Maximum Likelihood Event Estimation and List-mode Image Reconstruction on GPU Hardware

Luca Caucci, Lars R. Furenlid, *Member, IEEE,* and Harrison H. Barrett, *Fellow, IEEE*

***Abstract*—The scintillation detectors commonly used in SPECT and PET imaging and in Compton cameras require estimation of the position and energy of each gamma ray interaction. Ideally, this process would yield images with no spatial distortion and the best possible spatial resolution. In addition, especially for Compton cameras, the computation must yield the best possible estimate of the energy of each interacting gamma ray. These goals can be achieved by use of maximum-likelihood (ML) estimation of the event parameters, but in the past the search for an ML estimate has not been computationally feasible. Now, however, graphics processing units (GPUs) make it possible to produce optimal, real-time estimates of position and energy, even from scintillation cameras with a large number of photodetectors. In addition, the mathematical properties of ML estimates make them very attractive for use as list entries in list-mode ML image reconstruction. This two-step ML process—using ML estimation once to get the list data and again to reconstruct the object—allows accurate modeling of the detector blur and, potentially, considerable improvement in reconstructed spatial resolution.**

I. Introduction

In tomographic imaging of gamma-ray-emitting radiotracers with scintillation detectors, each gamma-ray photon that interacts with a detector produces a set of signals on photomultipliers (PMTs) or other photodetectors. From these signals the position and energy of each gamma-ray interaction are determined by some algorithm, known as the *event-estimation process* and then stored, either as a list of position and energy for each interaction events or as a binned array where each bin entry represents the number of events with position and energy within some interval. We refer to the first alternative as *list-mode* data storage and the second as *binned-mode* storage. In either case, we refer to the estimated position and energy as *attributes* of the event. Image reconstruction can then be performed either directly from the attribute list or from the binned data.

The current trend towards measuring more and more attributes of each event and building larger and larger 2-D and 3-D detectors causes the memory requirements for binned data to grow geometrically, making binned-data approaches challenging. For example, a Compton camera typically yields six attributes (coordinates x and y and energy E for each of two interactions) for every measured coincidence event. Similarly, a PET system using scintillation cameras with depth-of-interaction capability can yield up to ten attributes (x, y, z and E plus time of arrival on each detector) for each event.

Manuscript received November 18, 2009. This work was supported by NIH grants 5 R37 EB000803-18 and 5 P41 EB002035-10.

The authors are with the College of Optical Sciences, University of Arizona, 1630 E. University Blvd., Tucson, Arizona 85721 and also with the Center for Gamma-Ray Imaging, University of Arizona, 1609 N. Warren Ave., Tucson, Arizona 85719.

Fortunately, list-mode (LM) maximum-likelihood (ML) expectation-maximization (EM) image reconstruction algorithms are well developed and often used for PET, Compton cameras and other applications in nuclear medicine. There are, however, two problems with LMMLEM algorithms: they are computationally demanding, and for best performance they require an accurate probability model for the attributes. It is the purpose of this paper to show that both of these problems can be addressed effectively with graphics processing units (GPUs).

In Sec. II we present the basics of LMMLEM reconstruction, and in Sec. III we show in detail how to develop the probability model for a Compton camera based on thick, monolithic scintillation crystals; a similar model applies to monolithic-crystal PET systems. In Sec. IV we discuss the advantages of using ML methods to estimate the attributes used in the reconstruction algorithm, and we show how an important asymptotic property of ML estimates can be used to simplify the probability model. Thus we recommend a dual use of ML theory, first in the event-estimation step to give values for the attributes of each event and then in the reconstruction step to estimate the object distribution from the attribute list.

In Sec. V we discuss the computational requirements of this approach and present some preliminary results obtained by running a 3-D ML position estimation algorithm on a GPU supercomputer equipped with four NVIDIA GeForce 9800 GX2 computing processors. Our GPU implementation resulted in a 250× speedup with respect to an implementation of the same algorithm on a conventional cluster machine. These promising results show the enormous potential of commodity GPU hardware for medical image reconstruction.

II. MLEM and LMMLEM algorithms

Given the data vector $\boldsymbol{g} = \{g_1, \ldots, g_M\}$, where g_m represents the number of events stored in the m^{th} bin, the goal of an image reconstruction algorithm is to produce a vector $\hat{\boldsymbol{f}} = \{\hat{f}_1, \ldots, \hat{f}_N\}$ approximating the radiotracer distribution in the object [1], [2].

A popular iterative reconstruction algorithm in the medical literature is the maximum-likelihood (ML) expectation-maximization (EM) algorithm [3], [4]. If the data vector $\boldsymbol{g}$ follows Poisson statistics, the MLEM algorithm takes the form:

$$\hat{f}_n^{(k+1)} = \hat{f}_n^{(k)} \left\{ \frac{1}{\sum_{m=1}^{M} H_{mn}} \sum_{m=1}^{M} \frac{g_m H_{mn}}{\left[H \hat{\boldsymbol{f}}^{(k)}\right]_m} \right\}, \quad (1)$$

where the matrix H (with components H_{mn}) represents a discretized version of the system operator [2]. This way of

organizing the sensitivity data requires such a large amount of storage and computational capability [5] to make it impossible to record more than just a few attributes for each event.

Alternatively, the MLEM algorithm is applied to raw event estimates recorded in list mode [5]–[9], resulting in the list-mode (LM) MLEM algorithm, which we now detail for specific configurations of a Compton camera and a PET imager.

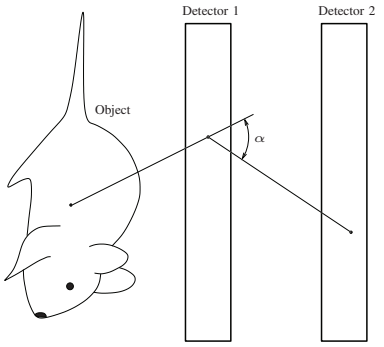

Fig. 1. Setup for a Compton camera system

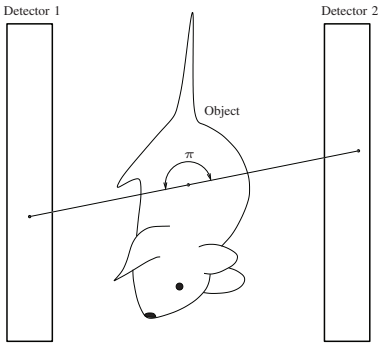

Fig. 2. Setup for a PET system

In the Compton camera of Fig. 1, two position-and-energy-sensitive thick detectors image the radiotracer distribution. The two detectors allow high sensitivity because collimation is accomplished in software rather than by placing physical collimators in front of the detectors. For the j^{th} event (with $j = 1, \ldots, J$), we can consider two 3-D positions of interaction estimates, $\hat{\boldsymbol{r}}_{j1}$ and $\hat{\boldsymbol{r}}_{j2}$ along with the estimate $\hat{\beta}_j$ of the cosine of the scatter angle $\hat{\alpha}_j$ (calculated from estimates $\hat{E}_{j1}$ and $\hat{E}_{j2}$ of the deposited energies). Given $\hat{\boldsymbol{r}}_{j1}$, $\hat{\boldsymbol{r}}_{j2}$, and $\hat{\beta}_j$ we define the j^{th} attribute vector as $A_j = (\hat{\boldsymbol{r}}_{j1}, \hat{\beta}_j, \hat{\boldsymbol{r}}_{j2})$.

In the PET setup shown in Fig. 2, two detectors are placed on opposite sides of the object. For the j^{th} event, two 3-D positions of interaction estimates are computed. Thus, the j^{th} attribute vector is $A_j = (\hat{\boldsymbol{r}}_{j1}, \hat{\boldsymbol{r}}_{j2})$. We will assume that coincidence windowing is used to pair events occurring in different detectors. We will also assume that, for each event, both photons have been collected. Energy estimation is not used as the vast majority of electron-positron decays result in two 511 keV gamma photons being emitted at an angle $\alpha \approx \pi$. For the sake of simplicity, we will assume $\alpha = \pi$.

In both cases, we can define the list $\boldsymbol{A} = \{A_1, \ldots, A_J\}$ of the J attribute vectors. Assume we subdivide the field of view in N voxels, centered at 3-D locations $\boldsymbol{r}_n$, for $n = 1, \ldots, N$. Let $\boldsymbol{f} = \{f_1, \ldots, f_N\}$ be the radiotracer distribution we want to estimate, where f_n is the number of photons emitted from voxel centered at $\boldsymbol{r}_n$. Furthermore, let S_n be the probability that a photon emitted from voxel at $\boldsymbol{r}_n$ is detected. Thus, S_n is referred to as the sensitivity for the n^{th} voxel [7]. The probability that a detected photon was emitted from voxel at $\boldsymbol{r}_n$ is then written as:

$$\Pr(\boldsymbol{r}_n|\boldsymbol{f}) = \frac{S_n f_n}{\sum_{n'=1}^{N} S_{n'} f_{n'}}.$$

If we denote with $\mathrm{pr}(A_j|\boldsymbol{r}_n)$ the probability density function (PDF) for measuring the attribute vector A_j given that a photon is emitted from the n^{th} voxel, then:

$$\mathrm{pr}(A_j|\boldsymbol{f}) = \sum_{n=1}^{N} \mathrm{pr}(A_j|\boldsymbol{r}_n)\Pr(\boldsymbol{r}_n|\boldsymbol{f}).$$

The likelihood of $\boldsymbol{f}$ given the list $\boldsymbol{A}$ of J statistically independent attribute events is:

$$\mathcal{L}(\boldsymbol{f}; \boldsymbol{A}) = \prod_{j=1}^{J} \mathrm{pr}(A_j|\boldsymbol{f}).$$

As [3] and [7] show, the iterative expression for LMMLEM reconstruction is [cf. (1)]:

$$\hat{f}_n^{(k+1)} = \hat{f}_n^{(k)} \left\{ \frac{1}{T} \sum_{j=1}^{J} \frac{\mathrm{pr}(A_j|\boldsymbol{r}_n)}{\sum_{n'=1}^{N} \mathrm{pr}(A_j|\boldsymbol{r}_{n'}) S_{n'} \hat{f}_{n'}^{(k)}} \right\}, \quad (2)$$

in which T is the total exposure time. The expression above emphasizes that the quantity $\mathrm{pr}(A_j|\boldsymbol{r}_n)$ plays a central role in the implementation of LMMLEM reconstruction algorithms.

It can be shown [7] that if the first estimate, $\hat{\boldsymbol{f}}^{(0)}$, satisfies positivity constraints, then all the subsequent estimates obtained using (2) satisfy positivity constraints as well and, under reasonable conditions [7], the algorithm has the desired property of converging to the global maximum of the likelihood function $\mathcal{L}(\boldsymbol{f}; \boldsymbol{A})$. Global convergence is proved by showing that the Hessian (with respect to $\boldsymbol{f}$) of $\log \mathcal{L}(\boldsymbol{f}; \boldsymbol{A})$ is negative definite [7].

III. Derivation of the PDF $\mathrm{pr}(A_j|\boldsymbol{r}_n)$

In this section, we provide a derivation of the expression for the probability density function $\mathrm{pr}(A_j|\boldsymbol{r}_n)$ for the Compton camera setup of Fig. 1. The same expression can easily be modified to accommodate the PET setup shown in Fig. 2.

Recall that $A_j = (\hat{\boldsymbol{r}}_{j1}, \hat{\beta}_j, \hat{\boldsymbol{r}}_{j2})$. By Bayes' theorem,

$$\mathrm{pr}(A_j|\boldsymbol{r}_n) = \int_{\text{det. 1}} \int_{-1}^{1} \int_{\text{det. 2}} \mathrm{pr}(\boldsymbol{r}_{j1}, \beta_j, \boldsymbol{r}_{j2}|\boldsymbol{r}_n) \times \mathrm{pr}(\hat{\boldsymbol{r}}_{j1}, \hat{\beta}_j, \hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j1}, \beta_j, \boldsymbol{r}_{j2}, \boldsymbol{r}_n)\, \mathrm{d}^3\boldsymbol{r}_{j1}\, \mathrm{d}\beta_j\, \mathrm{d}^3\boldsymbol{r}_{j2}, \quad (3)$$

where vectors $\boldsymbol{r}_{j1}$ and $\boldsymbol{r}_{j2}$ denote actual (i.e. not estimates) 3-D interaction within the detector. Similarly, β_j is the cosine of the actual scattering angle. Notations $\int_{\text{det. 1}}$ and $\int_{\text{det. 2}}$ denote integration over detectors 1 and 2, respectively.

Notice first that in $\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}, \hat{\beta}_j, \hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j1}, \beta_j, \boldsymbol{r}_{j2}, \boldsymbol{r}_n)$ there is no need to condition on $\boldsymbol{r}_n$ because estimates depend only on the interaction locations and the scattering angle. Furthermore,

estimates of $\hat{\boldsymbol{r}}_{j1}$ and $\hat{\boldsymbol{r}}_{j2}$ are statistically independent because they use two different sets of detector outputs. Thus:

$$\begin{aligned}&\mathrm{pr}(\hat{\boldsymbol{r}}_{j1},\hat{\beta}_j,\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j1},\beta_j,\boldsymbol{r}_{j2},\boldsymbol{r}_n)\\&\qquad=\mathrm{pr}(\hat{\boldsymbol{r}}_{j1},\hat{\beta}_j|\boldsymbol{r}_{j1},\beta_j)\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2}).\end{aligned}$$

If the detector is made of homogeneous material, then $\boldsymbol{r}_{j1}$ and β_j are statistically independent. However, $\boldsymbol{r}_{j2}$ depends on the location of the first interaction and the scattering angle. These observations lead to:

$$\mathrm{pr}(\boldsymbol{r}_{j1},\beta_j,\boldsymbol{r}_{j2}|\boldsymbol{r}_n)=\mathrm{pr}(\boldsymbol{r}_{j1}|\boldsymbol{r}_n)\mathrm{pr}(\beta_j)\mathrm{pr}(\boldsymbol{r}_{j2}|\boldsymbol{r}_{j1},\beta_j,\boldsymbol{r}_n).$$

Thus, the probability density function $\mathrm{pr}(A_j|\boldsymbol{r}_n)$ in (3) assumes the form:

$$\begin{aligned}&\mathrm{pr}(A_j|\boldsymbol{r}_n)\\&=\int_{\mathrm{det.\,1}}\int_{-1}^{1}\int_{\mathrm{det.\,2}}\mathrm{pr}(\boldsymbol{r}_{j1}|\boldsymbol{r}_n)\mathrm{pr}(\beta_j)\mathrm{pr}(\boldsymbol{r}_{j2}|\boldsymbol{r}_{j1},\beta_j,\boldsymbol{r}_n)\\&\quad\times\mathrm{pr}(\hat{\boldsymbol{r}}_{j1},\hat{\beta}_j|\boldsymbol{r}_{j1},\beta_j)\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})\,\mathrm{d}^3\boldsymbol{r}_{j1}\,\mathrm{d}\beta_j\,\mathrm{d}^3\boldsymbol{r}_{j2}.\quad(4)\end{aligned}$$

To compute $\mathrm{pr}(\boldsymbol{r}_{j1}|\boldsymbol{r}_n)$, consider a small volume element centered at $\boldsymbol{r}_{j1}$. It is convenient to take this volume element as a cube of side ε with one face perpendicular to the line connecting $\boldsymbol{r}_n$ to $\boldsymbol{r}_{j1}$. Then, the probability that a Compton interaction occurs in the cube is:

$$\Pr(\text{interaction in cube at }\boldsymbol{r}_{j1}|\boldsymbol{r}_n)=\frac{\varepsilon^3\mu_{\mathrm{C}_1}}{4\pi|\boldsymbol{r}_{j1}-\boldsymbol{r}_n|^2},$$

where μ_{C_1} is the Compton attenuation coefficient in detector 1. From the result above, it follows:

$$\mathrm{pr}(\boldsymbol{r}_{j1}|\boldsymbol{r}_n)=\frac{\mu_{\mathrm{C}_1}}{4\pi|\boldsymbol{r}_{j1}-\boldsymbol{r}_n|^2}.\qquad(5)$$

To derive an expression for $\mathrm{pr}(\boldsymbol{r}_{j2}|\boldsymbol{r}_{j1},\beta_j,\boldsymbol{r}_n)$, consider a cube of side ε center at $\boldsymbol{r}_{j2}$ and having one face perpendicular to the line connecting point $\boldsymbol{r}_{j2}$ with $\boldsymbol{r}_{j1}$. Conditioning on $\boldsymbol{r}_{j1}$ and $\boldsymbol{r}_n$ fixes a direction $\boldsymbol{s}_{1n}^{(j)}$ from the origin of the event $\boldsymbol{r}_n$ to the interaction location $\boldsymbol{r}_{j1}$. Conditioning on β_j imposes that the scattered photon lie on a cone of infinitesimal with around the direction $\boldsymbol{s}_{12}^{(j)}$, defined by $\boldsymbol{r}_{j1}$ and $\boldsymbol{r}_{j2}$. By passing to the limit for $\varepsilon\to 0$, we get:

$$\mathrm{pr}(\boldsymbol{r}_{j2}|\boldsymbol{r}_{j1},\beta_j,\boldsymbol{r}_n)=\frac{\mu_{\mathrm{P}_2}}{2\pi|\boldsymbol{r}_{j1}-\boldsymbol{r}_{j2}|^2}\delta\big(\boldsymbol{s}_{1n}^{(j)}\cdot\boldsymbol{s}_{12}^{(j)}-\beta_j\big),\quad(6)$$

in which μ_{P_2} is the photoelectric attenuation coefficient in detector 2, $\delta(\dots)$ denotes the delta function, and the unit vectors $\boldsymbol{s}_{1n}^{(j)}$ and $\boldsymbol{s}_{12}^{(j)}$ are given by:

$$\boldsymbol{s}_{1n}^{(j)}=\frac{\boldsymbol{r}_{j1}-\boldsymbol{r}_n}{|\boldsymbol{r}_{j1}-\boldsymbol{r}_n|},\qquad\boldsymbol{s}_{12}^{(j)}=\frac{\boldsymbol{r}_{j2}-\boldsymbol{r}_{j1}}{|\boldsymbol{r}_{j2}-\boldsymbol{r}_{j1}|}.\qquad(7)$$

Substituting (5) and (6) into (4) and performing the integration over β_j gives:

$$\begin{aligned}\mathrm{pr}(A_j|\boldsymbol{r}_n)=\int_{\mathrm{det.\,1}}\int_{\mathrm{det.\,2}}&\frac{\mu_{\mathrm{C}_1}\mu_{\mathrm{P}_2}}{8\pi^2|\boldsymbol{r}_{j1}-\boldsymbol{r}_n|^2|\boldsymbol{r}_{j1}-\boldsymbol{r}_{j2}|^2}\\&\times\mathrm{pr}_\beta\big(\boldsymbol{s}_{1n}^{(j)}\cdot\boldsymbol{s}_{12}^{(j)}\big)\mathrm{pr}\big(\hat{\boldsymbol{r}}_{j1},\hat{\beta}_j|\boldsymbol{r}_{j1},\boldsymbol{s}_{1n}^{(j)}\cdot\boldsymbol{s}_{12}^{(j)}\big)\\&\times\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})\,\mathrm{d}^3\boldsymbol{r}_{j1}\,\mathrm{d}^3\boldsymbol{r}_{j2}.\end{aligned}$$

So far, no approximations have been made. Assumptions that might be valid for silicon detectors are that position and energy estimates in the first detector are statistically independent, that spatial resolution does not depend of the amount of energy deposited, and, finally, that energy resolution is constant over the whole 3-D detector. With these assumptions:

$$\mathrm{pr}(\hat{\boldsymbol{r}}_{j1},\hat{\beta}_j|\boldsymbol{r}_{j1},\beta_j)=\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}|\boldsymbol{r}_{j1})\mathrm{pr}(\hat{\beta}_j|\beta_j),$$

leading to:

$$\begin{aligned}\mathrm{pr}(A_j|\boldsymbol{r}_n)=\int_{\mathrm{det.\,1}}\int_{\mathrm{det.\,2}}&\frac{\mu_{\mathrm{C}_1}\mu_{\mathrm{P}_2}}{8\pi^2|\boldsymbol{r}_{j1}-\boldsymbol{r}_n|^2|\boldsymbol{r}_{j1}-\boldsymbol{r}_{j2}|^2}\\&\times\mathrm{pr}_\beta\big(\boldsymbol{s}_{1n}^{(j)}\cdot\boldsymbol{s}_{12}^{(j)}\big)\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}|\boldsymbol{r}_{j1})\mathrm{pr}\big(\hat{\beta}_j|\boldsymbol{s}_{1n}^{(j)}\cdot\boldsymbol{s}_{12}^{(j)}\big)\\&\times\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})\,\mathrm{d}^3\boldsymbol{r}_{j1}\,\mathrm{d}^3\boldsymbol{r}_{j2}.\end{aligned}$$

If we further assume that $\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}|\boldsymbol{r}_{j1})$, $\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})$ are sharply peaked about $\hat{\boldsymbol{r}}_{j1}$ and $\hat{\boldsymbol{r}}_{j2}$, respectively, then:

$$\begin{aligned}\mathrm{pr}(A_j|\boldsymbol{r}_n)\approx&\frac{\mu_{\mathrm{C}_1}\mu_{\mathrm{P}_2}\ \mathrm{pr}_\beta\big(\hat{\boldsymbol{s}}_{1n}^{(j)}\cdot\hat{\boldsymbol{s}}_{12}^{(j)}\big)}{8\pi^2|\hat{\boldsymbol{r}}_{j1}-\boldsymbol{r}_n|^2|\hat{\boldsymbol{r}}_{j1}-\hat{\boldsymbol{r}}_{j2}|^2}\int_{\mathrm{det.\,1}}\int_{\mathrm{det.\,2}}\\&\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}|\boldsymbol{r}_{j1})\mathrm{pr}\big(\hat{\beta}_j|\boldsymbol{s}_{1n}^{(j)}\cdot\boldsymbol{s}_{12}^{(j)}\big)\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})\,\mathrm{d}^3\boldsymbol{r}_{j1}\,\mathrm{d}^3\boldsymbol{r}_{j2},\quad(8)\end{aligned}$$

where [cf. (7)]:

$$\hat{\boldsymbol{s}}_{1n}^{(j)}=\frac{\hat{\boldsymbol{r}}_{j1}-\boldsymbol{r}_n}{|\hat{\boldsymbol{r}}_{j1}-\boldsymbol{r}_n|},\qquad\hat{\boldsymbol{s}}_{12}^{(j)}=\frac{\hat{\boldsymbol{r}}_{j2}-\hat{\boldsymbol{r}}_{j1}}{|\hat{\boldsymbol{r}}_{j2}-\hat{\boldsymbol{r}}_{j1}|}.$$

If instead of a Compton camera we have the PET system of Fig. 2, then (8) becomes:

$$\begin{aligned}\mathrm{pr}(A_j|\boldsymbol{r}_n)\approx&\frac{\mu_{\mathrm{P}}\ \delta\big(\hat{\boldsymbol{s}}_{1n}^{(j)}\cdot\hat{\boldsymbol{s}}_{12}^{(j)}+1\big)}{8\pi^2|\hat{\boldsymbol{r}}_{j1}-\boldsymbol{r}_n|^2|\hat{\boldsymbol{r}}_{j1}-\hat{\boldsymbol{r}}_{j2}|^2}\\&\times\int_{\mathrm{det.\,1}}\int_{\mathrm{det.\,2}}\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}|\boldsymbol{r}_{j1})\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})\,\mathrm{d}^3\boldsymbol{r}_{j1}\,\mathrm{d}^3\boldsymbol{r}_{j2},\quad(8')\end{aligned}$$

because $\beta_0=\cos(\alpha_0)=-1$ for $\alpha_0=\pi$. The delta function that appears in (6) and (8′) is an angular delta function [2], defined as:

$$\int_{4\pi}\delta(\boldsymbol{s}-\boldsymbol{s}_0)t(\boldsymbol{s})\,\mathrm{d}\Omega_{\mathrm{s}}=t(\boldsymbol{s}_0),$$

in which the integral is performed over the solid angle 4π, Ω_{s} is the infinitesimal solid angle element, and $t(\boldsymbol{s})$ is a test function.

The two expressions for $\mathrm{pr}(A_j|\boldsymbol{r}_n)$ derived above might require experimentally determining $\mathrm{pr}(\beta_j)$, $\mathrm{pr}(\hat{\beta}_j|\beta_j)$, $\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}|\boldsymbol{r}_{j1})$, and $\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})$. Alternatively, assumptions can be made on the shape of those PDFs. For example, if we assume that $\mathrm{pr}(\hat{\beta}_j|\beta_j)$, $\mathrm{pr}(\hat{\boldsymbol{r}}_{j1}|\boldsymbol{r}_{j1})$, and $\mathrm{pr}(\hat{\boldsymbol{r}}_{j2}|\boldsymbol{r}_{j2})$ are all Gaussians, integrals in (8) and (8′) can be evaluated analytically.

IV. ML ESTIMATION OF POSITION OF INTERACTION

The integrands in (8) and (8′) contain quantities of the form $\mathrm{pr}(\hat{\boldsymbol{r}}|\boldsymbol{r})$, which expresses how well $\boldsymbol{r}$ is estimated from the detector outputs. One way to calculate $\hat{\boldsymbol{r}}$ is by means of ML estimation [10], [11]. Assume that $\boldsymbol{G}(\boldsymbol{r})$ represents noisy (i.e. random) detector outputs corresponding to an event occurring at location $\boldsymbol{r}$ in the detector. For example, $\boldsymbol{G}(\boldsymbol{r})$ might be a vector with K random components, each of them corresponding to a PMT output in a scintillation camera. In such a case, the ML estimate of $\boldsymbol{r}$ can be mathematically defined as:

$$\hat{\boldsymbol{r}}=\underset{\boldsymbol{r}_0}{\mathrm{argmax}}\,\mathrm{pr}\big[\boldsymbol{G}(\boldsymbol{r})|\boldsymbol{r}_0\big].\qquad(9)$$

If $\hat{\boldsymbol{r}}$ satisfies the condition above then $\mathrm{pr}(\hat{\boldsymbol{r}}|\boldsymbol{r})$ asymptotically approximates a multivariate normal PDF with mean $\boldsymbol{r}$ and covariance matrix given by the inverse of the Fisher information matrix $\mathbf{F}_{\hat{\boldsymbol{r}}|\boldsymbol{r}}$ for the estimation of $\boldsymbol{r}$ [2], [12], [13]:

$$\mathrm{pr}(\hat{\boldsymbol{r}}|\boldsymbol{r}) \to \frac{\sqrt{\det(\mathbf{F}_{\hat{\boldsymbol{r}}|\boldsymbol{r}})}}{(2\pi)^{3/2}} \exp\left[-\frac{1}{2}\big(\hat{\boldsymbol{r}}-\boldsymbol{r}\big)^{\mathsf{T}}\mathbf{F}_{\hat{\boldsymbol{r}}|\boldsymbol{r}}\big(\hat{\boldsymbol{r}}-\boldsymbol{r}\big)\right].$$

In a scintillation camera, multiple visible photons are produced for each gamma-ray event, and each visible photon leads to the generation of many photoelectrons. Thus, asymptotic properties of ML estimation hold true and the approximation above is a valid assumption.

The Fisher information matrix provides a way of quantifying an upper bound on how well information can be extracted from detector outputs [2]. If the noise in the components of $\boldsymbol{G}(\boldsymbol{r})$ follows a Poisson distribution, the $(m,n)^{\text{th}}$ component of the Fisher information matrix is given by [2]:

$$\big[\mathbf{F}_{\hat{\boldsymbol{r}}|\boldsymbol{r}}\big]_{mn} = \sum_{k=1}^{K} \frac{1}{\overline{G}_k(\boldsymbol{r})} \frac{\partial \overline{G}_k(\boldsymbol{r})}{\partial x_m} \frac{\partial \overline{G}_k(\boldsymbol{r})}{\partial x_n} \tag{10}$$

where $\boldsymbol{r} = (x_1, x_2, x_3)$ and $\overline{G}_k(\boldsymbol{r})$ is the mean of the k^{th} detector output for an interaction event occurring at $\boldsymbol{r}$. The derivatives that appear in (10) can be evaluated analytically by first interpolating calibration data with spline functions [14]. Also, spline interpolation can be used to deliver continuous values for $\hat{\boldsymbol{r}}$ when (9) is used.

To summarize, the overall reconstruction algorithm that gives an estimate $\hat{\boldsymbol{f}}$ of the radiotracer distribution consists of two separate ML-estimation steps:

1) For each event, use (9) to estimate the position of interaction;
2) Given the list of estimates previously calculated, iterate (2) to get $\hat{\boldsymbol{f}}$.

It is worth noting that ML methods provide a rigorous and—in some sense—optimal way to extract information from detector data. The optimality of ML estimation is used in the two steps above, giving an estimate $\hat{\boldsymbol{f}}$ that most likely produced the collected detector data.

V. Hardware configuration and results

For this research, we used a GPU supercomputer, which we assembled using commodity hardware. The hardware configuration included four NVIDIA GeForce 9800 GX2 graphics cards. The software configuration included the Linux operating system and the NVIDIA's CUDA software development kit. For a total cost of less than $3000, our GPU supercomputer is capable of about 4 TFLOPS.

The theory outlined in the previous sections requires the ability of performing 3-D ML estimation of event position from PMT data. A 3-D contracting-grid [10] algorithm was implemented for a CUDA-capable GPU device and run on our GPU machine. For the ℓ^{th} event occurring on a detector, a vector $\boldsymbol{G}_\ell = \{G_{\ell 1}, \ldots, G_{\ell K}\}$ of K PMT outputs is collected. The elements of $\boldsymbol{G}_\ell$ are integer numbers. Assuming Poisson statistics, the ML estimation problem can now be mathematically formulated as [2], [10]:

$$\begin{aligned}\hat{m}_\ell &= \underset{m\in\{1,\ldots,M\}}{\operatorname{argmax}} \left\{ \prod_{k=1}^{K} \big[\overline{G}_k(m)\big]^{G_{\ell k}} \frac{1}{[G_{\ell k}]!} e^{-\overline{G}_k(m)} \right\} \\ &= \underset{m\in\{1,\ldots,M\}}{\operatorname{argmax}} \Lambda_{\ell m}, \end{aligned} \tag{11}$$

in which the vector $\overline{\boldsymbol{G}}(m)$ is the mean PMT output for an event occurring at pixel m [15]. The goal of the contracting-grid algorithm is to compute $\hat{m}_\ell$ by avoiding the exhaustive search above. In our case, we considered a $69 \times 69 \times 25$-voxel 3-D detector and we assumed $K = 64$ PMT outputs for each event. The contracting-grid algorithm starts by considering a uniform $4 \times 4 \times 4$ coarse grid that covers the whole detector. The quantity $\Lambda_{\ell m}$ in (11) is evaluated at each point of the grid and the value of m that maximizes $\Lambda_{\ell m}$ on the grid is used as the center of the grid in the next iteration. The grid spacing is halved at each iteration. The algorithm terminates when the grid spacing is one voxel in each dimension.

In our GPU implementation, each event was assigned to a $4 \times 4 \times 4$ thread block. The threads in a thread block are responsible for loading into the shared memory the PMT outputs $\boldsymbol{G}_\ell$. Once the PMT data are loaded, each thread computes the likelihood $\Lambda_{\ell m}$ for a point on the grid. The 64 values of the likelihood are then searched for the maximum value. The center of the grid for the next iteration and the new grid spacing are calculated as outlined above, and shared with all the threads in the block using shared variables.

Each NVIDIA GeForce 9800 GX2 graphics card installed in our machine contains two CUDA-capable devices, which can be used independently, for a total of eight devices. We also implemented the same contracting-grid algorithm for a conventional cluster machine, which we used for our comparison. Results of the comparison are reported in Table I.

TABLE I
Comparison results

Architecture	Events/s	Speedup
AMD Phenom 9850 2.5 GHz CPU	989.67	—
NVIDIA GeForce 9800 GX2, 1 device	34224.59	34.58
NVIDIA GeForce 9800 GX2, 8 devices	246003.92	248.57

VI. Conclusions

In this paper the list-mode (LM) MLEM algorithm is applied to ML position and energy estimates for a Compton camera. An expression for the probability density function of the event attribute vector was given. The full implementation of the reconstruction algorithm requires two distinct estimation steps: the first is the 3-D ML estimation of event position from PMT data and the second consists in the reconstruction of the object distribution from the list of position estimates by using an iterative formula. We have written a GPU algorithm that implements the first estimation step and we ran our code on a GPU supercomputer assembled from commodity hardware. Our implementation could achieve a 250× speedup with respect to an implementation on a conventional cluster machine.

We are currently working on the GPU implementation of the iterative reconstruction step.

REFERENCES

[1] H. H. Barrett, "Objective assessment of image quality: Effects of quantum noise and object variability," *J. Opt. Soc. Am. A*, vol. 7, no. 7, pp. 1266–1278, Jul. 1990.

[2] H. H. Barrett and K. J. Myers, *Foundations of Image Science*. Hoboken, NJ: Wiley-Interscience, 2004.

[3] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. B Met.*, vol. 32, no. 1, pp. 1–38, 1977.

[4] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. 1, no. 2, pp. 113–122, Oct. 1982.

[5] H. H. Barrett, T. White, and L. C. Parra, "List-mode likelihood," *J. Opt. Soc. Am. A*, vol. 14, no. 11, pp. 2914–2923, Nov. 1997.

[6] A. Reader, R. Manavaki, S. Zhao, P. Julyan, D. Hastings, and J. Zweit, "Accelerated list-mode EM algorithm," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 1, pp. 42–49, Feb. 2002.

[7] L. C. Parra and H. H. Barrett, "List-mode likelihood: EM algorithm and image quality estimation demonstrated on 2-D PET," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 228–235, Apr. 1998.

[8] D. L. Snyder and D. G. Politte, "Image reconstruction from list-mode data in an emission tomography system having time-of-flight measurements," *IEEE Trans. Nucl. Sci.*, vol. 30, no. 3, pp. 1843–1849, Jun. 1983.

[9] A. Reader, K. Erlandsson, R. Ott, and M. Flower, "Attenuation and scatter correction of list-mode data driven iterative and analytic image reconstruction algorithms for rotating 3D PET systems," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 2218–2226, Dec. 1999.

[10] L. R. Furenlid, J. Y. Hesterman, and H. H. Barrett, "Real-time data acquisition and maximum-likelihood estimation for gamma cameras," in *14th IEEE-NPSS Real Time Conference*, Jun. 2005, pp. 498–501.

[11] J. Y. Hesterman, L. Caucci, M. A. Kupinski, H. H. Barrett, and L. R. Furenlid, "Maximum-likelihood estimation with a contracting-grid search algorithm," 2009, accepted for publication in *IEEE Trans. Nucl. Sci.*

[12] H. H. Barrett, W. C. J. Hunter, B. W. Miller, S. K. Moore, Y. Chen, and L. R. Furenlid, "Maximum-likelihood methods for processing signals from gamma-ray detectors," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 3, pp. 725–735, Jun. 2009.

[13] R. A. Fisher, "Theory of statistical estimation," *P. Camb. Philos. Soc.*, vol. 22, pp. 700–725, 1925.

[14] L. A. Piegl and W. Tiller, *The $\mathcal{NURBS}$ book*. Berlin, Germany: Springer, 1997.

[15] W. C. J. Hunter, H. H. Barrett, and L. R. Furenlid, "Calibration method for ML estimation of 3D interaction position in a thick gamma-ray detector," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 1, pp. 189–196, Feb. 2009.