Title

A generalization of the maximum likelihood expectation maximization (MLEM) method: Masked-MLEM

Journal

Physics in Medicine and Biology, 68(24)

ISSN

0031-9155

Authors

Zheng, Yifan

Frame, Emily

Caravaca, Javier

et al.

Publication Date

2023-12-21

DOI

10.1088/1361-6560/ad0900

Peer reviewed

HHS Public Access

Author manuscript

Phys Med Biol. Author manuscript; available in PMC 2024 December 11.

Published in final edited form as:

Phys Med Biol. ; 68(24): . doi:10.1088/1361-6560/ad0900.

A Generalization of the Maximum Likelihood Expectation Maximization (MLEM) Method: Masked-MLEM

Yifan Zheng[1,2,4], **Emily Frame**[2], **Javier Caravaca**[1], **Grant T. Gullberg**[1], **Kai Vetter**[2,3], **Youngho Seo**[1,2]

[1]Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA 94143, USA

[2]Department of Nuclear Engineering, University of California, Berkeley, CA 94720, USA

[3]Applied Nuclear Physics Group, Lawrence Berkeley National Laboratory, Berkeley, CA 94502, USA

Abstract

Objective.—In our previous work on image reconstruction for single-layer collimatorless scintigraphy, we developed the min-min weighted robust least squares (WRLS) optimization algorithm to address the challenge of reconstructing images when both the system matrix and the projection data are uncertain. Whereas the WRLS algorithm has been successful in two-dimensional (2D) reconstruction, expanding it to three-dimensional (3D) reconstruction is difficult since the WRLS optimization problem is neither smooth nor strongly-convex. To overcome these difficulties and achieve robust image reconstruction in the presence of system uncertainties and projection noise, we propose a generalized iterative method based on the maximum likelihood expectation maximization (MLEM) algorithm, hereinafter referred to as the Masked-MLEM algorithm.

Approach.—In the Masked-MLEM algorithm, only selected subsets ("masks") from the system matrix and the projection contribute to the image update to satisfy the constraints imposed by the system uncertainties. We validate the Masked-MLEM algorithm and compare it to the standard MLEM algorithm using experimental data obtained from both collimated and uncollimated imaging instruments, including parallel-hole collimated SPECT, 2D collimatorless scintigraphy, and 3D collimatorless tomography. Additionally, we conduct comprehensive Monte Carlo simulations for 3D collimatorless tomography to further validate the effectiveness of the Masked-MLEM algorithm in handling different levels of system uncertainties.

Main Results.—The Masked-MLEM and standard MLEM reconstructions are similar in cases with negligible system uncertainties, whereas the Masked-MLEM algorithm outperforms the standard MLEM algorithm when the system matrix is an approximation. Importantly, the Masked-MLEM algorithm ensures reliable image reconstruction across varying levels of system uncertainties.

[4]Author to whom any correspondence should be yifanzheng@berkeley.edu.

Submitted to: *Phys. Med. Biol.*

Significance.—With a good choice of system uncertainty and without requiring accurate knowledge of the actual system matrix, the Masked-MLEM algorithm yields more robust image reconstruction than the standard MLEM algorithm, effectively reducing the likelihood of erroneously reconstructing higher activities in regions without radioactive sources.

Keywords

Masked-MLEM; robust image reconstruction; system uncertainty; WRLS

1. Introduction

Generating high-quality images is an essential procedure for obtaining structural or functional information in many imaging modalities, such as computed tomography (CT), single photon emission computed tomography (SPECT), and positron emission tomography (PET) (Defrise & Gullberg 2006). In emission and transmission tomography, image reconstruction is performed using either analytical or iterative methods (Hsieh, Nett, Yu, Sauer, Thibault & Bouman 2013, Zeng 2001). Iterative reconstruction often yields better image quality and offers several advantages over analytical reconstruction. These include the flexibility for reconstruction using incomplete data or irregular sampling, and the capability to directly incorporate imaging physics and projection noise into the reconstruction process (Hsieh et al. 2013, Zeng 2001, Beister, Kolditz & Kalender 2012, Yan, Bui, Cong & Vese 2013). A popular iterative reconstruction algorithm aimed at finding the most likely image given the observed data is maximum likelihood expectation maximization (MLEM) (Shepp & Vardi 1982, Lange, Carson et al. 1984), or its variant, ordered subsets expectation maximization (OSEM) (Hudson & Larkin 1994). MLEM-based algorithms are particularly suitable for imaging problems with Poisson noise characteristics and enforce the non-negativity of reconstructed images, which are desirable properties in PET or SPECT reconstructions (Lange et al. 1984, Zeng 2001).

Mathematically, image reconstruction addresses a linear ill-posed inverse problem to recover an unknown image from incomplete and noisy measurements (Hsieh et al. 2013, Yan et al. 2013). Typically, this process is structured as a convex optimization problem using Tikhonov regularization, where the data fidelity term is a least-squares (LS) optimization problem and the additive regularization term imposes some prior knowledge on the image to be reconstructed (Tikhonov 1963, Panin, Zeng & Gullberg 1998, Yan et al. 2013). With a growing research interest in image priors, various regularization terms including total variation (TV), nonlocal means (NLM), and deep-learning-based priors, have been successfully integrated into iterative algorithms whereas assuming a known and accurate system response in the data fidelity term (Panin et al. 1998, Sawatzky, Brune, Wubbeling, Kosters, Schafers & Burger 2008, Chen, Qi, Wu, Xu & Zhou 2016, Gong, Catana, Qi & Li 2018). In reality, the system matrix used for image reconstruction suffers from uncertainties due to complicated physical effects (Liu, Wang, Gao, Tian, Chen, Hu & Shi 2012). Previous studies has investigated the impact of noise in the system matrix and shown that these uncertainties in the system matrix can propagate to the reconstructed images (Qi & Huesman 2004, Rafecas, Boning, Pichler, Lorenz, Schwaiger & Ziegler 2004). Therefore, in addition to incorporating projection statistical models and image priors, incorporating constraints on

system uncertainties is expected to further improve the fidelity of the reconstructed images and yield robust reconstructions.

Due to the large dimensionality and ill-posedness of the system matrix, considering the system uncertainties explicitly during the reconstruction process is difficult and computationally expensive. Because of this, only a few studies have investigated system uncertainties in image reconstruction (Qi & Huesman 2004, Rafecas, Boning, Pichler, Lorenz, Schwaiger & Ziegler 2004, Liu et al. 2012, Kucharczak, Loquin, Buvat, Strauss & Mariano-Goulart 2018, Zheng, Huh, Su, Wang, Lin, Vetter & Seo 2020, Lunz, Hauptmann, Tarvainen, Schonlieb & Arridge 2021). Instead of directly addressing the system uncertainties, previous studies have adopted alternative approaches to account for system uncertainties indirectly (Rico, Strauss & Mariano-Goulart 2009, Rico & Strauss 2010, Kucharczak et al. 2018, Lunz et al. 2021). By transforming the system uncertainties to the projection confident intervals, the non-additive interval-based EM algorithm has been developed to reconstruct image confidence intervals based on the projection confident intervals, providing an estimate of the uncertainties in the reconstructed images (Rico et al. 2009, Rico & Strauss 2010, Kucharczak et al. 2018). An alternative approach to transforming system uncertainties into projection uncertainties is by training a neural network to learn a correction for the projection data, which can effectively provide a correction for the system matrix (Lunz et al. 2021). Deep-learning-based approaches take advantage of the learning capabilities of neural networks to model the complex system uncertainties and the projection uncertainties, but they require a large amount of training data to effectively learn and generalize.

In our previous work, we developed a min-min weighted robust least squares (WRLS) algorithm to directly address the uncertainties present in both the projection data and the system matrix (Zheng et al. 2020). We demonstrated the success of the WRLS algorithm in two-dimensional (2D) collimatorless scintigraphy using the CVXPY toolkit for iterative reconstruction (Zheng et al. 2020, Diamond & Boyd 2016). However, the WRLS optimization problem is neither smooth nor strongly-convex (Boyd & Vandenberghe 2004). When applying some traditional gradient-based algorithms to solve the WRLS optimization problem, the non-smoothness will introduce difficulties in differentiating the objective function, and the lack of strong convexity will lead to multiple local optima, making it difficult to find the global optimum and resulting in slow convergence. As image dimensionality increases, the optimization landscape becomes more complex, and finding the global optimum is more challenging. For example, in three-dimensional (3D) tomographic reconstruction, the CVXPY toolkit is computationally memory-intensive and time-consuming to handle such a large-scale WRLS optimization problem even on a dedicated server. Without an efficient iterative solver, the applications of the WRLS algorithm will be limited.

In this study, we propose the Masked-MLEM algorithm as a generalized iterative MLEM approach to achieve robust image reconstruction in the presence of system uncertainties and projection noise. We apply the Masked-MLEM algorithm for image reconstruction using experimental data obtained from three distinct imaging modalities: parallel-hole collimated SPECT, 2D collimatorless scintigraphy, and 3D collimatorless tomography. Furthermore, we

validate the performance of the Masked-MLEM algorithm through Monte Carlo simulations using Geant4 for 3D collimatorless tomography with an ideal ring detector geometry (Agostinelli, Allison, Amako, Apostolakis, Araujo, Arce, Asai, Axen, Banerjee, Barrand et al. 2003). The images reconstructed using the Masked-MLEM algorithm are compared with those using the standard MLEM algorithm to examine the robustness of the Masked-MLEM algorithm.

2. Materials and Methods

2.1. Masked-MLEM

The Masked-MLEM algorithm is derived based on the WRLS algorithm, where "Masked" refers to selecting subsets in both the system matrix and the projection in order to satisfy the constraints on the system uncertainties, and using the selected subsets for the image update. To provide a comprehensive understanding of the Masked-MLEM algorithm under the assumption of Poisson noise in projection, we begin by revisiting the WRLS algorithm considering Gaussian noise in projection. In our previous work, the WRLS algorithm took into account the uncertainties in both the projection data y and the system matrix A by assuming a box uncertainty set such that element-wisely $A_{lb} \preceq A \preceq A_{ub}$, where A_{lb} and A_{ub} are the lower and upper bounds of A respectively (Zheng et al. 2020). It is noteworthy that in the context of matrices and vectors, the notation $\preceq$ represents element-wise less than or equal to, and the reverse relationship applies as well. Essentially, the box uncertainty set for A leads to a feasibility finding problem given by equation (1), whose goal is to seek a non-negative image x that satisfies the system uncertainty constraints without optimizing an objective function as we define it as a constant equal to 1, thereby incorporating the system uncertainties into the image reconstruction process.

$$\begin{aligned} &\min_{x \succeq 0} 1 \text{ s.t. } \quad A_{lb}x \preceq y \preceq A_{ub}x \\ \Leftrightarrow &\min_{x \succeq 0} 1 \text{ s.t. } \quad A_u x - y_u \preceq 0, \end{aligned} \tag{1}$$

where

$$A_u = \begin{bmatrix} A_{lb} \\ -A_{ub} \end{bmatrix}, \quad y_u = \begin{bmatrix} y \\ -y \end{bmatrix}.$$

By relaxing the strict equality constraint $Ax = y$ to the additional soft constraints in equation (1) based on the system uncertainties, the feasible reconstructed image space can be expanded and the reconstruction process has more flexibility to explore a larger solution space when the exact solution for $Ax = y$ may not be feasible.

To find the optimal image while accounting for system uncertainties, the iterative image update process involves penalizing those instances of x that violate the constraints in equation (1) (i.e., $A_u x - y_u \succ 0$), which forms the fundamental principle underlying both

the WRLS and Masked-MLEM algorithms. We first reformulate the WRLS optimization problem in (Zheng et al. 2020) as

$$\min_{x \succeq 0, \mu \succeq 0} \frac{1}{2} \| \mu \|^2_{\Sigma^{-1}} \quad \text{s.t.} \quad \begin{bmatrix} I \\ I \end{bmatrix} \mu \succeq A_u x - y_u, \tag{2}$$

where $\sum$ is a diagonal matrix representing an additive Gaussian noise in y, $I \in \mathbb{R}^{m \times m}$ is the identity matrix with m being the number of measurements, and μ is a slack variable introduced to make the objective function convex (Boyd & Vandenberghe 2004). Instead of applying any optimization techniques like the CVXPY toolkit to directly solve the WRLS optimization problem, we decompose it into smaller subproblems and address the challenges posed by non-smoothness and lack of strong convexity within these subproblems. Looking at equation (2), if we can find some x^* satisfying $A_u x^* - y_u \preceq 0$ (i.e., a feasible solution for equation (1)), then the optimal solutions for equation (2) will be $\mu^* = 0$ and x^*, i.e., x^* is the final reconstructed image. Otherwise, when some entries in the vector $A_u x - y_u$ are positive, the objective function in equation (2) is to minimize the ℓ_∞ norm of the positive entries in $A_u x - y_u$ (i.e., $\forall i = 1, \ldots, m$, $\max\{(a_{lb})_i^T x - y_i, y_i - (a_{ub})_i^T x, 0\}$, see (Zheng et al. 2020)). Since the ℓ_∞ norm optimization is neither smooth nor strongly-convex, solving equation (2) in large-scale image reconstruction problems is computationally intractable and prone to convergence issues.

To overcome these challenges, one optimization approach is to relax the ℓ_∞ norm with a smooth convex surrogate, such as the ℓ_2 norm (i.e., $\forall i = 1, \ldots, m$, $\sqrt{\max\{(a_{lb})_i^T x - y_i, 0\}^2 + \max\{y_i - (a_{ub})_i^T x, 0\}^2}$). This relaxation extends the objective function in equation (2) to a larger and smooth function. After relaxation, instead of optimizing the original objective function in equation (2), we optimize the upper bound of equation (2) given by equation (3) (see proof below).

$$\begin{aligned} &\min_{x \succeq 0} \frac{1}{2} \| \max\{A_u x - y_u, 0\} \|^2_{\Sigma^{-1}} \\ = &\min_{x \succeq 0} \frac{1}{2} \| M|A_u| x - M|y_u| \|^2_{\Sigma^{-1}}, \end{aligned} \tag{3}$$

where $M = \text{diag}\{A_u x - y_u > 0\}$ is a diagonal matrix with its ith diagonal element taking the value 1 if $(A_u x - y_u)_i \geq 0$, and 0 otherwise, $\forall i = 1, \ldots, 2m$. This matrix essentially functions as a "mask" that operates on A_u and y_u, facilitating the selection of the positive subsets in $A_u x - y_u$ while nullifying the non-positive subsets. Given that half of the elements in A_u and y_u are negative (see their definitions in equation (1)), which mathematically makes the constraints in equations (1) and (2) convex but is physically trivial within the context of the ℓ_2 norm in equation (3), we take the element-wise absolute values of MA_u and My_u as $M|A_u|$ and $M|y_u|$ to make the selected subsets in A_u and y_u non-negative.

Proof.—Equation (2) $\overset{\text{See (Zheng et al. 2020)}}{\Longleftrightarrow} \min_{x \geq 0} \sum_{i=1}^{m} \frac{1}{2\sigma_i} \max\left\{(a_{lb})_i^T x - y_i, y_i - (a_{ub})_i^T x, 0\right\}^2$ Let $v_1 = \max\left\{(a_{lb})_i^T x - y_i, 0\right\}$, $v_2 = \max\left\{y_i - (a_{ub})_i^T x, 0\right\}$ we have:

$\Rightarrow \max\left\{(a_{lb})_i^T x - y_i, y_i - (a_{ub})_i^T x, 0\right\} = \max\{v_1, v_2\} \overset{\ell_\infty \to \ell_2}{\leq} \sqrt{v_1^2 + v_2^2} \therefore$ Equation (2) $\overset{\ell_\infty \to \ell_2}{\leq} \min_{x \geq 0} \sum_{i=1}^{m} \frac{1}{2\sigma_i} \left(\max\left\{(a_{lb})_i^T x - y_i, 0\right\}^2 + \max\left\{y_i - (a_{ub})_i^T x, 0\right\}^2\right) \Leftrightarrow$ equation (3) ■

The form of equation (3) is a generalized data fidelity term that takes into account the system uncertainty A_u and the additive Gaussian noise in y. It can be solved efficiently by a number of iterative algorithms used for LS optimization, where the modification is to apply a mask to select the positive subsets in $A_u x - y_u$ for the image update at each iteration, as if the actual system matrix is $M|A_u|$ and the projection is $M|y_u|$ for LS optimization. The values of the unselected subsets in A_u and y_u are zeroed out at each iteration, and thus their values will not contribute to the image update. After the relaxation, the optimal solution to equation (3) will be different from equation (2). We have tested that the difference between the reconstructed images using equation (2) and equation (3) can be negligible considering the intrinsic image resolution and projection noise.

Inspired by equation (3), wherein we consider the actual system matrix as $M|A_u|$ and the projection as $M|y_u|$ in each iteration (representing a selected subset of the original A_u and y_u), when we assume Poisson noise in y, we derive the Masked-MLEM algorithm based on the MLEM algorithm as shown in Algorithm 1 (see Appendix A for derivation details). The Masked-MLEM algorithm is a generalization of the standard MLEM algorithm, i.e., MLEM reduces to Masked-MLEM (Dasgupta, Papadimitriou & Vazirani 2008). The validity of this reduction is straightforward: if we know how to solve Masked-MLEM, by setting $A_{lb} = A = A_{ub}$ and running Masked-MLEM, we will get the solution of MLEM.

In Algorithm 1, for each iteration in the outer loop N_{out}, we first check the condition whether $A_u x^s - y_u$ is element-wisely non-positive. When this condition is met, we have found the optimal solution for equation (3) and the returned x^s stands as the final reconstructed image. This step serves as a crucial sanity check and the actual improvements to the image come from the subsequent masked update phase when $A_u x^s - y_u \succ 0$. In this case, we select a mask M such that only the selected subsets in A_u and y_u will contribute to the iterative update of the image x^s. This controlled masking strategy can be viewed as a form of regularization and offers a mechanism to refine and optimize the image reconstruction process, directing computational efforts towards

Algorithm 1 Masked-MLEM (N_{out}, N_{in})

Initialize x^0 to be a positive vector.
for $s = 1, \ldots, N_{out}$ **do**
  if $A_u x^s - y_u \preceq 0$ **then**
    ▷ We find the optimal solution for equation (3).
    return x^s
  else
    $M \leftarrow \mathrm{diag}\{A_u x^s - y_u \succ 0\}$
    $x^k \leftarrow x^s$
    for $k = 1, \ldots, N_{in}$ **do**
      $x^{k+1} = \frac{x^k}{(M|A_u|)^T \mathbf{1}} (M|A_u|)^T \frac{M|y_u|}{M|A_u|x^k}$
      ▷ $|\cdot|$ is the element-wise absolute value.
    $x^s \leftarrow x^k$
return x^s

areas of significance within the given projection data and system uncertainties. Unlike some deep-learning-based denoising methods, the mask M is dynamically determined at each iteration rather than being randomly selected. This is because the dynamic adaptability of the mask M is specifically tailored to the system uncertainty constraints at each iteration. We have checked that relying on predefined or random masking patterns can lead to convergence challenges and inaccurate reconstructions, potentially guiding the optimization process towards incorrect gradient descent directions during image updates. To speed up reconstruction, we can use the same M to update the image for N_{in} iterations, rather than updating M whenever the image is updated, since updating M may be computationally expensive for huge matrices and the selected masks at subsequent iterations may be similar. We have checked that the Masked-MLEM algorithm converges when using appropriate A_u, N_{out}, and N_{in} by monitoring the loss curve during the reconstruction process. The reconstruction loss is defined as $\frac{\| x^{k+1} - x^k \|_2}{\| x^k \|_2}$, where x^{k+1} and x^k represent the reconstructed images at iteration $k+1$ and k respectively, and serves as the metric for convergence throughout the iterations. For the convergence of the Masked-MLEM algorithm, N_{in} should be determined empirically through experimentation and should take a small value (usually less than 10 depending on reconstruction problems), otherwise the mask M will be outdated and not suitable for future image updates.

The Masked-MLEM algorithm is also compatible with ordered subsets (OS), i.e., Masked-OSEM. If we update the image using OS from the mask M, it does not guarantee convergence since M varies with the image update. For the purpose of convergence, we use the OS to calculate masks, followed by the Masked-MLEM algorithm. The detailed Masked-OSEM algorithm is shown in Algorithm 2.

Algorithm 2 Masked-OSEM (N_{out}, N_{in}, N_{subset})

Initialize x^0 to be a positive vector.
for $s = 1, \ldots, N_{out}$ **do**
  $sensitivity \leftarrow 0$
  $sum \leftarrow 0$
  for $i = 1, \ldots, N_{subset}$ **do**
    if $\max\{(A_u)_i x^s - (y_u)_i, 0\} > 0$ **then**
      $M \leftarrow \mathrm{diag}\{(A_u)_i x^s - (y_u)_i \succ 0\}$
      $sensitivity \leftarrow sensitivity + (M|(A_u)_i|)^T \mathbf{1}$
      $sum \leftarrow sum + (M|(A_u)_i|)^T \frac{M|(y_u)_i|}{M|(A_u)_i|x^s}$
▷ $|\cdot|$ is the element-wise absolute value.
  if $sum = 0$ **then**
▷ We find the optimal solution for equation (3).
    return x^s
  $x^k \leftarrow x^s$
  for $k = 1, \ldots, N_{in}$ **do**
▷ Perform the EM step.
    $x^{k+1} = \frac{x^k}{sensitivity} \times sum$
  $x^s \leftarrow x^k$
return x^s

2.2. System Uncertainties

To apply the Masked-MLEM algorithm, we design the system matrix upper and lower bounds for both collimated and uncollimated imaging modalities. For the parallel-hole collimated SPECT, we use the standard ray-tracing algorithm to generate the system matrix A_S, which does not account for unwanted photon transmission through the collimator (Huesman, Gullberg, Greenberg & Budinger 1977, Siddon 1985, Gullberg, Zeng, Tsui & Hagius 1989). Because of the imperfect collimator performance of SPECT systems in reality, A_S is too idealistic and we assume that the point spread function (PSF) of the system response is a Gaussian function rather than an ideal Dirac delta function. We blur each PSF in A_S with a 5 by 5 Gaussian kernel and add the Gaussian smear to A_S as the penetration component, resulting in a new system matrix A_G. Thus the system matrix lower and upper bounds for the parallel-hole collimated SPECT are calculated by

$$\begin{aligned} A_{lb} &= A_S \\ A_{ub} &= A_S + \alpha(A_G - A_S), \end{aligned} \tag{4}$$

where α is a hyperparameter adjusting the magnitude of the Gaussian smear in A_S.

For collimatorless imaging, we follow the same approach as in (Zheng et al. 2020) to calculate the system matrix lower and upper bounds, and add more hyperparameters to make the bounds more flexible (see figure 1). Without any collimation, one viable analytical method to calculate the system matrix $\widetilde{A}$ for collimatorless gamma cameras is based on solid angles. The element $\tilde{a}_{ij}$ in $\widetilde{A}$, which represents the probability that photons from the image voxel j reach the detector pixel i, is calculated by

$$\tilde{a}_{ij} = \frac{p^2 r_{ij}}{4\pi R_{ij}^3 + 2p^2 r_{ij}},$$

(5)

where p is the detector pixel size, R_{ij} and r_{ij} are the distance and the normal distance between the image voxel j and the detector pixel i. The $2p^2r_{ij}$ in the denominator of equation (5) is an approximation introduced to address scenarios where the source is in close proximity to the facing detector pixel (i.e., $R_{ij} \simeq r_{ij} \simeq 0$), in which cases the solid angle equation should yield $\frac{1}{2}$ (Frame 2022). This $\widetilde{A}$ derived from solid angles is a weak approximation to the actual system matrix for collimatorless imaging, as it only takes into account the solid angle. A more accurate computation of $\widetilde{A}$ for collimatorless imaging is based on Monte Carlo simulations (Rafecas, Boning, Pichler, Lorenz, Schwaiger & Ziegler 2004, Rafecas, Mosler, Dietz, Pogl, Stamatakis, McElroy & Ziegler 2004, Yao, Ma & Shao 2009). In our simulation studies focusing on 3D collimatorless tomography, we employ this method by positioning a point source that emits 218 keV gamma rays at different locations within the imaging field of view (FOV) of the system and recording the resulting PSFs as the components of $\widetilde{A}$. However, this Monte Carlo-based system matrix calculation may still introduce uncertainties due to factors such as phantom attenuation and scattering.

The detailed procedures to obtain the system matrix lower and upper bounds based on $\widetilde{A}$ are shown in Algorithm 3, where ϵ, η, θ, and ζ are hyperparameters used to tune A_{lb} and A_{ub}. The matrices $\widehat{A}$, $A_{\max}$, A', $\overline{A}$, and W in Algorithm 3 are illustrated in figure 1 and follow the same definitions in (Zheng et al. 2020).

Algorithm 3 System uncertainty $(\epsilon, \eta, \theta, \zeta)$

Input: $\epsilon, \theta, \zeta \in [0,1]$, $\eta \in \left[-\min_j \left(\frac{(1+\nu_j)\min_i \widetilde{a}_{ij}}{\max_i \widetilde{a}_{ij} - (1+\nu_j)\min_i \widetilde{a}_{ij}}\right), 1\right]$ such that $A' \succeq 0$

$\hat{A} \in \mathbb{R}^{m\times n}$, $\hat{a}_{ij} = \begin{cases} \widetilde{a}_{ij}, & \text{if } \widetilde{a}_{ij} \geq (1-\epsilon)\max_i \widetilde{a}_{ij} \\ 0 & \text{otherwise} \end{cases}$

$A_{\max} \in \mathbb{R}^{m\times n}$, $(a_{\max})_{ij} = \max_i \widetilde{a}_{ij}$

$\nu \in \mathbb{R}^n$, $\nu_j = \frac{(1-\epsilon)\max_i \widetilde{a}_{ij}}{\max_i(\widetilde{a}_{ij} - \hat{a}_{ij})} - 1$ ▷ such that $\widetilde{A} + \nu(\widetilde{A} - \hat{A}) \preceq (1-\epsilon)A_{\max}$

$A' = \eta A_{\max} + (1-\eta)(\widetilde{A} + \nu(\widetilde{A} - \hat{A}))$

$\bar{A} = \theta\hat{A} + (1-\theta)A'$

$W = A' - \bar{A}$

$A_{lb} = \bar{A} - \zeta W$

$A_{ub} = \bar{A} + \zeta W$

return A_{lb}, A_{ub}

The main idea of Algorithm 3 is to change the characteristics of the PSF by assuming that the largest response in $\widetilde{A}$ is accurate, while considering that the uncertainties of the remaining $\tilde{a}_{ij}$s are linearly proportional to their respective magnitudes. The hyperparameter ϵ determines the proportion of the largest response in $\widetilde{A}$ that is assumed to be accurate and retained. The hyperparameter η determines the location of A' and the hyperparameter θ determines the location of $\overline{A}$ in figure 1. Both η and θ contribute to shaping the characteristics of the PSF. Additionally, the hyperparameter ζ controls the size of the uncertainty set. These hyperparameters jointly define a system uncertainty set and can be individually adjusted. The value of ϵ is determinate when we have knowledge of the ideal system response $\widehat{A}$ using a perfect parallel-hole collimator. If $\widetilde{A}$ is already close to the actual system matrix, a smaller θ can be set, and vice versa. Similarly, if the uncertainty set is

expected to be small, a smaller ζ can be selected accordingly. Among these hyperparameters, η is the most sensitive and typically needs to be set close to zero. Finding the optimal value for η often involves trying out various smaller values through experimentation. We have tested in our previous study that using these hyperparameters to adjust the system uncertainties works for collimatorless image reconstruction (Zheng et al. 2020).

For reconstruction in collimatorless imaging, the system matrix for the standard MLEM algorithm is $\tilde{A}$, and the system matrix uncertainty A_u for the Masked-MLEM algorithm is obtained by manually adjusting ϵ, η, θ, and ζ in Algorithm 3. Both $\tilde{A}$ and A_u set zero responses for the elements corresponding to the dead detector pixels or strips in collimatorless imaging.

2.3. Phantom Studies

We obtained experimental data from three imaging instruments involving both collimated and uncollimated imaging modalities to validate the Masked-MLEM algorithm: 1) Parallel-hole collimated SPECT; 2) 2D collimatorless scintigraphy; and 3) 3D collimatorless tomography. Additionally, we conducted comprehensive Monte Carlo simulations using Geant4 to generate realistic projection data for 3D collimatorless tomography to further validate the effectiveness of the Masked-MLEM algorithm in handling different levels of system uncertainties. Due to limited access to these imaging instruments, we used four unique phantoms filled with both different radionuclides and activities in different imaging modalities. The focus of this study is to compare the standard MLEM and the Masked-MLEM approaches, not the imaging instruments themselves.

2.3.1. SPECT—For SPECT imaging, a NEMA SPECT triple line source phantom (*NEMA SPECT Triple Line Source Phantom* ™ n.d.), filled with about 5 mCi/mL ^{99m}Tc pertechnetate by following a standard filling procedure, was scanned by the GE Infinia Hawkeye 4 SPECT/CT for 30 s per projection (*GoldSeal Infinia Hawkeye 4* n.d.). The line sources are positioned in a right triangle configuration with one at the center and the other two at a distance of 75 mm from the center. Each line source has a diameter of 1 mm and a useful height of 184 mm. The SPECT scanner uses a NaI detector consisting of 128 by 128 detector pixels and a parallel-hole collimator (LE-HR). To obtain the projection data, 10% energy window was used for the 140 keV photopeak of ^{99m}Tc and no scatter correction was applied. We used the projection data from 60 projection angles separated by 6° for image reconstruction. Due to the large dimensionality of the system response, we used 6 subsets for the standard OSEM and the Masked-OSEM reconstructions. The system matrix for the standard OSEM algorithm is A_S, and the system matrix uncertainty A_u for the Masked-OSEM algorithm is obtained by adjusting α in equation (4). We ran both the OSEM algorithm and Masked-OSEM algorithm for 50 iterations until convergence (the reconstruction loss was less than 0.004) with $N_{out} = 50$ and $N_{in} = 1$ for the Masked-OSEM reconstruction.

2.3.2. Collimatorless Scintigraphy—For 2D collimatorless scintigraphy, a fillable mouse phantom with its brain, liver, and tumor filled with 261 nCi, 261 nCi, and 239 nCi of ^{225}Ac aqueous solution, respectively, was put on the surface of a CZT detector and measured

for 20 min (*Fillable Mouse/rat phantom* 2022). The experiment setup is shown in figure 2. The dimensions of the mouse phantom are 97.86 mm in length, 27.48 mm in width, and 22.70 mm in height. The detector is composed of 5 stacked CZT crystals, with each crystal consisting of 16 by 16 pixels. This results in a total of 80 by 16 detector pixels. Each pixel has a size of $1.6\ \text{mm} \times 1.6\ \text{mm} \times 5.0\ \text{mm}$. The distance from the bottom of the mouse phantom to the surface of the detector was about 2 mm. The mouse phantom was scanned on only one projection and we used a 20% energy window for the characteristic X-ray peak centering at 90 keV and a 6% energy window for the 218 keV photopeak of ^{221}Fr, one daughter of ^{225}Ac, to obtain the projection data for both MLEM reconstruction and Masked-MLEM reconstruction. We ran both the MLEM algorithm and Masked-MLEM algorithm for 200 iterations until convergence (the reconstruction loss was less than 0.002) with $N_{out} = 200$ and $N_{in} = 1$ for the Masked-MLEM reconstruction.

2.3.3. Experimental Collimatorless Tomography—For experimental 3D collimatorless tomography, a fillable micro hollow sphere phantom filled with $0.53\ \text{kBq}/\mu\text{L}$ of ^{225}Ac aqueous solution in its three different-sized spheres, was rotated at 8 angles separated by 45° and scanned by a 3D position-sensitive HPGe double-sided strip detector (DSSD) for 30 min per projection (see figure 3) (Frame, Barnowski, Gunter, Mihailescu & Vetter 2022, Frame, Bobba, Gunter, Mihailescu, Bidkar, Flavell & Vetter 2023). The detector is part of a Compton camera and consists of 37 by 37 orthogonal stripes with an active pixel size of $2\ \text{mm} \times 2\ \text{mm} \times 15\ \text{mm}$ (Frame et al. 2022, Frame et al. 2023). The cylinder phantom has a diameter of 40 mm and a height of 82 mm, and was filled with water. The micro hollow spheres, from smallest to largest, have diameters of 4.3 mm, 6.2 mm, and 7.8 mm; therefore, the total activities of ^{225}Ac were approximately 0.6 mCi, 1.8 mCi, and 3.6 mCi, respectively. The distance from the center of the micro hollow sphere phantom to the surface of the detector was about 52.5 mm. The detailed experiment setup for measuring the micro hollow sphere phantom is shown in (Frame et al. 2023). Given that the 3D position-sensitive detector could record the depth of interactions, we summed the counts of the single interaction events as the projection data, whose interaction depths were within 3 mm from the surface of the first detector of Compton cameras, and energies were in the 18% energy window of the characteristic X-ray peak centering at 87 keV, the 1.1% energy window of the 218 keV photopeak of ^{221}Fr, and the 0.6% energy window of the 440 keV photopeak of ^{213}Bi. With the constraint on the interaction depth, the system matrix $\widetilde{A}$ is applicable for projections coming from a mixture of energies. We ran both the MLEM algorithm and Masked-MLEM algorithm for 500 iterations until convergence (the reconstruction loss was less than 0.002) with $N_{out} = 500$ and $N_{in} = 1$ for the Masked-MLEM reconstruction.

2.3.4. Simulated Collimatorless Tomography—For our simulated 3D collimatorless tomography study, we designed an ideal ring detector system tailored for achieving close proximity for small-animal-scale collimatorless imaging (see figure 4). This CZT-based ring detector features 64 pixels per ring and 16 longitudinal rings. The detector pixel size is $2\ \text{mm} \times 2\ \text{mm} \times 5\ \text{mm}$, aligning with state-of-the-art detection systems. The imaging FOV of the ring detector system spans a diameter of 40.7 mm. At the center of the ring detector system, we positioned a cylindrical water phantom with a diameter of 25 mm and

a thickness of 10 mm, i.e., the distance from the center of water phantom to the surface of the detector was about 20.4 mm. Within the phantom, two spheres each with a diameter of 7 mm are situated in the central transverse plane and symmetrically positioned 7 mm away from the phantom's center. The source particles, ^{225}Ac ions with their full decay chain, were uniformly sampled within the aqueous solution of the spheres. We simulated 10^7 decayed ^{225}Ac particles, corresponding to a total activity of 0.45 μCi or a concentration of 1.25 μCi/ml of ^{225}Ac within the spheres when measuring for 10 min.

In the Geant4 Monte Carlo simulations, we used the Livermore physics list (G4EmLivermorePhysics) along with the decay physics lists (G4DecayPhysics and G4RadioactiveDecayPhysics) (Ivanchenko, Apostolakis, Bagulya, Abdelouahed, Black, Bogdanov, Burkhard, Chauvie, Cirrone, Cuttone et al. 2011, Beaudoux, Blin, Barbrel, Kantor & Zacharatou 2019). To account for experimental noise, the detected energy spectrum was blurred by Gaussian noise with a standard deviation of 3 keV. We used an 8.3% energy window for the 218 keV photopeak of ^{221}Fr to obtain the projection data for both MLEM and Masked-MLEM reconstructions. Additionally, the projection data were smoothed by a 3×3 Gaussian filter for denoising. To perform image reconstruction, we employed both the solid angle model and the Monte Carlo method to calculate the system matrix $\widetilde{A}$. The solid angle model reflects a higher level of system uncertainty, whereas the Monte Carlo method represents a lower level of system uncertainty. This enables us to examine the effectiveness of the Masked-MLEM algorithm in addressing varying levels of system uncertainties. For the Monte Carlo method, 10^9 218 keV gamma rays were simulated as a point source at various positions within the imaging FOV, and the resulting projection data within the same energy window (the 8.3% energy window centering at 218 keV) were subsequently smoothed by the 3×3 Gaussian filter and then recorded as the PSFs for generating $\widetilde{A}$. In both cases, the system uncertainty A_u for the Masked-MLEM algorithm was determined based on the corresponding $\widetilde{A}$. We ran both the MLEM algorithm and Masked-MLEM algorithm for 1000 iterations until convergence (the reconstruction loss was less than 0.0005) with $N_{out} = 1000$ and $N_{in} = 1$ for the Masked-MLEM reconstruction.

For all three experiment studies and the simulation study, we perform 3D image reconstruction and subsequently apply a $3 \times 3 \times 3$ median filter to the 3D reconstructed images for denoising. In some cases, we collapse the 3D reconstructed images to the corresponding 2D images for comparison between the standard MLEM algorithm and the Masked-MLEM algorithm. For example, in collimatorless scintigraphy, we sum the 3D reconstructed image along the normal direction of the detector surface to obtain a 2D coronal image of the mouse phantom, since the depth reconstruction of the mouse phantom using only one projection is inaccurate. In this study, we do not consider any additional image regularizations and solely compare the reconstructed images obtained by the standard MLEM algorithm and the Masked-MLEM algorithm. To evaluate the quality of reconstructions, we normalize both the reconstructed images and the ground truth images by their respective maximum values, and utilize three figures of merit (FOMs) that are not influenced by the normalization factors: normalized root mean squared error (NRMSE), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM). NRMSE and PSNR

quantitatively measure the accuracy between the reconstructed images and the ground truth images, while SSIM is used to assess the similarity between them.

3. Results

In SPECT imaging, the reconstruction results of the triple line source phantom overlaid on a transverse slice of the CT image are shown in figure 5. We use $\alpha = 2$ in equation (4) to obtain A_u for the Masked-OSEM reconstruction. For an equitable visual assessment between the standard OSEM algorithm and the Masked-OSEM algorithm, we sum the 3D reconstructed images along the longitudinal direction to obtain a 2D transverse image of the triple line source phantom, as shown in the top row of figure 5. This approach allows us to visually compare the averaged transverse image of the triple line source phantom, mitigates noise in each transverse slice of the 3D reconstructed images, and avoids introducing any bias of selectively comparing specific reconstructed slices where the Masked-OSEM algorithm might outperform the standard OSEM algorithm. Notably, the FOMs are compared based on the original 3D reconstructed images. The reconstruction profiles along the x- and y-axes in the collapsed transverse image of the triple line source phantom are shown in the bottom row of figure 5.

For collimatorless imaging, the hyperparameters specified in Algorithm 3 are detailed in table 1 to obtain the system uncertainty A_u for the Masked-MLEM reconstruction. These hyperparameters have been carefully tuned through comprehensive experimentation to ensure their suitability for the respective collimatorless imaging modality. In the simulated collimatorless tomography study, we have verified that the hyperparameter values in table 1 are effective for reconstructing various phantoms by performing reconstructions on the simulated sphere phantom containing spheres of different sizes and positions.

In collimatorless scintigraphy, the projection data coming from the characteristic X-rays and the 218 keV photopeak of ^{221}Fr are shown in figure 6 (a). We reconstruct the 3D images of the mouse phantom using the standard MLEM algorithm and the Masked-MLEM algorithm based on the single-angle projection data from X-rays and ^{221}Fr, respectively. The 2D collapsed coronal images overlaid on a coronal slice of the mouse phantom's CT image are shown in figure 6 (c) and (d).

In experimental collimatorless tomography, the projection data of the micro hollow sphere phantom filled with ^{225}Ac at 8 angles are shown in figure 7. The representative transverse, coronal and sagittal slices from the center of the reconstructed spheres are shown in figure 8.

In simulated collimatorless tomography, the reconstruction results of the 2-sphere phantom are shown in figure 9. When using the solid angle model to calculate $\tilde{A}$, along with the corresponding hyperparameter values from table 1 to determine the system uncertainty A_u, the representative transverse and coronal slices from the center of the reconstructed spheres using the standard MLEM algorithm and the Masked-MLEM algorithm are shown in figure 9 (b) and (c), respectively. Similarly, when using the Monte Carlo method to calculate $\tilde{A}$, the reconstructions of the spheres using the standard MLEM algorithm and the Masked-MLEM algorithm are shown in figure 9 (d) and (e).

Table 2 shows the FOMs for the 3D reconstructed images of the triple line source phantom, the micro hollow sphere phantom, and the simulated 2-sphere phantom, as well as the FOMs for the 2D collapsed coronal images of the mouse phantom. These FOMs are computed for both the standard MLEM (OSEM) and the Masked-MLEM (Masked-OSEM) reconstructions.

4. Discussion

In this study, the Masked-MLEM algorithm is developed and validated on both collimated and uncollimated imaging instruments and compared with the standard MLEM algorithm. Here we are comparing the reconstruction results across different algorithms, rather than across different imaging instruments. The limited access to different phantoms, radionuclides, and total activities for evaluating the reconstruction performance on one imaging instrument is a limitation of the study.

In SPECT imaging, there is almost no difference between the reconstruction of the triple line source phantom using the standard OSEM algorithm and the Masked-OSEM algorithm (see figure 5 and table 2). Similar results were obtained in previous studies which selected the worst-case system matrix in its uncertainty set for robust reconstruction (Liu et al. 2012). The reasons may have the following two aspects. Firstly, the system matrix A_S used for the simple-structured triple line source phantom is potentially already optimal or suboptimal within the given uncertainty set. Secondly, our definition of the system matrix upper bound, which adds an adjustable Gaussian smear to A_S as the penetration component, may not accurately capture the true system uncertainties and there may be other more effective designs of the system uncertainties which require further exploration.

In contrast to SPECT imaging, the Masked-MLEM algorithm outperforms the standard MLEM algorithm in collimatorless imaging as shown in figures 6, 8, and 9. Image resolution is a limiting factor in collimatorless imaging because we do not have any collimation and only rely on solid angles, which sacrifices the image resolution for higher sensitivity. Therefore, the three key factors influencing collimatorless imaging are:

1. Proximity, i.e., how close the phantom is to the detector, which determines how ill-conditioned the system matrix is, thereby affecting the resolution of the reconstructed image. When the phantom is placed in close proximity to the detector, the system matrix becomes less ill-conditioned, resulting in improved image quality and reconstruction accuracy.
2. Projection Statistics, which reflects the noise level in projections and has a direct impact on the signal-to-noise ratio (SNR) of reconstructed images. Whereas collimatorless imaging captures photons coming from various angles, ensuring sufficient projection counts remains crucial for optimizing SNR.
3. Uniformity of Source Distribution, i.e., whether the source is uniformly distributed throughout the phantom or concentrated in specific regions within the phantom, which impacts the reconstruction accuracy. Notably, the uniform source distribution in this context does not imply a uniform concentration of radioactivity in source regions, but rather a consistent total activity across

all source regions. The impact on reconstruction accuracy arises from how the source distribution influences the consistency of acquired projection data in collimatorless imaging. When the source is uniformly distributed with comparable total activities in each source region, the projections measured from various angles represent the activity distribution in a balanced manner and align with the underlying source distribution, leading to more accurate reconstruction. In contrast, non-uniform source distribution with dominance in particular regions makes the projections more sensitive to those dominant regions. This imbalance can result in distorted or biased reconstructions where the features corresponding to the dominant regions are emphasized whereas other source regions of the phantom may be suppressed or misrepresented.

The collimatorless scintigraphy experiment is an example showing good proximity, good projection statistics, and a uniform source distribution. With the mouse phantom placed directly on the detector surface ensuring close proximity, and the source organs well-separated while containing similar total activities of ^{225}Ac (indicating a uniform source distribution), the projection data in figure 6 (a) distinctly highlight the source organs, which facilitates easier collimatorless image reconstruction and improves the reconstruction accuracy. In collimatorless scintigraphy, the Masked-MLEM algorithm achieves more accurate reconstruction of the brain, liver, and tumor of the mouse phantom with fewer artifacts and better FOMs compared to the standard MLEM algorithm (see figure 6 and table 2). Conversely, the standard MLEM reconstructions allocate higher activity values to adjacent organs that should not exhibit radioactivity, as evident in organs like the kidneys of the mouse phantom in figure 6. This primarily stems from the following two reasons. The foremost reason is that the system matrix $\tilde{A}$ used in the standard MLEM algorithm is solely derived from the solid angle model, which inherently introduces uncertainties into the reconstruction process. On the other hand, the second reason, though present, is not the primary one but does make both algorithms perform worse. Given that we used the single-angle projection for 3D image reconstruction followed by collapsing the 3D image to obtain the 2D coronal image, this approach results in a highly underdetermined optimization problem, which leads to reconstruction ambiguities and inaccuracies because multiple plausible solutions can fit the limited measurements. Whether using the MLEM algorithm or the Masked-MLEM algorithm, the images reconstructed from the ^{221}Fr projection are noisier than those from the projection of X-rays in figure 6, since the projection data of ^{221}Fr have fewer counts and thus worse statistics than those of X-rays in figure 6 (a). Therefore, it is important to obtain sufficient counts in collimatorless projections to improve the quality of image reconstruction.

Optimal proximity, good projection statistics, and a uniform source distribution are ideal conditions for collimatorless reconstruction; however, in practice, such conditions are not always met. The collimatorless tomography experiment is an illustrative example of this. In this experimental study, although the Masked-MLEM algorithm exhibits superior performance compared to the standard MLEM algorithm in figure 8, both algorithms struggle to achieve high-resolution reconstructions in figure 8 and attain satisfactory FOMs in table 2. This can be explained by the poor proximity of the micro hollow sphere phantom, and the non-uniform activity distribution among the spheres. In regard to the latter, the total

activities in the two largest spheres were much higher than those in the smallest one. Due to this non-uniform distribution and the poor image resolution of collimatorless tomography, only the largest two spheres appear in the reconstruction (see figure 8). Furthermore, due to the dead stripes in the detector (see figure 7), certain valuable projection data are missing, degrading the image quality. Though reduced reconstruction quality, we experimentally demonstrate the feasibility of performing proximity reconstruction with a Compton camera system, as discussed in (Caravaca, Huh, Gullberg & Seo 2022).

The simulated collimatorless tomography study demonstrates an idealized extreme scenario for small-animal-scale collimatorless imaging, featuring a perfect ring detector geometry, close proximity from all measurement angles, and a uniform source distribution within the phantom. This simulation study highlights the effectiveness of the Masked-MLEM algorithm in handling varying levels of system uncertainties. In figure 9 (b), the MLEM reconstruction using the system matrix $\widetilde{A}$ derived from the solid angle model fails to reconstruct the two spheres due to the higher level of system uncertainty. Conversely, the Masked-MLEM reconstruction in figure 9 (c), which utilizes the system uncertainties based on the same $\widetilde{A}$, successfully reconstructs the positions of both spheres, despite with some artifacts present between them. Compared to figure 9 (b), the MLEM reconstruction in figure 9 (d), using the Monte Carlo-calculated $\widetilde{A}$ with a lower level of system uncertainty, captures the structural details of the two spheres more effectively, although their positions slightly shift towards the boundaries of the imaging region of interest (ROI). Consequently, the Masked-MLEM reconstruction based on this refined $\widetilde{A}$ is similar to the MLEM reconstruction but more accurately positions the spheres (see figure 9 (e) and table 2). Due to the inherent limitations of image resolution in collimatorless imaging, both algorithms struggle to achieve well-defined boundaries for the spheres and fail to eliminate activities in the regions between them. It is noteworthy that the FOMs for the MLEM reconstruction using the solid angle model are much better than those for the other methods in this simulation study (see table 2). This is because the reconstructed images in figure 9 (b) assign zero activity to background regions, which enhances all three FOMs.

In both experimental and simulated collimatorless tomography studies, we notice a consistent shift in position towards the boundaries of the imaging ROI (see figures 8 and 9). This suggests that incorporating some form of collimation would be beneficial to enhance the accuracy of position reconstruction. How to further improve the image reconstruction quality in collimatorless imaging will be investigated in our future studies.

In the three experimental studies and one simulation study, the Masked-MLEM algorithm consistently produces robust image reconstructions by effectively minimizing the likelihood of erroneously reconstructing higher activities in regions without radioactive sources (see figures 5, 6, 8, and 9). Essentially, the Masked-MLEM algorithm can be understood as using the optimal system matrix A^* within the predefined system uncertainty set (i.e., $A_{lb} \preceq A^* \preceq A_{ub}$) and then performing the standard MLEM reconstruction. When A^* is close to or matches the actual system matrix and we have accurate knowledge of the actual system matrix, the resulting Masked-MLEM reconstruction will be nearly identical to the standard MLEM reconstruction using the actual system matrix. This can be seen in the

SPECT reconstruction and the simulated collimatorless tomography study when using the Monte Carlo-calculated $\widetilde{A}$ (see figures 5 and 9). In both examples, although the system matrix used in the standard MLEM algorithm is not exactly the actual matrix, it does not introduce substantial uncertainties, and thus the Masked-MLEM reconstructions are similar to or slightly superior to the standard MLEM reconstructions. Although it is preferable to have an accurate system matrix, in many cases, we are only able to obtain an approximation to the actual system matrix due to complicated factors such as scattering, attenuation, and nonuniform detector responses, as demonstrated in the simulated collimatorless tomography study where inherent system uncertainties persist despite using the Monte Carlo method to generate the system matrix. Additionally, computing an accurate system matrix can often be time-consuming or impractical. The Masked-MLEM algorithm offers the capability to reconstruct images reliably even without accurate knowledge of the actual system matrix. The power of this algorithm lies in leveraging the well-defined A_{lb} and A_{ub} based on the current system matrix $\widetilde{A}$, eliminating the need for time-consuming computations of an optimal system matrix. This allows for efficient image reconstructions that yield results comparable to or even superior to those obtained using the standard MLEM algorithm, regardless of the availability or accuracy of the system matrix.

The success of the Masked-MLEM algorithm depends heavily on the choice of the system uncertainty A_u. Although the Masked-MLEM algorithm will choose the best system matrix within its uncertainty set, it's important to note that expanding the system uncertainty set excessively is not advisable. On the other hand, the optimal choice of A_u should be as restrictive as possible, including all feasible system matrices while excluding all implausible system matrices. We have tested that η is the most sensitive hyperparameter for tuning A_u in collimatorless imaging, and all other hyperparameters can be kept the same for different reconstruction problems. In this study, A_u is largely defined from a mathematical point of view, and it will be better to understand some prior knowledge of the system uncertainties to define A_u according to the physical effects in real applications. The box uncertainty of the system matrix provides the flexibility to element-wisely design the system uncertainties and the power to include any model for system uncertainties, which has the potential to be widely used in other inverse optimization problems involving system uncertainties. Although loading the lower and upper bounds of the system matrix consumes additional storage memory, the computation time per iteration in the Masked-MLEM algorithm is slightly shorter than or comparable to that of the standard MLEM algorithm, which is acceptable.

5. Conclusion

In this study, we develop the Masked-MLEM algorithm as a generalization of the standard MLEM algorithm for robust image reconstruction. The Masked-MLEM algorithm incorporates a box uncertainty set in the system matrix and accounts for Poisson noise in projection data. We validate the Masked-MLEM algorithm using data from both collimated and uncollimated imaging instruments by carefully designing the corresponding system matrix uncertainties. In cases where the system matrix does not introduce substantial uncertainties, the Masked-MLEM and standard MLEM reconstructions are similar, as observed in SPECT imaging and the simulated collimatorless tomography using the Monte

Carlo-calculated system matrix. However, the Masked-MLEM algorithm outperforms the standard MLEM algorithm in situations where the system matrix is an approximation, as demonstrated in other collimatorless imaging studies. In both collimated and uncollimated imaging, the Masked-MLEM algorithm enhances the robustness of image reconstruction, yielding more reliable results in the presence of varying levels of system uncertainties, and effectively reducing the likelihood of erroneously reconstructing higher activities in regions without radioactive sources. Future studies will focus on refining system uncertainties for SPECT imaging and designing novel imaging modalities for further improving the reconstructed images in collimatorless imaging.

Acknowledgments

This work was supported by National Institute of Biomedical Imaging and Bioengineering under grant R01EB026331. All authors have no known conflicts of interest in terms of competing for financial interests or personal relationships that could have an influence or are relevant to the work reported in this paper. We would like to acknowledge many helpful discussions with Donald L. Gunter at LBNL regarding the derivation of the solid angle equation (see (5)).

Appendix

Appendix A. Masked-MLEM derivation

First we need to demonstrate that

$$M|y_u| \sim \text{Poisson}(M|A_u|x). \tag{A.1}$$

Although y appears twice in y_u (see equation (2)), for any element y_i in y, $i = 1,\ldots, m$, it can not appear twice in $M|y_u|$. Using proof by contradiction, we assume that y_i appears twice in $M|y_u|$, i.e.,

$$\begin{aligned}(a_{lb})_i^T x - y_i > 0\\ y_i - (a_{ub})_i^T x > 0.\end{aligned} \tag{A.2}$$

Since $x \geq 0$, and element-wisely $(a_{ub})_i \geq (a_{lb})_i$, and thus we have

$$(a_{ub})_i^T x \geq (a_{lb})_i^T x > y_i, \tag{A.3}$$

which leads to a contradiction with equation (A.2). This proves that $M|y_u|$ remains independent and identically distributed (i.i.d.) after applying the mask M. Thus $M|y_u|$ contains a subset of y.

Assuming Poisson noise in each y_i, we have

$$p(M|y_u \parallel x) \sim \prod_{(M|y_u|)_i \neq 0} \frac{(M|A_u|x)_i^{(M|y_u|)_i}}{(M|y_u|)_i!} e^{-(M|A_u|x)_i}. \tag{A.4}$$

Similar to the derivation of the MLEM algorithm, we minimize the negative log-likelihood function $-\log p(M|y_u \parallel x)$ as shown in equation (A.5).

$$\min_{x \geq 0} \sum_{(M|y_u|)_i \neq 0} (M|A_u|x)_i - (M|y_u|)_i \log(M|A_u|x)_i \tag{A.5}$$

The Lagrangian function of equation (A.5) $\left(\mathscr{L} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\right)$ is given by

$$\mathscr{L}(x, \lambda) = \sum_{(M|y_u|)_i \neq 0} (M|A_u|x)_i - (M|y_u|)_i \log(M|A_u|x)_i - \lambda^T x, \tag{A.6}$$

where $\lambda \geq 0$ is a Lagrange multiplier corresponding to the constraint $x \geq 0$.

Following the Karush-Kuhn-Tucker (KKT) conditions (Karush 1939, Kuhn & Tucker 2014, Boyd & Vandenberghe 2004, Yan et al. 2013), we have

- Primal feasibility: $x \geq 0$;
- Dual feasibility: $\lambda \geq 0$;
- Complementary slackness: $x^T \lambda = 0$;
- Stationarity: $\nabla_x \mathscr{L}(x, \lambda) = (M|A_u|)^T 1 - (M|A_u|)^T \frac{M|y_u|}{M|A_u|x} - \lambda = 0$.

The first three conditions imply $x_j \lambda_j = 0$, $j = 1, \ldots, n$. Using the stationarity condition to eliminate λ, we obtain the resulting iterative relationship given by

$$x^{k+1} = \frac{x^k}{(M|A_u|)^T 1} (M|A_u|)^T \frac{M|y_u|}{M|A_u|x^k}, \tag{A.7}$$

where k is the iteration number.

References

Agostinelli S, Allison J, Amako K. a., Apostolakis J, Araujo H, Arce P, Asai M, Axen D, Banerjee S, Barrand G et al. (2003). Geant4—a simulation toolkit, Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506(3): 250–303.

Beaudoux V, Blin G, Barbrel B, Kantor G & Zacharatou C (2019). Geant4 physics list comparison for the simulation of phase-contrast mammography (xpulse project), Physica Medica 60: 66–75. [PubMed: 31000088]

Beister M, Kolditz D & Kalender WA (2012). Iterative reconstruction methods in x-ray ct, Physica medica 28(2): 94–108. [PubMed: 22316498]

Boyd SP & Vandenberghe L (2004). Convex optimization, Cambridge university press.

Caravaca J, Huh Y, Gullberg GT & Seo Y (2022). Compton and proximity imaging of 225ac in vivo with a czt gamma camera: a proof of principle with simulations, IEEE Transactions on Radiation and Plasma Medical Sciences

Chen Z, Qi H, Wu S, Xu Y & Zhou L (2016). Few-view ct reconstruction via a novel non-local means algorithm, Physica Medica 32(10): 1276–1283. [PubMed: 27289353]

Dasgupta S, Papadimitriou CH & Vazirani UV (2008). Algorithms, McGraw-Hill Higher Education New York.

Defrise M & Gullberg GT (2006). Image reconstruction, Physics in Medicine & Biology 51(13): R139. [PubMed: 16790900]

Diamond S & Boyd S (2016). Cvxpy: A python-embedded modeling language for convex optimization, The Journal of Machine Learning Research 17(1): 2909–2913.

Fillable Mouse/rat phantom (2022). URL: https://bioemtech.com/fillable-mouse-rat-phantom/

Frame EA (2022). Three-Dimensional Coded Aperture and Compton Gamma-Ray Imaging for Near-Field Applications, PhD thesis, University of California, Berkeley.

Frame E, Barnowski R, Gunter D, Mihailescu L & Vetter K (2022). A dual-modality volumetric gamma-ray imager for near-field applications, IEEE Transactions on Nuclear Science 69(12): 2343–2359.

Frame E, Bobba K, Gunter D, Mihailescu L, Bidkar A, Flavell R & Vetter K (2023). Coded aperture and compton imaging for the development of 225ac-based radiopharmaceuticals, Medical Physics

GoldSeal Infinia Hawkeye 4 (n.d.). **URL:** https://www.gehealthcare.com/products/goldseal—refurbished-systems/goldseal-nuclear-medicine/goldseal-infinia-hawkeye-4**URL:**

Gong K, Catana C, Qi J & Li Q (2018). Pet image reconstruction using deep image prior, IEEE transactions on medical imaging 38(7): 1655–1665. [PubMed: 30575530]

Gullberg GT, Zeng GL, Tsui BM & Hagius JT (1989). An iterative reconstruction algorithm for single photon emission computed tomography with cone beam geometry, International Journal of Imaging Systems and Technology 1(2): 169–186.

Hsieh J, Nett B, Yu Z, Sauer K, Thibault J-B & Bouman CA (2013). Recent advances in ct image reconstruction, Current Radiology Reports 1(1): 39–51.

Hudson HM & Larkin RS (1994). Accelerated image reconstruction using ordered subsets of projection data, IEEE transactions on medical imaging 13(4): 601–609. [PubMed: 18218538]

Huesman R, Gullberg G, Greenberg W & Budinger T (1977). Users manual: Donner algorithms for reconstruction tomography

Ivanchenko V, Apostolakis J, Bagulya AV, Abdelouahed HB, Black R, Bogdanov A, Burkhard H, Chauvie S, Cirrone GAP, Cuttone G et al. (2011). Recent improvements in geant4 electromagnetic physics models and interfaces, Progress in nuclear science and technology 2: 898–903.

Karush W (1939). Minima of functions of several variables with inequalities as side constraints, M. Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago

Kucharczak F, Loquin K, Buvat I, Strauss O & Mariano-Goulart D (2018). Interval-based reconstruction for uncertainty quantification in pet, Physics in Medicine & Biology 63(3): 035014. [PubMed: 29192605]

Kuhn HW & Tucker AW (2014). Nonlinear programming, Traces and emergence of nonlinear programming, Springer, pp. 247–258.

Lange K, Carson R et al. (1984). Em reconstruction algorithms for emission and transmission tomography, J Comput Assist Tomogr 8(2): 306–16. [PubMed: 6608535]

Liu H, Wang S, Gao F, Tian Y, Chen W, Hu Z & Shi P (2012). Robust framework for pet image reconstruction incorporating system and measurement uncertainties, PloS one 7(3): e32224. [PubMed: 22427826]

Lunz S, Hauptmann A, Tarvainen T, Schonlieb C-B & Arridge S (2021). On learned operator correction in inverse problems, SIAM Journal on Imaging Sciences 14(1): 92–127.

NEMA SPECT Triple Line Source Phantom (n.d.). **URL:** https://capintec.com/product/nema-spect-triple-line-source-phantom/**URL:**

Panin V, Zeng G & Gullberg G (1998). Total variation regulated em algorithm, 1998 IEEE Nuclear Science Symposium Conference Record. 1998 IEEE Nuclear Science Symposium and Medical Imaging Conference (Cat. No. 98CH36255), Vol. 3, IEEE, pp. 1562–1566.

Qi J & Huesman RH (2004). Effect of errors in the system matrix on iterative image reconstruction, IEEE Symposium Conference Record Nuclear Science 2004, Vol. 5, IEEE, pp. 2854–2858.

Rafecas M, Boning G, Pichler BJ, Lorenz E, Schwaiger M & Ziegler SI (2004). Effect of noise in the probability matrix used for statistical reconstruction of pet data, IEEE Transactions on Nuclear Science 51(1): 149–156.

Rafecas M, Mosler B, Dietz M, Pogl M, Stamatakis A, McElroy DP & Ziegler SI (2004). Use of a monte carlo-based probability matrix for 3-d iterative reconstruction of madpet-ii data, IEEE Transactions on Nuclear Science 51(5): 2597–2605.

Rico A & Strauss O (2010). Imprecise expectations for imprecise linear filtering, International Journal of Approximate Reasoning 51(8): 933–947.

Rico A, Strauss O & Mariano-Goulart D (2009). Choquet integrals as projection operators for quantified tomographic reconstruction, Fuzzy Sets and Systems 160(2): 198–211.

Sawatzky A, Brune C, Wubbeling F, Kosters T, Schafers K & Burger M (2008). Accurate em-tv algorithm in pet with low snr, 2008 IEEE nuclear science symposium conference record, IEEE, pp. 5133–5137.

Shepp LA & Vardi Y (1982). Maximum likelihood reconstruction for emission tomography, IEEE transactions on medical imaging 1(2): 113–122. [PubMed: 18238264]

Siddon RL (1985). Fast calculation of the exact radiological path for a three-dimensional ct array, Medical Physics 12(2): 252–255. [PubMed: 4000088]

Tikhonov AN (1963). On the solution of ill-posed problems and the method of regularization, Doklady akademii nauk, Vol. 151, Russian Academy of Sciences, pp. 501–504.

Yan M, Bui AA, Cong J & Vese LA (2013). General convergent expectation maximization (em)-type algorithms for image reconstruction, Inverse Problems & Imaging 7(3): 1007.

Yao R, Ma T & Shao Y (2009). Derivation of system matrix from simulation data for an animal spect with slit-slat collimator, IEEE Transactions on Nuclear Science 56(5): 2651–2658.

Zeng GL (2001). Image reconstruction—a tutorial, Computerized medical imaging and graphics 25(2): 97–103. [PubMed: 11137785]

Zheng Y, Huh Y, Su Q, Wang J, Lin Y, Vetter K & Seo Y (2020). Collimatorless scintigraphy for imaging extremely low activity targeted alpha therapy (tat) with weighted robust least squares (wrls), International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 803–811.

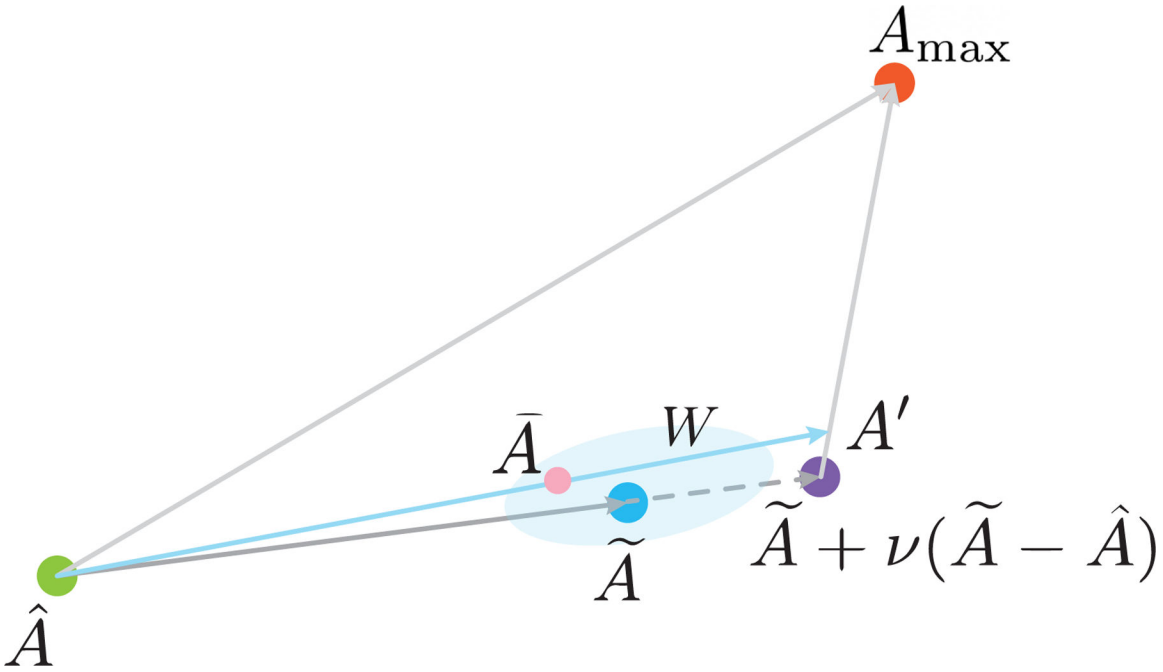

Figure 1.

System uncertainty set for collimatorless imaging (Zheng et al. 2020). The system matrix $\tilde{A}$ for collimatorless imaging may not be accurate and we assume the actual system matrix is in the neighborhood of $\tilde{A}$ (shown as the blue cloud). The system matrix $\hat{A}$ corresponds to the ideal scenario where a perfect parallel-hole collimator is used. The system matrix $A_{\max}$ defines the upper limit that the current system can achieve, i.e., it represents the scenario where every image voxel has an equally maximum probability of reaching every detector pixel. The system matrix A' determines the characteristics of the PSF and the range $\bar{A} \pm W$ is desired to be close to the actual system matrix.

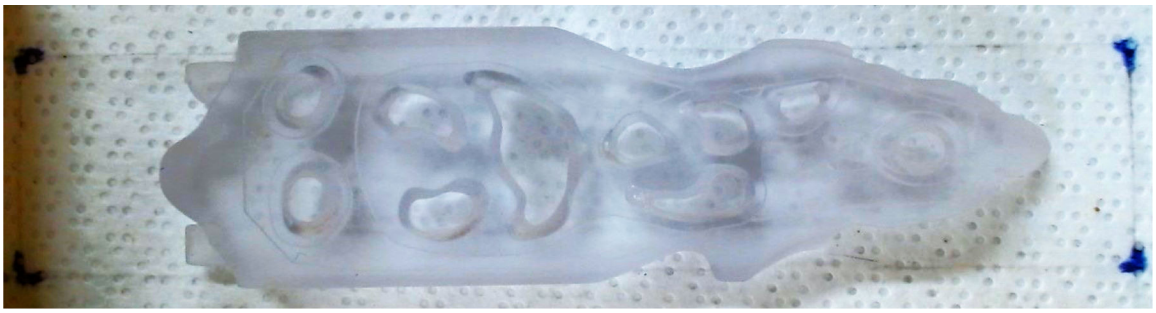

Figure 2.
Experiment setup of the mouse phantom for collimatorless scintigraphy.

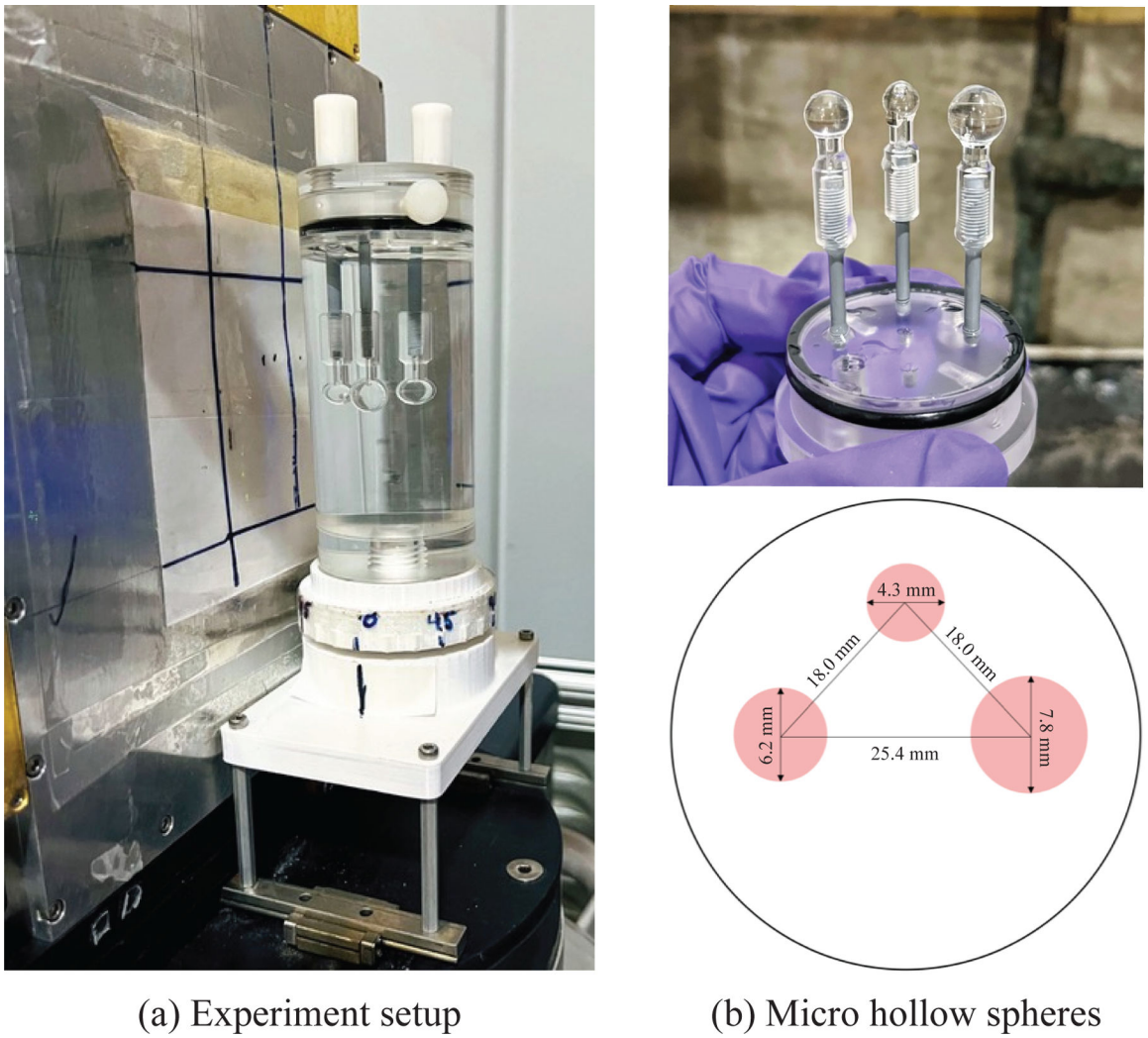

Figure 3.
(a) Experiment setup of the micro hollow sphere phantom for collimatorless tomography. (b) The micro hollow sphere phantom filled with 0.6 mCi, 1.8 mCi, and 3.6 mCi of ^{225}Ac from smallest to largest spheres (Frame et al. 2023).

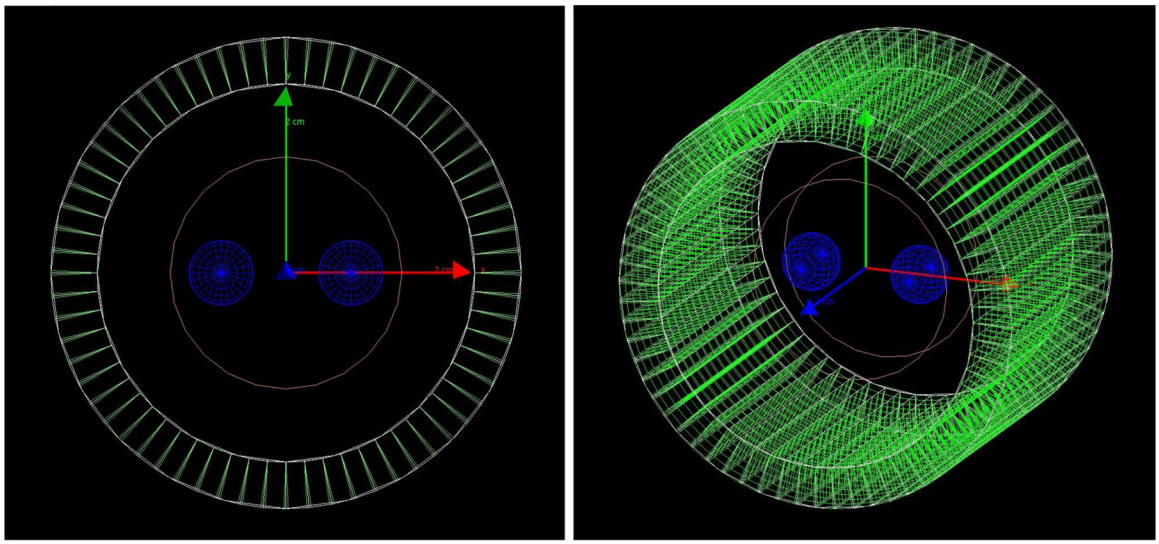

Figure 4.

Monte Carlo simulation setup using Geant4 for collimatorless tomography with the ring detector system: (a) Geant4 rendering from the x-y plane, and (b) 3D view. The CZT detector is shown in green and the cylindrical water phantom is shown in orange. Inside the phantom, two blue spheres contain ^{225}Ac source particles uniformly distributed in the aqueous solution.

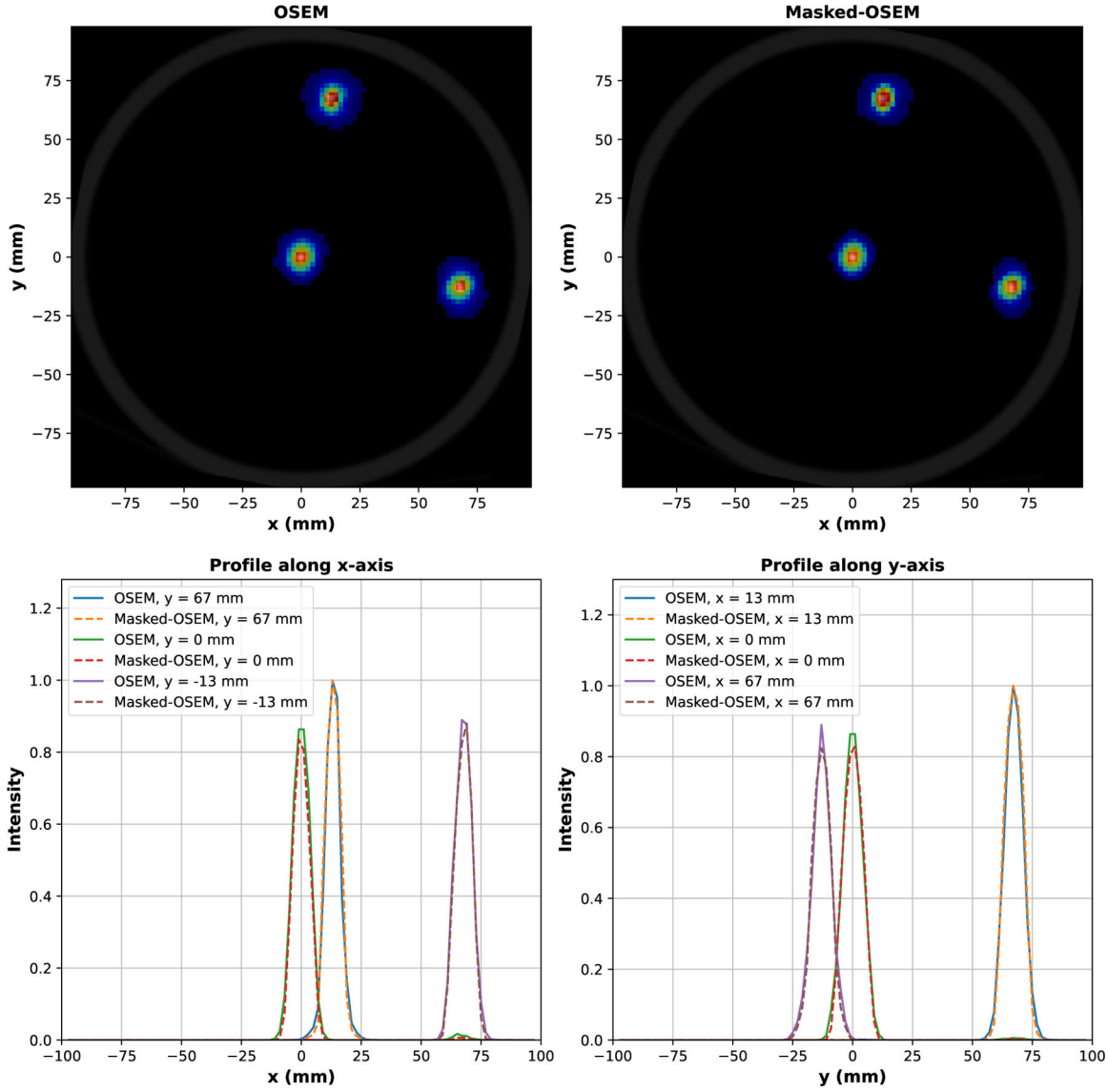

Figure 5.

Reconstruction of the triple line source phantom in SPECT imaging (shown in jet colormap). The top row shows the transverse images of the triple line source phantom obtained by longitudinally summing the 3D images reconstructed using the standard OSEM algorithm and the Masked-OSEM algorithm. These images are overlaid on a transverse slice of the CT image (shown in grayscale), which serves as the ground truth reference for the triple line source phantom. It is noteworthy that the reconstruction resolution of SPECT is inferior to the diameter of the line sources (1 mm), hence the depiction of the line sources in the CT image as the central gray dots. The bottom row shows the reconstruction profiles of the triple line source phantom in the top row.

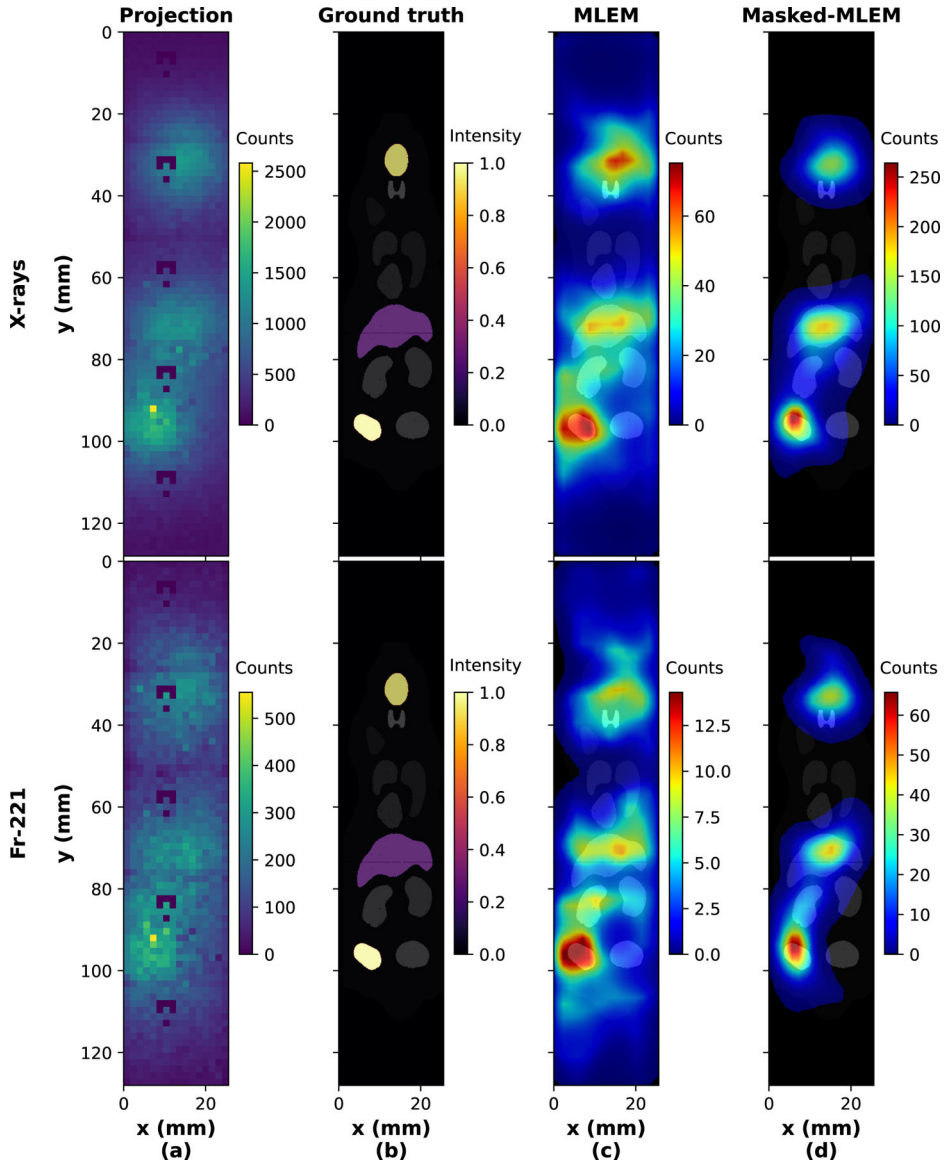

Figure 6.

Reconstruction of the mouse phantom in collimatorless scintigraphy with its brain, liver, and tumor filled with 261 nCi, 261 nCi, and 239 nCi of ^{225}Ac. (a) Projection data coming from the characteristic X-rays and the 218 keV photopeak of ^{221}Fr where the zero response in certain positions was due to the dead detector pixels. It is noteworthy that the dead detector pixels appeared at the same positions across all five CZT crystals, and were not the result of a multi-stage acquisition. (b) Normalized concentrations of the source organs as the ground truth. The last two columns show the 2D coronal images overlaid on the CT image and reconstructed using (c) the standard MLEM algorithm and (d) the Masked-MLEM algorithm, respectively.

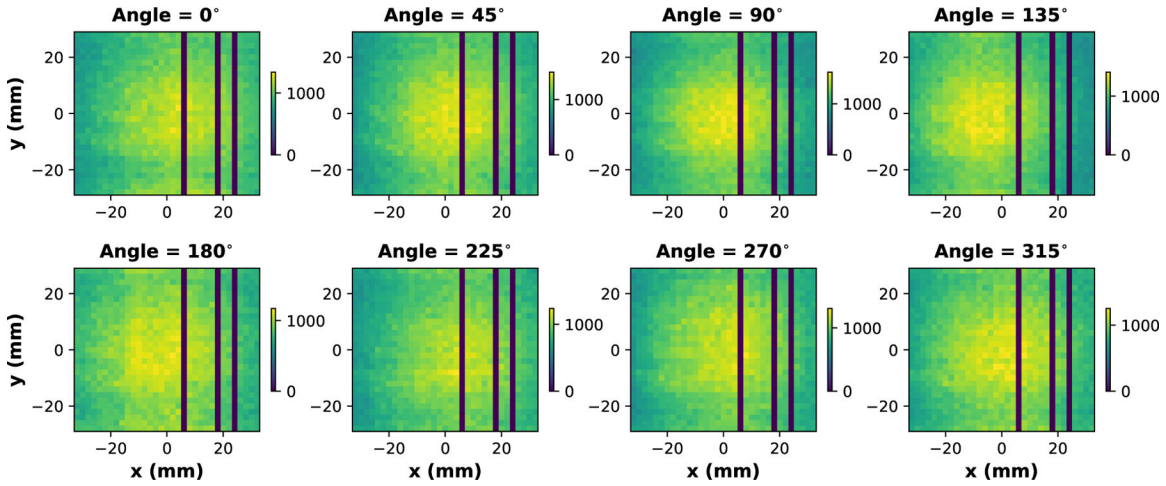

Figure 7.

Projection of the micro hollow sphere phantom filled with ^{225}Ac at 8 angles. The projection counts came from the single interaction events near the detector surface (within 3 mm) with energies emitted from the characteristic X-rays, the 218 keV photopeak of ^{221}Fr, and the 440 keV photopeak of ^{213}Bi. The zero response was caused by the dead detector strips.

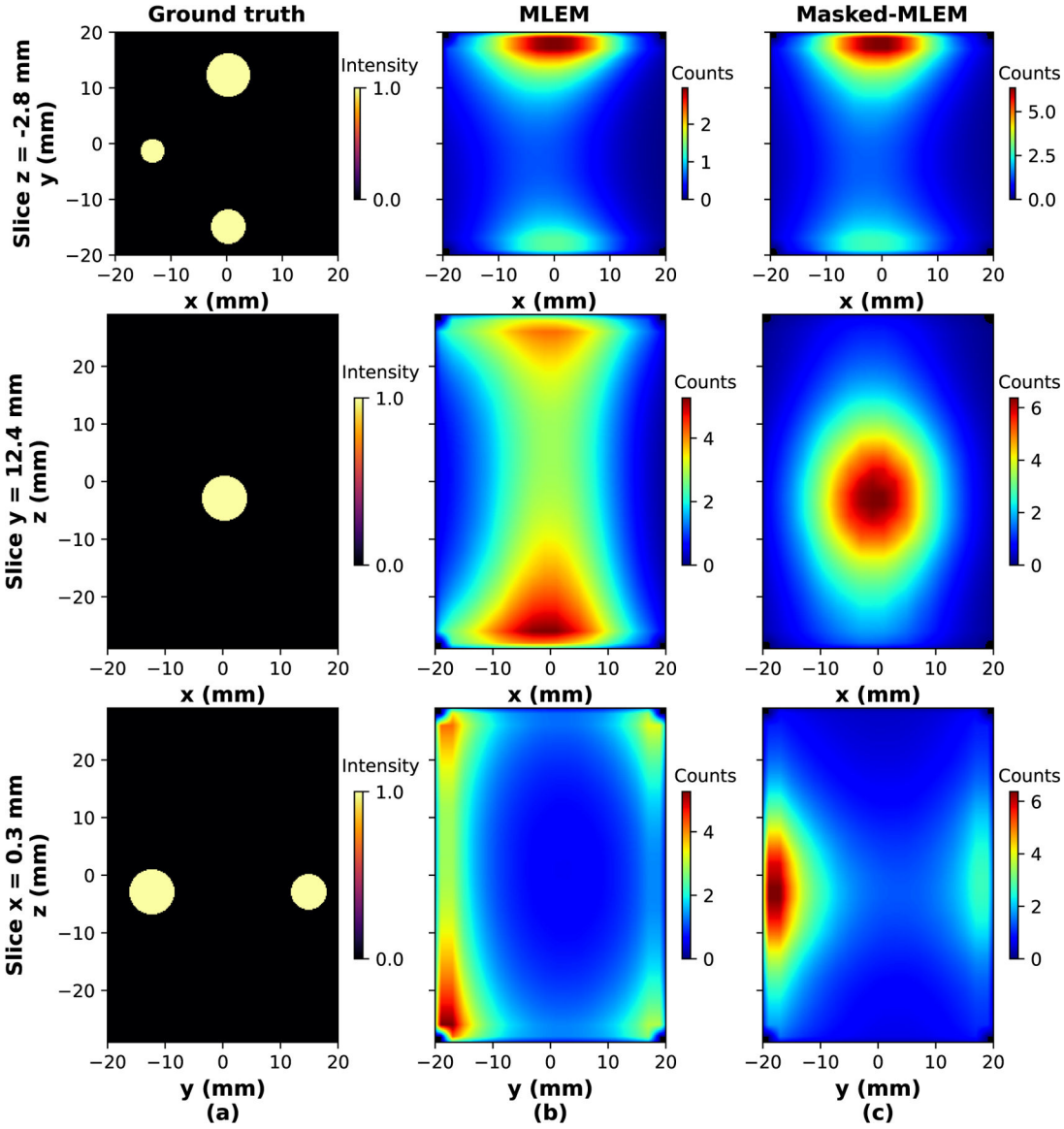

Figure 8.

(a) Ground truth and reconstructions of the micro hollow sphere phantom using (b) the standard MLEM algorithm and (c) the Masked-MLEM algorithm. The phantom was filled with 0.6 mCi, 1.8 mCi, and 3.6 mCi of ^{225}Ac in the smallest to the largest micro hollow spheres. From top to bottom, it shows the transverse, coronal and sagittal slices of the reconstructed spheres.

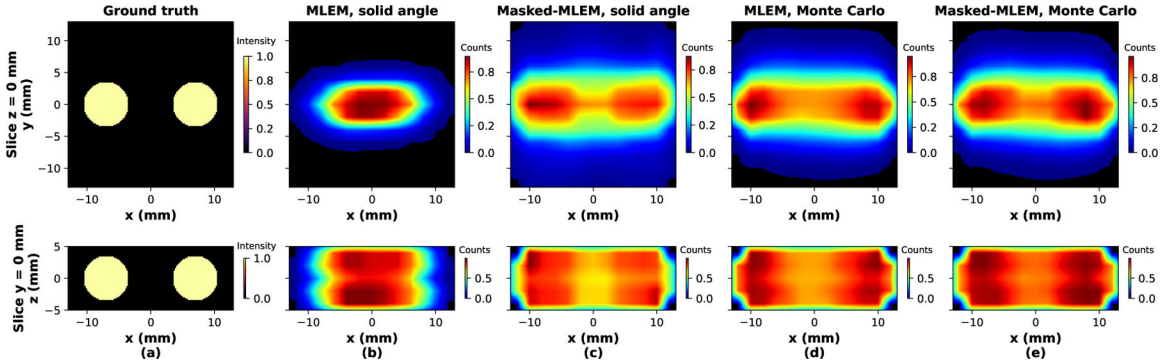

Figure 9.

Reconstruction of the simulated 2-sphere phantom containing 0.45 μCi of ^{225}Ac within the spheres when measuring for 10 min. From left to right, it shows: (a) Ground truth of the phantom; (b) MLEM and (c) Masked-MLEM reconstructions when using the solid angle model to calculate $\widetilde{A}$; (d) MLEM and (e) Masked-MLEM reconstructions when using the Monte Carlo method to calculate $\widetilde{A}$. The reconstructed spheres are shown in transverse and coronal slices from top to bottom.

Table 1.

Hyperparameters used in Algorithm 3 to determine the system lower and upper bounds for collimatorless imaging.

Hyperparameters	2D scintigraphy		Experimental 3D tomography	Simulated 3D tomography	
	X-rays	^{221}Fr		Solid angle	Monte Carlo
ϵ	0.04	0.04	0.04	0.004	0.004
η	0.02	0.04	0.5	0.065	−0.015
θ	0.2	0.2	0.2	0.1	0.1
ζ	0.5	0.5	0.5	0.5	0.5

Table 2.

FOMs for the 3D reconstructed images shown in figures 5, 8, and 9, as well as for the 2D collapsed images in figure 6.

Modality	Images	FOMs	MLEM (OSEM)	Masked-MLEM (Masked-OSEM)
SPECT	3D	NRMSE	0.71	0.71
		PSNR	28.43	28.50
		SSIM	0.95	0.96
Collimatorless scintigraphy	2D, X-rays	NRMSE	1.67	0.82
		PSNR	12.13	18.36
		SSIM	0.10	0.65
	2D, ^{221}Fr	NRMSE	1.67	0.85
		PSNR	12.16	17.97
		SSIM	0.10	0.67
Experimental collimatorless tomography	3D	NRMSE	3.48	2.81
		PSNR	12.68	14.54
		SSIM	0.03	0.03
Simulated collimatorless tomography	3D, solid angle	NRMSE	1.24	1.45
		PSNR	10.90	9.51
		SSIM	0.56	0.07
	3D, Monte Carlo	NRMSE	1.43	1.47
		PSNR	9.66	9.42
		SSIM	0.22	0.24