# Gamma-Ray Point-Source Localization and Sparse Image Reconstruction using Poisson Likelihood

Daniel Hellfeld, Tenzing H. Y. Joshi, Mark S. Bandstra, Reynold J. Cooper, Brian J. Quiter, and Kai Vetter

***Abstract*—Gamma-ray imaging attempts to reconstruct the spatial and intensity distribution of gamma-emitting radionuclides from a set of measurements. Generally, this problem is solved by discretizing the spatial dimensions and employing the maximum likelihood expectation maximization (ML-EM) algorithm, with or without some form of regularization. While the generality of this formulation enables use in a wide variety of scenarios, it is susceptible to overfitting, limited by the discretization of spatial coordinates, and can be computationally expensive. We present a novel approach to 3D gamma-ray image reconstruction for scenarios where sparsity may be assumed, for example radiological source search. In this work we first formulate a point-source localization (PSL) approach as an optimization problem where both position and source intensity are continuous variables. We then extend and generalize this formulation to an iterative algorithm called additive point-source localization (APSL) for sparse parametric image reconstruction. A set of simulated source search scenarios using a single non-directional detector are considered, finding improved image accuracy and computational efficiency with APSL over traditional grid-based approaches.**

***Index Terms*—radiological source search, source localization, Poisson likelihood, maximum likelihood, gamma-ray imaging**

## I. Introduction

THE reconstruction of the spatial and intensity distribution of gamma-emitting radionuclides from a set of Poisson-distributed measurements is employed for applications ranging from medical imaging to nuclear security. In the case of medical imaging, measurements are often from many static detectors viewing a stationary volume, whereas in the case of nuclear security, measurements may be from one or several detectors that have moved through an environment. In both cases, the position and orientation (pose) of the detectors relative to the scene must be known; in the former the detector geometry is known *a priori*, however in the latter case pose information may be provided by geospatial positioning (GPS), an inertial navigation system (INS), or from simultaneous localization and mapping (SLAM) [1], [2].

Here we consider the subset of scenarios where the source-term may be assumed sparse. In the case of radiological source search, the task is often to localize and quantify the activity of one or more compact sources of gamma-ray radiation. Under a point-source assumption, previous work in static regularly-spaced 2D detector networks has considered techniques such as triangulation using least-squares [3], hybrid grid-based maximum likelihood (ML) estimation and expectation maximization (EM) [4], and regularized ML-EM pre-conditioned Fisher's scoring iterations [5]. In the space of adaptive networks or moving detectors, approaches such as two-stage algorithms [6], adaptive likelihoods [7], and sequential Bayesian estimation using particle filters [8] have been explored. Other approaches in the literature include grid-refinement and iterative pruning [9], [10], [11], [12].

The existing approaches tend to localize a known number of sources (primarily in 2D), assume background is known or can be estimated prior to source localization, and rely on the discretization of the spatial and intensity domains either directly or for initialization of a direct solver. Grid-based methods are limited in accuracy and can be computationally intractable when searching for multiple sources [5]. Moreover, methods developed in 2D may not be easily extensible to 3D localization, due to increasing degeneracy and non-convexity in the solution space.

More recent work has approached the problem in the maximally general case (unknown background and no priors on the source distribution), using ML-EM to perform 2D and 3D image reconstruction of sparse and distributed sources with a variety of static and free-moving detector systems including hand-held [13], [14], ground-vehicle [15], and airborne platforms [16], [17]. Various regularization approaches to impose assumptions about the source distribution have also been studied [18], [19], [20]. While the generality of the ML-EM formulations enable use in a wide variety of scenarios, it is also susceptible to overfitting, limited by spatial discretization, and can be computationally expensive and memory intensive.

In this work, we first consider the general ML-EM approach in a discretized 3D space subject to Poisson gamma-ray counting statistics, with and without sparsity regularization, and demonstrate the limitations of such an approach. We then propose a reformulation of the problem to one in which the source model is confined to a single voxel, i.e., point-source localization (PSL). This approach is similar to prior work [4], though here we reconstruct in 3D and include background as a free parameter. We then consider the PSL

Manuscript received July 18, 2019. This material is based on work supported by the Department of Energy National Nuclear Security Administration through the Nuclear Science and Security Consortium under Award No. DE-NA0003180 and by the Defense Threat Reduction Agency under HDTRA 10027-28018. Distribution A: approved for public release, distribution is unlimited. This support does not constitute an express or implied endorsement on the part of the United States Government.

D. Hellfeld is with the Nuclear Engineering Department at the University of California, Berkeley, Berkeley, CA 94720 USA (email: dhellfeld@berkeley.edu).

T. H. Y Joshi, M. S. Bandstra, R. J. Cooper and B. J. Quiter are with the Applied Nuclear Physics Program at Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA.

K. Vetter is with the Nuclear Engineering Department at the University of California, Berkeley and the Applied Nuclear Physics Program at Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA.

problem in which both the spatial and intensity domains are continuous. Finally, we extend the continuous PSL formalism for general sparse image reconstruction, or additive point-source localization (APSL), where the image is considered the sum of multiple point-sources whose position and intensity are continuous in nature. APSL mitigates over-fitting in its iterative bottom-up nature and statistically-founded stopping criteria and, because of the inherent point-source assumption and continuous variables, results in images with improved accuracy and interpretability as compared with traditional grid-based approaches. Furthermore, we show that APSL offers this enhanced performance at a reduced computational burden.

This work serves as an introduction and simple demonstration of the APSL approach. The formalism is general in nature in that it can search for an unknown number of sources in an unknown background environment using static and dynamic detectors or detector arrays. A simulated source search scenario comprising a single free-moving detector with uniform directional sensitivity whose poses are derived using SLAM is considered here. The detector poses are subject to arbitrary rotations and translations in 3D as if being carried by a human operator. An isotropic detector was used for simplicity and clarity, though this case is challenging for both localization and quantification due to the inability to break solution degeneracy. An exhaustive study of the localization and imaging performance of APSL and the impact of non-uniform directional sensitivity and detector arrays will be addressed in subsequent works.

The structure of the paper is as follows: the traditional and sparsity regularized ML-EM approaches for a source search scenario are outlined in Sec. II, followed by the discrete and continuous PSL formalisms on the same scenario in Sec. III. The APSL algorithm is demonstrated on a multi-source localization scenario in Sec. IV. A summary and future work are presented in Sec. V−VI.

## II. Maximum Likelihood Approaches

### A. Poisson Likelihood

Gamma-ray measurements are governed by Poisson statistics since they involve independent counting of discrete events. The negative log-likelihood of a set of $I$ measurements $\mathbf{x}^{[I\times 1]} = [\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_I]^T$ in units of counts per unit integration time from mean-rates $\boldsymbol{\lambda}^{[I\times 1]}$ is

$$\ell(\mathbf{x}|\boldsymbol{\lambda}) = [\boldsymbol{\lambda} - \mathbf{x} \odot \log \boldsymbol{\lambda} + \log[\Gamma(\mathbf{x}+1)]]^T \cdot \mathbf{1}\,, \qquad (1)$$

where $\odot$ denotes element-wise multiplication and $\Gamma(\cdot)$ is the gamma function. The system matrix $\mathbf{V}^{[I\times J]}$ in units of inverse activity ($\mathrm{Bq}^{-1}$) describes the geometric and detector efficiency of the $I$ measurements relative to $J$ image voxels with intensities $\mathbf{w}^{[J\times 1]}$ in units of activity (Bq). The mean-rates $\boldsymbol{\lambda}$ in units of counts per unit integration time are the forward-projection of the voxel intensities

$$\boldsymbol{\lambda} = \mathbf{V} \cdot \mathbf{w} + b\boldsymbol{t}\,, \qquad (2)$$

where $b$ is a background rate, assumed herein to be constant, in a single detector and $\boldsymbol{t}^{[I\times 1]}$ are measurement time durations. Detection algorithms that operate in variable background environments are currently under development and could eliminate the need for the constant background assumption [21].

The notation presented here is for a single detector system, however, the formulation is easily extensible to multi-detector systems. In this case, the number of measurements $I$ will increase according to the number of detectors $D$ (i.e., $I \leftarrow I \times D$). Additionally, the background rate in each detector is treated as a free variable, such that $b \Rightarrow \boldsymbol{b}^{[D\times 1]}$ and $\boldsymbol{t} \Rightarrow \boldsymbol{t}^{[I\times D]}$, where only a single element of each row in $\boldsymbol{t}$ is non-zero and Eq. 2 takes the form $\boldsymbol{\lambda} = \mathbf{V} \cdot \mathbf{w} + \boldsymbol{t} \cdot \boldsymbol{b}$.

### B. Maximum Likelihood Expectation Maximization

ML-EM [22] is an iterative algorithm that solves for the maximum likelihood estimate of intensity and background by minimizing Eq. 1

$$\hat{\mathbf{w}}, \hat{b} = \underset{\mathbf{w},b}{\operatorname{argmin}}\ \ell(\mathbf{x}|\mathbf{w}, b)\,. \qquad (3)$$

Without background, the update equation at iteration $q+1$ is

$$\hat{\mathbf{w}}^{(q+1)} = \frac{\hat{\mathbf{w}}^{(q)}}{\boldsymbol{\varsigma}} \odot \mathbf{V}^T \cdot \frac{\mathbf{x}}{\mathbf{V} \cdot \hat{\mathbf{w}}^{(q)}}\,, \qquad (4)$$

where the sensitivity $\boldsymbol{\varsigma}^{[J\times 1]} = \mathbf{V}^T \cdot \mathbf{1}$ and $\hat{\mathbf{w}}$ is typically initialized ($q = 0$) with a flat image. With the inclusion of a constant background rate across all measurements, the updates can be separated into source intensity and background equations. First defining the comparator term as

$$\boldsymbol{\xi}^{(q)} = \frac{\mathbf{x}}{\mathbf{V} \cdot \hat{\mathbf{w}}^{(q)} + \hat{b}^{(q)}\boldsymbol{t}}\,, \qquad (5)$$

the updates become

$$\begin{aligned} \hat{\mathbf{w}}^{(q+1)} &= \frac{\hat{\mathbf{w}}^{(q)}}{\boldsymbol{\varsigma}} \odot \left[\mathbf{V}^T \cdot \boldsymbol{\xi}^{(q)}\right]\,, \\ \hat{b}^{(q+1)} &= \frac{\hat{b}^{(q)}}{T}\left[\boldsymbol{t}^T \cdot \boldsymbol{\xi}^{(q)}\right]\,, \end{aligned} \qquad (6)$$

where the total measurement time $T = \boldsymbol{t}^T \cdot \mathbf{1}$ and $\hat{b}^{(0)}$ is typically the median of $\mathbf{x}$. This approach is statistically founded and highly general, but, as we will show, can result in overfitting in sparse and underdetermined scenarios. Furthermore, the calculation of $\mathbf{V}$ can be computationally expensive and memory intensive when $I \times J \gg 1$.

To demonstrate the limitations of ML-EM for sparse image reconstruction, we consider a simulated isotropic (non-directional) 100% efficient detector (effective area of 5 $\mathrm{cm}^2$) at points along a short experimental human-walked trajectory (270 poses) in a 20 $\mathrm{m}^2$ area. The trajectory was tracked with a LiDAR-IMU sensor system and the Google Cartographer SLAM algorithm [23]. The elevation (Z) of the path varied slightly, roughly $\pm 30$ cm about the XY plane (Z = 0). The choice of a simulated detector and measured trajectory was made here for the ease of single (and multiple) point-source injection studies while capturing the variation in detector-source distance (in 3D) observed in real free-moving source search scenarios. A 5 $\mu$Ci (0.185 MBq) point-source was simulated in the XY plane with a closest approach of $\sim$45 cm

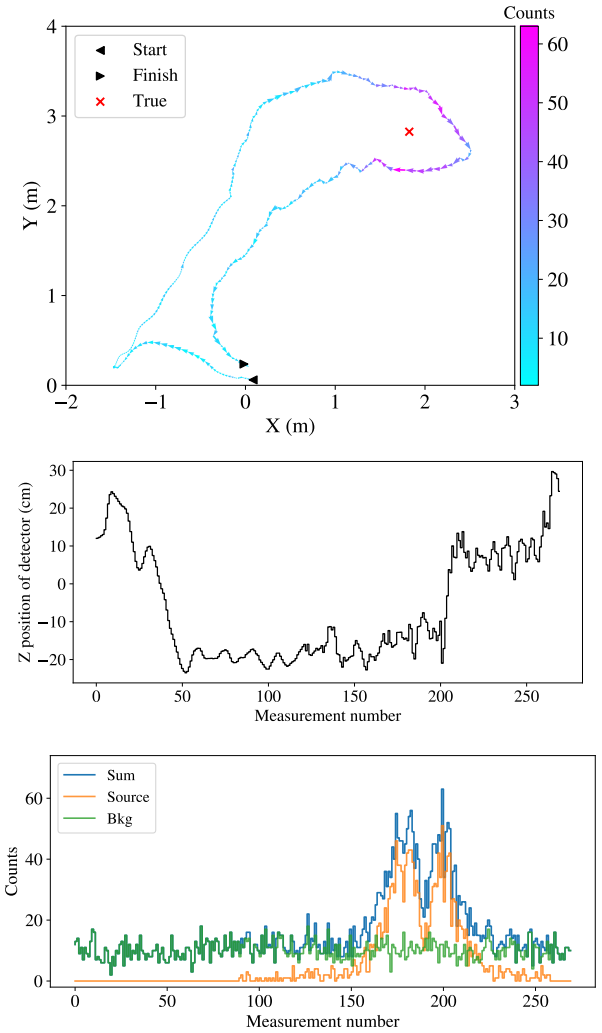


Fig. 1. (Color online) (Top) Measurement path in the XY plane, colorized with the total simulated counts at each measurement. The path is a human-walked trajectory tracked with a SLAM system. The simulated measurements remained static at each position for a duration of 0.1 s. The source position is shown with a red "x". (Middle) Measurement path along the Z dimension. (Bottom) Simulated signal and background at each measurement.

and without regard to any spatial discretization for ML-EM. A background count rate of 100 counts per second was assumed and each simulated measurement had a duration of 0.1 s.

The top pane of Fig. 1 shows the path of the detector in the XY plane with heading described with arrows, colorized with the total counts at each measurement. Larger arrows indicate a faster speed in the trajectory, though the simulations were done statically at each position. The source position is shown with a red "x". The middle pane shows the position of the detector in Z at each measurement. Notice the detector was positioned at Z $\approx -14$ cm for the majority of the path near the source. The bottom pane shows the simulated signal and background ("bkg") at each measurement.

The directional sensitivity of the detector is represented through the computation of $\mathbf{V}$. A non-directional detector was used in this work to limit the reconstruction to "proximity imaging" (relying only on flux modulation from the inverse square distance to the source) and focus on the development and performance of the algorithms that follow.

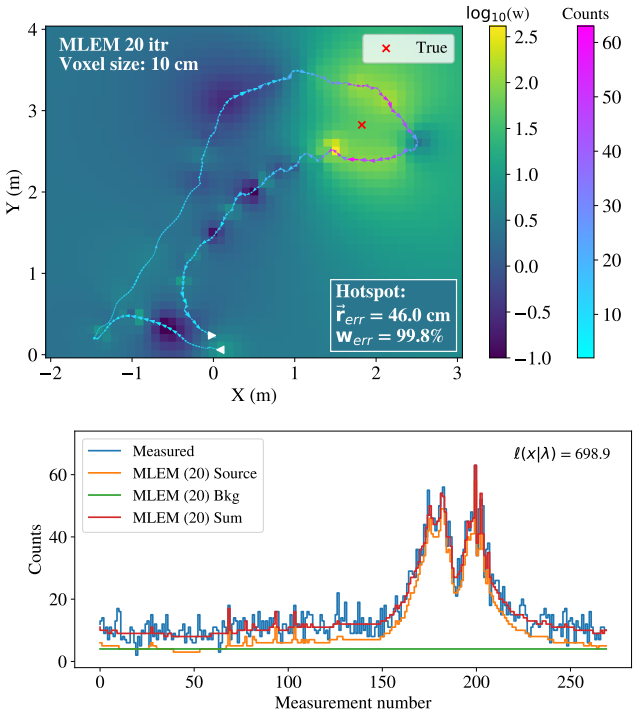


Fig. 2. (Color online) (Top) ML-EM reconstruction $\hat{\mathbf{w}}$ (20 iterations), in $\log_{10}$ scale, of a slice through the XY plane (Z = 0). A 5 $\mu$Ci (0.185 MBq) source was placed at the red "x" at Z = 0. (Bottom) Forward projection of ML estimates into count space $\hat{\boldsymbol{\lambda}}$ compared to the measurement $\mathbf{x}$. The total negative log-likelihood of the fit is shown in the upper right corner.

The 3D image space was discretized into cubic voxels (10 cm length), with a Z extent from -3 to 3 m (total of $\sim 1.28 \times 10^5$ voxels). The image space was centered around Z = 0 as the path of the detector was roughly centered around this plane.

Figure 2 shows a slice of the ML-EM reconstruction $\hat{\mathbf{w}}$ (20 iterations), in $\log_{10}$ scale, along the XY plane (Z = 0). The number of iterations reflects a $1.5 \times 10^{-3}$ fractional change in negative log-likelihood and thus a reasonable estimate for convergence. Twenty iterations were used in all subsequent reconstructions for consistency (in computation) between methods. Also shown is the forward projection of the ML estimates into measurement space ($\hat{\boldsymbol{\lambda}} = \mathbf{V} \cdot \hat{\mathbf{w}} + \hat{b}\boldsymbol{t}$) compared to the measurement, $\mathbf{x}$. The optimized negative log-likelihood (hereby referred to as the "loss") of the fit is shown in the upper right corner of the lower plot. As the problem is underdetermined, the ML-EM solution significantly overfits, placing intensities in many voxels across the whole image space, ultimately just fitting to the noise. This effect increases with additional iterations.

The system matrix calculation and ML-EM reconstruction were run on a quad-core 2.7 GHz Intel Core i7 processor with runtimes of 22 s and 0.7 s, respectively. The same hardware was used for all reconstructions in this work.

While it minimizes the loss, the solution fails to localize the source and places most of the intensity near the path of the detector. The bias towards the measurement path results in a localization error (Euclidian distance between the true location and the highest intensity voxel center, $\vec{r}_{\text{err}} = ||\vec{r}_{\text{true}} -$

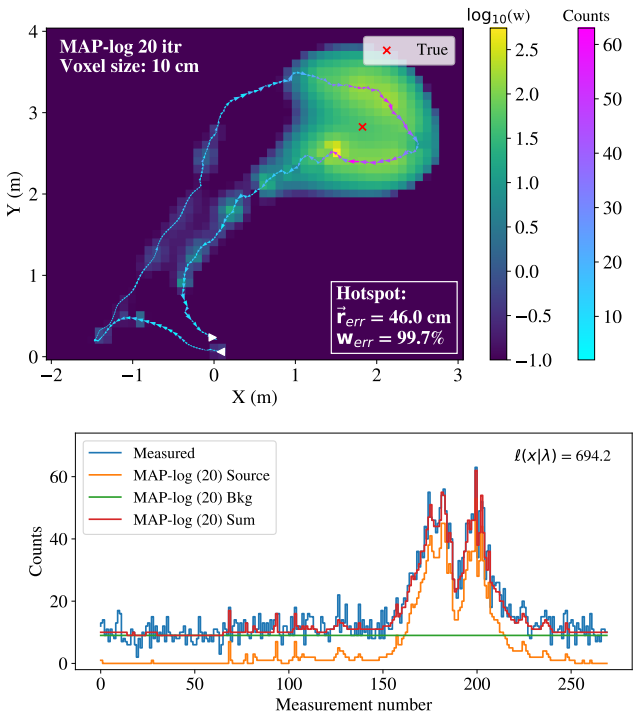


Fig. 3. (Color online) (Top) MAP reconstruction $\hat{\mathbf{w}}$ (20 iterations of EM) with the log prior in Eq. 8 ($\delta = 1$ and $\rho = 0.01$), in $\log_{10}$ scale, of a slice through the XY plane. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{\lambda}}$ compared to the measured signal $\mathbf{x}$. The total negative log-likelihood of the fit is shown in the upper right corner.

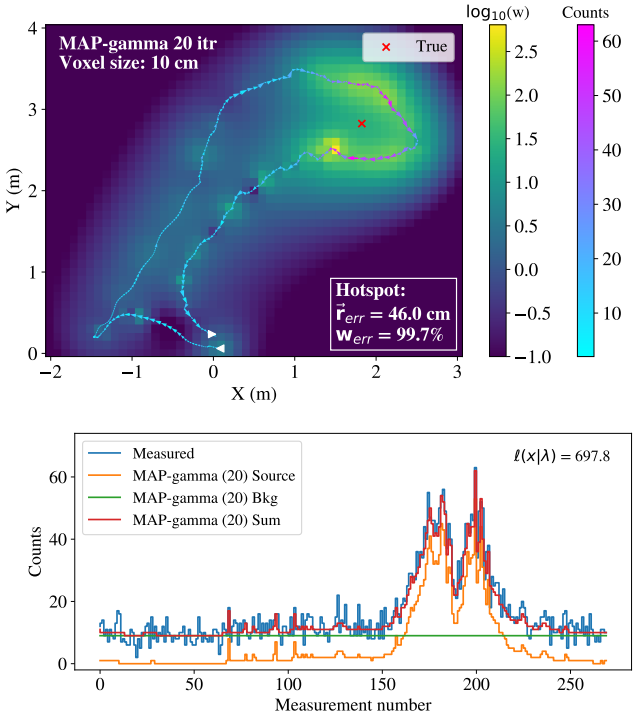


Fig. 4. (Color online) (Top) MAP reconstruction $\hat{\mathbf{w}}$ (20 iterations of EM) with the gamma prior in Eq. 9 ($\alpha = 1.01, \beta = 1$, and $\rho = 10^{-3}$), in $\log_{10}$ scale, of a slice through the XY plane. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{\lambda}}$ compared to the measurement. The total negative log-likelihood of the fit is shown in the upper right corner.

$\vec{r}_{j=\text{argmax}(\hat{\mathbf{w}})}||$) of 46 cm ($\sim$100% of closest approach). To compensate for a much closer source, the individual source intensities are much lower than the true value. However, the ML-EM background estimate is considerably lower than the true rate, resulting in more source intensity being placed throughout the 3D image space (total of $\sim$19 $\mu$Ci).

### C. Maximum A Posteriori

*A priori* knowledge about the image distribution can be incorporated into Eq. 3 with

$$\hat{\mathbf{w}}, \hat{b} = \underset{\mathbf{w},b}{\text{argmin}}\ \ell(\mathbf{x}|\mathbf{w}, b) + \rho R(\mathbf{w})\,, \tag{7}$$

where $R$ is a convex regularizer function or penalty on the intensities, $\rho$ controls the strength of the regularization, and $\hat{\mathbf{w}}$ is the maximum *a posteriori* (MAP) estimate.

Lingenfelter et al. [18] studied the effects of various penalties to enforce sparsity in the image intensity distributions. The sparsity-enforcing $l_0$ and $l_1$ norms were shown to produce solutions that were equivalent or scaled versions of the ML-EM solution. The authors proposed a non-convex penalty based on the sum of the logarithm of all image intensity values resulting in a sparser solution than traditional ML-EM:

$$R_{\log}(\mathbf{w}) = \sum_{j=1}^{J} \log\left(\frac{\mathrm{w}_j}{\delta} + 1\right)\,, \tag{8}$$

where $\delta$ is a scale parameter.

Figure 3 shows the MAP solution $\hat{\mathbf{w}}$ (20 iterations of EM) to the data in Fig. 1 using the log prior ($\delta = 1$ and $\rho = 0.01$) as well as the comparison of $\hat{\boldsymbol{\lambda}}$ to the measurement. The reconstruction runtime and likelihood convergence were similar to ML-EM.

Hyperparameter optimization was done with a coarse grid search over $\delta$ and $\rho$ ($\delta, \rho \in 10^{\{-3,-2,-1,0,1,2\}}$), maximizing the fraction of the intensity near the source (within half the distance of closest approach). The optimal parameters were below the suggested range of $\delta$ and $\rho$ in [18] due to the difference in scale in the data. While the solution is successful in suppressing much of the activity surrounding the trajectory and producing a better background estimate than ML-EM, the log prior still suffers from overfitting, fails to localize the source, and biases the intensity near the path.

Another prior used in image reconstruction, developed primarily for emission and transmission tomography [24], [25], is the gamma prior

$$R_\Gamma(\mathbf{w}) = \sum_{j=1}^{J} \left[\frac{\alpha_j \mathrm{w}_j}{\beta_j} - (\alpha_j - 1)\log \mathrm{w}_j\right]\,, \tag{9}$$

where $\alpha_j$ and $\beta_j$ correspond to the mean ($\beta_j$) and variance ($\beta_j^2/\alpha_j$) of the gamma density for each voxel. This follows naturally for the underlying Poisson intensities as the gamma distribution is the conjugate prior of the Poisson likelihood. More importantly, a gamma distribution with a constant low mean and variance in each voxel will enforce sparsity as it penalizes the addition of source intensity far from the mean.

Figure 4 shows the MAP solution $\hat{\mathbf{w}}$ (20 iterations of EM) to the data in Fig. 1 using the gamma prior ($\alpha = 1.01, \beta = 1$, and $\rho = 10^{-3}$) and the fit in count space. Again, the recon-

struction runtime and likelihood convergence were similar to ML-EM. A grid search approach similar to above was used for hyperparameter optimization ($\alpha$, $\beta$, and $\rho$). Similar to the log regularizer, the gamma prior suppresses intensity surrounding the trajectory, though it still places intensity along the track and fails to localize the source.

In both cases presented here, regularizing the reconstruction using prior information fails to improve source localization and quantification. Ultimately both problems are still severely ill-posed. However, if the unknown source is known to be a point-source (i.e., a source occupying a single voxel), such as in radiological source search, significant improvements can be made by reformulating the optimization problem in Eq. 3.

## III. Point-Source Localization

### A. Discrete Space

Under the point-source assumption, the optimal background rate and point-source intensities for each voxel can be solved using ML-EM with the following replacements

$$\begin{aligned}
\boldsymbol{\lambda} &\Rightarrow \boldsymbol{\Lambda}^{[I\times J]} = [\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_J]\,, \\
\mathbf{w} &\Rightarrow \mathbf{W}^{[J\times J]} = \mathrm{diag}(\mathbf{w})\,, \\
\mathbf{x} &\Rightarrow \mathbf{X}^{[I\times J]} = [\mathbf{x}, \ldots, \mathbf{x}]\,, \\
b &\Rightarrow \boldsymbol{b}^{[1\times J]} = [b_1, \ldots, b_J]\,, \\
\boldsymbol{t} &\Rightarrow \boldsymbol{T}^{[I\times J]} = [\boldsymbol{t}, \ldots, \boldsymbol{t}]\,.
\end{aligned}$$

Equation 2 then becomes

$$\boldsymbol{\Lambda} = \mathbf{V}\cdot\mathbf{W} + \boldsymbol{b}\odot\boldsymbol{T}. \tag{10}$$

The voxel in discrete space that alone best describes the data (minimum loss) may then be identified by solving Eq. 3 for each voxel. This can be thought of as individual ML-EM problems in each voxel, each solving for the optimal weight (and background) that best explains the measured data. While Eqs. 6 and 10 can be refactored for computational efficiency, the reconstruction will be slower than traditional ML-EM because more computations are required. In the point-source scenario, an analytical solution exists for the source intensities in the absence of background

$$\mathbf{W} = \mathbf{X}^T\cdot\mathbf{1}\,/\,\mathbf{V}^T\cdot\mathbf{1}\,. \tag{11}$$

This solution can be computed exactly with little computational power, however it does not hold in the case of the unknown background rate presented here. The analytical solution is used as an improved initial image (over a flat image) to increase the speed of convergence in the iterative approach. While only 5 iterations were required for the minimum loss voxel to achieve the same likelihood convergence as the ML approaches above, this was not true in every voxel in the image space. Therefore 20 iterations were used to be consistent with the other approaches.

Figure 5 shows the ML-EM loss for each single-voxel model of a slice through the XY plane. The reconstruction runtime was 25.2 s. The distribution agrees with intuition as the detector is non-directional, leading to many voxels outside of the track that can fit sections of the data (one of the two

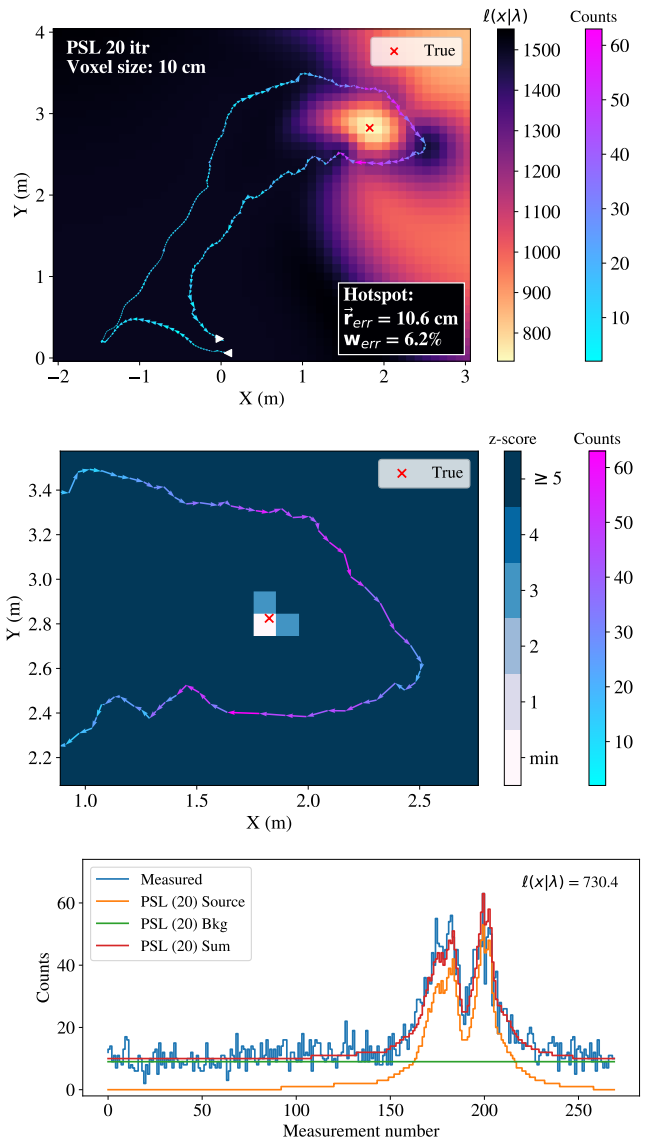


Fig. 5. (Color online) (Top) ML-EM loss (20 iterations), in $\log_{10}$ scale, for each single-voxel model (PSL) of a slice through the XY plane. (Middle) Likelihood ratio test statistic of the loss' to the minimum loss, represented with a z-score (zoomed in near the true source location). (Bottom) Forward projected mean-rates of the minimum loss single-voxel model $\hat{\boldsymbol{\lambda}}$ compared to the measurement $\mathbf{x}$.

peaks in the data) albeit with a higher loss. It is only the voxels near the true location that capture both peaks in the data, reflecting the lower loss.

The minimum loss voxel was one voxel above the true source location in Z (resulting in the position error $\sim$ 10 cm). This is not surprising as a non-directional detector was used and there was little change in elevation over the measurement path. Therefore some degree of degeneracy may exist in the maximum likelihood solution. Note the activity estimate for a point-source in this voxel was 4.7 $\mu$Ci ($<$ 10% error).

While the minimum loss voxel can be used directly for the single-source localization (assuming the detection of the source above background has already been done with some other means, e.g., gross counts or spectral), spatial confidence intervals can be computed to bound the true source location

with some degree of certainty (e.g., 95%). Under the assumption that the minimum loss voxel contains the true source location, confidence intervals can be computed around the voxel by comparing to the surrounding likelihoods with a likelihood ratio test.

The comparison of each loss $\ell_j$ to the minimum loss $\ell_{\text{min}}$ remains in the interior of the parameter space, and thus Wilks' Theorem [26] states that the test statistic of twice the negative log-likelihood difference

$$z = 2(\ell_j - \ell_{\text{min}}) \,, \tag{12}$$

will be asymptotically distributed like $\chi^2_k$, where the number of degrees of freedom $k$ is the difference in the number of free parameters between $\ell_{\text{min}}$ and $\ell_j$. The parameters of $\ell_{\text{min}}$ are all fixed, resulting in a difference of 5 free parameters (one source intensity, one background rate, and three spatial coordinates of the source), therefore $z \sim \chi^2_5$. For a system of $N_d$ detectors, the solution would generalize to $k = 4 + N_d$.

Note that with degenerate maximum likelihood estimates (as can be the case with non-directional systems and limited movement in one or more of the spatial dimensions) the number of degrees of freedom $k \leq 4 + N_d$. Simulations or resampling methods would be required to determine the shape of the distribution and the true value of $k$, which may not be feasible in an operational search setting. However, assuming a larger $k$ will always result in a more conservative confidence interval. While some degree of degeneracy may exist in this problem presented here, a value of $k = 5$ was used for the single detector system.

The equivalent Gaussian sigma (z-score) of a random variable $x$ distributed like $\chi^2_k$ is

$$\text{z-score} = \sqrt{2}\ \text{erf}^{-1}\left[\Phi_{\chi^2_k}(x)\right] \tag{13}$$

where $\text{erf}^{-1}(\cdot)$ is the inverse error function and $\Phi_{\chi^2_k}(\cdot)$ is the cumulative distribution function (CDF) of the $\chi^2_k$ distribution. The middle image of Fig. 5 shows the confidence intervals around the minimum loss voxel in units of z-scores (zoomed in near the source location). It is clear that the localization is sharply peaked with high confidence ($\geq 5\sigma$), demonstrating the superior localization capabilities of the approach.

The calculation of confidence intervals for the source intensity is more complex as it is coupled to the reconstructed position and background estimates. A more appropriate approach in this case would be to compute covariances from the Fisher information matrix (Cramer-Rao bound) or by determining the distributions empirically by gridding over intensity and background for the positions within some given confidence limit. However, this is outside the scope of this work.

The bottom image of Fig. 5 shows the minimum loss single-voxel model fit in count space. The solution is no longer overfit (resulting in a larger loss compared to the ML-EM and MAP approaches). The simple reformulation of the problem drastically improves both the localization and quantification performance, though, if the reconstruction is in discrete space, at the cost of a higher computational burden.

### *B. Continuous Space*

The minimization of Eq. 1 can be reformulated as an optimization problem not only in continuous intensity, but also in the continuous 3D spatial coordinates of the point-source ($\vec{r_s}$)

$$\underset{(\text{w}_s, \vec{r_s}, b)}{\text{argmin}}\ \ell(\mathbf{x}|\text{w}_s, \vec{r_s}, b) \,, \tag{14}$$

where the $i^{\text{th}}$ instance of Eq. 2 from a source at position $\vec{r_s}$ is given by

$$\lambda_i = v_{is}\text{w}_s + bt_i \,, \tag{15}$$

and the system response for measurement $i$ from a point-source at $\vec{r_s}$ (neglecting attenuation) is

$$v_{is} \approx \frac{\eta(\vec{r_s}, \vec{r_i})t_i}{|\vec{r_i} - \vec{r_s}|^2} \,, \tag{16}$$

where $\eta(\vec{r_s}, \vec{r_i})$ is the angular response of the detector at position $\vec{r_i}$ to a point-source at position $\vec{r_s}$. For an isotropic detector $\eta$ = constant.

In this formulation, the response is calculated only where needed in the optimization. This removes the need to compute the entire 3D system matrix, $\mathbf{V}$, significantly reducing the computational and memory burden of the reconstruction. Furthermore, the reconstruction is no longer limited to the size of voxels used to discretize the image space.

Note that solving for source intensities and source positions are independently convex problems, but solving for them simultaneously in Eq. 14 is no longer convex. Several methods exist to overcome this issue, but for this work a traditional convex optimization algorithm was used along the optimal ML-EM intensity (and background) manifold. In other words, the optimization algorithm was allowed to search only the positions of the point sources and at each step (source position) ML-EM was performed to find the optimal source intensity and constant background.

The BOBYQA derivative-free algorithm [27] for bound constrained optimization in the NLopt [28] package within the parallel optimization Python library PYGMO [29] was used here with 20 ML-EM iterations at each step. Several other algorithms (both derivative-free and gradient-based) are available, but BOBYQA produced the best results with the fastest convergence in this problem. The results are shown in Fig. 6. The point-source position is no longer limited by the voxel size, producing a position error $< 2$ cm, and the error in the reconstructed point-source intensity was $< 1\%$ of the true intensity.

The reconstruction runtime was 160 ms, considerably faster than the previous approaches. This time also includes the calculation of the system matrix, but in this case the response was only computed for $\sim$50 positions, equivalent to $\sim$50 columns of $\mathbf{V}$. This is compared to the $J = 1.28 \times 10^5$ columns for the voxelized space in the previous section. The 50 columns in this case indicate the number of positions tested in the BOBYQA optimization routine before convergence. Note that as the number of measurements and voxels increases, the full calculation of $\mathbf{V}$ becomes significantly expensive and memory intensive (in some cases, requiring more memory than available RAM, thus further slowing the reconstruction as $\mathbf{V}$

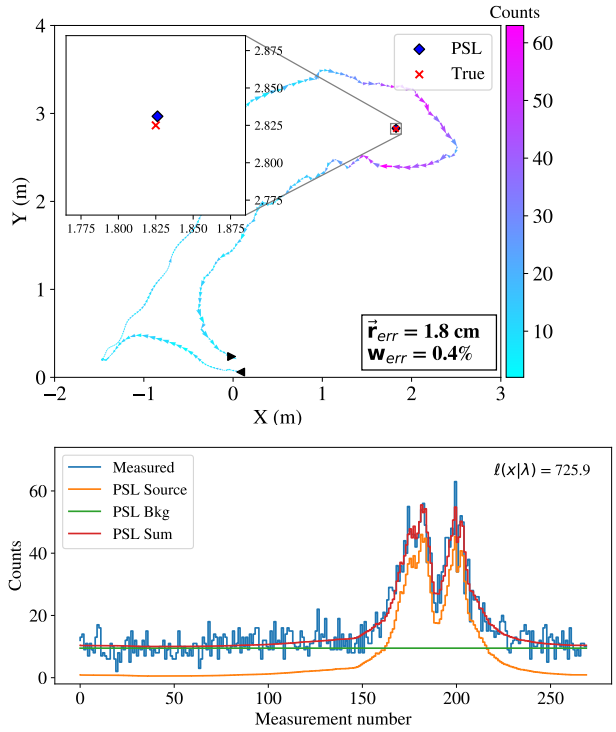


Fig. 6. (Color online) (Top) PSL solution (blue diamond) of a single point-source to best describe the measured data, solved for in continuous 3D space along the maximum likelihood source intensity and background manifold using a conventional derivative-free optimization algorithm (BOBYQA). (Bottom) Forward projected mean-rates $\boldsymbol{\lambda}$ compared to the measurement $\mathbf{x}$.

must be computed on the fly), further limiting the runtime performance of spatially-discretized approaches. The continuous PSL formulation avoids this problem entirely, scaling linearly with the number of measurements.

To explore the performance of the continuous PSL approach in this source scenario, the simulation was run 10,000 times with source activities randomly sampled between 1-10 $\mu$Ci and a constant background rate of 100 counts per second. Figure 7 shows a box plot of the position and source intensity errors against the true source strength, binned at 1 $\mu$Ci intervals. For each interval, the mean signal-to-noise-ratio ($\overline{\text{SNR}}$) is shown as an additional horizontal axis. The SNR of a single simulation is defined as $\max(S_i/\sqrt{S_i + B_i})$ where $S_i$ and $B_i$ are the signal and background at measurement $i$, respectively. The position errors are expressed as a percent of closest approach (45 cm). The median in each bin is denoted with an orange horizontal line, the boxes range from the first to third quartile, and the arms extend out to the 10$^{\text{th}}$ and 90$^{\text{th}}$ percentiles. Both the median and spread of each error decreases with source strength, as expected. The median source intensity error is below 10% across all activities and the median position error is less than 40% of the distance of closest approach.

Figure 8 shows the correlation between the position and source intensity errors for $>$ 99% of the total $10^4$ simulations (outlier reconstructions are ignored). Recall the detector position remained close to Z $\approx -14$ cm while near the source, which was placed at Z $= 0$. The isotropic response of the detector and the little movement in Z results in degenerate solutions in likelihood that extend along Z. Furthermore, the

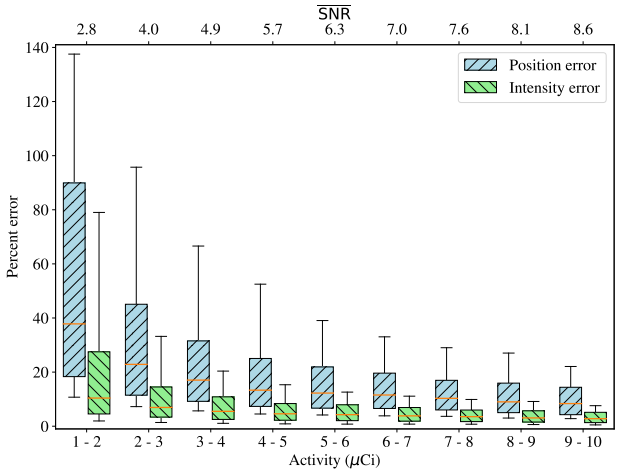


Fig. 7. (Color online) Box plot of the position and source intensity errors against the true source strength, binned at each $\mu$Ci, for $10^4$ simulations with source intensities randomly sample between 1-10 $\mu$Ci. Position error as defined in Sec. II-B is shown here in percent of closet approach. The median in each bin is denoted with an orange line, the boxes range from the first to third quartile, and the arms extend out to the 10$^{\text{th}}$ and 90$^{\text{th}}$ percentiles. The mean SNR over each interval is shown as an additional horizontal axis.

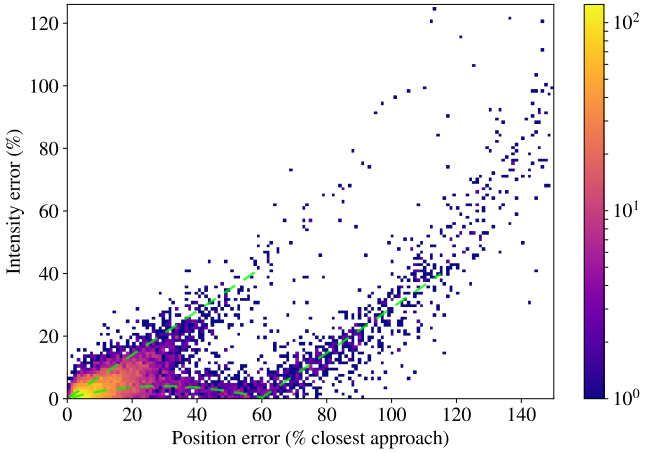


Fig. 8. (Color online) 2D histogram of position and source intensity errors for $>$ 99% of the total $10^4$ simulations with source activities between 1-10 $\mu$Ci. Less than 1% of events were outside of the defined bounds and are not shown. Bin widths are 1% in each direction. The degeneracy of the solution space is highlighted by green dashed lines and a discussion in provided in the text.

position below the measurement path equal to the offset of the path to the true source location (Z $\approx -28$ cm or 60% position error) represents a degenerate solution in source intensity. These effects can be seen in the plot in three distinct regions (shown with dashed green lines). First, a diagonal band extends out from the origin, corresponding to reconstructed positions above the true location in Z. A second diagonal band extends out from a position error of 60%, corresponding to reconstructed positions below the degenerate source intensity solution in Z. Finally, a slightly curved band is observed between the origin and this point, representing the positions between the true source location and the degenerate source intensity solution on the opposite side of the path. The band peaks at a position error of 30% (plane of the measurement path, Z $\approx -14$ cm), where the source intensity estimate is lower to compensate for the closer source. Ultimately, while

some degeneracy is observed in the solution space, enough variation existed in the detector position that the majority of reconstructions were near the origin in Fig. 8, correctly localizing and quantifying the source, particularly for high SNR sources.

## IV. Additive Point-Source Localization

The additive nature of Poisson variables facilitates the inclusion of $M$ known, constant source rate contributions, $\boldsymbol{\mu}^{[I\times 1]}$ into Eq. 15, rewritten as

$$\lambda_i = v_{is}\mathrm{w}_s + \mu_i + bt_i\,, \tag{17}$$

where $\mu_i = \sum_{m=1}^{M} v_{im}\mathrm{w}_m$. Now Eq. 14 can be reformulated to localize an additional source

$$\underset{(\mathrm{w}_s,\vec{r_s},b)}{\operatorname{argmin}}\ \ell(\mathbf{x}|\mathrm{w}_s,\vec{r_s},b,\boldsymbol{\mu})\,. \tag{18}$$

While one can attempt to solve Eq. 18 in discrete space as in Sec. III-A by successively solving for sources one at a time, this approach does not allow for the re-optimization of source intensities and positions after a new source is found. In the continuous space formulation, a re-optimization can be done in position and intensity across multiple sources, allowing all the free parameters to vary at once. This capability is crucial in the multi-point-source reconstruction problem as the addition of another source can affect the overall likelihood fit of the previous source configuration.

Algorithm 1 is proposed here to iteratively reconstruct a sparse parametric image of $N$ sources, $\mathbf{S} = \{(\mathrm{w}_1,\vec{r_1}),...,(\mathrm{w}_N,\vec{r_N});b\}$, in the continuous space formulation, where $N$ is also treated as an unknown. After each new source is identified, the re-optimization of source positions and intensities can be done in two ways:

1) In the fashion of [30], alternate between
   a) Fix intensities and re-optimize positions using conventional optimization methods (e.g., BOBYQA).
   b) Fix positions and re-optimize intensities and backgrounds using ML-EM.
2) Re-optimize source positions using conventional optimization methods along the optimal intensity and background manifold (as done in Sec. III-B).

A model selection criterion is then used to test the new model ($N$+1 sources) compared to the old model ($N$ sources). The Bayesian Information Criterion (BIC) [31] was used here and is given by

$$\mathrm{BIC} = \log(I)k + 2\ell(\mathbf{x}|\hat{\boldsymbol{\lambda}})\,, \tag{19}$$

where $k = 4N + 1$ is the number of parameters estimated by the model (the intensity and XYZ position of each source and the constant background rate in the detector) and $I$ is the number of measurements. Again, for a multi-detector system $k$ would generalize to $k = 4N + N_\mathrm{d}$. The preferred model is the one with the minimum BIC value. The BIC penalizes the acceptance of a new model based on the number of parameters used in the model (i.e., the addition of another source must significantly improve the model in order to be accepted). The BIC was used here, as opposed to another model selection criterion such as the Akaike Information Criterion (AIC) [32], as it applies a stricter penalty, therefore restricting the addition of many sources and reducing overfitting.

**Algorithm 1** Additive Point-Source Localization

```
Initialize reconstruction. S = {(); b = median(x)}
converged = False
Solve Eq. 14, append to S
while not converged do
    S_old = S
    Solve Eq. 18 for additional source, append to S
    Re-optimize source positions, intensities and background, update S
    Test for acceptance of S relative to S_old using BIC
    if accepted then
        Clean S: drop low weight or weakly contributing sources and
        collapse nearby sources
        Re-optimize current state of source positions, intensities and
        background, update S
    else
        S = S_old
        converged = True
```

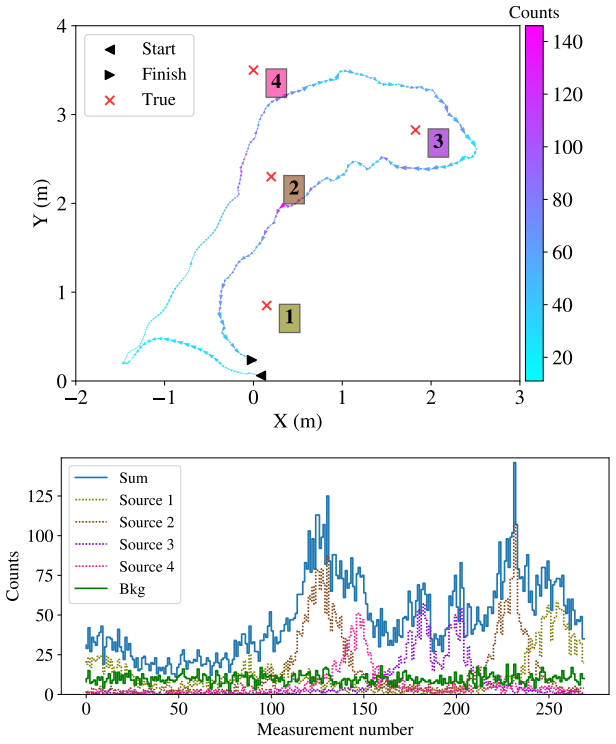


Fig. 9. (Color online) (Top) Measurement scenario similar to Fig. 1 now with multiple sources in the XY plane. The source positions are shown with red x's as well as numbers in colored boxes. (Bottom) Simulated signal source components and background at each measurement.

If the new model is accepted, a cleaning procedure is invoked to drop low intensity and weakly contributing sources as well as combine spatially close sources (defined to be $< 10$ cm in this case). If the source configuration is changed in the cleaning, a re-optimization is performed again. If the new model is rejected, the algorithm is stopped. We note the APSL algorithm correctly stopped at one source in the source scenario presented above.

The multi-source performance of the APSL algorithm is explored using the simulated detector and measurement path from above (see Fig. 9). In addition to the 5 $\mu$Ci source used previously, three additional sources were placed in the

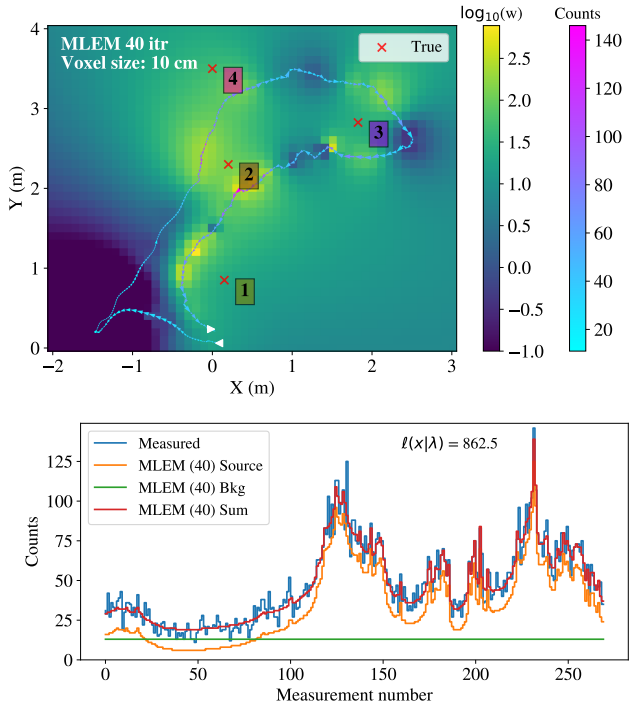


Fig. 10. (Color online) (Top) ML-EM reconstruction $\hat{\mathbf{w}}$ (40 iterations), in $\log_{10}$ scale, of a slice through the XY plane (Z = 0) for the scenario presented in Fig. 9. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{\lambda}}$ compared to the measurement $\mathbf{x}$.

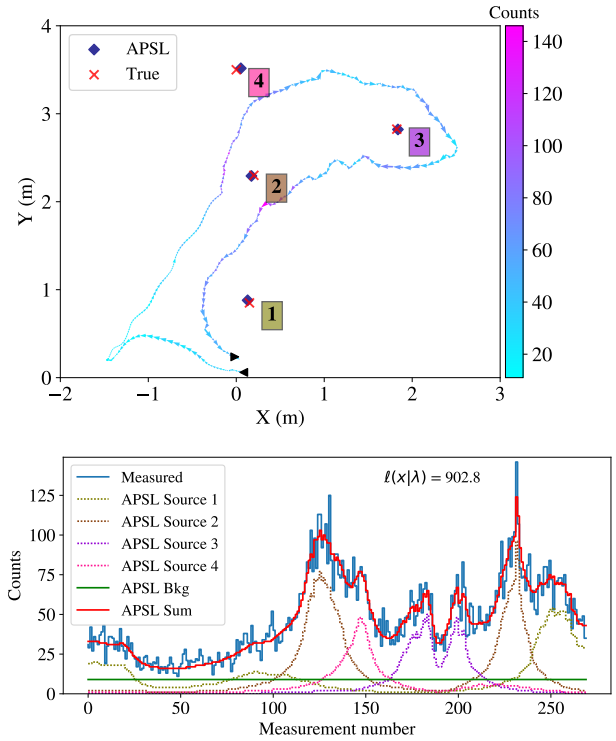


Fig. 11. (Color online) (Top) APSL solution (blue diamonds) in the XY plane (Z = 0) for the scenario presented in Fig. 9. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{\lambda}}$ (individual source and background components) compared to the measurement $\mathbf{x}$.

XY plane near the trajectory, with activities ranging from 6-8 $\mu$Ci and closest standoffs of 35-55 cm. The bottom pane of of Fig. 9 shows the simulated signal source components and background at each measurement.

The unregularized ML-EM reconstruction, in $\log_{10}$ space, along the XY plane and the corresponding fit in count space are shown in Fig. 10. Forty iterations were required in this case to achieve the same likelihood convergence stated in Sec. II-B. The system matrix calculation runtime was 22 s and the reconstruction runtime was 1.2 s. ML-EM continues to produces an overfit solution and fails to correctly localize any of the sources. Again, the overfitting behavior worsens with more iterations. Similar to Sec. II-C, the inclusion of sparsity regularizers in ML-EM did not improve the results.

The APSL reconstruction results are shown in Fig. 11 and the errors are shown in Table I. The BOBYQA algorithm was used with 20 ML-EM iterations at each step in the position optimization (additional iterations were not necessary and did not affect the result). The total reconstruction runtime was 12.8 s. Again, this time includes the calculation of the columns in $\mathbf{V}$ needed in the optimization. APSL converges to the correct number, location, and intensity of the four unique sources, without the use of pre-conditioning as in [5], and accurately estimates the constant background rate.

The reconstruction errors in position are all $< 10$ cm in the XY plane, with slightly higher errors in the Z dimension. Larger errors are observed for the sources outside of the measurement path (#1 and #4). This is expected due to the limited sensitivity to these regions and the degeneracy in the solution space discussed previously. In general, it is shown that

TABLE I
APSL RECONSTRUCTION ERRORS IN FIG. 11.

| | Source 1 | Source 2 | Source 3 | Source 4 |
|---|---|---|---|---|
| $\mathbf{w}_{\text{true}}$ ($\mu$Ci) | 8.0 | 6.0 | 5.0 | 6.0 |
| $\hat{\mathbf{w}}$ ($\mu$Ci / % err) | 7.3 / 9.9 | 5.6 / 7.2 | 4.8 / 3.9 | 7.7 / 22.3 |
| $\hat{\bar{\boldsymbol{r}}}_{\text{err, XY}}$ (cm) | 3.6 | 2.9 | 1.3 | 5.5 |
| $\hat{\bar{\boldsymbol{r}}}_{\text{err, Z}}$ (cm)$^{\dagger}$ | 10.5 | 5.8 | 5.4 | 15.4 |

$^{\dagger}$ Generally, $\hat{\bar{\boldsymbol{r}}}_{\text{err, Z}} > \hat{\bar{\boldsymbol{r}}}_{\text{err, XY}}$ due to the non-directional detector used and that the path primarily moved in XY, without much change in Z.

larger position errors correlate with larger intensity errors, as was shown in Fig. 8.

While this example uses a simple model with nearby sources and relatively large individual SNRs, several peaks overlap to produce an overall complex source term. Furthermore, the measured detector path captures the variation in detector position we would expect in an operational source search scenario, adding variation to the signal as well as breaking some of the degeneracy in the ML solution space. Ultimately, the results highlight the success of the APSL algorithm in the deconvolution of the signal into the correct individual source components and background.

### A. *Separation of Spatially Close Sources*

The ability to separate sources with an isotropic detector is dependent on the measurement path, the statistics collected, and the degeneracy of the ML solution space. However, it is expected that, given a path near two sources, an isotropic detector should be capable of resolving two sources (in space

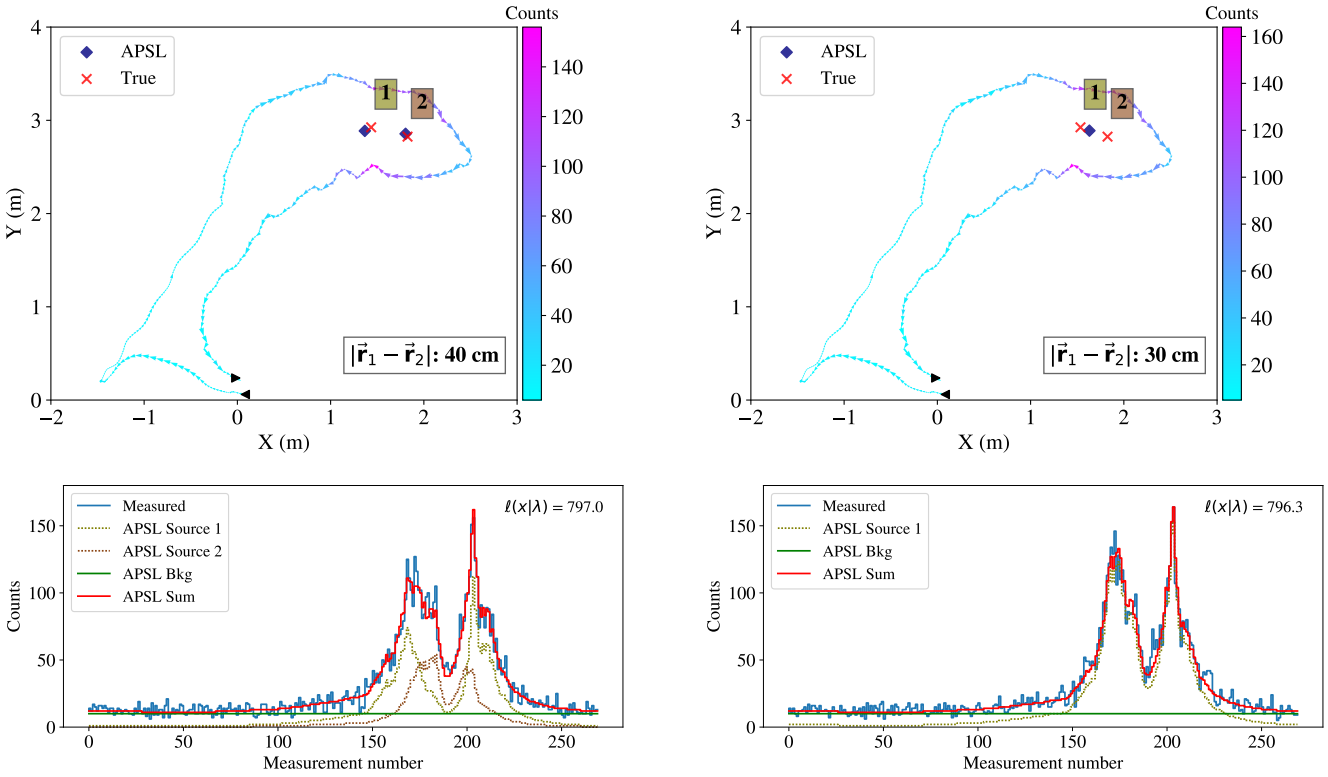


Fig. 12. (Color online) (Top) APSL solution (blue diamonds) in the XY plane (Z = 0) for two sources (#1: 9 $\mu$Ci, #2: 5 $\mu$Ci) with separation of 40 cm (left) and 30 cm (right) placed in the XY plane. The distance of closest approach in the measurement path is 40 cm in both cases. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{\lambda}}$ (individual source and background components) compared to the measurement $\mathbf{x}$ for the images in the top panel. Reconstruction errors are discussed in the text.

and intensity) that are spatially separated by at least the distance of closest approach.

To demonstrate the resolving capability of APSL when using an isotropic detector, two unequal activity sources (9 $\mu$Ci and 5 $\mu$Ci) are simulated in the XY plane near the measurement path, with a 40 cm distance to closest approach. We consider two scenarios in which the source separations are 40 cm and 30 cm. APSL is run on each scenario and the results are shown in Figure 12.

In the case where the source separation is the distance of closest approach, APSL is successful in resolving two sources, with XY position errors $<$ 10 cm, Z position errors $<$ 15 cm, and intensity errors $<$ 5%. In the case where the source separation is less than the distance of closest approach, APSL places a single source with a combined activity (13.8 $\mu$Ci) between the two sources. A slight bias is observed towards the stronger source (#1), though this is expected as this source contributes more to the measured signal and a better fit to this source can result in a lower overall loss.

## V. Summary

Using a simple simulation of a single point-source and a non-directional detector moved over an experimental SLAM-tracked trajectory, we have shown that a reformulation of the sparse inverse problem can improve source localization and quantification performance. In sparse scenarios, the traditional and sparsity-enforcing regularized ML-EM approaches suffer from overfitting and fail to localize point-sources. Ultimately, these approaches are attempting to solve an inherently underdetermined problem. The reformulation of the problem with a point-source assumption (PSL) is well-posed and was shown to accurately localize point sources and quantify their intensity, though at the expense of computational complexity when reconstructing in a discretized spatial domain. Solving PSL as an optimization over continuous intensity and position increased localization and quantification performance and improved computational efficiency and memory consumption.

The continuous PSL formulation was extended and generalized to APSL for multiple point-source localization or sparse parametric image reconstruction. To demonstrate the concept, four point-sources were simulated near the same experimental measurement path, forming a complex source term. ML-EM failed to localize any of the sources or provide a useful initialization for a gradient-based approach, even with sparsity regularization. APSL successfully deconvolved the signal into the correct individual source and background components (including the number, location and strength of the sources) at a reduced computational and memory burden.

Finally, APSL successfully resolved (in both space and intensity) two sources of unequal activity in close proximity. The resolving capability drops quickly near the distance of closest approach, but detectors or detector arrays that are better able to determine the direction of photon incidence are expected to better resolve closer sources.

## VI. Future Work

In this manuscript we have presented the APSL approach for sparse parametric image reconstruction, using simulated data for demonstration purposes. Further work is needed to rigorously characterize the algorithm and validate the scenarios in which it may be used. We will pursue this work with experimental data, evaluating the performance, limitations, and scalability of the approach. This effort will include investigations into complex source terms such as weak, large standoff, and shielded sources as well as more constrained measurement paths relevant to radiological source search (e.g., straight lines or single pass) that may be subject to more degeneracy in the solution space. The impact of imaging detector systems and/or arrays (i.e., anisotropic detectors) on localization accuracy and computational burden will also be explored.

While a constant background assumption was made here and may be appropriate for the small search space and short measurement duration, this may not be appropriate in wide-area urban search scenarios [33], [34], [35]. Therefore a treatment for variable background rates must be incorporated into the algorithm for larger search spaces and longer measurements. For example, advanced radiological detection and identification algorithms designed for dynamic background environments [21] could be used to model spatially or temporally-varying background counts in the measured signal. The counts attributed only to an identified source (or sources) could then be used as an input to the APSL framework, mitigating the need for a varying background term. This approach would also facilitate the localization of specific radioisotopes.

Finally, relevant for field operations, APSL will be developed for streaming operation, informing the user of potential source locations in real-time and directing the path to decrease the time to detect and localize sources. With the addition of contextual sensors, the image space can be constrained to static surfaces and volumes [14], [36], [37] or to moving objects in the 3D scene model, improving the computational speed of the reconstruction as well as providing additional context to the user in real-time.

## References


[1] H. Durrant-Whyte and T. Bailey, "Simultaneous Localization and Mapping: Part I," *IEEE Robot. Autom.*, vol. 13, no. 2, 2006.

[2] ——, "Simultaneous Localization and Mapping: Part II," *IEEE Robot. Autom.*, vol. 13, no. 3, 2006.

[3] N. S. V. Rao *et al.*, "Identification of Low-Level Point Radiation Sources Using a Sensor Network," *in Proc. Int Conf. Inf. Process. Sens. Networks*, pp. 493–504, 2008.

[4] B. Deb, J. A. F. Ross, A. Ivan, and M. J. Hartman, "Radioactive Source Estimation Using a System of Directional and Non-directional Detectors," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 6, pp. 3281–3290, 2011.

[5] B. Deb, "Iterative Estimation of Location and Trajectory of Radioactive Sources with a Networked System of Detectors," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 2, pp. 1315–1326, 2013.

[6] U. Bhattacharyya and C. Baum, "Estimating the Location of a Nuclear Source in a Three-Dimensional Environment Using a Two-stage Adaptive Algorithm," *in Proc. of IEEE Comput. Comm. Workshop Conf.*, 2018.

[7] E. A. Miller *et al.*, "Adaptively Reevaluated Bayesian Localization (ARBL): A Novel Technique for Radiological Source Localization," *Nucl. Instrum. Meth. A*, vol. 785, pp. 332–338, 2015.

[8] B. Ristic, M. Morelande, and A. Gunatilaka, "Information Driven Search for Point Sources of Gamma Radiation," *Signal Process.*, vol. 90, no. 4, pp. 1225–1239, 2010.

[9] J. Chin, D. K. Yau, N. S. V. Rao, Y. Yang, C. Y. T. Ma, and M. Shankar, "Accurate Localization of Low-level Radioactive Source Under Noise and Measurement Errors," *in Proc. ACM Conf. Embed. Network Sens. Sys.*, pp. 183–196, 2008.

[10] R. Vilim and R. Klann, "RadTrac: A System for Detecting, Localizing, and Tracking Radioactive Sources in Real Time," *Nucl. Tech.*, vol. 168, no. 1, pp. 61–73, 2009.

[11] M. K. Sharma, A. B. Alajo, and H. K. Lee, "Three-dimensional Localization of Low Activity Gamma-ray Sources in Real-time Scenarios," *Nucl. Instrum. Meth. A*, vol. 813, no. 21, pp. 132–138, 2016.

[12] G. Cordone, R. R. Brooks, S. Sen, N. S. V. Rao, C. Q. Wu, M. L. Berry, and K. M. Grieme, "Improved Multi-Resolution Method for MLE-based Localization of Radiation Sources," *in Proc. Int. Conf. Inf. Fusion*, 2017.

[13] C. G. Wahl *et al.*, "The Polaris-H Imaging Spectrometer," *Nucl. Instrum. Meth. A*, vol. 784, pp. 377–381, 2015.

[14] A. Haefner, R. Barnowski, P. N. Luke, M. Amman, and K. Vetter, "Handheld Real-time Volumetric 3-D Gamma-ray Imaging," *Nucl. Instrum. Meth. A*, vol. 857, pp. 42–49, 2017.

[15] J. C. Curtis *et al.*, "Simulation and Validation of the Mobile Urban Radiation Search (MURS) Gamma-ray Detector Response," *Nucl. Instrum. Meth. A*, 2018.

[16] T. H. Y. Joshi *et al.*, "Measurement of the Energy Dependent Angular Response of the ARES Detector System and Application to Aerial Imaging," *IEEE Trans. Nucl. Sci.*, vol. 64, no. 7, pp. 1754–1760, 2017.

[17] R. Pavlovsky *et al.*, "3-D Radiation Mapping in Real-Time with the Localization and Mapping Platform LAMP from Unmanned Aerial Systems and Man-Portable Configurations," *arXiv:1901.05038*, 2018.

[18] D. J. Lingenfelter, J. A. Fessler, and Z. He, "Sparsity Regularization for Image Reconstruction with Poisson Data," *in Proc. SPIE*, vol. 7246, 2009.

[19] Z. Harmany, R. F. Marcia, and R. M. Willet, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms - Theory and Practice," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1084–1096, 2012.

[20] X. Zhang, Y. Lu, and T. Chan, "A Novel Sparsity Reconstruction Method from Poisson Data for 3D Bioluminescence Tomography," *J. Sci. Comput.*, vol. 50, no. 3, pp. 519–535, 2012.

[21] K. J. Bilton *et al.*, "Non-negative Matrix Factorization of Gamma-ray Spectra for Background Modeling, Detection, and Source Identification," *IEEE Trans. Nucl. Sci.*, vol. 68, no. 5, pp. 827–837, 2019.

[22] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. R. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, 1977.

[23] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-Time Loop Closure in 2D LIDAR SLAM," *in Proc. IEEE Int. Conf. Robot. Autom.*, pp. 1271–1278, 2016.

[24] K. Lange, M. Bahn, and R. Little, "A Theoretical Study of Some Maximum Likelihood Algorithms for Emission and Transmission Tomography," *IEEE Trans. Med. Imaging*, vol. 6, no. 2, pp. 106–114, 1987.

[25] C. Byrne, "Likelihood Maxmimuzation for List-Mode Emission Tomographic Image Reconstruction," *IEEE Trans. Med. Imaging*, vol. 20, no. 10, pp. 1084–1092, 2001.

[26] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Ann. Math. Stat.*, vol. 9, no. 1, pp. 60–62, 1938.

[27] M. J. D. Powell, "The BOBYQA Algorithm for Bound Constrained Optimization Without Derivatives," *Department of Applied Mathematics and Theoretical Physics, Cambridge University Technical Report NA2009/06*, 2009.

[28] S. G. Johnson, "The NLopt Nonlinear-Optimization Package," 2008. [Online]. Available: http://ab-initio.mit.edu/nlopt

[29] F. Biscani *et al.*, "Esa/Pagmo2: Pagmo 2.10," 2019. [Online]. Available: https://doi.org/10.5281/zenodo.2529931

[30] N. Boyd, G. Schiebinger, and B. Recht, "The Alternating Descent Conditional Gradient Method for Sparse Inverse Problems," *SIAM J. Optim.*, vol. 27, no. 2, pp. 616–639, 2017.

[31] G. Schwarz, "Estimating the Dimension of a Model," *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, 1974.

[32] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.

[33] T. J. Aucott, M. S. Bandstra, V. Negut, D. H. Chivers, R. J. Cooper, and K. Vetter, "Routine Surveys for Gamma-Ray Background Characterization," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 2, pp. 1147–1150, 2013.

[34] R. D. Penny, "Improved Radiological/Nuclear Source Localization in Variable NORM Background: An MLEM Approach with Segmentation Data," *Nucl. Instrum. Meth. A*, vol. 784, pp. 319–325, 2015.

[35] E. Bai, A. Heifetz, P. Raptis, S. Dasgupta, and R. Mudumbai, "Maximum Likelihood Localization of Radioactive Sources Against a Highly Fluctuating Background," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 6, pp. 3274–3281, 2015.

[36] R. Barnowski, A. Haefner, L. Mihailescu, and K. Vetter, "Scene Data Fusion - Real-time Standoff Volumetric Gamma-ray Imaging," *Nucl. Instrum. Meth. A*, vol. 800, pp. 65–69, 2015.

[37] D. Hellfeld, P. Barton, D. Gunter, A. Haefner, L. Mihailescu, and K. Vetter, "Omnidirectional 3D Gamma-ray Imaging with a Free-moving Spherical Active Coded Aperture," *in Proc. IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2017.