Model-Free Stochastic Localization of CBRN Releases[*]

R. Taylor Locke[†] and Ioannis Ch. Paschalidis,[‡] *Senior Member, IEEE*

Abstract—We present a novel two-stage methodology for locating a Chemical, Biological, Radiological, or Nuclear (CBRN) source in an urban area using a network of sensors. In contrast to earlier work, our approach does not solve an inverse dispersion problem but relies on data obtained from a simulation of the CBRN dispersion to obtain probabilistic descriptors of sensor measurements under a variety of CBRN release scenarios. At its first stage, subsequent sensor observations under nominal, CBRN event-free conditions are assumed to be independent and identically distributed and we rely on the method of types to detect a CBRN event. Conditional on such an event, subsequent sensor observations are assumed to follow a Markov process. Using composite hypothesis testing we map sensor measurements to a source location chosen out of a discrete set of possible locations. We leverage large deviation techniques to obtain a bound on the localization probability of error and propose several methodologies for fusing sensor data to arrive at a localization decision, including a distributed one. We also address the problem of optimally placing sensors to minimize the localization probability of error. Our techniques are validated numerically using two different CBRN release simulators.

Index Terms—Source detection, source localization, composite hypothesis testing, large deviations, optimization, sensor placement.

I. INTRODUCTION

CONCERN regarding *Chemical, Biological, Radiological, or Nuclear (CBRN)* terrorism is steadily increasing. This is due to gains in the technological capabilities of producing existing and potentially new, more lethal CBRN agents and delivery mechanisms. Urban areas are of particular concern since these areas tend to have large population densities and are centers for large-scale commerce and politics [1].

In this context, it becomes critical to detect and locate a CBRN source. Early detection and accurate localization enable effective emergency response. Localization is important even for weak CBRN sources as it can help prevent a harmful release and facilitate investigating the origins of the source (forensics).

Invariably, all existing approaches rely on a network of sensors deployed in the urban area and able to measure CBRN particulate concentrations. They differ however, in how they process the sensor data. A related problem is that of placing a given set of sensors so as to optimize a performance metric related to detection and localization.

Earlier work (reviewed later in more detail) has mostly concentrated on solving a difficult inverse dispersion problem in order to locate the source. Such an approach requires an explicit model of CBRN dispersion and, typically, simple analytical models are being assumed to render the inverse problem tractable. Yet, simple dispersion models are not accurate in urban areas for a variety of reasons. First, a typical city tends to be characterized by irregular geometry, many different types of structures and buildings (each with its own surface texture), and highly dynamic population fluxes, thus, giving rise to an *inhomogeneous* dispersion terrain. Second, weather patterns, which drive CBRN dispersion, are highly variable and interact with the various city structures to create complex phenomena (micro-climate effects, urban canyons, etc.). All this complexity can not be captured analytically and, coupled with the inherent uncertainty, suggests a stochastic approach.

In this paper we develop such an approach that uses *simulation* of CBRN dispersion under a variety of source locations and dispersion scenarios to develop probabilistic descriptors of sensor measurements. Localization is performed solely through consultation of the sensor measurements, not through an explicit physical dispersion model. Hence, we describe our approach as *model-free*. Our approach consists of two stages: (a) detection, and (b) localization. To detect a source we assume that subsequent measurements at any given sensor are independent and identically distributed (iid). This is justified since, in

† R. T. Locke is with the Division of Systems Engineering, Boston University, 15 Saint Mary's Street, Brookline, MA 02446, USA, and with the MIT Lincoln Laboratory, 244 Wood Street, Lexington, MA 02421, e-mail: rtl@bu.edu.

‡ I. Ch. Paschalidis is with the Department of Electrical and Computer Engineering and the Division of Systems Engineering, Boston University, 8 Saint Mary's Street, Boston, MA 02215, USA, e-mail: yannisp@bu.edu, url: http://ionia.bu.edu/.

the absence of a source, sensor measurements are driven by noise. Using the method of types ([2], [3]) we compare sequences of measurements with the probabilistic descriptors derived from the simulation to determine the presence of a source. We develop a rigorous detection test and show asymptotic Neyman-Pearson optimality.

Positive source detection triggers the 2nd – localization – stage of our approach. During that stage, and conditional on the presence of particulates from a source, subsequent measurements at a sensor are dependent. We make a Markovian assumption and develop an approach to compare measurements from the sensors to the probabilistic descriptors obtained from the simulation. Localization is formulated as a *composite hypothesis testing problem*. We devise several versions of a localization algorithm – a distributed, a centralized, and a hybrid – using different forms of data fusion from the multiple sensors. We derive a bound on the probability of error by extending our earlier work from [4], which also considered a localization problem but in a very different application domain. Further, we use this bound to formulate the problem of placing sensors so as to minimize the probability of error. As our numerical results demonstrate such an optimization can lead to a dramatic improvement in the error probability compared to ad-hoc sensor placement. The end result of our approach is a placement of the sensors and an *asymptotic guarantee* on the corresponding probability of error.

One will notice some similarities between the theoretical material presented in [4] and in this paper. We recast theorems and definitions here in a new application both for the sake of completeness as well as to provide background for our (Markovian) extensions to the previous (iid) work.

We summarize our contributions below:

1) We propose a new paradigm for CBRN source detection and localization that bypasses many of the problems existing "inverse dispersion" approaches face in an inhomogeneous and uncertain urban setting.
2) We extend our earlier asymptotic performance analysis of composite hypothesis testing in [4] from the iid to a Markovian setting.
3) We demonstrate how to place sensors so as to minimize the localization probability of error.
4) We validate our approach using two different, and independently developed, dispersion simulation engines: the Quick Urban & Industrial Complex (QUIC) Dispersion Modeling System [5] and a new simulator based on the *Lattice Boltzmann method (LBM)* that we develop for our purposes. The latter simulator is of independent interest as it is known to be a naturally parallelizable algorithm [6]. This quality places large scenarios

computationally within reach. These simulators simply simulate the physics of dispersion from some point release under given weather conditions and make none of the distributional assumptions our approach posits. Although it is certainly true that QUIC and LBM make certain assumptions and approximations to simplify the physics of dispersion, our approach decouples these assumptions from the process of release detection and source localization by relying solely on measurements obtained from sensors during particulate propagation. As such, our approach naturally extends to any simulator or even to measurements obtained from physical experimentation. Different scenarios will call for different dispersion simulators. If accuracy is less of an issue, one could select a faster, less accurate simulator. If the area under observation contains heavy foliage, a simulator with satisfactory deposition accuracy should be selected. For situations in which particulate deposition is important due to physicochemical properties of the dispersed agent or sensor performance limitations, say, the computationally efficient method of estimating boundary properties presented in [7] could be of use. A set of illustrative numerical results demonstrate excellent performance of our methods. The results show promise for a practical CBRN urban detection and localization system based on this work.

The remainder of the paper is organized as follows. We discuss related work in Section II. We formulate our problem in Section III where we also introduce some of our notation. We deal with the detection stage in Section IV and the localization stage in Section V. The placement problem is discussed in Section VI. The two dispersion simulators and a host of numerical results are presented in Section VII. Final remarks are in Section VIII. The Appendix contains technical proofs of our results.

Notational Conventions: Throughout the paper all vectors are assumed to be column vectors. We use lower case boldface letters to denote vectors and for economy of space we write $\mathbf{x} = (x_1, \ldots, x_R)$ for the column vector $\mathbf{x}$. $\mathbf{x}'$ denotes the transpose of $\mathbf{x}$ and $\mathbf{0}$ the vector of all zeroes. We use upper case boldface letters to denote matrices. Finally, we use $|\mathscr{A}|$ to denote the cardinality of a set $\mathscr{A}$.

II. RELATED WORK

A sequential measurement based hypothesis testing paradigm for CBRN release detection is presented in [8]. There, the authors make the assumption that concentration samples adhere to a zero-mean Gaussian error and formulate a maximum likelihood test used for release

detection. Our detection approach is similar in that it is a hypothesis test. However, we make no assumptions on the error a CBRN sensor makes, opting instead to learn a probabilistic law of nominal, release-free sensor observations. Whenever a sensor's observations sufficiently depart, according to a Hoeffding test, from the known nominal behavior, we declare that a CBRN release has occurred.

Existing CBRN source localization approaches ([9], [10], [11]) observe CBRN agent concentrations and solve the inverse problem of tracing dispersion backward in time and space to the source of the release. As discussed, limitations of this methodology stem from the irregular and dynamic phenomena typically found in urban areas. In [9] the presence of challenging geographies and wind turbulence is accommodated by incorporating Monte Carlo simulation of fluid dispersion. However, all of these works suffer from the difficulty of determining, without a detection process, the point in time in which the event started. Barring this information, these inverse problem approaches are vulnerable to erroneous localizations.

Simulation-aided localization of contaminant releases on a continental scale is presented in [12]. In [13], the inverse problem is replaced by computing the posterior probabilities of a finite set of release locations and selecting the release location with maximal likelihood. The computed likelihoods are based on assumptions on both the distribution of concentration observations at sensor locations and a particular dispersion model. Our approach is similar to these, in that localization is based entirely on sensor observations and probabilistic descriptors of CBRN releases obtained from simulation, with the important difference, however, that the only assumptions we make on concentration observations under release conditions is that they adhere to a first-order Markov chain.

While our approach does depend on numerical or physical simulation of agent dispersion to build a statistical understanding of CBRN release evolution, once this understanding is established the actual dispersion model is no longer used. Hence, we refer to our localization approach as model-free. Additionally, while we employ two specific dispersion simulators in the numerical evaluation in Section VII, any dispersion simulation, numerical or physical, can be used. We suggest using the simulation scheme that provides the most accurate depiction of particulate releases for the particular application. This is even more important in applications where turbulence plays a greater role in particulate displacement.

Our localization approach is inspired by [4]. There, a wireless sensor deployed in an indoor environment is located through sequential hypothesis testing by using

signal strength measurements corresponding to packets transmitted by the sensor and received by a set of stationary clusterheads. [4] also solves the problem of optimally placing clusterheads to reduce erroneous localizations. The key difference between that work and the work presented here, aside from vastly different application areas, is that subsequent observations made by the clusterheads in the wireless sensor localization problem can be considered iid. Under a CBRN release scenario, if a CBRN sensor observes large particle concentrations it is likely to observe large particle concentrations in its next sample. To that end, in this work we extend the theory presented in [4] to allow for dependencies in subsequent sensor observations.

Our optimal sensor placement approach also stems from the work presented in [4], which poses clusterhead placement as an optimization problem whose objective is to minimize the probability of localization error. This differs from the more traditional sensor placement methodologies [14], [15], whose objectives are the facilitation and maintenance of general observability under the possibility of sensor faults. To our knowledge there are no other existing CBRN sensor placement techniques that are motivated by analytically derived localization error probabilities.

III. PROBLEM FORMULATION

Consider a CBRN sensor network deployed in an urban setting. Due to the nature of plume dispersion and existing CBRN attack response techniques, extreme precision on the locations of releases and sensors is not needed. Rather, general locations (e.g., corner of X St. and Y Ave.) suffice for disaster avoidance and response measures. It naturally follows that release and potential sensor locations can be discretized to conform to potentially irregular grids. In the following, we assume N possible CBRN release scenarios represented by the set $\mathscr{L} = \{L_1, \ldots, L_N\}$ and M possible CBRN sensor locations represented by the set $\mathscr{B} = \{B_1, \ldots, B_M\}$. For applications in which greater precision is required, the set of discrete locations can be extended to include more locations. However, this accrues the cost of requiring more simulations. The elements of $\mathscr{L}$ correspond to a release location and a set of release associated characteristics potentially including wind direction, variability in wind speed, and release mass, among others. Wind, in particular, is a critical component of a release scenario as it is the primary mode of particulate transport. For a release scenario L_j, the corresponding set of release characteristics will be denoted $\Omega_j = \{\theta_j^1, \ldots, \theta_j^{|\Omega_j|}\}$.

Let $\mathbf{y}^k$ denote the vector of sensor observations made by a sensor at location B_k. Each sensor outputs a CBRN agent concentration estimation and potentially local wind observations. Since different concentrations result in

different levels of casualties, counter measures, and responses, observed concentrations can naturally be discretized to concentration levels (e.g., $[0, x], [x, y_{LCT50}]$, etc., where y_{LCT50} represents a concentration that would lead to an accumulated dosage that produces a 50% chance of survival). Each $\mathbf{y}^k$ is mapped to a symbol in a finite alphabet denoted by $\Sigma = \{\sigma_1, \ldots, \sigma_{|\Sigma|}\}$. A sequence of n sensor observations at location B_k is denoted $\mathbf{y}^{k,n} = (\mathbf{y}_1^k, \ldots, \mathbf{y}_n^k)$. It is assumed throughout the sequel that an initial state of $\mathbf{y}_0^k$ is known a priori.

While not under an attack scenario, a sensor's agent concentration approximation fluctuations are attributed to measurement noise. Accordingly, when there is no CBRN attack, the sensor's readings can be assumed to be iid.

Under a release scenario, however, the agent disperses gradually, resulting in concentration observations that change much slower than a sensor's sampling rate. That is, if a certain concentration level is observed by a sensor at a certain time step, one can expect to see a similar concentration observation in the next time step. To capture this dependency, we model concentration level at a sensor when there is an attack as a first-order Markov chain.

The first problem we wish to solve is to detect a CBRN event, given a set of observations $\{\mathbf{y}^{1,n}, \ldots, \mathbf{y}^{K,n}\}$ from K sensors placed at locations $B_1, \ldots, B_K$. Once a CBRN event is identified, the next problem is to determine where the event originated. We also wish to solve the related problem of selecting the locations from $\mathscr{B}$ where we can place K sensors (with $K < M$) such that we minimize the probability of localization error. The solution to this last problem is accomplished through analysis of theoretical performance guarantees.

IV. RELEASE DETECTION

In this section we describe our approach to the problem of release detection. We first construct empirical probability laws for the vector $\mathbf{y}^k$ under nominal conditions and then compare these laws with empirical probability measures based on in situ sensor observations. When a measure differs significantly from the established probability laws, a release is declared and the localization engine is triggered.

Allow $M_1(\Sigma)$ to denote the set of all probability vectors with support defined by the state alphabet $\Sigma = \{\sigma_1, \ldots, \sigma_{|\Sigma|}\}$. Consider the empirical probability measure of a sequence $\mathbf{y}^{k,n} = (\mathbf{y}_1^k, \ldots, \mathbf{y}_n^k)$,

$$\mathscr{E}^{\mathbf{y}^{k,n}}(\sigma_r) = \frac{1}{n} \sum_{t=1}^{n} \mathbf{1}\{\mathbf{y}_t^k = \sigma_r\}, \quad r = 1, \ldots, |\Sigma|, \ (1)$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function. Let $\mathscr{E}^{\mathbf{y}^{k,n}} = (\mathscr{E}^{\mathbf{y}^{k,n}}(\sigma_1), \ldots, \mathscr{E}^{\mathbf{y}^{k,n}}(\sigma_{|\Sigma|}))$. First we construct the empirical probability law $\boldsymbol{\mu}^k \in M_1(\Sigma)$ for all sensor

locations $B_k \in \mathscr{B}$ from observed sequences $\mathbf{y}^{k,N} = (\mathbf{y}_1^k, \ldots, \mathbf{y}_N^k)$ with sufficiently large N by setting $\boldsymbol{\mu}^k = \mathscr{E}^{\mathbf{y}^{k,N}}$.

When the sensor network is operational, a sequence $\mathbf{y}^{k,n} = (\mathbf{y}_1^k, \ldots, \mathbf{y}_n^k)$ is observed for some $n \ll N$. An empirical probability measure $\boldsymbol{\nu}_n^k = \mathscr{E}^{\mathbf{y}^{k,n}}$ is constructed via (1). To determine if the observations in $\mathbf{y}^{k,n}$ are anomalous, we employ the so-called *method of types*. The following theorem from [3] is useful.

Theorem IV.1 *For every $\boldsymbol{\nu} \in M_1(\Sigma)$ let*

$$I_1(\boldsymbol{\nu}) = H(\boldsymbol{\nu}|\boldsymbol{\mu}^k),$$

where $H(\boldsymbol{\nu}|\boldsymbol{\mu}^k)$ is the relative entropy of the probability vector $\boldsymbol{\nu}$ with respect to $\boldsymbol{\mu}^k$:

$$H(\boldsymbol{\nu}|\boldsymbol{\mu}^k) = \sum_{r=1}^{|\Sigma|} \boldsymbol{\nu}(\sigma_r) \log \frac{\boldsymbol{\nu}(\sigma_r)}{\boldsymbol{\mu}^k(\sigma_r)}.$$

Then, for any set Γ of probability vectors in $M_1(\Sigma)$

$$-\inf_{\boldsymbol{\nu} \in \Gamma^\circ} I_1(\boldsymbol{\nu}) \leq \liminf_{n \to \infty} \frac{1}{n} \log \mathbf{P}_{\boldsymbol{\mu}^k}\left[\mathscr{E}^{\mathbf{y}^{k,n}} \in \Gamma\right] \leq$$

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbf{P}_{\boldsymbol{\mu}^k}\left[\mathscr{E}^{\mathbf{y}^{k,n}} \in \Gamma\right] \leq -\inf_{\boldsymbol{\nu} \in \Gamma} I_1(\boldsymbol{\nu}),$$

where Γ° denotes the interior of Γ.

Theorem IV.1 rigorously identifies a distance metric that can be used in what is known as the Hoeffding's rule [3]:

$$\mathscr{S}(\mathbf{y}^{k,n}) = \begin{cases} \text{no attack,} & \text{if } I_1(\mathscr{E}^{\mathbf{y}^{k,n}}) < \eta, \\ \text{attack,} & \text{otherwise.} \end{cases} \quad (2)$$

Here, η is a threshold that can be determined as $\eta = -\log \frac{\epsilon}{n}$, where ϵ is a tolerable false alarm rate [4].

In order to assess the validity and optimality of tests of the form (2), we need to define what notion of optimality we hope to attain. One often adopted measure of any decision test's optimality is to consider the likelihood of making an error. The two types of error the Hoeffding test in (2) can make are: declaring an anomaly when in fact the activity recorded in $\mathbf{y}^{k,n}$ was nominal (Type-I error), or not declaring an anomaly when $\mathbf{y}^{k,n}$ departs from nominal behavior (Type-II error). We denote the probabilities of these two errors by $\alpha_{k,n}$ and $\beta_{k,n}$, respectively.

A particular challenge arises when trying to evaluate $\beta_{k,n}$. While the probability law depicting nominal behavior $\boldsymbol{\mu}^k$ can be used to evaluate $\alpha_{k,n} = \mathbf{P}_{\boldsymbol{\mu}^k}[I_1(\mathscr{E}^{\mathbf{y}^{k,n}}) \geq \eta]$, computation of $\beta_{k,n}$ requires considering all probability vectors from the set $M_1(\Sigma) \setminus \{\boldsymbol{\mu}^k\}$. With this in mind, we introduce the following concept of asymptotic optimality known as the Neyman-Pearson criterion.

Definition 1

A test $\mathscr{S}$ is considered optimal for a given $\eta > 0$ if, among all tests that satisfy

$$\limsup_{n \to \infty} \frac{1}{n} \alpha_{k,n} \leq -\eta$$

it maximizes $-\limsup_{n \to \infty} \frac{1}{n} \beta_{k,n}$ *uniformly over all probability vectors in $M_1(\Sigma) \setminus \{\boldsymbol{\mu}^k\}$.*

The following theorem, thanks to Hoeffding [16] and appearing in [3], establishes the optimality of (2).

Theorem IV.2 *The anomaly detector presented in (2) is optimal under Definition 1.*

The effect of the first stage conclusion is two-fold. First, it establishes a reasonable indication that a CBRN event has actually happened and that release localization analysis is within order. In addition, it provides an estimation on the time of the CBRN release. In practice, the time of attack is taken to be the time at which the first high concentration observation within $\mathbf{y}^{k,n}$ is sampled. This is an important contribution to the second stage problem, as knowledge of the time of attack decreases the amount of parametrizations in Ω_j that need to be considered. For localization, the window of observations beginning at the time of the release and of length determined by an ideal response time is used.

V. Source Localization

Allow $M_2(\Sigma \times \Sigma)$ to denote the set of all discrete time Markov transition probability matrices on the state alphabet $\Sigma = \{\sigma_1, \ldots, \sigma_{|\Sigma|}\}$. For each sensor location and release scenario pair (B_k, L_j), associate a series of first-order Markov transition probability matrices $\boldsymbol{\Pi}^k_{\theta_j}$, for all $\theta_j \in \Omega_j$, defined as

$$\boldsymbol{\Pi}^k_{\theta_j} = \left\{ \pi^k_{\theta_j}(\sigma_v, \sigma_u) \right\}^{|\Sigma|}_{u,v=1}, \tag{3}$$

where $\pi^k_{\theta_j}(\sigma_v, \sigma_u) = \mathbf{P}^k_{\theta_j}[\mathbf{y}^k_{t+1} = \sigma_u | \mathbf{y}^k_t = \sigma_v]$. These matrices depict the CBRN particulate concentration state evolution over all considered parametrizations of release scenarios and available sensor locations. Using (3), the probability of observing a sequence of sensor observations $\mathbf{y}^{k,n}$ can be expressed as

$$P^k_{\theta_j}(\mathbf{y}^{k,n}) = p^k_{\theta_j}(\mathbf{y}^k_0) \prod^n_{t=1} \pi^k_{\theta_j}(\mathbf{y}^k_{t-1}, \mathbf{y}^k_t), \tag{4}$$

where $p^k_{\theta_j}(\cdot)$ denotes the stationary probability distribution on the observed concentration states at sensor location B_k under release parametrization θ_j of scenario L_j.

Under the assumption of a Markov source, the empirical measure of a sequence $\mathbf{y}^n = (\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n)$ of sensor observations (note we now assume that there is a known initial concentration state $\mathbf{y}_0$) takes the form

$$\mathscr{E}^{\mathbf{y}^n}_2(\sigma_v, \sigma_u) = \frac{1}{n} \sum^n_{i=1} \mathbf{1}\{\mathbf{y}_{i-1} = \sigma_v, \mathbf{y}_i = \sigma_u\}. \tag{5}$$

For any release location and release scenario pair (B_k, L_j) we can observe $\mathbf{y}^{k,N}$ under release conditions θ_j with sufficiently large N and obtain a (Markovian) empirical probability law as follows:

$$\boldsymbol{\Pi}^k_{\theta_j} = \left\{ \mathscr{E}^{\mathbf{y}^{k,N}}_2(\sigma_u | \sigma_v) \right\}_{\sigma_u, \sigma_v \in \Sigma},$$

where $\mathscr{E}^{\mathbf{y}^n}_2(\sigma_u | \sigma_v)$ are the empirical transition probabilities observed in $\mathbf{y}^n$, defined as

$$\mathscr{E}^{\mathbf{y}^n}_2(\sigma_u | \sigma_v) = \frac{\mathscr{E}^{\mathbf{y}^n}_2(\sigma_v, \sigma_u)}{\sum^{|\Sigma|}_{r=1} \mathscr{E}^{\mathbf{y}^n}_2(\sigma_v, \sigma_r)}.$$

The empirical distribution $\boldsymbol{\mathscr{E}}^{\mathbf{y}^n}_2$ is said to be *shift invariant* if the empirical marginal probabilities of each state are equal. That is, $\boldsymbol{\mathscr{E}}^{\mathbf{y}^n}_2$ is shift invariant if

$$\mathscr{E}^{\mathbf{y}^n}_2(\sigma_v) = \sum^{|\Sigma|}_{u=1} \mathscr{E}^{\mathbf{y}^n}_2(\sigma_v, \sigma_u) = \sum^{|\Sigma|}_{u=1} \mathscr{E}^{\mathbf{y}^n}_2(\sigma_u, \sigma_v).$$

The next theorem, a version of Sanov's Theorem for Markov chains proven in [3], provides a convex relative entropy measure for Markov chains.

Theorem V.1 *For every shift invariant $\mathbf{Q} \in M_2(\Sigma \times \Sigma)$ let*

$$I_2(\mathbf{Q}|\boldsymbol{\Pi}) = \sum^{|\Sigma|}_{v=1} q(\sigma_v) H_2(q(\cdot|\sigma_v)|\pi(\sigma_v, \cdot)),$$

where H_2 is the relative entropy

$$H_2(q(\cdot|\sigma_v)|\pi(\sigma_v, \cdot)) = \sum^{|\Sigma|}_{u=1} q(\sigma_u | \sigma_v) \log \frac{q(\sigma_u | \sigma_v)}{\pi(\sigma_v, \sigma_u)}.$$

Then for any set $\Gamma \subset M_2(\Sigma \times \Sigma)$ of shift invariant Markov sources

$$- \inf_{\mathbf{Q} \in \Gamma^\circ} I_2(\mathbf{Q}|\boldsymbol{\Pi}^k_{\theta_j}) \leq \liminf_{n \to \infty} \frac{1}{n} \log \mathbf{P}^k_{\theta_j}\left[\boldsymbol{\mathscr{E}}^{\mathbf{y}^{k,n}}_2 \in \Gamma \right] \leq$$

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbf{P}^k_{\theta_j}\left[\boldsymbol{\mathscr{E}}^{\mathbf{y}^{k,n}}_2 \in \Gamma \right] \leq - \inf_{\mathbf{Q} \in \Gamma} I_2(\mathbf{Q}|\boldsymbol{\Pi}^k_{\theta_j}),$$

where Γ° denotes the interior of Γ and $\mathbf{P}^k_{\theta_j}$ is a probability evaluated assuming that $\mathbf{y}^{k,n}$ is drawn according to $\boldsymbol{\Pi}^k_{\theta_j}$.

Consider now the subproblem of using the sensor readings from a sensor at location B_k to determine if the CBRN release corresponds to scenario L_i or L_j. The

Generalized Likelihood Ratio Test (GLRT) compares the normalized generalized log-likelihood ratio

$$X_{ijk}(\mathbf{y}^{k,n}) = \frac{1}{n} \log \frac{\sup_{\theta_i \in \Omega_i} P_{\theta_i}^k(\mathbf{y}^{k,n})}{\sup_{\theta_j \in \Omega_j} P_{\theta_j}^k(\mathbf{y}^{k,n})}$$

to a threshold λ and declares L_i to be the release scenario whenever

$$\mathbf{y}^{k,n} \in \mathscr{S}_{ijk,n}^{GLRT} = \{\mathbf{y}^n | X_{ijk}(\mathbf{y}^n) \geq \lambda\}, \qquad (6)$$

and L_j otherwise. Selection of the value for λ in (6) is performed through evaluation of decision error probabilities and is discussed later in this section.

The GLRT described in (6) can make two types of error. Namely, the decision test can declare L_j to be the release scenario when in actuality the release scenario was L_i and vice versa. The probabilities of these errors are represented as,

$$\alpha_{ijk,n}^{GLRT}(\theta_j) = \mathbf{P}_{\theta_j}^k[\mathbf{y}^{k,n} \in \mathscr{S}_{ijk,n}^{GLRT}],$$
$$\beta_{ijk,n}^{GLRT}(\theta_i) = \mathbf{P}_{\theta_i}^k[\mathbf{y}^{k,n} \notin \mathscr{S}_{ijk,n}^{GLRT}].$$

We will refer to the quantities

$$\limsup_{n \to \infty} \frac{1}{n} \log \alpha_{ijk,n}^{GLRT}(\theta_j),$$
$$\limsup_{n \to \infty} \frac{1}{n} \log \beta_{ijk,n}^{GLRT}(\theta_i),$$

as the exponents of the Type-I and Type-II error probabilities, respectively.

Certainly, some values of λ in (6) are better than others. To evaluate the GLRT's performance, we will use a generalized version of Definition 1, which we will call the generalized Neyman-Pearson criterion.

Definition 2
The decision rule $\{\mathscr{S}_{ijk,n}\}$ is optimal if it satisfies

$$\limsup_{n \to \infty} \frac{1}{n} \log \alpha_{ijk,n}^{\mathscr{S}}(\theta_j) < -\lambda, \quad \forall \theta_j \in \Omega_j \quad (7)$$

and maximizes $-\limsup_{n \to \infty} \frac{1}{n} \log \beta_{ijk,n}^{\mathscr{S}}(\theta_i)$ *uniformly over all $\theta_i \in \Omega_i$.*

The following theorem, whose proof is in Appendix A, establishes the optimality of the GLRT according to Definition 2.

Theorem V.2 *The GLRT with a threshold λ is optimal under Definition 2 if and only if*

$$\inf_{\mathbf{Q} \in \mathscr{C}_{ijk}} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k) \geq \inf_{\mathbf{Q} \in \mathscr{A}_{ijk}} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k), \quad \forall \theta_i \in \Omega_i,$$
(8)

where $\mathscr{A}_{ijk} = \{\mathbf{Q}| \inf_{\theta_j \in \Omega_j} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_j}^k) < \lambda\}$ *and* $\mathscr{C}_{ijk} = \{\mathbf{Q}| \inf_{\theta_j \in \Omega_j} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_j}^k) - \inf_{\theta_i \in \Omega_i} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k) < \lambda \leq \inf_{\theta_j \in \Omega_j} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_j}^k)\}$.

Furthermore, assuming (8) is satisfied

$$\limsup_{n \to \infty} \frac{1}{n} \log \alpha_{ijk,n}(\theta_j)^{GLRT} \leq -\lambda,$$

for all $\theta_j \in \Omega_j$, and

$$\limsup_{n \to \infty} \frac{1}{n} \log \beta_{ijk,n}^{GLRT}(\theta_i) \leq - \inf_{\mathbf{Q} \in \mathscr{A}_{ijk}} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k),$$

for all $\theta_i \in \Omega_i$.

Now, in the event that the GLRT is not asymptotically optimal (i.e., condition (8) does not hold) information on the error probabilities can still be gleamed. Define the set $\mathscr{D}_{ijk} = \{\mathbf{Q}| \inf_{\theta_j \in \Omega_j} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_j}^k) - \inf_{\theta_i \in \Omega_i} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k) < \lambda\}$. Note $\mathbf{y}^{k,n} \in \mathscr{S}_{ijk,n}^{GLRT}$ if and only if $\mathscr{E}_2^{\mathbf{y}^{k,n}} \notin \mathscr{D}_{ijk}$. From Theorem V.1 it follows:

$$\limsup_{n \to \infty} \frac{1}{n} \log \alpha_{ijk,n}^{GLRT}(\theta_j) \qquad (9)$$
$$\leq - \inf_{\mathbf{Q} \notin \mathscr{D}_{ijk}} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_j}^k), \quad \forall \theta_j \in \Omega_j$$

$$\limsup_{n \to \infty} \frac{1}{n} \log \beta_{ijk,n}^{GLRT}(\theta_i) \qquad (10)$$
$$\leq - \inf_{\mathbf{Q} \in \mathscr{D}_{ijk}} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k), \quad \forall \theta_i \in \Omega_i.$$

In addition, we have the following lemma, proven in Appendix B.

Lemma V.3 *If the GLRT is not optimal, the constraint*

$$\limsup_{n \to \infty} \frac{1}{n} \log \alpha_{ijk,n}^{GLRT}(\theta_j) \leq -\lambda, \quad \forall \theta_j \in \Omega_j$$

still holds.

We can therefore determine an asymptotic bound on the exponent of the Type-II error probability by the nonlinear program

$$Z_{ijk}(\lambda, \theta_i) = \begin{array}{ll} \min_{\mathbf{Q}} & I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k) \\ \text{s.t.} & g(\mathbf{Q}; \Omega_i, \Omega_j) \leq \lambda, \\ & \mathbf{Aq} \leq \mathbf{0}, \end{array} \quad (11)$$

where $g(\mathbf{Q}; \Omega_i, \Omega_j) = \min_{\theta_j \in \Omega_j} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_j}^k) - \min_{\theta_i \in \Omega_i} I_2(\mathbf{Q}|\mathbf{\Pi}_{\theta_i}^k)$. The constraints enforcing $\mathbf{Q}$ to be shift invariant and correspond to valid transition probabilities for a Markov chain are linear and are represented by the system of linear inequalities $\mathbf{Aq} \leq \mathbf{0}$, where $\mathbf{q}$ is a "vectorization" of $\mathbf{Q}$.

Using (11), $Z_{ijk}(\lambda) = \min_{\theta_i \in \Omega_i} Z_{ijk}(\lambda, \theta_i)$ is the exponent of the worst case Type-II error probability of the GLRT (i.e., declaring the release scenario is L_j when the actual release scenario is L_i) when we use information from sensor k. $Z_{ijk}(\cdot)$ is a non-increasing, non-negative function, with $\lim_{\lambda \to \infty} Z_{ijk}(\lambda) = 0$. Therefore, there must exist $\lambda_{ijk}^* \geq 0$ such that $Z_{ijk}(\lambda_{ijk}^*) = \lambda_{ijk}^*$, which represents the exponent of equal Type I and Type II error probabilities in the GLRT. Since there is no guarantee

that $\lambda_{ijk}^* = \lambda_{jik}^*$ (i.e., information from a sensor at location B_k does not provide symmetric decision making via the GLRT), we define

$$d_{ijk} = \max\{\lambda_{ijk}^*, \lambda_{jik}^*\}, \qquad (12)$$

and this represents a symmetric error probability exponent. In particular, if λ_{ijk}^* is the maximizer in (12) we let $(\bar{i}, \bar{j}) = (i, j)$, otherwise we let $(\bar{i}, \bar{j}) = (j, i)$. In either case, we use the GLRT test that compares $X_{\bar{i}\bar{j}k}(\mathbf{y}^{k,n})$ to d_{ijk}. The above leads to the following proposition.

Proposition V.4 *Suppose the sensor at B_k uses the GLRT and compares $X_{\bar{i}\bar{j}k}(\mathbf{y}^{k,n})$ to d_{ijk}. Then, the maximum probability of error satisfies*

$$\limsup_{n \to \infty} \frac{1}{n} \log P_{ijk,n}^{(e)} \le -d_{ijk},$$

where $P_{ijk,n}^{(e)}$ denotes the maximum probability of error when deciding, via the GLRT, between release scenarios L_i and L_j using information from a sensor placed at location B_k.

Thus, for every sensor location in $\mathscr{B}$ we have an upper bound on the GLRT's probability of error when deciding between any two scenarios in $\mathscr{L}$. This result will prove useful when it comes to selecting locations to place a limited number of sensors.

A. Fusing information from multiple sensors

So far all of the results presented for the source localization problem considered the binary GLRT in (6) when the real problem is to select a location that appears within the scenarios in $\mathscr{L}$. One approach is to make a series of decisions of the form (6) until a single release scenario has been chosen.

Assume, without loss of generality, that the sensors are placed at locations $B_1, \ldots, B_K$. Using the observations from sensor B_{k^*}, where $k^* = \operatorname{argmax}_k d_{12k}$, a decision is made between scenarios L_1 and L_2 via the GLRT with threshold d_{12k^*}, as established in (6) and (12). This scenario is then compared with L_3 by the GLRT, and so on, until $N - 1$ decisions have been made and only one scenario from $\mathscr{L}$ is accepted. The location associated with this selected scenario is declared the CBRN source location. We will refer to this localization paradigm as the *sequential GLRT approach*.

Alternatively, localization can be performed using all sensor information in unison by selecting the location with greatest likelihood in a single decision. We will refer to this decision paradigm as the *centralized maximum likelihood approach*. Unfortunately, the amount of information needed to fully fit a complete joint probability mass function for all sensor observation combinations makes it practically impossible. Instead,

joint log-likelihood values under scenario L_j can be approximated via (4) and

$$P(L_j) = \sum_{k=1}^{K} \log \sup_{\theta_j \in \Omega_j} P_{\theta_j}^k(\mathbf{y}^{k,n}). \qquad (13)$$

Localization is then performed by selecting the location associated with $L^* \in \mathscr{L}$ which maximizes the value in (13).

The sequential GLRT and centralized maximum likelihood localization techniques can be combined. In what we will call the *hybrid approach* to localization, a sequence of decisions is made, as in the sequential GLRT approach, but each decision employs information from more than one sensors. Assume, again, without loss of generality, that sensors are placed at locations $B_1, \ldots, B_K$. Using observations from m sensors at locations $B_{k_1^*}, \ldots, B_{k_m^*}$ corresponding to the m largest values of d_{12k}, $k = 1, \ldots, K$, scenario L_1 is declared if

$$P(L_1) = \sum_{t=1}^{m} \log \sup_{\theta \in \Omega_1} P_{\theta}^{k_t^*}(\mathbf{y}^{k_t^*, n})$$

$$> P(L_2) = \sum_{t=1}^{m} \log \sup_{\theta \in \Omega_2} P_{\theta}^{k_t^*}(\mathbf{y}^{k_t^*, n}),$$

and L_2 is declared otherwise. The chosen scenario from this decision is then compared to scenario L_3 in the same manner. This process continues until only one release scenario remains. In the extreme case in which only one sensor is used in each test within the sequence, this approach is equivalent to the sequential GLRT with every threshold set to 0. In the other extreme case, in which every deployed sensor is used in each test within the sequence, this approach is equivalent to the centralized maximum likelihood approach.

VI. SENSOR PLACEMENT

Now that performance guarantees have been established for each sensor location in $\mathscr{B}$, the question of where to place $K < M$ sensors can be addressed. Using the bound in Proposition V.4, we can select the K elements of $\mathscr{B}$ such that the worst case error when differentiating between any two release scenarios in $\mathscr{L}$ is minimized. This is done, as in [4], via the mixed integer linear program (MILP)

$$\begin{aligned}
\max \quad & \epsilon \\
\text{s.t.} \quad & \sum_{k=1}^{M} x_k = K, \\
& \sum_{k=1}^{M} y_{ijk} = 1, \ i, j = 1, \ldots, N, i < j, \\
& y_{ijk} \le x_k, \ \forall i, j, i < j, k = 1, \ldots, M, \qquad (14) \\
& \epsilon \le \sum_{k=1}^{M} d_{ijk} y_{ijk}, \ \forall i, j, i < j, \\
& y_{ijk} \ge 0, \ \forall i, j, k, i < j, \\
& x_k \in \{0, 1\}, \ \forall k.
\end{aligned}$$

In (14), the x_k are the indicator functions for placement of a sensor at location B_k. Since only K sensors

are available, the number of positive indicator variables is limited in the constraint $\sum_{k=1}^{K} x_k = K$. The constraint $\epsilon \leq \sum_{k=1}^{K} d_{ijk} y_{ijk}$, where d_{ijk} is defined in (12), ensures that for every release scenario pair in $\mathscr{L}$, there exists a d_{ijk} at least as large as ϵ. Although shown to be NP-hard, (14) can be solved efficiently for large sets $\mathscr{B}$ and $\mathscr{L}$ using a special purpose algorithm [17].

Let $\mathscr{Y}$ be any subset of $\mathscr{B}$ with cardinality K. Further let $\mathbf{x}(\mathscr{Y})$ denote the vector of decision variables where $x_k(\mathscr{Y})$ denotes the presence of B_k in $\mathscr{Y}$. Define

$$\epsilon(\mathscr{Y}) = \min_{i<j} \max_{\{k|x_k(\mathscr{Y})=1\}} d_{ijk}.$$

This term simply represents the worst case pair-wise error probability exponent present in the placement solution $\mathscr{Y}$. Allowing $\mathbf{x}^*$, $\mathbf{y}^*$, and ϵ^* to represent the optimal decision variables in (14), we have the following result from [17], but now with d_{ijk}'s corresponding to the Markovian rather than the iid case.

Proposition VI.1 *For any sensor placement $\mathscr{Y}$ we have*

$$\epsilon^* \geq \epsilon(\mathscr{Y}).$$

Moreover,

$$\epsilon^* = \min_{i<j} \max_{\{k|x_k^*=1\}} d_{ijk} \qquad (15)$$

and the optimal solution satisfies

$$y_{ijk}^* = \begin{cases} 1, & \text{if } k = argmax_{\{k|x_k^*=1\}} d_{ijk}, \\ 0, & \text{otherwise}, \end{cases}$$

for all i, j, $i < j$, and all k, where at most one y_{ijk}^* is set to 1 for a given (i,j) pair.

For the sequential GLRT decision making scheme, a performance guarantee for sensor placement determined via solution to (14) is found in the following proposition. The proof is provided in Appendix C.

Proposition VI.2 *Let $\mathbf{x}^*$, y^*, and ϵ^* be an optimal solution to the MILP in (14). Place sensors at locations B_k such that $x_k^* = 1$ and for every (L_i, L_j) scenario pair, select one sensor with index k_{ij}^* such that $y_{ijk_{ij}^*} = 1$. Then, the worst case probability of error for the sequential GLRT decision rule, $P_n^{(e),opt}$, satisfies*

$$\limsup_{n \to \infty} \frac{1}{n} P_n^{(e),opt} \leq -\epsilon^*.$$

VII. NUMERICAL EXPERIMENTATION

To demonstrate the performance of the sensor placement approach of Section VI and the two-stage CBRN release detection and localization of Section III, we simulated several release scenarios in two illustrative environments. In the first environment – the toy environment – CBRN releases were simulated in the *Quick Urban & Industrial Complex (QUIC)* Dispersion Modeling System [5] developed at the Los Alamos National Laboratory. QUIC first solves the fluid dynamics problem of determining local wind eddies throughout a modeled three-dimensional, outdoor setting using the methods of Röckle [18]. Using the fluid flow solution, QUIC simulates the travel of CBRN particulates via a Lagrangian random walk. Previously, the QUIC codes have been tested and validated for real-world situations [5]. A three-dimensional illustration of the toy environment some time after a QUIC simulated point release is found in Figure 1.

Additionally, CBRN releases were simulated using a simulator we developed based on the *Lattice Boltzmann Method (LBM)*. LBM evolved from the numerical fluid modeling technique *Lattice Gas Automata (LGA)*, in which parcels of air adhere to microscopic laws which dictate their movement. Macroscopic values of flow velocities and densities are then derived by the underlying microscopic properties propagated by the algorithm [19]. Unfortunately, LGA often falls victim to instability in the face of statistical noise [20]. LBM extends LGA by considering air parcel movement more notionally by modeling microscopic air parcel velocities as distributions in the *Lattice Boltzmann Equation (LBE)*. It has been shown that under reasonable starting conditions LBM provides accurate approximations to fluid flows. Further, the macroscopic Navier-Stokes equations can be recovered from the microscopic LBE [6].
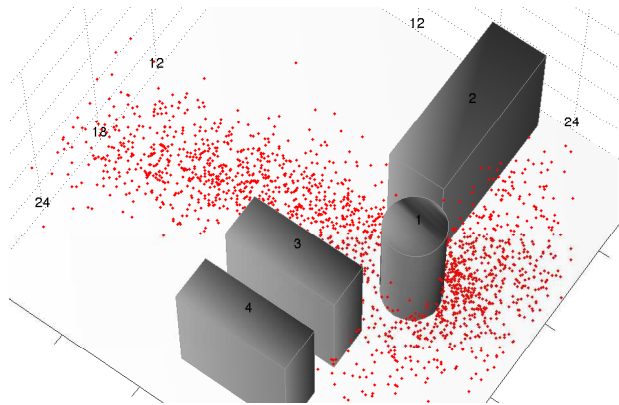

Fig. 1. Toy environment QUIC representation. The dots represent dispersed particulates some time after a point release. The release location is between the cylinder and the top-most block in the figure.

Wind is the main driver of particulate transportation. As such, fluid flow accuracy should be of prime concern when selecting a dispersion simulator. This becomes even more important if the environment of interest contains, as in the scenarios considered here, obstructions

or irregular geometries. These characteristics produce irregularities in the environment's wind field. Both QUIC and LBM accept nominal wind conditions which are translated into varying wind eddies distributed throughout the environment.

It should be noted that Markov chains have been shown to be good approximations of Lagrangian random walk based dispersion models [21], such as the dispersion modeling used by QUIC. One should expect to see our localization approach to perform well on QUIC generated data. Our LBM model, however, is not susceptible to the same ailment.

We stress that these two independently developed simulators make none of the distributional assumptions (Markovian particulate concentrations) our localization methods postulate. They simply consider a point release and simulate the physics of dispersion in the given environment and under a given weather pattern. Hence, the results we produce assess not only the performance of the methods but also the validity of our assumptions.

In the second environment –the city grid environment– CBRN releases were also simulated in QUIC and LBM. This environment consists of geometries more typical of dense urban areas. It was within this environment that we compared the agent concentration profiles of the two different dispersion simulators at a sampler located down wind of the point release. This comparison appears in Figure 2. These models have different discretizations of the three-dimensional model and hence produce concentration values that differ in scale. This has been accommodated in Figure 2 by reporting the percentage of all observed concentrations reported by a single sensor downwind from a release. The LBM simulator produces a much smoother agent concentration evolution than the QUIC simulator. Thus, noise within the evaluation of the proposed methodologies when the LBM data are used is primarily due to our sensor false alarm model rather than the dispersion model. Further dissimilarities between the two concentration profiles also stem from differences in the way surface conditions are treated in the two simulated environments. QUIC employs boundary conditions to reflect interaction with a concrete surface. The LBM representation, on the other hand, uses only the so-called "bounce back" boundary conditions [6].

A. Toy Environment

To demonstrate the performance of the proposed techniques, first a simple three-dimensional environment with several large obstructions (i.e., buildings in an urban setting) was modeled in QUIC and LBM. CBRN point releases originating from three release locations were simulated with wind originating from NNW, N, or NNE. 225 potential sensor locations are considered on an
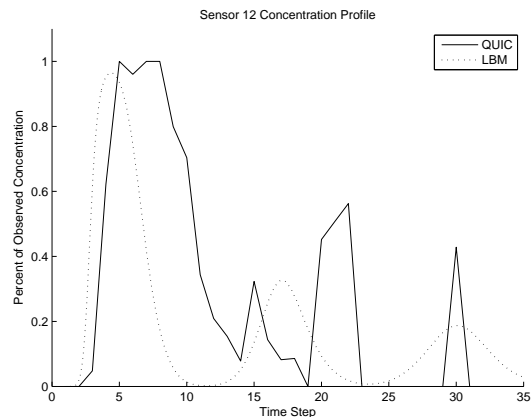

Fig. 2. Evolution of concentration at a sensor downwind from the source of the CBRN event.

evenly spaced grid near ground level. Figure 3 illustrates a cross section of the environment with the locations of the obstructions (larger shapes) and releases (denoted by an "x").
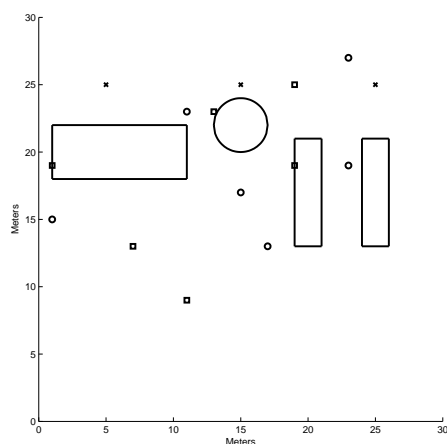

Fig. 3. Toy environment with optimal sensor placement solution.

The probability laws were determined empirically via Monte Carlo simulation (using (1) and (5)) for each combination of wind direction, release location, and sensor location. First, sensor concentration observations are encoded into binary outputs indicating the presence of non-nil concentration. Considering the binary nature of the sensor readings, sensor measurement noise is introduced via Bernoulli trials which result in flipping sensor observations. The resulting noisy sensor readings provide a wealth of observations from which probability laws can be established, allowing for solution of (11) and the determination of the d_{ijk} in (12). For evalu-

ation of the release detection and source localization methodologies, sensors were sampled once per second for 50 seconds, resulting in an observation window of 50 binary indicators of the presence of a CBRN agent. In the case of both the QUIC and LBM data, by the time six sensors are placed within the environment there is no further improvement in the asymptotic worst case localization error. The solution depicted for the case in which QUIC data were used is shown in Figure 3 as the small circles. The small squares correspond to the optimal sensor placement determined using LBM data.

A test set is constructed for each release location by first selecting a wind direction according to a "wind rose" which describes the likelihood of each wind direction. Then, using the simulations from which the probability laws were derived, test sensor observations are generated.

1) Release detection evaluation: For each scenario in the test set, a release was declared whenever at least one sensor's observations resulted in declaring a CBRN attack according to the test in (2). To gain an understanding on the CBRN detector's performance, this process was conducted for many different values of η to facilitate the construction of the *Receiver Operating Characteristic (ROC)* curve. When a reasonable sensor error rate of 1/8 was used in both the case in which QUIC data were used in the evaluation and when LBM data were used, the detector attained perfect detection performance, with 100% probability of detection and 0% probability of false alarm. To better understand how robust our detection methodology is to sensor noise, we repeated the experiment with different sensor error rates. Figure 4 shows the probability of detection achieved when 0% probability of false alarm is demonstrated as a function of sensor error rate. As shown, there is a precipitous decline in detection performance by the time roughly one third of the sensor measurements are erroneous.

2) Source localization evaluation: The ability of all three localization methodologies we outlined in Section V was also evaluated on each test scenario. Table I displays the accuracy of each approach.

TABLE I
ACCURACY OF LOCALIZATION METHODOLOGIES IN THE TOY ENVIRONMENT

	QUIC Data	LBM Data
Sequential	99.7%	98.3%
Centralized	100%	99.7%
Hybrid	100%	99.7%

When the hybrid localization approach is used, gains in localization accuracy begin to disappear after a certain number of sensors are used in each decision within the sequence of maximum likelihood tests. This accuracy
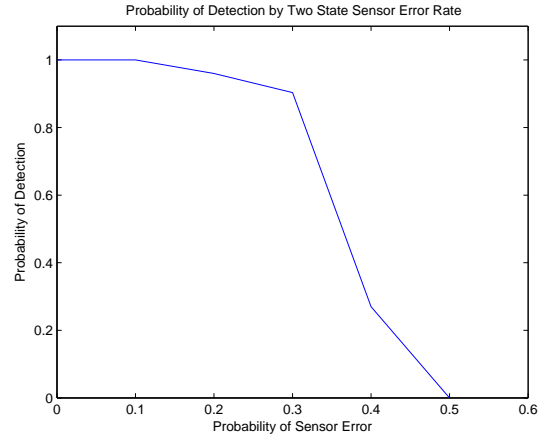

Fig. 4. Probability of detection as the error rate for the two state CBRN sensor grows. Reported values correspond to zero observed false alarms.

is equal to the accuracy attained by the centralized maximum likelihood approach. When QUIC data were used for evaluation, perfect performance was observed when only one sensor was used in each decision in the sequence. This actually corrected the one mistake found when the sequential GLRT localization technique was used. When LBM data were used for evaluation, it was not until three sensors were employed for each decision within the sequence of maximum likelihood tests that the same performance as the maximum likelihood approach was observed.

B. City Grid Environment

To further evaluate the proposed CBRN detector and locator, a three-dimensional environment inspired by a dense urban grid pattern was modeled in QUIC as well as using LBM. The environment consists of a city grid, four blocks-by-four blocks. Each block is 100 meters-by-100 meters with 10 meter-wide throughways. Each block's height is drawn randomly from the uniform distribution ranging from 20 to 60 meters. Sensors are allowed to be placed at any intersection and five intersections are considered as potential release locations. The shape of the grid, as well as the location of simulated releases, is shown in Figure 5.

CBRN event scenarios considered are the 40 unique combinations of five different release locations within the grid with wind blowing at 1 m/s or 5 m/s and originating from the four cardinal directions. The two selected wind speeds represent nominal wind conditions, but do not translate into a uniform wind field. As the simulation proceeds, local wind eddies vary according to conditions incited by the presence of buildings within the grid.
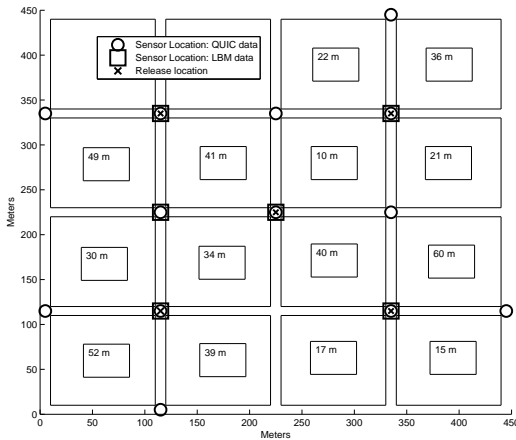
Fig. 5. City environment with optimal sensor placement solution.

As in Subsection VII-A, the probability laws according to (1) and (5) for each combination of CBRN scenario and sensor location are determined by Monte Carlo simulation using data generated from both QUIC and LBM simulations. This time, however, not only is the binary sensor model with Bernoulli false alarms considered, but a real valued sensor model with additive white noise is considered as well. These sensors are modeled as

$$\hat{C} = C + N(0, \sigma_\epsilon), \qquad (16)$$

where C is the actual concentration of particulate present at the sensor's location, $N(0, \sigma_\epsilon)$ denotes a normally distributed random variable with mean 0 and standard deviation σ_ϵ, and $\hat{C}$ denotes the sensor's reported concentration observation. Values of $\hat{C}$ were then translated into a discretized representation of five concentration states. First a value of σ_ϵ which produces an equivalent false alarm as the binary sensor model was used. To illustrate the power of increasing the number of concentration states used in the analysis, the normal random variate in (16) was replaced with $N(0, 10\sigma_\epsilon)$.

For the binary sensor case, if data generated via LBM were used for analysis, little improvement in worst case error probability is observed after six sensors are deployed. However, when data generated via QUIC were used for analysis, it is not until 16 sensors are deployed that no further improvement is made in the bound on the worst case error probability.

For the real-valued sensor case with mild noise, if data generated via LBM were used for analysis, once five sensors are placed in the environment no further improvement in worst case error probability is observed. These five sensor locations correspond to the five release locations in the environment. When QUIC data were

employed for analysis, no further improvement in the worst case error probability is observed after 13 sensors have been deployed. The optimal placements of real-valued sensors with mild noise are presented in Figure 5.

In the case of real valued sensor models with large noise, analysis utilizing LBM data shows no further improvement after eight sensors are deployed. When QUIC data is used, no further performance is guaranteed after the same 13 sensors are deployed. This solution is the same as the one found in the mild noise case.

A test set for the city grid was constructed for each of the release locations by first selecting a wind direction (all four cardinal directions are chosen with equal likelihood) and then selecting a wind speed (both wind speeds are selected with equal likelihood). The test set sensor readings were generated in the same way as the sensor readings used to determine the probability laws.

1) Release detection evaluation: Following the procedure in Subsection VII-A1, for each scenario in the test set, a release was considered detected whenever at least one sensor's observations lead to the decision of "attack" via the test in (2). The resulting ROC curves are depicted in Figure 6. Whether QUIC or LBM simulations were used to generate scenario data, the detector attains nearly perfect performance. This holds true regardless of which sensor model is used.
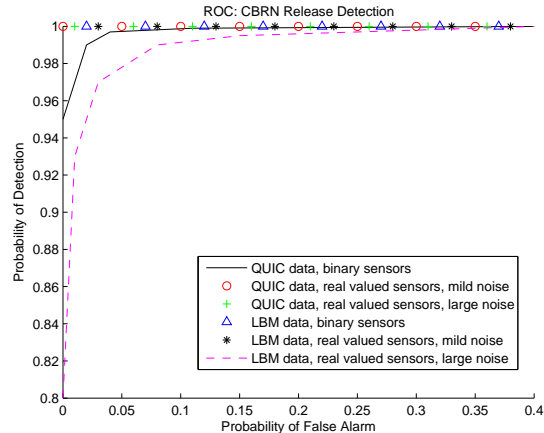

Fig. 6. ROC curve of CBRN detector in the city grid environment.

2) Source localization evaluation: The proposed CBRN localization techniques performed well on both the QUIC and LBM generated data sets, even when real valued sensors with large additive noise are used. The performances observed are listed in Table II.

While all three methodologies perform phenomenally in both simulated environments, the hybrid and centralized maximum likelihood techniques are the apparent champions, both consistently outperforming the sequential localization technique. Yet, the sequential technique

has the advantage of being fully distributed which is desirable in implementations involving a large number of sensors.

In some instances, the hybrid approach outperformed the centralized approach, as noted in the binary sensor and large noise real valued sensor scenarios using LBM data. In these cases, reducing the number of sensors used in a maximum likelihood decision prevented a single sensor's observations skewing a computed probability into an incorrect localization.

C. City Grid Environment with Sensor Constraints

In Subsection VII-B, the optimal solution to (14) was evaluated when K was determined by performance gain analysis. However, it may be the case that a limited number of sensors are available for deployment. To assess the performance of the proposed CBRN detector and source locators in a sensor constrained scenario, we repeated the analysis in Subsection VII-B, this time with $K = 5$, indicating that five sensors are available for deployment. The optimal placements of binary sensors under this constraint can be found in Figure 7 using the QUIC and LBM generated data and binary sensor models.
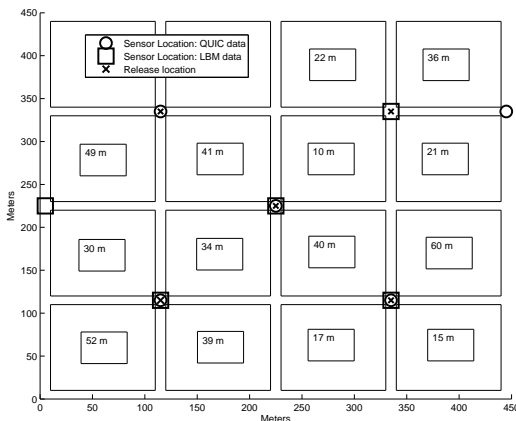

Fig. 7.   Optimal placement of 5 sensors in the city environment.

When the number of sensors deployed decreases, the number of potentially invisible releases (i.e., a release that is not detected by a single sensor) increases. An intuitive approach to address this concern is to span the deployed sensors out as much as possible. It appears, in Figure 7, that the optimal placement solutions adhere to this philosophy. Evaluations using the real valued sensor models produced optimal sensor placements corresponding to the five simulated release locations.

1) Release detection evaluation: There is a significant decrease in performance due to the increased presence of invisible releases in the test sets, as illustrated in Figure 8. However, this degradation is confined solely to analysis which uses a binary sensor model and CBRN dispersion modeled using QUIC.
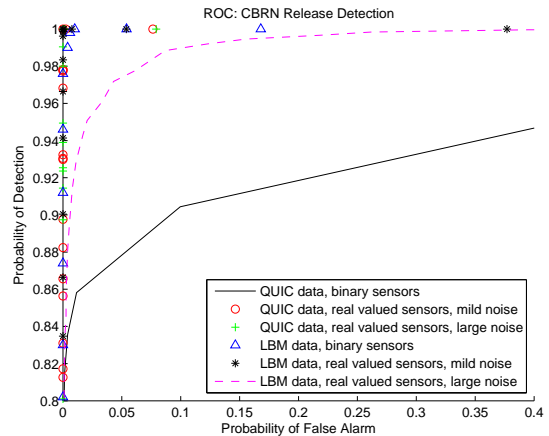

Fig. 8.   ROC curve of CBRN detector in the city grid environment for the case of using only five sensors.

2) Source localization evaluation: Localization accuracy for all three approaches is listed in Table III. We actually see an increase in performance when the number of sensors are limited in the LBM scenarios with real valued sensors and large additive noise. This is due to removing sensor information that previously led to erroneous localizations, a phenomenon similar to the reason we observed better performance from the hybrid localization technique over the centralized approach.

An apparent follow up question to the localization performance analysis is, how well does the optimal solution to the sensor placement MILP compare to other, non-optimal solutions? To answer this question, placements in which the sensor locations are chosen uniformly randomly were evaluated on the same test set as the optimal placement solution. The percentage of test cases that were correctly localized appear in Figure 9. Clearly, the optimal placement solution outperforms any of the randomly chosen sensor placements.

VIII. CONCLUSIONS

We have presented a new two-stage methodology to the related problems of detecting CBRN events and then determining the location of their origin. Our approach bypasses many of the problems "inverse dispersion" approaches face in an inhomogeneous and uncertain urban setting.

TABLE II
ACCURACY OF LOCALIZATION METHODOLOGIES IN THE CITY GRID ENVIRONMENT

	QUIC Data Binary sensors	QUIC Data Real valued sensors Mild noise	QUIC Data Real valued sensors Large noise	LBM Data Binary Sensors	LBM Data Real valued sensors Mild noise	LBM Data Real valued sensors Large noise
Sequential	98.42%	93.50%	94.66%	98.98%	94.18%	83.90%
Centralized	99.98%	100%	99.98%	99.82%	100%	99.48%
Hybrid	99.98%	100%	99.32%	99.84%	100%	99.60%

TABLE III
ACCURACY OF LOCALIZATION METHODOLOGIES IN THE CITY GRID ENVIRONMENT WHEN ONLY FIVE SENSORS ARE AVAILABLE FOR DEPLOYMENT

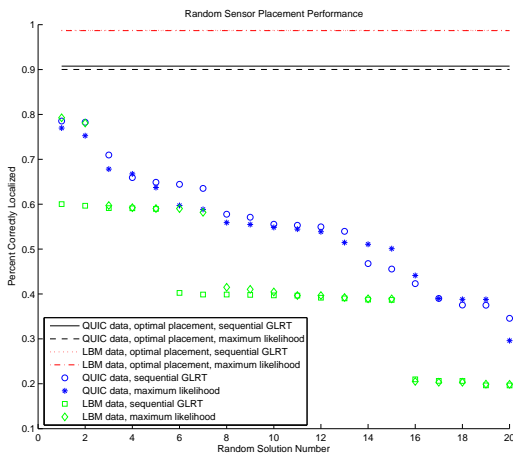|	QUIC Data Binary sensors	QUIC Data Real valued sensors Mild noise	QUIC Data Real valued sensors Large noise	LBM Data Binary Sensors	LBM Data Real valued sensors Mild noise	LBM Data Real valued sensors Large noise
Sequential	90.76%	93.40%	93.90%	98.72%	94.18%	86.48%
Centralized	90.02%	96.74%	96.62%	98.66%	100%	99.52%
Hybrid	90.02%	97.88%	99.32%	97.40%	100%	99.60%

Fig. 9. Evaluation of 20 randomly chosen placements of 5 sensors.

To detect attacks, relative entropy is used to compare current sensor observations to known nominal probability laws derived from simulations. Once a CBRN event is detected, a localization procedure is used to determine where the CBRN event started. Our localization techniques include a sequence of GLRT decisions using information from a single sensor, a maximum likelihood selector in which information from all deployed sensors are used, and a hybrid of these two approaches consisting of a sequence of maximum likelihood decisions using information from more than one sensor.

The proposed CBRN detection approach demonstrated promise within a numerical evaluation with data generated in two different dispersion simulators. Our approaches for source localization performed well also. The sequential GLRT localization methodology is vulnerable to committing a localization error early on in the decision making process, yet it offers obvious implementation advantages in large deployments. The centralized maximum likelihood localization approach remedies this by employing all deployed sensor information in a single decision, however, the hybrid approach committed, in total, slightly fewer localization errors in our experiments.

APPENDIX A
PROOF OF THEOREM V.2

Our approach to show optimality of the GLRT will be to show that it is a subset of decision rules which satisfy the criterion of Definition 2. To that end we first present a lemma similar to a result first produced by Hoeffding [16].

Lemma A.1 *The sequence of tests*

$$\mathscr{S}_{ijk,n}^* = \{\mathbf{y}^{k,n} | \inf_{\theta_j \in \Omega_j} I_2(\mathscr{E}_2^{\mathbf{y}^{k,n}} | \mathbf{\Pi}_{\theta_j}^k) \geq \lambda\} \qquad (17)$$

is optimal under Definition 2.

Proof: Consider the sequence of tests

$$\hat{\mathscr{S}}_{ijk,n}(\theta_j) = \{\mathbf{y}^{k,n} | I_2(\mathscr{E}_2^{\mathbf{y}^{k,n}} | \mathbf{\Pi}_{\theta_j}^k) \geq \lambda\}.$$

By Lemma 1 of [22], for all $\theta_j \in \Omega_j$, the constraint

$$\limsup_{n \to \infty} \frac{1}{n} \log \hat{\alpha}_{ijk,n}(\theta_j) < -\lambda$$

is satisfied and the value

$$-\limsup_{n \to \infty} \frac{1}{n} \log \hat{\beta}_{ijk,n}(\theta_i)$$

is maximized uniformly over all $\theta_i \in \Omega_i$, where $\hat{\alpha}_{ijk,n}(\theta_j)$ and $\hat{\beta}_{ijk,n}(\theta_i)$ indicate the Type-I and Type-II errors of test $\hat{\mathscr{S}}_{ijk,n}(\theta_j)$. Thus, the constraint (7) is satisfied for test $\mathscr{S}_{ijk,n}^*$, since $\mathbf{y}^{k,n} \in \mathscr{S}_{ijk,n}^*$ if and only if

$$\mathbf{y}^{k,n} \in \bigcup_{\theta_j \in \Omega_j} \hat{\mathscr{S}}_{ijk,n}(\theta_j).$$

Further, by the same reasoning and using the union bound, it must be the case that the quantity $-\limsup_{n\to\infty}\frac{1}{n}\log\beta^*_{ijk,n}(\theta_i)$ is maximized uniformly over all $\theta_i\in\Omega_i$. ∎

Now, allow $\mathbf{y}^{k,n}\in\mathscr{S}^{GLRT}_{ijk,n}$. Then we have,

$$
\begin{aligned}
\lambda \leq& \frac{1}{n}\log\sup_{\theta_i\in\Omega_i}P^k_{\theta_i}(\mathbf{y}^{k,n})-\frac{1}{n}\log\sup_{\theta_j\in\Omega_j}P^k_{\theta_j}(\mathbf{y}^{k,n})\\
=&-H(\mathscr{E}^{\mathbf{y}^{k,n}}_2)-\inf_{\theta_i\in\Omega_i}I_2(\mathscr{E}^{\mathbf{y}^{k,n}}_2|\mathbf{\Pi}^k_{\theta_i})\\
&-\frac{1}{n}\log\sup_{\theta_j\in\Omega_j}P^k_{\theta_j}(\mathbf{y}^{k,n})\\
\leq&-H(\mathscr{E}^{\mathbf{y}^{k,n}}_2)-\frac{1}{n}\log\sup_{\theta_j\in\Omega_j}P^k_{\theta_j}(\mathbf{y}^{k,n})\\
=&\inf_{\theta_j\in\Omega_j}I_2(\mathscr{E}^{\mathbf{y}^{k,n}}_2|\mathbf{\Pi}^k_{\theta_j}),
\end{aligned}
$$

where $H(\cdot)$ is the *empirical entropy*, defined as

$$
H(\mathscr{E}^{\mathbf{y}^{k,n}}_2)=-\sum_{u=1}^{|\Sigma|}\sum_{v=1}^{|\Sigma|}\mathscr{E}^{\mathbf{y}^{k,n}}_2(\sigma_v,\sigma_u)\log\mathscr{E}^{\mathbf{y}^{k,n}}_2(\sigma_u|\sigma_v).
$$

The first and second equalities stem from the definitions of H and I_2. The second inequality is the result of the non-negativity of I_2. Given (17), it follows that $\mathscr{S}^{GLRT}_{ijk,n}\subseteq\mathscr{S}^*_{ijk,n}$ and Lemma A.1 establishes that the constraint (7) is satisfied for the GLRT.

For any $\theta_i\in\Omega_i$ note

$$
\begin{aligned}
\beta^{GLRT}_{ijk,n}(\theta_i)=&\mathbf{P}_{\theta_i}[\mathbf{y}^{k,n}\notin\mathscr{S}^{GLRT}_{ijk,n}]\\
=&\mathbf{P}_{\theta_i}[\mathbf{y}^{k,n}\notin\mathscr{S}^*_{ijk,n}]\\
&+\mathbf{P}_{\theta_i}[\mathbf{y}^{k,n}\in\mathscr{S}^*_{ijk,n},\mathbf{y}^{k,n}\notin\mathscr{S}^{GLRT}_{ijk,n}].
\end{aligned}
$$

If $\mathbf{y}^{k,n}\notin\mathscr{S}^{GLRT}_{ijk,n}$,

$$
\begin{aligned}
\lambda>&\frac{1}{n}\log\sup_{\theta_i\in\Omega_i}P^k_{\theta_i}(\mathbf{y}^{k,n})-\frac{1}{n}\log\sup_{\theta_j\in\Omega_j}P^k_{\theta_j}(\mathbf{y}^{k,n})\\
=&-H(\mathscr{E}^{\mathbf{y}^{k,n}}_2)-\inf_{\theta_i\in\Omega_i}I_2(\mathscr{E}^{\mathbf{y}^{k,n}}_2|\mathbf{\Pi}^k_{\theta_i})\\
&+H(\mathscr{E}^{\mathbf{y}^{k,n}}_2)+\inf_{\theta_j\in\Omega_j}I_2(\mathscr{E}^{\mathbf{y}^{k,n}}_2|\mathbf{\Pi}_{\theta_j})\\
=&\inf_{\theta_j\in\Omega_j}I_2(\mathscr{E}^{\mathbf{y}^{k,n}}_2|\mathbf{\Pi}_{\theta_j})-\inf_{\theta_i\in\Omega_i}I_2(\mathscr{E}^{\mathbf{y}^{k,n}}_2|\mathbf{\Pi}_{\theta_i}).
\end{aligned}
$$

Since $\mathbf{y}^{k,n}\in\mathscr{S}^*_{ijk,n}$ implies $\lambda\leq\inf_{\theta_j\in\Omega_j}I_2(\mathscr{E}^{\mathbf{y}^{k,n}}_2|\mathbf{\Pi}^k_{\theta_j})$, when $\mathbf{y}^{k,n}\in\mathscr{S}^*_{ijk,n}$ and $\mathbf{y}^{k,n}\notin\mathscr{S}^{GLRT}_{ijk,n}$ we have $\mathscr{E}^{\mathbf{y}^{k,n}}_2\in\mathscr{C}_{ijk}$. By Theorem V.1, for all $\theta_i\in\Omega_i$,

$$
-\limsup_{\theta_i\in\Omega_i}\frac{1}{n}\log\mathbf{P}^k_{\theta_i}[\mathbf{y}^{k,n}\in\mathscr{S}^*_{ijk,n},\mathbf{y}^{k,n}\notin\mathscr{S}^{GLRT}_{ijk,n}]
$$
$$
\geq\inf_{\mathbf{Q}\in\mathscr{C}_{ijk}}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_i}).
$$

Therefore,

$$
-\limsup_{n\to\infty}\frac{1}{n}\log\beta^{GLRT}_{ijk,n}(\theta_i)
$$

$$
\begin{aligned}
=&\min\Bigg\{-\limsup_{n\to\infty}\frac{1}{n}\log\beta^*_{ijk,n}(\theta_i),\\
&-\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{P}^k_{\theta_i}[\mathbf{y}^{k,n}\in\mathscr{S}^*_{ijk,n},\mathbf{y}^{k,n}\notin\mathscr{S}^{GLRT}_{ijk,n}]\Bigg\}\\
\geq&\min\{\inf_{\mathbf{Q}\in\mathscr{A}_{ijk}}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_i}),\inf_{\mathbf{Q}\in\mathscr{C}_{ijk}}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_i})\}\\
=&\inf_{\mathbf{Q}\in\mathscr{A}_{ijk}}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_i}).
\end{aligned}
$$

It follows that the Type-II error probability decreases exponentially to zero faster than the Type-II error probability of test $\mathscr{S}^*_{ijk,n}$. Lemma A.1 has already established that $\mathscr{S}^*_{ijk,n}$ is optimal under Definition 2. We can therefore conclude that the GLRT is optimal under Definition 2.

APPENDIX B
PROOF OF LEMMA V.3

If the GLRT is not optimal, for all $\mathbf{Q}\notin\mathscr{D}_{ijk}$, $\inf_{\theta_j\in\Omega_j}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_j})-\inf_{\theta_i\in\Omega_i}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_i})\geq\lambda$. Equivalently,

$$
\inf_{\theta_j\in\Omega_j}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_j})\geq\lambda+\inf_{\theta_i\in\Omega_i}I_2(\mathbf{Q}|\mathbf{\Pi}^k_{\theta_i})\geq\lambda,
$$

where the last inequality follows from the non-negativity of I_2. Therefore, using Lemma A.1, it follows that

$$
\limsup_{n\to\infty}\frac{1}{n}\log\alpha^{GLRT}_{ijk}(\theta_j)\leq-\lambda,\quad\forall\theta_j\in\Omega_j.
$$

APPENDIX C
PROOF OF PROPOSITION VI.2

The sensor with index k^*_{ij} will use the GLRT and achieve a maximum probability of error with exponent no smaller than $d_{ijk^*_{ij}}$. For every i and $j\neq i$ define $E_n(i,j)$ as the event that the GLRT employed by the sensor at $B_{k^*_{ij}}$ will decide L_j when L_i is true. For all $\delta_n>0$ and large enough n we have

$$
\begin{aligned}
\mathbf{P}_{\boldsymbol{\theta}_i}[\text{error}]\leq&\mathbf{P}_{\boldsymbol{\theta}_i}[\cup_{j\neq i}E_n(i,j)]\\
\leq&\sum_{j\neq i}e^{-n(d_{ijk^*_{ij}}+\delta_n)}\leq(N-1)e^{-n(\epsilon^*+\delta_n)}.
\end{aligned}
$$

The second inequality above is due to Proposition V.4 and the last inequality above is due to (15). Since the bound above holds for all i we obtain the desired result.

REFERENCES

[1] G. Schweitzer and E. Sharber, A.C., *Countering urban terrorism in Russia and the United States: proceedings of a workshop*. National Academies Press, 2006. [Online]. Available: http://books.google.com/books?id=qlW4eDiSvpgC
[2] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, New York, 1991.
[3] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. NY: Springer-Verlag, 1998.
[4] I. C. Paschalidis and D. Guo, "Robust and distributed stochastic localization in sensor networks: Theory and experimental results," *ACM Trans. Sensor Networks*, vol. 5, no. 4, pp. 34:1–34:22, 2009.

[5] E. R. Paryak and M. J. Brown, "QUIC URB v.1.1: theory and users guide," Los Alamos National Laboratory, Tech. Rep., 2007.

[6] S. Chen and G. D. Doolen, "Lattice Boltzmann method for fluid flows," *Annual Review of Fluid Mechanics*, vol. 30, no. 1, pp. 329–364, 1998.

[7] A. Atalla and A. Jeremic, "Estimation of boundary properties using stochastic differential equations," in *Electrical and Computer Engineering, 2009. CCECE '09. Canadian Conference on*, 2009, pp. 385–389.

[8] M. Ortner and A. Nehorai, "A sequential detector for biochemical release in realistic environments," *Signal Processing, IEEE Transactions on*, vol. 55, no. 8, pp. 4173–4182, 2007.

[9] M. Ortner, A. Nehorai, and A. Jeremic, "Biochemical transport modeling and Bayesian source estimation in realistic environments," *Signal Processing, IEEE Transactions on*, vol. 55, no. 6, pp. 2520–2532, 2007.

[10] A. Atalla and A. Jeremic, "Localization of chemical sources using stochastic differential equations," in *Acoustics, Speech and Signal Processing (ICASSP), 2008. IEEE International Conference on*, 2008, pp. 2573–2576.

[11] M. Ortner and A. Nehorai, "Biochemical transport modeling, estimation and detection in realistic environments," in *Acoustics, Speech and Signal Processing (ICASSP), 2008. IEEE International Conference on*, 2008, pp. 5169–5172.

[12] L. Delle Monache, J. K. Lundquist, B. Kosović, G. Johannesson, K. M. Dyer, R. D. Aines, F. K. Chow, R. D. Belles, W. G. Hanley, S. C. Larsen, G. A. Loosmore, J. J. Nitao, G. A. Sugiyama, and P. J. Vogt, "Bayesian inference and Markov chain Monte Carlo sampling to reconstruct a contaminant source on a continental scale," *Journal of Applied Meteorology and Climatology*, vol. 47, no. 10, pp. 2600–2613, 2008.

[13] A. Keats, E. Yee, and F. Lien, "Bayesian inference for source determination with applications to complex urban environment," *Atmospheric Environment*, vol. 41, pp. 465–479, 2007.

[14] M. S. Daskin, *Network and discrete location: models, algorithms, and applications*, ser. Wiley-Interscience series in discrete mathematics and optimization. Wiley, 1995.

[15] G. Hoblos, M. Staroswiecki, and A. Aitouche, "Optimal design of fault tolerant sensor networks," in *International Conference on Control Applications*, vol. 9, no. 50, 2000, pp. 467–472.

[16] W. Hoeffding, N. Fisher, and P. Sen, *The collected works of Wassily Hoeffding*, ser. Springer series in statistics: Perspectives in statistics. Springer-Verlag, 1994. [Online]. Available: http://books.google.com/books?id=f7xLrFTtKpYC

[17] S. Ray, W. Lai, and I. C. Paschalidis, "Statistical location detection with sensor networks," *Joint special issue IEEE/ACM Trans. Networking and IEEE Trans. Information Theory*, vol. 52, no. 6, pp. 2670–2683, 2006.

[18] Röckle, "Bestimmung der strömungsverhlltnisse im bereich komplexer bebauungsstrukturen," Ph.D. dissertation, Technischen Hochschule, Darmstadt, Germany, 1990.

[19] U. Frisch, B. Hasslacher, and Y. Pomeau, "Lattice-gas automata for the Navier-Stokes equation," *Phys. Rev. Lett.*, vol. 56, pp. 1505–1508, 1986.

[20] P. Lallemand and L.-S. Luo, "Theory of the lattice Boltzmann method: Dispersion, dissipation, isotropy, Galilean invariance, and stability," *Phys. Rev. E*, vol. 61, pp. 6546–6562, 2000.

[21] B. Legg and M. R. Raupach, "Markov-chain simulation of particle dispersion in inhomogeneous flows: the mean drift velocity induced by a gradient in eulerian velocity variance," *Boundary-Layer Meteorology*, vol. 24, pp. 3–13, 1982.

[22] O. Zeitouni, J. Ziv, and N. Merhav, "When is the generalized likelihood ratio test optimal?" *Information Theory, IEEE Transactions on*, vol. 38, no. 5, pp. 1597–1602, 1992.

R. Taylor Locke has an M.S. in Applied Mathematics from the Rensselaer Polytechnic Institute (RPI) and a Ph.D. in Systems Engineering from Boston University (BU). He is currently at the MIT Lincoln Laboratory as Technical Staff in the Informatics and Decision Support Group.

Ioannis Ch. Paschalidis (M'96, SM'06) is a Professor and Distinguished Faculty Fellow at Boston University with appointments in the Department of Electrical and Computer Engineering and the Division of Systems Engineering. He is a Co-Director of the Center for Information and Systems Engineering (CISE). He completed his graduate education at the Massachusetts Institute of Technology (MIT) receiving an M.S. (1993) and a Ph.D. (1996), both in Electrical Engineering and Computer Science. In September 1996 he joined Boston University where he has been ever since. He has held visiting appointments with MIT, and the Columbia University Business School. His current research interests lie in the fields of systems and control, networking, applied probability, optimization, operations research, computational biology, and medical informatics.