x}, \mathbf{y}) = \Lambda_{\mathbf{x}}^{(n)}(\mathbf{x}) + \Lambda_{\mathbf{y}}^{(n)}(\mathbf{y}) + C^{(n)} \tag{13}$$

where

$$\Lambda_{\mathbf{x}}^{(n)}(\mathbf{x}) = 2 \sum_{j=1}^{J} \sum_{k \in N_j} h_{jk}^{(n)}(x_j)$$

$$\Lambda_{\mathbf{x}}^{(n)}(\mathbf{x}) = 2 \sum_{j=1}^{J} \sum_{k \in N_j} h_{jk}^{(n)}(x_j) \tag{14}$$

and

$$\Lambda_{\mathbf{y}}^{(n)}(\mathbf{y}) = 2 \sum_{j=1}^{J} \sum_{k \in N_j} g_{jk}^{(n)}(y_j)$$

$$\Lambda_{\mathbf{y}}^{(n)}(\mathbf{y}) = 2 \sum_{j=1}^{J} \sum_{k \in N_j} g_{jk}^{(n)}(y_j) \tag{15}$$

where $C^{(n)}$ is a constant independent of $\mathbf{x}$ and $\mathbf{y}$

$$h_{jk}^{(n)}(x) = \omega_{jk} \gamma_x \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right)$$
$$\times \left( x - \frac{x_j^{(n)} + x_k^{(n)}}{2} \right)^2 \tag{16}$$

and

$$g_{jk}^{(n)}(y) = \omega_{jk} \gamma_y \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right)$$
$$\times \left( y - \frac{y_j^{(n)} + y_k^{(n)}}{2} \right)^2. \tag{17}$$

### G. Surrogate Function for the Log Likelihood Functions

De Pierro derived a surrogate function for the log likelihood function [7], [19]. Chang *et al.* later adopted this surrogate function to develop a convergent PML algorithm for PET image reconstruction [10]. This surrogate function, $L_{\mathbf{x}}^{(n)}$, is given by

$$L_{\mathbf{x}}^{(n)}(\mathbf{x})$$
$$= \sum_{i=1}^{I} \left\{ \sum_{j=1}^{J} P_{ij} \frac{p_i x_j^{(n)}}{\sum\limits_{k=1}^{J} P_{ik} x_k^{(n)}} \log x_j - \sum_{j=1}^{J} P_{ij} x_j \right\}$$
$$+ C_{1,L_x}^{(n)} \tag{18}$$

where the constant $C_{1,L_x}^{(n)}$ is independent of $x_j$. In (18), $P_{ij}$ denotes the probability that a photon emitted from the $j$th voxel of image $\mathbf{x}$ is detected in the $i$th detector bin, and $p_i$ is the observed projection data from image $\mathbf{x}$ in the $i$th detector bin. It has been shown [7], [10], that

$$\text{C1)} \quad L_{\mathbf{x}}^{(n)}(\mathbf{x}^{(n)}) = L_{\mathbf{x}}(\mathbf{x}^{(n)}) \qquad \text{and} \tag{19}$$

$$\text{C2)} \quad L_{\mathbf{x}}^{(n)}(\mathbf{x}) \leq L_{\mathbf{x}}(\mathbf{x}) \quad \text{for all } \mathbf{x}. \tag{20}$$

Note that (18) does not include additive terms to model crosstalk, but these could be easily incorporated.

### H. Surrogate Function for the Objective Function

The goal is to derive a separable surrogate function for the objective function in (1). We have obtained the surrogate functions for both the log likelihood functions and the prior. The surrogate for the objective function in (1) is thus the sum of these two surrogate functions, i.e.,

$$
\begin{aligned}
\Phi^{(n)}(\mathbf{x}, \mathbf{y}) = &- L_{\mathbf{x}}^{(n)}(\mathbf{x}) - L_{\mathbf{y}}^{(n)}(\mathbf{y}) \\
&+ \beta \Lambda_{\mathbf{x}}^{(n)}(\mathbf{x}) + \beta \Lambda_{\mathbf{y}}^{(n)}(\mathbf{y}) \\
&+ \beta C^{(n)} - C_{1,L_x}^{(n)} - C_{1,L_y}^{(n)}.
\end{aligned} \tag{21}
$$

Considering equations (8) and (9) and (19) and (20), it can be proved that the surrogate objective function $\Phi^{(n)}(\mathbf{x}, \mathbf{y})$ in (21) satisfies the following conditions:

$$\text{C1)} \quad \Phi^{(n)}(\mathbf{x}^{(n)}, \mathbf{y}^{(n)}) = \Phi(\mathbf{x}^{(n)}, \mathbf{y}^{(n)}) \tag{22}$$

and

$$\text{C2)} \quad \Phi^{(n)}(\mathbf{x}, \mathbf{y}) \geq \Phi(\mathbf{x}, \mathbf{y}) \text{ for all } \mathbf{x} \text{ and } \mathbf{y}. \tag{23}$$

Substituting the surrogate functions of the log likelihood functions and the priors into (1), we obtain the separable surrogate function of $\Phi(\mathbf{x}, \mathbf{y})$, i.e.,

$$\Phi^{(n)}(\mathbf{x}, \mathbf{y}) = \Phi_{\mathbf{x}}^{(n)}(\mathbf{x}) + \Phi_{\mathbf{y}}^{(n)}(\mathbf{y}) + C^{(n)} + C_{1,L_x}^{(n)} + C_{1,L_y}^{(n)} \tag{24}$$

where

$$
\begin{aligned}
&\Phi_{\mathbf{x}}^{(n)}(\mathbf{x}) \\
&= -L^{(n)}(\mathbf{x}) + \beta \Lambda^{(n)}(\mathbf{x}) \\
&= -\sum_{i=1}^{I} \left\{ \sum_{j=1}^{J} p_i \frac{P_{ij} x_j^{(n)}}{\sum\limits_{k=1}^{J} P_{ik} x_k^{(n)}} \log x_j - \sum_{k=1}^{J} P_{ik} x_k \right\} \\
&\quad + 2\beta \sum_{j=1}^{J} \sum_{k \in N_j} h_{jk}^{(n)}(x_j)
\end{aligned} \tag{25}
$$

and

$$
\begin{aligned}
&\Phi_{\mathbf{y}}^{(n)}(\mathbf{y}) \\
&= -L^{(n)}(\mathbf{y}) + \beta \Lambda^{(n)}(\mathbf{y})
\end{aligned}
$$

$$
\begin{aligned}
&= -\sum_{i=1}^{I} \left\{ \sum_{j=1}^{J} q_i \frac{H_{ij} y_j^{(n)}}{\sum\limits_{k=1}^{J} H_{ik} y_k^{(n)}} \log y_j - \sum_{k=1}^{J} H_{ik} y_k \right\} \\
&\quad + 2\beta \sum_{j=1}^{J} \sum_{k \in N_j} g_{jk}^{(n)}(y_j).
\end{aligned} \tag{26}
$$

The surrogate functions (25) and (26) are separable in terms of $\mathbf{x}$ and $\mathbf{y}$. Thus one can minimize $\Phi^{(n)}(\mathbf{x}, \mathbf{y})$ simply by minimizing $\Phi_{\mathbf{x}}^{(n)}(\mathbf{x})$ and $\Phi_{\mathbf{y}}^{(n)}(\mathbf{y})$ separately. In (26), $H_{ij}$ is the probability that a photon emitted from the $j$th voxel of image $\mathbf{y}$ is detected in the $i$th detector bin, and $q_i$ is the observed projection data from image $\mathbf{y}$ in the $i$th detector bin. Note that $H_{ij}$ differs from $P_{ij}$ because the attenuation and scatter are different for the two tracers.

### I. Derivation of the Convergent PML Algorithm

From the decoupled objective functions, one can solve for $\mathbf{x}$ and $\mathbf{y}$ by minimizing (25) and (26) separately. First consider the minimization problem for $\Phi_{\mathbf{x}}^{(n)}(\mathbf{x})$ in (25). Substituting (16) into (25), $\Phi_{\mathbf{x}}^{(n)}(\mathbf{x})$ becomes

$$
\begin{aligned}
&\Phi_{\mathbf{x}}^{(n)}(\mathbf{x}) \\
&= -\sum_{i=1}^{I} \sum_{j=1}^{J} p_i \frac{P_{ij} x_j^{(n)}}{\sum_{k=1}^{J} P_{ik} x_k^{(n)}} \log x_j + \sum_{i=1}^{I} \sum_{k=1}^{J} P_{ik} x_k \\
&\quad + 2\beta \sum_{j=1}^{J} \sum_{k \in N_j} \omega_{jk} \gamma_x \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right) x_j^2 \\
&\quad - 2\beta \sum_{j=1}^{J} \sum_{k \in N_j} \omega_{jk} \gamma_x \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right) \\
&\quad \times \left( x_j^{(n)} + x_k^{(n)} \right) x_j \\
&\quad + \frac{1}{2}\beta \sum_{j=1}^{J} \sum_{k \in N_j} \omega_{jk} \gamma_x \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right) \\
&\quad \times \left( x_j^{(n)} + x_k^{(n)} \right)^2 .
\end{aligned} \tag{27}
$$

Simplifying, we have

$$
\begin{aligned}
\Phi_{\mathbf{x}}^{(n)}(\mathbf{x}) &= -\sum_{j=1}^{J} E_j^{(n)} \log x_j + \sum_{j=1}^{J} G_j^{(n)} x_j + \sum_{j=1}^{J} F_j^{(n)} x_j^2 \\
&= \sum_{j=1}^{J} \left[ -E_j^{(n)} \log x_j + G_j^{(n)} x_j + F_j^{(n)} x_j^2 \right]
\end{aligned} \tag{28}
$$

where

$$E_j^{(n)} = \sum_{i=1}^{I} p_i \frac{P_{ij} x_j^{(n)}}{\sum\limits_{k=1}^{J} P_{ik} x_k^{(n)}}, \tag{29}$$

$$F_j^{(n)} = 2\beta \sum_{k \in N_j} \omega_{jk} \gamma_x \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right) \tag{30}$$

and

$$G_j^{(n)} = \sum_{i=1}^{I} P_{ij} - 2\beta \sum_{k \in N_j} \omega_{jk} \gamma_x$$
$$\times \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right) \left( x_j^{(n)} + x_k^{(n)} \right). \quad (31)$$

Now we can find the optimal solution for $x_j$, that is

$$x_j^{(n+1)} = \arg \min_{x_j \geq 0} s_j^{(n)}(x_j) \quad j = 1, 2, \ldots J \quad (32)$$

where

$$s_j^{(n)}(t) = -E_j^{(n)} \log t + G_j^{(n)} t + F_j^{(n)} t^2. \quad (33)$$

Since the second derivative of $s_j^{(n)}(t)$ is non-negative, i.e.,

$$\frac{\partial^2 s_j^{(n)}(t)}{\partial t^2} = E_j^{(n)} \frac{1}{t^2} + 2F_j^{(n)} > 0 \quad (34)$$

the function $s_j^{(n)}(t)$ is convex. Finding the solution for (32) is equivalent to solving the following equation:

$$\frac{\partial s_j^{(n)}(t)}{\partial t} = -\frac{E_j^{(n)}}{t} + G_j^{(n)} + 2F_j^{(n)} t = 0. \quad (35)$$

The solution to (32) is thus

$$x_j^{(n+1)} = \frac{-G_j^{(n)} + \sqrt{G_j^{(n)2} + 8F_j^{(n)} E_j^{(n)}}}{4F_j^{(n)}}. \quad (36)$$

Similarly, we derived the solution for $y_j^{(n+1)}$

$$y_j^{(n+1)} = \frac{-G_j'^{(n)} + \sqrt{G_j'^{(n)2} + 8F_j'^{(n)} E_j'^{(n)}}}{4F_j'^{(n)}} \quad (37)$$

where

$$E_j'^{(n)} = \sum_{i=1}^{I} q_i \frac{H_{ij} y_j^{(n)}}{\sum\limits_{k=1}^{J} H_{ik} y_k^{(n)}} \quad (38)$$

$$F_j'^{(n)} = 2\beta \sum_{k \in N_j} \omega_{jk} \gamma_y \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right) \quad (39)$$

and

$$G_j'^{(n)} = \sum_{i=1}^{I} H_{ij} - 2\beta \sum_{k \in N_j} \omega_{jk} \gamma_y$$
$$\times \left( x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)} \right) \left( y_j^{(n)} + y_k^{(n)} \right). \quad (40)$$

(36)–(40) are the MAP algorithm. Note that in deriving (36) and (37), we excluded the negative solutions.

### J. Accelerated PML Algorithms Using Ordered Subsets

The algorithm derived in the previous section converges to the MAP solution; however, it is slow without using ordered subset techniques. Direct application of ordered subsets would accelerate the reconstruction process, but the resulting algorithm would not be guaranteed to converge. We have thus derived a convergent ordered-subset algorithm based on the work in [12], [20] and COSEM [5], [15]–[17], both of which are particular examples of incremental EM algorithm [13]. In [21], incremental optimization transfer, as a generalization of the incremental EM algorithm, was applied to transmission tomography. The incremental EM algorithm is a generalization of ordinary EM, generalizing from expectation-maximization to maximization-maximization. That is, the E step in the ordinary EM algorithm is replaced by a maximization procedure. This variant of the EM algorithm is applicable to the case of statistically independent observations, and thus to the case of emission tomography. The algorithm divides the observed data and complete data into partitions (ordered-subsets), and there are two sub-iterations for each iteration. In the first sub-iteration, optimization is achieved by updating the conditional sufficient statistics with complete data on one partition only. In the second step, the parameters, in our case the image estimates, are chosen to optimize the conditional expectation of the likelihood of all the complete data. For more details, refer to the discussions in [12], [13], and [20]. In emission tomography, the elements of the complete data are the number of emissions from each voxel detected by one particular detector bin, which is also a component of the sufficient statistics vector; the observed data is the number of emissions from all pixels detected in one detector bin; and the parameters to be estimated are the activity in each voxel. As a particular example of incremental EM algorithm, COSEM formulates the reconstruction problem as an optimization problem with respect to both the image estimate and complete data, which is solved by alternate updating of the image estimates and complete data. In particular, in the COSEM-MAP framework, the objective function for the complete data is formulated using ordered-subsets as [15], [21]

$$E_{\text{cmp-MAP}}(\mathbf{C}, \mathbf{x}, \mathbf{v})$$
$$= E_{\text{cmp-ML}}(\mathbf{C}, \mathbf{x}, \mathbf{v}) + \beta \Psi(\mathbf{x})$$
$$= \sum_{ij} P_{ij} x_j + \sum_{l=1}^{L} \sum_{i \in S_l} \sum_{j=1}^{N} C_{ij} \log \frac{C_{ij}}{P_{ij} x_j}$$
$$- \sum_{l=1}^{L} \sum_{i \in S_l} \sum_{j=1}^{N} C_{ij}$$
$$+ \sum_{i} v_i \left( \sum_{j} C_{ij} - p_i \right) + \beta \Psi(\mathbf{x}) \quad (41)$$

where $C_{ij}$ is the $(i,j)$th element of the complete data, $\nu_i$ is a Lagrange multiplier to add the constraint $\sum_j C_{ij} = p_i$, where $p_i$ is the projection data in the $i$th bin, $\Psi(\mathbf{x})$ is the potential function of the prior and $s_l$ is the $l$th subset. In [14], a separable

surrogate was derived for a quadratic prior. The optimization of (41) uses an alternating descent approach and performs the reconstruction using two update equations: 1) a complete data update and 2) an image estimate update.

Here we extend (41) to the joint estimation problem for the cross-tracer prior. Rewriting (1) using the surrogate function for the cross-tracer prior we get

$$
\begin{aligned}
E(\mathbf{C},\mathbf{C}',\mathbf{x},\mathbf{y},\mathbf{v},\mathbf{v}') = & E_{\text{cmp-ML}}(\mathbf{C},\mathbf{x},\mathbf{v}) \\
& + E_{\text{cmp-ML}}(\mathbf{C}',\mathbf{y},\mathbf{v}') \\
& + \Lambda(\mathbf{x},\mathbf{y})
\end{aligned} \tag{42}
$$

where $\mathbf{C}$ and $\mathbf{C}'$ are the complete data matrices for images $\mathbf{x}$ and $\mathbf{y}$, respectively, and $\mathbf{v}$ and $\mathbf{v}'$ are the Lagrange multipliers for the constraints $\sum_j C_{ij} = p_i$ and $\sum_j C'_{ij} = q_i$, respectively.

The first step of the optimization is the complete data update. Taking partial derivatives with respect to the elements in $\mathbf{C}$ and $\mathbf{C}'$, solving for the maximum, and enforcing the constraints results in

$$
C_{ij}^{(n,l)} = p_i \frac{P_{ij}x_j^{(n,l-1)}}{\sum_k P_{ik}x_k^{(k,l-1)}}, \forall\, i \in S_l, \forall\, j \tag{43}
$$

$$
C_{ij}^{(n,l)} = C_{ij}^{(n,l-1)}, \forall\, i \notin S_l, \forall\, j \tag{44}
$$

$$
C'^{(n,l)}_{ij} = q_i \frac{H_{ij}x_j^{(n,l-1)}}{\sum_k H_{ik}x_k^{(n,l-1)}}, \forall\, i \in S_l, \forall\, j \tag{45}
$$

and

$$
C'^{(n,l)}_{ij} = C'^{(n,l-1)}_{ij}, \forall\, i \notin S_l, \forall\, j \tag{46}
$$

where the superscript $(n,l)$ denotes the $n$th iteration and $l$th subset.

The second step in the optimization is updating the image estimates, $\mathbf{x}$ and $\mathbf{y}$, with the complete data fixed. To do this, we replaced the prior with the separable surrogate function. Following the derivation in [14] and [17], in this step we used the objective

$$
\begin{aligned}
& E_{obj}(\mathbf{x}^{(n,l)},\mathbf{y}^{(n,l)}) \\
& = -\sum_{m=1}^{L}\sum_{i \in S_m}\sum_{j=1}^{N} C_{ij}^{(n,l)} \log P_{ij}x_j + \sum_{ij} P_{ij}x_j \\
& \quad -\sum_{m=1}^{L}\sum_{i \in S_m}\sum_{j=1}^{N} C'^{(n,l)}_{ij} \log H_{ij}y_j + \sum_{ij} H_{ij}y_j \\
& \quad + 2\beta\sum_{j=1}^{J}\sum_{k \in N_j} h_{jk}^{(n,l)}(x_j) + C_{\mathbf{x}}^{(n,l)} \\
& \quad + 2\beta\sum_{j=1}^{J}\sum_{k \in N_j} g_{jk}^{(n,l)}(x_j) + C_{\mathbf{y}}^{(n,l)}
\end{aligned} \tag{47}
$$

where $g_{jk}^{(n,l)}(x)$ and $h_{jk}^{(n,l)}(x)$ are defined similar to those in (16) and (17), and $C_x^{(n,l)}$ and $C_y^{(n,l)}$ are constants that do not depend on $\mathbf{x}^{(n,l)}$ and $\mathbf{y}^{(n,l)}$. Minimization of (47) can thus be achieved by minimizing the terms involving $\mathbf{x}$ and $\mathbf{y}$ separately to obtain

the solution to $\mathbf{x}^{(n,l)}$ and $\mathbf{y}^{(n,l)}$. In particular, taking the partial derivatives of (47) with respect to $x_j$ and $y_j$ separately, equating to zero, and solving, we obtain

$$
\begin{cases}
x_j^{(n,l)} = \dfrac{-G_j^{(n,l-1)}+\sqrt{\left(G_j^{(n,l-1)}\right)^2+16\beta F_j^{(n,l-1)}E_j^{(n,l)}}}{8\beta F_j^{(n,l-1)}} \\[4mm]
y_j^{(n,l)} = \dfrac{-G'^{(n,l-1)}_j+\sqrt{\left(G'^{(n,l-1)}_j\right)^2+16\beta F'^{(n,l-1)}_j E'^{(n,l)}_j}}{8\beta F'^{(n,l-1)}_j}
\end{cases} \tag{48}
$$

where

$$
E_j^{(n,l)} = \sum_i C_{ij}^{(n,l)}
$$

$$
F_j^{(n,l-1)} =
$$
$$
\sum_{k \in N_j} \left\{ \omega_{jk}\gamma_x \left( x_j^{(n,l-1)} - x_k^{(n,l-1)}, y_j^{(n,l-1)} - y_k^{(n,l-1)} \right) \right\}
$$

$$
G_j^{(n,l-1)} = \sum_i P_{ij} -
$$
$$
2\beta\sum_{k \in N_j} \omega_{jk}\gamma_x \left( x_j^{(n,l-1)} - x_k^{(n,l-1)}, y_j^{(n,l-1)} - y_k^{(n,l-1)} \right)
$$
$$
\left( x_j^{(n,l-1)} + x_k^{(n,l-1)} \right)
$$

$$
E'^{(n,l)}_j = \sum_i C'^{(n,l)}_{ij}
$$

$$
F'^{(n,l-1)}_j =
$$
$$
\sum_{k \in N_j} \left\{ \omega_{jk}\gamma_y \left( x_j^{(n,l-1)} - x_k^{(n,l-1)}, y_j^{(n,l-1)} - y_k^{(n,l-1)} \right) \right\}
$$

and

$$
G'^{(n,l-1)}_j =
$$
$$
\sum_i H_{ij} - 2\beta\sum_{k \in N_j} \omega_{jk}\gamma_y
$$
$$
\left( x_j^{(n,l-1)} - x_k^{(n,l-1)}, y_j^{(n,l-1)} - y_k^{(n,l-1)} \right)\left( y_j^{(n,l-1)} + y_k^{(n,l-1)} \right).
$$

In summary, the dual-isotope reconstruction using the cross-tracer prior based on COSEM framework is implemented in two steps. Step 1 updates the complete data using (43)–(46), and Step 2 updates the images using (48).

## II. EVALUATION EXPERIMENTS

### A. Comparison Using Sample Images

We have investigated the convergence of the proposed algorithm and the qualitative properties of the reconstructed images using a simulation study. Projection data were simulated using the 4D NCAT torso phantom. A defect was placed on the myocardial wall. Low noise projection data were simulated using the SimSET Monte Carlo simulation code. The projection data were scaled and Poisson noise was added to obtain projections with a clinically realistic noise level. We used these data to perform experiments demonstrating the convergence of the algorithm and the properties of the cross-tracer prior.

## B. Image Fidelity Study

To investigate the properties and benefits of the cross-tracer prior we performed a simulation experiment using a phantom designed to simulate myocardial perfusion imaging while being simple enough to provide insight into the effects of the reconstruction hyperparameters. The phantom consisted of an elliptical cylinder 16 cm long with major and minor axes of 31 and 21 cm, respectively. A spherical shell with outside and inside diameters of 8 and 5 cm, respectively, was created to simulate the myocardium. The shell was offset 4 cm in the left lateral direction and 2 cm in the anterior direction from the center of the cylinder. The shell had an activity 5 times as high as the background. We created a simulated perfusion defect with a 50% reduction of myocardial activity in the region of intersection of a 4 cm diameter sphere centered in the middle of the shell wall at the left lateral position. We created a pair of phantoms representing a reversible defect: one without the defect (rest) and one with the defect (stress). The phantoms were digitized into $64 \times 64 \times 32$ cm matrices with a voxel size of 0.5 cm. We also created an attenuation map the same size as the cylinder with an attenuation coefficient of $0.15 \text{ cm}^{-1}$. We used the same analytic projector as used in the reconstruction to generate projections for a standard $180°$ cardiac orbit with 64 projection views. We modeled an LEHR collimator (including spatially varying blurring) and a circular radius of rotation of 16 cm. The projections were scaled to a total of 100 k counts in the central slice of the phantom and Poisson noise was simulated. The images were reconstructed using OS-EM (up to 20 iterations with 16 subsets per iteration), the proposed MAP algorithm with the cross-tracer prior (CTMAP), and the same algorithm modified to use only the hyperbolic prior applied to each of the images separately. In other words, we used the prior

$$\Lambda(\mathbf{x}) = \sum_{j}^{J} \sum_{k \in N_j} \omega_{jk} \sqrt{1 + \left( \frac{x_j - x_k}{\delta} \right)^2} - 1. \qquad (49)$$

for the $\mathbf{x}$ an $\mathbf{y}$ images, which were reconstructed separately. A total of 100 iterations (16 subsets per iteration) were used for the MAP reconstruction to insure convergence. All reconstructions were performed with attenuation and collimator-detector response modeling. The OS-EM images were postreconstruction filtered using Butterworth filters with an order 8 and cutoff frequencies from 0.10 to 0.30 pixel$^{-1}$ in increments of 0.02 pixels$^{-1}$. For the MAP methods, we investigated values of $\beta$ equal to 5, 2, and 1 times $10^n$, with $n$ in the range 0 to $-4$, for a total of 12 values of $\beta$. For each $\beta$ we used values of $\delta^{1/2}$ equal to 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, and 1.0, and used the same value of $\delta$ for both images (i.e., set $\eta = \delta$ for CTMAP).

After reconstruction and (for OS-EM) filtering, we computed the voxel-by-voxel mean-square error with respect to the phantom to provide a measure of image fidelity.

## C. Task-Based Three-Class Observer Study

To evaluate the proposed algorithms in the context of reconstructing dual-isotope images, we conducted a three-class observer study to evaluate the classification performance in distinguishing between patients with no defect, with a reversible

defect, or with a fixed defect. In this study, we used a population of stress ($^{99m}$Tc) and rest ($^{201}$Tl) images that realistically modeled patient variations.

In this study we used a previously developed phantom population [4], [22] based on the 4D NCAT phantom [23], which realistically modeled the anatomical and organ uptake variations seen in clinical data. The phantom population included 24 NCAT anatomies with variations in gender, patient size, heart angle, size, and position, and lung volume [24]. We modeled variations in perfusion defect sizes and locations. Please refer to [4], [22], [25] for the details of the phantom population generation.

For each of the 24 anatomies, we separately simulated projections of five organs (heart, lung, body, liver, and kidney) using a combined SimSET and angular response function (ARF) simulation code [26]. Attenuation, scatter in the body, the distant-dependent collimator-detector response, and penetration and scatter in the collimator were modeled in the simulation. Data for $^{201}$Tl were simulated including the seven major emission energies. The energy window for $^{99m}$Tc was 15% wide, centered at 140.5 KeV; that for $^{201}$Tl was 20% wide, centered at 70 KeV. Square projection bins and cubic reconstructed voxels with a side length of 0.48 cm were used for both $^{99m}$Tc and $^{201}$Tl images. The projections consisted of 64 views over $180°$ from left posterior oblique (LPO) to right anterior oblique (RAO). Since the phantoms modeled uptakes observed in real patients, and the simulations modeled the sensitivity and acquistion times for clinical cardiac acquisitions, the count levels in the data were clinically realistic for both isotopes. The means $\pm$ standard deviations of the total counts in the sinogram averaged over the 48 slices reconstructed were $193 \pm 38$ k counts for $^{99m}$Tc and $110 \pm 26$ k counts for $^{201}$Tl.

Using the individual projections, we generated a population of dual-isotope projection data that included realistic variations in organ uptake by summing the organ projections. Defects at three locations in the myocardium were simulated. The projection data were scaled to model $^{99m}$Tc and $^{201}$Tl injected activities of 16 and 3 mCi, respectively. This procedure produced 1296 pairs of $^{99m}$Tc/$^{201}$Tl projection datasets (24 phantoms $\times$ 3 defect locations $\times$ 6 uptake ratio combinations $\times$ 3 defect statuses, i.e., normal, reversible defect, and fixed defect). Note that we did not include crosstalk in these simulations because the implementation of the cross-tracer MAP reconstruction algorithm did not include the ability to compensate for crosstalk. The inclusion of crosstalk in the data combined with use of crosstalk compensation would result in increased noise in the reconstructed images, especially for $^{201}$Tl. This should provide a greater advantage for the cross tracer prior, but this requires future in depth investigation.

We reconstructed the 1,296 pairs of $^{99m}$Tc/$^{201}$Tl images using the proposed algorithm. In the reconstruction process, attenuation and detector response effects (AD) were modeled. As a comparison, we performed OS-EM reconstruction with the same compensations. In both cases we used 20 iterations with 16 updates per iteration.

We then applied standard postprocessing methods to the reconstructed images. For OS-EM reconstruction, the images were reoriented to short axis, filtered with a Butterworth filter,

and windowed to obtain short axis (SA) $^{99m}$Tc/$^{201}$Tl image pairs. For each reoriented 3D stress ($^{99m}$Tc) image, a $32 \times 32$ image containing the centroid of the defect (or the same short-axis slice in the corresponding defect-absent image) was extracted from the postprocessed short-axis slices. The same short-axis slice in the corresponding rest ($^{201}$Tl) image was also selected. These $^{99m}$Tc/$^{201}$Tl SA image pairs served as the input for the three-class mathematical observer studies. For the cross-tracer MAP reconstruction, the same postprocessing methods were used with the exception that the Butterworth filtering was omitted.

We then applied a three-class channelized Hotelling observer (3-CHO) to the $^{99m}$Tc/$^{201}$Tl image pair, resulting in one feature vector for each image pair. The details for the 3-CHO and the channels used for the 3-CHO are described in [4].

Using half the image pairs for each reconstruction and postprocessing combination, we trained a 3-CHO, which was then tested using the remaining half of the images. The volume under a three-class ROC surface (VUS) was used as a figure-of-merit for task performance.

## III. RESULTS

### A. Convergence of the Two Algorithms

We verified the convergence of the convergent and accelerated PML algorithms for the cross-tracer prior, introduced in Sections I-I and I-J, using the following parameters: $\delta = \eta = \sqrt{10}$ and $\beta = 0.05$.

We evaluated the value of the objective function in (1) at each iteration for the convergent PML algorithm, as shown in Fig. 2(a). Also shown is the (negative) change of objective function on a semi-log scale. In particular, we computed $\ln(-(\phi(i+1) - \phi(i)))$, where $\phi(i)$ is the objective function evaluated at the $i$th iteration. These data show that, as expected, the objective function decreases monotonically.

Using the same set of parameters, we tested the convergence of the accelerated PML algorithms. As noted in [12], [15], [20], the energy function defined in (42) decreases monotonically with each iteration, and the algorithm reaches the maximum *a posterior* (MAP) solution of (1), though a monotonic decrease of the objective function in (1) is not guaranteed. Despite this, in our simulations we did observe a monotonic decrease in the objective function of (1). Fig. 3 shows the value of the objective function for three values of the number of subsets per iteration (8, 4, and 2), corresponding to 8, 16, and 32 angles per subset, respectively. It can be seen that the accelerated algorithm converged faster with a smaller number of angles per subset (i.e., more subsets per iteration). In addition, the accelerated algorithm achieved the same ultimate value of the objective function.

We have also studied the change of mean value of three slices of the reconstructed images with iterations to demonstrate the convergence of the convergent PML algorithm. The results are shown in Fig. 4. It can be seen that convergence is achieved after 200 iterations. A comparison of images and profiles through the Tc-99m stress images after 200 and 300 iterations is shown in Fig. 5. These data demonstrate that both the value of the objective function and the reconstructed images converge.
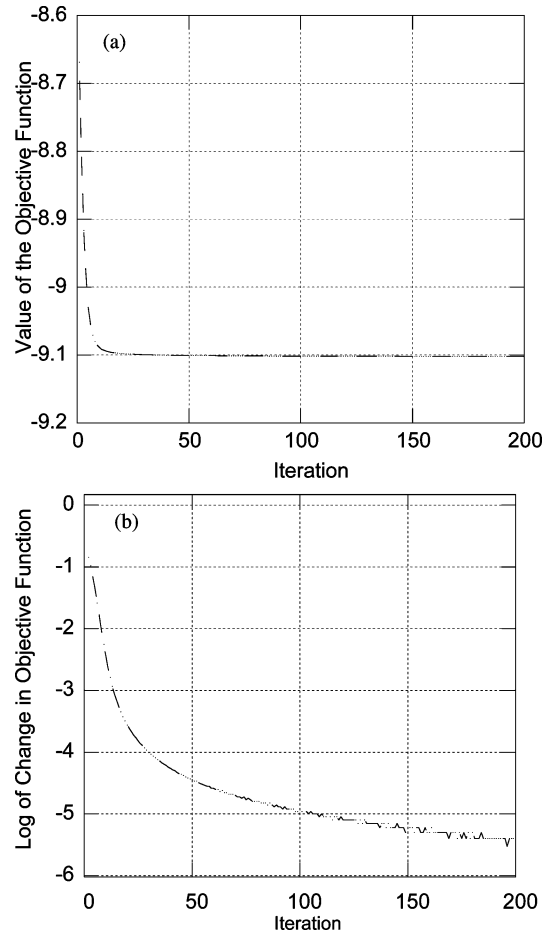


Fig. 2. (a) Plot of the objective function and (b) logarithm of the (negative) change of the objective function of the convergent PML algorithm as a function of iteration.

### B. Effects of the Cross-Tracer Prior—Comparison of Sample Images

Using the parameters in the previous section, we reconstructed a pair of images that contained reversible defects to demonstrate the effects of the cross-tracer prior compared to a conventional edge-preserving prior. The conventional edge-preserving prior used is that shown in (49). This reconstruction algorithm will be referred to as single-tracer MAP (STMAP). We used the same value of $\beta$ and $\delta$ for both the rest and stress images and the same values used in the previous section ($\beta = 0.05, \delta = \sqrt{10}$). We used a set of projection data with a reversible defect (defect present in the stress image, but not in the rest image) with a reduction of 35% of the myocardial activity concentration. We generated and reconstructed 50 sets of simulated Poisson noise realizations and computed both the mean and standard deviation images.

Fig. 6 shows coronal slices through the phantom, mean, and sample noisy images for both isotopes and both reconstruction algorithms. Note that the mean images contain some noise artifacts, partly because of the residual noise in the MC simulated data. From these images we see that edges corresponding to large activity discontinuities, such as at the boundary of the myocardium and chamber and liver and background, are sharp with both algorithms. Fig. 7 shows profiles through the mean images
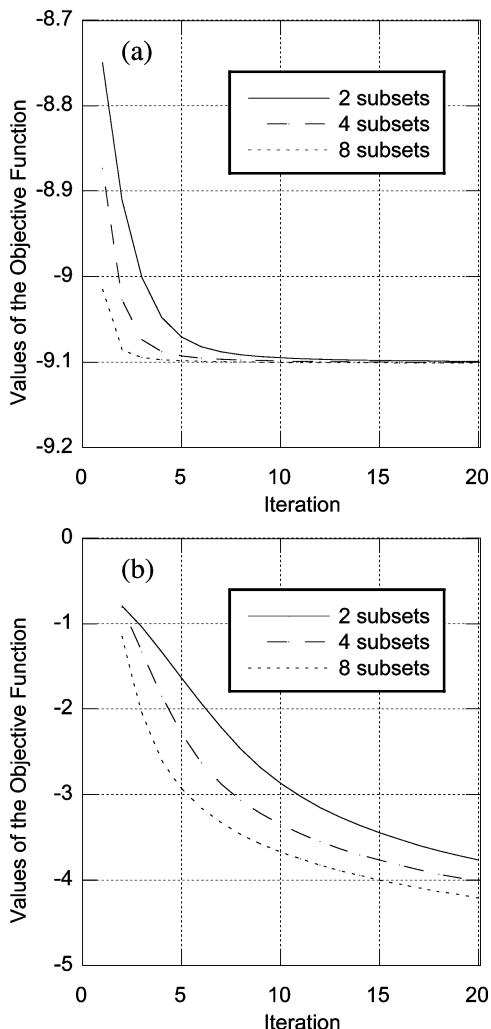
Fig. 3.   Plot of the value of objective functions (a) and logarithm of the (negative) change of the objective functions of the accelerated PML algorithm as a function of iteration for 2, 4, and 8 subsets per iteration.
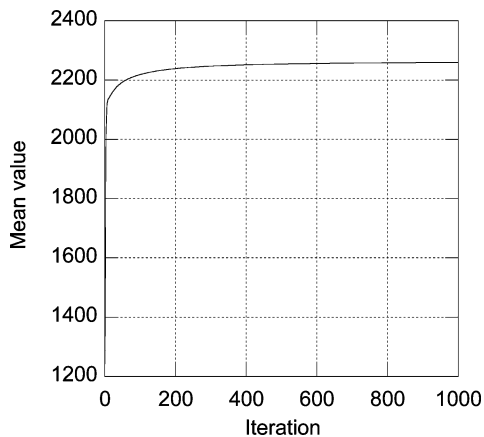


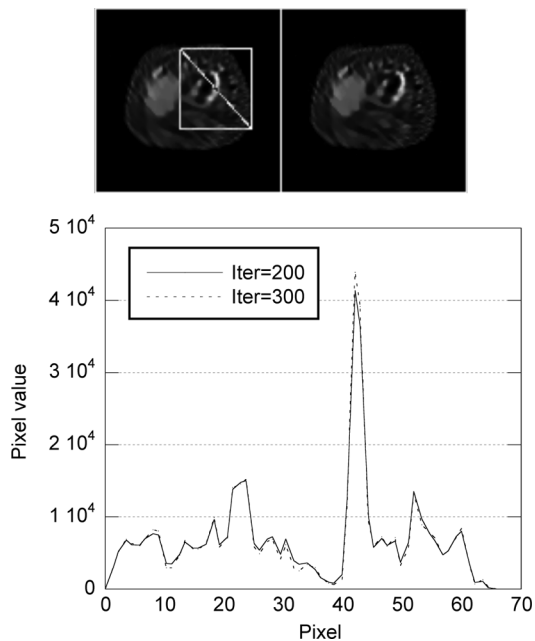Fig. 4.   The mean value of the reconstructed images with iterations.
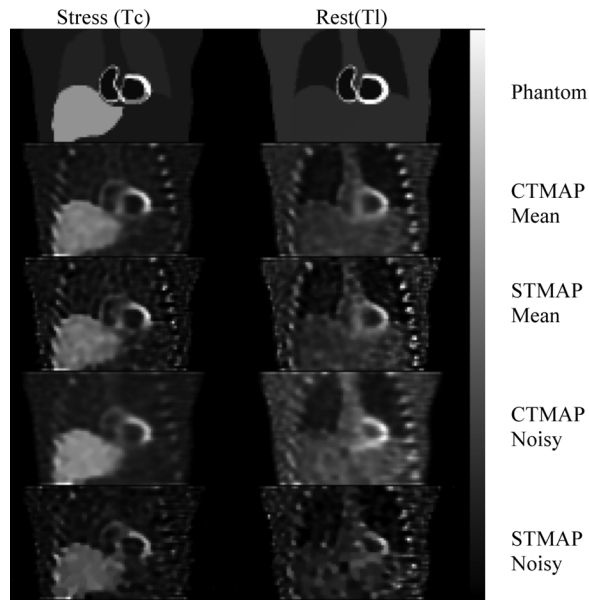


Fig. 5.   Profiles of stress images at iteration 200 and 300.



Fig. 6.   Comparison of images reconstructed using the cross-tracer (CTMAP) and single-tracer (STMAP) algorithms. The rows are, top to bottom, the phantom, mean of 50 noisy CTMAP and STMAP reconstructions, and sample noisy CTMAP and STMAP reconstructions. The columns are (left to right) images of the rest and stress distributions. A 35% contrast myocardial perfusion defect was present in the stress distribution, but there was no defect in the rest distribution. All reconstructions were obtained using 16 subsets, 100 iterations, $\beta = 0.05$, and $\delta = 10$.

that illustrate the similar sharpness. In Fig. 6, the sample noisy images from CTMAP appear considerably less noisy inside uniform regions like the liver. This is further corroborated by the profile through the standard deviation image, shown in Fig. 7,

where the standard deviation for CTMAP is on the order of 50% less than for STMAP.

Also note that the images in Fig. 6 contain reversible defects, i.e., a defect-present in the stress image, and no defect in the rest image. Although the cross-tracer prior couples the smoothness on both images, the existence of a defect in the stress image did not induce a shadow in the rest image, as can be seen by
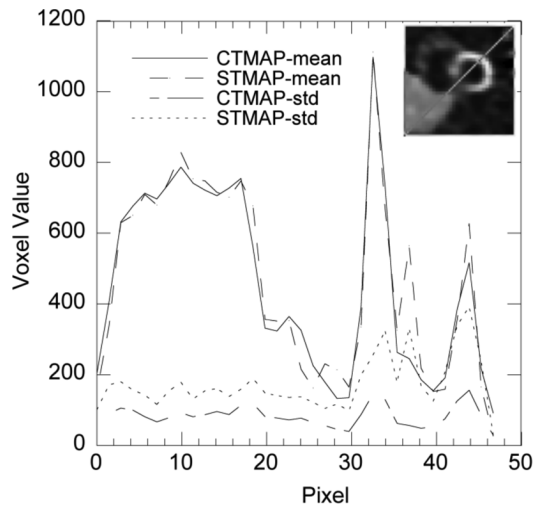
Fig. 7. Profiles through the mean stress image in Fig. 6 as well as standard deviation (std) images (not shown in Fig. 6) for the cross-tracer (CTMAP) and single-tracer (STMAP) algorithms at the position indicated in the inset image.



Fig. 8. Reconstructed stress images using STMAP (100 iterations, 16 subsets). The rows are for (top to bottom) $\beta = 0.2$, 0.5, 1.0, 2.0, and 5.0. The columns (left to right) are for $\delta^{-1/2} = 0.01$, 0.02, 0.05, 0.1, and 0.2.

comparing the mean rest images from CTMAP and STMAP in the defect region in the inferior wall of the left ventricle.

### C. Effects of the Cross-Tracer Prior—Results From Fidelity Study

Figs. 8 and 9 show the slice containing the center of the heart reconstructed with STMAP and CTMAP, respectively, for a range of $\beta$ and $\delta$ values. For the same values of $\beta$ and $\delta$, the cross-tracer reconstruction does a better job of preserving edges. This allows using a higher value of $\beta$, and thus increasing the smoothing in flat regions with the same edge preservation. For example, among all the combinations of $\beta$ and $\delta$, the one having the lowest MSE for CTMAP was $\beta = 1.0$ and $\delta = 0.1^{-0.5}$, while for STMAP the best was $\beta = 0.5$, $\delta = 0.2^{-0.5}$. Further, using these optimal combinations of priors, the MSE values for the stress and rest images were lower for CTMAP (stress 266.0, rest 275.0) than for STMAP (stress 297.2, rest 309.8). Both of these were substantially lower than the combination of OS-EM iteration and postreconstruction Butterworth filter cutoff frequency (five iterations and 0.20 pixel$^{-1}$) giving the lowest MSE (stress 406.3, rest 416.7). The rest and stress images for the optimal parameter values for each of the methods are shown, in comparison with the phantom images, in Fig. 10. Note in particular that the cross tracer prior did not induce shadow defects nor did it eliminate the real defect when appropriate values of the hyperparameters were used. These results indicate that, at least in terms of image fidelity, the MAP methods are better than OS-EM and, for the range of parameters investigated, the cross-tracer prior provides some advantage over a single-tracer prior having the same functional form.

### D. Three-Class Observer Study

For cross-tracer MAP reconstruction, we used $\delta = \eta = \sqrt{2}$. These values were chosen based on qualitative evaluation of the image quality and were not formally optimized. We did, however, optimize the weight of the prior, $\beta$, by computing the value of the VUS for four different values of $\beta$ in the range 0.001 to 0.4. The resulting VUS values are shown in Fig. 11. Fig. 12
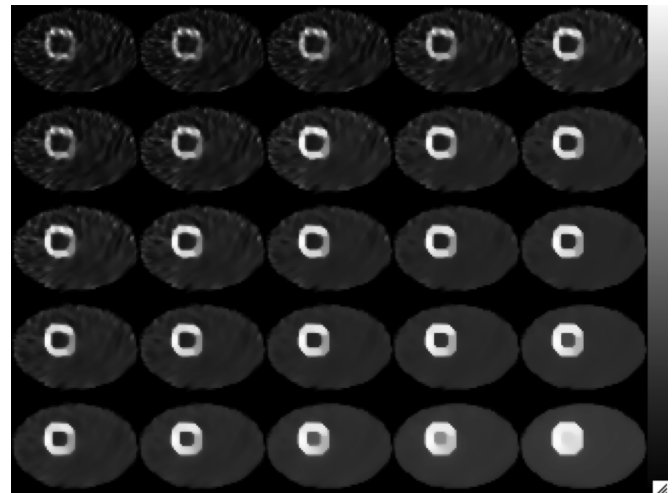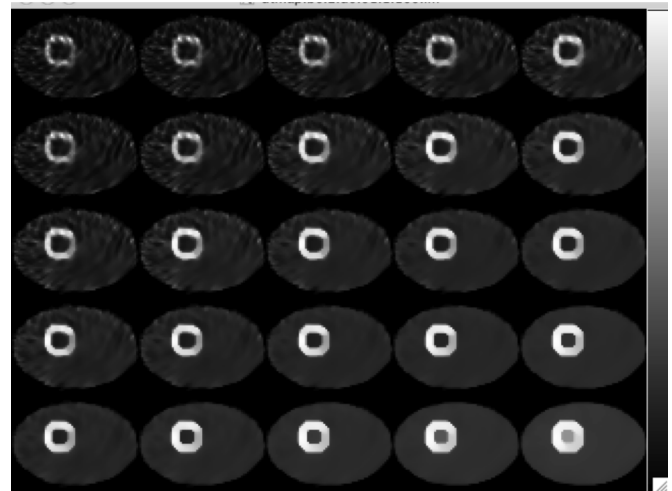


Fig. 9. Reconstructed stress images using CTMAP (35 iterations, 16 subsets). The rows are for (top to bottom) $\beta = 0.2$, 0.5, 1.0, 2.0, and 5.0. The columns (left to right) are for $\delta^{-1/2} = 0.01$, 0.02, 0.05, 0.1, and 0.2.
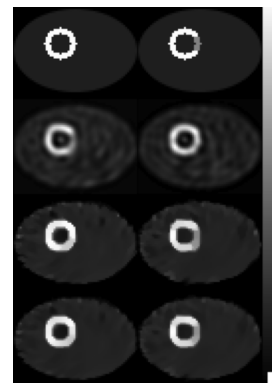


Fig. 10. Optimal Images: Columns (left to right) rest and stress. Rows (top to bottom): true image; OS-EM (5 iterations, 16 subsets, cutoff $= 0.2$ pixel$^{-1}$); STMAP ($\beta = 0.5$, $\delta^{1/2} = 0.2$); and CTMAP, ($\beta = 1.0$, $\delta^{1/2} = 0.1$).

shows the VUS values for OS-EM reconstruction with an order eight postreconstruction Butterworth filter as a function of the
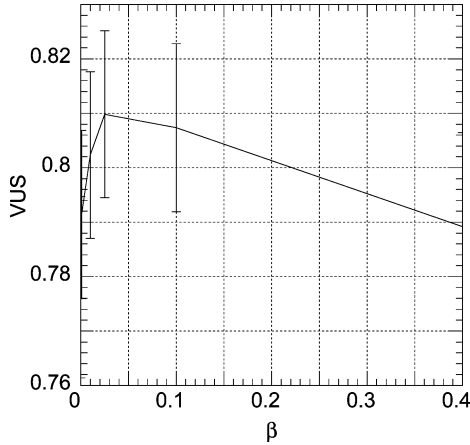
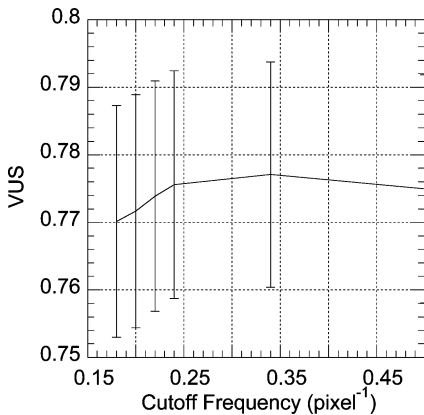Fig. 11. VUS values for cross-tracer MAP reconstruction.



Fig. 12. VUS values for OS-EM reconstruction with postreconstruction Butterworth filtering.

TABLE III
COMPARISON OF CROSS-TRACER MAP AND OS-EM WITH FILTERING USING OPTIMAL PARAMETERS

|  | VUS | Stddev |
|---|---|---|
| Cross-tracer MAP | 0.818 | 0.015 |
| OS-EM with filtering | 0.761 | 0.016 |

cutoff frequency. Table III shows a comparison of cross-tracer MAP and OS-EM with filtering using optimal parameters. It can be seen that cross-tracer MAP significantly improved the overall image quality as measured by the VUS. The $p$ value for this difference using a two-tailed z-test was smaller than 0.001, indicating that this difference was highly statistically significant.

Fig. 13 shows some sample reconstructed images. The parameters for the reconstruction are also shown. It can be seen that the cross-tracer reconstruction improved the contrast in the $^{201}$Tl image.

## IV. CONCLUSION

In simultaneous dual isotope nuclear medicine imaging protocols, two or more images are simultaneously obtained from a patient to characterize different functional or anatomical features. The use of simultaneous acquisition means that these images are perfectly registered in space and time. To exploit com-
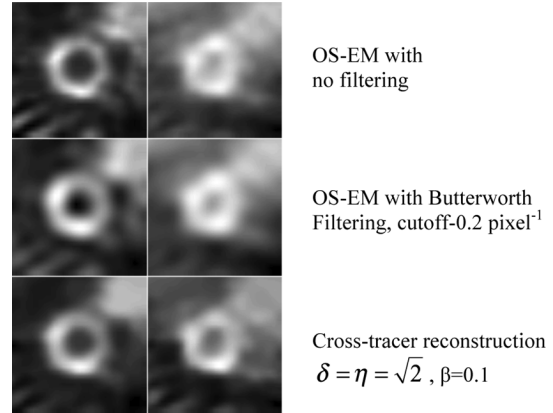


Fig. 13. Sample reconstructed images. All images were obtained using seven iterations.

monalities in these images, we have proposed a cross-tracer prior that couples the voxel values in two perfectly registered images. We then developed convergent MAP algorithms, with and without ordered subsets acceleration, to optimize the resulting objective function. The algorithms were applied to dual-isotope MPS imaging using data that did not include crosstalk, and were compared to OS-EM with respect to its performance in a three-class diagnostic task. The addition of crosstalk and crosstalk compensation to the evaluation would have been desirable. However, crosstalk combined with crosstalk compensation would increase the noise of the $^{201}$Tl images. Thus, one might hypothesize that this would increase the advantage of the cross-tracer prior. However, full evaluation of this remains a task for future work.

Results of both image resolution and noise and image fidelity studies showed that the proposed cross-tracer prior was superior to a single-tracer prior with the same functional form but without the cross-tracer coupling. Task-based evaluation studies showed that the proposed cross-tracer prior and the convergent algorithm substantially improved image quality for dual-isotope MPS images compared to OS-EM reconstruction with post-reconstruction filtering. In future work we will investigate more complete optimization of the hyperparameters of the cross-tracer prior and implementation of scatter and crosstalk compensation in the algorithm.

## APPENDIX

*Surrogate for the Cross-Tracer Prior:* The cross-tracer prior, $\lambda(x, y)$, is given by

$$\lambda(x, y) = \sqrt{1 + \left(\frac{x}{\delta}\right)^2 + \left(\frac{y}{\eta}\right)^2} - 1. \quad (A1)$$

Define $\bar{\lambda}(x, y)$ as a quadratic function

$$\bar{\lambda}(x, y) = a + \frac{1}{2}bx^2 + \frac{1}{2}cy^2 \quad (A2)$$

with the coefficients $a$, $b$, and $c$ to be determined such that

$$\begin{cases} \bar{\lambda}(x, y) \geq \lambda(x, y) \text{ for all the } x \\ \bar{\lambda}(x_0, y_0) = \lambda(x_0, y_0). \end{cases} \quad (A3)$$

These conditions imply that $\bar{\lambda}$ and $\lambda$ have a common tangent at $(x_0, y_0)$. Thus we obtain

$$\left.\frac{d\bar{\lambda}(x,y)}{dx}\right|_{(x_0,y_0)} = bx_0 = \left.\frac{d\lambda(x,y)}{dx}\right|_{(x_0,y_0)} = \lambda'_x(x_0,y_0) \tag{A4}$$

and

$$\left.\frac{d\bar{\lambda}(x,y)}{dy}\right|_{(x_0,y_0)} = cy_0 = \left.\frac{d\lambda(x,y)}{dy}\right|_{(x_0,y_0)} = \lambda'_y(x_0,y_0). \tag{A5}$$

Solving (A3), (A4) and (A5), for $a$, $b$, and $c$, and using the condition $\bar{\lambda}(x_0,y_0) = \lambda(x_0,y_0)$, we obtain

$$b = \frac{\lambda'_x(x_0,y_0)}{x_0} \tag{A6}$$

$$c = \frac{\lambda'_y(x_0,y_0)}{y_0} \tag{A7}$$

and

$$a = \lambda(x_0,y_0) - \frac{1}{2}x_0\lambda'_x(x_0,y_0) - \frac{1}{2}y_0\lambda'_y(x_0,y_0). \tag{A8}$$

Now we check the other condition in (A3), i.e., whether $\bar{\lambda}(x,y) \geq \lambda(x,y)$, for all $x$ and $y$. Denoting the difference of $\bar{\lambda}(x,y)$ and $\lambda(x,y)$, as $z(x,y)$, we see that

$$\begin{aligned}
z(x,y) &= \bar{\lambda}(x,y) - \lambda(x,y) \\
&= \lambda(x_0,y_0) - \frac{1}{2}x_0\lambda'_x(x_0,y_0) - \frac{1}{2}y_0\lambda'_y(x_0,y_0) \\
&\quad + \frac{1}{2}\frac{\lambda'_x(x_0,y_0)}{x_0}x^2 + \frac{1}{2}\frac{\lambda'_y(x_0,y_0)}{y_0}y^2 - \lambda(x,y).
\end{aligned} \tag{A9}$$

Substituting (A1) into (A9), A.II provides a proof that $z(x,y)$ achieves its minimum at $(x_0, y_0)$. From (A6)–(A8), we see that $\bar{\lambda}(x,y)$ is

$$\begin{aligned}
\bar{\lambda}(x,y) &= \left[\lambda(x_0,y_0) - \frac{1}{2}x_0\lambda'_x(x_0,y_0) - \frac{1}{2}y_0\lambda'_y(x_0,y_0)\right] \\
&\quad + \frac{1}{2}\frac{\lambda'_x(x_0,y_0)}{x_0}x^2 + \frac{1}{2}\frac{\lambda'_y(x_0,y_0)}{y_0}y^2. 
\end{aligned} \tag{A10}$$

We now define

$$\gamma_y(x,y) = \frac{\lambda'_y(x,y)}{y} \text{ and } \gamma_x(x,y) = \frac{\lambda'_x(x,y)}{x}. \tag{A11}$$

Equation (A10) can thus be rewritten as

$$\begin{aligned}
\bar{\lambda}(x,y) &= \left[\lambda(x_0,y_0) - \frac{1}{2}x_0\lambda'_x(x_0,y_0) - \frac{1}{2}y_0\lambda'_y(x_0,y_0)\right] \\
&\quad + \frac{1}{2}\gamma_x(x_0,y_0)x^2 + \frac{1}{2}\gamma_y(x_0,y_0)y^2 \\
&= \frac{1}{2}\gamma_x(x_0,y_0)\left[x^2 - 2xx_0 + x_0^2\right] + \lambda'_x(x_0,y_0)x \\
&\quad - \frac{1}{2}\lambda'_x(x_0,y_0)x_0 \\
&\quad + \frac{1}{2}\gamma_y(x_0,y_0)\left[y^2 - 2yy_0 + y_0^2\right] + \lambda'_y(x_0,y_0)y
\end{aligned}$$

$$\begin{aligned}
&\quad - \frac{1}{2}\lambda'_y(x_0,y_0)y_0 \\
&\quad + \lambda(x_0,y_0) - \frac{1}{2}x_0\lambda'_x(x_0,y_0) - \frac{1}{2}y_0\lambda'_y(x_0,y_0) \\
&= \frac{1}{2}\gamma_x(x_0,y_0)\left[x - x_0\right]^2 \\
&\quad + \frac{1}{2}\gamma_y(x_0,y_0)\left[y - y_0\right]^2 \\
&\quad + \lambda'_x(x_0,y_0)(x - x_0) + \lambda'_y(x_0,y_0)\left[y - y_0\right] \\
&\quad + \lambda(x_0,y_0). 
\end{aligned} \tag{A12}$$

Now let

$$x = x_j - x_k \text{ and } y = y_j - y_k. \tag{A13}$$

Similarly,

$$x_0 = \hat{x}_j - \hat{x}_k \text{ and } y_0 = \hat{y}_j - \hat{y}_k. \tag{A14}$$

From the relationship

$$\begin{aligned}
&[x_j - x_k - (\hat{x}_j - \hat{x}_k)]^2 \\
&= \left[\frac{1}{2}(2x_j - 2\hat{x}_j) + \frac{1}{2}(2\hat{x}_k - 2x_k)\right]^2 \\
&\leq \frac{1}{2}(2x_j - 2\hat{x}_j)^2 + \frac{1}{2}(2x_k - 2\hat{x}_k)^2 
\end{aligned} \tag{A15}$$

we can find a surrogate function for $\bar{\lambda}(x_j - x_k, y_j - y_k; \hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k)$, that is

$$\begin{aligned}
&\bar{\lambda}(x_j - x_k, y_j - y_k; \hat{x}_j - \hat{x}_k, y_j - y_k) \\
&\leq \lambda'_x(\hat{x}_j - \hat{x}_k, y_j - y_k)\left[x_j - x_k - (x_j - x_k)\right] \\
&\quad + \lambda'_y(\hat{x}_j - \hat{x}_k, y_j - y_k)\left[y_j - y_k - (y_j - y_k)\right] \\
&\quad + \lambda(\hat{x}_j - \hat{x}_k, y_j - y_k) \\
&\quad + \frac{1}{4}\gamma_x(\hat{x}_j - \hat{x}_k, y_j - y_k) \\
&\quad \times \left[(2x_j - 2x_j)^2 + (2x_k - 2x_k)^2\right] \\
&\quad + \frac{1}{4}\gamma_y(\hat{x}_j - \hat{x}_k, y_j - y_k) \\
&\quad \times \left[(2y_j - 2y_j)^2 + (2y_k - 2\hat{y}_k)^2\right]. 
\end{aligned} \tag{A16}$$

So, we now have a new surrogate function, $\phi(x_j, x_k, y_j, y_k)$, given by

$$\begin{aligned}
&\phi(x_j, x_k, y_j, y_k) \\
&= \lambda'_x(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k)\left[x_j - x_k - (\hat{x}_j - \hat{x}_k)\right] \\
&\quad + \lambda'_y(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k)\left[y_j - y_k - (\hat{y}_j - \hat{y}_k)\right] \\
&\quad + \lambda(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k) \\
&\quad + \frac{1}{4}\gamma_x(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k) \\
&\quad \times \left[(2x_j - 2\hat{x}_j)^2 + (2x_k - 2\hat{x}_k)^2\right] \\
&\quad + \frac{1}{4}\gamma_y(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k) \\
&\quad \times \left[(2y_j - 2\hat{y}_j)^2 + (2y_k - 2\hat{y}_k)^2\right]. 
\end{aligned} \tag{A17}$$

Rearranging (A17), we obtain,

$$
\begin{aligned}
\phi&(x_j, x_k, y_j, y_k) \\
&= \gamma_x(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k) \\
&\quad \times \left\{ \left[ x_j - \frac{1}{2}(x_j + x_k) \right]^2 + \left[ x_k - \frac{1}{2}(x_j + x_k) \right]^2 \right\} \\
&\quad + \gamma_y(\hat{x}_j - \hat{x}_k, y_j - y_k) \\
&\quad \times \left\{ \left[ y_j - \frac{1}{2}(y_j + y_k) \right]^2 + \left[ y_k - \frac{1}{2}(y_j + y_k) \right]^2 \right\} \\
&\quad + C(\hat{x}_j - \hat{x}_k, y_j - y_k) \quad\quad\quad\quad\quad\quad\quad \text{(A18)}
\end{aligned}
$$

where

$$
\begin{aligned}
C&(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k) \\
&= -\gamma_x(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k) \frac{1}{2}(\hat{x}_j - \hat{x}_k)^2 \\
&\quad - \gamma_y(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k) \frac{1}{2}(\hat{y}_j - \hat{y}_k)^2 \\
&\quad + \lambda(\hat{x}_j - \hat{x}_k, \hat{y}_j - \hat{y}_k)
\end{aligned}
$$

is a constant that is independent of $x_j$, $x_k$, $y_j$, and $y_k$. Now we substitute (A18) into (4), and the surrogate function of the prior term is

$$
\begin{aligned}
\Lambda^{(n)}&(\mathbf{x}, \mathbf{y}) \\
&= 2 \sum_{j}^{J} \sum_{k \in N_j} \omega_{jk} \left\{ \gamma_x(x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)}) \left[ x_j - \tfrac{1}{2}(x_j^{(n)} + x_k^{(n)}) \right]^2 \right\} \\
&\quad + 2 \sum_{j}^{J} \sum_{k \in N_j} \omega_{jk} \left\{ \gamma_y(x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)}) \left[ y_j - \tfrac{1}{2}(y_j^{(n)} + y_j^{(n)}) \right]^2 \right\} \\
&\quad + \sum_{j}^{J} \sum_{k \in N_j} \omega_{jk} C(x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)}) \quad\quad \text{(A19)}
\end{aligned}
$$

where $C\left(x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)}\right)$ is a constant that is independent of $\mathbf{x}$ and $\mathbf{y}$. Note that in (A19) we used $x_j^{(n)}$ and $y_j^{(n)}$ to replace $\hat{x}_j$ and $\hat{x}_j$, respectively, for notational purposes. In particular, the superscript $(n)$ denotes the estimate after the $n$th iteration. Simplifying $\Lambda^{(n)}(\mathbf{x}, \mathbf{y})$, we have

$$
\begin{aligned}
\Lambda^{(n)}(\mathbf{x}, \mathbf{y}) &= 2 \sum_{j=1}^{J} \sum_{k \in N_j} h_{jk}^{(n)}(x_j) \\
&\quad + 2 \sum_{j=1}^{J} \sum_{k \in N_j} g_{jk}^{(n)}(y_j) + C^{(n)} \quad\quad \text{(A20)}
\end{aligned}
$$

where

$$
\begin{aligned}
h_{jk}^{(n)}(t) &\triangleq \omega_{jk} \gamma_x(x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)}) \\
&\quad \times \left[ t - \frac{1}{2}\left( x_j^{(n)} + x_k^{(n)} \right) \right]^2 \\
g_{jk}^{(n)}(t) &\triangleq \omega_{jk} \gamma_y(x_j^{(n)} - x_k^{(n)}, y_j^{(n)} - y_k^{(n)}) \\
&\quad \times \left[ t - \frac{1}{2}\left( y_j^{(n)} + y_k^{(n)} \right) \right]^2
\end{aligned}
$$

and $C^{(n)}$ is a constant term that does not depend on $x_j$, $x_k$, $y_j$, and $y_k$.

$z(x, y)$ *in (A9) Achieves its Minimum at* $(x_0, Y_0)$: Substituting (A1) into (A9) and simplifying, $z(x, y)$ becomes (A21), shown at the bottom of the page.

In (A21), the denominator is greater than zero. Since $s = \sqrt{t}$ is a monotonically increasing function, proving the numerator is greater than zero is equivalent to proving that

$$
\begin{aligned}
\left[ 2 + \left( \frac{x_0}{\delta} \right)^2 + \left( \frac{y_0}{\eta} \right)^2 \right]^2 \\
\geq 4 \left( 1 + \left( \frac{x_0}{\delta} \right)^2 + \left( \frac{y_0}{\eta} \right)^2 \right) \left( 1 + \left( \frac{x}{\delta} \right)^2 + \left( \frac{y}{\eta} \right)^2 \right)
\end{aligned}
$$
(A22)

Simplifying (A22), we see that the inequality in (A22) is equivalent to

$$
\left[ \left( \left( \frac{x_0}{\delta} \right)^2 + \left( \frac{y_0}{\eta} \right)^2 \right) - \left( \left( \frac{x}{\delta} \right)^2 + \left( \frac{y}{\eta} \right)^2 \right) \right]^2 \geq 0. \quad \text{(A23)}
$$

Since the left-hand side of (A23) is a quadratic function, it is true for any $x$ and $y$, and thus the numerator of (A21) is greater than zero, and

$$
z(x, y) \geq 0. \quad\quad\quad\quad \text{(A24)}
$$

$$
z(x, y) = \frac{2 + \left( \frac{x_0}{\delta} \right)^2 + \left( \frac{y_0}{\eta} \right)^2 - 2\sqrt{\left( 1 + \left( \frac{x_0}{\delta} \right)^2 + \left( \frac{y_0}{\eta} \right)^2 \right) \left( 1 + \left( \frac{x}{\delta} \right)^2 + \left( \frac{y}{\eta} \right)^2 \right)}}{2\sqrt{\left( 1 + \left( \frac{x_0}{\delta} \right)^2 + \left( \frac{y_0}{\eta} \right)^2 \right)}}
$$
(A21)

## REFERENCES

[1] X. Song, E. C. Frey, W. T. Wang, Y. Du, and B. M. W. Tsui, "Validation and evaluation of model-based crosstalk compensation method in simultaneous Tc-99m stress and Tl-201 rest myocardial perfusion SPECT," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 1, pp. 72–79, Feb. 2004.

[2] W. T. Wang, E. C. Frey, B. M. W. Tsui, C. Tocharoenchai, and W. H. Baird, "Parameterization of Pb X-ray contamination in simultaneous Tl-201 and Tc-99m dual-isotope imaging," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 3, pp. 680–692, Jun. 2002.

[3] D. J. Kadrmas, E. C. Frey, and B. M. W. Tsui, "Simultaneous technetium-99m/thallium-201 SPECT imaging with model-based compensation for cross-contaminating effects," *Phys. Med. Biol.*, vol. 44, pp. 1843–1860, Jul. 1999.

[4] X. He, X. Y. Song, and E. C. Frey, "Application of three-class ROC analysis to task-based image quality assessment of simultaneous dual-isotope myocardial perfusion SPECT (MPS)," *IEEE Trans. Med. Imag.*, vol. 27, no. 11, pp. 1556–1567, Nov. 2008.

[5] I.-T. Hsiao, A. Rangarajan, and G. R. Gindi, "Provably convergent OSEM-like reconstruction algorithm for emission tomography," in *SPIE Med. Imag. 2002: Image Process.*, 2002, pp. 10–19.

[6] A. R. De Pierro and M. E. B. Yamagishi, "Fast EM-like methods for maximum "a posteriori" estimates in emission tomography," *IEEE Trans. Med. Imag.*, vol. 20, no. 4, pp. 280–288, Apr. 2001.

[7] A. R. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 132–137, Mar. 1995.

[8] S. Ahn and J. A. Fessler, "Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 613–626, May 2003.

[9] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *J. Computat. Graphical Stat.*, vol. 9, pp. 1–20, 2000.

[10] J. H. Chang, J. M. M. Anderson, and J. R. Votaw, "Regularized image reconstruction algorithms for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 23, no. 9, pp. 1165–1175, Sep. 2004.

[11] J. Browne and A. R. De Pierro, "A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 687–699, Oct. 1996.

[12] A. Gunawardana, "The information geometry of EM variants for speech and image processing," Ph.D. dissertation, Dept. Electrical Computer Eng., Johns Hopkins Univ., Baltimore, MD, USA, 2001.

[13] R. M. Neal and G. E. Hinton, "A view of the EM algorithm that justifies incremental, sparse, and other variants," in *Proceedings of the NATO Advanced Study Institute on Learning in Graphical Models*, M. I. Jordan, Ed. Norwell, MA: Kluwer Academic, 1998, pp. 355–368.

[14] I. T. Hsiao, A. Rangarajan, and G. Gindi, "A new convergent map reconstruction algorithm for emission tomography using ordered subsets and separable surrogates," in *Proc. 2002 IEEE Int. Symp. Biomed. Imag.*, 2002, pp. 409–412.

[15] I. T. Hsiao, A. Rangarajan, P. Khurd, and G. Gindi, "An accelerated convergent ordered subsets algorithm for emission tomography," *Phys. Med. Biol.*, vol. 49, pp. 2145–2156, Jun. 2004.

[16] I. T. Hsiao, P. Khurd, A. Rangarajan, and G. Gindi, "An overview of fast convergent ordered-subsets reconstruction methods for emission tomography based on the incremental EM algorithm," *Nucl. Instrum. Methods Phys. Res. Section A-Accelerators Spectrometers Detectors Associated Equipment*, vol. 569, pp. 429–433, Dec. 2006.

[17] A. Rangarajan, P. Khurd, I. T. Hsiao, and G. Gindi, Convergence proofs for COSEM-ML and COSEM-MAP algorithms Tech. Rep. MIPL-03–1, 2003.

[18] H. Erdogan and J. A. Fessler, "Joint estimation of attenuation and emission images from PET scans," in *IEEE Med. Imag. Conf.*, 2000, pp. 1672–1675.

[19] A. R. De Pierro, "On the relation between the ISRA and the EM algorithm for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 12, no. 2, pp. 328–333, Jun. 1993.

[20] A. Gunawardana and W. Byrne, "Convergence theorems for generalized alternating minimization procedures," *J. Mach. Learn. Res.*, vol. 6, pp. 2049–2073, 2005.

[21] S. Ahn, J. A. Fessler, D. Blatt, and A. O. Hero, "Convergent incremental optimization transfer algorithms: Application to tomography," *IEEE Trans. Med. Imag.*, vol. 25, no. 3, pp. 283–96, Mar. 2006.

[22] X. He, E. C. Frey, J. M. Links, K. L. Gilland, W. P. Segars, and B. M. W. Tsui, "A mathematical observer study for the evaluation and optimization of compensation methods for myocardial SPECT using a phantom population that realistically models patient variability," *IEEE Trans. Nucl. Sci.*, vol. 51, pp. 218–224, Feb. 2004.

[23] W. P. Segars, B. M. Tsui, D. S. Lalush, E. C. Frey, M. A. King, and D. Manocha, "Development and application of the new dynamic nurbs-based Cardiac-Torso (NCAT) phantom," *J. Nucl. Med.*, vol. 42, pp. 7p–7p, May 2001.

[24] A. B. Barclay, R. L. Eisner, and E. V. DiBella, PET thorax model database [Online]. Available: http://www.emory.edu/CRL/abb/thorax-model/Contents.html

[25] X. He, J. M. Links, K. L. Gilland, B. M. W. Tsui, and E. C. Frey, "Comparison of 180 degrees and 360 degrees acquisition for myocardial perfusion spect with compensation for attenuation, detector response, and scatter: Monte Carlo and mathematical observer results," *J. Nucl. Cardiol.*, vol. 13, pp. 345–353, May-Jun. 2006.

[26] X. Song, W. P. Segars, Y. Du, B. M. W. Tsui, and E. C. Frey, "Fast modelling of the collimator-detector response in Monte Carlo simulation of SPECT imaging using the angular response function," *Phys. Med. Biol.*, vol. 50, pp. 1791–1804, Apr. 2005.