Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2011, Article ID 980953, 17 pages
doi:10.1155/2011/980953

Research Article

Low-Complexity, Distributed Characterization of Interferers in Wireless Networks

Vibhav Kapnadak,[1] Murat Senel,[2] and Edward J. Coyle[3]

[1] *Network Systems Engineering, AT&T Labs, 2600 Camino Ramon, CA 94583, USA*
[2] *Robert Bosch LLC, Research and Technology Center North America, 4009 Miranda Avenue, Palo Alto, CA 94304, USA*
[3] *School of Electrical and Computer Engineering, Georgia Institute of Technology, 777 Atlantic Drive NW, Atlanta, GA 30332-0250, USA*

Correspondence should be addressed to Edward J. Coyle, ejc@gatech.edu

Received 8 February 2011; Revised 21 May 2011; Accepted 25 May 2011

We consider a large-scale wireless network that uses sensors along its edge to estimate the characteristics of interference from neighboring networks or devices. Each sensor makes a noisy measurement of the received signal strength (RSS) from an interferer, compares its measurement to a threshold, and then transmits the resulting bit to a cluster head (CH) over a noisy communication channel. The CH computes the maximum likelihood estimate (MLE) of the distance to the interferer using these noise-corrupted bits. We propose and justify a low-complexity threshold design technique in which the sensors use nonidentical thresholds to generate their bits. This produces a dithering effect that provides better performance than previous techniques that use different non-identical thresholds or the case in which all the sensor motes use an identical non-optimal threshold. Our proposed technique is also shown (a) to be of low complexity compared to previous non-identical threshold approaches and (b) to provide performance that is very close to that obtained when all sensors use the identical, but unknown, optimal threshold. We derive the Cramér-Rao bound (CRB) and also show that the MLE using our dithered thresholds is asymptotically both efficient and consistent. Simulations are used to verify these theoretical results.

1. Introduction

Large-scale deployments of wireless LANs in unlicensed RF bands are often subject to interference from many sources. We have encountered this problem in the e-Stadium wireless testbed [1], which enables football fans in the stadium at Purdue to access multimedia content related to the game [2] via 802.11b and 3G networks. The locations and coverage areas of the 802.11b access points (APs) in the stadium are shown in Figure 1. The two interfering APs also shown in the figure may disrupt eStadium services for fans sitting along the outer edge of the stands or in the tailgating area north of the stadium. The locations and settings of these interfering access points or devices change over time. The eStadium testbed must sense these changes and adapt its channel assignments and power levels to ensure that its users experience satisfactory Quality of Service (QoS).

The eStadium testbed includes clusters of wireless sensors that are distributed along the concourse area of the stadium [3]. They currently gather and process information from this area and make it available to fans and security personnel during games. We propose that these sensors perform the additional task of characterizing the sources of interference with the 802.11b-based portion of the eStadium testbed.

The information gathered about interferers by these sensors is not directly available to the eStadium APs because of their directional antennas or shadowing by the stadium's structure. A Smart WiFi system [4] would thus not be able to sense and adjust to the presence of these external interferers. This proposed use of the sensors, combined with algorithms to alter the 802.11b channel assignments and power settings, is thus a cognitive networking approach to enabling the coexistence of many systems in unlicensed bands.

The contributions of this paper include the following.

(a) Analysis and improvement of MLE approaches in which sensors conserve energy by minimizing the amount of data they transmit. Each sensor thresholds its noisy measurement of the interferer's signal strength, adds the resulting bit to a packet carrying

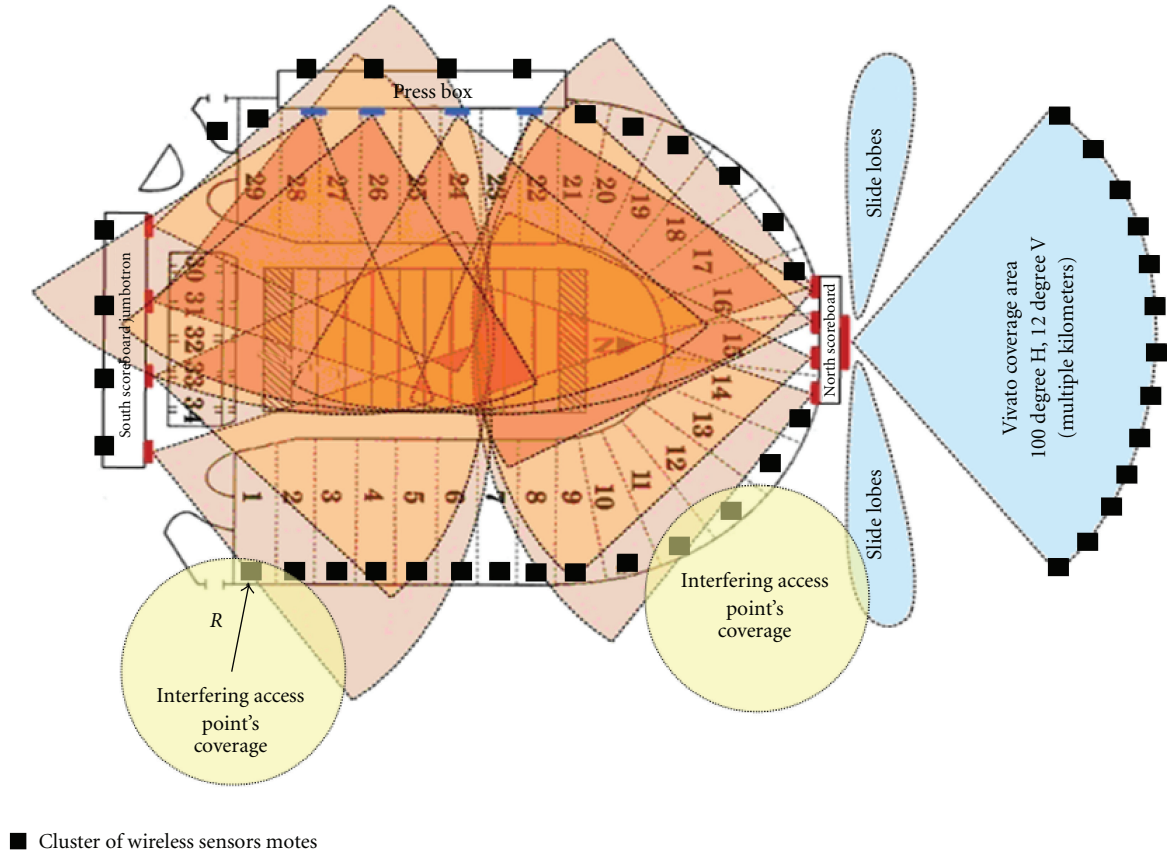


Figure 1: The eStadium testbed at Purdue's Ross Ade stadium includes (1) twelve 802.11 APs with directional antennas that provide coverage, shaded in orange, of the outdoor seating area (2) ten APs with omni antennas that cover the indoor seating areas, (3) a Vivato network that serves the parking area north of the stadium, and (4) clusters of sensors in the concourse area of the stadium. Interfering APs are generally WiFi networks operating in the vicinity of the stadium.

other data, and transmits the packet to the CH via a noisy channel. We consider and justify the case in which each sensor uses a different threshold and compare it with cases in which every sensor uses the same threshold.

(b) Proofs of the asymptotic efficiency and consistency of this new MLE approach and derivation of the Cramér-Rao bound (CRB). Given that the thresholds in our case are non-identical, the standard proofs of asymptotic consistency and efficiency do not hold.

(c) Verification that the MLE based on the use of non-identical thresholds by the sensors performs: (i) as well as one based on the use of identical *optimal* thresholds, even when the number of sensors is small, (ii) significantly better than the one based on the use of identical non-optimal thresholds, and (iii) is of lower complexity and has a performance that is better than or as good as existing non-identical threshold techniques. Because the identical optimal threshold case requires prior knowledge of the true value of the unknown parameter, and existing non-identical techniques involve the solution of a complex optimization problem, our non-identical threshold approach is much more practical and of lower complexity in terms of the design of the thresholds.

In Section 2, we review the most relevant prior work in this area. Sections 3 and 4 describe the system model and the formulation of the sensor fusion problem. The performance of the algorithm is evaluated and compared with existing techniques in Section 5, and conclusions are provided in Section 7.

2. Related Work

In the context of transmitter location estimation for wireless networks, the authors in [5–7] propose expectation maximization (EM) algorithms to locate multiple transmitters using received power levels at arbitrarily placed receivers. The authors in [8] consider a cognitive wireless network in which secondary users measure signal strengths from primary users to estimate the distance to the primary users. They propose estimation algorithms under different channel fading conditions. In [9], the authors study an 802.11 cognitive mesh network (COMNET) in which mesh clients (MC) calculate their distances from the primary stations with triangulation-based localization techniques that use RSS measurements. These localization techniques assume

that the MCs are synchronized and that they are located on the edge of the secondary 802.11 network; otherwise, they may not be able to detect the interference from the primary stations. In the above papers, the authors do not consider the use of thresholded measurements or the transmission of measurements over noisy channels.

In the context of distributed estimation in wireless sensor networks, the authors in [10, 11] consider 1-bit quantization using identical and non-identical thresholds and obtain the maximum-likelihood estimate of an unknown parameter. They account for measurement noise that is Gaussian or has an unknown distribution but do not consider other sources of error, such as noise in the communication channel. A similar situation is examined in [12], where multiple-bit quantized information is used to locate a target in a sensor field under error-free channel conditions. A nonparametric approach to density estimation using a sensor network was introduced in [13]. The authors considered the case where 1-bit quantized data is transmitted to the CH using random thresholds from a pmf. As with the previous papers, however, processing by sensors and the wireless channel between the sensors and the CH are considered to be error free, which is typically not the case in realistic situations. In [14], the authors study channel aware target localization using 1-bit quantized data. They analyze the effects on the root mean squared error (RMSE) of the MLE due to communication channel impairments. They consider the BSC, Rayleigh fading with coherent reception and non-coherent reception. However, in their work, they do not address the design of the 1-bit quantization algorithm for transmission over noisy communication channels.

In [15, 16], the authors consider a 1-bit quantization framework for noisy Gaussian communication links. However, it is assumed that either all sensors use the *same* threshold or use one of two thresholds. More recently, in [17], the authors have proposed an MLE-based distributed estimation algorithm using 1-bit quantized data that is transmitted over a BSC. The context in which the estimation algorithm is analyzed is for secure data transmission in wireless sensor networks, but it is assumed that all the sensors use *identical* thresholds.

Several other papers, including [18–21], also study the distributed estimation problem, but their approach is based on the best linear unbiased estimate (BLUE) at the CH instead of the MLE. This approach is typically not appropriate or optimal when the received data at the CH is a nonlinear function of the unknown parameter.

Our approach is unique because it (a) exploits and justifies the use of different thresholds by the sensors when quantizing their noisy measurements and (b) accounts for noisy measurements by the sensors and error-prone processing and communications between the sensors and the CH. Our analysis of the MLE in this new approach will show that uniformly-spaced thresholds produces a dithering effect that leads to near-optimal performance when the only information available about the parameter being estimated is its support. This is very important because previous approaches based on the use of an identical threshold by all sensors work well *only* when the chosen threshold is either very near the unknown true value, or the number of sensors is very large. Furthermore, our non-identical threshold design algorithm is significantly less complex and performs as good as or better than existing non-identical threshold techniques, such as the one in [10].

3. System Model and Problem Motivation

3.1. System Characteristics and Considerations. The following characteristics of the wireless sensor network and sensor fusion algorithm are assumed in this paper. They are motivated by the eStadium project but are relevant to many other systems operating in unlicensed bands.

(a) *Architecture.* The system consists of a network of wireless APs and a single-hop cluster of N spatially distributed wireless sensors deployed along the edge of the wireless network. Each sensor measures the RSS from an interfering AP, compares it to its threshold, and relays the resulting bit to the CH. The CH fuses the bits it receives to produce an estimate of the signal strength and distance to the interfering AP.

(b) *Measurement Errors.* Due to large-scale fading and shadowing effects, the RSS $X(k)$ observed at sensor k is assumed to be log normally distributed, as in [22, 23]. The mean $\mu(R)$ of this distribution is a function of the distance R from the transmitter to the receiver while the variance σ^2 is independent of R. We assume that $\mu(R)$ follows a path-loss-exponent model; hence,

$$\mu(R) = K - 10\alpha \log(R), \tag{1}$$

where α is the propagation law exponent, and K is the close-in reference power, that is, the power very close to the transmitter. In rest of the paper, the symbol μ will be used interchangeably with $\mu(R)$. On a dB scale, the received RSS at the sensors can be expressed as

$$X_k = \mu(R) + n_k, \quad k = 1 \cdots N, \tag{2}$$

where $\mu(R)$ is the mean and $n_k \sim \mathcal{N}(0, \sigma_n^2)$ is i.i.d. Gaussian noise that is independent of μ. The term σ_n is commonly referred to as the dB spread of the log-normal shadowing and is typically between 4 dB and 12 dB. We refer to this noise as the measurement noise because it corrupts the SNR measurement of each sensor, and we assume that its characteristics are known by the CH. Furthermore, we assume that the sensors are close enough to each other relative to the distance from the interferer that the received RSS at each sensor has the same mean; however, relative to each other, they are spaced far enough to ensure that the noise processes affecting different sensors are independent. The case of dependent noise is left for future work.

(c) *Energy Limits.* Energy efficiency is critical in the design of battery-powered sensor networks, so sensor k thresholds the RSS value it observes to produce only one bit, denoted by b_k, that is to be sent to the CH. The b_k's are transmitted over

a parallel access, noisy communication channel to the CH. These bits may be transmitted individually over the wireless channel or may be combined with other data into packets transmitted over an 802.15.4 (ZigBee) channel [24].

(d) *Communication Channel.* The wireless channel between each sensor, and the CH is modeled as a binary symmetric Channel (BSC) with a crossover probability of ϵ [25]. We assume that the value of ϵ is perfectly estimated at the CH; this can be achieved by the use of pilot training data transmitted by the sensors to the CH during the initialization process. We further assume that the communication channel is static for a period of T seconds; hence, ϵ does not change during the estimation process. For completeness, in Section 5, we also analyze the performance degradation of our algorithms due to a mismatch between the estimated value of ϵ and the true value of ϵ. The choice of a symmetric channel is not required for our analysis but does simplify comparisons. Collisions of bits/packets in the communication channel are assumed to result in their loss and subsequent retransmission. The BSC model thus applies to transmitted bits/packets that are not involved in collisions. Alternatively, one could assume that the sensors' transmissions are scheduled via a collision-free protocol, such as the one in [26], or that they use the TDMA-based frame that is available under the 802.15.4 (ZigBee) protocol [24].

(e) *Reliability and Trustworthiness.* Noisy communications may not be the only source of errors affecting the sensors' decisions before they reach the CH. The sensors may make processing or storage errors. They may be compromised and intentionally report incorrect results with some probability in order to compromise performance without being detected. These additional error sources can be aggregated with the communication errors, resulting in a larger crossover probability. These chapter's results can thus be used to characterize the sensitivity of fusion algorithms to these error sources.

(f) *Industrial, Scientific, and Medical (ISM) Band.* The true mean value of the RSS distribution (see (b), above), denoted by μ_0, is an unknown parameter that lies in the interval $[\mu_l, \mu_u]$. The lower bound, μ_l, corresponds to the maximum range at which the sensors are able to receive and decode a packet from the interferer. The upper bound, μ_u, is related to the maximum transmission power defined by the incumbent network's communication standard. In the eStadium network, these bounds are determined by a typical 802.11b AP operating in the 2.4 GHz band. For a specific example, we use the specs of a Cisco Aironet 1200 series AP [27] with an omni-directional antenna. Its maximum transmission power of 100 mW at a distance of $R = 1$ m from the transmitter corresponds to the upper bound; that is, $\mu_u = -10$ dB or 20 dBm. For the lower bound, the maximum distance at which a typical commercial 802.11 device can receive and decode a packet from an Aironet AP is $R = 200$ m, which corresponds to $\mu_l = -67$ dB or -37 dBm. Figure 2(b) shows the relationship between μ and R for the signal received from a typical AP.

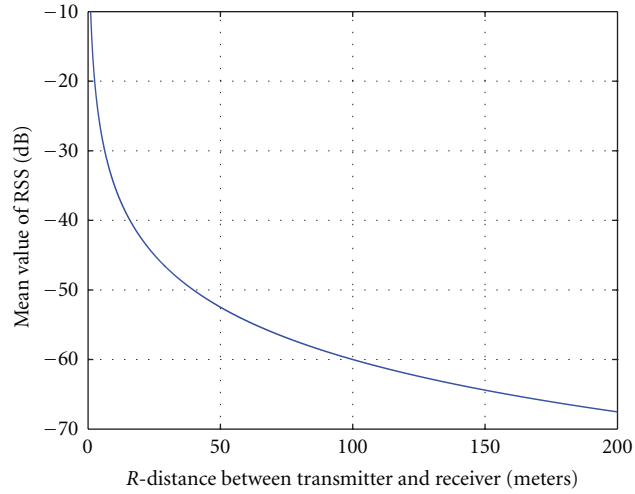


Figure 2: Plot shows the relationship between μ and R with $K = -10$ dB, and $\alpha = 2.5$.

(g) *Periodic Updates.* We assume that the estimates μ and R are updated every T seconds to capture any changes in the set of interferers. T is assumed to be greater than the time required for all sensors to collect a measurement and transmit their bits to the CH. In this paper, we focus on the performance of estimates for a single time period of length T. The better the estimate produced in this period, the faster an iterative algorithm based on it will converge. An iterative algorithm that builds on the results is presented in [28].

In summary, sensor k transmits a single bit of information $b(k)$

$$b_k = \begin{cases} +1 & \text{if } X_k > \tau_k \\ -1 & \text{otherwise,} \end{cases} \tag{3}$$

where τ_k is the threshold used at sensor k to quantize the RSS X_k to one bit. The set of $b(k)'s$ are transmitted to the CH, which uses them to estimate μ and then estimate R via (1). Figure 3 shows the system model and the signal processing block diagram of the cluster of sensors.

3.2. Optimal Identical Threshold Design. The optimal threshold can be defined to be the threshold that minimizes the Cramér-Rao bound (CRB) for any unbiased estimator of μ. This optimization criteria is similar to that proposed in [10, 11] for the error-free communication scenario. In this case, all the sensor motes use the same optimal threshold derived by solving the optimization problem. The thresholds can be expressed as $\tau_k = \tau^{\text{opt}}$ for $k = 1 \cdots N$, where τ^{opt} denotes the optimal threshold. The CRB for any unbiased estimator μ, denoted by $\overline{\mu}$ and derived in Section 5, is given by

$$\text{Var}(\overline{\mu}) \geq \mathcal{I}(\boldsymbol{\tau}, \mu_0)^{-1}, \tag{4}$$

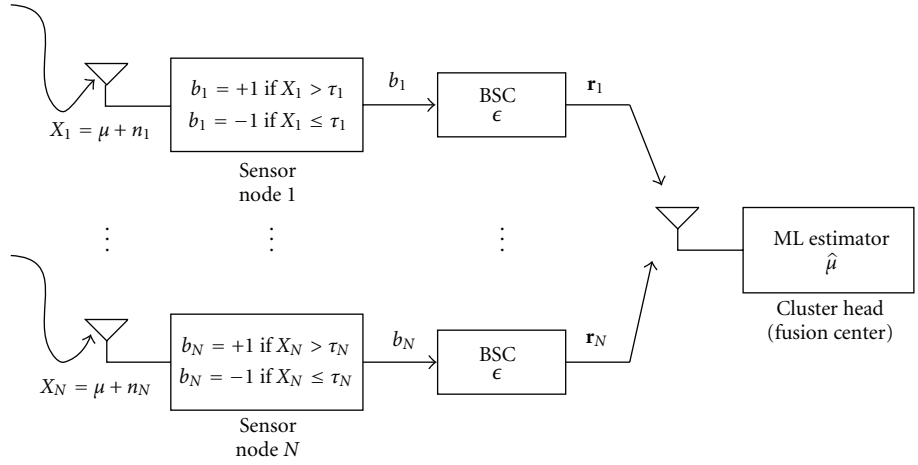


Figure 3: System block diagram for a cluster of wireless sensors. The sensor' RSS measurements are quantized using the dithered quantization technique. The resulting 1-bit values are transmitted over a BSC to the CH, which calculates the MLE of R.

where $\mathcal{I}(\boldsymbol{\tau}, \mu_0)$ is called the Fisher information (FI). As derived later in the paper, the FI is expressed as

$$\mathcal{I}(\boldsymbol{\tau}, \mu_0) = \sum_{k=1}^{N} \mathcal{I}(\tau_k, \mu_0), \tag{5}$$

where

$$\mathcal{I}(\tau_k, \mu_0) = \frac{\left((1-2\epsilon) f_n(\tau_k - \mu_0)\right)^2}{\epsilon + (1-2\epsilon)\overline{F}_n(\tau_k - \mu_0)\left(1 - \epsilon - (1-2\epsilon)\overline{F}_n(\tau_k - \mu_0)\right)} \tag{6}$$

is the FI for sensor k, $\boldsymbol{\tau} = [\tau_1 \cdots \tau_N]$, μ_0 is the true value of μ, and $\overline{F}_n(\cdot)$ and $f_n(\cdot)$ are the complementary cumulative distribution function (ccdf) and pdf of the measurement noise n, respectively. Due to lack of space, we only outline the general technique to obtain the critical points of the CRB function. Calculating the gradient of the r.h.s of (4) with respect to τ_k for $k = 1 \cdots N$, it can easily be shown that when $\tau_k = \mu_0$ for each k, the N equations are simultaneously equal to zero, and the Hessian of the CRB function is positive definite. Hence, the optimal threshold is $\tau^{\text{opt}} = \mu_0$. Note that this choice of threshold is not feasible because μ_0 is not known in advance.

An interesting observation is that the optimal identical threshold choice does not depend on σ_n or ϵ. It is only a function of the true unknown value μ_0. Furthermore, the performance of the identical threshold scheme is very sensitive to the choice of threshold, as shown in Section 6. Even a slight deviation of the threshold value from the true value of μ can result in significant performance degradation. Hence, our goal is to design a quantization scheme which is independent of the true value of μ and has a performance comparable to the optimal identical threshold case.

3.3. Threshold Design for Dithered Quantization. Because use of the optimal identical threshold for the 1-bit quantization step requires prior knowledge of the unknown parameter, we propose a framework in which each sensor uses a different threshold. This ensures that at least a few thresholds will be close to the true value of the unknown parameter. Furthermore, using non-identical thresholds at the sensors for the 1-bit quantization produces a dithering effect that reduces the bias in the estimator (as shown in Section 5). We thus refer to this technique as dithered quantization. Since it is well known that the uniform distribution has maximum entropy among all continuous distributions with compact support [25], we assume that μ is uniformly distributed over the interval $[\mu_l, \mu_u]$ and hence space the quantization thresholds equally over this interval. This leads to the thresholds being assigned for sensor mote k according to the following equation:

$$\tau_k = \mu_l + \frac{k(\mu_u - \mu_l)}{N+1}. \tag{7}$$

3.4. Binning-Based Nonidentical Threshold Design. Another approach to the use of non-identical threshold was proposed in [10]. They consider designing the threshold vector $\boldsymbol{\tau} = \{\tau_k, k \in \mathbb{Z}\}$ and the associated frequency vector $\boldsymbol{\rho} = \{\rho_k, k \in \mathbb{Z}\}$ to minimize the weighted asymptotic variance. The frequency for threshold τ_k is defined as $\rho_k = N_k/N$, where N_k is the total number of sensors transmitting binary information using threshold τ_k, and N is the total number of sensors in the network. Although the problem studied in [10] was for the error-free communication channel scenario, it can be modified for the BSC case. Hence, for our setup, this optimization problem for a uniform weighting function between $[\mu_l, \mu_u]$ denoted by $\mathcal{W}(\mu) = 1/(\mu_u - \mu_l)$ can be formulated as the following second-order cone program (SOCP) with auxiliary variable t as

$$\begin{aligned} \boldsymbol{\rho}^{\star} &= \underset{(t,\boldsymbol{\rho})}{\operatorname{argmin}}\, t, \\ s.t. &\|\mathbf{s} - \mathbf{P}\boldsymbol{\rho}\| \leq t, \\ &\boldsymbol{\rho} \geq \mathbf{0}, \quad \boldsymbol{\rho}^T \mathbf{1} = 1, \end{aligned} \tag{8}$$

where $\mathbf{s} := [S(\mu_0) \cdots S(\mu_M)]^T$, $\boldsymbol{\rho} := [\rho_1 \cdots \rho_L]^T$, L is the number of thresholds, M is the number of discrete points at which the function $S(\mu)$ is evaluated, the function $S(\mu)$ is given by

$$S(\mu) = \mathcal{K}\mathcal{W}^{1/2}(\mu), \qquad \mathcal{K} = \frac{\int_{-\infty}^{+\infty} ((1-2\epsilon) f_n(u))^2 / \Big(\Big(\epsilon + (1-2\epsilon)\overline{F}_n(u)\Big)\Big(1-\epsilon-(1-2\epsilon)\overline{F}_n(u)\Big)\Big) du}{\int_{-\infty}^{+\infty} \mathcal{W}^{1/2}(\mu)\, d\mu}, \tag{9}$$

and $\mathbf{P}$ is a $M \times L$ matrix with its (i, j)th entry given by

$$[\mathbf{P}]_{ij} = \frac{\Big((1-2\epsilon) f_n\Big(\tau_j - \mu_i\Big)\Big)^2}{\Big(\epsilon + (1-2\epsilon)\overline{F}_n\Big(\tau_j - \mu_i\Big)\Big)\Big(1-\epsilon-(1-2\epsilon)\overline{F}_n\Big(\tau_j - \mu_i\Big)\Big)}. \tag{10}$$

Using numerical techniques to approximate the numerator of $\mathcal{K}$, $S(\mu)$ can be expressed as

$$S(\mu) = \frac{\int_{-\infty}^{+\infty} ((1-2\epsilon) f_n(u))^2 / \Big(\Big(\epsilon + (1-2\epsilon)\overline{F}_n(u)\Big)\Big(1-\epsilon-(1-2\epsilon)\overline{F}_n(u)\Big)\Big) du}{|\mu_u - \mu_l|}. \tag{11}$$

For the rest of the paper, we refer to this technique as the binning-based non-identical threshold approach. In Section 5, we compare the performance of all three threshold design schemes.

3.5. Effect of Network Topology on Threshold Design. Wireless sensor networks change with time because sensors can fail as their batteries die, or new sensors may be added to an existing network. Our threshold design is robust to these changes, since when the number of sensors is large, the thresholds are densely distributed over the interval, the loss of a few sensors will not result in a significant loss of information. When the number of sensors is small, the thresholds are coarsely spread over the interval. A few failures might thus result in a significant decrease in the number of bits received by the CH. The CH can then reassign thresholds by communicating with the sensors over the wireless channel. When a new sensor joins the network, it can request a threshold from the CH, choose one based on its unique MAC address, or the CH can reassign all sensors' thresholds according to the algorithm in Section 3.3. On the other hand, the binning-based non-identical threshold technique requires the solution to the optimization problem (8) each time the thresholds need to be assigned. Hence, compared to our proposed approach, the binning-based approach is of much higher computational complexity and therefore cannot easily adapt to changes in the network topology.

4. Sensor Fusion Algorithm

We use maximum-likelihood estimation (MLE) techniques to estimate μ and then use the invariance property of MLE's [29] obtain an estimate for R. For sensor mote k, we denote $\mathcal{R}_k$ to be the random variable (r.v) for the received bit and r_k to be the value of the r.v. Since the received bits $\mathcal{R}_k$ are independent, nonidentically distributed (i.n.i.d) Bernoulli random variables (rv's),

$$\begin{aligned} \Pr(\mathcal{R}_k = 1 \mid \mu) &= \epsilon + (1-2\epsilon)\overline{F}_n(\tau_k - \mu), \\ \Pr(\mathcal{R}_k = -1 \mid \mu) &= (1-\epsilon) - (1-2\epsilon)\overline{F}_n(\tau_k - \mu), \end{aligned} \tag{12}$$

where $\overline{F}_n(\cdot)$ is the complementary cumulative distribution function (ccdf) of the measurement noise $n \sim \mathcal{N}(0, \sigma_n^2)$. Denoting the likelihood function of each received bit by $f(r_k \mid \mu)$, the likelihood function for the received vector of observations $\mathbf{r} = [r_1, \ldots, r_N]$ becomes

$$\begin{aligned} f(\mathbf{r} \mid \mu) &= \prod_{k=1}^{N} f(r_k \mid \mu) \\ &= \prod_{k=1}^{N} \Pr(\mathcal{R}_k = 1 \mid \mu)^{\mathbb{1}_{\{r_k=1\}}} \Pr(\mathcal{R}_k = -1 \mid \mu)^{\mathbb{1}_{\{r_k=-1\}}}. \end{aligned} \tag{13}$$

Let $\mathcal{H}_k(\mu) = \epsilon + (1-2\epsilon)\overline{F}_n(\tau_k - \mu)$, then the log-likelihood function is given by

$$\begin{aligned} &\ln f(\mathbf{r} \mid \mu) \\ &\quad = \sum_{k=1}^{N} [\mathbb{1}_{\{r_k=1\}} \ln(\mathcal{H}_k(\mu)) + \mathbb{1}_{\{r_k=-1\}} \ln(1 - \mathcal{H}_k(\mu))]. \end{aligned} \tag{14}$$

Thus, the MLE $\hat{\mu}$ is obtained by solving the following constrained maximization problem,

$$\begin{aligned} &\underset{\mu}{\text{argmax}} \quad \sum_{k=1}^{N} \left[\mathbb{1}_{\{r_k=1\}} \ln\left(\mathcal{H}_k(\mu)\right) + 1_{\{r_k=1\}} \ln\left(\mathbb{1} - \mathcal{H}_k(\mu)\right)\right], \\ &\text{s.t.} \quad \mu_l \le \mu \le \mu_u, \end{aligned} \tag{15}$$

and because the function relating μ to R in (1) is one to one, the MLE $\hat{R}$ can be obtained directly from $\hat{\mu}$. Note that we have assumed that the CH has perfect knowledge of the dB spread of the log-normal shadowing, that is, σ_n^2 and the cross-over probability ϵ. The former can usually be obtained offline by conducting field measurements in the area of deployment and then performing regression analysis using the experimental data [30]. As for ϵ, this is typically known a priori depending on the modulation and coding scheme used. For example, for transmission of information over fading channels, the probability of bit error (p_e) is well known when standard digital modulation (BSPK, QPSF, 64-QAM, etc.) schemes are used with error control coding strategies such as Turbo or LDPC codes. Hence, p_e can be used as the value for ϵ.

5. Performance Analysis of Dithered Quantization Approach

This section contains (i) proofs of the asymptotic consistency and efficiency of the MLE for μ and R using the proposed dithered quantization approach (ii) derivation of the Cramér-Rao bound (CRB) for the variance of any unbiased estimator for μ and R, and (iii) Monte Carlo simulations that validate the asymptotic results and compare the performance of the dithered quantization technique with the identical threshold-based quantization approach. Since the constrained ML optimization in (15) is analytically intractable, there is no closed-form expression for the estimates or the distribution of the estimates. Hence, to study the bias of the MLE's, we prove that the estimators are asymptotically strongly consistent and hence asymptotically unbiased. Furthermore, although the $\mathcal{R}(k)$'s are i.n.i.d Bernoulli r.v's, we prove that MLE's are asymptotically normally distributed and efficient. Monte Carlo simulations confirm these results and, as shown later, demonstrate that these desirable asymptotic behavior is effectively achieved with as few as $70 \sim 100$ sensors.

5.1. Asymptotic Properties. For the following results, we define $\hat{\mu}_N$, $\hat{R}_N$ to be the MLE of μ and R, respectively, obtained with N sensors using the dithered quantization technique.

Proposition 1. *The MLE $\hat{\mu}_N$ is unique and asymptotically strongly consistent* $\Pr(\lim_{N\to\infty}\hat{\mu}_N = \mu_0) = 1$, *where μ_0 represents the true value of μ.*

Proof. See Appendix A. □

Proposition 2. *$\hat{\mu}_N$ is asymptotically efficient* $(\hat{\mu}_N - \mu_0) \underset{\mathcal{D}}{\rightarrow} \mathcal{N}(0, 1/(\mathcal{I}(\mu_0)))$, *where μ_0 represents the true value of μ, $\mathcal{I}(\cdot)$ is the Fisher information (FI), and $\underset{\mathcal{D}}{\rightarrow}$ means convergence in distribution.*

Proof. See Appendix B. □

The following two corollaries are immediate consequences of the invariance property of functions of MLE's and the one-to-one mapping between μ and R.

Corollary 1. *The MLE $\hat{R}_N$ is unique and asymptotically strongly consistent* $\Pr(\lim_{N\to\infty}\hat{R}_N = R_0) = 1$, *where R_0 represents the true value of R.*

Corollary 2. *$\hat{R}_N$ is asymptotically efficient* $(\hat{R}_N - R_0) \underset{\mathcal{D}}{\rightarrow} \mathcal{N}(0, 1/(\mathcal{I}(R_0)))$, *where R_0 represents the true value of R.*

5.2. Cramér-Rao Bound. In this subsection, we derive the CRB for any unbiased estimator of μ denoted by $\overline{\mu}$. The MLE $\hat{\mu}$ derived previously may be biased for small N, but since we have proved that it is asymptotically unbiased, the CRB is useful as a lower bound of the variance of $\hat{\mu}$ for medium-to-large numbers of sensor motes.

Theorem 1. *The CRB for any unbiased estimator $\overline{\mu}$ obtained using the dithered quantization technique with the BSC model is*

$$\operatorname{Var}(\overline{\mu}) \ge \left[\sum_{k=1}^{N}\left\{\frac{\left((1-2\epsilon) f_n(\tau_k-\mu)\right)^2}{\left(\epsilon + (1-2\epsilon)\overline{F}_n(\tau_k-\mu)\right)\left(1-\epsilon-(1-2\epsilon)\overline{F}_n(\tau_k-\mu)\right)}\right\}\right]^{-1}. \tag{16}$$

Proof. See Appendix C. □

Using the CRB for $\overline{\mu}$ and observing that (1) is a one-to-one transformation, the CRB for any unbiased estimator of R, denoted by $\overline{R}$, can be derived using the invariance property of MLE's and functions of MLE's. The following lemma relates the variance of $\overline{\mu}$ to the variance of $\overline{R}$.

Lemma 1. *If $R = g(\mu)$, then the variance of any unbiased estimate of R is given by*

$$\operatorname{Var}\left(\overline{R}\right) \ge \frac{\left(\partial g/\partial\mu\right)^2}{\mathcal{I}(\mu_0)}. \tag{17}$$

Proof. A proof of this lemma can be found in [29, Chapter 3, Appendix 3A]. □

Corollary 3. *The CRB for any unbiased estimator for $\overline{R}$ using the dithered quantization technique with the BSC model is*

$$\mathrm{Var}\left(\overline{R}\right) \geq \frac{\left((\ln 10)10^{(K-\mu)/10\alpha}/10\alpha\right)^2}{\mathcal{I}(\mu)}, \tag{18}$$

where

$$\mathcal{I}(\mu) = \sum_{k=1}^{N}\left\{\frac{\left((1-2\epsilon)f_n(\tau_k-\mu)\right)_k^2}{\left(\epsilon+(1-2\epsilon)\overline{F}_n(\tau_k-\mu)\right)\left(1-\epsilon-(1-2\epsilon)\overline{F}_n(\tau_k-\mu)\right)}\right\} \tag{19}$$

is the FI.

Proof. Using Lemma 1 and the relationship between R and μ given by $R = 10^{(K-\mu)/10\alpha}$, the result follows. □

Figures 4(a) and 4(b) show the behavior of the square root of the CRB of $\hat{R}$ when we vary the number of sensors, N, and the crossover probability, ϵ, respectively. The CRB is a monotonically increasing function of ϵ for $0 \leq \epsilon < 0.5$ and, due to its symmetry about $\epsilon = 0.5$, is a monotonically decreasing function for $0.5 < \epsilon \leq 1$. As $\epsilon \to 0.5$ the CRB tends to ∞ because, from (1), the Fisher information converges to 0. From Figure 4(a), we observe that the CRB is a monotonically decreasing function of N, which is expected because more sensors means more observations, thus, lower CRB values. Although MLE's tend to have a significant bias for small sample sizes, we have proved that the MLE for our scheme is asymptotically strongly consistent and efficient. Therefore, the analytic derivations above for the CRB for $\overline{\mu}$ and $\overline{R}$ hold in the asymptotic region. The simulation results in the following subsection show that the asymptotic behavior is achieved even with a small number of sensors.

5.3. Numerical Simulation Results. To evaluate the performance of the dithered quantization scheme and compare it with the identical and the binning-based non-identical threshold approaches, we now conduct numerical Monte Carlo simulations. In these simulations (i) the number of sensors is varied from 10 to 200 and (ii) both a low ϵ of 0.05 and a high ϵ of 0.1 are considered. The value of R_0 is set to 50 m, which corresponds to $\mu_0 \approx -52.5$ dB and K is set to -10 dB, the maximum transmission power allowed for 802.11b. We use the typical value of 2.5 for the path-loss exponent α. The standard deviation of the RSS distribution observed at the sensor motes is set to 4 dB; that is, $\sigma_n = 4$ dB.

The results obtained were averaged over 2000 realizations of the measurement noise process. It can easily be shown (see Appendix D) that the log-likelihood function is *not* a concave function of μ for both non-identical threshold-based approaches. Hence, we implemented an 1-dimensional iterated grid search algorithm [31] to locate an approximate maximal value of the log-likelihood function denoted by $\mathcal{L}$. We employed an equidistance grid of size n points given as:

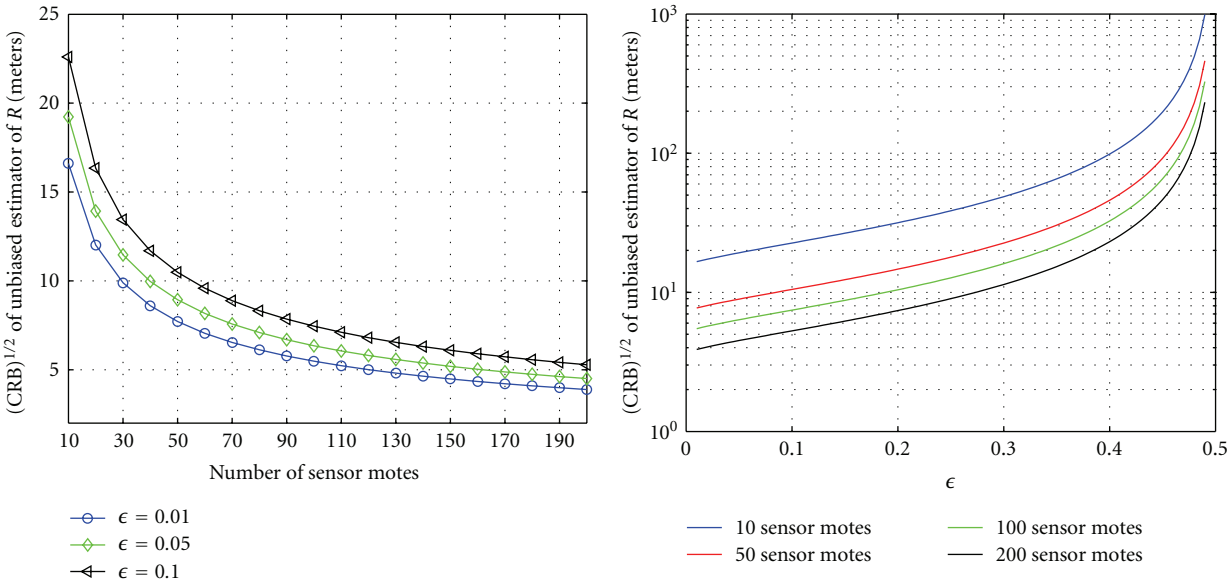
$$\mathcal{S} = \left\{x_k \mid x_k = \mu_l + \frac{k(\mu_u - \mu_l)}{n-1}, \quad k = 0 \cdots n-1\right\}. \tag{20}$$

Using the grid $\mathcal{S}$, the values $\mathcal{L}(x_1), \mathcal{L}(x_2) \cdots \mathcal{L}(x_n)$ are used in the grid search algorithm. An approximate local maximum is located by choosing the grid point x_k that has the largest value for $\mathcal{L}(x_k)$. Using this initialization point, we then used Matlab's sequential quadratic programming (SQP) method to perform the local optimization to determine the MLE.

To characterize the computational complexity of the optimization algorithm, we note that the local optimization using SQP has quadratic convergence rate. The computational complexity of the iterative grid search algorithm can be expressed as $\mathcal{C}(n,r) = rn^m$, where m is the dimension of the parameter space, n is the grid size, and r is the number of iterations. In our simulation experiments, we found that with $r = 10$ and a grid size of $n = 100$ the algorithm converged to an approximate local maximal point quickly. Hence, the computational complexity of the grid search for our case is $O(n)$. In our Matlab implementation of the grid search algorithm, we were able to reduce the complexity to $O(n \log n)$, by the use of sorting algorithms to find the grid point x_k that would correspond to the largest value of the log-likelihood function $\mathcal{L}(x_k)$. In a practical setting, the search grid will be precomputed, and the algorithm will be executed on the CH. Since the CH typically has more processing power than other motes, in a practical setting the convergence of the algorithm would be very fast.

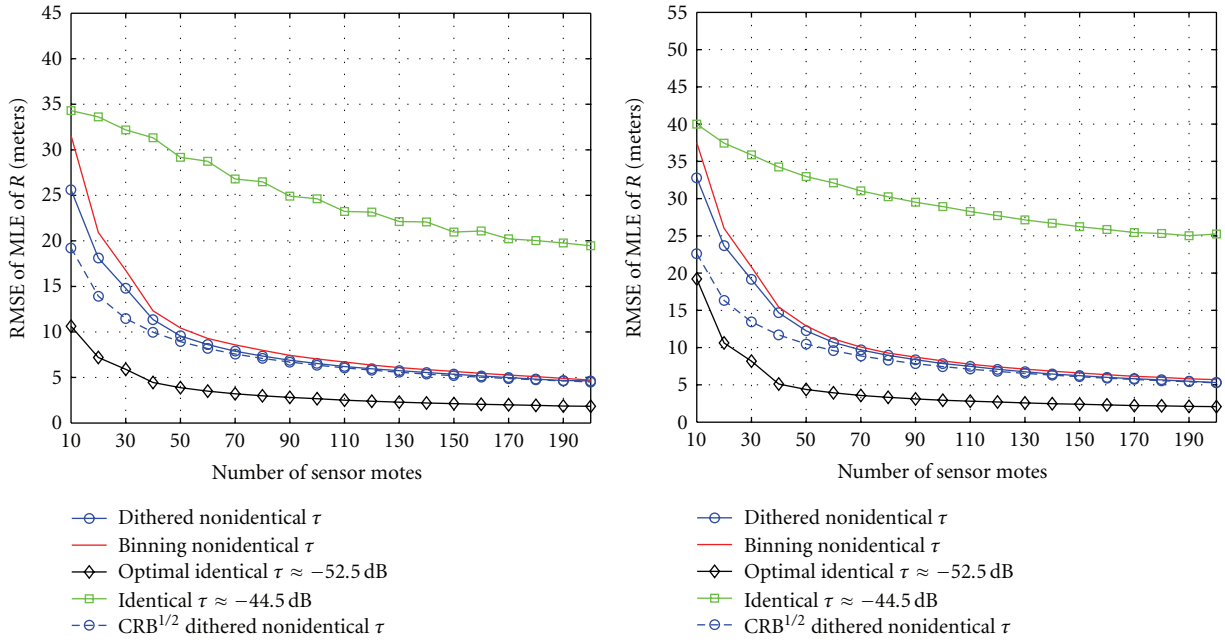
The non-identical thresholds for the dithered quantization technique are generated using the technique described in Section 3, and, for the binning-based technique the SOCP 7 is solved in Matlab using cvx software [32]. As mentioned in [10], the threshold spacing for the binning technique is set to $2\sigma_n$, and hence we have $\tau_{k+1} - \tau_k = 8$ dB, therefore the number of threshold bins $B = (\mu_u - \mu_l)/8 = 8$. For the case where an identical threshold is used by all sensor motes, we consider two scenarios. In the first case, the optimal threshold is used; that is, $\tau = \mu_0$, which corresponds to $\tau \approx -52.5$ dB. In the second case, an identical nonoptimal threshold is used, this corresponds to choosing τ above/below the true mean value of the RSS distribution; that is, $\tau = (\mu_0 + \Delta)$ dB. For simulation purposes, we set $\Delta = 8$; this corresponds to $\tau \approx -44.5$ dB. Similar simulation results are obtained for the case when $\tau > \tau^{\mathrm{opt}}$.

In Figures 5(a) and 5(b), we plot the root mean squared error (RMSE) of $\hat{R}$ for the low and high ϵ cases. We observe that the RMSE using the dithered quantization approach is lower for a small number of sensor motes compared to the non-optimal identical threshold approach and the binning-based non-identical approaches. As the number of sensor motes is increased the RMSE of the binning-based non-identical technique converges to the dithered quantization technique. This is expected since for the binning based approach as the number of sensor motes is increased the threshold frequency (ρ_k) increases which leads to the improvements in the performance.



(a) (b)

FIGURE 4: $(\mathrm{CRB})^{1/2}$ of $\overline{R}$ with $R_0 = 50$ m, $K = -10$ dB, $\alpha = 2.5$, and $\sigma_n = 4$ dB.(a) shows that the analytic expression for the square root of the CRB function is monotonically decreasing function with respect to the number of sensor motes. (b) shows that the analytic expression for the square root of the CRB function is monotonically increasing for the range $0 \leq \epsilon < 0.5$ and as ϵ approaches 0.5 the function approaches ∞ as the Fisher information converges to 0.



(a) (b)

FIGURE 5: RMSE of $\hat{R}$ with $R_0 = 50$ m, $K = -10$ dB, and $\alpha = 2.5$, $\sigma_n = 4$ dB. In (a) $\epsilon = 0.05$ and in (b) $\epsilon = 0.1$. In both cases the dithered quantization technique has a lower RMSE compared to the non-optimal and binning-based thresholding techniques particularly for a low number of sensor motes.

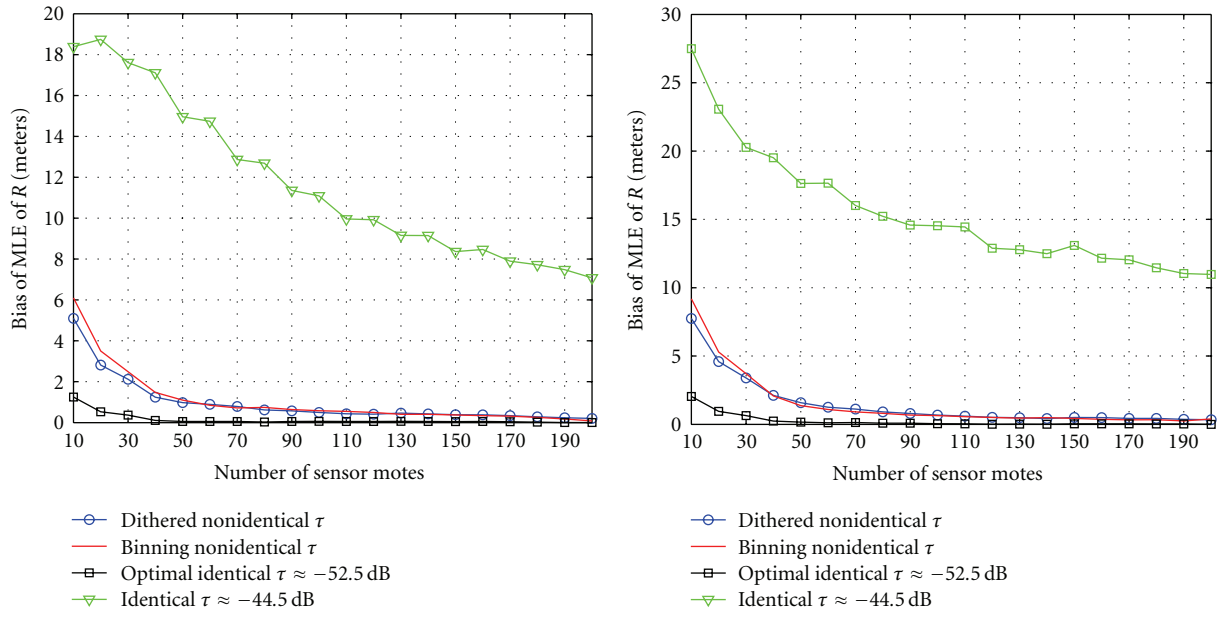


FIGURE 6: Bias of $\hat{R}$ with $R_0 = 50$ m, $K = -10$ dB, and $\alpha = 2.5$, $\sigma_n = 4$ dB. In (a) $\epsilon = 0.05$ and in (b) $\epsilon = 0.1$. In both cases the dithered quantization technique has a lower bias compared to the non-optimal and binning-based thresholding techniques especially for a low number of sensor motes.

Figures 6(a) and 6(b) show the bias of $\hat{R}$ versus the number of sensor motes. We notice that the bias of the MLE using the dithered quantization technique is lower than the non-optimal estimators. Furthermore, we also observe that the bias decreases at an exponential rate with respect to the number of sensor motes and approaches close to 0 with about 100 sensor motes. Figures 5 and 6 also show that the asymptotic properties of the MLE are achievable with about 50 sensor motes for low values of ϵ and about 70 sensor motes for high values of ϵ.

In Figures 7(a) and 7(b), we compare the performance loss incurred using the different thresholding techniques. We define the $\text{CRB}_{\text{Loss}} = \text{CRB}_{\text{non-opt}} - \text{CRB}_{\text{opt}}$, where $\text{CRB}_{\text{non-opt}}$ is the CRB of any unbiased estimator using either the dithered quantization, binning-based or identical non-optimal thresholding techniques, and CRB_{opt} is the CRB of any unbiased estimator using the optimal identical threshold. This metric serves as a good benchmark as, for large number of sensor motes, the MLE's are unbiased, and hence the MSE is equal to the CRB. Figures 7(a) and 7(b) show the CRB_{Loss} versus the number of sensors. We observe that for low and high ϵ values the loss is significantly less for the dithered quantization technique compared with the other schemes, particularly for a small number of sensor motes which is important from a practical point of view.

Figure 8 shows the robustness of our scheme due to a mismatch in the true value of ϵ and the value of ϵ used by the CH for the ML estimation. We notice that even if there is a mismatch, the performance degradation is not significant. The performance of having a perfectly estimated value of ϵ can be replicated using more sensor motes under the condition of a mismatch in ϵ. Furthermore, the gap between the performance between the perfectly estimated ϵ and the mismatched ϵ decreases as the number of sensor motes increases.

In a realistic situation in which the mean of the RSS distribution at the sensor nodes is unknown priori, there is no way to choose the optimal τ for the identical threshold case. It will thus be difficult to achieve good/reliable estimates using non-optimal identical thresholds. In this type of practical scenario, if both approaches (non-optimal identical and non-identical thresholds) use the same number of sensors, and each sensor samples the RSS once, then using the non-identical threshold approach would produce a much more accurate estimate.

Although both non-identical thresholding approaches require fewer sensors than the non-optimal identical threshold technique to achieve near-optimal performance, we observe that compared to the binning approach our thresholding scheme is of low complexity, as it does not involve solving any complex optimization problem to design the thresholds. Furthermore, the dithered quantization technique performs better in terms of having a lower RMSE, variance, and bias compared with the binning and non-optimal identical thresholding techniques, particularly when the number of sensors is small, which is the typical case in a practical scenario.

In summary, the dithered quantization technique has significant advantages over other techniques in terms of either the number of sensors required to achieve a given accuracy or the accuracy of estimates achievable with a fixed number of sensors.

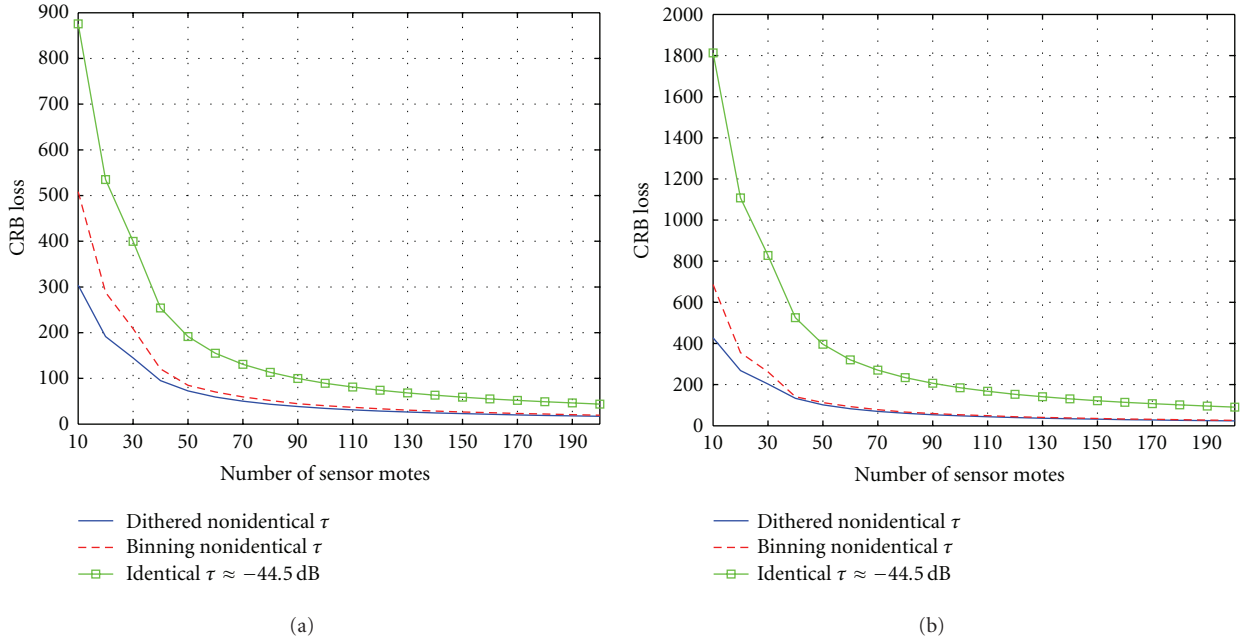


FIGURE 7: $(\text{CRB})^{1/2}$ with $R_0 = 50$ m, $K = -10$ dB, $\alpha = 2.5$, and $\sigma_n = 4$ dB. In (a) $\epsilon = 0.05$ and in (b) $\epsilon = 0.1$, in both cases the loss incurred using the dithered quantization technique is significantly less compared to the other techniques, particularly for low number of sensor motes.

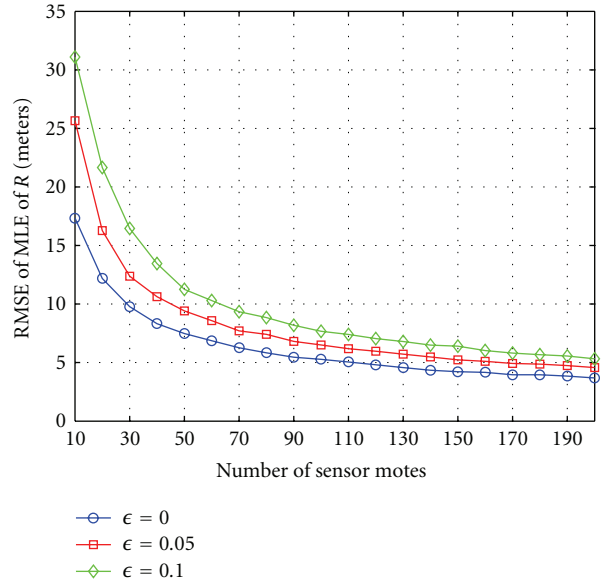


FIGURE 8: shows the RMSE of $\hat{R}$ using dithered quantization technique with $R_0 = 50$ m, $K = -10$ dB, $\alpha = 2.5$, and $\sigma_n = 4$ dB for various values of ϵ. The plot shows the robustness of our estimation scheme due to a mismatch in the true value of ϵ and the estimated value of ϵ used by the CH.

6. Drawbacks of the Identical Threshold Approach

In this section, we study the disadvantages of using the identical threshold approach and the effect on $\hat{R}$ using this technique. As observed in previous work [10, 11], with the error-free communication channel, the disadvantage of the identical threshold approach is that the CRB for $\hat{\mu}$ grows exponentially with $[(\tau - \mu_0)^2/\sigma_n]^2$. A similar observation can be seen for the CRB of $\overline{R}$ when the identical threshold is

used under a BSC model. Using the tight Chernoff bound for the ccdf of the Gaussian distribution (i.e., $\overline{F}_n(\tau-\mu) \leq e^{-(\tau-\mu)^2/2}/2$), the CRB for $\overline{R}$ can be bounded as follows:

$$\begin{aligned}
&\mathrm{CRB}(\epsilon,\tau,R_0)\\
&= G\frac{\left[\epsilon+(1-2\epsilon)\overline{F}_n(\tau-\mu_0)\right]\left[(1-\epsilon)-(1-2\epsilon)\overline{F}_n(\tau-\mu_0)\right]}{N\left[(1-2\epsilon)f_n(\tau-\mu_0)\right]^2}\\
&\leq G\frac{\pi\sigma_n^2 e^{(\tau-\mu_0)^2/2\sigma_n^2}}{2N(1-2\epsilon)^2}\\
&\quad\times\left[\left(2\epsilon+(1-2\epsilon)e^{-(\tau-\mu_0)^2/2\sigma_n^2}\right)\right.\\
&\qquad\left.\times\left(2(1-\epsilon)-(1-2\epsilon)e^{-(\tau-\mu_0)^2/2\sigma_n^2}\right)\right],
\end{aligned} \tag{21}$$

where $G = ((\ln 10)10^{(K-\mu_0)/10\alpha}/10\alpha)^2$. Figure 9(a) shows the exponential increase with respect to $(\tau-\mu_0)/\sigma_n$ for $\epsilon = 0.01$ and $\epsilon = 0.1$ for the CRB of $\hat{R}$. We notice that the Chernoff bound is tight and that a small deviation of τ from the true unknown value μ_0 will result in a significant loss in performance. Although a similar result has previously been widely reported for the error-free channel case [10, 11], to the best of our knowledge, no one has studied the performance of the MLE's in terms of the MSE and variance effects using the identical threshold approach. Hence, to further analyze the effects of using this technique, Figure 9(b) shows the MSE and variance plot as $(\tau-\mu_0)/\sigma_n$ is varied. As with the CRB plot, the performance degradation in terms of the MSE is also exponential as τ deviates from μ_0. However, we notice the existence of multiple critical points. This effect is because the MSE can be decomposed into the $(\text{bias})^2$ and variance. The figure shows that for small deviations from μ_0 the $(\text{bias})^2$ is nearly zero, and the variance increases exponentially; however, when $\tau \gg \mu_0$, the variance starts to decrease, but the bias increases linearly with $(\tau-\mu_0)$. These combined effects reiterate the fact that using identical thresholds could potentially result in estimators with very high bias and MSE.

7. Conclusion

In this paper, we proposed and analyzed a sensor fusion algorithm for interference characterization in wireless networks. To minimize energy usage, each wireless sensor in the network transmits only 1-bit of quantized, noisy RSS information to the FC over a BSC.

Unlike previous approaches that used identical thresholds for all sensors or use non-identical thresholds based on solving a complex optimization problem, we propose a simple and low complexity threshold design technique of uniformly distributing the thresholds over the range of possible values of the mean of the RSS distribution. This approach performs (i) as well as the optimal identical threshold approach and (ii) significantly better than the identical non-optimal threshold design approaches and has a performance that is more accurate or as good as previously proposed non-identical threshold techniques. The optimal identical threshold case is highly unrealistic because the optimal threshold cannot be known in advance; our non-identical threshold technique is much more practical, reliable, and of low complexity design compared to previously proposed techniques based on either using non-identical or identical non-optimal thresholds.

Appendices

A. Asymptotic Consistency of MLE

Proof. We start with the following assumptions that are restated from [33, page 443-444] for completeness and are easy to verify the following.

(A0) The distributions $F(\mathcal{R} \mid \mu)$ of the observations are distinct.

(A1) The distributions $F(\mathcal{R} \mid \mu)$ have common support.

(A2) The observations are $\mathbf{r} = (r_1,\ldots,r_N)$, where the r_k are independent with probability density $f(r_k \mid \mu)$ with respect to the underlying probability measure.

(A3) The parameter space Ω contains an open set ω of which the true parameter value μ_0 is an interior point.

Next, we prove the following lemma.

Lemma 2. *For the independent nonidentically distributed (i.n.i.d) random variables $\mathcal{R}_k$'s we have*

$$\Pr\left\{\lim_{N\to\infty}\ln f\left(\overline{\mathcal{R}} \mid \mu\right) < \ln f\left(\overline{\mathcal{R}} \mid \mu_0\right)\right\} = 1, \quad \forall \mu \neq \mu_0, \tag{A.1}$$

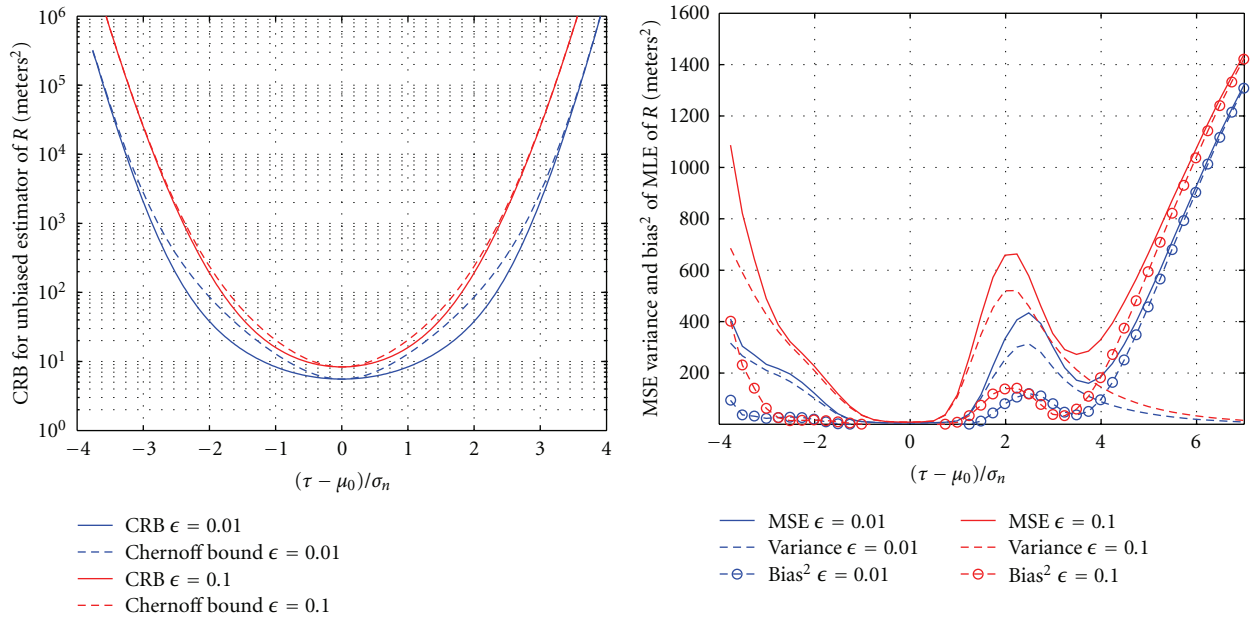
where μ_0 is the true value of μ.

Proof. To prove the lemma, it is equivalent to prove the following:

$$\Pr\left\{\lim_{N\to\infty}\left[\frac{1}{N}\sum_{k=1}^{N}\ln\frac{f(\mathcal{R}_k \mid \mu)}{f(\mathcal{R}_k \mid \mu_0)} < 0\right]\right\} = 1, \quad \forall \mu \neq \mu_0. \tag{A.2}$$

Next, we show that the strong law of large numbers (SLLNs) can be applied to the term $(1/N)\sum_{k=1}^{N}\ln f(\mathcal{R}_k \mid \mu)/f(\mathcal{R}_k \mid \mu_0)$. To show this, we check the Kolmogorov sufficient conditions for the SLLN to hold for independent r.v.'s. Let $Y_k = \ln f(\mathcal{R}_k \mid \mu)/(f(\mathcal{R}_k\mu_0))$, $p_k = \Pr(\mathcal{R}_k = 1 \mid \mu)$, $q_k = \Pr(\mathcal{R}_k = -1 \mid \mu) = (1-p_k)$, $p_{k_0} = \Pr(\mathcal{R}_k = 1 \mid \mu_0)$ and $q_{k_0} = \Pr(\mathcal{R}_k = -1 \mid \mu_0)$. We further assume w.l.o.g. that $\epsilon \leq 0.5$, and using (12) we have the following inequalities:

$$\begin{aligned}
\epsilon \leq p_k \leq 1-\epsilon,\\
1-\epsilon \leq q_k \leq \epsilon.
\end{aligned} \tag{A.3}$$



(a) (b)

FIGURE 9: In (a) the CRB and the Chernoff bound for $\hat{R}$ is shown for varyious values of τ at $R_0 = 50\,\text{m}$ ($\mu_0 \approx -52.5\,\text{dB}$), $N = 100$ and $\sigma_n = 4\,\text{dB}$. Observe that the performance degradation is exponential as $(\tau - \mu_0)/\sigma_n$ increases for different crossover probabilities. (b) shows the MSE, variance and (bias)^2 for $\hat{R}$ with varying τ with $R_0 = 50\,\text{m}$ ($\mu_0 \approx -52.5\,\text{dB}$) $N = 100$ and $\sigma_n = 4\,\text{dB}$. The MSE also has an exponential growth with respect to $(\tau - \mu_0)/\sigma_n$; however, we notice the peculiar behavior of multiple critical points due to the bias and variance effect of the MLE of R.

Similar inequalities hold for p_{k_0} and q_{k_0}, respectively. The variance of Y_k can be bounded as follows:

$$\begin{aligned}
\text{Var}(Y_k) &\le \mathbb{E}_{\mu_0}\left[{Y_k}^2\right] \\
&= p_{k_0}\left(\ln\frac{p_k}{p_{k_0}}\right)^2 + q_{k_0}\left(\ln\frac{q_k}{q_{k_0}}\right)^2 \\
&\le (1-\epsilon)\left(\ln\frac{(1-\epsilon)}{\epsilon}\right)^2 + \epsilon\left(\ln\frac{\epsilon}{(1-\epsilon)}\right)^2 \\
&= 1.
\end{aligned} \tag{A.4}$$

Therefore, $\lim_{N\to\infty}\sum_{k=1}^{N}\text{Var}(Y_k)/k^2 \le \lim_{N\to\infty}\sum_{k=1}^{N}(1/k^2) < \infty$, hence, the SLLN holds for Y_k's, and we have

$$\Pr\left\{\lim_{N\to\infty}\frac{1}{N}\left[\sum_{k=1}^{N}\ln\frac{f(\mathcal{R}_k \mid \mu)}{f(\mathcal{R}_k \mid \mu_0)} - \sum_{k=1}^{N}\mathbb{E}_{\mu_0}\left(\ln\frac{f(\mathcal{R}_k \mid \mu)}{f(\mathcal{R}_k \mid \mu_0)}\right) = 0\right]\right\} = 1, \tag{A.5}$$

where $\mathbb{E}_{\mu_0}(\cdot)$ represents the expectation using the pdf when $\mu = \mu_0$. Now by Jensen's inequality we have

$$\begin{aligned}
&\mathbb{E}_{\mu_0}\left[\ln\frac{f(\mathcal{R}_k \mid \mu)}{f(\mathcal{R}_k \mid \mu_0)}\right] \\
&< ln\mathbb{E}_{\mu_0}\left[\frac{f(\mathcal{R}_k \mid \mu)}{f\left(\mathcal{R}_k \mid \mu_0\right)}\right] \\
&= \ln\left[\frac{\Pr(\mathcal{R}_1 = 1 \mid \mu)}{\Pr(\mathcal{R}_1 = 1 \mid \mu_0)}\Pr(\mathcal{R}_1 = 1 \mid \mu_0)\right. \\
&\qquad \left. + \frac{\Pr(\mathcal{R}_1 = -1 \mid \mu)}{\Pr(\mathcal{R}_1 = -1 \mid \mu_0)}\Pr(\mathcal{R}_1 = -1 \mid \mu_0)\right] \\
&= \ln[\Pr(\mathcal{R}_1 = 1 \mid \mu) + \Pr(\mathcal{R}_1 = -1 \mid \mu)] \\
&= 0.
\end{aligned} \tag{A.6}$$

Using the result (A.6) in (A.5), statement (A.2) is proved, and hence the lemma follows. □

Next we consider a δ-neighborhood about μ_0 such that $(\mu_0 - \delta, \mu_0 + \delta) \in \omega$ and define $S_N = \{\mathbf{r} : \ln f(\mathbf{r} \mid \mu_0 - \delta) < \ln f(\mathbf{r} \mid \mu_0)$ and $\ln f(\mathbf{r} \mid \mu_0) > \ln f(\mathbf{r} \mid \mu_0 + \delta)\}$. Hence, by Lemma 2 we have,

$$\Pr\left\{\lim_{N\to\infty} S_N \mid \mu_0\right\} = 1 \tag{A.7}$$

and, therefore, for all $\mathbf{r} \in S_N$, there exists $\hat{\mu}_N \in (\mu_0 - \delta, \mu_0 + \delta)$ s.t. the likelihood function is maximized in the interval. Since for any $\delta > 0$ small enough, there exists a sequence $\hat{\mu}_N = \hat{\mu}_N(\delta)$ of roots to the equation $\partial \ln f(\mathbf{r} \mid \mu)/\partial\mu = 0$ s.t. $\Pr\{\lim_{N\to\infty} \mid \hat{\mu}_N - \mu_0 \mid < \delta \mid \mu_0\} = 1$. Next we choose a sequence which does not depend on δ, by observing that the likelihood function is a continuous function, and the limit of a sequence of roots is also another root of the equation $\partial \ln f(\mathbf{r} \mid \mu)/\partial\mu = 0$. Therefore, there exists a root μ_N^* that is closest to μ_0; hence, $\Pr(\lim_{N\to\infty} \mid \mu_N^* - \mu_0 \mid < \delta \mid \mu_0) = 1$ and since δ was arbitrary chosen the proposition follows. □

B. Asymptotic Efficiency of MLE

Proof. The key step is to show that the central limit theorem (CLT) holds for the i.n.i.d log-likelihood functions of the $\mathcal{R}_k$'s. As before, we state the following regularity conditions from [33, page 440-441] for completeness.

(a) The parameter space Ω is an open interval (not necessarily finite).

(b) The distributions of $F(\mathcal{R}_k \mid \mu)$ have common support, so that the set $A = \{r_k : f(r_k \mid \mu) > 0\}$ is independent of μ.

(c) For every $r_k \in A$, the density $f(r_k \mid \mu)$ is twice differentiable with respect to μ, and the second derivative is continuous in μ

(d) $\mathbb{E}[\partial \ln f(\mathcal{R}_k \mid \mu)/\partial\mu] = 0$ and $\mathbb{E}[-\partial^2 \ln f(\mathcal{R}_k \mid \mu)/\partial\mu^2] = \mathbb{E}[(\partial \ln f(\mathcal{R}_k \mid \mu)/\partial\mu)^2] = \mathcal{I}(\mu)$.

(e) The Fisher information $0 < \mathcal{I}(\mu) < \infty$.

(f) For any given $\mu_0 \in \Omega$, there exists a positive number c and a function $M(r_k)$ (both of which may depend on μ_0) s.t.

$$\left| \frac{\partial^2 \ln f(r_k \mid \mu)}{\partial\mu^2} \right| \le M(r_k), \quad \forall r_k \in A,\ \mu_0 - c < \mu < \mu_0 + c, \tag{B.1}$$

and $\mathbb{E}_{\mu_0}[M(\mathcal{R}_k)] < \infty$.

Since most of the conditions listed above are easy to verify, we provide proofs only for (d) and (f). To prove the first part of (d) we need to show that $\mathbb{E}_{\mu_0}[\partial \ln f(\mathcal{R}_k \mid \mu)/\partial\mu] = 0$. Using the definitions for p_k and q_k from Appendix A, we have $p_k' = (1-2\epsilon) f_n(\tau_k - \mu)$, $q_k' = -(1-2\epsilon) f_n(\tau_k - \mu)$, $p_k'' = (1-2\epsilon) f_n(\tau_k - \mu)/\sigma_n^2$, and $q_k'' = p_k''$, where the notation $f' = \partial f/\partial\mu$ and $f'' = \partial^2 f/\partial\mu^2$. Therefore, using (12), we have

$$\begin{aligned}
\mathbb{E}_{\mu_0}\left[\frac{\partial \ln f(\mathcal{R} \mid \mu)}{\partial\mu}\right] &= p_{k_0}\left(\frac{p_k'}{p_k}\right) + q_{k_0}\left(\frac{q_k'}{q_k}\right) \\
&= p_{k_0}\left(\frac{p_k'}{p_k}\right) + (1 - p_{k_0})\left(\frac{p_k'}{(1-p_k)}\right) \\
&= 0.
\end{aligned} \tag{B.2}$$

Similarly we have

$$\begin{aligned}
\mathbb{E}_{\mu_0}\left[\left(\frac{\partial \ln f(\mathcal{R} \mid \mu)}{\partial\mu}\right)^2\right] &= \frac{p_{k_0} {p_k'}^2}{{p_k}^2} + \frac{q_{k_0} {q_k'}^2}{{q_k}^2}, \\
\mathbb{E}\left[\frac{-\partial^2 \ln f(\mathcal{R}_k \mid \mu)}{\partial\mu^2}\right] &= -\left(\frac{p_{k_0}\left(p_k'' p_k - {p_k'}^2\right)}{{p_k}^2} + \frac{q_{k_0}\left(q_k'' q_k - {q_k'}^2\right)}{{q_k}^2}\right) \\
&= -\frac{p_{k_0} p_k''}{p_k} + \frac{q_{k_0} p_k''}{q_k} + \frac{p_{k_0} {p_k'}^2}{{p_k}^2} + \frac{q_{k_0} {q_k'}^2}{{q_k}^2} \\
&= \frac{p_{k_0} {p_k'}^2}{{p_k}^2} + \frac{q_{k_0} {q_k'}^2}{{q_k}^2}.
\end{aligned} \tag{B.3}$$

To check condition (f) using the result from Appendix D we have for all $r_k \in A$, $\mu_0 - c < \mu < \mu_0 + c$

$$\begin{aligned}
\left| \frac{\partial^2 \ln f(r_k \mid \mu)}{\partial\mu^2} \right| &\le \mathbb{1}_{\{r_k = 1\}}\left(\frac{(\mu_u - \mu_l)(1-2\epsilon)}{\sqrt{2\pi}\sigma_n^3}\right) \\
&\quad + \mathbb{1}_{\{r_k = -1\}}\left(\frac{(\mu_u - \mu_l)(1-2\epsilon)}{\sqrt{2\pi}\sigma_n^3}\right).
\end{aligned} \tag{B.4}$$

If we let $M(r_k) = \mathbb{1}_{\{r_k=1\}}((\mu_u - \mu_l)(1-2\epsilon)/\sqrt{2\pi}\sigma_n^3) + \mathbb{1}_{\{r_k=-1\}}((\mu_u - \mu_l)(1-2\epsilon)/\sqrt{2\pi}\sigma_n^3)$, and hence

$$\begin{aligned}
\mathbb{E}_{\mu_0}[M(R_k)] &= p_{k_0}\left(\frac{(\mu_u - \mu_l)(1-2\epsilon)}{\sqrt{2\pi}\sigma_n^3}\right) \\
&\quad + q_{k_0}\left(\frac{(\mu_u - \mu_l)(1-2\epsilon)}{\sqrt{2\pi}\sigma_n^3}\right) < \infty.
\end{aligned} \tag{B.5}$$

Next we prove the CLT for independent nonidentically distributed (i.n.i.d) r.v.'s, which will be used to show the asymptotic efficiency result.

Lemma 3. *For the independent, nonidentically distributed (i.n.i.d) log-likelihood function* $\ln f(\overline{\mathcal{R}} \mid \mu_0)$*, we have*

$$\frac{1}{\sqrt{N}} \frac{\partial}{\partial\mu_0} \ln f\left(\overline{\mathcal{R}} \mid \mu_0\right) \underset{\mathcal{D}}{\to} \mathcal{N}\left(0, \frac{\mathcal{I}(\mu_0)}{N}\right), \tag{B.6}$$

where μ_0 *is the true value of* μ, $\mathcal{I}(\mu_0) = \sum_{k=1}^{N} (1-2\epsilon)^2 f_n^2(\tau_k - \mu_0)/p_{k_0} q_{k_0}$ *is the Fisher information matrix,* $p_{k_0} = \Pr(\mathcal{R}_k = 1 \mid \mu_0)$, $q_{k_0} = \Pr(\mathcal{R}_k = -1 \mu_0)$, *and* $\underset{\mathcal{D}}{\to}$ *means convergence in distribution.*

Proof. For ease of notation let $Y_k = \partial/\partial\mu_0 \ln f(\mathcal{R}_k \mid \mu_0)$, $\sigma_{Y_k}^2 = \mathbb{E}_{\mu_0}[Y_k^2]$, due to the regularity conditions, we have $\mathbb{E}_{\mu_0}[Y_k] = 0$. Therefore, the l.h.s of (B.6) can be written as

$$\frac{1}{\sqrt{N}} \frac{\partial}{\partial\mu_0} \ln f\left(\overline{\mathcal{R}} \mid \mu_0\right) = \frac{\sqrt{N}}{N} \sum_{k=1}^{N} Y_k. \tag{B.7}$$

Since the Y_k's are not i.i.d., to apply the CLT we must verify the Lindeberg sufficiency condition. Hence we need to show that for all $\epsilon > 0$,

$$L(N) = \frac{1}{S_N^2}\sum_{k=1}^{N}\mathbb{E}_{\mu_0}\left\{|Y_k|^2\right\}\mathbb{1}_{\{|Y_k|>\epsilon S_N\}} \underset{N\to\infty}{\to} 0, \tag{B.8}$$

where $S_N^2 = \sum_{k=1}^{N}\sigma_{Y_k}^2$. It can be shown that $\sigma_{Y_k}^2 = (1-2\epsilon)^2 f_n^2(\tau_k-\mu_0)/p_{k_0}q_{k_0}$, where $p_{k_0} = \Pr(\mathcal{R}_k = 1 \mid \mu_0)$ and $q_{k_0} = \Pr(\mathcal{R}_k = -1 \mid \mu_0)$. Using this we have

$$S_N^2 = \sum_{k=1}^{N}\frac{(1-2\epsilon)f_n^2(\tau_k-\mu_0)}{p_{k_0}q_{k_0}} \overset{(a)}{\geq} 4(1-2\epsilon)\sum_{k=1}^{N}f_n^2(\tau_k-\mu_0), \tag{B.9}$$

where (a) in (B.9) follows from the fact that $1/p_{k_0}q_{k_0} \geq 4$ for $p_{k_0}, q_{k_0} \in (0,1)$. Therefore,

$$\begin{aligned} L(N) &= \frac{1}{S_N^2}\sum_{k=1}^{N}\int_{\{|Y_k|>\epsilon S_N\}}|Y_k|^2 dF(Y_k) \\ &= \frac{1}{S_N^2}\sum_{k=1}^{N}\int_{\mathbb{R}}|Y_k|^2\mathbb{1}_{\{|Y_k|>\epsilon S_N\}}dF(Y_k). \end{aligned} \tag{B.10}$$

We observe that from (B.9) that $S_N^2 \underset{N\to\infty}{\to} \infty$ hence

$$\lim_{N\to\infty}L(N) = \lim_{N\to\infty}\frac{1}{S_N^2}\sum_{k=1}^{N}\int_{\mathbb{R}}|Y_k|^2\mathbb{1}_{\{|Y_k|>\epsilon S_N\}}dF(Y_k) = 0, \tag{B.11}$$

where (B.11) follows from the fact that $\lim_{N\to\infty}\int_{\mathbb{R}}|Y_k|^2\mathbb{1}_{\{|Y_k|>\epsilon S_N\}}dF(Y_k) = 0$ by the Lebesgue dominated convergence theorem. Hence, the Lindeberg conditions are satisfied and the result follows. □

For rest of the proof, let $\mathcal{L}'(\mu) = \sum_{k=1}^{N}(\partial/\partial\mu)\ln f(\mathcal{R}_k \mid \mu)$, $\mathcal{L}''(\mu) = \sum_{k=1}^{N}(\partial^2/\partial\mu^2)\ln f(\mathcal{R}_k \mid \mu)$. Then, by the mean value theorem, $\exists\ \lambda \in (0,1)$ with $\mu_N^* = \mu_0 + \lambda(\hat{\mu}_N - \mu_0)$ for some $\mu_N^* \in (\mu_0, \hat{\mu}_N)$ such that $\mathcal{L}''(\hat{\mu}_N)$:

$$\mathcal{L}'(\hat{\mu}_N) = \mathcal{L}'(\mu_0) + (\hat{\mu}_N - \mu_0)\mathcal{L}''(\mu_N^*). \tag{B.12}$$

Since $\hat{\mu}_N$ is the MLE, the l.h.s. of (B.12) is equal to 0. We thus obtain

$$\sqrt{N}(\hat{\mu}_N - \mu_0) = \frac{(1/\sqrt{N})\mathcal{L}'(\mu_0)}{-(1/N)\mathcal{L}''(\mu_N^*)}. \tag{B.13}$$

Now since $\hat{\mu}_N$ is a strongly consistent estimator of μ_0, we have $\mu_N^* \underset{\mathcal{P}}{\to} \mu_0$, where $\underset{\mathcal{P}}{\to}$ means convergence in probability. This implies $-(1/N)\mathcal{L}''(\mu_N^*) \underset{\mathcal{P}}{\to} -(1/N)\mathcal{L}''(\mu_0)$, by the fact that if $\mu_N^* \underset{\mathcal{P}}{\to} \mu_0$, then $g(\mu_N^*) \underset{\mathcal{P}}{\to} g(\mu_0)$ for a continuous function $g(\cdot)$. Using arguments similar to those in Appendix A, it can be shown that the Kolmogorov sufficient conditions for the SLLN are satisfied for $(1/N)\mathcal{L}''(\mu_0)$, hence $-(1/N)\mathcal{L}''(\mu_0) \underset{\mathcal{P}}{\to} \mathcal{I}(\mu_0)/N$. By the CLT proved earlier, we have $(1/\sqrt{N})\mathcal{L}'(\mu_0) \underset{\mathcal{D}}{\to} \mathcal{N}(0, \mathcal{I}(\mu_0)/N)$. Hence by Slutsky's theorem, we thus have $\sqrt{N}(\hat{\mu}_N - \mu_0) \underset{\mathcal{D}}{\to} \mathcal{N}(0, N/\mathcal{I}(\mu_0))$ and the proposition follows. □

C. Cramér-Rao Bound

Proof. Using the definition of Fisher information matrix, we have

$$\begin{aligned} -\mathbb{E}_{\mu}\left[\frac{\partial^2 \ln f\left(\overline{\mathcal{R}} \mid \mu\right)}{\partial\mu^2}\right] &= -\sum_{k=1}^{N}\mathbb{E}_{\mu_0}\left[\frac{\partial^2 \ln f(\mathcal{R}_k \mid \mu)}{\partial\mu^2}\right] \\ &= -\sum_{k=1}^{N}\left[\frac{\partial^2 \ln f(\mathcal{R}_k = 1 \mid \mu)}{\partial\mu^2}\Pr(\mathcal{R}_k = 1 \mid \mu)\right. \\ &\qquad \left. + \frac{\partial^2 \ln f(\mathcal{R}_k = -1 \mid \mu)}{\partial\mu^2}\Pr(\mathcal{R}_k = -1 \mid \mu)\right]. \end{aligned} \tag{C.1}$$

Differentiating $\ln f(\mathcal{R}_k = 1 \mid \mu)$ and $\ln f(\mathcal{R}_k = -1 \mid \mu)$ with respect to μ twice, we have:

$$\begin{aligned} \frac{\partial^2 \ln f(\mathcal{R}_k = 1 \mid \mu)}{\partial\mu^2} &= \frac{-(1-2\epsilon)^2 f_n^2(\tau_k-\mu) + \Pr(\mathcal{R}_k = 1 \mid \mu)(1-2\epsilon)f_n'(\tau_k-\mu)}{P(\mathcal{R}_k = 1 \mid \mu)^2}, \\ \frac{\partial^2 \ln f(\mathcal{R}_k = -1 \mid \mu)}{\partial\mu^2} &= \frac{-(1-2\epsilon)^2 f_n'(\tau_k-\mu)\Pr(\mathcal{R}_k = -1 \mid \mu) - (1-2\epsilon)^2 {f_n}^2(\tau_k-\mu)}{\Pr(\mathcal{R}_k = -1 \mid \mu)^2}. \end{aligned} \tag{C.2}$$

Using (C.2) and (C.1) can be expressed as

$$\mathcal{I}(\mu_0) = -\sum_{k=1}^{N}\left\{\frac{((1-2\epsilon)f_n(\tau_k-\mu))^2}{\left(\epsilon+(1-2\epsilon)\overline{F}_n(\tau_k-\mu)\right)\left(1-\epsilon-(1-2\epsilon)\overline{F}_n(\tau_k-\mu)\right)}\right\}. \tag{C.3}$$

Hence Theorem 1 is proved. □

D. Concavity of Log-Likelihood Function

Lemma 4. *The log-likelihood function* $\ln f(\mathbf{r} \mid \mu)$ *is not a concave function.*

Proof. Let $G(\mu) = \sum_{k=1}^{N}(\mathbb{1}_{\{r_k=1\}}\ln\alpha_k(\mu) + \mathbb{1}_{\{r_k=-1\}}\ln\beta_k(\mu))$, where $\alpha(\mu) = \epsilon + (1-2\epsilon)\overline{F}_n(\tau_k-\mu)$, $\beta(\mu) = (1-\epsilon) - (1-2\epsilon)\overline{F}_n(\tau_k-\mu)$, and w.l.o.g we assume that $\epsilon < 0.5$. From the properties of concave functions, we know that if $\alpha(\mu)$ and $\beta(\mu)$ are concave and positive, then $G(\mu)$ is concave. Consider the terms $\alpha_k(\mu)$ and $\beta_k(\mu)$, the first and second derivatives of these terms can be expressed as

$$\begin{aligned}
\alpha_k'(\mu) &= (1-2\epsilon)f_n(\tau_k-\mu),\\
\alpha_k''(\mu) &= \frac{(\tau_k-\mu)\alpha_k'(\mu)}{\sigma_n^2},\\
\beta_k'(\mu) &= -(1-2\epsilon)f_n(\tau_k-\mu) = -\alpha_k'(\mu),\\
\beta_k''(\mu) &= \frac{-(\tau_k-\mu)\alpha_k'(\mu)}{\sigma_n^2}.
\end{aligned} \tag{D.1}$$

Using the above equations, we can express the second derivative of the log-likelihood function as

$$\begin{aligned}
G''(\mu) = \sum_{k=1}^{N}\Big(&\mathbb{1}_{\{r_k=1\}}\left[A_k(\mu)(\tau_k-\mu) - (B_k(\mu))^2\right]\\
&+\mathbb{1}_{\{r_k=-1\}}\left[-C_k(\mu)(\tau_k-\mu) - (D_k(\mu))^2\right]\Big),
\end{aligned} \tag{D.2}$$

where $A_k(\mu) = \alpha_k'(\mu)/\sigma_n^2\alpha_k(\mu)$, $B_k(\mu) = \alpha_k'(\mu)/\alpha_k(\mu)$, $C_k(\mu) = \alpha_k'(\mu)/\sigma_n^2\beta_k(\mu)$, and $D_k(\mu) = \alpha_k'(\mu)/\beta_k(\mu)$. For $G(\mu)$ to be concave, a sufficient condition is that for all $\mu \in [\mu_l, \mu_u]$, $G''(\mu) \le 0$, it can be seen from (D.2) that depending on the values of τ_k and the received bits r_k the sufficiency condition may not be satisfied. For example, consider the scenario with $N = 2$ in this case $\tau_1 \approx -29.2$ dB, and $\tau_2 \approx -48.4$ dB. With the received vector $\mathbf{r} = [r_1 r_2] = [1\,1]$, the second derivative of the log-likelihood function can be expressed as

$$\begin{aligned}
G''(\mu) &= \frac{\alpha_1'(\mu)(\tau_1-\mu)}{\sigma_n^2\alpha_1(\mu)} - \left(\frac{\alpha_1'(\mu)}{\alpha_1(\mu)}\right)^2\\
&\quad+\frac{\alpha_2'(\mu)(\tau_2-\mu)}{\sigma_n^2(1-\alpha_2(\mu))} - \left(\frac{\alpha_2'(\mu)}{(1-\alpha_2(\mu))}\right)^2.
\end{aligned} \tag{D.3}$$

An equivalent sufficiency condition to show that $G(\mu)$ is concave is if $-G''(\mu) \ge 0$, for all $\mu \in [\mu_l, \mu_u]$. Using the definitions of $\alpha_k(\mu)$ and $\alpha_k'(\mu)$, we can obtain the following upper bounds

$$\begin{aligned}
\alpha_k'(\mu) &\le (1-2\epsilon)f_n(0) = \frac{(1-2\epsilon)}{\sqrt{2\pi}\sigma_n},\\
\frac{1}{(1-\epsilon)} &\le \frac{1}{\alpha_k(\mu)} \le \frac{1}{\epsilon}.
\end{aligned} \tag{D.4}$$

Then $-G''(\mu)$ can be upper bounded as follows:

$$\begin{aligned}
-G''(\mu) &= \frac{\alpha_1'(\mu)(\mu-\tau_1)}{\sigma_n^2\alpha_1(\mu)} + \left(\frac{\alpha_1'(\mu)}{\alpha_1(\mu)}\right)^2 + \frac{\alpha_2'(\mu)(\mu-\tau_2)}{\sigma_n^2(1-\alpha_2(\mu))}\\
&\quad+\left(\frac{\alpha_2'(\mu)}{(1-\alpha_2(\mu))}\right)\\
&\le \frac{(1-2\epsilon)(\mu-\tau_1)}{\epsilon\sqrt{2\pi}\sigma_n^3} + \left(\frac{1-2\epsilon}{\sqrt{2\pi}\sigma_n\epsilon}\right)^2\\
&\quad+\frac{(1-2\epsilon)(\mu-\tau_2)}{\epsilon\sqrt{2\pi}\sigma_n^3} + \left(\frac{1-2\epsilon}{\sqrt{2\pi}\sigma_n\epsilon}\right)^2\\
&= \frac{(1-2\epsilon)}{\epsilon\sigma_n^2\sqrt{\pi}}\left(\frac{2\mu-(\tau_1+\tau_2)}{\sqrt{2}\sigma_n} + \frac{(1-2\epsilon)}{\epsilon\sqrt{\pi}}\right)\\
&\approx \frac{(1-2\epsilon)}{\epsilon\sigma_n^2\sqrt{\pi}}\left(\frac{2\mu+77.6}{\sqrt{2}\sigma_n} + \frac{(1-2\epsilon)}{\epsilon\sqrt{\pi}}\right).
\end{aligned} \tag{D.5}$$

It can be seen from (D.5) that for $\mu_l \le \mu \le \sqrt{2}\sigma_n(2\epsilon-1)/(2\epsilon\sqrt{\pi}) - 38.8$ the second derivative of the log-likelihood function $-G'' < 0$ and hence $G(\mu)$ will not be concave for this range of μ. With $\epsilon = 0.1$, we have that for $\mu_l \le \mu < -51.6$ dB $G''(\mu) > 0$. Therefore, we can conclude that log-likelihood function is not concave for all $\mu \in [\mu_l, \mu_u]$. □

References

[1] X. Zhong, H. H. Chan, T. J. Rogers, C. P. Rosenberg, and E. J. Coyle, "The development and eStadium testbeds for research and development of wireless services for large-scale sports venues," in *Proceedings of the 2nd International IEEE/Create-Net Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TRIDENTCOM '06)*, pp. 340–348, March 2006.

[2] A. Ault, J. V. Krogmeier, S. R. Dunlop, and E. J. Coyle, "eStadium: the mobile wireless football experience," in *Proceedings of the 3rd International Conference on Internet and Web Applications and Services (ICIW '08)*, pp. 644–649, June 2008.

[3] X. Zhong and E. J. Coyle, "eStadium: a wireless "living lab" for safety and infotainment applications," in *Proceedings of the 1st International Conference on Communications and Networking in China (ChinaCom '06)*, 2007.
[4] A. Hills, "Smart Wi-Fi," *Scientific American*, vol. 285, no. 10, pp. 86–94, 2005.
[5] J. K. Nelson and M. R. Gupta, "An EM technique for multiple transmitter localization," in *Proceedings of the 41st Annual Conference on Information Sciences and Systems (CISS '07)*, pp. 610–615, March 2007.
[6] J. K. Nelson, M. U. Hazen, and M. R. Gupta, "Global optimization for multiple transmitter localization," in *Proceedings of the Military Communications Conference (MILCOM '06)*, pp. 1–7, October 2006.
[7] T. Roos, P. Myllymaki, and H. Tirri, "A statistical modeling approach to location estimation," *IEEE Transactions on Mobile Computing*, vol. 1, no. 1, pp. 59–69, 2002.
[8] X. Liu and S. Shankar, "Sensing-based opportunistic channel access," *Mobile Networks and Applications*, vol. 11, no. 4, pp. 577–591, 2006.
[9] K. R. Chowdhury and I. F. Akyildiz, "Cognitive wireless mesh networks with dynamic spectrum access," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 168–181, 2008.
[10] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks—part I: gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131–1143, 2006.
[11] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks—part II: unknown probability density function," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2784–2796, 2006.
[12] R. Niu and P. K. Varshney, "Target location estimation in sensor networks with quantized data," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4519–4528, 2006.
[13] A. Dogandžić and B. Zhang, "Nonparametric probability density estimation for sensor networks using quantized measurements," in *Proceedings of the 41st Annual Conference on Information Sciences and Systems (CISS '07)*, pp. 759–764, March 2007.
[14] O. Ozdemir, R. Niu, and P. K. Varshney, "Channel aware target localization with quantized data in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1190–1202, 2009.
[15] T. C. Aysal and K. E. Barner, "Constrained decentralized estimation over noisy channels for sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1398–1410, 2008.
[16] T. C. Aysal and K. E. Barner, "Blind decentralized estimation for bandwidth constrained wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, Article ID 4524301, pp. 1466–1471, 2008.
[17] T. C. Aysal and K. E. Barner, "Sensor data cryptography in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, Article ID 4470151, pp. 273–289, 2008.
[18] Z. Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2210–2219, 2005.
[19] S. Cui, J. J. Xiao, A. J. Goldsmith, Z. Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, 2007.
[20] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, 2006.
[21] J. J. Xiao and Z. Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564–3575, 2005.
[22] E. Visotsky, S. Kuffher, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proceedings of the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, pp. 338–345, 2005.
[23] A. Ghasemi and E. S. Sousa, "Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing," *IEEE Communications Letters*, vol. 11, no. 1, pp. 34–36, 2007.
[24] IEEE802.15.4, "Ieee standard 802 part 15.4 wireless medium access control (mac) and physical layer (phy) specifications for low rate wireless personal area networks (wpans)," 2006.
[25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991.
[26] S. Bandyopadhyay, Q. Tian, and E. J. Coyle, "Spatio-temporal sampling rates and enery efficiency in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 6, pp. 1339–1352, 2005.
[27] Cisco, "Cisco aironet 1200 series access point data sheet," [Online], 2006, http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps430/ps4076/product_data_sheet09186a00800937a6.html.
[28] V. Kapnadak, M. Senel, and E. J. Coyle, "Distributed iterative quantization for interference characterization in wireless networks," *Digital Signal Processing*. In press.
[29] S. M. Kay, *Fundamentals of Statistical Signal Processing-Estimation Theory*, vol. 1, PrenticeHall PTR, Upper Saddle River, NJ, USA, 1993.
[30] T. S. Rappaport, *Wireless Communications- Principles and Practices*, PrenticeHall PTR, Upper Saddle River, NJ, USA, 2002.
[31] R. A. Thisted, *Elements of Statistical Computing*, vol. 1, Chapman and Hall, 1988.
[32] "Cvx: Matlab software for disciplined convex programming," http://cvxr.com/cvx/.
[33] E. Lehmann and G. Casella, *Theory of Point Estimation*, Springer, 1998.