Relative Location Estimation in Wireless Sensor Networks

Neal Patwari, Alfred O. Hero III,
Matt Perkins, Neiyer S. Correal and Robert J. O'Dea

Neal Patwari and Alfred O. Hero III are with the University of Michigan, Dept. of EECS, Ann Arbor MI, USA. E-mail: [npatwari, hero]@eecs.umich.edu. Matt Perkins, Neiyer S. Correal, and Robert J. O'Dea are with Motorola Labs, Plantation, Florida, USA. E-mail: [M.Perkins, N.Correal, Bob.O'Dea]@Motorola.com

Abstract

This article explores self-configuration in wireless sensor networks. Self-configuration is a general class of estimation problems which we explore via the Cramér-Rao bound (CRB). Specifically, the sensor location estimation problem is explored for sensors that measure range via received signal strength (RSS) or time-of-arrival (TOA) between themselves and neighboring sensors. A small fraction of sensors in the network have known location while the remaining locations must be estimated. We derive CRBs and maximum-likelihood estimators (MLEs) for sensor location estimation under a Gaussian and log-normal model for the TOA and RSS measurements, respectively. The variance bounds are calculated for example networks of devices. Then we report on an extensive TOA and RSS measurement campaign in an indoor office area which shows MLE performance in a real channel. Finally, relative location estimation algorithms are implemented in a wireless sensor network testbed. The wireless sensor devices are deployed in indoor and outdoor environments to demonstrate the accuracy and real-time operation of the location system. Results from the measurements and testbed experiments demonstrate 1 m RMS location errors using TOA, and 1 m to 2 m RMS location errors using RSS.

I. Introduction

In this paper, we consider location estimation in networks in which a small proportion of devices, called reference devices, have *a priori* information about their coordinates. We assume that all devices, regardless of their absolute coordinate knowledge, estimate the range between themselves and their neighboring devices. Such location estimation is termed 'relative location' because the range estimates collected are predominantly between pairs of devices of which neither has absolute coordinate knowledge. These devices without a priori information we call blindfolded devices. Compared to other location estimation systems that have appeared in the literature, eg., cellular location estimation [3][22][20] or local positioning systems (LPS) [26] [25], this estimation problem is complicated by its multi-dimensional nature. Because a device's location is estimated based on ranges to other unknown location devices, the problem is not seperable as it is when each device estimates ranges only to known position devices. However, as will be shown in this article, relative location estimation enables greater accuracy as more devices are added into the network, even when new devices range to just a few close neighbors. Greater accuracy in the network is possible without increasing the burden of installation of more known-location reference devices.

To be implemented, relative location systems require a network with devices capable of peer-to-peer ranging, an ad-hoc networking protocol, and a distributed or centralized location estimation algorithm. For peer-to-peer ranging, TDOA is usually not considered since ad-hoc devices aren't likely to be accurately synchronized. However, TOA ranging has been implemented using two-way or 'round-trip' time-of-arrival measurements [12] [8]. Inquiry-response protocols and careful calibration procedures are presented to allow devices to measure the total delay between an original inquiry and the returned response. Ranging is also possible using RSS measurement, which can be measured from reception of any transmission in the network. In a frequency-hopping radio, RSS measurements can be averaged over frequency to reduce frequency-selective fading error. RSS is attractive from the point of view of device complexity, but is traditionally seen as a coarse measure of range. In this article we will show that by using relative location estimation in dense networks, RSS can lead to accurate location estimates.

The recent literature has seen many distributed [1] [21] [24] and centralized location algorithms [6] [15] designed for locationing in wireless sensor networks. The distributed methods are similar in that a particular device's estimation algorithm depends on the total number of ranges to known-location devices it has measured. If an insufficient number, the solution to its coordinates is underdetermined, and the device makes arbitrary assumptions about its blindfolded neighbors, essentially creating a local coordinate system [24]. If the number of ranges to known-location devices is equal to the dimension of the coordinates, then geometry can readily be applied to calculate a location. If the solution is overdetermined, the device can use a residual weighting algorithm from [3] or least-squared error estimator [21] to reduce the error in the solution. Also, a residual value can be kept with each location estimate to indicate its reliability [1]. After obtaining a location solution, a device becomes a known-location device. Devices iteratively repeat their location estimation protocol as more neighbors estimate and broadcast their location. The local coordinate systems used to solve the underdetermined solutions are merged with other local or global coordinate systems as possible. These distributed algorithms offer the promise of enabling location estimation in networks that have no central processor available. However, they can require many iterations to converge, adding to the communication requirements of the network, especially if sensor movement is possible. Also, the convergence of these iterative methods is not guaranteed [1] [21].

Centralized algorithms assume that the application permits deployment of a central processor

to perform the location estimation [6] [13] [15]. In [6], each range between devices is represented a geometric constraint, and the set of all constraints is solved by convex optimization. Both [13] and [15] provide MLEs for sensor location estimation, when observations are angle-of-arrival [13] and when observations are RSS [15].

This article focuses on the accuracy possible using such relative location estimation algorithms. The radio channel is notorious for its impairments [14] [9], thus accurate sensor location is by no means a given. In wireless sensor networks, "sensing data without knowing the sensor location is meaningless" [18]. The CRBs presented in this article provide a means to determine if the location accuracy necessary for a particular application is possible.

This article begins by considering the category of network self-calibration estimators, which includes as a subset location estimation. Section II explores the CRB for unbiased estimators in this class of problems. We show in particular that the CRB bound decreases as the number of devices in the network increases. Next, we present models of the radio channel for measured RSS and TOA. These models will be verified from measurements in Section VI. The models allow the development of CRBs for location estimation in Section IV. Examples are given for which the inverse of the Fisher information matrix can be solved analytically. The CRB is shown in a special case to match results for traditional location systems such as LPS or cellular location estimation. MLEs are then derived in Section V for relative location estimators. Next, in Section VI measurements of TOA and RSS in a peer-to-peer network are used to test the MLEs. Finally, real-time operation of relative location using RSS is demonstrated in Section VII.

II. Network Estimation Bounds

We start by deriving a CRB for a general class of network self-calibration estimators. In this class of estimation problems, observations made between pairs of devices provide information about the unknown relative parameters of each device, and a subset of devices know their own parameters. An estimator is used to determine the state of all of the device paramenters. For example, distributed clock synchronization in a wireless sensor network could be achieved by devices observing pair-wise timing offsets with just a small number of synchronous devices. As another example, parameters of transducers vary in manufacturing, and accurate characterization of each device is often costly. But with characterization of a fraction of the devices, a network can self-calibrate if it is able to observe statistics of device parameters. Such a procedure could be

used to estimate the received signal strength indicator (RSSI) circuit characteristics of wireless sensors in a network.

In these problems, devices in a network make observations regarding parameters θ that describe the devices in the network. For simplicity, assume that each device has one parameter, that m reference devices have known parameters θ_i, for $i = -m+1 \ldots 0$ and n blindfolded devices do not know their parameters θ_i, for $i = 1 \ldots n$. Devices make pair-wise observations X_{ij},

$$X_{i,j} \sim f_{x_{i,j}|\theta_i,\theta_j}(X_{i,j}|\theta_i,\theta_j) \quad (1)$$

$$l_{i,j} = \log f_{x_{i,j}|\theta_i,\theta_j}(X_{i,j}|\theta_i,\theta_j) \quad (2)$$

We allow for the case when devices make incomplete observations. For example, two devices may be out of range, or a limited channel may not allow the capacity needed for each pair of devices in the network to make observations. Specifically, let $H(i) \subseteq \{-m+1, \ldots, n\}$ be the set of devices with which device i makes pair-wise observations. We assume that $i \notin H(i)$ since a device cannot make a pair-wise observation with itself. By symmetry, if $j \in H(i)$ then $i \in H(j)$. We assume by reciprocity that $X_{i,j} = X_{j,i}$. In practice, if it is possible to make independent observations on the links from i to j, and from j to i, then we assume that a scalar sufficient statistic for the two observations can be found and assigned to $X_{i,j}$ and $X_{j,i}$. Finally, we assume $X_{i,j}$ are independent for $j < i$. Then the log of the joint conditional pdf is

$$l(X|\theta) = \sum_{i=-m+1}^{n} \sum_{\substack{j \in H(i) \\ j<i}} l_{i,j} \quad (3)$$

The Fisher information matrix (FIM) is defined as,

$$\mathbf{F} = -E\left[\nabla_\theta (\nabla_\theta l(X|\theta))^T\right] = \begin{bmatrix} f_{1,1} & \cdots & f_{1,n} \\ \vdots & \ddots & \vdots \\ f_{n,1} & \cdots & f_{n,n} \end{bmatrix} \quad (4)$$

The diagonal and off-diagonal elements of $\mathbf{F}$ are derived in Appendix A and are given by

$$f_{k,l} = \begin{cases} \sum_{j \in H(k)} E\left[\left(\frac{\partial}{\partial \theta_k} l_{kj}\right)^2\right], & k = l \\ \mathrm{I}_{H(k)}(l) E\left[\left(\frac{\partial}{\partial \theta_k} l_{k,l}\right)\left(\frac{\partial}{\partial \theta_l} l_{k,l}\right)\right], & k \neq l \end{cases} \quad (5)$$

Here, $\mathrm{I}_{H(k)}(l)$ is an indicator function which is one when $l \in H(k)$ and zero when it is not. Equivalently,

$$f_{k,l} = \begin{cases} \sum_{j \in H(k)} E\left[\frac{\partial^2}{\partial \theta_k^2} l_{kj}\right], & k = l \\ \mathrm{I}_{H(k)}(l) E\left[\frac{\partial^2}{\partial \theta_k \partial \theta_l} l_{k,l}\right], & k \neq l \end{cases} \quad (6)$$

A. Conditions for a decreasing CRB

Intuitively, we can see that as more devices are used in the location estimator, the accuracy increases for all of the devices in the network. For n devices in the network, there are $O(n)$ parameters to estimate, but there are $O(n^2)$ ranges to use in their estimation. The analysis of this section proves that given certain conditions, the CRB decreases as devices are added to the network. An additional device introduces an additional parameter to be estimated but allows more observations regarding the parameters of interest. Specifically, consider the Fisher information matrices, $\mathbf{F}$ as given in (4), and $\mathbf{G}$ for the network with $n+1$ blindfolded devices.

Theorem 1: Let $[[\mathbf{G}^{-1}]]_{ul}$ be the upper left $n \times n$ block of the inverse of $\mathbf{G}$. The properties:

1) $\mathbf{F}^{-1} - [[\mathbf{G}^{-1}]]_{ul} \geq 0$ in the positive semi-definite sense, and
2) $\text{tr}\, \mathbf{F}^{-1} > \text{tr}\, [[\mathbf{G}^{-1}]]_{ul}$,

both hold if the following conditions are true:

1) $l_{i,j} = l_{j,i}, \forall i, j \in \{1 \ldots n+1\}$,
2) $\frac{\partial}{\partial \theta_{n+1}} l_{k,n+1} = a \frac{\partial}{\partial \theta_k} l_{k,n+1}$ for some constant a not a function of X, $\forall k \in \{1 \ldots n\}$, and
3) Device $n+1$ makes pairwise observations between itself and at least one blindfolded device and at least two devices, in total.

Condition (2) is held for distributions that depend on parameters θ_k and θ_j only as a function of $\theta_k - \theta_j$ or θ_k/θ_j. The Gaussian and log-normal distributions discussed in later sections of this article will be shown to meet this condition. Property (1) implies that the additional $(n+1)$st device does not impair the estimation of the original n parameters. Furthermore, property (2) implies that the sum of the CRB variance bounds for the n parameters strictly decreases when the conditions are met. Thus when a device enters a network and makes pairwise observations with at least one blindfolded device and at least two devices in total, the bound on the average variance of the original n devices is reduced. Note that properties (1) and (2) of Theorem 1 are trivially satisfied by the data processing theorem when the number of parameters is fixed.

The proof of Theorem 1 is shown in Appendix A. Before we can show how this applies specifically to location estimation, we must describe models for RSS and TOA measurements.

III. Radio Channel Models for RSS and TOA

In this section, we develop models for RSS and TOA observations which are used as measurements of range. We show that for the particular measurement system and testbed devices

presented in this article, RSS and TOA measurements can be described as log-normal and Gaussian distributed random variables, respectively.

The multipath channel is modeled in the literature by a sum of attenuated, phase-shifted, and time-delayed impulses [9] [23] [19]. We denote the channel between devices i and j as $h_{i,j}(t)$, thus

$$h_{i,j}(t) = \sum_{l=0}^{K} \alpha_{i,j}(l) e^{j\phi_{i,j}(l)} \delta(t - \tau_{i,j}(l)) \tag{7}$$

where $\alpha_{i,j}(l)$, $\phi_{i,j}(l)$, and $\tau_{i,j}(l)$, are the amplitude, phase, and time delay of the lth multipath component. Here we assume that $l = 0$ indicates the line-of-sight (LOS) component, regardless of how attenuated it might be, and $\tau_{i,j}(0) = cd_{i,j}$ is called the LOS TOA, where c is the speed of light, and $d_{i,j}$ is the straight-line distance between device i and j, given in 2-D by

$$d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \tag{8}$$

The direct-sequence spread-spectrum (DS-SS) channel measurement system used in Section VI is very common in the wideband channel measurement literature [2][5][19]. It uses an unmodulated pseudo-noise (PN) code to measure the impulse response within a wide but finite bandwidth. The transmitter (TX) produces a PN signal at chip rate R_C and center frequency f_c. The receiver samples a downconverted signal at rate R_S and correlates it with a local PN signal. The norm of the result, called the power-delay profile (PDP), is given by $S_{i,j}(k)$, where

$$\begin{aligned} S_{i,j}(k) &= |c_{i,j}(k)|^2 \\ c_{i,j}(k) &= \sum_{l=0}^{K} \alpha_{i,j}(l) e^{j\phi_{i,j}(l)} R_{PN}\left(\frac{k}{R_S} - \tau_{i,j}(l)\right) + n(k). \end{aligned} \tag{9}$$

where $R_{PN}(\tau)$ is the autocorrelation of the PN signal, and $n(k)$ is i.i.d. Gaussian noise. $R_{PN}(\tau)$ is close to zero outside of the region $-1/R_C < \tau < 1/R_C$.

In our measurements, the noise $n(k)$ is not a major factor, since we keep the SNR > 25 dB for all measurements in our campaign. Nevertheless, estimation of $\tau_{i,j}(0)$ from measurement of the PDP is complicated by several factors:

- The finite bandwidth of the measurement system changes multipath impulses in (7) to the $2/R_C$ wide peaks in (9).
- Non-LOS multipath typically do arrive within $2/R_C$ of the LOS TOA, and they jointly with the LOS component determine the shape of the first peak of the PDP.

- If the LOS path is attenuated compared to the early-arriving multipath, it can become indistinquishable.

Due to these early-arriving multipath, we estimate $\tau_{i,j}(0)$ by template-matching [17], in which samples of the leading edge of the PDP are compared to an oversampled template of $R_{PN}(\tau)$. The TOA estimate $\tilde{T}_{i,j}$ is the delay that minimizes the squared-error between the samples of the PDP and the template. Due to the fact that the non-LOS multipath are delayed in time compared to the LOS, $\tilde{T}_{i,j}$ usually has a positive bias. In Section VI-B we will estimate and subtract out the bias to get the unbiased TOA estimator $T_{i,j}$. The resulting estimator is shown experimentally to be a zero mean Gaussian random variable with variance σ_T^2.

Received power, $P_{i,j}$, is measured by two different methods in this article.

1) During the measurement experiment, $P_{i,j}$ is estimated from the PDP of (9). It has been shown that a wideband estimate of received power is obtained by summing the powers of the multipath in the PDP [19].
2) The testbed devices directly measure received power via an RSS sensor in the receiver. The devices are frequency-hopping radios and average RSS across each of 8 center frequencies in a 28 MHz bandwidth, as described in Section VII.

Both methods provide a wideband average RSS that reduce the frequency-selective fading effects. During the measurements and the testbed, time and reciprocal channel averaging were both used. Time averaging takes advantage of fading caused by movement in the channel, while reciprocal channel averaging (averaging the measurement at i from j and the measurement at j from i) helps to reduce device calibration errors.

However, neither method can reduce shadowing (also called medium-scale fading [9]) effects. Shadowing is caused by obstructions in the channel between the TX and RX. Shadowing is random function of placement, since a TX and RX placed the same distance apart in a different area would have a different shadow fading effect. Shadowing loss is often reported to be a log-normal random variable [4][9][19], which leads to the log-normal shadowing model,

$$\begin{aligned} P_{i,j}(\text{dBm}) &= \bar{P}_{i,j}(\text{dBm}) + Z_{i,j}(\text{dB}) \\ \bar{P}_{i,j}(\text{dBm}) &= P_0(\text{dBm}) - 10n\log_{10}\left(\frac{d_{i,j}}{d_0}\right) \end{aligned} \tag{10}$$

where $P_{i,j}$ is the power received at device i transmitted by device j, $\bar{P}_{i,j}(\text{dBm})$ is the mean power in dBm, and $Z_{i,j}(\text{dB})$ is the shadowing gain (loss) which is Gaussian when expressed in

dB. The mean received power is a function of $P_0(\text{dBm})$, the received power in dBm at a reference distance d_0, the path loss exponent n, and the distance $d_{i,j}$. We denote the linear power in mW as $P_{i,j}$ which is calculated as $P_{i,j} = 10^{P_{i,j}(\text{dBm})/10}$. Because we are able to reduce other types of fading by averaging but unable to counter shadowing, we lump all of the measured fading effects into the variable $Z_{i,j}(\text{dB})$, and we show via measurements in Section VI that $Z_{i,j}(\text{dB})$ is still well-modeled as a Gaussian (in dB) random variable.

To simplify analysis, we assume that the fading errors $Z_{i,j}(\text{dB})$ or the TOA errors $T_{i,j} - \tau_{i,j}(0)$ are i.i.d. random variables on different pair-wise links. However, in [16], it was shown that some correlations exist in the RSS case. This motivates channel measurements such as those in Sections VI and VII to verify system performance in real channels.

IV. Cramer-Rao Bound for Co-ordinate Estimation

In this section, we consider the potential performance of sensor network location systems which use RSS or TOA measurements between pairs of devices. Note that neither $P_{i,j}$ nor $T_{i,j}$ are ergodic random variables. That is, their time averages will not be the same as their ensemble averages because obstructions in the measured environment do not significantly change over time. The CRB derived in this section gives a lower bound on the ensemble variance of unbiased estimators of the coordinates of each device. That is, if the same wireless sensor network with the same device coordinates is implemented in many different areas, the variance of each coordinate across implementations will be lower bounded by the CRB. Such a CRB allows researchers to determine if the performance of relative location estimation can possibly meet a particular application's requirements.

Consider a network composed of m reference devices and n blindfolded devices. Here, the unknown parameter vector θ is the vector of all of the unknown coordinates of the devices. In the RSS case, $X_{i,j} = P_{i,j}$, the linear power measured at device i transmitted from device j. We assume that the path loss exponent n, P_0, and d_0 are known constants. Thus the density of $P_{i,j}$ is

$$\begin{aligned} f_{P_{i,j}|\theta}(P_{i,j}|\theta) &= \frac{10}{\sqrt{2\pi\sigma_{dB}^2}\log 10}\frac{1}{P_{i,j}}\exp\left[-\frac{b}{8}\left(\log\frac{d_{i,j}^2}{\tilde{d}_{i,j}^2}\right)^2\right] \\ b &= \left(\frac{10n}{\sigma_{dB}\log 10}\right)^2, \quad \tilde{d}_{i,j} = d_0\left(\frac{P_0}{P_{i,j}}\right)^{\frac{1}{n}}, \end{aligned} \tag{11}$$

where $d_{i,j}$, given in (8), contains the dependence on the coordinates. We have defined the random variable $\tilde{d}_{i,j}$ to help see the physical meaning behind the measured power. It has units of meters and is actually the MLE of range $d_{i,j}$ given received power $P_{i,j}$.

In the TOA case, $X_{i,j} = T_{i,j}$, the time delay measured at device i transmitted from device j. The density of $T_{i,j}$ is

$$f_{T_{i,j}|\theta}(T_{i,j}|\theta) = \frac{1}{\sqrt{2\pi\sigma_T^2}} \exp\left(-\frac{(T_{i,j} - d_{i,j}/c)^2}{2\sigma_T^2}\right) \tag{12}$$

where c is the speed of light, and $d_{i,j}$ is as defined in (8).

A. One-Dimensional TOA Example

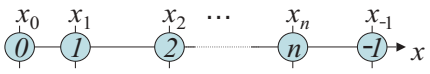

Fig. 1. An example in Section IV-A in which n blindfolded devices are known to be on a linear track. Devices -1 and 0 are two of m reference devices in this network.

As an example, consider location estimation using TOA measurements, when the devices are limited to being located on a 1-D linear track, as shown in Fig. 1. This might have application for locating products on an assembly line. Consider m reference devices and n blindfolded devices, with coordinates $\theta = \{x_1, \ldots, x_n\}$. The distribution of the observations is given by (12) with $d_{i,j} = |x_j - x_i|$. The partial derivatives of the log-likelihood of each measurement are given by

$$\begin{aligned} \frac{\partial}{\partial x_j} l_{i,j} &= \frac{1}{\sigma_T^2 c}(T_{i,j} - d_{i,j}/c) \\ \frac{\partial^2}{\partial x_j^2} l_{i,j} &= -\frac{\partial^2}{\partial x_j \partial x_i} l_{i,j} = \frac{-1}{\sigma_T^2 c^2} \end{aligned}$$

which is a constant w.r.t. the random variable $T_{i,j}$. The FIM for the estimation of θ is then calculated from (6) to be

$$\mathbf{F}_T = \frac{1}{\sigma_T^2 c^2} \begin{bmatrix} n+m-1 & -1 & \cdots & -1 \\ -1 & n+m-1 & & -1 \\ \vdots & & \ddots & \vdots \\ -1 & -1 & \cdots & n+m-1 \end{bmatrix}. \tag{13}$$

For $m \geq 1$, the matrix is invertible, and the inverse is

$$\mathbf{F}_T^{-1} = \sigma_T^2 c^2 \begin{bmatrix} \frac{m+1}{m(n+m)} & \frac{1}{m(n+m)} & \cdots & \frac{1}{m(n+m)} \\ \frac{1}{m(n+m)} & \frac{m+1}{m(n+m)} & & \frac{1}{m(n+m)} \\ \vdots & & \ddots & \vdots \\ \frac{1}{m(n+m)} & \frac{1}{m(n+m)} & \cdots & \frac{m+1}{m(n+m)} \end{bmatrix}. \tag{14}$$

Thus the the variance of an unbiased estimator for x_i must have variance

$$\sigma_{x_i}^2 \geq \sigma_T^2 c^2 \frac{m+1}{m(n+m)}. \tag{15}$$

As we would expect, the variance $\sigma_{x_i}^2$ is reduced more quickly by adding reference devices than by adding blindfolded devices. However, for large m, the difference between increasing m and n is negligible.

B. Two-dimensional Location Estimation

Although one-dimensional examples are interesting because they can lead to analytical expressions, in this article, we are mainly interested in two-dimensional location estimation. The 2-D unknown parameter vector is

$$\theta = \{x_1, \ldots, x_n, y_1, \ldots, y_n\}. \tag{16}$$

Consider the Fisher information matrices for the RSS and TOA cases, $\mathbf{F}_R$ and $\mathbf{F}_T$, respectively. Each device has two parameters, but the FIM in (4) assumes that each device has one parameter. From the definition of the FIM in (4) we can see however that the FIM for the 2-D case will have a similar form if partitioned into blocks,

$$\mathbf{F}_R = \begin{bmatrix} \mathbf{F}_{Rxx} & \mathbf{F}_{Rxy} \\ \mathbf{F}_{Rxy}^T & \mathbf{F}_{Ryy} \end{bmatrix}, \quad \mathbf{F}_T = \begin{bmatrix} \mathbf{F}_{Txx} & \mathbf{F}_{Txy} \\ \mathbf{F}_{Txy}^T & \mathbf{F}_{Tyy} \end{bmatrix} \tag{17}$$

where the blocks $\mathbf{F}_{Rxx}$ and $\mathbf{F}_{Txx}$ are given by (5) using only the x parameter vector $\theta' = \{x_1, \ldots, x_n\}$, and the blocks $\mathbf{F}_{Ryy}$ and $\mathbf{F}_{Tyy}$ are given by (4) using only the y parameter vector $\theta'' = \{y_1, \ldots, y_n\}$. The off-diagonal blocks $\mathbf{F}_{Rxy}$ and $\mathbf{F}_{Txy}$ have elements defined by,

$$f_{k,l} = \begin{cases} \sum_{j \in H(k)} E\left[\left(\frac{\partial}{\partial x_k} l_{k,j}\right)\left(\frac{\partial}{\partial y_k} l_{k,j}\right)\right], & k = l \\ \mathrm{I}_{H(k)}(l) E\left[\left(\frac{\partial}{\partial x_k} l_{k,l}\right)\left(\frac{\partial}{\partial y_l} l_{k,l}\right)\right], & k \neq l \end{cases} \tag{18}$$

The elements of the sub-matrices are derived in Appendix C from the distribution of the measurements given either in (11) or (12). For the RSS case, the elements are given by

$$
\begin{aligned}
((F_{Rxx}))_{k,l} &= \begin{cases} b\sum_{i\in H(k)} \frac{(x_k-x_i)^2}{[(x_k-x_i)^2+(y_k-y_i)^2]^2} & k=l \\ -b\,\mathrm{I}_{H(k)}(l)\frac{(x_k-x_l)^2}{[(x_k-x_l)^2+(y_k-y_l)^2]^2} & k\neq l \end{cases} \\
((F_{Rxy}))_{k,l} &= \begin{cases} b\sum_{i\in H(k)} \frac{(x_k-x_i)(y_k-y_i)}{[(x_k-x_i)^2+(y_k-y_i)^2]^2} & k=l \\ -b\,\mathrm{I}_{H(k)}(l)\frac{(x_k-x_l)(y_k-y_l)}{[(x_k-x_l)^2+(y_k-y_l)^2]^2} & k\neq l \end{cases} \\
((F_{Ryy}))_{k,l} &= \begin{cases} b\sum_{i\in H(k)} \frac{(y_k-y_i)^2}{[(x_k-x_i)^2+(y_k-y_i)^2]^2} & k=l \\ -b\,\mathrm{I}_{H(k)}(l)\frac{(y_k-y_l)^2}{[(x_k-x_l)^2+(y_k-y_l)^2]^2} & k\neq l \end{cases}
\end{aligned} \tag{19}
$$

Compared to elements of $\mathbf{F}_R$, which have d^4 terms in the denominator, elements of $\mathbf{F}_T$ have d^2 terms in the denominator. This is shown explicitly in the results for the elements of the sub-matrices,

$$
\begin{aligned}
((F_{Txx}))_{k,l} &= \begin{cases} \frac{1}{c^2\sigma_T^2}\sum_{i\in H(k)} \frac{(x_k-x_i)^2}{(x_k-x_i)^2+(y_k-y_i)^2} & k=l \\ -\frac{1}{c^2\sigma_T^2}\mathrm{I}_{H(k)}(l)\frac{(x_k-x_l)^2}{(x_k-x_l)^2+(y_k-y_l)^2} & k\neq l \end{cases} \\
((F_{Txy}))_{k,l} &= \begin{cases} \frac{1}{c^2\sigma_T^2}\sum_{i\in H(k)} \frac{(x_k-x_i)(y_k-y_i)}{(x_k-x_i)^2+(y_k-y_i)^2} & k=l \\ -\frac{1}{c^2\sigma_T^2}\mathrm{I}_{H(k)}(l)\frac{(x_k-x_l)(y_k-y_l)}{(x_k-x_l)^2+(y_k-y_l)^2} & k\neq l \end{cases} \\
((F_{Tyy}))_{k,l} &= \begin{cases} \frac{1}{c^2\sigma_T^2}\sum_{i\in H(k)} \frac{(y_k-y_i)^2}{(x_k-x_i)^2+(y_k-y_i)^2} & k=l \\ -\frac{1}{c^2\sigma_T^2}\mathrm{I}_{H(k)}(l)\frac{(y_k-y_l)^2}{(x_k-x_l)^2+(y_k-y_l)^2} & k\neq l \end{cases}
\end{aligned} \tag{20}
$$

The CRBs for the RSS and TOA cases are $\mathbf{F}_R^{-1}$ and $\mathbf{F}_T^{-1}$, respectively. If we define $\hat{x}_i$ and $\hat{y}_i$ as estimators of the two unknowns for device i, x_i and y_i, then we can write the bound on the variance of the location estimate for device i as

$$\sigma_i^2 \leq \mathrm{Var}(\hat{x}_1) + \mathrm{Var}(\hat{y}_1). \tag{21}$$

Although we do not have an analytical form for the matrix inverse in general for n blindfolded devices, we can for $n = 1$ show and verify the result.

C. Traditional Location System Example

Consider the simple case when device 1 is a blindfolded device and devices $-m+1 \ldots 0$ are reference devices. Assume that device 1 makes pairwise measurements with all m reference devices. This example is the same as non-relative location systems (eg. cellular or LPS) that

have been studied in the literature, and a bound for the variance of the location estimator has been derived in the TOA case [22]. In the RSS case,

$$E\left[(\hat{x}_1 - x_1)^2 + (\hat{y}_1 - y_1)^2\right] \geq \sigma_1^2 = \frac{F_{Rxx} + F_{Ryy}}{F_{Rxx}F_{Ryy} - F_{Rxy}^2} \tag{22}$$

from which we get the result that

$$\sigma_1^2 = \frac{1}{b} \frac{\sum_{i=-m+1}^{0} d_{1,i}^{-2}}{\sum_{i=-m+1}^{-1} \sum_{j=i+1}^{0} \left(\frac{d_{1\perp i,j} d_{i,j}}{d_{1,i}^2 d_{1,j}^2} \right)^2} \tag{23}$$

where the distance $d_{1\perp i,j}$ is the shortest distance from the point (x_1, y_1) to the line between device i and device j. For the TOA case, the result is,

$$\sigma_1^2 = c^2 \sigma_T^2 m \left[\sum_{i=-m+1}^{-1} \sum_{j=i+1}^{0} \left(\frac{d_{1\perp i,j} d_{i,j}}{d_{1,i} d_{1,j}} \right)^2 \right]^{-1} \tag{24}$$

The ratio $d_{1\perp i,j} d_{i,j} / (d_{1,i} d_{1,j})$ has been called the geometric conditioning $\mathcal{A}_{i,j}$ of device 1 with respect to references i and j [22]. $\mathcal{A}_{i,j}$ is the area of the parallelogram formed by the vectors from device 1 to reference i and from device 1 to reference j, normalized by the lengths of the two vectors. Thus the geometric dilution of precision (GDOP), defined as σ_1/σ_T, can be written as

$$GDOP = \sqrt{\frac{m}{\sum_{i=-m+1}^{-1} \sum_{j=i+1}^{0} \mathcal{A}_{i,j}^2}} \tag{25}$$

which matches the result in [22] and verifies our analysis for the TOA case.

Note that if the distance between reference devices is increased without changing the geometric conditioning, and we assume σ_T^2 is constant with range, then the TOA bound remains constant. In the RSS case, however, the CRB scales proportionally with distance between reference devices when the geometry is the same and σ_{dB}^2 is constant with range. This difference between TOA and RSS is also true for $n > 1$. As a result, TOA is a good measurement method even for large, sparse networks. However, if a network is scaled down to make a small, dense network of devices, at some density, RSS can perform as well as TOA.

Contour plots of the CRB for both the RSS and the TOA cases are shown in Fig. 2. The minimum value in Fig. 2(a) is 0.27. Since the CRB scales with size in the RSS case, the standard deviation of location estimates in a traditional RSS system operating in a channel with $\sigma_{dB}/n = 1.7$ is limited to about 27% of the distance between reference devices. This performance has prevented use of RSS in many existing location systems, but Section VII will demonstrate

reasonable error performance for relative location systems using RSS. The TOA CRB result in Fig. 2(b) does not scale with distance between reference devices. However, it does scale with $c\sigma_T$, and $c\sigma_T = 1$ was chosen for ease of calculation.

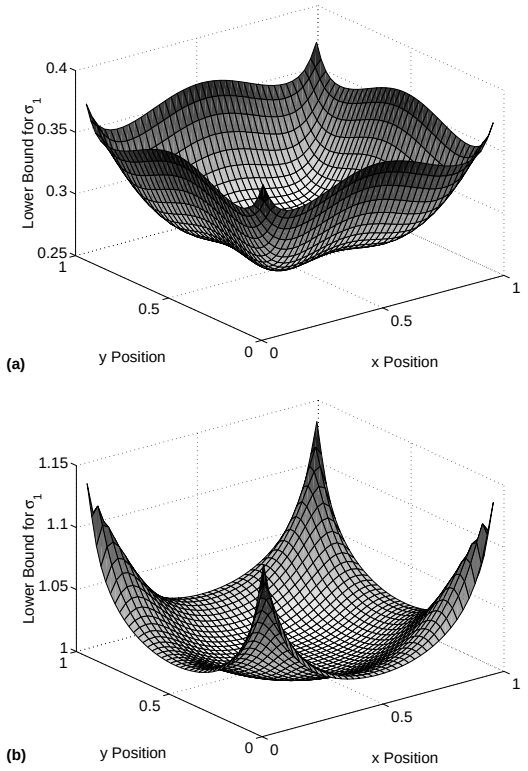

Fig. 2. σ_1 in meters with four reference devices in the corners of a 1m by 1m area plotted as a function of the coordinates of the single blindfolded device, with (a) RSS with log-normal errors with $\sigma_{dB}/n = 1.7$, or (b) TOA with Gaussian errors and $c\sigma_T = 1$.

V. Maximum Likelihood Relative Location Estimation

We now present maximum likelihood estimators for the relative location estimation system. For the TOA case, the MLE of θ is given by

$$\hat{\theta}_T = \arg\min \sum_{i=-m+2}^{n} \sum_{\substack{j \in H(i) \\ j<i}} \left(cT_{i,j} - d_{i,j}\right)^2 \tag{26}$$

The maximum likelihood estimator for the two-dimensional RSS case is shown in [15] to be,

$$\hat{\theta}_{\tilde{R}} = \arg\min \sum_{i=-m+2}^{n} \sum_{j \in H(i)} \left(\ln \frac{\tilde{d}_{i,j}^2}{d_{i,j}^2}\right)^2 \tag{27}$$

Unlike in the TOA case, the RSS MLE is readily shown to be biased. Consider that for $n = 1$ and $m = 1$, the range between the two devices will be estimated to be equal to $\tilde{d}_{i,j}$. Using the pdf in (11), the mean of $\tilde{d}_{i,j}$ is given by

$$E[\tilde{d}_{i,j}] = C\, d_{i,j}, \quad \text{where } C = \exp\left[\frac{1}{2}\left(\frac{ln(10)}{10}\frac{\sigma_{dB}}{n}\right)^2\right]. \tag{28}$$

A bias-reduced MLE,

$$\hat{\theta}_R = \arg\min \sum_{i=-m+2}^{n} \sum_{j \in H(i)} \left(\ln \frac{\tilde{d}_{i,j}^2}{C^2 d_{i,j}^2}\right)^2 \tag{29}$$

is preferred. The bias reduction has been shown to improve estimator performance in simulations and in the testbed experiments presented in Section VII.

However, even with the reduction there is bias in the coordinate estimates. Consider $m = 4$ and $n = 1$. Place the reference devices at the corners of a 1 m by 1 m square and the blindfolded device within the square. This is the case for which we calculated the CRB in Fig. 2. We calculate via simulation the bias norm of $\hat{x}_1$ and display it in Fig. 3. The bias is high near the edges of the square area.

The gradient of the bias can be used to calculate the achievable variance of the biased estimator [10]. Fig. 3 shows that the bias is changing sharply at the corners of the square, and in fact, the simulated bias gradient norm is almost equal to 1 at the corners. Intuitively, this matches the multiplicative nature of the MLE in (29), which has a cost proportional to $[\ln(\tilde{d}_{i,j}^2/C^2/d_{i,j}^2)]^2$. For example, the cost of positioning the devices such that $d_{i,j}^2 = 2\tilde{d}_{i,j}^2$ is the same whether $\tilde{d}_{i,j}$ is very small or very large. But if $\tilde{d}_{i,j}$ is very small, then the estimator has very little freedom (in terms of area) to position device i with respect to device j. In the limit, as blindfolded device 1 approaches reference device i, $\tilde{d}_{1,i} \to 0$, and the MLE will locate device 1 on top of device j. In this case, we expect the variance of the MLE to approach zero, even though the CRB for an unbiased estimator is greater than zero (as shown in Fig. 2(a)). This is the effect of the high (close to 1) bias gradient norm at the corners of Fig. 3.

Note that the bias of the MLE is zero near the center of the square. As a rule of thumb for $n > 1$, the MLE generally has low bias for blindfolded devices near the center of their neighbors.

VI. Channel Measurement Experiment

Multipoint-to-multipoint (M2M) wideband channel measurements are conducted at the Motorola facility in Plantation, Florida in order to provide an example of the errors that the MLEs

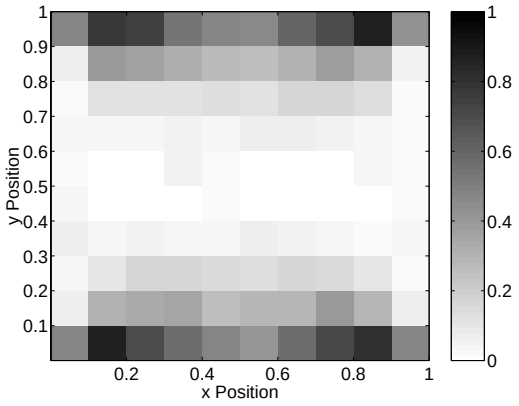

Fig. 3. Simulated bias gradient norm of $\hat{x}_1$ calculated from (29) for the RSS case with $n = 1$ and $m = 4$ with reference devices placed in the corner of a 1m by 1m square area.

Fig. 4. Photo of measurement area looking above cubicle walls.

presented in the previous section would see in an indoor office environment. The measurement area is a 14m by 13m area partitioned by 1.8m high cubicle walls, with hard partitioned offices, external glass windows and cement walls on the outside of the area. There are also metal and concrete support beams within and outside of the area. Offices are occupied with desks, bookcases, metal and wooden filing cabinets, computers, and test & measurement equipment. Forty-four device locations are identified and marked with tape.

The measurement system uses a wideband direct-sequence spread-spectrum (DS-SS) transmitter (TX) and receiver (RX) (Sigtek model ST-515). They are operated synchronously using two Datum ExacTime GPS and rubidium-based oscillators. The transmitter and receiver equipment are battery-powered and are placed on carts. The TX outputs an unmodulated pseudo-noise (PN) code signal with a 40 MHz chip rate and code length 1024. The center frequency f_c is 2443

MHz, and the transmit power P_t is 10 mW. The RX records I and Q samples at a rate of 120 MHz (for a total of 240 MSamples / sec) and forwards the samples into a PC for processing. The TX and RX are triggered by a 1 pulse per second signal from the Datum oscillators. Both use 2.4 GHz sleeve dipole antennas kept at a height of 1m above the floor. The antennas have an omnidirectional pattern in the horizontal plane and a measured antenna gain of 1.1 dBi.

A. Measurement Procedure

The Datum oscillators at the TX and RX are carefully synchronized throughout each measurement day. After an initial GPS synch, GPS is disconnected and the rubidium oscillators provide very stable 1 pps signals. The frequencies of the two rubidium oscillators are off very slightly, thus the 1PPS signals drift linearly with time. The actual rate of change is on the order of nanoseconds per hour and is accurately determined with periodic time calibration measurements. The result of the linear behavior of the rubidium oscillators and the time calibration measurements is that, given the time-of-day of any particular RX measurement, the effect of the oscillator drift can be cancelled. A time base with accuracy of approximately 1-2 ns is achieved.

For the M2M measurements, the channel between each pair of device locations is measured. First, the TX is placed at location 1 while the RX is moved and measurements are made at locations 2 through 44. Then the TX is moved to location 2, as the RX is moved to locations 1 and 3 through 44. At each combination of TX and RX locations, the RX records five wideband channel measurements for averaging purposes. A total of 44*43*5 = 9460 wideband channels are measured. Both TOA and RSS are estimated as described in Section III. In post-processing, the calibration corrections are made. Then, averaging is done using the five measurements taken with the RX at i and the TX at j. Due to reciprocity, these are averaged with the five measured with the RX at j and the TX at i. The TOA measurements $T_{i,j}$ are thus the arithmetic mean of the 10 measured TOAs. Due to the log-normal distribution of the RSS measurements, $P_{i,j}$ and $P_{j,i}$ are set to the geometric mean of the 10 wideband received power measurements.

B. Measurement Results

First, the assumption about the log-normal distribution of the RSS error is verified. As shown in Fig. 5, the RSS measurements matched the model in (10) with $n = 2.30$, and $\sigma_{dB} = 3.92$ dB. The reference distance d_0 was assumed to be 1m. The low path loss exponent is attributed to the fact that many measured channels were obstructed only by cubicle partitions, which do

not cause much shadowing loss. In addition, reflections from the ceiling see few obstructions since cubicle walls are only 1.8m high. The low variance seems to predominantly be a result of the homogeneity and small size of the measured cubicle area. The results are similar to some reported in [19]. The quantile-quantile plot in Fig. 6(a) compares the empirical distribution of $Z_{i,j}$ (dB) with a Gaussian distribution. It shows a close fit between the -2 and +2 quantiles, although at the tails it is heavier than the Gaussian distribution.

The RSS measurements are input to the MLE in (29) and the estimated device locations are compared to the actual locations in Fig. 8(a). The RMS location error over all 40 blindfolded devices is 2.30m. Notice that the devices close to the center are located more accurately than the devices on the edges. Devices 11 and 12 have especially high error because they are estimated to be far away from the deployment area. This is expected, due to devices at the edges having fewer neighbors to aid in their location estimate. To show that the devices close to the center of the network have the best location estimates, we plot in Fig. 7(a) the average location error vs. distance from a device to its nearest reference device. The devices in the center will have the furthest distance to the nearest reference device. Thus Fig. 7 shows that the error decreases as distance from the nearest reference device increases.

The empirical distribution of TOA estimation error, $\hat{T}_{i,j} - T_{i,j}$ (ns), is also compared to the Gaussian distribution and found to be heavy-tailed, as shown in Fig. 6(b). The estimation error has a mean of 10.9 ns and a $\sigma_T = 6.1$ ns. The mean is subtracted out, and the resulting unbiased estimated TOA is input to the MLE in (26). The location estimates are shown in Fig. 8(b). The performance using TOA ranging is better than with RSS ranging - the RMS location error in this case is 1.15m.

To verify the CRB, we would need several measurement campaigns in different office areas but with the same device geometry, as discussed in Section IV. Due to resources and time, such a verification could not be conducted. Nevertheless, it is interesting to report the CRB for the measured network. For the RSS and TOA cases, the square root of the trace of the CRBs, $\left(\sum_{i=1}^{40} \sigma_i^2\right)^{1/2}$, are 0.76m and 0.69m, respectively.

VII. Testbed Experimentation

For further testing of relative location estimation and to provide an easy means for M2M radio channel data collection, we developed and fabricated at Motorola Labs a testbed of 12 prototype wireless sensor devices. These devices operate in a peer-to-peer mode and measure received

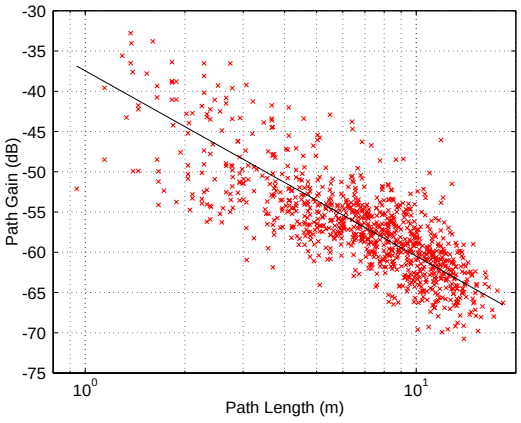

Fig. 5. Measured wideband path gain (x) as a function of path length. Linear fit (—) is with $d_0 = 1$m, $n = 2.3$, and $\sigma_{dB} = 3.92$.

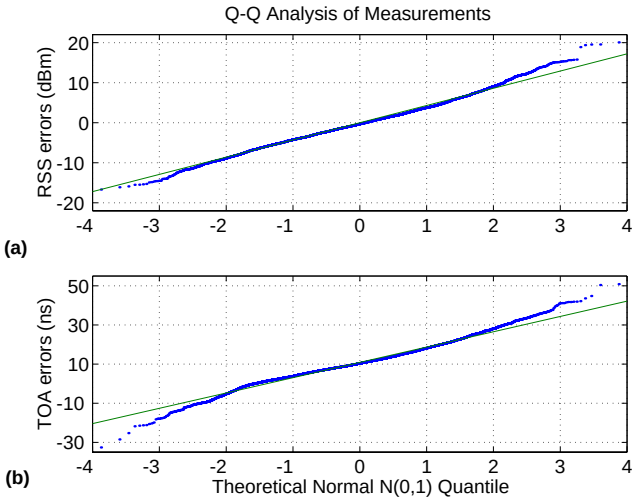

Fig. 6. Q-Q plot of sorted (a) measured dB shadowing errors $Z_{i,j}$ (dB), and (b) measured TOA estimation errors $\hat{T}_{i,j} - T_{i,j}$ (ns), compared to a standard Gaussian quantile.

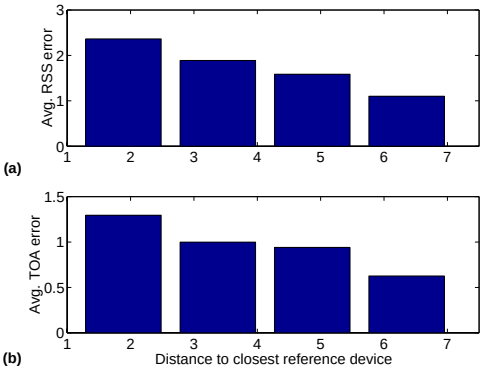

Fig. 7. Mean location estimation error vs. distance from the closest reference device for (a) RSS and (b) TOA measurements. Data from 40 blindfolded devices are grouped into 4 bins.

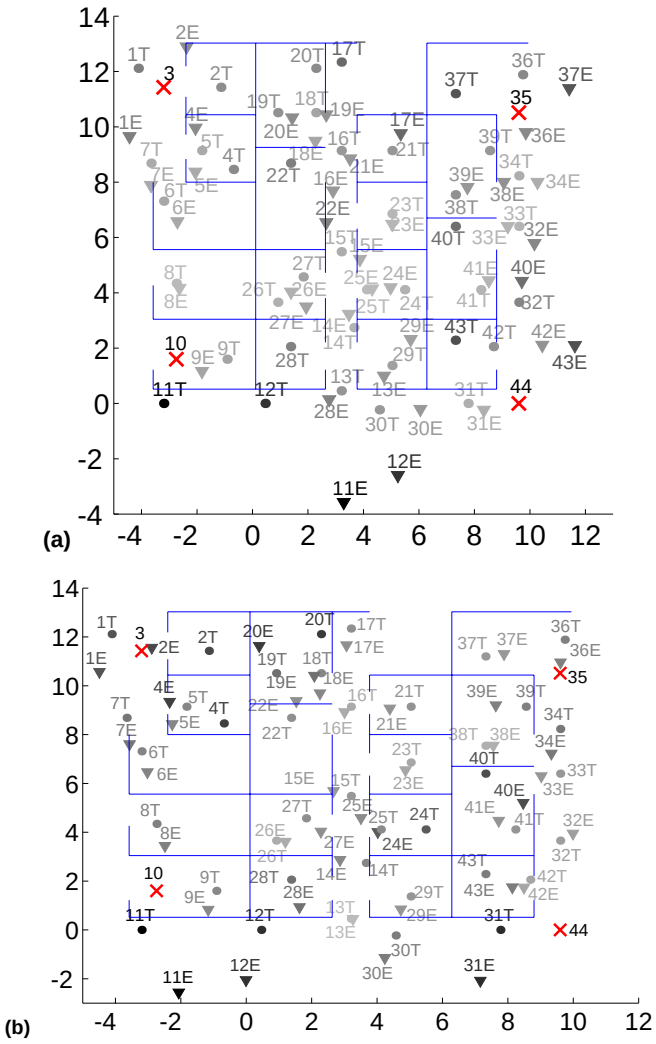

Fig. 8. True (•#T) and estimated (▽#E) location using (a) RSS and (b) TOA measurements for 40 blindfolded devices. High and low errors are indicated by dark and light gray text, respectively. Four reference devices (X) are near the corners of the area.

power as described in Section III. Each device's TX power level and RSS sensor characteristic was measured before deployment of the testbed. The devices operate in the 900-928 MHz band (allocated for ISM use in the US) and can be set to one of eight center frequencies within the band. The separation between the 8 frequencies is approximately equal to the coherence bandwidth of the channel (~ 4 MHz). The devices are narrowband using FSK with a 50 kHz data rate. While one device transmits a packet, its neighbors measure the RSS. Packet transmissions are infrequent and packets are short, thus the channel is almost always silent. Devices are asynchronous and use a carrier-sense multiple access (CSMA) protocol to prevent collisions.

Thus receiver measurements are not affected by multi-user interference in this testbed. Every two seconds, each device creates a packet of measured RSS data and transmits it to a central 'listening' device. The listening device uploads the data to a processing algorithm running on a laptop computer. The algorithm has access to the measured transmit power and receiver characteristic of each device in the network and thus can determine the path loss values accurately. The algorithm stores the ranges for each pair of devices, each frequency that was measured, and each measurement over time.

The algorithm first averages the measurements. Time averaging is set to average the most recent four RSS measurements. Frequency averaging across the 8 center frequencies reduces the small-scale (frequency selective) fading effects. Reciprocal channel averaging helps to reduce any device calibration errors that might exist. When the locations of the reference devices are input into the computer, the algorithm uses the averaged RSS measurements as input to the bias-reduced MLE of (29). The maximum of the likelihood equation is found using a conjugate gradient algorithm, which takes less than one second on the Pentium laptop. More than 60 pair-wise ranges are received by the laptop each second, and an updated location estimate is calculated each second. The location estimates for the blindfolded devices are displayed in real time on a map using a Visual Basic GUI. Thus real time tracking of slow movement (eg., walking) is possible.

Note that an estimate of the path loss exponent must be entered into the algorithm. However, the testbed of wireless devices can also serve as a quick and accurate channel characterization system. When all of the device locations are known and entered into the computer, the algorithm uses the path loss vs. path length data to estimate the path loss exponent, n_{pl} [7]. With 12 devices doing pair-wise path loss measurements across 8 frequencies, the algorithm uses 880 measurements to characterize the channel in the area of deployment. During the testbed experiments, we first use the testbed to estimate the path loss exponent, then remove the coordinates of the blindfolded devices from the computer and operate the relative location estimation algorithm using the estimated n_{pl}. In the next sections, the results of two such experiments are presented.

A. Parking Lot Area

Testbed devices are placed in a 9 m by 9 m area in a 3 m grid in an empty area of the parking lot at the Motorola facility in Plantation, Florida, as seen in Fig. 9. The devices are kept at a height of 0.35 m above the ground, to keep the devices both above the ground plane and visible

Fig. 9. Parking lot testbed experiment. Devices are located on top of 16 upside-down recycling bins.

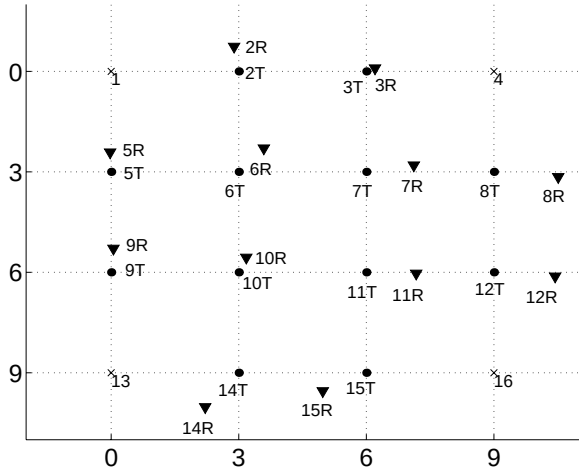

Fig. 10. Actual blindfolded device locations (T) and relative location estimates (R) for the parking lot test, showing a 1.02 meter RMS location error. Reference device locations are indicated with an 'x'.

to drivers. One side of the test area borders a wooded area. The path loss exponent is estimated to be 3.2. Four reference devices are placed at the corners of the area, and 7 blindfolded devices are placed in the grid. Devices record pair-wise ranges and transmit them back to a laptop computer, which estimates their location coordinates. The blindfolded devices are then moved to different positions in the grid for a new trial. 16 trials are run. The RMS location errors for the individual trials range from 0.9 m to 2.4 m. Overall, the RMS location error is 1.70 m. However, by moving 7 blindfolded devices around between positions, we record enough spot-to-spot ranges to see what would happen if there were 12 blindfolded devices, one in each spot on the grid. In an off-line calculation, we use the recorded range data to calculate what would have happened

in this case, and we find that the RMS location error would have been 1.46 m. Similarly, if we extended the duration of the time averaging from 4 ranges to 32 ranges, we would see the location estimates shown in Fig. 10, and we would reduce the RMS location error to 1.02 m. We can see that in this environment, where we expect that shadow fading is not severe, time averaging is effective at improving location estimates.

B. Residential Home

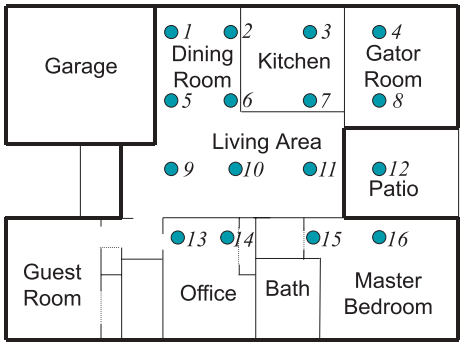

Fig. 11. Map of the Perkins home which shows how the grid of devices span several rooms.

Fig. 12. Residential home testbed experiment. Three devices are shown (one is hidden between the couch and the table). Neiyer supervises the operation of the location algorithm on the laptop.

The performance of the system is tested in the Perkins home, a single-family, ranch-style house in Sunrise, Florida (Fig. 11). An identical 9 m by 9 m grid is used in this test, however, the area spanned across several rooms: office, bedroom, kitchen, dining area, living room, outdoor patio, and Gator room. The obstructions, some shown in Fig. 12, include the indoor walls, furnishings, and appliances, and exterior walls and windows. Here, there are four devices used

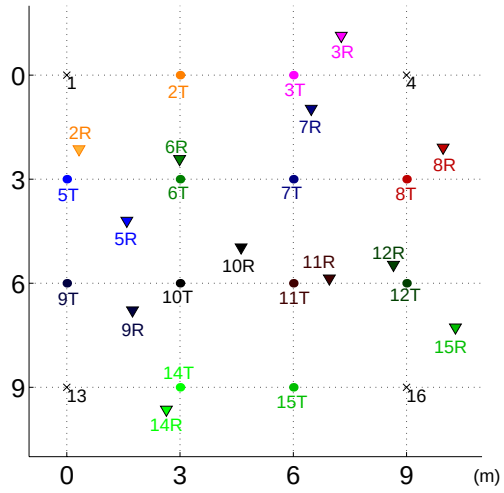

Fig. 13. Actual blindfolded device locations (T) and relative location estimates (R) for the in-house location test. Reference device locations are indicated with an 'x'.

as reference devices in the corners of the grid and 8 other devices used as blindfolded devices. In this environment, the estimated path loss exponent is 4.0. In the 16 individual trials, the RMS location errors range from 1.0 m to 2.7 m, and the overall RMS error is 1.9 m. If all device ranges are used together, as described in the previous section, we see the results shown in Fig. 13, in which the RMS error is 2.1 m. This error doesn't reduce significantly when the duration of time-averaging is increased from 4 to 32 ranges. The error is predominantly due to device #15, which had an error of 4.5 m. As shown in Fig. 11, device #15 is actually 3 m away from device #14. There is significant shadowing caused by three interior walls that form the office closet and master bedroom closet that lie in between two devices (fading of $Z_{14,15} = -22$ dB in (10)), and as a result the range estimate between the two is found to be 10.5 m. Unfortunately, this shadowing can't be countered by time or frequency averaging. To its credit, even though a 22 dB channel modeling error existed in the network, the location error of device #15 was limited to 4.5 m.

VIII. Conclusions

The motivation of this article has been to show with what accuracy wireless sensor networks can estimate sensor locations. The results have been promising. First, location estimation variances in ad hoc networks are shown to decrease as more devices are added to the network. Next, CRBs can be derived for simple examples, and readily calculated for arbitrary numbers and geometries of devices. MLEs have been presented and used in several real channels, both

for TOA and RSS measurements. Sensor location estimation with about 1 m RMS error has been demonstrated using TOA measurements. However, despite the reputation of RSS as a coarse means to estimate range, it is also able to achieve an accuracy of about 1 m RMS in a testbed experiment. Fading outliers can still impair the RSS relative location system, implying the need for a robust estimator. The results presented in this article should help wireless sensor networking researchers determine if the accuracy possible in relative location estimation can meet the requirements of their application.

IX. Acknowledgments

We would like to acknowledge the contribution of Miguel Roberts, who assisted with the measurement system.

Appendix

A. CRB for Network Self-Calibration

The diagonal elements, $f_{k,k}$, of $\mathbf{F}$ given in (4) are given by

$$\begin{aligned} f_{k,k} &= E\left[\left(\frac{\partial}{\partial\theta_k} l(X|\theta)\right)^2\right] = E\left[\left(\sum_{j\in H(k)} \frac{\partial}{\partial\theta_k} l_{k,j}\right)^2\right] \\ f_{k,k} &= \sum_{j\in H(k)} \sum_{p\in H(k)} E\left[\left(\frac{\partial}{\partial\theta_k} l_{k,j}\right)\left(\frac{\partial}{\partial\theta_k} l_{k,p}\right)\right] \end{aligned}$$

Since $X_{k,j}$ and $X_{k,p}$ are independent random variables, and

$$E\left[\frac{\partial}{\partial\theta_k} l_{k,j}\right] = 0, \tag{30}$$

the expectation of the product is only nonzero for $p = j$. Thus $f_{k,k}$ simplifies to the $k = l$ result in (5). The off-diagonal elements similarly simplify,

$$\begin{aligned} f_{k,l} &= E\left[\left(\sum_{j\in H(k)} \frac{\partial}{\partial\theta_k} l_{k,j}\right)\left(\sum_{p\in H(l)} \frac{\partial}{\partial\theta_l} l_{l,p}\right)\right] \\ f_{k,l} &= \sum_{j\in H(k)} \sum_{p\in H(l)} E\left[\left(\frac{\partial}{\partial\theta_k} l_{k,j}\right)\left(\frac{\partial}{\partial\theta_l} l_{l,p}\right)\right] \end{aligned} \tag{31}$$

Here, due to independence and zero mean of the two terms, the expectation of the product will be zero unless both $p = k$ and $j = l$. Thus the $k \neq l$ result in (5).

B. Proof of Theorem 1

To prove Theorem 1, partition $\mathbf{G}$ into blocks:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{ul} & \mathbf{g}_{ur} \\ \mathbf{g}_{ll} & g_{lr} \end{bmatrix} \tag{32}$$

Here, $\mathbf{G}_{ul}$ is an $n \times n$ matrix, $\mathbf{g}_{ur} = \mathbf{g}_{ll}^T$ are $n \times 1$ vectors with kth element,

$$\mathbf{g}_{ur}(k) = \mathrm{I}_{H(n+1)}(k) E\left[\left(\frac{\partial}{\partial\theta_k} l_{k,n+1}\right)\left(\frac{\partial}{\partial\theta_{n+1}} l_{k,n+1}\right)\right] \tag{33}$$

where $H(n+1) \subseteq \{-m+1, \ldots, n\}$ is set of devices with which device $n+1$ makes pair-wise observations. g_{lr} is the scalar Fisher information for the $n+1$st parameter,

$$g_{lr} = \sum_{j \in H(n+1)} E\left[\left(\frac{\partial}{\partial\theta_{n+1}} l_{n+1,j}\right)^2\right]. \tag{34}$$

Let $l_n(X|\theta_n)$ be the joint log-likelihood function as defined in (3) for the n parameter case and $l_{n+1}(X|\theta_{n+1})$ be the joint log-likelihood function for the $n+1$ parameter case. For the $n+1$ parameter case,

$$\begin{aligned} l_{n+1}(X|\theta_{n+1}) &= \sum_{i=-m+1}^{n+1} \sum_{\substack{j \in H(i) \\ j<i}} l_{i,j} \\ &= \sum_{i=-m+1}^{n} \sum_{\substack{j \in H(i) \\ j<i}} l_{i,j} + \sum_{j \in H(n+1)} l_{n+1,j} \\ &= l_n(X|\theta_n) + \sum_{j \in H(n+1)} l_{n+1,j}. \end{aligned} \tag{35}$$

Looking at the 2nd partial derivatives of $\sum_{j \in H(n+1)} l_{n+1,j}$ w.r.t. θ_k and θ_l, where $k, l \leq n$, since $l_{n+1,j}$ is a function only of parameters θ_{n+1} and θ_j,

$$\frac{\partial^2}{\partial\theta_k \partial\theta_l} \sum_{j \in H(n+1)} l_{n+1,j} = \begin{cases} \mathrm{I}_{H(n+1)}(k) \frac{\partial^2}{\partial\theta_k^2} l_{n+1,k}, & l = k \\ 0, & l \neq k \end{cases} \tag{36}$$

Eqs. (35) and (36) show that the $\mathbf{G}_{ul}$ matrix is different from to the $\mathbf{F}$ matrix from (4) only in the diagonal elements. The off-diagonal elements in $\mathbf{G}_{ul}$ equal their counterparts in $\mathbf{F}$, specifically, $[[\mathbf{G}_{ul}]]_{i,j} = f_{i,j} \forall i \neq j$. This is clear since (36) is zero for $l \neq k$. The diagonal elements of $\mathbf{G}_{ul}$ each have one additional term compared to their counterparts in $\mathbf{F}$, that is, $[[\mathbf{G}_{ul}]]_{k,k} = f_{k,k} + h_k$ where h_k is defined as,

$$h_k = \mathrm{I}_{H(n+1)}(k) E\left[\left(\frac{\partial}{\partial\theta_k} l_{n+1,k}\right)^2\right] \tag{37}$$

Then, defining $\mathbf{h} = \{h_1, \ldots, h_n\}$, we have that $\mathbf{G}_{ul} = \mathbf{F} + \mathrm{diag}(\mathbf{h})$. Now, compare the CRB for the covariance matrix of the first n devices in the n and $n+1$ blindfolded device cases, given by $\mathbf{F}^{-1}$ and $[[\mathbf{G}^{-1}]]_{ul}$, respectively. Here, $[[\mathbf{G}^{-1}]]_{ul}$ is the upper left $n \times n$ submatrix of G^{-1}, which can be written

$$[[\mathbf{G}^{-1}]]_{11} = \left\{\mathbf{G}_{ul} - \mathbf{g}_{ur} g_{lr}^{-1} \mathbf{g}_{ll}\right\}^{-1} = \{\mathbf{F} + \mathbf{J}\}^{-1} \tag{38}$$

where

$$\mathbf{J} = \mathrm{diag}(\mathbf{h}) - \frac{\mathbf{g}_{ur}\mathbf{g}_{ur}^t}{g_{lr}} \tag{39}$$

Both $\mathbf{F}$ and $\mathbf{J}$ are Hermitian. We know that $\mathbf{F}$ is positive semidefinite. Let $\lambda_k(\mathbf{F}), k = 1 \ldots n$ be the eigenvalues of $\mathbf{F}$ and $\lambda_k(\mathbf{F} + \mathbf{J}), k = 1 \ldots n$ be the eigenvalues of the sum, both listed in increasing order, then if we can show that $\mathbf{J}$ is positive semidefinite, then it is known [11] that:

$$0 \leq \lambda_k(\mathbf{F}) \leq \lambda_k(\mathbf{F} + \mathbf{J}), \forall k = 1 \ldots n \tag{40}$$

Since the eigenvalues of an inverse of a matrix are the inverses of the eigenvalues of the original matrix,

$$\lambda_k\left(\{\mathbf{F} + \mathbf{J}\}^{-1}\right) \leq \lambda_k(\mathbf{F}^{-1}), \forall k = 1 \ldots n \tag{41}$$

If in addition, we can show that $\mathrm{tr}\mathbf{J} > 0$, then $\mathrm{tr}\mathbf{F} + \mathbf{J} > \mathrm{tr}\mathbf{F}$, and therefore $\sum_{k=1}^{n} \lambda_k(\mathbf{F} + \mathbf{J}) > \sum_{k=1}^{n} \lambda_k(\mathbf{F})$. This with (40) implies that $\lambda_j(\mathbf{F} + \mathbf{J}) > \lambda_j(\mathbf{F})$ for at least one $j \in 1 \ldots n$. Thus in addition to (41),

$$\lambda_j\left(\{\mathbf{F} + \mathbf{J}\}^{-1}\right) < \lambda_j(\mathbf{F}^{-1}), \text{for at least one} j \in 1 \ldots n \tag{42}$$

which implies that

$$\mathrm{tr}\left(\{\mathbf{F} + \mathbf{J}\}^{-1}\right) < \mathrm{tr}\mathbf{F}^{-1} \tag{43}$$

which shows that the mean of the bounds on the unbiased estimator variances for the n parameters has strictly decreased.

1) Showing positive semidefiniteness and positive trace of $\mathbf{J}$*:* First, we show that the diagonal elements of $\mathbf{J}$ are all non-negative. These elements, $[[\mathbf{J}]]_{k,k}$ are given by

$$[[\mathbf{J}]]_{k,k} = h_k - \frac{\mathbf{g}_{ur}^2(k)}{g_{lr}}. \tag{44}$$

If $k \notin H(n+1)$ then $h_k = 0$ and $\mathbf{g}_{ur}(k) = 0$, thus the diagonal element equals zero. Otherwise if $k \in H(n+1)$ then the diagonal elements are,

$$[[\mathbf{J}]]_{k,k} = E\left[\left(\frac{\partial}{\partial\theta_k} l_{n+1,k}\right)^2\right] - \frac{E^2\left[\left(\frac{\partial}{\partial\theta_k} l_{n+1,k}\right)\left(\frac{\partial}{\partial\theta_{n+1}} l_{n+1,k}\right)\right]}{\sum_{j\in H(n+1)} E\left[\left(\frac{\partial}{\partial\theta_{n+1}} l_{n+1,j}\right)^2\right]}. \tag{45}$$

Because of the reciprocity condition (2) of Theorem 1, the numerator of the fraction is equal to the square of the $j = k$ term in the sum in the denominator. Thus

$$[[\mathbf{J}]]_{k,k} \geq E\left[\left(\frac{\partial}{\partial\theta_k} l_{n+1,k}\right)^2\right] - E\left[\left(\frac{\partial}{\partial\theta_k} l_{n+1,k}\right)\left(\frac{\partial}{\partial\theta_{n+1}} l_{n+1,k}\right)\right] = 0 \tag{46}$$

The equality will hold if k is the only member of the set $H(n+1)$. If $H(n+1)$ has other members, then $[[\mathbf{J}]]_{k,k}$ will be strictly greater than zero. This proves that with at least two members in the set $H(n+1)$, tr$\mathbf{J} > 0$.

Next, we show that $\mathbf{J}$ is diagonally dominant [11], i.e.,

$$|[[\mathbf{J}]]_{k,k}| \geq \sum_{j=1, j\neq k}^{n} |[[\mathbf{J}]]_{k,j}| \tag{47}$$

$$|h_k - \frac{\mathbf{g}_{ur}^2(k)}{g_{lr}}| \geq \sum_{j=1, j\neq k}^{n} \frac{|\mathbf{g}_{ll}(k)\mathbf{g}_{ll}(j)|}{g_{lr}} \tag{48}$$

Because have shown the diagonal elements are ≥ 0 we can remove the absolute value from the left hand side and add $\frac{\mathbf{g}_{ur}^2(k)}{g_{lr}}$ to both sides. Next, $g_{lr} > 0$ since $H(n+1) \neq \emptyset$, so multiply both sides by g_{lr}. Then we must prove that

$$g_{lr} h_k \geq |\mathbf{g}_{ll}(k)| \sum_{j=1}^{n} |\mathbf{g}_{ll}(j)| \tag{49}$$

We start the proof by noting that if $k \notin H(n+1)$ then $h_k = 0$ and $\mathbf{g}_{ll}(k) = 0$, and the equality holds. If $k \in H(n+1)$, then

$$g_{lr} h_k = E\left[\left(\frac{\partial}{\partial\theta_k} l_{k,n+1}\right)^2\right] \sum_{j\in H(n+1)} E\left[\left(\frac{\partial}{\partial\theta_{n+1}} l_{n+1,j}\right)^2\right] \tag{50}$$

Because of the reciprocity condition (2) of Theorem 1,

$$E^2\left[\left(\frac{\partial}{\partial\theta_k} l_{k,n+1}\right)^2\right] = E^2\left[\left(\frac{\partial}{\partial\theta_{n+1}} l_{k,n+1}\right)\left(\frac{\partial}{\partial\theta_k} l_{k,n+1}\right)\right] \tag{51}$$

Thus

$$\begin{aligned} g_{lr} h_k &= \left|E\left[\left(\frac{\partial}{\partial\theta_{n+1}} l_{k,n+1}\right)\left(\frac{\partial}{\partial\theta_k} l_{k,n+1}\right)\right]\right| \sum_{j\in H(n+1)} \left|E\left[\left(\frac{\partial}{\partial\theta_{n+1}} l_{j,n+1}\right)\left(\frac{\partial}{\partial\theta_j} l_{j,n+1}\right)\right]\right| \\ &= |\mathbf{g}_{ll}(k)| \left\{ \sum_{\substack{j\in H(n+1) \\ j\geq 1}} |\mathbf{g}_{ll}(j)| + \sum_{\substack{j\in H(n+1) \\ j\leq 0}} \left|E\left[\left(\frac{\partial}{\partial\theta_{n+1}} l_{j,n+1}\right)\left(\frac{\partial}{\partial\theta_j} l_{j,n+1}\right)\right]\right| \right\} \end{aligned} \tag{52}$$

Since $\mathbf{g}_{ll}(j) = 0$ if $j \notin H(n+1)$ we can change include in the first sum all $j \in 1 \ldots n$. The first and second sums in the brackets correspond to the terms due to observations with blindfolded

and reference devices, respectively. Thus

$$|g_{lr}h_k| \geq |\mathbf{g}_{ll}(k)| \sum_{j=1}^{n} |\mathbf{g}_{ll}(j)| \tag{53}$$

which is what we needed to prove. If device $n+1$ makes a pairwise comparison with atleast one reference device, the inequality in (53) will be strictly greater than zero. If it makes no observations with reference nodes, then the equality in (53) would hold.

Diagonal dominance implies $\mathbf{J}$ is positive semidefinite. Strict diagonal dominance, which would hold if $H(n+1) \supset 1 \ldots n$ (device $n+1$ makes observations with all blindfolded devices and atleast one reference device), implies $\mathbf{J}$ is positive definite. Positive definiteness assures that the CRB matrix will strictly decrease. The looser positive semidefinite condition implies the inequality in (41).

C. CRB for Location Estimation

For the elements of $\mathbf{F}_R$, we first explicitly state the log likelihood function for the observation between device i and j,

$$l_{i,j} = \log\left(\frac{10}{\sqrt{2\pi\sigma_{dB}^2}\log 10}\frac{1}{P_{i,j}}\right) - \frac{b}{8}\left(\log\frac{d_{i,j}^2}{\hat{d}_{i,j}^2}\right)^2. \tag{54}$$

Recall that $d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. Thus taking the partial of $l_{i,j}$ w.r.t. x_j

$$\frac{\partial}{\partial x_j} l_{i,j} = -b\left(\log\frac{d_{i,j}^2}{\hat{d}_{i,j}^2}\right)\frac{x_j - x_i}{d_{i,j}^2}, \tag{55}$$

Note that $\frac{\partial}{\partial x_j} l_{i,j} = -\frac{\partial}{\partial x_i} l_{i,j}$, thus the log-normal distribution of RSS measurements meets condition (2) of Theorem 1. The second partial derivatives will differ based on whether or not $i = j$ and if the partial is taken w.r.t. y_i or x_i. For example,

$$\begin{aligned}
\frac{\partial^2}{\partial x_j \partial y_j} l_{i,j} &= -b\frac{(x_i - x_j)(y_i - y_j)}{d_{i,j}^4}\left[-\log\left(\frac{d_{i,j}^2}{\hat{d}_{i,j}^2}\right) + 1\right] \\
\frac{\partial^2}{\partial x_j \partial y_i} l_{i,j} &= -b\frac{(x_i - x_j)(y_i - y_j)}{d_{i,j}^4}\left[\log\left(\frac{d_{i,j}^2}{\hat{d}_{i,j}^2}\right) - 1\right]
\end{aligned} \tag{56}$$

All of the 2nd partial derivatives depend on the term, $\log(d_{i,j}^2/\hat{d}_{i,j}^2)$, which has an expected value of zero. We use these simplified terms in (5) for each block of the $\mathbf{F}_R$ matrix given in (17), simplifies considerably and the final FIM takes the form in (19). For the TOA case,

$$l_{i,j} = \left(-\log\sqrt{2\pi\sigma_T^2} - \frac{(T_{i,j} - d_{i,j}/c)^2}{2\sigma_T^2}\right) \tag{57}$$

taking the partial w.r.t. x_j

$$\frac{\partial}{\partial x_j} l_{i,j} = -\frac{1}{\sigma_T^2}\left(\frac{cT_{i,j}}{d_{i,j}} - 1\right)(x_j - x_i), \tag{58}$$

Note that in the TOA case we also have $\frac{\partial}{\partial x_j} l_{i,j} = -\frac{\partial}{\partial x_i} l_{i,j}$, meeting condition (2) of Theorem 1. Two examples of the second partial derivatives are given by,

$$\begin{aligned}
\frac{\partial^2}{\partial x_j \partial y_j} &= -\frac{1}{\sigma_T^2 c^2}\frac{cT_{i,j}}{d_{i,j}}\frac{(x_i - x_j)(y_i - y_j)}{(x_i - x_j)^2 + (y_i - y_j)^2} \\
\frac{\partial^2}{\partial x_j \partial x_i} &= -\frac{1}{\sigma_T^2 c^2}\left[\frac{cT_{i,j}}{d_{i,j}} - 1 - \frac{cT_{i,j}}{d_{i,j}}\frac{(x_i - x_j)^2}{d_{i,j}^2}\right]
\end{aligned}$$

The 2nd partial derivatives depend on the term, $cT_{i,j}/d_{i,j}$, which has an expected value of 1. These simplified terms for each block of the $\mathbf{F}_R$ matrix take the form in (20).

References

[1] J. Albowicz, A. Chen, and L. Zhang. Recursive position estimation in sensor networks. In *IEEE International Conference on Network Protocols*, pages 35–41, Nov 2001.

[2] N. Benvenuto. Distortion analysis on measuring the impulse response of a system using a crosscorrelation method. Technical Report 10, AT&T Bell Laboratories Technical Journal, Dec. 1984.

[3] P.-C. Chen. A non-line-of-sight error mitigation algorithm in location estimation. In *IEEE Wireless Communications and Networking Conference*, pages 316–320, Sept. 1999.

[4] A. J. Coulson, A. G. Williamson, and R. G. Vaughan. A statistical basis for lognormal shadowing effects in multipath fading channels. *IEEE Transactions on Vehicular Technology*, 46(4):494–502, April 1998.

[5] D. M. Devasirvatham. Time delay spread and signal level measurements of 850 MHz radio waves in building environments. *IEEE Trans. on Ant. & Prop.*, AP-34(11):1300–1305, Nov. 1986.

[6] L. Doherty, K. S. pister, and L. E. Ghaoui. Convex position estimation in wireless sensor networks. In *IEEE INFOCOM*, volume 3, pages 1655 –1663, 2001.

[7] G. Durgin, T. Rappaport, and H. Xu. Measurements and models for radio path loss and penetration loss in and around homes and trees at 5.85 GHz. *IEEE Journal on Sel. Areas in Comm.*, 46(11):1484–1496, Nov. 1998.

[8] R. Fleming and C. Kushner. Low-power, miniature, distributed position location and communication devices using ultra-wideband, nonsinusoidal communication technology. Technical report, Aetherwire Inc., Semi-Annual Technical Report, ARPA Contract J-FBI-94-058, July 1995.

[9] H. Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7):943–968, July 1993.

[10] A. O. Hero, J. A. Fessler, and M. Usman. Exploring estimator bias-variance tradeoffs using the uniform CR bound. *IEEE Transactions on Signal Processing*, 44(8):2026–2041, Aug. 1996.

[11] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge Univ. Press, New York, 1990.

[12] D. McCrady, L. Doyle, H. Forstrom, T. Dempsy, and M. Martorana. Mobile ranging with low accuracy clocks. *IEEE Trans. on Microwave Theory and Techniques*, 48(6):951–957, June 2000.

[13] R. L. Moses, D. Krishnamurthy, and R. Patterson. An auto-calibration method for unattended ground sensors. In *ICASSP*, volume 3, pages 2941–2944, May 2002.

[14] K. Pahlavan, P. Krishnamurthy, and J. Beneat. Wideband radio propagation modeling for indoor geolocation applications. *IEEE Communications Magazine*, pages 60–65, April 1998.

[15] N. Patwari, R. J. O'Dea, and Y. Wang. Relative location in wireless networks. In *IEEE VTC*, volume 2, pages 1149–1153, May 2001.

[16] N. Patwari, Y. Wang, and R. J. O'Dea. The importance of the multipoint-to-multipoint indoor radio channel in ad hoc networks. In *IEEE Wireless Communications and Networking Conference*, volume 2, pages 608–612, March 2002.

[17] B. B. Peterson, C. Kmiecik, R. Hartnett, P. M. Thompson, J. Mendoza, and H. Nguyen. Spread spectrum indoor geolocation. *Journal of the Institute of Navigation*, 45(2):97–102, Summer 1998.

[18] J. M. Rabaey, M. J. Ammer, J. L. J. da Silva, D. Patel, and S. Roundy. Picoradio supports ad hoc ultra-low power wireless networking. *IEEE Computer Magazine*, pages 42–48, July 2000.

[19] T. Rappaport. *Wireless Communications: Principles and Practice*. Prentice-Hall Inc., New Jersey, 1996.

[20] J. H. Reed, K. J. Krizman, B. D. Woerner, and T. S. Rappaport. An overview of the challenges and progress in meeting the E-911 requirement for location service. *IEEE Communications Magazine*, pages 30–37, April 1998.

[21] C. Savarese, J. M. Rabaey, and J. Beutel. Locationing in distributed ad-hoc wireless sensor networks. In *ICASSP*, pages 2037–2040, May 2001.

[22] M. A. Spirito. On the accuracy of cellular mobile station location estimation. *IEEE Transactions on Vehicular Technology*, 50(3):674–685, May 2001.

[23] G. L. Turin, F. D. Clapp, T. L. Johnston, S. B. Fine, and D. Lavry. A statistical model of urban multipath propagation. *IEEE Transactions on Vehicular Technology*, VT-21(1):1–9, Feb. 1972.

[24] S. Čapkun, M. Hamdi, and J.-P. Hubaux. GPS-free positioning in mobile ad-hoc network. In 34^{th} *IEEE Hawaii International Conference on System Sciences (HICSS-34)*, Jan. 2001.

[25] A. Ward and A. H. A. Jones. A new location technique for the active office. *IEEE Personal Communications*, 4(5):42–47, Oct. 1997.

[26] J. Werb and C. Lanzl. Designing a positioning system for finding things and people indoors. *IEEE Spectrum*, 35(9):71–78, Sept. 1998.