Iterative Estimation of Location and Trajectory of Radioactive Sources With a Networked System of Detectors

Budhaditya Deb

***Abstract*—We consider the problem of estimating the parameters (location and intensity) of multiple radioactive sources using a system of radiation detectors. The problem formulated as maximum likelihood estimation (MLE) requires the optimization of a high-dimensional objective function and presents significant computational challenges. We propose Fisher's scoring iterations approach (a special case of Newton's iterative method) for finding the MLE. While being computationally scalable, an inherent problem with this approach is finding good initial estimates specifically when multiple sources are present. We propose an expectation maximization (EM) based approach which finds the approximate distribution of the source intensity in space. Peaks in this distribution are used as initial estimates of the parameters to bootstrap the iterative MLE procedure. Next, we consider the problem of estimating the trajectory of a moving and maneuvering source. Since a priori motion model cannot be assumed, the trajectory is approximated as a set of points which again presents a high dimensional estimation problem. The trajectory estimation is posed as a constrained weighted least squares problem which is iteratively solved using the Interior Point Method (IPM). Simulation results are presented which illustrate the behavior and performance of our proposed approaches.**

***Index Terms*—Estimation, expectation-maximization algorithms, iterative methods, maximum likelihood estimation, radiation imaging, signal analysis-motion estimation.**

I. Introduction

DETECTION of contraband radioactive sources using a spatially dispersed network of detectors has recently emerged as an active topic of research. Detectors register gamma photons at a rate based on the intensity and the distance between the source and the detector. By fusing information from spatially diverse detectors, source parameters can be estimated. While it is debatable whether a system of low efficiency network of detectors is better suited to detection than fewer detectors with high efficiency [1], [2], spatial diversity of dispersed sensor resources clearly benefits localization.

Manuscript received February 09, 2012; revised August 10, 2012 and January 04, 2013; accepted February 04, 2013. Date of publication March 22, 2013; date of current version April 10, 2013. This work was supported in part by the U.S. Department of Homeland Security, Grant No HSHQDC-09-C-00125.

The author was with the Advanced Communication Systems Lab, GE Global Research, Niskayuna, NY 12309 USA. He is currently with the Advanced Networking group at Raytheon-BBN Technologies, Cambridge, MA 02138 USA (e-mail: bdeb@bbn.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2013.2247060

A network of detectors introduces new algorithmic challenges on multi-sensor information fusion and source estimation. Approaches involve trilateration [3], [4], weighted least squares (WLS) [5], and maximum likelihood estimation (MLE) [6]–[8]. WLS and MLE involve optimization of a high dimensional objective function for which closed form solution does not exist for the source location problem. Prior work which uses grid based searches to find the MLE [6], [7] show them to be computationally infeasible for more than two sources for typical computing environment of a laptop or workstation (single core or single process thread). Iterative approaches have been used for tracking a source once the initial location of a source is known [5] but it is not clear how one may bootstrap an iterative algorithm for multiple unknown sources using a similar approach.

Bayesian approaches have also been proposed [9]–[12] which estimate parameters by computing the expectation of the posterior distribution of the parameters. Since, a closed form for the expectation does not exist for the radiation location problem, a variety of numerical techniques involving importance sampling and particle filters have been used for approximating the posterior distribution and its expectation. Bayesian estimation in theory can guarantee minimum mean square error in estimates but the performance depends on the accuracy of the sampling process. When informative prior distribution about the parameters is not available (importance sampling typically uses the prior as the sampling distribution), the number of samples required can exhibit complexity exponential in the number of parameters being estimated. To reduce the sample size, [9] and [10] employ an iterative method called progressive correction which reduces sample size at the cost of increased computation cost through iterative correction and re sampling steps. In the absence of informative priors, it is debatable whether Bayesian estimation has benefits over MLE when computation complexity is taken into consideration.

The algorithmic challenge is how to deal with the *curse of dimensionality* of the parameter space to be explored. Existing literature has typically looked at sensor systems with tens of detectors which have been used to locate around 1 to 4 sources simultaneously (e.g., Fig. 1). Even at this scale the estimation problem involves significant computational challenge for real time systems. This paper aims at providing computationally scalable algorithms for ML estimation in scenarios where prior work may prove to be computationally infeasible. The approaches may also be considered as low complexity alternatives when informative priors for Bayesian estimation are not available.

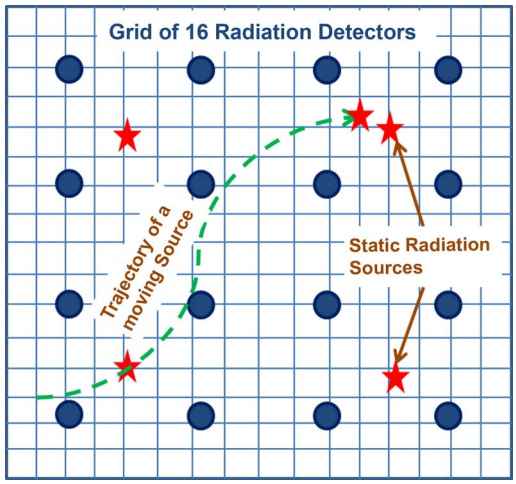

Fig. 1. Scenario with sensor grid and multiple source.

First, we consider the problem of estimating the parameters (location and intensity) of *static* radio-active sources. Our approach is based on the gradient based Fischer's scoring iterations (a special case of Newton's iterative method) to compute the MLE. While gradient based schemes have low complexity, they may diverge or converge to local maxima especially if the initial estimates to start the iterations are not sufficiently close to the global maxima. Initialization with good starting points is a challenge specifically with multiple sources. We propose an expectation maximization (EM) based approach which approximates the activity distribution of the monitored region, the peaks of which represent the number and possible locations of the source. The initialization through EM is an important contribution in this paper since it makes the gradient based scheme robust for multiple sources. The EM approach can itself be considered as a low complexity source estimation method.

Second, we consider the problem of estimating the trajectory of a moving and maneuvering source through the monitored field. For a maneuvering target, since we cannot associate a parametric motion model (e.g., initial position, speed, and direction for straight line paths assumed in [1], [10], and [12]), the trajectory is represented as a point-wise approximation. The problem is then of estimating multiple points which best define the trajectory given the observations. We formulate it as a constrained weighted least squares optimization problem and propose an iterative solution using the Interior Point Method (IPM). IPM reduces the constrained optimization to an unconstrained one which can then be optimized using the standard Newton's iterative method.

II. MLE for Locating Multiple Sources

Fig. 1 illustrates the scenario under consideration. N sensors, located at positions $\{x_n, y_n\}$ receive counts C_n^t from multiple sources during time interval t in the photo-peak region of a specific isotope under consideration. The detectors are assumed to have isotropic response and do not provide any directional information when placed far from each other. For illustration and simulations we consider the Cs-137 photo-peak region of 662-keV. The source location and the intensity for M sources are modeled by an unknown parameter vector $\theta = \{X_1^S, Y_1^S, \lambda_1^S, \ldots, X_{\mathrm{M}}^S, Y_{\mathrm{M}}^S, \lambda_{\mathrm{M}}^S\}$. The background is assumed to be spatially uniform with rate λ_B in the photo-peak region considered. Given the detector photo-peak efficiency μ (which also includes the branching factor for the photo-peaks of different isotopes), the air attenuation coefficient ρ, detector cross section area A, and a geometric efficiency described by the inverse-square distance law, the incident count-rate from the mth source at the nth detector is given by(1) and total count-rate λ_n in (2):

$$\lambda_n^m = \frac{\mu A \lambda_m^S}{4\pi d_{n,m}^2} e^{-\rho d_{n,m}} \tag{1}$$

$$\lambda_n = \lambda_B + \sum_{m=1}^{M} \lambda_n^m \tag{2}$$

where $d_{n,m}^2 = (X_m^S - x_n)^2 + (Y_m^S - y_n)^2$.

The likelihood of observing C_n^t counts given the parameters θ is given by (3), the joint log- likelihood for N sensors by (4) and the ML estimate by (5):

$$P\left(C_n^t | \theta_1, \ldots, \theta_M\right) = \frac{e^{-\lambda_n t}(\lambda_n t)^{C_n^t}}{C_n^t!} \tag{3}$$

$$L(\theta_1, \ldots, \theta_M | C_1^t, \ldots C_N^t) = ln \prod_{n=1}^{N} P\left(C_n^t | \theta_1, \ldots, \theta_M\right) \tag{4}$$

$$\hat{\theta} = Argmax[L(\theta_1, \ldots, \theta_M | C_1^t, \ldots C_N^t)]. \tag{5}$$

We can extend the above formulation for different detector sensitivities and sensitivity based on angular orientation of sensors. Instead of fixed values, the factors would depend on the detector number and location of a point on the grid with respect to a detector. Other known attenuation factors in the monitored region can also be incorporated without changing the estimation problem. However, if some of the model complexity resides in the estimation parameters, for example with non-isotropic source and non-uniform shielding, then our proposed model would not be able to account for them. These are subjects for future work.

Since the optimization of the likelihood function cannot be solved in closed form, we solve it numerically. For a single source a grid search may be used to find the MLE. We divide the region into a grid (e.g., 1 m grid points in a $100 \times 100\ \mathrm{m}^2$ region for location and 0 to 10 mCi in 0.1 mCi steps for source intensity) and compute the joint-likelihood value for each of these points. Then we simply find the point with the maximum value. For multiple sources it is intractable to find optimal parameters using the grid search. For example to search for two sources with similar precision, we would have to search on a grid with 10^{10} points. Enhancements on the basic grid search may be also employed. For example, MATLAB's *fminsearch* (which uses the simplex method) is used in [6] for optimization. While better than a fixed grid search, it does not scale with the number of parameters and was shown to work only for up to two sources in [6].

In cases where an analytical solution does not exist, iterative numerical methods are used for optimizing a nonlinear function. We consider Newton's method for optimization which can provide extremely accurate estimation of the parameters with low

complexity and fast convergence. A drawback with Newton's method is that a good initial estimate of the parameters is required for convergence to global maxima. In fact, inability to find consistent start points may render the method ineffective. We first present the iterative procedure and then present our approach to find the initial estimate.

A. Newton's Iterative Method for Finding the MLE

Newton's method for optimizing the joint log-likelihood function is given by the iterative steps in (6):

$$\left.\begin{array}{l}\Delta\theta = -\left[\frac{\partial^2 L(\theta|C^t)}{\partial\theta^2}\right]^{-1}\frac{\partial L(\theta|C^t)}{\partial\theta}\Big|_{\theta=\theta(k)} \\ \theta(k+1) = \theta(k) + \delta^*\Delta\theta \\ where,\ \delta^*\ is\ a\ multiplier\ found\ through\ line\ search.\end{array}\right\}. \tag{6}$$

Above, k denotes the iteration number. To find the optimal multiplier δ^* via line search, we compute the likelihood values with $0 < \delta \leq 1$ at predefined intervals of 0.1 and select δ with the highest likelihood value. The above formulation is written in vector form in (7) where $\theta_i\{i = 1 \ldots 3M\} \equiv \{X_1^S, Y_1^S, \lambda_1^S, \ldots, X_{\mathrm{M}}^S, Y_{\mathrm{M}}^S, \lambda_{\mathrm{M}}^S\}$ denotes the parameter vector:

$$\theta(k+1) = \theta(k) + \delta^* H(\theta)^{-1} G(\theta)\Big|_{\theta=\theta(k)}. \tag{7}$$

Above, H- is the $3M \times 3M$ Hessian matrix of 2nd order derivatives and G is the $3M \times 1$ matrix of 1st order derivatives. The derivatives are computed using the current parameter estimates and the observations. For the joint log-likelihood function of (4), the 1st and 2nd order derivatives are given by (8) and (9):

$$\begin{aligned} G(\theta_i) &= \sum_{n=1}^{N} \frac{\partial(C_n^t ln(\lambda_n t)}{\partial\theta_i} - \frac{\partial\lambda_n t}{\partial\theta_i} \\ &= \sum_{n=1}^{N} \frac{\partial\lambda_n}{\partial\theta_i}\left(\frac{C_n^t}{\lambda_n} - t\right) \end{aligned} \tag{8}$$

$$H(\theta_i, \theta_j) = \sum_{n=1}^{N} \frac{\partial^2\lambda_n}{\partial\theta_i\partial\theta_i}\left(\frac{C_n^t}{\lambda_n} - t\right) - \frac{\partial\lambda_n}{\partial\theta_i}\frac{\partial\lambda_n}{\partial\theta_j}\frac{C_n^t}{\lambda_n^2}. \tag{9}$$

Since observations may vary according to the distribution of the variables, the derivatives may fluctuate leading to unstable iterations. It has been shown that if instead the 2nd order derivative is replaced by its expected value (taken over the all possible counts), it leads to more stable iterations [17]. By replacing the Hessian by its expectation (also known as the Fisher Information), we get the *Fisher's scoring iterations*. Since the expected value of $E[C_n^t] = t\lambda_n$ the 2nd order terms in $H(\theta_i, \theta_j)$ cancel out and we get the form in (10):

$$E\left[H(\theta_i, \theta_j)\right] = -\sum_n \frac{\partial\lambda_n}{\partial\theta_i}\frac{\partial\lambda_n}{\partial\theta_j}\frac{1}{\lambda_n}. \tag{10}$$

With the elimination of the 2nd order terms, the Fisher's scoring iterations are given by

$$\theta(k+1) = \theta(k) - \left(J_\lambda^T W J_\lambda\right)^{-1}\left(J_\lambda^T W\right)F. \tag{11}$$

Above, $J_\lambda(n, j) = \partial\lambda_n/\partial\theta_i$ is the $N \times 3M$ Jacobian matrix, W is an $N \times N$ diagonal matrix with elements $1/\lambda_n$ and $F_n = C_n^t - \lambda_n t$ denotes the residual between the signal model and the observations at the nth node.

B. Weighted Least Squares and Equivalence to MLE

For the problem under consideration, it can also be shown that the iterative solution for MLE is equivalent to the iterative solution of a properly weighted least squares estimate [5], [17]. Since the WLS presents a considerably simpler formulation, it would be useful when we turn our attention to tracking for moving sources. Here we briefly outline the equivalence.

In the least squares formulation, we only assume a signal model to fit the observed data in the *least squares* sense. For the radiation localization problem, LS formulation can be defined as follows: Let $F_n(C_n^t, \theta) = C_n^t - \lambda_n t$ denote the residuals (difference between assumed signal model and observations) and w_n a weight assignment for the nth observation. WLS parameter estimate $\hat{\theta}$ is given by the following optimization problem:

$$\hat{\theta} = Argmin \sum_{n=1}^{N} w_n F_n^2(C_n^t, \theta). \tag{12}$$

Since the above nonlinear function does not have a closed form solution, we can use Newton's method as before:

$$\theta(k+1) = \theta(k) + H(\theta)^{-1} G(\theta)\Big|_{\theta=\theta(k)}. \tag{13}$$

Above G and H denote the 1st and 2nd order gradients for the WLS objective function. The Newton iterations can be further simplified through the *Gauss-Newton* approximation. The Hessian term is approximated by ignoring the 2nd order derivatives. The gradient and the Hessian terms are given by (14) and (15):

$$G(\theta_i) = 2\sum_{n=1}^{N} w_n F_n \frac{\partial F_n}{\partial\theta_i} \tag{14}$$

$$H(\theta_i, \theta_j) \approx 2\sum_{n=1}^{N} w_n \frac{\partial F_n \partial F_n}{\partial\theta_i\partial\theta_j}. \tag{15}$$

In WLS, optimality is achieved when weights are assigned inversely proportional to the variance in the observations (or any consistent estimator of the variance). For the Poisson random variable, since the variance is equal to the mean, the optimal weight allocation is $w_n = 1/\lambda_n$. The WLS iterations are then given by (16). It is clear that the iterations are equivalent to the Fischer's scoring method since $J_F(n, i) = -\partial\lambda_n/\partial\theta_i$. In the subsequent discussions iterative WLS and MLE are used to denote the equivalent iterative procedures:

$$\theta(k+1) = \theta(k) + \left(J_F^T W J_F\right)^{-1}\left(J_F^T W\right)F. \tag{16}$$

C. CRLB for the Location Estimation Problem

To measure the performance of our estimation procedure, we look at the variance between the estimate and the true value of parameters and compare it to the theoretical Cramer Rao lower bound (CRLB).

CRLB in the simplest form states that the variance for any unbiased estimator is at least as high as the inverse of the Fisher

Information Matrix (FIM) as given by (17). Then if we have an unbiased estimator, CRLB states that $Cov_\theta(Est_\theta(C^t)) \geq I(\theta)^{-1}$. The FIM terms in (17) can be computed using the Hessian terms. Under certain regularity conditions (valid for our likelihood function) the expected value of the Hessian terms is equal to the negative of the Fisher Information Matrix, i.e., $I(\theta_i, \theta_j) = -E[H(\theta_i, \theta_j)]$. Thus we can compute the FIM using the Jacobians similar to (11) in (17):

$$\begin{aligned} I(\theta_i, \theta_j) &= E\left[\frac{\partial L(\theta, C^t)}{\partial \theta_i}\frac{\partial L(\theta, C^t)}{\partial \theta_j}\right] \\ &= -E\left[\frac{\partial^2 L(\theta, C^t)}{\partial \theta_i \partial \theta_j}\right] = \frac{\partial \lambda_n}{\partial \theta_i}\frac{\partial \lambda_n}{\partial \theta_j}\frac{1}{\lambda_n} \\ &= \left(J_\lambda^T W J_\lambda\right). \end{aligned} \tag{17}$$

D. Algorithm Complexity

We briefly analyze the computational complexity of the approach. In particular we compare the complexity relative to other related literature such as references [9]–[12] which employ Bayesian estimation methods.

In the iterative MLE solution, the most expensive operation is the inversion of a $3M \times 3M$ matrix. Thus the iterative MLE grows as $O(NKM^3)$ where N is the number of nodes, K the number of iterations, and M the number of sources. In Bayesian methods parameters are estimated from the posterior probability (likelihood times the prior probability) for which a closed form solution does not exist for this problem. Numerical approximation methods using Importance Sampling are used for estimation. Here each sample of parameters is represented by a 3xM dimensional particle (or a sample vector) for which the prior probability and likelihood are computed. Thus an equivalent Bayesian approach grows as $O(NKPM)$ where P is the number of samples and M the number of sources.

For the iterative MLE, the matrix inversion becomes more expensive and more iterations are required for convergence as the problem scales .In contrast Bayesian approaches become intractable due to the complexity of representing the posterior distribution through samples. The number of samples required for good representation of the posterior grows exponentially with the number parameters. The iterative MLE algorithm has a*power law* scaling while the Bayesian estimation scales *exponentially*. Thus iterative MLE would scale better as the dimensionality increases. A comparative evaluation is part of future work.

III. Finding Initial Estimates Using the Expectation Maximization Algorithm

Recall that convergence of Newton's iterations for optimization depends on many factors such as the function is twice differentiable with continuous derivatives and the initial estimate is sufficiently close to the solution. A short discussion with references related to conditions and assumptions for Newton's method can be found here [20].While the conditions on derivatives are usually met for problems where Newton's method is applied, finding initial estimates can still be a challenge. Fig. 2 illustrates the problem. An initial point is sufficiently close if the function can be approximated by Taylor series expansion in the interval between the point and the maxima such that higher order derivatives can be ignored. As can be seen from Fig. 2, the assumption does not always depend on the actual distance between the initial point and the maxima. Thus we cannot emphasize too strongly the need for good initial estimates in the convergence of Newton's iterative method.

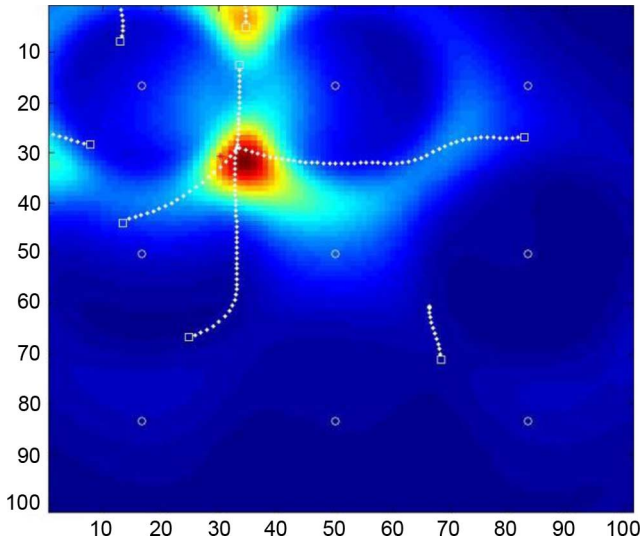

Fig. 2. Plot showing the log-likelihood map for a single source and the iterative computation of MLE starting at different points. Circles represent the detectors, squares represent initial estimates and dashed lines represent the successive estimates starting with the initial estimate. Axes represent distances in meters.

For a single source, several heuristics can be considered for finding the initial estimates. These include 1) estimates from a coarse grid search, 2) multiple random start points, and 3) initial estimate around the node with the highest number of counts. However, the methods are not well suited for multiple sources. Any grid-based search quickly becomes intractable as the number of parameters increase. The third approach scales linearly with parameters but is inconsistent when searching for multiple sources. For example, we frequently observed a high number of counts in multiple nodes encircling a strong source while a weaker source at another location had fewer counts for nodes around it. Finally each of the above approaches assumes that we know the number of sources a priori which limits the utility of these approaches.

A. Approximate Source Distribution Using the Expectation Maximization Algorithm

Our proposed approach is based on estimating the approximate activity distribution in the monitored region. This is in contrast to the formulation in (1)–(5) where we assumed a fixed number of point sources. We divide the monitored area into a 2-D grid and attempt to estimate the approximate source strength over this grid. The "hot-spots" in the estimated source distribution are potential locations of sources and starting points for the iterative MLE method.

To estimate the source distribution, we turn to the EM approach due Dempster, Laird and Rubin (DLR) [13]. EM estimation has been used for estimating Poisson mixture densities with

Compton and PET detectors [14]–[16] with directional information in observation matrix. Directional information has non-uniform projection probability which creates a unique mapping of source pixels in the observed volume. The effect is that the problem leads to a Poisson mixture density estimation to which the EM algorithm of Dempster *et al.* can be applied, for example in [14] for PET detectors. For the non-directional detector array embedded inside the area under observation, we need to consider certain simplifying assumptions to reduce our problem to a similar Poisson mixture density form. The final form of our EM algorithm is equivalent to Shepp *et al.* [14] albeit with a different transition probability which derives from the simplifying assumptions.

Let $j = 1 \ldots R \in \mathbb{R}^2$ be the set of 2-D grid points of the monitored field and $\lambda_j \in [\lambda_1, \ldots, \lambda_R]$ the source strength at these grid points. Let $\lambda_{n,j} \in [\lambda_{n1}, \ldots, \lambda_{nR}]$ be the incident rate at n due to λ_j attenuated by $\alpha_{n,j} = \mu A/\pi d_{n,j}^2 e^{-\rho d_{n,j}}$, with $d_{n,j}$ the distance between node n and location j. Each node detects C_n^t counts for the exposure duration of t seconds.

We note here that we may apply Shepp's iterative solution on the above by using the attenuation factor $\alpha_{n,j}$ as the transition probability matrix. However, since $\alpha_{n,j}$ monotonically decreases with distance, we get a trivial source distribution with low rate point sources on pixels nearest to each detector and zero elsewhere. Starting with an initial distribution estimate which prevents the above (for example with strong source strength midway between detectors) and regularization did not solve this problem.

While the trivial distribution is mathematically a valid explanation of the observed counts, what we are looking for is a distribution with few peaks denoting point sources with some minimum activity (e.g., 0.1 mCi), and the rest of the field with some negligible activity denoting the background counts. Since the detectors do not have any directionality, we assume further that these peaks are well separated in space such that only one peak or point contributes to counts at any detector.

Let $\Psi(C, \lambda) = e^{-\lambda}\lambda^C/C!$ denote the Poisson PDF. If $\Lambda_n = \sum_j \alpha_{nj}\lambda_j$ is the total incident rate at node n, the likelihood of counts is given by $P(C_n^t|\lambda) = \Psi(C_n^t, \Lambda_n t)$. However, if multiple sources are relatively well separated, most of the counts will be contributed by a single source located near each node. In other words, if j is the location of the nearest point source to n, we can approximate the sum $\sum_k \alpha_{nk}\lambda_k \approx \alpha_{nj}\lambda_j$. Let $z_n : \{z_{n1}, \ldots, z_{nR}\}$ be the indicator variable to denote this location for each node:

$$z_{nj} = \begin{cases} 1 & \text{if counts originated in pixel } j \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

By our assumption only one term in z_n is non-zero. This allows us to remove the summed rate term Λ_n in $\Psi(C_n^t, \Lambda_n t)$ and convert $\Psi(C_n^t, \Lambda_n t)$ to (19):

$$P(C_n^t, z_n|\lambda) = \sum_j z_{nj}\Psi(C_n^t, \alpha_{nj}\lambda_j t). \tag{19}$$

The vector z_n indicating the location of the source cannot be observed and is a *latent variable*. By including z_n and taking the joint likelihood over all N nodes we form the complete likelihood function in (20):

$$P_{CD}(C_1^t, \ldots, C_N^t, z_1, \ldots, z_N|\lambda) = \prod_n \sum_j z_{nj}\Psi(C_n^t, \alpha_{nj}\lambda_j t). \tag{20}$$

We can further convert (20) to a purely product form since for any node only one term in z_n is non-zero. Denoting the vector of counts by C, and latent variables z_{nj} by Z, (20) can also be written as

$$P\left(C, Z|\lambda\right) = \prod_n \prod_j \Psi\left(C_n^t, \alpha_{nj}\lambda_j t\right)^{z_{nj}}. \tag{21}$$

Our next objective is to maximize (21) through the iterative EM algorithm. Before we introduce the EM steps, we first simplify (21) to a more convenient form as a product of Poisson and multinomial distributions in (22) to (24). First, the probability that a single gamma ray particle from location j hits the sensor n is given by

$$P\left(gamma = j\right) = \frac{\alpha_{nj}\lambda_j}{\sum\limits_j \alpha_{nj}\lambda_j} = \beta_{nj}. \tag{22}$$

Next, using the β_{nj} definition in (22), the probability of registering a total of $C^t = \sum_n C_n^t$ counts at N nodes given the distribution λ is given by the multinomial probability function in (23):

$$\Omega(C, Z|\lambda) = \left(C^t\right)! \prod_n \prod_j \left(\frac{\left(\beta_{nj}\right)^{C_n^t}}{C_n^t!}\right)^{z_{nj}}. \tag{23}$$

Let $\Lambda = \sum_n \sum_j \alpha_{nj}\lambda_j$ be the total rate after attenuation at all the detectors. Finally, (21) can now be written as a product of the multinomial function in (23) and a Poisson probability of total counts by simple manipulation:

$$P_{CD}\left(C, Z|\lambda\right) = \Psi\left(C^t, t\Lambda\right)\Omega\left(C, Z|\lambda\right). \tag{24}$$

Next we maximize the log-likelihood $L_{CD}(\lambda|C^t, Z) = lnP_{CD}(C^t, Z|\lambda)$ over the latent variables Z using the EM algorithm which consists of two iterative steps.

Expectation Step: Let λ^k be the current estimate of the source distribution. In the expectation step we compute the expectation of the complete log-likelihood for λ^{k+1} over all values of the latent variable z, with λ^k as the prior probability:

$$\begin{aligned} Q\left(\lambda^{k+1}|\lambda^k\right) &= E_z\left[L_{CD}\left(\lambda^{k+1}|C, Z\right)|\lambda^k\right] \\ &= \sum_z P\left(Z|C, \lambda^k\right) L_{CD}\left(\lambda^{k+1}|C, Z\right). \end{aligned} \tag{25}$$

In the above, $P(Z|C, \lambda^k)$ denotes the prior probability of Z. Since the probability $P(Z|C, \lambda)$ is not known, the current estimate of the distribution is used to compute the expectation. In the expectation, we can simply replace z_{nj} by its expected value which is equivalent to conditional probability given the

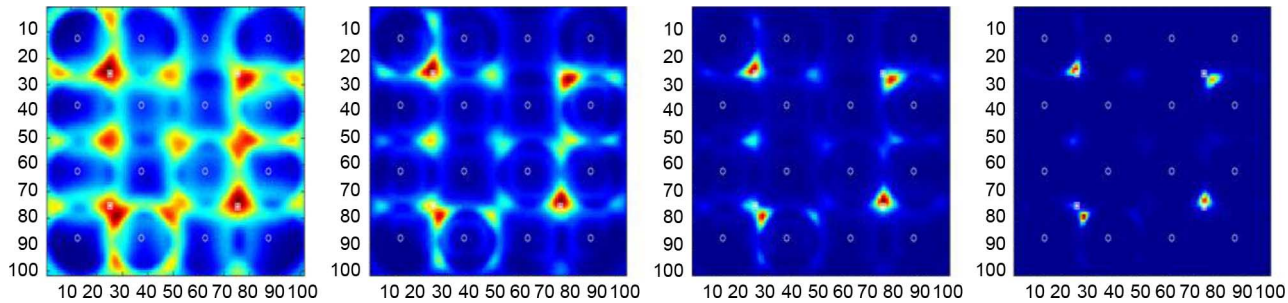

Fig. 3. Source distribution estimate at Iterations 1–4 (left to right) with 1 mCi Cs source at 5 s exposure Axes represent distances in meters.

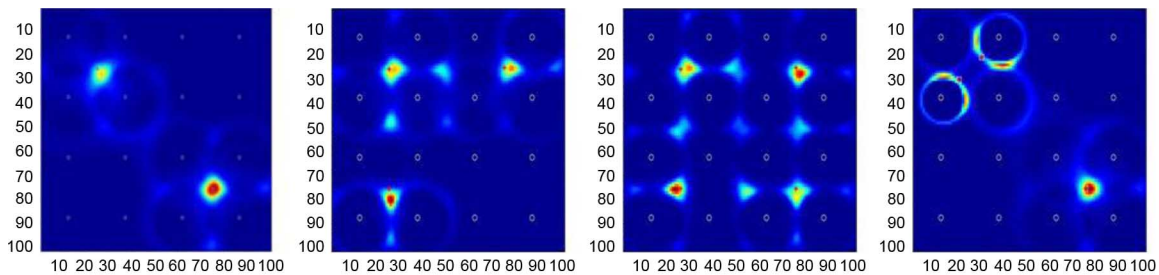

Fig. 4. Source distribution using Approximate EM 1 mCi Cs source at 5 s exposure for (a) 2 sources, (b) 3 sources, (c) 4 sources, and (d) 3 sources with two sources close in close proximity. Axes represent distances in meters.

observed counts and the current estimate of the source distribution:

$$\begin{aligned}\hat{z}_{nj}^k &= P(z_{nj}|C_n^t, \lambda^k) \\ &= \frac{P(z_{nj}=1)p(C_n^t|j)}{\sum_j P(z_{nj}=1)p(C_n^t|\alpha_{nj}\lambda_j^k)} \\ \hat{z}_{nj}^k &= \frac{\beta_{nj}^k \Psi(C_n^t, t\alpha_{nj}\lambda_j^k)}{\sum_j \beta_j^k \Psi(C_n^t, \alpha_{nj}\lambda_j^k)} \\ &= \frac{\alpha_{nj}\lambda_j^k \Psi(C_n^t, t\alpha_{nj}\lambda_j^k)}{\sum_j \alpha_{nj}\lambda_j^k \Psi(C_n^t, t\alpha_{nj}\lambda_j^k)}. \end{aligned} \tag{26}$$

From (24) and (26) the expectation in (25) is given in (27). The superscripts denote whether the terms use λ^{k} or $\lambda^{\mathrm{k+1}}$:

$$Q(\lambda^{k+1}|\lambda^k) = \sum_{n=1}^{N}\sum_j \left(\hat{z}_{nj}^k C_n^t \beta_{nj}^{k+1}\right) + ln\Psi\left(C^t, t\Lambda^{k+1}\right). \tag{27}$$

Maximization Step: In the maximization step, we find the new value of the parameter which maximizes the conditional expectation. We take the derivative of the $Q(\lambda^{k+1}|\lambda^k)$ with respect to the λ^{k+1} (keeping λ^k fixed), and set it to zero to find the new value of the parameters:

$$\lambda^{k+1} = \underset{\lambda}{Argmax}\, Q(\lambda^{k+1}|\lambda^k) \tag{28}$$

$$\begin{aligned}\frac{\partial Q}{\partial \lambda_j^{k+1}} &= \frac{\sum_{n=1}^{N} C_n^t \hat{z}_{nj}^k}{\lambda_j^{k+1}} - \frac{\sum_{n=1}^{N} C_n^t \hat{z}_{nj}^k \alpha_{nj}}{\sum_j \lambda_j^{k+1}\alpha_{nj}} \\ &\quad + \frac{C^t \sum_n \alpha_{nj}}{\sum_n \sum_j \alpha_{nj}\lambda_j} - t\sum_n \alpha_{nj} = 0. \end{aligned} \tag{29}$$

The two inner terms cancel out since both represent the proportion of counts received from location j given the source distribution. After simplification and using $\hat{z}_{nj}^k$ from (26) we get the EM iterative update step in (30):

$$\lambda_j^{k+1} = \frac{\lambda_j^k}{t\sum_n \alpha_{nj}} \sum_{n=1}^{N} \frac{C_n^t \alpha_{nj} \Psi(C_n^t, t\alpha_{nj}\lambda_j^k)}{\sum_j \alpha_{nj}\lambda_j^k \Psi(C_n^t, t\alpha_{nj}\lambda_j^k)}. \tag{30}$$

After each iteration we need to correct the *range* of updated distribution λ^{k+1} such that it reflects the number of counts received at each node. This is required to account for the approximations assumed in the estimation process. This step may also be considered as a form of *regularization* of the EM estimates.

Let $\tilde{j} = \underset{j}{argmax}\{\lambda\}$ be the pixel with the highest intensity and $\tilde{n} = \underset{n}{argmin}\{Dist(n, \tilde{j})\}$ be the node which is closest to it. Since we assumed that nodes get counts from a source close to it, we assume that node $\tilde{n}$ received all its counts from location $\tilde{j}$. Thus the source rate at pixel $\tilde{j}$ under these assumptions should be $C_{n^{\tilde{j}}}/\alpha_{n^{\tilde{j}}}$. Hence the corrected source distribution is given by (31). The correction factor may also be computed by considering the counts of all the nodes and doing a least squares fit. However, in our evaluation, correction using the attenuation from the maximum intensity spot to a single node was sufficient:

$$\lambda^{k+1} \rightarrow \frac{\lambda^{k+1} C_{\tilde{n}}^t}{t\lambda_{\tilde{j}}^{k+1}\alpha_{\tilde{n}\tilde{j}}}. \tag{31}$$

Fig. 3 shows the source intensity estimate at different iterations for a scenario with 4 sources. Fig. 4 shows the source intensity estimate for scenarios involving 2, 3, and 4 sources and for a 3-source scenario in which 2 sources are close to each other.

When the sources are distant, the approximate EM can accurately estimate the number of sources, the location and intensity

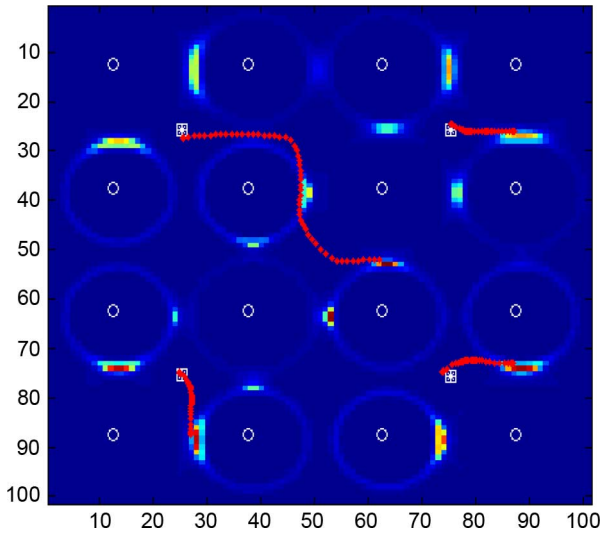

Fig. 5. Iterative MLE corrections with EM start points. 1 mCi Cs sources at 5 s exposure at four corners Axes represent distances in meters.

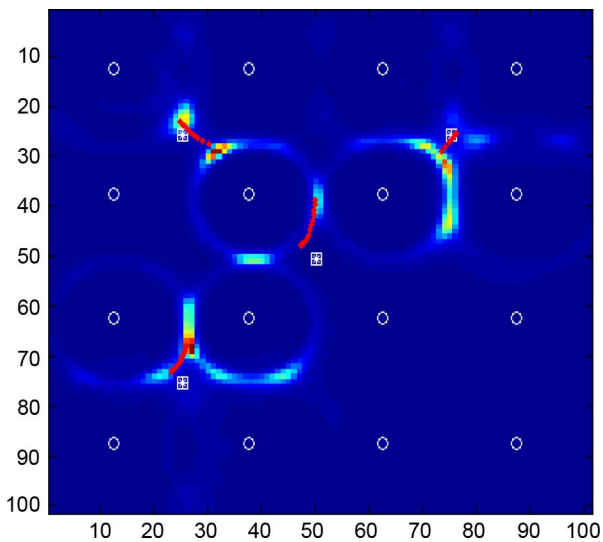

Fig. 6. Iterative MLE corrections with EM start points. Three 1 mCi Cs sources at the corners and one at the center after 5 s of exposure. Axes represent distances in meters.

of the sources. The approximate EM is less effective when multiple sources are in close proximity to each other. The method also creates artifacts (in the form of circular rings) in the estimation since the detectors are non-directional. However since it provides a good initial estimate, the EM solution can be refined using the iterative MLE method. Figs. 5 and 6 show the iterative MLE refinement with start points from the EM algorithm. Sources are shown with a white square and solution of iterative WLS is shown with red dots. All the cases shown locate the source accurately.

The peaks in the source distribution represent the likely locations of the point sources. The summed rate in a small circular region followed by thresholding is used for detecting hot-spots and for identifying the potential number of sources present. The threshold for selecting the hot-spots is 0.1 mCi. We start by selecting the highest intensity spot as the location of the first source. Next, we create a window around this estimate and select the highest intensity location outside this region.

Since the EM procedure is independent of the number of sources, the complexity is only dependent on the grid size and the number of iterations. To find initial estimates a grid size of 20×20 m^2 and 5 iterations proved sufficient (the images in Figs. 1–6 were generated for a 100×100 grid for illustration).

IV. Estimating Trajectory of a Moving Source

Next, we introduce our algorithm to estimate the trajectory of a moving and maneuvering source. The trajectory estimation problem requires the estimation of the entire path through which a radiation source moves through the monitored field. Thus the problem is to estimate multiple points which best fit the observed data. In contrast, the problem of tracking involves the estimation of the current location (or a single point) of a moving source. While tracked points may be considered as the approximate trajectory, joint estimate of a set of points should provide a more accurate estimation of the trajectory. Reference [18] considers a similar approach for linking point estimates of locations for a moving source. Bayesian approaches for tracking typically use Markov chain approximation to reduce computational complexity (current estimate based only on the previous estimate of the source location).

Our method is based on formulating the problem as a constrained WLS optimization. Recall that in WLS we need a signal model to formulate the objective function. When the target is moving in a straight line, the trajectory is trivial and hence simple motion model based on speed, direction and initial position can be used to parameterize the model. Such a model was assumed in [1], [10], and [12] where the signal model (total incident rate) involved integral over a straight line path. For a maneuvering target, such a simplistic model cannot be assumed. While more complex models involving splines may be used for a maneuvering target, we adopt a simpler approach which proves to be effective.

We define the source trajectory by a set of points defined over a time window (T to $T + T_w$). We do not impose a specific trajectory model. Instead we reduce the problem dimensionality by assuming certain restrictions in motion. In order for the set of points to be a valid trajectory the distance between two consecutive trajectory points needs to satisfy the maximum velocity assumed for a moving source for example 2 m/s for pedestrian motion or 20–50 m/s for a vehicle. Other constraints such as the change in direction (not considered in this paper) may also be imposed for any practical maneuvers by a source. The problem is then formulated as a constrained-weighted least squares in which we estimate the set of points which best fit the observation.

We assume that sensors register counts c_n^t during the tth time period (time period is 1 sec in our evaluation). Note the difference in the definition for the static case (1) in previous sections where C_n^t denoted the cumulative counts during the entire exposure period of t seconds. Let θ denote the set of points defining the trajectory and the source of intensity λ_S in a time window T_w, $F(\theta)$ the WLS objective function and $G^t(\theta)$ the velocity constraints. We assume that the target is approximately static during each time interval of 1 s such that any point wise estimate

X^t, Y^t corresponds to the counts c_n^t for that duration. The constrained optimization problem is then formulated in (32)–(35):

$$\theta = \left(X^T, Y^T, \ldots, X^{T+T_w}, Y^{T+T_w}, \lambda_S\right) \tag{32}$$

$$F(\theta) = \sum_{t=T}^{T+T_w} \sum_{n=1}^{N} \frac{1}{\lambda_n^t}\left(c_n^t - \lambda_n^t\right)^2 \tag{33}$$

$$G^t(\theta) \equiv Max_{Speed} - Dist\left(X^t, Y^t, X^{t+1}, Y^{t+1}\right) \geq 0 \tag{34}$$

$$Minimize\ F(\theta)$$
$$Subject\ to\ G^t(\theta) \geq 0. \tag{35}$$

Above, $\lambda_n^t = \mu A \lambda_s / 4\pi d_{n,t}^2 e^{-\rho d_{n,t}}$ is a function of sensor and source distance $d_{n,t}^2 = (X^t - x_n)^2 + (Y^t - y_n)^2$ at time t. The above is a nonlinear constrained optimization problem of high dimensionality for which we would like to use iterative methods. Iterative solutions for constrained optimization problems are conveniently formulated using the IPM [19]. In IPM the idea is to re-formulate the constrained optimization problem as an unconstrained one which can be solved using Newton's iterative method.

We first represent the constraints inside a *Barrier Function* which acts as a penalty for crossing the boundary of the feasibility region. Typically a log-barrier function is used which has the property that it is infinite for parameters outside the feasibility region. This is given by

$$\varphi(\theta) = -\sum_t \log\ G^t(\theta). \tag{36}$$

Next the constrained optimization function is approximated as an unconstrained function comprising of the objective and the barrier functions as follows:

$$Minimize\ F(\theta) + \frac{1}{\alpha}\varphi(\theta). \tag{37}$$

It is clear that as $\alpha \to \infty$ the above reduces to the unconstrained objective function itself. The Newton's iterations in the Interior Point Method consist of inner and outer iterations. The inner loop minimizes the transformed unconstrained function while the outer loop progressively increases the parameter α to improve the approximation. Thus at each step the problem is optimized in the interior of a feasible region while progressively arriving at the unconstrained solution.

The IPM iterative steps are shown in Fig. 7. In the iterations the two second-order derivative terms are computed separately. The Hessian matrix $\nabla^2 F(\theta)$ is approximated using the Gauss-New ton approximation ($\nabla^2 F \approx J_F^T J$). Each of the terms $F(\theta_t) = \sum_{n=1}^{N} 1/\lambda_n^t (c_n^t - \lambda_n^t)$ and its derivatives are independent in time and hence can be computed separately. The Hessian terms $\nabla^2 \varphi(\theta)$ for the log-barrier function are relatively simpler to compute and no approximations are required.

Consider a new time period, T, when the detectors register a new set of observations, and a new point estimate of location of the source is estimated using the EM algorithm. We first create an initial parameter vector using the last $T_w - 1$ estimates and the new estimate at the Tth interval $\theta^0 = \{X_0^{T-T_w+1}, Y_0^{T-T_w+1} \ldots X_0^T, Y_0^T\}$.

Find an initial feasible solution set θ^0
Begin: $\theta = \theta^0, \alpha = \alpha^0$, *m=number of constraints*
While outer stopping criterion not met
  While inner stopping criterion not met
    $\Delta\theta = -\left(\nabla^2 F(\theta) + \frac{1}{\alpha}\nabla^2\phi(\theta)\right)^{-1}\left(\nabla F(\theta) + \frac{1}{\alpha}\nabla\phi(\theta)\right)$
    find multiplier δ^**through line search*
    $\theta = \theta + \delta^*\Delta\theta$
  Until $d\theta < \varepsilon\_inner$
  $\alpha = 2\alpha$
Until $\frac{m}{\alpha} < \varepsilon\_outer$

Fig. 7. Iterative solution of the constrained WLS formulation for tracking with the Interior Point Method.

Before we conduct the iterations of IPM, we need to make sure that θ^0 is a feasible solution set and does not violate the constraints, i.e., $G_{tW}(\theta) < 0$. To ensure this, we consider a pair of consecutive estimates $\{X_0^t, Y_0^t, X_0^{t+1}, Y_0^{t+1}\}$ and adjust their values to satisfy the constraints. In our case, we need to ensure that the distance between two consecutive points is less than the maximum speed. Equation (38) shows the adjustment for the X coordinate which ensures (when the same adjustment is conducted for the Y coordinate) that the distance between two consecutive points is less than the maximum speed:

$$X_0^t \to \frac{X_0^{t+1} + X_0^t}{2} - \vec{u}\frac{Max_{Speed}}{\sqrt{2}}$$
$$X_0^{t+1} \to \frac{X_0^{t+1} + X_0^t}{2} + \vec{u}\frac{Max_{Speed}}{\sqrt{2}}$$
$$where\ \vec{u} = \frac{X^{t+1} - X^t}{abs(X^t - X^{t+1})}. \tag{38}$$

The above process is repeated for all the T_w point estimates to get a feasible initial set. Since λ_S is constant for the entire duration, the average value across the time period for each point estimate is used as the initial estimate. The constrained optimization is then conducted using the double loop shown in Fig. 7.

Figs. 8–13 illustrate the trajectory estimate for different sizes of time windows considered for optimization. The red points denote the true trajectory while blue dots represent the estimated trajectory. Fig. 8 illustrates the case with window size $T_w = 1$ s which is essentially the unconstrained estimates at each second. For the graphs we considered a moving target 1 mCi with speed of 1 m/s. The maximum velocity is assumed to be 2 m/s. The estimated trajectory progressively improves as the window size increases. It is evident that larger the number of points jointly optimized to estimate the trajectory, the better is the estimate. It becomes infeasible to optimize in real time beyond a window size of 100 points. However, as Fig. 21 shows, we get marginal improvements beyond a window size of 20 s for the scenario considered.

V. Results

For evaluations, we assume 0.0762 mNaI based scintillating detectors. The detector efficiency, the air attenuation coefficient and background rate (3.3 cps) is based on measurements for the Cs-137 photo-peak region of 662 keV. We simulate a $100 \times 100\ \mathrm{m}^2$ area with sensors placed in a square grid formation as shown in Fig. 1 and with sources randomly located in each trial.

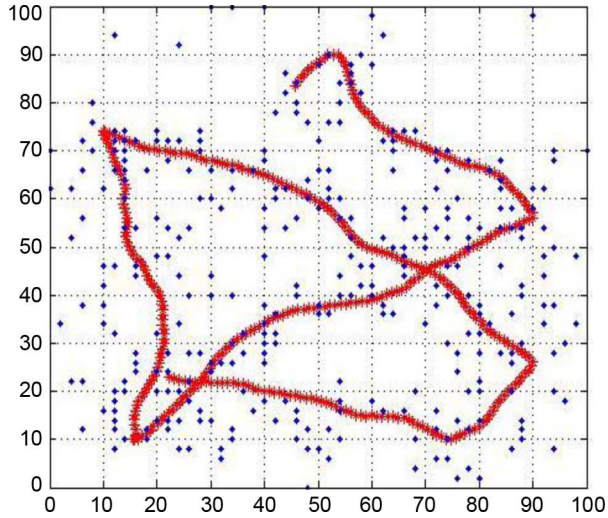

Fig. 8. Trajectory with Window Size = 1 s (Estimated trajectories for a 1 mCi Cs source moving at 1 m/s through a grid of 16 sensors. Axes represent distances in meters).

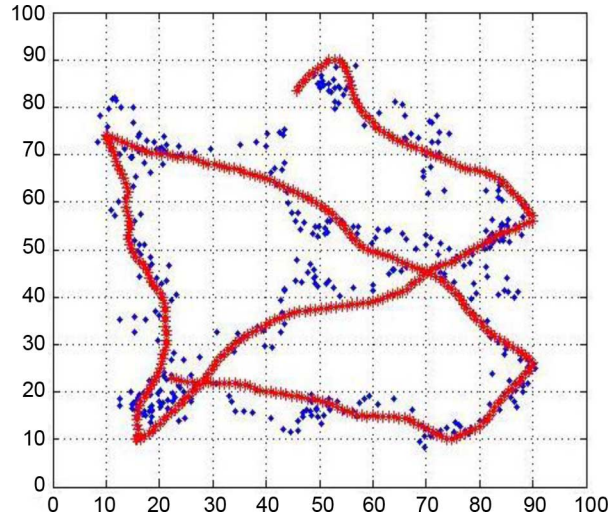

Fig. 10. Trajectory with Window Size = 3 s (Estimated trajectories for a 1 mCi Cs source moving at 1 m/s through a grid of 16 sensors. Axes represent distances in meters).

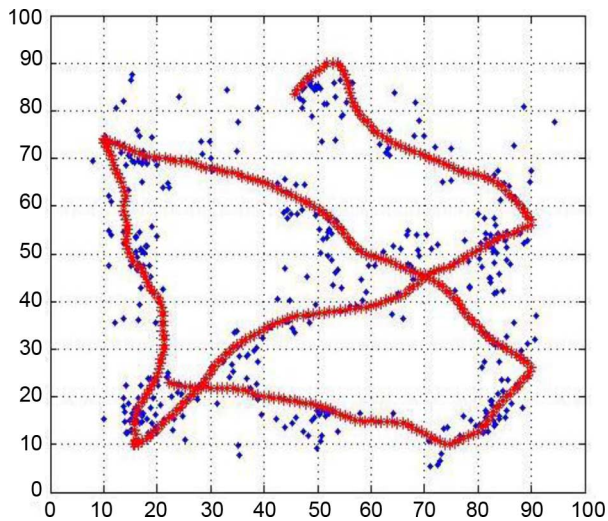

Fig. 9. Trajectory with Window Size = 2 s (Estimated trajectories for a 1 mCi Cs source moving at 1 m/s through a grid of 16 sensors. Axes represent distances in meters).

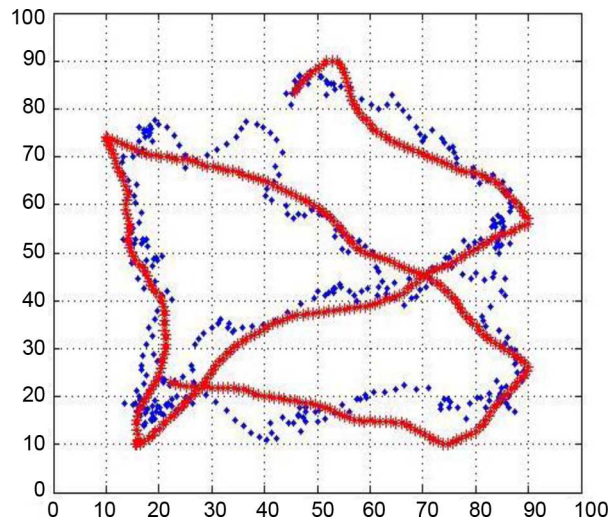

Fig. 11. Trajectory with Window Size = 5 s (Estimated trajectories for a 1 mCi Cs source moving at 1 m/s through a grid of 16 sensors. Axes represent distances in meters).

Photo-peak counts are generated for each detector as a Poisson random variable. Plots generated show the estimation errors as a function of exposure time. It is assumed that spectra from sensors are collected at 1 s intervals at a central server which computes the parameter estimates within the next 1 s interval. The simulated data and the algorithms have been implemented in MATLAB and run on a laptop with 2.5-GHz processor as a single thread. The estimation algorithms for up to 4 sources take less than 1 s of execution time. Since the data collection and estimation frequency is 1 s, the algorithm can adequately keep up with the data collection speed.

The location error is calculated as the average distance between the true and the estimated location of the source. For multiple sources, the average is computed over the errors for all the sources. For rate estimation error, the mean of the absolute deviation (or difference) between true and estimated rate is computed.

Fig. 14 shows the location error for the EM and the iterative WLS solution (with the EM estimate as starting points) for scenarios involving 2, 3 and 4 sources. For these plots, the static sources are placed as shown in Fig. 1. While EM provides a good initial estimate, we can see that WLS can refine the location accuracy. Next in Fig. 15, the iterative WLS refinement is compared with the theoretical CRLB for the fixed configuration of sources placed in corners of the sensor grid as shown in Fig. 1. The iterative WLS refinement matches the CRLB in less than 10 seconds of exposure.

Next we evaluate the average performance of the approach over trials involving random locations for sources. Fig. 16 shows the average location error with 16 detectors and 2, 3 and 4 randomly located sources over 10 000 trials. Fig. 17 shows the average error rate for 2 randomly located as the number of sensors is increased from 9 to 25 nodes. It also shows that both sources are located with equivalent precision across multiple

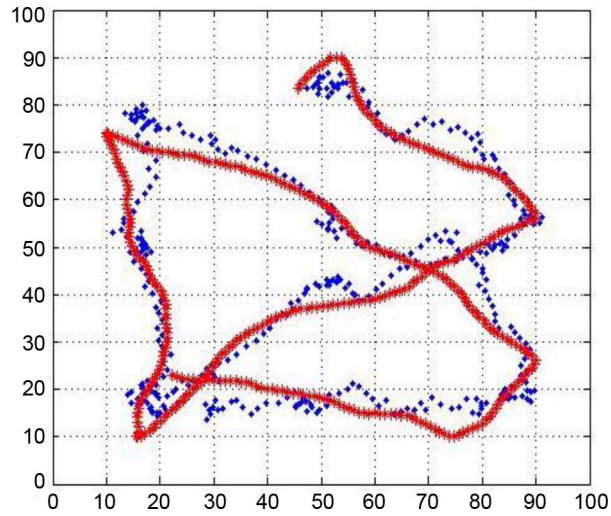

Fig. 12. Trajectory with Window Size $= 10$ s (Estimated trajectories for a 1 mCi Cs source moving at 1 m/s through a grid of 16 sensors. Axes represent distances in meters).

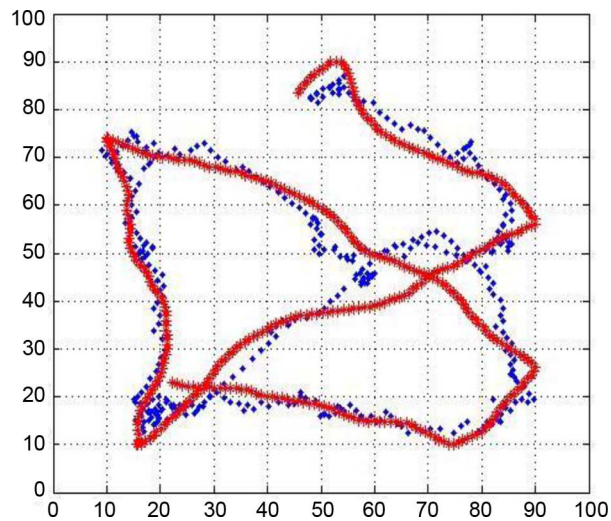

Fig. 13. Trajectory with Window Size $= 20$ s (Estimated trajectories for a 1 mCi Cs source moving at 1 m/s through a grid of 16 sensors. Axes represent distances in meters).

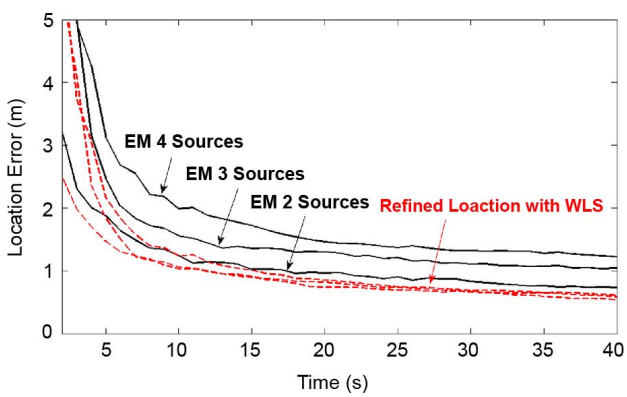

Fig. 14. Average location error (m) for EM and iterative WLS for the configuration in Fig. 1 with 16 nodes and 1 mCi sources.

trials. Fig. 18 also plots the error in rate estimates for these sources.

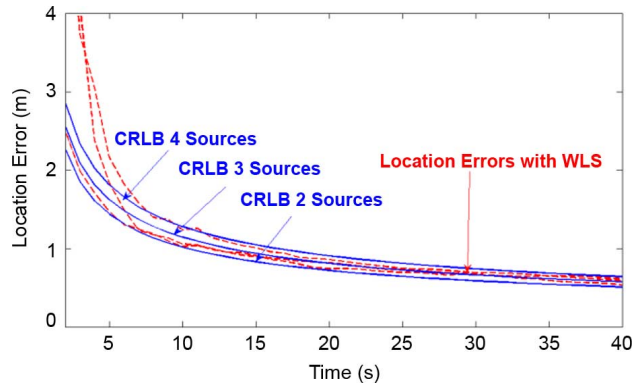

Fig. 15. Average location error (m) for iterative WLS vs CRLB for sensor configuration in Fig. 1 with 16 nodes, and 1 mCi sources.

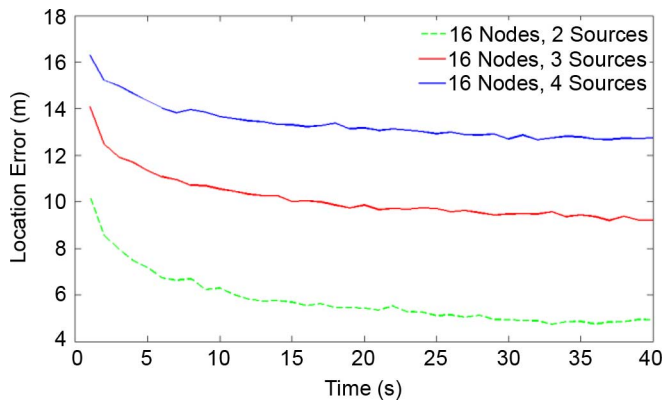

Fig. 16. Average location error (m) for the configuration in Fig. 1 with 16 nodes and 1 mCi sources.

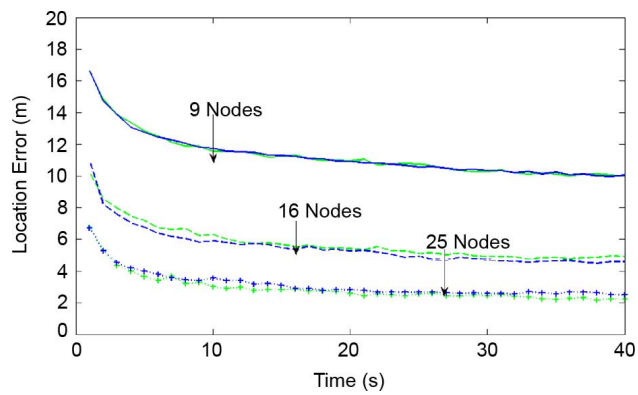

Fig. 17. Average location error (m) with different number of nodes in a 100×100 m^2 area and two 1 mCi sources.

Comparing Fig. 15 (well-separated sources) and Fig. 16 (completely random locations), we see that there is a large difference in the error rates. This is because if two or more sources fall within close vicinity of each other, the non-directional detectors are unable to distinguish between the two and hence report them as one single source with a larger rate. This effect is a limitation of the non-directional detectors since we require at least 4 nodes encircling a single source to locate the position and the rate uniquely.

The above location estimates are for cases when our hot-spot based detection algorithm (from EM estimate) correctly detects

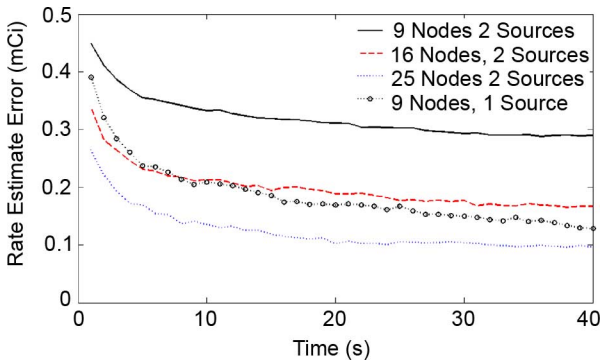

Fig. 18. Average rate estimate error (mCi) with two sources in a 100×100 m^2 area with 9 nodes and two 1 mCi sources.

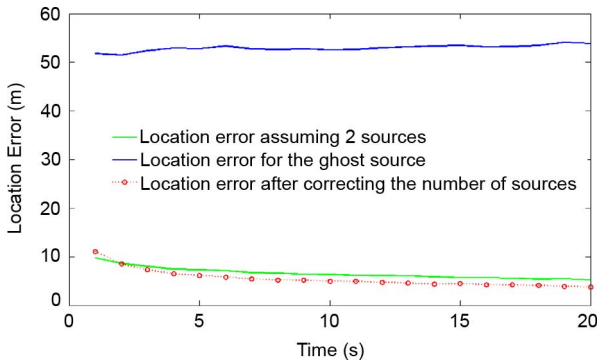

Fig. 19. Average location error with over-fitting (estimate 2 sources when only 1 is present) in a 100×100 m^2 area with 9 nodes and one 1 mCi source.

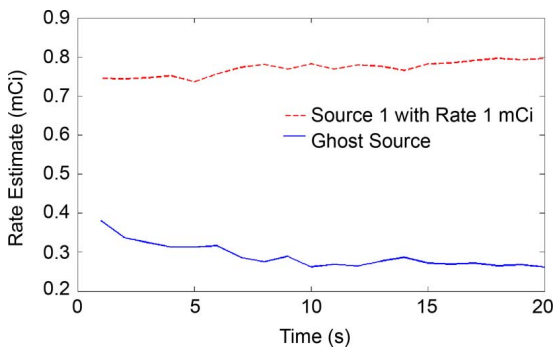

Fig. 20. Average rate estimate with over-fitting (estimate 2 sources when only 1 is present) in a 100×100 m^2 area with 9 nodes and one 1 mCi source.

the number of sources present. With a false detection, the observations would be over-fitted for a ghost source. Fig. 19 (location) and Fig. 20 (activity) evaluate the estimation errors when 2 sources were detected for a single source scenario. The average location error in Fig. 19 is around 50 m for the ghost source which means that random points across the 100×100 m field are being detected at each second for the location of the source. Also, Fig. 20 shows that the ghost source is estimated as a low rate source. The location error after correcting the number of sources is also plotted in Fig. 19. It is interesting see that the location error of the actual source in the over fitted-case is similar to the correctly fitted case.

The problem of source detection and mode selection is subject of future work. The results however suggest that we may be able to infer the number of sources from the resulting solution. First, a low estimate of the source rate can indicate over-fitting. A simple likelihood-ratio test for the presence of a point source vs. background can be conducted at these estimated locations. Second, if the location estimate wildly fluctuates across observations we may detect the presence of a ghost source.

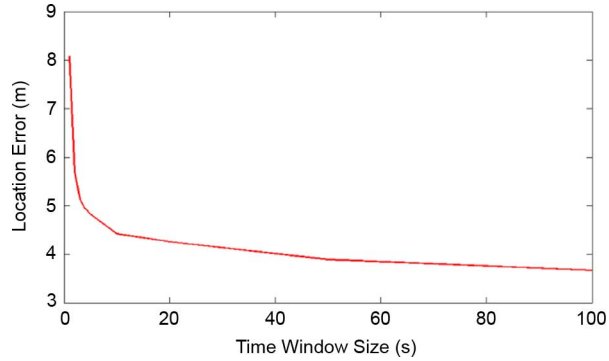

Fig. 21. Average location error of the tracking algorithm as a function of the window size in a 100×100 m^2 area with 9 sensors and one 1 mCi source.

Finally, we present the average performance of the trajectory estimation algorithm in Fig. 21 which plots the average location error between the estimated and true trajectory of the moving source with different time windows. For the plot, we considered a 1 mCi source with random motion direction and speed in the range of 0–2 m/s. The maximum velocity was assumed to be 2 m/s corresponding to typical pedestrian motion. The result shows that we get significant improvements as the time window is increased from 1 s to 20 s after which it flattens out.

VI. Conclusions

In this paper we propose computationally scalable algorithms for locating multiple sources and estimating the trajectory of a maneuvering target. Both these problems present estimation scenarios with high dimensionality where trivial approaches may prove intractable. While Bayesian approaches can potentially provide lower variance in estimation, practical implementation through sampling based approximations are limited by computational complexity and availability of informative prior distribution of parameters.

Our approach is based on iterative gradient based optimization of the likelihood functions to find the MLE. The iterative MLE algorithm has a *power law* scaling while the Bayesian estimation has *exponential* scaling. In practice, gradient based approaches suffer from convergence issues if a good initial estimate is not available to start the iterations. We solve this problem by proposing an EM algorithm to estimate the approximate distribution of source activity in a monitored region. The proposed approaches in this paper offer low complexity alternatives to a Bayesian approach. A comprehensive algorithm comparison in terms of estimation performance and computation complexity is planned as future work.

References

[1] R. J. Nemzek, J. S. Dreicer, and D. C. Torney, "Distributed sensor networks for detection of mobile radioactive sources," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 4, pp. 1693–1700, Aug. 2004.

[2] D. L. Stephens and A. J. Peurrung, "Detection of moving radioactive sources using sensor networks," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 5, pp. 2273–2278, Oct. 2004.
[3] S. V. NageswaraRao, J. C. Chin, D. K. Yau, S. Srivathsan, S. S. Iyengar, Y. Yang, J. C. Hou, X. Xu, and S. Sahni, "Identification of low-level point radiation sources using a sensor network," in *Proc. Int. Conf. Inf. Process. in Sensor Networks*, 2008.
[4] S. V. NageswaraRao, J. C. Chin, D. K. Y. Yau, Y. Yang, J. C. Hou, X. Xu, and S. Sahni, "Localization under random measurements with application to radiation sources," in *Proc. Int. Conf. Inf. Fusion*, 2008.
[5] J. W. Howse, L. O. Ticknor, and K. R. Muske, "Least squares estimation techniques for position tracking of radioactive sources," *Automatica*, vol. 37, no. 11, pp. 1727–17, Nov. 2001.
[6] A. Gunatilaka, B. Ristic, and R. Gailis, "On localization of a radiological point source," *IEEE Inform., Decision, Control*, pp. 236–241, Feb. 2007.
[7] M. Chandy, C. Pilotto, and R. McLean, "Networked sensing system for detecting people carrying radioactive material," in *Proc. 5th Int. Conf. Network Sensing Syst.*, 2008.
[8] R. T. Klann, S. C. de la Barrera, P. L. Vilim, and I. A. Ross, "Tracking of weak radioactive sources in crowded venues," in *Proc. IEEE Nucl. Sci. Symp.*, 2009, pp. 995–1001.
[9] M. Moreland, B. Ristic, and A. Gunatilaka, "Detection and parameter estimation of multiple radioactive sources," in *Proc. Int. Conf. Inf. Fusion*, 2007.
[10] M. Morelande and BrankoRistic, "Radiological source detection and localization using Bayesian techniques," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4220–4231, Nov. 2009.
[11] A. H. Liu, J. J. Bunn, and K. M. Chandy, "Sensor networks for the detection and tracking of radiation and other threats in cities," in *Proc. 10th Int. Conf. Inf. Process. Sensor Networks (IPSN)*, 2011.
[12] S. M. Brennan, A. B. Maccabe, A. M. Mielke, and D. C. Torney, "Radiation detection with distributed sensor networks," *IEEE Comput.*, vol. 37, no. 8, pp. 57–59, Aug. 2004.
[13] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. Series B*, 1977.
[14] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 1, no. 2, pp. 113–122, Oct. 1982.
[15] S. J. Wilderman, N. H. Clinthorne, J. A. Fessler, and W. L. Rogers, "List-mode likelihood reconstruction of Compton scatter camera images in nuclear medicine," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, 2000, pp. 1716–1720.
[16] 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, Dec. 2011.
[17] A. Charnes, E. L. Frome, and P. L. Yu, "The equivalence of generalized least squares and maximum likelihood estimates in the exponential family," *J. Amer. Statist. Assoc. Theory Meth. Section*, vol. 71, no. 353, pp. 214–222, Mar. 1976.
[18] R. B. Vilim, R. T. Klann, and S. C. de la Barrera, "Tracking of weak radioactive sources in crowded venues," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, 2009.
[19] N. Karmarkar, "A new polynomial-time algorithm for linear programming," in *Proc. 16th Annu. ACM Symp. Theory of Comput.- STOC*, 1984.
[20] T. J. Ypma, "Historical development of the Newton-Raphson method," *SIAM Rev.*, vol. 37, no. 4, pp. 531–551, 1995.