Geo-Spatial Location Estimation for Internet of Things (IoT) Networks with One-Way Time-of-Arrival via Stochastic Censoring

Sai Ganesh Nagarajan[1], Pengfei Zhang [1] and Ido Nevat[1],
[1] Institute for Infocomm Research (I^2R), Singapore.

Abstract

We develop new algorithms for Geo-Spatial location estimation for Internet of Things (IoT) Networks by utilizing a One Way Time of Arrival (OW-TOA) approach. We first demonstrate the shortcomings of current OW-TOA location estimation algorithms in cases where some of the Anchor Nodes (ANs) are outside the communication range with the source. These shortcomings are manifested in ambiguous level-sets of the likelihood function, resulting in inaccurate source localization. We then show that in addition to using the conventional noisy ranging measurements from nodes which are in the communication range, we can make use of the audibility information (which indicates whether a node is able/unable to communicate with the target) that each node has. By leveraging on this available information, we considerably improve the localization performance, by mitigating the well known ambiguity problem which arises when a few AN are inaudible. The first algorithm we develop is the *Joint source location and time offset* (J-LoT) Maximum Likelihood Estimation (MLE) and the second is the *Differential source location* (D-LoT) MLE. Our algorithms require no additional hardware and the novelty lies in using the available audibility information to resolve the likelihood ambiguity. Finally, we develop the Cramér-Rao Bounds (CRLB) of the source location estimate for both algorithms. Extensive simulations show the significant benefits of our algorithms compared to the conventional MLE algorithms.

Keywords: Wireless sensor networks, Internet of Things, Localization, Audibility, Time of Flight Ranging.

I. Introduction

The term "Internet-of-Things" (IoT) describes several technologies and research disciplines in which the Internet extends into the physical world [1]–[4]. For the IoT paradigm to be successful it will need to be able to track the location (and the movement) of these objects, thus providing location-based services, which we denote as Location of Things (LoT). The location of the object can be used to provide a wide range of novel services, for example smart-homes [5], consumer behaviour [6], logistics [7], military application [8] and many more. So far, location based services have been widely used in outdoor environments and for

navigation services of cars, airplanes etc [9]. For example, the widely used Global Position System (GPS) allows users to identify their coordinates on a map with an accuracy of approximately 10m [10]. In the LoT vision, tags can be connected to any physical element, from smart-phones, electric appliances to people and pets and should be able to operate in both indoor and outdoor environments with sub-meter accuracy [11]. Technologies such as GPS or WiFi based solutions are unable to provide such accuracy due to their low resolution and susceptibility to non-line-of-sight phenomena [12], [13]. This problem drives researchers to find alternative technologies for accurate localization for power constrained devices. In addition, in IoT systems, the user has no full access to the observation, since the network components (in our case, the ANs), may not belong to the user, but he/she may be able to use them (for free or for a fee). If the ANs are devices which were purchased from vendors (eg. [14]–[18]), the user will not have access to RSS measurements. This is a special feature of future IoT networks where data is shared, but may be only partially shared, at the discretion of the owner of the data. To illustrate this scenario, fig. 1 shows the special features which our model contains and why this is a practical and suitable model for IoT applications.

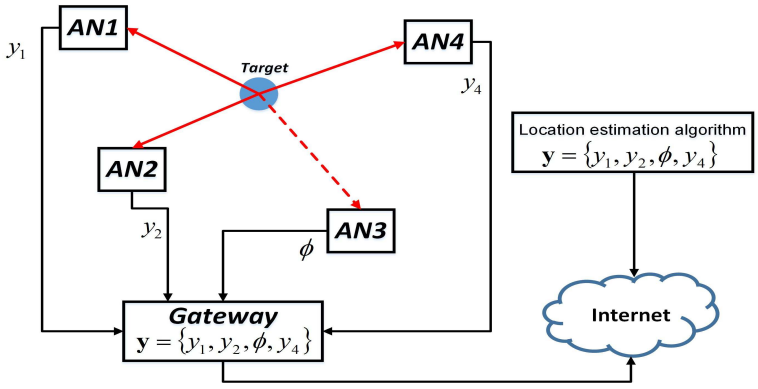


Fig. 1: IoT scenario for localization application

One promising approach that aims at achieving sub-centimeter accuracy is based on Ultra-Wideband (UWB) ranging [12], [19]–[21]. Ultra Wide Band (UWB) technology is becoming ubiquitous these days due to its low power and low costs. For these reasons the IEEE has recognized the need to standardize UWB technology for use in personal area networks and established the IEEE 802.15.4a standard specifying a new UWB physical layer for sensor networks. These properties enable the vision of Internet of Things, which is about determining the location of objects with sub-centimeter accuracy. In this regard, in papers [4], [22]–[27], they talk about the ubiquity of UWB technology in IoT applications, where UWB tags are attached to the "things". Another reason for us to consider UWB transceivers for IoT applications is due to the availability of these low cost, off-the shelf commercial products which are widely used, such as [14]–[18]. The process of ranging involves estimation of the distance between any two nodes. There are

several techniques available for ranging such as Angle of Arrival (AOA), Received Signal Strength (RSS), Time of Arrival (TOA), Time Difference of Arrival (TDOA) and a combination of these [28]–[31].

While RSS ranging is common in many applications because it uses existing infrastructure, the ranging accuracy of a RSS based techniques can be in practice highly dependent on the channel parameters and distance between the two objects making it less suitable for indoor environments [32]. On the other hand, time based ranging methods which are based on Time-of-Flight (ToF), exploit the fine delay resolution property of wideband signals and have great potential for providing sub-centimeter resolution [19]–[21].

A. Ranging Methodologies

The distance of separation between the target and the anchor nodes can be obtained via Time of Arrival ranging techniques such as, One-Way Time of Arrival (OW-TOA) , Two-Way Time of Arrival (TW-TOA) and Time Difference of Arrival (TDOA).

In OW-TOA, a node A sends a packet to another node B, with the timestamp (w.r.t A's clock) information embedded. The receiving node B, then calculates the difference between the current time (w.r.t B's clock) and the timestamp in the packet to estimate the Time of Arrival and hence the distance of separation. The caveat here is that two nodes have different clocks which are not synchronised and this introduces errors in the ranging. Details of which, will be seen in the next section.

In TW-TOA, node A sends a signal and the receiving node B, processes this signal and sends an ACK or a response to the node A. Once node A receives the response, the Round Trip Time (RTT) is estimated as the twice the propagation delay between the two nodes plus a known processing delay of the node B. While synchronisation of the nodes is not an issue here ,errors can still accumulate over the processing delay due to factors like bit synchronisation and channel estimation delays [33].

TDOA-based localization systems do not rely on absolute distance estimates between pairs of nodes. Typically there are two ways of obtaining the TDOA measurements. Firstly, multiple signals are broadcast from synchronized fixed anchor nodes located at known locations and the tag measures the TDOA. In the second method, a signal is broadcast by a tag and is received by several anchor nodes. These anchor nodes share their estimated TOA and compute the TDOA. The anchors are generally synchronised through a wired network connection [34]. To calculate the location of the tag, at least three anchors with known position and two TDOA measurements are required [35].

In a LoT environment, the performance of a mobile tag is contingent upon the power it consumes and

hence the OW-TOA is used as a power efficient and secure method of ranging,see for example [36], [37].

Thus in this paper we focus on OW-TOA and TDOA methodologies for ranging.

B. Source Localisation with missing observations

One major concern which complicates the use of OW-TOA technologies in conventional systems is the fact that at least three ANs are required to have high quality localization performance. Such systems use only information which is explicitly obtained from AN which are within the communication range to the source, but do not use information which implicitly obtained through the fact that an AN which is not in the range of communication indeed does provide additional information regarding the location of the source. From statistical signal processing point of view, location estimation can be cast as a regression problem . While in the conventional approach, missing observations from AN outside the communication range have been neglected (ie. treating them as Missing Completely at Random (MCAR)) and estimate the missing distances by finding geometrical patterns in the sensor network using Multidimensional Scaling(MDS) [38], [39] or using a fingerprinting based approach [40], our novel approach treats these missing observations as Missing Not at Random (MNAR) [41] and leverages this information when estimating the location. It is important to note that while we have access to noisy ranging observations from audible nodes, we do not have direct access to Received Signal Strength (RSS) measurements, but only knowledge of the AN is audible or not (that is, has or doesn't have an observation). The reason for that is that we are using of-the-shelf devices that provide the ranging information, but no RSS information, see for example Decawave [14].

C. How audibility mitigates an ambiguous likelihood situation

Before we proceed with the rest of the paper, we would like to present a simple case as shown in Fig. 2 where the information on audibility helps to resolve an ambiguous likelihood situation, where the conventional approach of ignoring the missing observations incorrectly estimates the location. This would help in justifying our motivation to include the information from the missing values.

D. Difference between proposed method and TOA-RSS approach

There is a fundamental difference between hybrid TOA-RSS methods and the method we present in this paper. The difference is that in our model **there is no access to RSS measurements**. We illustrate this important difference in figures 3 and 4. The left figure presents our LOT system, where only ranging information is shared by the owner of the network with the user. In contrast, the right figure presents the Hybrid-RSS type system, where the user has information of both ranging as well as RSS information.

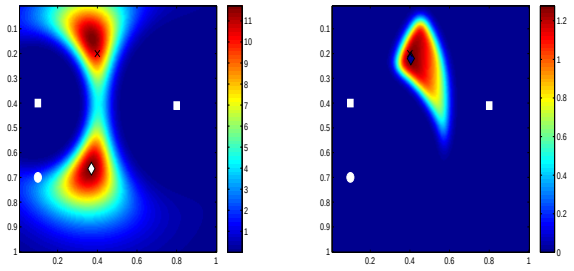


Fig. 2: An example of 3 ANs and a single target. The target's location is marked $\mathbf{x}$. The ANs are marked with $\square$ (audible) and $\bigcirc$ (inaudible), and the MLE is marked $\diamond$. **Left panel:** Likelihood surface of conventional approaches which ignore missing observations and result in ambiguity. **Right panel:** The proposed approach which incorporates missing observations from the inaudible AN and resolves the ambiguity of the likelihood surface.

As we pointed out previously,the reason that we do not have direct access to RSS measurements is that we are using off-the-shelf devices that provide the ranging information, but no RSS information. These products provide range measurements which form the future infrastructure of IoT location applications. However, it is impossible to obtain RSS measurements from these devices. Therefore, our framework provides a novel approach to incorporating audibility information which is a binary indicator, as presented in Point (3), in our system model.

E. Contributions

1) We develop a novel formulation for the LoT localization problem as a *Missing Not at Random* non-linear estimation problem, thus taking into account the valuable information of audibility that has been neglected in previous works to come up with a modified Maximum Likelihood function for:

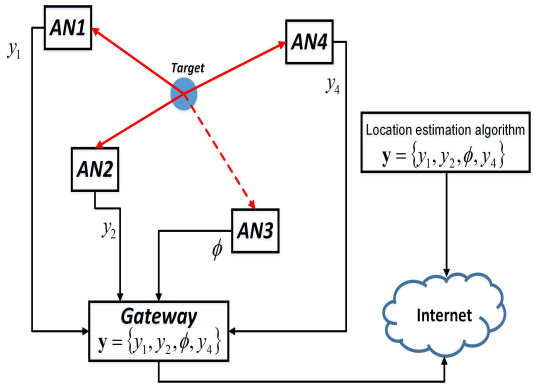


Fig. 3: LOT proposed model

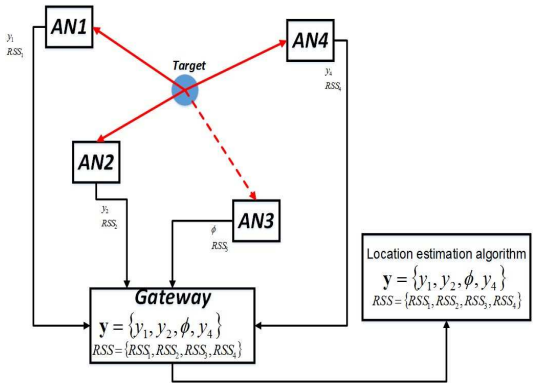


Fig. 4: hybrid TOA-RSS model

- The joint estimation of the location of the tag and the time offset term (due to OW-TOA), which we term as Joint-Location of Things (J-LoT).
- The estimation of the location of the tag under the Time Difference of Arrival (TDOA) approach, which we term as Differential-Location of Things (D-LoT).

2) We derive the spatial Cramér-Rao Lower Bound (CRLB) for the localization error for the both J-LoT and D-LoT algorithms.

II. Model Assumptions and Formal Definitions

In this section we present the definitions for the One-Way Time-of-Arrival (OW-TOA) ranging protocol , the connectivity model which is based on audibility properties, and a short overview of statistical regression with missing observations.

We will denote throughout the paper the Gaussian density function by

$$\mathcal{N}\left(x;\mu,\sigma^2\right) := \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}\left(x-\mu\right)^2\right)$$

and the standard Gaussian density and distribution functions by

$$\phi\left(x\right) := \mathcal{N}\left(x;0,1\right),$$
$$\Phi\left(x\right) := \int_{-\infty}^{x}\mathcal{N}\left(x;0,1\right)\mathrm{d}x.$$

We also denote the multivariate Normal distribution as

$$\mathcal{MVN}\left(\mathbf{y};\boldsymbol{\mu},\Sigma\right) = \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}}\exp\left(-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right).$$

Definition 1 (Vector notation)**.**
We define the zero-vector as $\vec{\mathbf{0}} = [0,0,\ldots,0]$.

A. One Way Time of Arrival (OW-TOA) Ranging Protocol

One Way Time of Arrival (OW-TOA) is a very popular method to estimate the point-to-point distance between two communicating devices. This is achieved based on the measurements of the time that a signal is required to travel from one device to the other. An example of the OW-TOA protocol can be found in [36] which aims at reducing the communication load and explores the broadcast property of WSN. The Answer to Reset(ATR) protocol makes all the other anchors listen to the ranging packets and record timestamps locally, when one anchor and the target node exchange their ranging packets.

Definition 2 (One-Way Time of Flight Ranging Model)**.** *We build the OW-TOA ranging model with the following assumptions:*

1) *There are N Anchor Nodes whose locations are fixed and known. We assume that the clocks of these ANs are synchronised through wired connection,see for example [34]*
2) *The mobile tags have a clock of their own which is not synchronised with any of the AN's in their clock offsets. But we assume here that the clocks do not suffer from clock drift.*
3) *Let t represent the "absolute" time. Let $C_m(t) = (1+\delta_m)t + o_m$ represent the local time w.r.t the absolute time t for the mobile tag,for this type of model [42]. $C_i(t) = (1+\delta_i)t + o_i$ models the local time w.r.t the absolute time t for the i^{th} Anchor Node, where δ_m,δ_i are the clock drifts ,o_m and o_i represent the clock offsets of the mobile tag and the i^{th} Anchor Node respectively.*
4) *The OW-TOA ranging scheme is as follows: Let the i^{th} anchor node send a packet with the timestamp t_a (time sent w.r.t its local clock) embedded, and once the mobile tag receives this packet at t_b (w.r.t its local clock), the propagation delay is given by:*

$$\begin{aligned}\tilde{t_p} &= C_m(t_b) - C_i(t_a) \\ &= (1+\delta_m)t_b + o_m - (1+\delta_i)t_a - o_i \\ &= (t_b - t_a) + \delta_m t_b - \delta_i t_a + o_m - o_i \\ &= t_p + \delta_m t_b - \delta_i t_a + o_m - o_i\end{aligned}$$

5) *Plugging in our assumptions about the clock drifts , we set δ_m and δ_i to be 0 for the mobile tag and all the anchor nodes. Also we assume that due to the wired time synchronisation for the AN's, o_i , the clock offset for all the anchor nodes are equal to a constant o_a. Hence the estimated propagation delay is: $\tilde{t}_p = t_p + o_m - o_a$ for all the anchor nodes, now writing this in terms of distance measurements by multiplying both sides by the speed of light c, we obtain for the i^{th} anchor node:*

$$y_i = d_i + ct_\delta + \epsilon_i$$

where d_i is the true distance between the tag and the i^{th} anchor node,ct_δ is the error term in the distance measurement due to the relative clock offset between the tag and the i^{th} AN and ϵ_i be a random variable which is formed by summing up individual component time delays and is commonly modelled as a Normal random variable ,$\epsilon_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right)$ where μ_i and σ_i are device dependent [43], [44].

B. Connectivity Model via Audibility

In order for the AN to obtain ranging measurements from the source, the two devices need to be connected so that they can exchange messages. This can be modelled via a Power Loss Model [45] and Audibility, as follows:

Definition 3 (Power Loss Model). *The power received by node A located at $\theta_A = [x_a \;\; y_a]$ from node B which is located at $\theta_B = [x_b \;\; y_b]$ is given by*

$$P_R = P_T - PL_0 - 10\alpha \log \frac{d(A,B)}{d_0} + w, \tag{1}$$

where P_T is the transmitted power by node B, α is the path-loss exponent,PL_0 is the path loss at the reference distance d_0.
$d(A,B) := \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$ is the Euclidean distance between nodes A and B and $w \sim \mathcal{N}(\mu_w, \sigma_w)$ represents the shadowing effect.

Definition 4 (Audibility). *Node A and node B are audible if*

$$P_R = P_T - 10\alpha \log \frac{d(A,B)}{d_0} + w > \lambda, \tag{2}$$

and inaudible otherwise, where λ is a pre-defined threshold representing the receiver's sensitivity.Here the constant PL_0 is subsumed in the system parameter λ.

C. Estimation with Missing Observations- Stochastic Censoring Model

Here we present a short introduction to regression with missing data and we adopt Rubin's definitions of missingness which is the standard approach for any discussion on this topic in regression settings. A more in depth information can be found in [41], [46]. According to Rubin, the tendency for missing data is a random variable that has a distribution. This implies that each variable potentially yields a pair of scores: an underlying Y value that may or may not be observed (corresponding to the observation) and a corresponding $\mathbf{\Gamma}$ (binomial random variable) value that denotes whether Y is observed or is missing (e.g., $\mathbf{\Gamma} = 1$ if Y is observed and $\mathbf{\Gamma} = 0$ if Y is missing). Under a *Missing Not at Random* (MNAR) mechanism, the data and the probability of missingness have a joint distribution:

$$p(y, \mathbf{\Gamma}|\theta; \boldsymbol{\psi}), \tag{3}$$

where y is the observation, $\mathbf{\Gamma}$ is the corresponding missing data indicator, θ is a set of parameters to be estimated (i.e. the tag's location), and $\boldsymbol{\psi}$ contains parameters that describe the propensity for missing data (e.g. the power loss model parameters).

First, under the MAR mechanism, which is commonly adopted in current algorithms, the missing observations are ignored and the model (3) simplifies and becomes

$$p(y, \mathbf{\Gamma}|\theta; \boldsymbol{\psi}) = p(y|\mathbf{\Gamma}, \theta)\, \mathrm{P}(\mathbf{\Gamma}; \boldsymbol{\psi}) \propto p(y|\theta). \tag{4}$$

and it is unnecessary to estimate the parameters that dictate missingness ψ. For this reason, an MAR mechanism is often referred to as *ignorable missingness*. According to Rubin [46], data are MAR if the

probability of a missing value is related to other measured variables, but unrelated to the underlying values of the variable that are missing.

In contrast, data are MNAR (Missing Not at Random) if the probability of missing data is systematically related to the values that are missing. The MNAR mechanism requires a model that includes all parameters of the joint distribution, not just those that are of substantive interest. It is important to note that the term "mechanism" is not intended to convey a causal relationship, but is a probabilistic explanation for how the missing values are related to variables in the data set.

Relating the MAR and MNAR models to the LoT problem, previous algorithms have been built based on the MAR model given in (4). This model ignores the reason for missingness, in our case, audibility of the tag. The algorithm we develop in this paper will make use of the MNAR model presented in (5), meaning that we take into account the reason for missingness of observations which adds important information regarding the location of the tag, leading to improved accuracy of the algorithm. In practical terms related to LoT problems, this means that our statistical analysis must incorporate a submodel that describes the propensity for missing data (e.g. audibility that predicts Γ) as the following factorization:

$$p\left(y, \mathbf{\Gamma}|\theta; \boldsymbol{\psi}\right) = p\left(y|\mathbf{\Gamma}, \theta\right) \mathrm{P}\left(\mathbf{\Gamma}|\theta; \boldsymbol{\psi}\right), \tag{5}$$

where $p\left(y|\mathbf{\Gamma}, \theta\right)$ is the likelihood model for the available observations and $\mathrm{P}\left(\mathbf{\Gamma}|\theta; \boldsymbol{\psi}\right)$ is the likelihood model for the audibility model Γ.

In this regard we make use of a type of model for the missingness of observations often found in clinical trials, surveys, economics [47], [48] called Stochastic Censoring models. Essentially in censored models the mechanism leading to the missing observation may not be under the control of the statistician , but often well understood which in our case , is a widely accepted and understood model for the audibility of sensor nodes given by Equation 4 and thus in our analysis we need to account for this information in order to avoid biased results. Thus if we were to apply the Stochastic Censoring models to the problem at hand, according to Definition 4 we have $\boldsymbol{y}$ the observation vector given by:

$$y = \begin{cases} y_i, \text{ if } P_T - 10\alpha \log \frac{d(A,B)}{d_0} + w > \lambda, \\ \varnothing, \text{ Otherwise} \end{cases} \tag{6}$$

In summary,relating the MAR and MNAR models to the Geo-Spatial Location Estimation problem we are solving, all previous algorithms have been built based on the MAR model. Those models ignore the reason for missingness, in our case, the audibility range of the tag. This implicit information can be informative regarding regions where the target is not likely to be. In contrast, the algorithm we develop in this paper makes use of the MNAR model, meaning that we take into account the reason for missingness of observations which adds important information regarding the location of the tag, leading to improved accuracy of the algorithm.

III. Geo-Spatial Location Estimation System Model

We consider a network with N Anchor Nodes (ANs) whose locations are known and one source node with an unknown location. We now present the system model:

1) A single target (node) is located at $\theta = [x_u \;\; y_u] \in \chi \subseteq \mathbb{R}^2$, which is deterministic unknown.
2) Assume N Anchor Nodes (ANs) deployed in a 2-D plane whose locations are $\mathbf{x}_i = [x_i \;\; y_i] \in \chi \subseteq \mathbb{R}^2$.
3) Let $\mathbf{\Gamma} \in \{0,1\}^N$ be the binary indicator vector, where $\mathbf{\Gamma}_i$ specifies if the tag is audible by the i-th AN according to Definition 4, meaning that

$$\begin{cases} \mathbf{\Gamma}_i = 1, \text{ if } P_T - 10\alpha \log \frac{d(A,B)}{d_0} + w > \lambda, \\ \mathbf{\Gamma}_i = 0, \text{ Otherwise} \end{cases} \tag{7}$$

where $w \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma_w^2)$ is the wireless channel shadowing effect, and λ is a pre-defined threshold representing the receiver's sensitivity.

4) Let $\mathbf{y}$ be the observation vector of the ANs, where the i-th observation is given by

$$y_i = \begin{cases} d(\theta, \mathbf{x}_i) + ct_\delta + \epsilon_i, & \text{if } \mathbf{\Gamma} i = 1 \text{ (Audible node)} \\ \varnothing, & \text{Otherwise (Inaudible node)} \end{cases} \tag{8}$$

where $\epsilon_i \overset{i.i.d}{\sim} \mathcal{N}(\mu_i, \sigma_i^2)$ and $\varnothing$ denotes "no observation". Although the algorithms proposed in the paper are valid for line-of-sight conditions,we would like to emphasise here that, the algorithm which is based on the Stochastic Censoring model to incorporate audibility information can be decoupled from the original ranging algorithm, in the sense that the ranging algorithm can be replaced by a model to account for the Non Line of Sight conditions,say for example a heteroscedastic ranging model similar to the one in [49]. We would still be able to extend it to NLoS conditions without much difficulty.

We now present the following definitions:

1) *Missing observations*: let $\mathbf{y}$ be partitioned into two sets $\mathbf{y}_o$ and $\mathbf{y}_m$, to denote the sets of ranging observations that are observed and missing, respectively. Also let $N_o = |\mathbf{y}_o|_0$ and $N_m = |\mathbf{y}_m|_0$ be the cardinality of y_o and y_m.
2) *Observation states*: let $\Gamma_{s_i}, \; i = \{1, 2, 3, \cdots, 2^N\}$ enumerate the 2^N permutations of occurrences in the elements of Γ. An example of which is shown in Table I

IV. Algorithm 1: Joint Source Location and Time Difference Maximum Likelihood Estimation (J-LoT)

As seen in definition II-A, an important characteristic of the OW-TOA method is that power efficiency in transmission comes at the cost of an inevitable time offset or a clock difference. There are many ways

to deal with this and in this section we will jointly estimate this time offset term along with the location of the tag. We now formulate the joint MLE for the unknown location of the target θ and the time offset t_δ , given set of distance and audibility observations $\mathbf{y}_{1:N}, \Gamma_{1:N}$. Here, $\mathbb{1}(\Gamma_n = 1)$ is the notation for the indicator function, which returns 1 if $\Gamma_n = 1$ and 0 otherwise.

Lemma 1. *The Joint Location of Things (J-LoT) MLE is given as the solution to the following optimisation problem:*

$$\begin{aligned}\left(\widehat{\theta}, \widehat{t_\delta}\right) = \arg\max_{\theta, t_\delta} &\sum_{n=1}^{N} \log \mathcal{N}\left(\mathbf{y}_n; d(\theta, \mathbf{x}_n) + ct_\delta + \mu_n, \sigma_n^2\right) \mathbb{1}(\Gamma_n = 1) \\ &+ \sum_{n=1}^{N} \log\left(1 - \Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d(\theta, \mathbf{x}_n)}{d_0}\right)\right)\right) \mathbb{1}(\Gamma_n = 1) \\ &+ \sum_{n=1}^{N} \log\left(\Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d(\theta, \mathbf{x}_n)}{d_0}\right)\right)\right) \mathbb{1}(\Gamma_n = 0)\end{aligned}$$

Proof: See Appendix A. ■

As seen in Lemma 1, the MLE formulation of this algorithm results in 3 components:

1) The first component represents the contribution of the conditionally independent distance observations from the audible ANs.
2) The second component represents the audibility information from the nodes which are audible.
3) The third component represents the inaudiblity information from the nodes which are inaudible.

V. Algorithm 2: Differential Source Location Maximum Likelihood Estimation (D-LoT)

We now formulate the MLE for the unknown location of the target, θ, given the set of distance and audibility observations $\mathbf{y}_{1:N}, \Gamma_{1:N}$.

Let $\boldsymbol{d}(\theta, \mathbf{x_o})$ be the corresponding distance vector between the target and all the audible ANs.Let $\boldsymbol{\mu_o}$ and $\boldsymbol{\mu_\delta}$ be the mean vectors for the time offset term and the ranging error(ϵ) for the audible ANs.Also let the observed state be given by Γ_{s_o}.

Γ_{s_i}	Binary indicator vector
Γ_{s_0}	$[0, 0, 0]$
Γ_{s_1}	$[0, 0, 1]$
Γ_{s_2}	$[0, 1, 0]$
Γ_{s_3}	$[0, 1, 1]$
Γ_{s_4}	$[1, 0, 0]$
Γ_{s_5}	$[1, 0, 1]$
Γ_{s_6}	$[1, 1, 0]$
Γ_{s_7}	$[1, 1, 1]$

TABLE I: Enumeration of $\boldsymbol{\Gamma_s}$, the states for $N = 3$

Now to deal with the time offset term, we subtract the first observation from all of the rest, i.e

$$\widetilde{\mathbf{y}_{ok}} = d\left(\theta, \mathbf{x}_{\boldsymbol{o_k}}\right) + \epsilon_k - \left(d\left(\theta, \mathbf{x}_{\boldsymbol{o_1}}\right) + \epsilon_1\right) \;\; k \in \{2, \ldots, N_o\} \tag{9}$$

where N_o is the number of audible nodes and the elements in $\mathbf{y}_{\boldsymbol{o}}$ are re-indexed from $\{1, 2, 3 \ldots, N_o\}$ and we choose the first element out of this observed vector as our reference node and subtract it from the rest, as shown in the equation above. We see from the above equation that the errors are correlated and hence the likelihood is represented as a multivariate normal distribution.

Lemma 2. *The Differential Location of Things (D-LoT) MLE is given as the solution to the following optimisation problem:*

$$\begin{aligned}\widehat{\theta} = \arg\max_{\theta} \log \mathcal{MVN}\left(\widetilde{\mathbf{y}_{\boldsymbol{o}}}; \tilde{\boldsymbol{d}}\left(\theta, \mathbf{x}_o\right) + \tilde{\boldsymbol{\mu}}_{\boldsymbol{o_\delta}} + \tilde{\boldsymbol{\mu}}_{\boldsymbol{o}}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{o}}\right) \mathbb{1}\left(\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{\boldsymbol{s_o}}\right) \\ + \sum_{n=1}^{N} \log\left(1 - \Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \mathbb{1}\left(\Gamma_n = 1\right) \\ + \sum_{n=1}^{N} \log\left(\Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \mathbb{1}\left(\Gamma_n = 0\right)\end{aligned}$$

where , $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{o}} \in \mathbb{R}^{N_o - 1 \times N_o - 1}$, $\tilde{\boldsymbol{\mu}}_{\boldsymbol{o}} \in \mathbb{R}^{N_o - 1}$,$\tilde{\boldsymbol{\mu}_{\boldsymbol{o_\delta}}} \in \mathbb{R}^{N_o - 1}$ *and* $\tilde{\boldsymbol{d}}\left(\theta, \mathbf{x}_o\right) \in \mathbb{R}^{N_o - 1}$, *and*

$$\begin{aligned}\tilde{\Sigma}_{o_{m,n}} &= \begin{cases} \sigma_{o_1}^2 + \sigma_{o_m}^2 & \text{if } m = n \\ \sigma_{o_1}^2 & \text{if } m \neq n \end{cases} \\ \tilde{\boldsymbol{\mu}_{\boldsymbol{o_\delta}}} &= \vec{\mathbf{0}}, \\ \tilde{\mu_{o_k}} &= \mu_{o_k} - \mu_{o_1}, \\ \tilde{d}\left(\theta, \mathbf{x}_{o_k}\right) &= d\left(\theta, \mathbf{x}_{o_k}\right) - d\left(\theta, \mathbf{x}_{o_1}\right).\end{aligned}$$

where $m, n = \{2, 3 \ldots, N_o\}$ and $k = \{2, 3 \ldots, N_o\}$ which are then re-indexed to $\{1, 2 \ldots, N_o - 1\}$ after performing the difference operation as shown in Equation 9

Proof: See Appendix B. ■

The interpretation for the components of the likelihood remain the same as the previous section, where the second and the third component represents the information from the audible and the inaudible nodes respectively and the first component is now a set of correlated differences of distance observations obtained from the nodes that are audible.

VI. Cramér-Rao Lower Bound on the performance of the proposed frameworks

In this Section we derive the spatial Cramér-Rao lower Bound for both the algorithms presented previously.

A. Cramér-Rao Lower Bound for the J-LoT Model

We now derive the CRLB for the J-LoT algorithm. Let $\theta_f = [\theta_x, \theta_y, t_\delta]$ be the vector of unknown parameters to be estimated.

For any unbiased estimator of the location, the CRLB is:

$$\mathbb{E}_{\mathbf{y}_{1:N},\Gamma_{1:N}}\left[\left(\widehat{\theta}-\theta\right)^2\right] \geq F^{-1}, \tag{10}$$

where F is the Fisher Information Matrix (FIM), which is given by:

$$F = -\mathbb{E}_{\mathbf{y}_{1:N},\Gamma_{1:N}}\left[\frac{\mathrm{d}^2 \log p\left(y_{1:N}, \Gamma_{1:N}|\theta, t_\delta\right)}{\mathrm{d}\theta_f^2}\right]. \tag{11}$$

Theorem 1. *The Fisher Information Matrix F of the J-LoT framework is given by:*

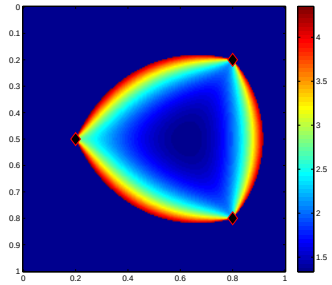
$$F = \begin{bmatrix} \Omega\left(\frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)}, \frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)}, 1, 1\right) & \Omega\left(\frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)}, \frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)}, 1, 1\right) & \Omega\left(\frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)}, 1, c, 0\right) \\ \Omega\left(\frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)}, \frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)}, 1, 1\right) & \Omega\left(\frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)}, \frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)}, 1, 1\right) & \Omega\left(1, \frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)}, c, 0\right) \\ \Omega\left(\frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)}, 1, c, 0\right) & \Omega\left(1, \frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)}, c, 0\right) & \Omega\left(1, c, c, 0\right) \end{bmatrix}$$

where

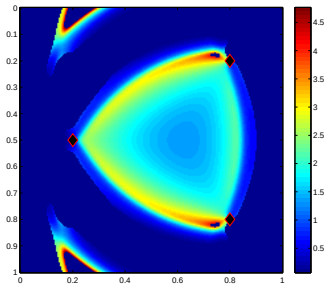
$$\Omega\left(z_1, z_2, z_3, z_4\right) = \sum_{n=1}^{N}\left(\frac{z_3 \bar{\Phi}\left(s\left(\theta, \mathbf{x}_n\right)\right)}{\sigma_n^2} + z_4 100\alpha^2\phi^2\left(s\left(\theta, \mathbf{x}_n\right)\right)\left(\frac{1}{\Phi\left(s\left(\theta, \mathbf{x}_n\right)\right)} + \frac{1}{\bar{\Phi}\left(s\left(\theta, \mathbf{x}_n\right)\right)}\right)\right) z_1 z_2$$

$$s\left(\theta, \mathbf{x}_n\right) = \left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)$$

Proof: See Appendix C for proof of Theorem 1. ■

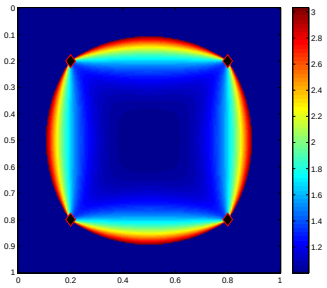


(a) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 3$,$\sigma_i^2 = 1$

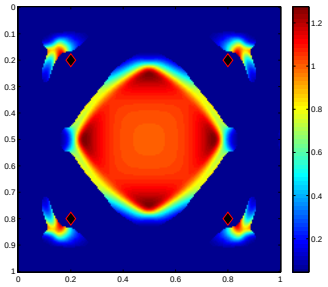


(b) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 15$,$\sigma_i^2 = 1$

Fig. 5: A spatial map of the CRLB for J-LoT for the 3 ANs whose locations are indicated by $\diamond$



(a) $N = 4$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 3$,$\sigma_i^2 = 1$



(b) $N = 4$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 15$,$\sigma_i^2 = 1$

Fig. 6: A spatial map of the CRLB for J-LoT for the 4 ANs whose locations are indicated by $\diamond$

B. Cramér-Rao Lower Bound for the D-LoT Model

We now derive the CRLB for the D-LoT algorithm. Let $\theta = [\theta_x, \theta_y]$ where F is the Fisher Information Matrix (FIM), which is given by:

$$F = -\mathbb{E}_{\widetilde{\mathbf{y}_o},\mathbf{\Gamma}}\left[\frac{\mathrm{d}^2 \log p\left(\widetilde{y}_{o2:N_o}, \Gamma_{1:N}|\theta\right)}{\mathrm{d}\theta^2}\right]. \tag{12}$$

Theorem 2. *The Fisher Information Matrix F of the D-LoT framework is given by:*

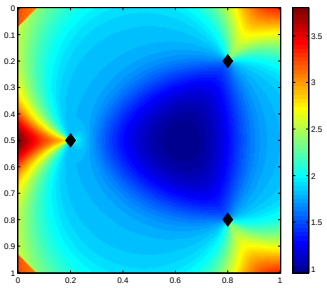
$$F = \begin{bmatrix} \Omega\left(\theta_x, \theta_x, x_n, x_n\right) & \Omega\left(\theta_x, \theta_y, x_n, y_n\right) \\ \Omega\left(\theta_x, \theta_y, x_n, y_n\right) & \Omega\left(\theta_y, \theta_y, y_n, y_n\right) \end{bmatrix}$$

where

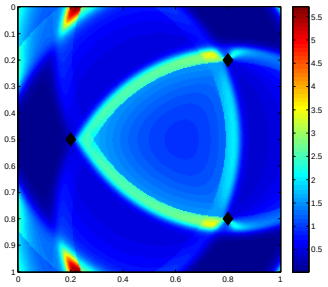
$$\begin{aligned}
\Omega\left(z_1, z_2, z_3, z_4\right) = &\sum_{o=1}^{2^N-N-1}\left[\frac{\partial \boldsymbol{\omega}_{\boldsymbol{o}}^T}{\partial z_1}\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{o}}^{-1}\frac{\partial \boldsymbol{\omega}_{\boldsymbol{o}}}{\partial z_2}\right] P\left(\mathbf{\Gamma} = \mathbf{\Gamma}_{\boldsymbol{s_o}}\right) \\
&+ 100\alpha^2 \sum_{n=1}^{N} \frac{\left(z_1 - z_3\right)\left(z_2 - z_4\right)}{d^2\left(\theta, \mathbf{x}_n\right)} \phi^2\left(s\left(\theta, \mathbf{x}_n\right)\right)\left(\frac{1}{\Phi\left(s\left(\theta, \mathbf{x}_n\right)\right)} + \frac{1}{\bar{\Phi}\left(s\left(\theta, \mathbf{x}_n\right)\right)}\right), \\
P\left(\mathbf{\Gamma} = \mathbf{\Gamma}_{\boldsymbol{s_o}}\right) = &\prod_{i=1}^{N}\left(\bar{\Phi}\left(s\left(\theta, \mathbf{x}_i\right)\right)\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=1\right)}\left(\Phi\left(s\left(\theta, \mathbf{x}_i\right)\right)\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=0\right)}, \\
s\left(\theta, \mathbf{x}_n\right) = &\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right) \\
\boldsymbol{\omega}_{\boldsymbol{o}} = &\tilde{\boldsymbol{d}}\left(\theta, \mathbf{x}_{\boldsymbol{o}}\right) + \tilde{\boldsymbol{\mu}}_{\boldsymbol{\delta}} + \tilde{\boldsymbol{\mu}}_{\boldsymbol{o}}.
\end{aligned}$$

Proof: See Appendix D for proof of Theorem 2. ■

The spatial maps as shown in Figs 5 to 8, are generated by uniformly varying the target locations across a 1×1 room and evaluating the CRLB at these locations. These maps help us to characterize the performance of our localisation algorithms and also help design the placement of the ANs, so as to optimize the system

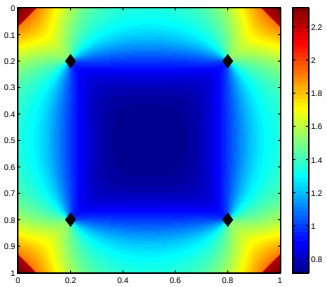


(a) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 3$,$\sigma_i^2 = 1$

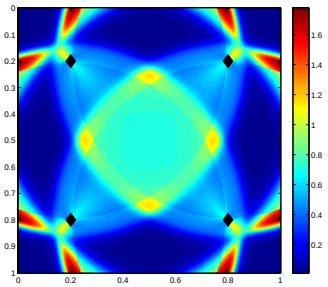


(b) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 15$,$\sigma_i^2 = 1$

Fig. 7: A spatial map of the CRLB for D-LoT for the 3 ANs whose locations are indicated by $\diamond$



(a) $N = 4$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 3$,$\sigma_i^2 = 1$



(b) $N = 4$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 15$,$\sigma_i^2 = 1$

Fig. 8: A spatial map of the CRLB for D-LoT for the 4 ANs whose locations are indicated by $\diamond$

performance. We shall see a comparison of the CRLB to the MSE in section VIII,which would be useful in making design decisions.

VII. Audibility analysis

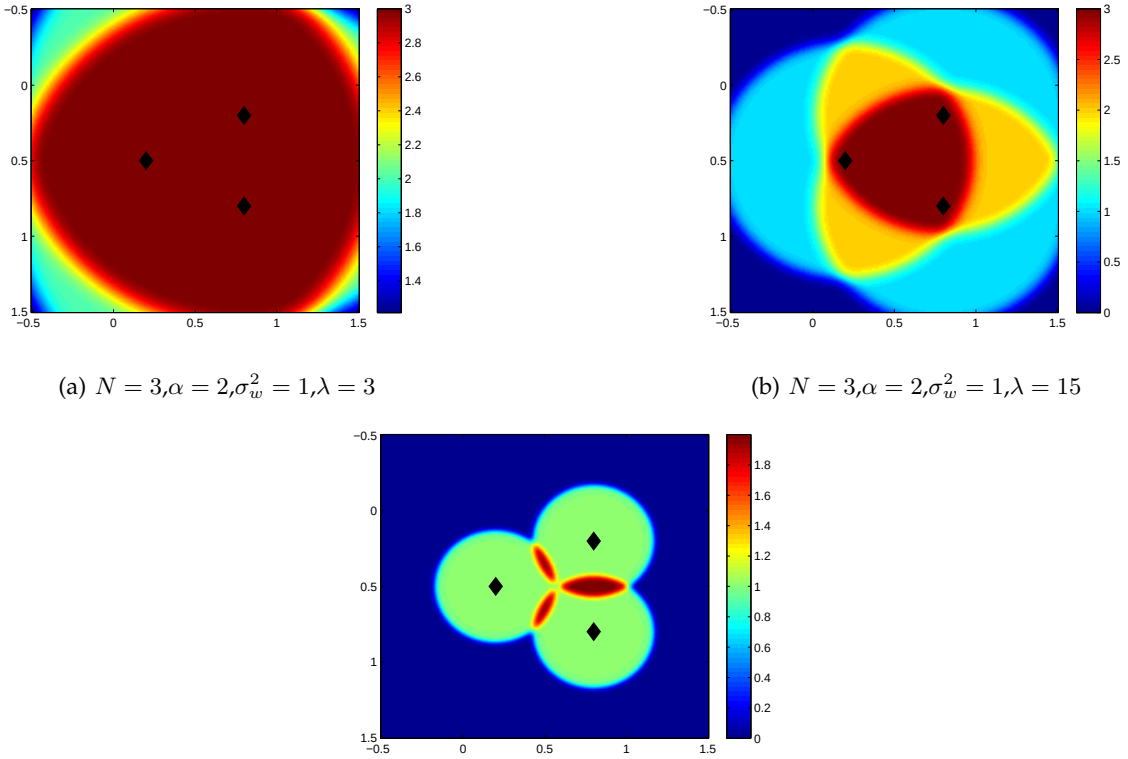
In this section we provide analysis regarding the audibility properties of the network. This helps understand the performance that can be achieved and aids design the the LOT system in a way which provides maximal coverage. To this end we define $\mathcal{K}$ as a random variable which represents the number of audible nodes at each location, conditional on the locations of the ANs. We can now calculate the probability that l ANs are audible at location θ. For example, the probability that exactly two ANs are audible when

there are $N = 3$ ANs is given as follows:

$$\begin{aligned}
&\mathrm{P}\left(\mathcal{K} = 2; T, P_T, \theta, \mathbf{x}_{1:N}, N = 3\right) \\
&= \mathrm{P}\left(\Gamma_1 = 1\right) \mathrm{P}\left(\Gamma_2 = 1\right) \mathrm{P}\left(\Gamma_3 = 0\right) \\
&+ \mathrm{P}\left(\Gamma_1 = 1\right) \mathrm{P}\left(\Gamma_2 = 0\right) \mathrm{P}\left(\Gamma_3 = 1\right) \\
&+ \mathrm{P}\left(\Gamma_1 = 0\right) \mathrm{P}\left(\Gamma_2 = 1\right) \mathrm{P}\left(\Gamma_3 = 1\right) \\
&= \left(1 - \Phi\left(s\left(\theta, \mathbf{x}_1\right)\right)\right) \left(1 - \Phi\left(s\left(\theta, \mathbf{x}_2\right)\right)\right) \Phi\left(s\left(\theta, \mathbf{x}_3\right)\right) \\
&+ \left(1 - \Phi\left(s\left(\theta, \mathbf{x}_1\right)\right)\right) \Phi\left(s\left(\theta, \mathbf{x}_2\right)\right) \left(1 - \Phi\left(s\left(\theta, \mathbf{x}_3\right)\right)\right) \\
&+ \Phi\left(s\left(\theta, \mathbf{x}_1\right)\right) \left(1 - \Phi\left(s\left(\theta, \mathbf{x}_2\right)\right)\right) \left(1 - \Phi\left(s\left(\theta, \mathbf{x}_3\right)\right)\right).
\end{aligned}$$

Lemma 3. *The average number of audible nodes at any location is given by the first moment of $\mathcal{K}$, as follows:*

$$\mathbb{E}\left[\mathcal{K} | T, P_T, \theta, \mathbf{x}_{1:N}, N\right] = \sum_{n=1}^{N} nP\left(\mathcal{K} = n | T, P_T, \theta, \mathbf{x}_{1:N}\right).$$



(a) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 3$

(b) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 15$

(c) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\lambda = 30$

Fig. 9: A map of the average number of audible nodes for the 3 ANs whose locations are indicated by $\diamond$

As we can see from the above the area of coverage indicated by the average number of audible nodes decreases rapidly as λ changes from 3 till 30. In our simulations we consider this entire range from $\lambda = 3$ to $\lambda = 30$ which illustrates the extremes in terms of the area of coverage. These audibility maps are vital to understand the results, as they illustrate the extent of coverage as a function of the threshold λ. The novelty of our approach lies in the fact that we establish a trade-off between the extent of coverage and the localisation error, by leveraging on the audibility information,which we shall see in the next section.

VIII. Simulation Results

In this section we present extensive simulation results as evidence that our framework is able to outperform the existing conventional methodology of source localisation in various scenarios.We compare the MSE with the CRLB derived in earlier sections, which helps us to the characterize the system performance.In addition we present a heuristic (partly geometrical-statistical) algorithm and we compare it to the conventional approach and our stochastic censoring approach.Finally, we discuss some important characteristics of our algorithms and their interpretation in different contexts.

A. Simulation Setup

The simulation for the two algorithms consist of the same environment, a 1×1 room, with 3 and 4 ANs whose locations are predetermined. The other parameters are, $\alpha = 2, 4$, $\mu_w = 0$, $\sigma_w^2 = 1, 2$, $\sigma_i^2 = 0.01, 1$ and $P_T = 10$. We vary the outage threshold, uniformly in $\lambda \in \{3, 6, \ldots, 30\}$ to study how the information about audibility affects the performance. We uniformly chose $J = 100$ target locations on a grid in this room and $J = 50$ realisations of the ranging observations and audibility observations were generated for each location.

B. Performance Evaluation of the Proposed Algorithms

The performance metric to evaluate the quality of the proposed algorithms is the Mean Squared Error (MSE), defined as:

$$\text{MSE} = \frac{1}{J}\frac{1}{K}\sum_{j=1}^{J}\sum_{k=1}^{K}\left\|\theta_j - \widehat{\theta}_{j,k}\right\|^2. \tag{13}$$

Remark I: All the simulation results we present are conditional on the fact that at least one AN is audible and it is due to the following reasons:

1) For the conventional algorithm, estimating the location when all the nodes are inaudible is statistically meaningless. This is because when we have no observations the MLE would result in a deterministic output, which has no statistical interpretation. Since an estimator has to be a function of the observations, we have to condition on at least one audible node.

2) Without conditioning on at least one audible node, the geometry of the space we consider would have a significant impact on the results. For example, in cases where no observations are obtained, for a circular shaped room or a rectangular shaped room, different localization results would be obtained. In order to eliminate this undesired effect and to have a geometry free results, we need this conditioning.

Let us consider the scenarios, $N = 3, 4$, $\alpha = 2, 4$, $\sigma_w^2 = 1, 4$ and $\sigma_i^2 = 0.01, 1$. Our aim is to show that the two algorithms (J-LoT and D-LoT) out-perform the conventional algorithm in any combination of the above mentioned parameters.The purpose of this comparison is to show that significant improvement can be obtained (in terms of localization accuracy), and to motivate the use of audibility information in addition to ranging measurements. We show a few instances below for each algorithm. When we talk about the conventional approach, we mean that the approach calculates the location estimates based on only the available distance observations from the audible ANs and thus ignoring information from the inaudible nodes like [50]–[54]

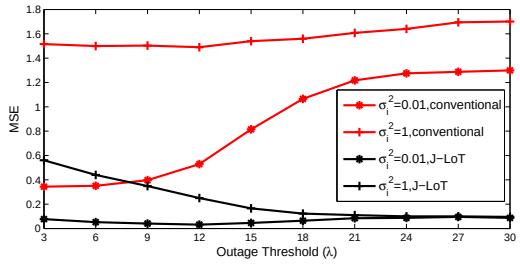
C. J-LoT: Performance comparison with Conventional approach

We begin by presenting the Conventional MLE objective function.

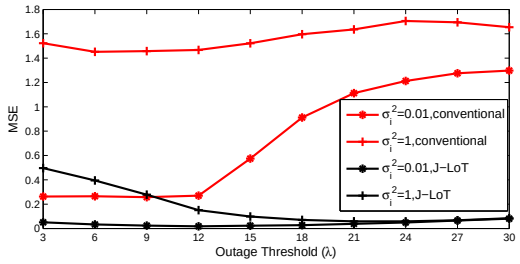
Lemma 4. *The conventional MLE for the joint estimation of the time offset and the location is given by:*

$$\left(\widehat{\theta}_{conv}, \widehat{t}_{\delta_{conv}}\right) = \arg\max_{\theta, t_\delta} \sum_{n=1}^{N} \log \mathcal{N}\left(\mathbf{y}_n; d\left(\theta, \mathbf{x}_n\right) + ct_\delta + \mu_n, \sigma_n^2\right) \times \mathbb{1}\left(\Gamma_n = 1\right)$$

This algorithm is referred to as "conventional" in the figures below. As we can clearly see from Figs. 10 and 11, J-LoT outperforms the conventional approach by a significant margin. Also this is more pronounced in the case of high variance ranging errors and this emphasizes importance of including the information on audibility in the estimation. **Remark II:** We have to exercise caution while interpreting the results shown below. The results may seem to suggest that the MSE decreases as we increase λ, but we have to bear in mind that we calculate the estimate, conditional on the fact that there is at least one audible AN and we know that λ dictates the regions of audibility as shown in the previous section. Thus, as λ increases, the area where one can find more than a single audible node keeps shrinking as seen in Fig. 9.In the wireless communication literature the value of λ is related to the transmit power P_T, and to the sensitivity of the receiver, see for example [55], [56].In simple terms, the lower the value of λ the more sensitive the receiver and therefore more expensive. In this paper we do not choose the value of λ but rather treat it as a system parameter.

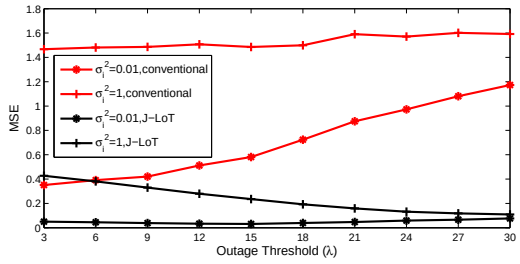


(a) $N = 3, \alpha = 2, \sigma_w^2 = 1$

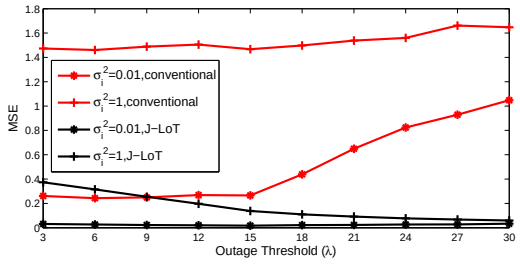


(b) $N = 4, \alpha = 2, \sigma_w^2 = 1$

Fig. 10: Comparison of how MSE varies with outage threshold (λ), for J-LoT and its conventional counterpart, for different ranging variances



(a) $N = 3, \alpha = 4, \sigma_w^2 = 4$



(b) $N = 4, \alpha = 4, \sigma_w^2 = 4$

Fig. 11: A comparison of how mse varies with outage threshold (λ), for J-LoT and its conventional counterpart,simulated for different ranging variances

D. D-LoT: Performance comparison with Conventional approach

We begin by presenting the Conventional MLE objective function.

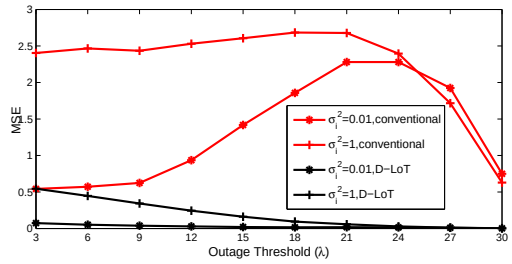
Lemma 5. *The conventional MLE for the TDOA approach is given by:*

$$\widehat{\theta}_{conv} = \arg\max_{\theta} \log \mathcal{MVN}\left(\widetilde{\mathbf{y_o}}; \tilde{\boldsymbol{d}}\left(\theta, \mathbf{x}_o\right) + \tilde{\boldsymbol{\mu}}_{\tilde{\boldsymbol{o}}_{\boldsymbol{\delta}}} + \tilde{\boldsymbol{\mu}}_{\boldsymbol{o}}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{o}}\right)$$
$$\times \mathbb{1}\left(\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{\boldsymbol{s_o}}\right)$$

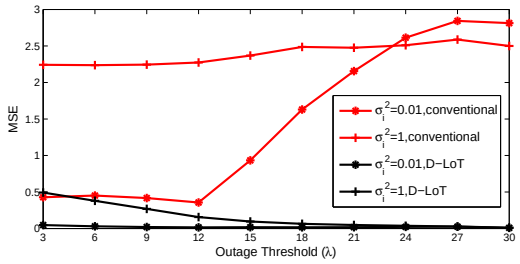
This is referred to as "conventional" in the following figures. As we are able to see from Figs 12 and 13 D-LoT significantly outperforms the conventional approach for all the settings and especially for high ranging error variance. We reiterate the fact that the results shown here are conditional on the fact that there are at least 2 audible ANs(requirement for any TDOA approach) and we have to be careful in interpreting the results as explained in the previous subsection.

To summarize the observations from D-LOT and J-LOT, we have:

1) The parameter λ significantly affects the number of audible nodes as seen from the audibility maps in fig. 9 in the previous section. $\lambda = 3$ covers a large area where 3 nodes are audible and this area

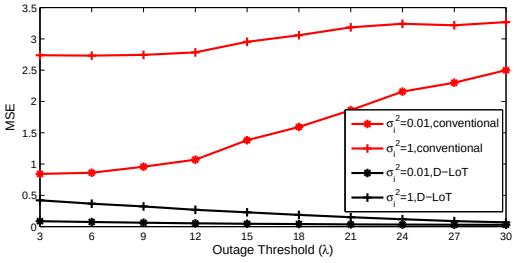


(a) $N = 3, \alpha = 2, \sigma_w^2 = 1$

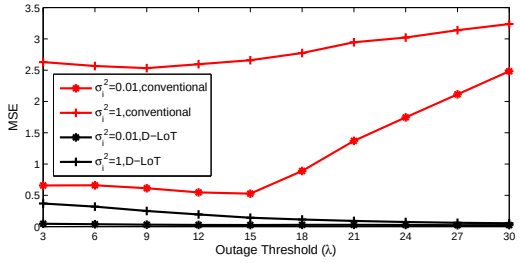


(b) $N = 4, \alpha = 2, \sigma_w^2 = 1$

Fig. 12: Comparison of how MSE varies with outage threshold (λ), for D-LoT and its conventional counterpart, for different ranging variances



(a) $N = 3, \alpha = 4, \sigma_w^2 = 4$



(b) $N = 4, \alpha = 4, \sigma_w^2 = 4$

Fig. 13: A comparison of how the MSE varies with outage threshold (λ), for D-LoT and its conventional counterpart, for different ranging variances

considerably shrinks as λ increases to 30. Our simulations are designed to cover this wide spectrum of λ.

2) We condition on at least one node being audible,because when all the nodes are inaudible, it is statistically meaningless to get the location estimate for the conventional algorithm. Also the geometry of the space under consideration would significantly impact the results without this conditioning.
3) The audibility maps are useful for interpreting the results in terms of choosing λ which can directly equate to reducing the sensitivity of the receiver or the transmitted power, thereby reducing the cost of the system. Our approach is able to exploit the audibility/inaudibility information in conjunction with these low cost transceivers or with lower Transmission Power without compromising the accuracy of the location estimate.

Remark III: We note that the MSE of the conventional method is above 1 in the above figures although the target is within 1×1 room. This is because while the size of the room is 1×1, our search region was a square of 2×2 and therefore it is possible to have such errors. This was intentionally done in order not to constrain our results by the geometry of the room. Otherwise, the shape of the room would have biased

our results.

E. Comparison between the CRLB and the MSE for J-LoT and D-LoT

The CRLB provides a theoretical lower bound on system's performance and helps understand the lower bound on the localization error (these of course depend on system's parameters). Therefore, the CRLB is a vital tool for engineers wanting to design a localisation system.

We would like to shed some light on the localisation accuracy as given by the CRLB and MSE as a function of λ as shown in Fig. 14 for J-LOT and D-LOT. The simulation results are plotted when $N = 4, \alpha = 4, \sigma_w = 2$. We observed that the localization error is close to the CRLB for high SNR (when $\epsilon = 0.01$). The localization error is far from CRB for low SNR (when $\epsilon = 1$).Thus the CRLB is mostly achievable when SNR is high. However,for any model in which the observations are are not an affine transformation of the parameters (θ), no efficient estimator exists [57].Our model, just as all other distance based ranging models, contains the Euclidean distance between the AN and the target. As such, this is a highly non-linear transformation. This means, that for all the localisation algorithms which are based on a similar model, no efficient estimator exist.

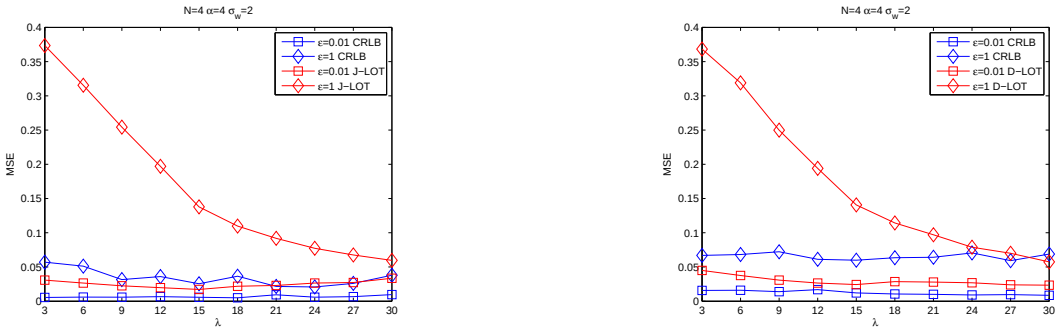
Moreover we observed that the audibility information helps affect the localization accuracy of CRLB as well as our localization accuracy. We can see from the trend that when λ increases from 3 to 30, the CRLB starts to decrease , and at $\lambda = (15 - 18)$ the CRLB reaches a minimum where audibility/inaudibility gives us significant information and beyond this, as the threshold increases the number of inaudible nodes(as seen from the audibility maps) overwhelms the system and the accuracy starts to degrade. Similar results are obtained for other values of N(the number of ANs). We also notice that, for D-LOT, when $\lambda = 30, \epsilon = 1$, the CRLB is slightly above our simulation result. This is because for high threshold, most regions in the network are inaudible. Our result is conditioned on at least one node audible, as explained previously in Remark I,but the CRLB does not condition on at least one node being audible and also it is unaware of the geometry or the bounds of the room and hence the CRLB is generally large for high values of λ.

Thus we are able to establish a trade off between the maximum allowable MSE and the minimum area of coverage that is to be ensured. It would be of interest to exploit this trade off and optimize the value of λ, given some allowed maximal localization error, but we leave this aspect for future work.

F. Comparison between a heuristic approach, the conventional and the LoT algorithms

We present a statistical-geometrical heuristic algorithm to resolve ambiguity and we describe it below:

1) Calculate the MLE according to the conventional method using only the observations from the audible nodes (as seen in Lemma 4 and Lemma 5). The MLE then returns multiple solutions (multiple maxima locations) due to the ambiguity of the likelihood surface.



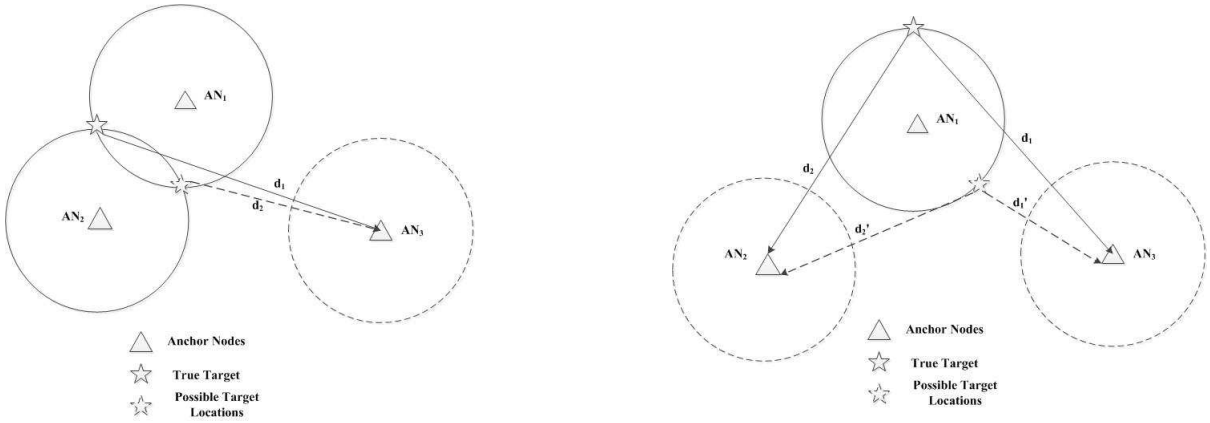
(a) $N = 4$,$\alpha = 4$,$\sigma_w^2 = 4$, J-LOT

(b) $N = 4$,$\alpha = 2$,$\sigma_w^2 = 4$, D-LOT

Fig. 14: Comparison between CRLB and simulation MSE for $N = 4, \alpha = 4, \sigma_w = 2$ for J-LOT and D-LOT

2) Calculate the sum of the distances to all inaudible nodes, and pick the solution from the MLE which maximises this.

We present a simple illustration of the heuristic algorithm for 3 AN's in fig. 15, where fig 15a represents the case with 1 inaudible node.Here we calculate the distances to the inaudible anchor node AN3 from the 2 solutions of the MLE(shown by intersecting circles). Whereas in fig 15b, we have nodes AN2 and AN3 to be inaudible and we have (multiple) possible target locations on the circle surrounding AN1. Here we calculate the sum of the distances to the inaudible nodes, and estimate the target to be the one that maximises this sum. To explain the intuition behind this heuristic approach, we present our stochastic



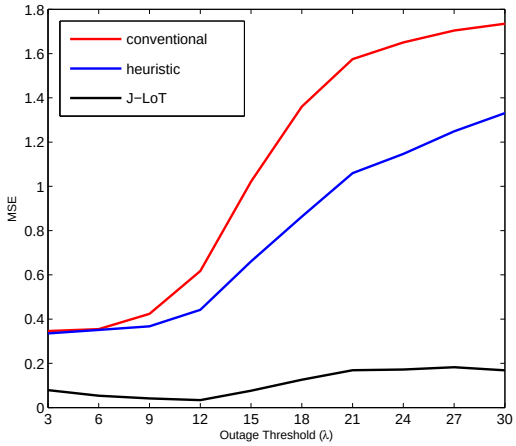
(a) 3 ANs with 1 inaudible node

(b) 3 ANs with 2 inaudible nodes

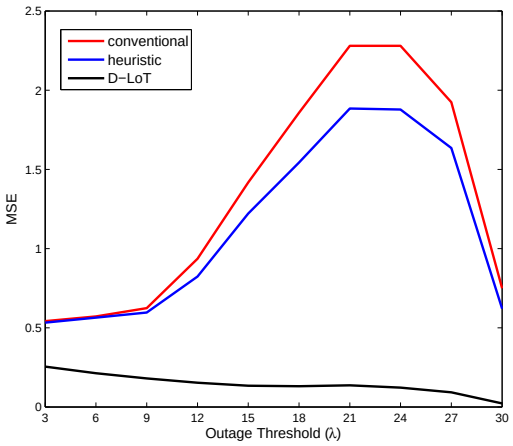
Fig. 15: Comparison of the MSE between the heuristic algorithm, the conventional approach and the J-LoT and a similar comparison for the D-LoT algorithms

censoring model given by equation 6 .The path loss model in the first equation of the model is explicitly a function of the distance between the tag and the AN, and hence we intuitively expect an AN to be

inaudible if it is further away from the tag, which is the main idea behind the heuristic algorithm. We present the MSE comparison in Fig 16.



(a) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\sigma_i^2 = 0.01$,J-LoT algorithm



(b) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$,$\sigma_i^2 = 0.01$,D-LoT algorithm

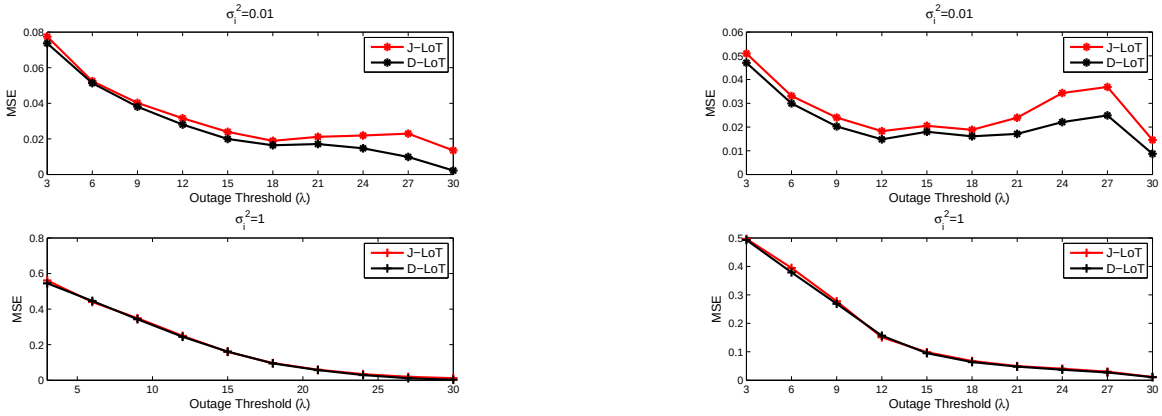
Fig. 16: Comparison of the MSE between the heuristic algorithm, the conventional approach and the J-LoT and a similar comparison for the D-LoT algorithms

1) In our statistical model, we are able to account for effects like shadowing specified by w which is very common in these environments and the heuristic is ignorant of such random effects.
2) The MLE for our model, as seen in Lemma 1 incorporates the audibility information from both the audible and the inaudible nodes in conjunction with the ranging information,however the heuristic algorithm does not incorporate the information from the nodes that are audible.We also observe that the MSE of the conventional MLE and the heuristic algorithms converge to the same value for small values of λ. This makes sense since in for small values of λ the ANs are audible and the heuristic algorithm is the same as the conventional MLE.

G. Performance comparison between J-LoT and D-LoT algorithms

Now that we have seen that J-LoT and D-LoT outperform their conventional counterparts, let us move on to compare the two algorithms. We consider the scenarios, $N = 3, 4, \alpha = 2, \sigma_w^2 = 1$ and $\sigma_i^2 = \{0.01, 1\}$. As seen in Fig. 17 , the performance of the two algorithms is almost indistinguishable and hence boils down to the very general trade-offs between the two algorithms. On the one hand, the J-LoT requires the estimation of one extra nuisance parameter, thus increasing the state space from two to three parameters $\theta_f = [\theta_x, \theta_y, t_\delta]$. This will result in an increase of computational complexity in comparison with the D-LoT algorithm. On the other hand, the D-LoT algorithm introduces dependent observations which makes the

inference more complicated and also requires at least two audible ANs, while the J-LoT can perform even if a single AN is audible.



(a) $N = 3$,$\alpha = 2$,$\sigma_w^2 = 1$ (b) $N = 4$,$\alpha = 2$,$\sigma_w^2 = 1$

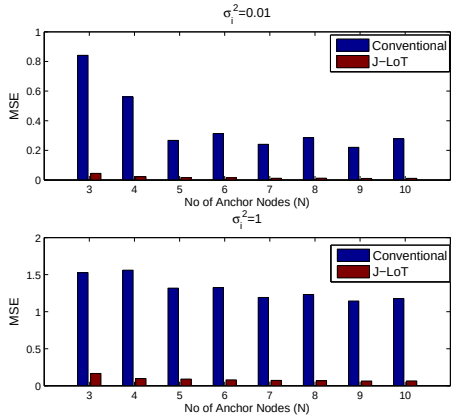
Fig. 17: A comparison of how mse varies with outage threshold (λ), for J-LoT and D-LoT, simulated for different ranging variances

H. Comparison between the number of ANs

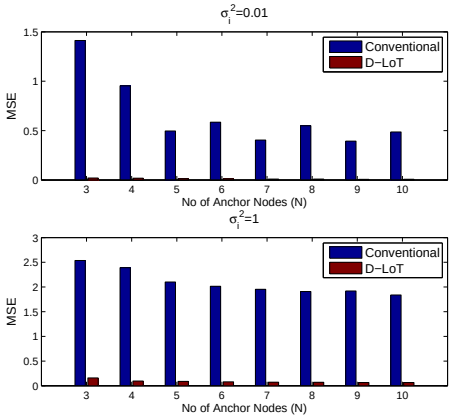
Let us consider the following scenario to compare how the performance varies across the number of ANs: $N = [3, 4, 5 \ldots, 10]$, $\alpha = 2$, $\sigma_w^2 = 1$, $\sigma_i^2 = \{1, 0.01\}$ and $\lambda = 15$. Patterns similar to that of Fig. 18 were observed in other scenarios as well. This seems to suggest that even though the conventional approach does significantly improve with increasing number of ANs, our approach still performs much better, albeit with the difference between the two approaches decreasing with increasing number of ANs.

IX. Conclusion

In this paper we developed a novel approach for Geo-Spatial tagging in the Internet of Things (IoT). We first demonstrated how conventional OW-TOA localization algorithms may produce incorrect location estimates in the presence of inaudible ANs. We then showed that we can make use of the audibility information (which indicates whether an AN is able or unable to communicate with the tags). By leveraging on this available information, we re-formulated the localization problem as a statistical non-linear estimation problem under the joint estimation approach (J-LoT) and the Time Difference of Arrival approach (D-LoT). Our methodology provided considerable improvement of the localization performance by mitigating the well-known ambiguity problem which arose when only a few ANs were audible. In addition, we derived the spatial Cramér-Rao lower Bound of the source location estimate under the proposed framework. Finally



(a) $\alpha = 2$,$\sigma_w^2 = 1$ and $\sigma_i^2 = 1, 0.01$ and $\lambda = 15$



(b) $\alpha = 2$,$\sigma_w^2 = 1$ and $\sigma_i^2 = 1, 0.01$ and $\lambda = 15$

Fig. 18: Comparison of the MSE vs the number of ANs for J-LoT and D-LoT, against their conventional counterparts

we showed that these results can also be interpreted in terms of a design problem to choose particular parameters of the system constrained by the requirements of the environment.

APPENDIX A
PROOF OF LEMMA 1

$$
\begin{aligned}
\left(\widehat{\theta}, \widehat{t}_o\right) &= \arg\max_{\theta, t_\delta} p\left(y_{1:N}, \Gamma_{1:N} | \theta, t_\delta\right) \\
&= \arg\max_{\theta, t_\delta} p\left(y_{1:N} | \Gamma_{1:N}, \theta, t_\delta\right) \mathrm{P}\left(\Gamma_{1:N} | \theta, t_\delta\right) \\
&= \arg\max_{\theta, t_\delta} \prod_{n=1}^{N} p\left(y_n | \Gamma_n, \theta, t_\delta\right) \mathrm{P}\left(\Gamma_n | \theta, t_\delta\right) \\
&= \arg\max_{\theta, t_\delta} \sum_{n=1}^{N} \left(\log p\left(y_n | \Gamma_n, \theta, t_\delta\right) + \log \mathrm{P}\left(\Gamma_n | \theta, t_\delta\right)\right) \\
&= \arg\max_{\theta, t_\delta} \sum_{n=1}^{N} \log \mathcal{N}\left(y_n; d\left(\theta, \mathbf{x}_n\right) + c t_\delta + \mu_n, \sigma_n^2\right) \mathbb{1}\left(\Gamma_n = 1\right) \\
&\qquad + \sum_{n=1}^{N} \log \mathrm{P}\left(P_{R_n} > \lambda | \theta\right) \mathbb{1}\left(\Gamma_n = 1\right) \\
&\qquad + \sum_{n=1}^{N} \log \mathrm{P}\left(P_{R_n} < \lambda | \theta\right) \mathbb{1}\left(\Gamma_n = 0\right) \\
&= \arg\max_{\theta, t_\delta} \sum_{n=1}^{N} \log \mathcal{N}\left(y_n; d\left(\theta, \mathbf{x}_n\right) + c t_\delta + \mu_n, \sigma_n^2\right) \mathbb{1}\left(\Gamma_n = 1\right) \\
&\qquad + \sum_{n=1}^{N} \log \left(1 - \Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \\
&\qquad \times \mathbb{1}\left(\Gamma_n = 1\right) \\
&\qquad + \sum_{n=1}^{N} \log \left(\Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \\
&\qquad \times \mathbb{1}\left(\Gamma_n = 0\right)
\end{aligned}
\tag{14}
$$

APPENDIX B
PROOF OF LEMMA 2

$$
\begin{aligned}
\widehat{\theta} &= \arg\max_{\theta} p\left(\widetilde{\mathbf{y}_{\boldsymbol{o}2:N_o}}, \Gamma_{1:N}|\theta\right) \\
&= \arg\max_{\theta} p\left(\widetilde{y_{o}}_{2:N_o}|\Gamma_{1:N}, \theta\right) \mathrm{P}\left(\Gamma_{1:N}|\theta\right) \\
&= \arg\max_{\theta} \log p\left(\widetilde{\mathbf{y}_{\boldsymbol{o}2:N_o}}|\Gamma_{1:N}, \theta\right) \\
&\qquad + \sum_{n=1}^{N} \log\left(1 - \Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \\
&\qquad \times \mathbb{1}\left(\Gamma_n = 1\right) \\
&\qquad + \sum_{n=1}^{N} \log\left(\Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \\
&\qquad \times \mathbb{1}\left(\Gamma_n = 0\right) \\
&= \arg\max_{\theta} \log \mathcal{MVN}\left(\widetilde{\mathbf{y}_{\boldsymbol{o}}}; \tilde{\boldsymbol{d}}\left(\theta, \mathbf{x}_o\right) + \boldsymbol{\mu}_{\boldsymbol{o\delta}}^{\sim} + \boldsymbol{\mu}_{\boldsymbol{o}}^{\sim}, \boldsymbol{\Sigma}_{\boldsymbol{o}}^{\sim}\right) \mathbb{1}\left(\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{\boldsymbol{s}_o}\right) \\
&\qquad + \sum_{n=1}^{N} \log\left(1 - \Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \\
&\qquad \times \mathbb{1}\left(\Gamma_n = 1\right) \\
&\qquad + \sum_{n=1}^{N} \log\left(\Phi\left(\lambda - \left(P_T - 10\alpha \log \frac{d\left(\theta, \mathbf{x}_n\right)}{d_0}\right)\right)\right) \\
&\qquad \times \mathbb{1}\left(\Gamma_n = 0\right).
\end{aligned}
\tag{15}
$$

APPENDIX C
PROOF OF THEOREM 1

A. Fisher Information Matrix

The Fisher Information Matrix for the first term:

$$
\begin{aligned}
F_n^{(1)} &= -\mathbb{E}_{y_n,\Gamma_n}\left[\frac{\mathrm{d}\left(\log\mathcal{N}\left(y_n; d\left(\theta,\mathbf{x}_n\right)+ct_\delta+\mu_n,\sigma_n^2\right)\mathbb{1}\left(\Gamma_n=1\right)\right)}{\mathrm{d}\theta_f^2}\right] \\
&= \frac{1}{\sigma_n^2}\Bigg(-\mathbb{E}_{y_n,\Gamma_n}\Bigg[-\mathbb{1}\left(\Gamma_n=1\right)\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f} \\
&\qquad\qquad \times\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f}\Bigg]\Bigg) \\
&+ \frac{1}{\sigma_n^2}\left(-\mathbb{E}_{y_n,\Gamma_n}\left[\mathbb{1}\left(\Gamma_n=1\right)\left(y_n-\beta_n\right)\frac{\mathrm{d}^2\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f^2}\right]\right) \\
&= \frac{1}{\sigma_n^2}\sum_{\Gamma_n}\Bigg(\int_{-\infty}^{+\infty}\mathbb{1}\left(\Gamma_n=1\right)\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f} \\
&\qquad\qquad \times\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f}\mathcal{N}\left(y_n;\beta_n,\sigma_n^2\right)dy_n\Bigg)\mathrm{P}\left(\Gamma_n=1\right) \\
&+ \frac{1}{\sigma_n^2}\sum_{\Gamma_n}\Bigg(\int_{-\infty}^{+\infty}-\mathbb{1}\left(\Gamma_n=1\right)\left(y_n-\beta_n\right)\frac{\mathrm{d}^2\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f^2} \\
&\qquad\qquad \times\mathcal{N}\left(y_n;\beta_n,\sigma_n^2\right)dy_n\Bigg)\mathrm{P}\left(\Gamma_n=1\right) \\
&= \frac{1}{\sigma_n^2}\sum_{\Gamma_n}\mathbb{1}\left(\Gamma_n=1\right)\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f}\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f} \\
&\qquad\qquad \times\mathrm{P}\left(P_{R_n}>\lambda\right)
\end{aligned}
$$

The FIM of the first term is calculated by summing over all the ANs, as follows:

$$
\begin{aligned}
F^{(1)} = \sum_{n=1}^{N}\frac{1}{\sigma_n^2}&\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f}\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f} \\
&\times\left(1-\Phi\left(\lambda-\left(P_T-10\alpha\log\frac{d\left(\theta,\mathbf{x}_n\right)}{d_0}\right)\right)\right)
\end{aligned}
$$

where $\beta_n = d\left(\theta,\mathbf{x}_n\right)+ct_\delta+\mu_n$

$$
\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f}\left(\frac{\mathrm{d}\left(d\left(\theta,\mathbf{x}_n\right)+ct_\delta\right)}{\mathrm{d}\theta_f}\right)^T =
$$

$$
\begin{bmatrix}
\frac{(\theta_x-x_n)^2}{d^2(\theta,\mathbf{x}_n)} & \frac{(\theta_x-x_n)(\theta_y-y_n)}{d^2(\theta,\mathbf{x}_n)} & c\frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)} \\
\frac{(\theta_x-x_n)(\theta_y-y_n)}{d^2(\theta,\mathbf{x}_n)} & \frac{(\theta_y-y_n)^2}{d^2(\theta,\mathbf{x}_n)} & c\frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)} \\
c\frac{(\theta_x-x_n)}{d(\theta,\mathbf{x}_n)} & c\frac{(\theta_y-y_n)}{d(\theta,\mathbf{x}_n)} & c^2
\end{bmatrix}.
$$

The FIM for the n-th element in the second term is given by:

$$
\begin{aligned}
F_n^{(2)} &= -\mathbb{E}_{y_n,\Gamma_n}\left[\frac{\mathrm{d}^2\left(\log\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\mathbb{1}\left(\Gamma_n=0\right)\right)}{\mathrm{d}\theta^2}\right] \\
&= -\mathbb{E}_{y_n,\Gamma_n}\Big[\Big(\frac{-s\left(\theta,\mathbf{x}_n\right)\phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)}{\Phi^2\left(s\left(\theta,\mathbf{x}_n\right)\right)} \\
&+ \frac{-\phi^2\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}}{\Phi^2\left(s\left(\theta,\mathbf{x}_n\right)\right)}\Big)\frac{\mathrm{d}s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta}\mathbb{1}\left(\Gamma_n=0\right)\Big] \\
&- \mathbb{E}_{y_n,\Gamma_n}\left[\frac{\phi\left(s\left(\theta,\mathbf{x}_n\right)\right)}{\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)}\frac{\mathrm{d}^2 s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta^2}\mathbb{1}\left(\Gamma_n=0\right)\right] \\
&= \Big(\frac{s\left(\theta,\mathbf{x}_n\right)\phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)}{\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)} \\
&+ \frac{\phi^2\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}}{\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)}\Big)\frac{\mathrm{d}s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta} \\
&- \phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}^2 s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta^2}.
\end{aligned}
$$

The FIM of the second term is calculated by summing over all the ANs, as follows:

$$
\begin{aligned}
F^{(2)} = \sum_{n=1}^{N}\Big(\Big(&\frac{s\left(\theta,\mathbf{x}_n\right)\phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)}{\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)} \\
&+ \frac{\phi^2\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}}{\Phi\left(s\left(\theta,\mathbf{x}_n\right)\right)}\Big)\frac{\mathrm{d}s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta}\Big) \\
&- \sum_{n=1}^{N}\left(\phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}^2 s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta^2}\right).
\end{aligned}
$$

In a similar way the FIM of the third term is given by:

$$
\begin{aligned}
F^{(3)} = \sum_{n=1}^{N}\Big(\Big(&\frac{-s\left(\theta,\mathbf{x}_n\right)\phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}\bar{\Phi}\left(s\left(\theta,\mathbf{x}_n\right)\right)}{\bar{\Phi}\left(s\left(\theta,\mathbf{x}_n\right)\right)} \\
&+ \frac{\phi^2\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}s(\theta,\mathbf{x}_n)}{\mathrm{d}\theta}}{\bar{\Phi}\left(s\left(\theta,\mathbf{x}_n\right)\right)}\Big)\frac{\mathrm{d}s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta}\Big) \\
&+ \sum_{n=1}^{N}\left(\phi\left(s\left(\theta,\mathbf{x}_n\right)\right)\frac{\mathrm{d}^2 s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta^2}\right).
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\mathrm{d}^2 s\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta_f^2} = \frac{\mathrm{d}\frac{10\alpha}{d(\theta,\mathbf{x}_n)}}{\mathrm{d}\theta_f}\frac{\mathrm{d}d(\theta,\mathbf{x}_n)}{\mathrm{d}\theta_f} &= \frac{-10\alpha}{d^2\left(\theta,\mathbf{x}_n\right)}\left(\frac{\mathrm{d}d\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta_f}\right)^2 \\
&+ \frac{10\alpha}{d\left(\theta,\mathbf{x}_n\right)}\frac{\mathrm{d}^2 d\left(\theta,\mathbf{x}_n\right)}{\mathrm{d}\theta_f^2},
\end{aligned}
$$

and

$$\frac{\mathrm{d}^2 d\left(\theta, \mathbf{x}_n\right)}{\mathrm{d}\theta_f^2} =$$

$$\frac{1}{d^3\left(\theta, \mathbf{x}_n\right)} \times$$

$$\begin{bmatrix} \left(\theta_y - y_n\right)^2 & -\left(\theta_y - y_n\right)\left(\theta_x - x_n\right) & 0 \\ -\left(\theta_y - y_n\right)\left(\theta_x - x_n\right) & \left(\theta_x - x_n\right)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\frac{\mathrm{d}s\left(\theta, \mathbf{x}_n\right)}{\mathrm{d}\theta_f} = \frac{10\alpha}{d\left(\theta, \mathbf{x}_n\right)} \frac{\mathrm{d}d\left(\theta, \mathbf{x}_n\right)}{\mathrm{d}\theta_f}.$$

$$\frac{\mathrm{d}d\left(\theta, \mathbf{x}_n\right)}{\mathrm{d}\theta_f} = \begin{bmatrix} \dfrac{\theta_x - x_n}{d\left(\theta, \mathbf{x}_n\right)} & \dfrac{\theta_y - y_n}{d\left(\theta, \mathbf{x}_n\right)} & 0 \end{bmatrix}^T.$$

The complete Fisher information matrix is thus given by combining all the terms $F = F^{(1)} + F^{(2)} + F^{(3)}$, which is written in a compact form as specified in Theorem 1

APPENDIX D
PROOF OF THEOREM 2

Lemma 6. *The FIM for a N-variate multivariate normal distribution has a special form. Let $\mu(\theta) = [\mu_1(\theta), \mu_2(\theta), \cdots, \mu_N(\theta)]^T$, and let $\Sigma(\theta)$ be the covariance matrix. Then the typical element $F_{m,n}, 0 \leq m, n < K$ of the FIM for $X \sim N(\mu(\theta), \Sigma(\theta))$ is:*

$$F_{m,n} = \frac{\partial \mu^T}{\partial \theta_m} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_n} + \frac{1}{2} \mathrm{tr}\left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_m} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_n}\right) \tag{16}$$

where $(\cdots)^T$ denotes the transpose of a vector, $\mathrm{tr}(\cdots)$ denotes the trace of a square matrix, and

$$\frac{\partial \mu}{\partial \theta_m} = \left[\frac{\partial \mu_1}{\partial \theta_m} \frac{\partial \mu_2}{\partial \theta_m} \cdots \frac{\partial \mu_N}{\partial \theta_m}\right]^T$$

$$\frac{\partial \Sigma}{\partial \theta_m} = \begin{bmatrix} \frac{\partial \Sigma_{1,1}}{\partial \theta_m} & \frac{\partial \Sigma_{1,2}}{\partial \theta_m} & \cdots & \frac{\partial \Sigma_{1,N}}{\partial \theta_m} \\ \frac{\partial \Sigma_{2,1}}{\partial \theta_m} & \frac{\partial \Sigma_{2,2}}{\partial \theta_m} & \cdots & \frac{\partial \Sigma_{2,N}}{\partial \theta_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \Sigma_{N,1}}{\partial \theta_m} & \frac{\partial \Sigma_{N,2}}{\partial \theta_m} & \cdots & \frac{\partial \Sigma_{N,N}}{\partial \theta_m} \end{bmatrix}.$$

Note that a special, but very common case is the one when $\Sigma(\theta) = \Sigma$ is constant. Then

$$F_{m,n} = \frac{\partial \mu^T}{\partial \theta_m} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_n} \tag{17}$$

Using the results of the above defined Lemma, we proceed with the derivation of the CRLB for the D-LoT framework as follows:

According to (12)

$$F = -\mathbb{E}_{\widetilde{\mathbf{y}_o},\mathbf{\Gamma}}\left[\frac{\mathrm{d}^2 \log p\left(\widetilde{\mathbf{y}_{o2:N_o}}, \Gamma_{1:N}|\theta\right)}{\mathrm{d}\theta^2}\right]$$

$$F = -\mathbb{E}_{\widetilde{\mathbf{y}_o},\mathbf{\Gamma}}\left[\frac{\mathrm{d}^2 \log p\left(\widetilde{\mathbf{y}_{o2:N_o}}|\Gamma_{1:N}, \theta\right) p\left(\Gamma_{1:N}|\theta\right)}{\mathrm{d}\theta^2}\right]$$

$$F = -\mathbb{E}_{\widetilde{\mathbf{y}_o},\mathbf{\Gamma}}\left[\frac{\mathrm{d}^2 \log p\left(\widetilde{\mathbf{y}_{o2:N_o}}|\Gamma_{1:N}, \theta\right)}{\mathrm{d}\theta^2}\right] - \mathbb{E}_{\widetilde{\mathbf{y}_o},\mathbf{\Gamma}}\left[\frac{\mathrm{d}^2 \log p\left(\Gamma_{1:N}, \theta\right)}{\mathrm{d}\theta^2}\right].$$

The FIM of the second term in the above mentioned sum is same as the one derived in the previous section. We use the results of Lemma 6, to derive the first term of the above defined CRLB denoted by $F^{(1)}$, which involves a Multivariate Normal distribution, under the special case that Σ is a constant. Let us denote $\boldsymbol{\omega_o} = \tilde{\boldsymbol{d}}\left(\theta, \mathbf{x_o}\right) + \tilde{\boldsymbol{\mu}}_{\boldsymbol{\delta}} + \tilde{\boldsymbol{\mu}}_{\boldsymbol{o}}$

$$\begin{aligned}
F^{(1)} &= -\mathbb{E}_{\widetilde{\mathbf{y}_o},\mathbf{\Gamma}}\left[\frac{\mathrm{d}^2 \log p\left(\widetilde{y}_{o2:N_o}|\Gamma_{1:N}, \theta\right)}{\mathrm{d}\theta^2}\right] \\
&= -\mathbb{E}_{\widetilde{\mathbf{y}_o},\mathbf{\Gamma}}\left[\frac{\mathrm{d}^2 \log \mathcal{MVN}\left(\widetilde{\boldsymbol{y_o}}; \tilde{\boldsymbol{d}}\left(\theta, \mathbf{x}_o\right) + \tilde{\boldsymbol{\mu}}_{\delta} + \tilde{\boldsymbol{\mu}}_{\boldsymbol{o}}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{o}}\right) \mathbb{1}\left(\mathbf{\Gamma} = \mathbf{\Gamma}_{\boldsymbol{s_o}}\right)}{\mathrm{d}\theta^2}\right] \\
&= -\sum_{o=1}^{2^N-N-1}\left[\int_{-\infty}^{+\infty} \frac{\mathrm{d}^2 \log \mathcal{MVN}\left(\widetilde{\boldsymbol{y_o}}; \boldsymbol{\omega_o}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{o}}\right) \mathbb{1}\left(\mathbf{\Gamma} = \mathbf{\Gamma}_{\boldsymbol{s_o}}\right)}{\mathrm{d}\theta^2} \right. \\
&\qquad \left. \times\, p\left(\widetilde{y}_{o2:N_o}|\Gamma_{1:N}, \theta\right) \mathrm{d}\widetilde{\mathbf{y}_o}\right] \\
&= \sum_{o=1}^{2^N-N-1} \tilde{\boldsymbol{\Lambda}}_{\boldsymbol{o}} \mathbb{1}\left(\mathbf{\Gamma} = \mathbf{\Gamma}_{\boldsymbol{s_o}}\right) P\left(\mathbf{\Gamma} = \mathbf{\Gamma}_{\boldsymbol{s_o}}\right) \\
&= \sum_{o=1}^{2^N-N-1} \tilde{\boldsymbol{\Lambda}}_{\boldsymbol{o}} \prod_{i=1}^{N} P\left(\Gamma_{s_{o_i}} = 1\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=1\right)} P\left(\Gamma_{s_{o_i}} = 0\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=0\right)} \\
&= \sum_{o=1}^{2^N-N-1} \tilde{\boldsymbol{\Lambda}}_{\boldsymbol{o}} \prod_{i=1}^{N} P\left(P_{R_i} > \lambda\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=1\right)} P\left(P_{R_i} < \lambda\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=0\right)} \\
&= \sum_{o=1}^{2^N-N-1} \tilde{\boldsymbol{\Lambda}}_{\boldsymbol{o}} \prod_{i=1}^{N} \left(\bar{\Phi}\left(s\left(\theta, \mathbf{x}_i\right)\right)\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=1\right)} \left(\Phi\left(s\left(\theta, \mathbf{x}_i\right)\right)\right)^{\mathbb{1}\left(\Gamma_{s_{o_i}}=0\right)},
\end{aligned}$$

where $s\left(\theta, \mathbf{x}_i\right) = \left(\lambda - \left(P_T - 10\alpha \log \frac{d(\theta, \mathbf{x}_i)}{d_0}\right)\right)$
and $\tilde{\mathbf{\Lambda}}_{\boldsymbol{o_{m,n}}} = \left[\frac{\partial(\boldsymbol{\omega_o})^T}{\partial \theta_m} \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{o}}^{-1} \frac{\partial(\boldsymbol{\omega_o})}{\partial \theta_n}\right]$ (Using Lemma 6)
and $m, n = \{x, y\} \in \chi \subseteq \mathbb{R}^2$ which is the parameter space. Also there are $2^N - N - 1$ observable states for this approach as it requires at least two observations to calculate the location estimate after performing the difference operation.

$$\frac{\partial\left(\boldsymbol{\omega_o}\right)}{\partial \theta_x} = \left[\begin{array}{cccc} \frac{(\theta_x - x_{o_2})}{d\left(\theta, \mathbf{x}_{o_2}\right)} - \frac{(\theta_x - x_{o_1})}{d\left(\theta, \mathbf{x}_{o_1}\right)} & \cdots & \frac{\left(\theta_x - x_{o_{N_o}}\right)}{d\left(\theta, \mathbf{x}_{o_{N_o}}\right)} - \frac{(\theta_x - x_{o_1})}{d\left(\theta, \mathbf{x}_{o_1}\right)} \end{array}\right]^T,$$

and

$$\frac{\partial\left(\boldsymbol{\omega_o}\right)}{\partial \theta_y} = \left[\begin{array}{cccc} \frac{(\theta_y - y_{o_2})}{d\left(\theta, \mathbf{x}_{o_2}\right)} - \frac{(\theta_y - y_{o_1})}{d\left(\theta, \mathbf{x}_{o_1}\right)} & \cdots & \frac{\left(\theta_y - y_{o_{N_o}}\right)}{d\left(\theta, \mathbf{x}_{o_{N_o}}\right)} - \frac{(\theta_y - y_{o_1})}{d\left(\theta, \mathbf{x}_{o_1}\right)} \end{array}\right]^T$$

The complete Fisher information matrix is thus given by combining all the terms $F = F^{(1)} + F^{(2)} + F^{(3)}$, which is written in a compact form as specified in Theorem 2

REFERENCES

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[2] O. Vermesan, P. Friess, P. Guillemin, S. Gusmeroli, H. Sundmaeker, A. Bassi, I. S. Jubert, M. Mazura, M. Harrison, M. Eisenhauer *et al.*, "Internet of things strategic research roadmap," *O. Vermesan, P. Friess, P. Guillemin, S. Gusmeroli, H. Sundmaeker, A. Bassi, et al., Internet of Things: Global Technological and Societal Trends*, vol. 1, pp. 9–52, 2011.

[3] D. Bandyopadhyay and J. Sen, "Internet of things: Applications and challenges in technology and standardization," *Wireless Personal Communications*, vol. 58, no. 1, pp. 49–69, 2011.

[4] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges," *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497–1516, 2012.

[5] I. Chatzigiannakis, J. P. Drude, H. Hasemann, and A. Kröller, "Developing smart homes using the internet of things: How to demonstrate your system," in *Distributed, Ambient, and Pervasive Interactions*. Springer, 2014, pp. 415–426.

[6] M. Swan, "Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0," *Journal of Sensor and Actuator Networks*, vol. 1, no. 3, pp. 217–253, 2012.

[7] D. Uckelmann, M. Harrison, and F. Michahelles, *Architecting the internet of things*. Springer Science & Business Media, 2011.

[8] W.-f. SUN, Z. WANG, H.-y. MA, Y. ZHAO, and Q.-r. LI, "Key technologies and typical military application of iot [j]," *Internet of Things Technologies*, vol. 4, p. 033, 2012.

[9] J. Schiller and A. Voisard, *Location-based services*. Elsevier, 2004.

[10] B. Rao and L. Minakakis, "Evolution of mobile location-based services," *Communications of the ACM*, vol. 46, no. 12, pp. 61–65, 2003.

[11] K. Gao, Q. Wang, and L. Xi, "Controlling moving object in the internet of things*," *IJACT: International Journal of Advancements in Computing Technology*, vol. 4, no. 5, pp. 83–90, 2012.

[12] E. Arias-de Reyna and P. M. Djuric, "Indoor localization with range-based measurements and little prior information," *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1979–1987, 2013.

[13] W. R. Jung, S. Bell, A. Petrenko, and A. Sizo, "Potential risks of wifi-based indoor positioning and progress on improving localization functionality," in *Proceedings of the Fourth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*. ACM, 2012, pp. 13–20.

[14] "Decawave web-site," 09 2015. [Online]. Available: http://www.decawave.com/

[15] "Pixie web-site," 09 2015. [Online]. Available: https://www.getpixie.com/

[16] 2015. [Online]. Available: http://onsolution.com.au/wireless-tags/

[17] "Lapa: Bluetooth item finder," 09 2015. [Online]. Available: http://lapa-app.com/

[18] "Tile: Bluetooth tracking gadjet," 09 2015. [Online]. Available: https://www.thetileapp.com/

[19] J. C. Adams, W. Gregorwich, L. Capots, and D. Liccardo, "Ultra-wideband for navigation and communications," in *Aerospace Conference, 2001, IEEE Proceedings.*, vol. 2. IEEE, 2001, pp. 2–785.

[20] C.-C. Chong, F. Watanabe, and H. Inamura, "Potential of uwb technology for the next generation wireless communications," in *2006 IEEE Ninth International Symposium on Spread Spectrum Techniques and Applications,*. IEEE, 2006, pp. 422–429.

[21] I. Guvenc and Z. Sahinoglu, "Threshold-based toa estimation for impulse radio uwb systems," in *2005 IEEE International Conference on Ultra-Wideband, 2005. ICU 2005.* IEEE, 2005, pp. 420–425.

[22] M. Gorlatova, P. Kinget, I. Kymissis, D. Rubenstein, X. Wang, and G. Zussman, "Energy harvesting active networked tags (enhants) for ubiquitous object networking," *Wireless Communications, IEEE*, vol. 17, no. 6, pp. 18–25, 2010.

[23] D. Dardari, R. D. Errico, C. Roblin, A. Sibille, and M. Z. Win, "Ultrawide bandwidth rfid: The next generation?" *Proceedings of the IEEE*, vol. 98, no. 9, pp. 1570–1582, 2010.

[24] Y. Zhou, C. L. Law, and J. Xia, "Ultra low-power uwb-rfid system for precise location-aware applications," in *Wireless Communications and Networking Conference Workshops (WCNCW), 2012 IEEE*. IEEE, 2012, pp. 154–158.

[25] C. C. Cruz, J. R. Costa, and C. A. Fernandes, "Hybrid uhf/uwb antenna for passive indoor identification and localization systems," *IEEE Transactions on Antennas and Propagation,*, vol. 61, no. 1, pp. 354–361, 2013.

[26] F. Guidi, N. Decarli, D. Dardari, C. Roblin, and A. Sibille, "Performance of uwb backscatter modulation in multi-tag rfid scenario using experimental data," in *Ultra-Wideband (ICUWB), 2011 IEEE International Conference on*. IEEE, 2011, pp. 484–488.

[27] I. F. Akyildiz and J. M. Jornet, "The internet of nano-things," *Wireless Communications, IEEE*, vol. 17, no. 6, pp. 58–63, 2010.

[28] Z. A. lu, S. Gezici, and G. Ismail, *Ultra-Wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*. Cambridge University Press, 2008.

[29] S. Gezici and H. V. Poor, "Position estimation via ultra-wide-band signals," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 386–403, 2009.

[30] S. Gezici, "A survey on wireless position estimation," *Wireless personal communications*, vol. 44, no. 3, pp. 263–282, 2008.

[31] E. Xu, Z. Ding, and S. Dasgupta, "Source localization in wireless sensor networks from signal time-of-arrival measurements," *IEEE Transactions on Signal Processing,*, vol. 59, no. 6, pp. 2887–2897, 2011.

[32] P. Veeranath, D. Rao, S. Vathsal, and N. Bhasker, "Reducing multipath effects in indoor channel for analysis of gps/pseudolite signal acquisition," *vol*, vol. 3, pp. 1–6, 2013.

[33] B. Zhen, H.-B. Li, and R. Kohno, "Clock management in ultra-wideband ranging," in *Mobile and Wireless Communications Summit, 2007. 16th IST*. IEEE, 2007, pp. 1–5.

[34] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 404–426, 2009.

[35] B. T. Fang, "Simple solutions for hyperbolic and related position fixes," *IEEE Transactions on Aerospace and Electronic Systems,*, vol. 26, no. 5, pp. 748–753, 1990.

[36] Y. Wang, X. Ma, and G. Leus, "Robust time-based localization for asynchronous networks," *IEEE Transactions on Signal Processing,*, vol. 59, no. 9, pp. 4397–4410, 2011.

[37] M. M. Saad, C. Bleakley, M. Walsh, and T. Ye, "High accuracy location estimation for a mobile tag using one-way uwb signaling," in *Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS), 2012*. IEEE, 2012, pp. 1–8.

[38] Y. Shang, W. Ruml, Y. Zhang, and M. P. Fromherz, "Localization from mere connectivity," in *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*. ACM, 2003, pp. 201–212.

[39] F. Xiao, C. Sha, L. Chen, L. Sun, and R. Wang, "Noise-tolerant localization from incomplete range measurements for wireless sensor networks," in *2015 IEEE Conference on Computer Communications (INFOCOM),*, April 2015, pp. 2794–2802.

[40] C. Beder and M. Klepal, "Fingerprinting based localisation revisited: A rigorous approach for comparing rssi measurements coping with missed access points and differing antenna attenuations," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN),*, Nov 2012, pp. 1–7.

[41] R. J. Little and D. B. Rubin, *Statistical analysis with missing data*. John Wiley & Sons, 2014.

[42] F. Sivrikaya and B. Yener, "Time synchronization in sensor networks: a survey," *Network, IEEE*, vol. 18, no. 4, pp. 45–50, 2004.

[43] T. Sathyan, D. Humphrey, and M. Hedley, "Wasp: A system and algorithms for accurate radio localization using low-cost hardware," *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, vol. 41, no. 2, pp. 211–222, 2011.

[44] D. D. McCrady, L. Doyle, H. Forstrom, T. Dempsey, and M. Martorana, "Mobile ranging using low-accuracy clocks," *IEEE Transactions on Microwave Theory and Techniques,*, vol. 48, no. 6, pp. 951–958, 2000.

[45] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. prentice hall PTR New Jersey, 1996, vol. 2.

[46] D. B. Rubin, "Inference and missing data," *Biometrika*, vol. 63, no. 3, pp. 581–592, 1976.

[47] F. D. Nelson, "Censored regression models with unobserved, stochastic censoring thresholds," *Journal of econometrics*, vol. 6, no. 3, pp. 309–327, 1977.

[48] E. Marubini and M. G. Valsecchi, *Analysing survival data from clinical trials and observational studies*. John Wiley & Sons, 2004, vol. 15.

[49] B. Alavi and K. Pahlavan, "Modeling of the toa-based distance measurement error using uwb indoor radio measurements," *Communications Letters, IEEE*, vol. 10, no. 4, pp. 275–277, 2006.

[50] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correal, and R. J. O'dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing,*, vol. 51, no. 8, pp. 2137–2148, 2003.

[51] Y.-T. Chan, H. Y. C. Hang, and P.-c. Ching, "Exact and approximate maximum likelihood localization algorithms," *IEEE Transactions on Vehicular Technology,*, vol. 55, no. 1, pp. 10–16, 2006.

[52] X. Li, "Rss-based location estimation with unknown pathloss model," *IEEE Transactions on Wireless Communications,*, vol. 5, no. 12, pp. 3626–3633, 2006.

[53] J. Zheng and Y.-C. Wu, "Joint time synchronization and localization of an unknown node in wireless sensor networks," *IEEE Transactions on Signal Processing,*, vol. 58, no. 3, pp. 1309–1320, 2010.

[54] G. Wang and K. Yang, "A new approach to sensor node localization using rss measurements in wireless sensor networks," *IEEE Transactions on Wireless Communications,*, vol. 10, no. 5, pp. 1389–1395, 2011.

[55] R. Verdone, F. Fabbri, and C. Buratti, "Area throughput for csma based wireless sensor networks," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*. IEEE, 2008, pp. 1–6.

[56] P. C. Pinto and M. Z. Win, "A unifying framework for local throughput in wireless networks," *arXiv preprint arXiv:1007.2814*, 2010.

[57] A. Host-Madsen, "On the existence of efficient estimators," *IEEE Transactions on Signal Processing,*, vol. 48, no. 11, pp. 3028–3031, 2000.