

# Error bounds for wireless localization in NLOS environments

Omotayo Oshiga, Ali Nyangwarimam Obadiah

Department of Electrical and Electronics Engineering, Nile University of Nigeria, Nigeria

---

## Article Info

### Article history:

Received Jan 15, 2020

Revised Apr 18, 2020

Accepted Apr 30, 2020

---

### Keywords:

Cramèr-Rao lower bound

Error analysis

Non-parametric Estimation

Position error bound

Wireless sensor networks

---

## ABSTRACT

An efficient and accurate method to evaluate the fundamental error bounds for wireless sensor localization is proposed. While there already exist efficient tools like Cramèr-Rao lower bound (CRLB) and position error bound (PEB) to estimate error limits, in their standard formulation they all need an accurate knowledge of the statistic of the ranging error. This requirement, under Non-Line-of-Sight (NLOS) environments, is impossible to be met a priori. Therefore, it is shown that collecting a small number of samples from each link and applying them to a non-parametric estimator, like the gaussian kernel (GK), could lead to a quite accurate reconstruction of the error distribution. A proposed Edgeworth Expansion method is employed to reconstruct the error statistic in a much more efficient way with respect to the GK. It is shown that with this method, it is possible to get fundamental error bounds almost as accurate as the theoretical case, i.e. when a priori knowledge of the error distribution is available. Therein, a technique to determine fundamental error limits—CRLB and PEB—onsite without knowledge of the statistics of the ranging errors is proposed.

Copyright © 2020 Institute of Advanced Engineering and Science.

All rights reserved.

---

## Corresponding Author:

Omotayo Oshiga

Nile University of Nigeria, Abuja, Nigeria

Email: ooshiga@nileuniversity.edu.ng

---

## 1. INTRODUCTION

Recently, wireless sensor localization have been widely used for positioning and navigation with various applications in health, transport, environment and other commercial services [1, 2, 3, 4]. As we know, WSNs comprises numerous of wirelessly connected sensors, as a result sensor positioning has become an important problem. The global positioning system (GPS) currently available is expensive, and therein relatively few sensors are equipped with GPS receivers called reference devices, whereas the other sensors are blind-folded devices (nodes). Several methods have been proposed to estimate the positions of sensor nodes in WSN, a problem known as *Node Localization* [5, 6].

Inherently, obtaining the lower bound on location errors in relation to every node is an essential and basic problem within the positioning context of WSN. As a result, the most commonly used tool is the Cramèr-Rao lower bound [7, 8, 9, 10], describing the average mean square error (i.e. the *distance* between the true and estimated node location). Also, it establishes the minimum root mean square error theoretically achievable with an unbiased estimator and it is commonly used as a designing tool, in the sense that it offers a bench mark against which estimation algorithms can be compared with. Another popular tool is the position error bound [11, 12, 13] which illustrates the *confidence region* where a node should be located with a certain confidence interval. It is important to note that both the CRLB and the PEB are obtained from the fisher information matrix. Since they both rely on the knowledge of the distribution of the ranging error, which in turn depends on environmental and technological factors, obtaining their formulation *a priori* is almost impossible, especially in WSNs affected by mainly Non-Line-of-Sight (NLOS).

Mainly, there exists two methods for evaluating the distribution of ranging error measurements—the parametric method which are used for specific and explicit distributions such as Gaussian, Exponential, Rayleigh etc. and non-parametric method are used for all other distributions without explicit expression. The feasible solution is to approximate the distribution statistics of the ranging errors on-site, by collecting ranging samples from each target-anchor link and then estimating the lower bound on the location errors even before target localization. One immediate application of on-site estimation of error statistics is that this can be used to inform cooperative localization algorithms on which nodes to cooperate with to reduce the commulative localization error for any target.

To this end, the well known *maximum likelihood parametric approach* is going to fail, given that in general there is no a priori knowledge on the error distribution. A truly non-parametric approach is therefore required in this case; in particular the *kernel method* is very appreciated for its capability to reconstruct empirical distributions from samples, and in particular its Gaussian kernel (GK) realization. Numerous works have been done on error analyses for wireless localization with most efforts based on Line-of-Sight conditions [14, 15, 16], which lead to severe degradations as NLOS conditions are more appropriate for an accurate wireless localization. Various localization algorithms and performance analyses for NLOS environment have been proposed [15, 16, 17, 18]. The parametric exponential distribution-based CRLB model in [15] can not be used for other parametric distributions to simulate NLOS ranging errors. The CRLB in [16] was derived for NLOS environment using on a single reflection model, and can not be used in a situation where most signals arrive at the receiver after multi-reflections. The CRLB with or without NLOS statistics was derived for NLOS situation in [17]. For the case without NLOS statistics, the authors computed the CRLB in a mixed NLOS/LOS environment and proved that the CRLB for a mixed NLOS/ LOS environment depends only on LOS signals, while for the case with NLOS statistics, the authors only provided a definition of CRLB.

In this article, the GK method utilised to obtain the on-site the statistic of the ranging errors is reproduced and both the CRLB and PEB are then rewritten, along with their performance analysis in various forms. Compared with the previous performance studies for LOS and NLOS conditions, the contributions of this article are as follows:

- a. A mathematical description of the system model and standard error bounds are formulated, which depicted that the ranging model and bounds derived are applicable to any distribution of ranging errors. For easy modelling of NLOS conditions, the nakagami distribution model was used Section 2.
- b. A Gaussian kernel (GK) method was introduced and a mathematical formulation of its lower bounds were obtained to derive the statistical distribution of the errors similar to [18] Section 3. Also, a newly proposed Edgeworth expansion (EE) method was introduced and a mathematical formulation of its lower bounds were obtained to derive the statistical distribution of the errors Section 4.
- c. A thorough and complete analyses of CRLBs and PEBs for the GK and EE methods, which upholdss the proposed EE method by exhibiting that it indeed comes very close in achieving the fundamental lower bound in terms of location error. Its greater efficiency is further proved by the much lower number of samples needed to reach the same level of accuracy as the GK technique Section 5.

## 2. SYSTEM MODEL AND FORMULATION

### 2.1. System model

Consider a network of  $N$  nodes in an  $\eta$ -dimensional Euclidean space, out of which blindfolded devices indexed  $1, \dots, N_t$  have no knowledge of their location (henceforth *targets*), while devices indexed  $N_t + 1, \dots, N_t + N_a$  are *anchors*, i.e. reference devices of *a priori* known location. For the sake of clarity, we shall hereafter scrutinize the case of when  $\eta = 2$ , with the remark that the analysis to follow can be straightforwardly extended to  $\eta > 2$ .

The localization problem consists of estimating the location of target nodes, given the knowledge on the location of anchor nodes, and a set of measures of distances amongst devices typically affected by errors [8]. To elaborate, let the position of the  $i$ -th device be denoted by  $(x_i, y_i)$ , such that the coordinate vector of the target to be approximated is described as

$$\Theta \triangleq [\boldsymbol{\theta}_x, \boldsymbol{\theta}_y] = [x_1, \dots, x_{N_t}, y_1, \dots, y_{N_t}] \quad (1)$$

Likewise, we describe the anchors' coordinate vector by

$$\Phi \triangleq [\phi_x, \phi_y] = [x_{N_t+1}, \dots, x_{N_t+N_a}, y_{N_t+1}, \dots, y_{N_t+N_a}] \quad (2)$$

It is well known that when two nodes are able to exchange information, they are able to estimate the mutual distances between themselves, a process referred to as *ranging*. Consistently, ranging measurements are always affected by noise and often they are not obtained over a LOS link between nodes. In NLOS scenarios, an additional ranging error referred to as *bias* in the form of a positive deviation from the true mutual distance appears. Under these assumptions, the ranging model applicable to a pair of devices  $i$ -th and  $j$ -th is given by

$$\tilde{d}_{ij} = d_{ij} + n_{ij} + b_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} + v_{ij} \quad (3)$$

where  $\tilde{d}_{ij}$  is the measured distance,  $d_{ij}$  is the true distance,  $n_{ij}$  is an additive white Gaussian noise with mean  $\mu = 0$  and variance  $\sigma_{ij}^2$ ,  $b_{ij}$  is the bias, and the *residual noise*  $v_{ij}$  where the noise and bias are modelled jointly.

## 2.2. Standard error bound formulations

Here, the fisher information matrix (FIM)  $\mathbf{J}$  [9] as the fundamental matrix to obtain both the CRLB and PEB are clearly formulated, with the aim of clearly introducing the notations and methods to be employed in the Sections 3. and 4. where the gaussian kernel (GK) [18] and edgeworth expansion (EE) (proposed) [19] error bounds will be formulated and discussed.

Let  $\tilde{\mathbf{d}}$  be the range measurements (measured distances) vector denoted as

$$\tilde{\mathbf{d}} \triangleq \{\tilde{d}_{ij}\} \quad (4)$$

where  $i, j = 1 \dots N$  for  $i \neq j$ . Let  $\hat{\theta}$  be an estimate of the vector parameter  $\theta$  and  $\mathbb{E}[\hat{\theta}]$  as the expected value of  $\hat{\theta}$ . The CRLB matrix relates to the Fisher information matrix  $\mathbf{J}$  [9] as

$$\mathbb{E}[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] \succeq \mathbf{J}^{-1} \quad (5)$$

The Fisher information matrix  $\mathbf{J}$  is accordingly given as

$$\mathbf{J} \triangleq \mathbb{E} \left[ \frac{\partial \ln f(\tilde{\mathbf{d}}|\theta)}{\partial \theta} \left( \frac{\partial \ln f(\tilde{\mathbf{d}}|\theta)}{\partial \theta} \right)^T \right] \quad (6)$$

The log of the joint conditional probability density function (PDF) is

$$\ln f(\tilde{\mathbf{d}}|\theta) = \sum_{i=1}^N \sum_{\substack{j \in H(i) \\ j < i}} l_{ij} \quad (7)$$

where  $l_{ij} = \ln f(\tilde{d}_{ij}|(x_i, y_i, x_j, y_j))$ . Substituting  $l_{ij}$  in (7) and in (6), the FIM is then denoted by [14]

$$\mathbf{J} \triangleq \begin{bmatrix} \mathbf{J}_{xx} & \mathbf{J}_{xy} \\ \mathbf{J}_{xy} & \mathbf{J}_{yy} \end{bmatrix} \quad (8)$$

where

$$[\mathbf{J}_{xx}]_{kl} = \begin{cases} \sum_{j \in H(k)} \mathbb{E} \left[ \left( \frac{\partial l_{kj}}{\partial x_k} \right)^2 \right] \\ e_{kl} \mathbb{E} \left[ \frac{\partial l_{kl}}{\partial x_k} \frac{\partial l_{kl}}{\partial x_l} \right] \end{cases}, \quad [\mathbf{J}_{xy}]_{kl} = \begin{cases} \sum_{j \in H(k)} \mathbb{E} \left[ \frac{\partial l_{kj}}{\partial x_k} \frac{\partial l_{kj}}{\partial y_l} \right] \\ e_{kl} \mathbb{E} \left[ \frac{\partial l_{kl}}{\partial x_k} \frac{\partial l_{kl}}{\partial y_l} \right] \end{cases},$$

$$[\mathbf{J}_{yy}]_{kl} = \begin{cases} \sum_{j \in H(k)} \mathbb{E} \left[ \left( \frac{\partial l_{kj}}{\partial y_k} \right)^2 \right] & k = l \\ e_{kl} \mathbb{E} \left[ \frac{\partial l_{kl}}{\partial y_k} \frac{\partial l_{kl}}{\partial y_l} \right] & k \neq l \end{cases}$$

and  $k, l = 1 \dots n$  are the blindfolded (target) nodes.  $\mathbf{J}_{xx}$ ,  $\mathbf{J}_{yy}$ ,  $\mathbf{J}_{xy}$ , and  $\mathbf{J}$  are of sizes  $n \times n$  and  $2n \times 2n$ , respectively.

### 2.3. Modeling range measurements

The statistics of the measured distances between nodes- adopting the most recognised propagation models in mobile and wireless communication in the literature [21, 22], has been modeled after the nakagami distribution (ND). The nakagami distribution was selected to fit empirical data and is known to provide a closer match to most measurement data than either the Gaussian, Rayleigh or Rician distributions. Beyond its empirical justification, the nakagami distribution is often used for the following reasons. First, the nakagami distribution can model environmental conditions that are either more or less severe than Rayleigh fading. When the nakagami shape factor is 1, the nakagami distribution becomes the Rayleigh distribution, and when the nakagami shape factor is 1/2, it becomes a one-sided Gaussian distribution. Second, the Rice distribution can be closely approximated using the close form relationship between the Rice factor and the nakagami shape factor. Due to the empirical data and work done in [21], the nakagami distribution was chosen to model the NLOS conditions for ranging measurements.

The PDF of the residual noise  $v_{ij}$ , to evaluate the performance of both the gaussian kernel and edge-worth expansion methods, will therein be

$$f_{v_{ij}}(v_{ij}) = \frac{2m_{ij}^{m_{ij}}}{\Gamma(m_{ij})\Omega_{ij}^{m_{ij}}} v_{ij}^{2m_{ij}-1} \exp\left(-\frac{m_{ij}}{\Omega_{ij}} v_{ij}^2\right) \quad (9)$$

where  $m_{ij}$  and  $\Omega_{ij}$  are the shape and controlling spread parameters of the Nakagami distribution.

### 2.4. Bounds derivation using nakagami distributions

Given the obtained ranging model's PDF, it is now attainable to derive a new formula for the FIM. From (9), take its natural logarithm and substitute the result into  $\frac{\partial l_{kl}}{\partial x_k}$ ,  $\frac{\partial l_{kl}}{\partial y_k}$ ,  $\frac{\partial l_{kl}}{\partial x_l}$  and  $\frac{\partial l_{kl}}{\partial y_l}$  yields

$$\begin{aligned} \frac{\partial l_{kl}}{\partial x_k} &= \frac{x_k - x_l}{d_{kl}} \left( \frac{2m_{kl}v_{kl}}{\Omega_{kl}} - \frac{2m_{kl} - 1}{v_{kl}} \right), & \frac{\partial l_{kl}}{\partial y_k} &= \frac{y_k - y_l}{d_{kl}} \left( \frac{2m_{kl}v_{kl}}{\Omega_{kl}} - \frac{2m_{kl} - 1}{v_{kl}} \right), \\ \frac{\partial l_{kl}}{\partial x_l} &= -\frac{x_k - x_l}{d_{kl}} \left( \frac{2m_{kl}v_{kl}}{\Omega_{kl}} - \frac{2m_{kl} - 1}{v_{kl}} \right), & \frac{\partial l_{kl}}{\partial y_l} &= -\frac{y_k - y_l}{d_{kl}} \left( \frac{2m_{kl}v_{kl}}{\Omega_{kl}} - \frac{2m_{kl} - 1}{v_{kl}} \right) \end{aligned} \quad (10)$$

and therefore

$$\begin{aligned} [\mathbf{J}_{xx}]_{kl} &= \begin{cases} \sum_{j \in H(k)} A_{kj} \frac{(x_k - x_j)^2}{d_{kj}^2} \\ e_{kl} A_{kl} \frac{(x_k - x_l)^2}{d_{kl}^2} \end{cases}, & [\mathbf{J}_{xy}]_{kl} &= \begin{cases} \sum_{j \in H(k)} A_{kj} \frac{(x_k - x_j)(y_k - y_j)}{d_{kj}^2} \\ e_{kl} A_{kl} \frac{(x_k - x_l)(y_k - y_l)}{d_{kl}^2} \end{cases}, \\ [\mathbf{J}_{yy}]_{kl} &= \begin{cases} \sum_{j \in H(k)} A_{kj} \frac{(y_k - y_j)^2}{d_{kj}^2} \\ e_{kl} A_{kl} \frac{(y_k - y_l)^2}{d_{kl}^2} \end{cases} \end{aligned}$$

where

$$A_{kl} = \mathbb{E} \left[ \left( \frac{2m_{kl}v_{kl}}{\Omega_{kl}} - \frac{2m_{kl} - 1}{v_{kl}} \right)^2 \right] = \int_{-\infty}^{\infty} \left( \frac{2m_{kl}v_{kl}}{\Omega_{kl}} - \frac{2m_{kl} - 1}{v_{kl}} \right)^2 f_{v_{kl}}(v_{kl}) dv_{kl} \quad (11)$$

### 3. ERROR ESTIMATION VIA GAUSSIAN KERNEL

In [18], using the gaussian kernel (GK) method the error bound formulation was obtained to model the PDF of the positive deviation - bias  $b_{ij}$ . This step was taken to ensure that the white Gaussian noise  $n_{ij}$  and positive bias  $b_{ij}$  in the ranging errors were to be modeled independently. In this article, the work in [18] was well modified and improved upon where the residual noise was modeled independently so as to enable both the noise and bias to be modeled jointly as the residual noise  $v_{ij}$ , as it is well known in the literature the impossibility of separating LOS noise from NLOS bias in a wireless environment.

#### 3.1. Error distribution reconstruction

The PDF of the residual noise  $v_{ij}$  is obtained from samples of ranging measurements. This is an estimation of the true distribution by building a sum of kernels (which are derived from an exponentially decaying function) of the collected ranging samples, whose efficiency and accuracy depends on the total number of collected samples  $P$ . Between the  $i$ -th and  $j$ -th nodes,  $Sv_{ijq}$  is defined as the  $q$ -th sample over the link, the non-parametric Gaussian Kernel technique estimates the PDF of the residual noise as

$$f_{v_{ij}}(v_{ij}) = \frac{1}{\sqrt{2\pi}Ph_{ij}} \sum_{q=1}^P \exp\left(-\frac{(v_{ij} - Sv_{ijq})^2}{2h_{ij}^2}\right) \quad (12)$$

where  $\exp(-)$  is the Gaussian kernel exponential function and the smoothing constant  $h_{ij}$  is the width of this Gaussian kernel function given as  $1.06\sigma_s P^{-1/5}$  ( $\sigma_s$  is the sample standard deviation of the residual noise).

#### 3.2. Bounds derivation using gaussian kernel

Following the same approach as in Subsection 2.4., from (12), the natural logarithm can be substituted into  $\frac{\partial l_{kl}}{\partial x_k}, \frac{\partial l_{kl}}{\partial y_k}, \frac{\partial l_{kl}}{\partial x_l}$  and  $\frac{\partial l_{kl}}{\partial y_l}$  obtaining

$$\begin{aligned} \frac{\partial l_{kl}}{\partial x_k} &= \frac{x_k - x_l}{d_{kl}} \frac{g_{kl}(v_{kl})}{f_{v_{kl}}(v_{kl})}, \quad \frac{\partial l_{kl}}{\partial y_k} = \frac{y_k - y_l}{d_{kl}} \frac{g_{kl}(v_{ij})}{f_{v_{kl}}(v_{kl})}, \\ \frac{\partial l_{kl}}{\partial x_l} &= -\frac{x_k - x_l}{d_{kl}} \frac{g_{kl}(v_{ij})}{f_{v_{kl}}(v_{kl})}, \quad \frac{\partial l_{kl}}{\partial y_l} = -\frac{y_k - y_l}{d_{kl}} \frac{g_{kl}(v_{kl})}{f_{v_{kl}}(v_{kl})} \end{aligned} \quad (13)$$

where

$$g_{kl}(v_{kl}) = \frac{1}{\sqrt{2\pi}Ph_{ij}} \sum_{t=1}^P \exp\left(-\frac{(v_{kl} - Sb_{klt})^2}{2h_{kl}^2}\right) \frac{v_{kl} - Sb_{klt}}{h_{kl}^2} \quad (14)$$

and the elements of the Fisher Information are similar with (11) except for the coefficient:

$$A_{kl} = \mathbb{E} \left[ \left( \frac{g_{kl}(v_{kl})}{f_{v_{ij}}(v_{kl})} \right)^2 \right] = \int_{-\infty}^{\infty} \frac{g_{kl}(v_{kl})^2}{f_{v_{kl}}(v_{kl})} dv_{kl} = \frac{1}{\sigma_{kl}^2} \quad kl \in \text{LOS} \quad \int_{-\infty}^{\infty} \frac{g_{kl}(v_{kl})^2}{f_{v_{kl}}(v_{kl})} dv_{kl}. \quad k, l \in \text{NLOS} \quad (15)$$

where  $kl \in \text{LOS}$  and  $kl \in \text{NLOS}$  represent propagation conditions between nodes  $k$  and  $l$ .

### 4. ERROR ESTIMATION VIA EDGEWORTH EXPANSION

The ranging error approximation technique presented in the previous section, though robust, is constrained by the enormous amount of samples required to obtain a fair accuracy of the approximates of the distribution of a given set of samples. In the following, we introduce a more efficient and general method, based on Edgeworth expansion, with two main advantages: a much smaller number of samples are required for approximation and the possibility to model both the additive Gaussian noise and the positive bias jointly. While the prospect of reducing the number of samples required to obtain a fair accuracy can not be overemphasized, it is essential to state that, in wireless channels, the positive bias and Gaussian ranging errors cannot

be separated from each other. Therein, we describe the process of reconstructing the ranging error distribution from samples, and then the convergence and monotonicity of moments from samples is shown, thereby proving a clear improvement in accuracy with respect to the Gaussian kernel technique, and finally the proposed formulation of PEB and CRLB are shown.

#### 4.1. Error distribution reconstruction

The Edgeworth expansion which is an improved version on the central limit theorem (CLT) is a true asymptotic expansion of the PDF of a gaussian variable  $\hat{x} = (x - \mu)/\sigma$  in the powers of the mean  $\mu$ . EE is a formal series of functions that has the characteristics of truncating a series after a finite number of terms, which is sufficient enough to provide an accurate estimation to this function, therein the estimation error is monitored [19].

The EE as a non-parametric approximator can be used for estimating the PDF of given ranging errors from their sample moments  $\alpha_w$  [19]. The EE is given as

$$f(x) = \mathcal{N}(\mu, \sigma^2) \left[ 1 + \sum_{s=1}^{\infty} \sigma^s \sum_{\{k_w\}} He_{s+2r}(\hat{x}) \prod_{w=1}^s \frac{1}{k_w!} \left( \frac{S_{w+2}}{(w+2)!} \right)^{k_w} \right] \quad (16)$$

where,

$\mathcal{N}(\mu, \sigma^2)$  is the PDF of a normal distribution with mean  $\mu$  and variance  $\sigma^2$ ,  $S_{w+2} = \frac{\kappa_{w+2}}{\kappa_2^{w+1}}$ ,  $\kappa_w$  are the cumulants obtained from the sample moments  $\alpha_w$  as

$$\kappa_s = s! \sum_{\{k_w\}} (-1)^{(r-1)} (r-1)! \prod_{w=1}^s \frac{1}{k_w!} \left( \frac{\alpha_w}{w!} \right)^{k_w} \quad (17)$$

The set  $\{k_w\}$  consists of all non-negative (positive and zero) integer solutions of the Diophantine set of equations  $s = k_1 + 2k_2 + \dots + sk_s$  and  $r = k_1 + k_2 + \dots + k_s$ . The Chebyshev-Hermite polynomial  $He_n(\hat{x})$  is

$$He_s(\hat{x}) = s! \sum_{k=0}^{s/2} \frac{(-1)^k \hat{x}^{s-2k}}{k!(s-2k)!2^k} \quad (18)$$

and the mean and variance of the ranging errors are  $\mu = \alpha_1$  and  $\sigma^2 = \kappa_2$ , respectively [20, 21, 22].

The sample moments from the ranging errors are  $\alpha_w = 1/n \sum_{i=1}^n X_i^w$ , where  $X_i$  are the ranging errors and  $w = 1, 2, 3, \dots$  are the orders of the moment. To determine the number of orders of moment  $\alpha_w$  required to estimate a given sample, the standard error  $s$  of the samples is calculated using  $\sigma_s^2/\sqrt{P}$ , where  $s$  must be  $\leq 0.3$ , for each order  $w$ . The Edgeworth Expansion is used to model the residual noise  $v_{ij}$ , hence, the estimated PDF of the residual noise  $f_{v_{ij}}(v_{ij})$  is

$$f_{v_{ij}}(v_{ij}) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left(-\frac{(v_{ij} - \mu)^2}{2\sigma_{ij}^2}\right) \left( 1 + \sum_{s=1}^{\infty} \sigma^s \sum_{\{k_w\}} A_s He_{s+2r}(\hat{x}) \right) \quad (19)$$

where  $A_s = \prod_{w=1}^s \frac{1}{k_w!} \left( \frac{S_{w+2}}{(w+2)!} \right)^{k_w}$  and  $v_{ij} = \tilde{d}_{ij} - d_{ij}$ .

#### 4.2. Efficiency and convergence of sample moments of the EE method

To illustrate the effectiveness and efficiency of the Edgeworth Method, it is mandatory to demonstrate the convergence of its sample moments as the number of samples  $P$  increases [23, 24], and therein compare it with the Gaussian Kernel. Using the Nakagami Distributed random variables as seen in Figure 1, the true moments  $\gamma_w$  of a ND ( $m = 1, \Omega = 1$ ) is compared with sample moments  $\alpha_w$  [23] for different number of samples and moment orders  $w = 1, \dots, 4$ . The deviation  $\hat{e}_w$  of the sample moments from the true moment

is obtained as the absolute of  $(\alpha_w - \gamma_w)/\gamma_w$ . Also, the monotonicity of the sample moments is seen in Figure 2. The Kulback Leiber Divergence of the PDFs obtained using the two methods (GK and EE) from the true/theoretical Nakagami PDF is shown in Figure 3. For KLD=0.01, the proposed EE required less than 300 samples, while the GK needs approximately 500 samples to obtain similar results. This delta increases even more for lower level of divergency: to reach a KLD=0.0075, the GK method needs approximately twice the number of samples than the EE method. As a result, the Edgeworth method is a good choice for approximation the distribution of ranging errors from samples.

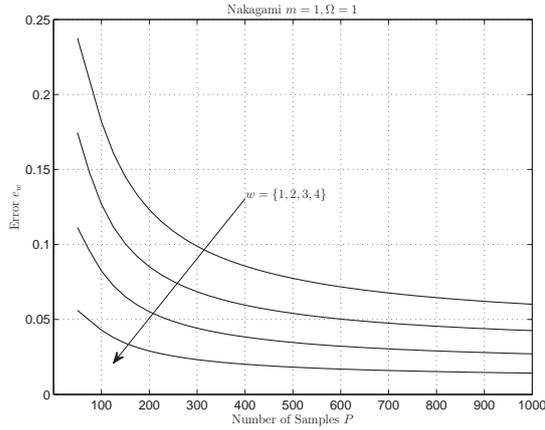


Figure 1. Convergence of sample moments  $\alpha_w$

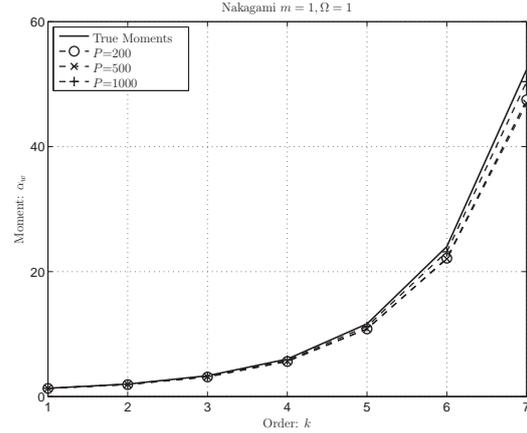


Figure 2. Monotonicity of sample moments  $\alpha_w$

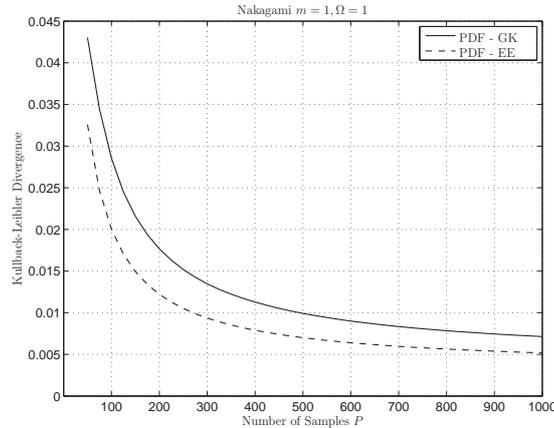


Figure 3. CKL Divergence of the two estimators

#### 4.3. Bounds derivation using edgeworth expansion

From the formulation of the approximated PDF of the residual error in (19), the creation of the Fisher information is the same with the GK: This yields

$$\begin{aligned} \frac{\partial l_{kl}}{\partial x_k} &= \frac{x_k - x_l}{d_{kl}} \left( \frac{v_{kl} - \mu}{\sigma_{kl}^2} - \frac{g_{kl}(v_{kl})}{\sigma_{kl} f_{v_{kl}}(v_{kl})} \right), \quad \frac{\partial l_{kl}}{\partial y_k} = \frac{y_k - y_l}{d_{kl}} \left( \frac{v_{kl} - \mu}{\sigma_{kl}^2} - \frac{g_{kl}(v_{kl})}{\sigma_{kl} f_{v_{kl}}(v_{kl})} \right), \\ \frac{\partial l_{kl}}{\partial x_l} &= -\frac{x_k - x_l}{d_{kl}} \left( \frac{v_{kl} - \mu}{\sigma_{kl}^2} - \frac{g_{kl}(v_{kl})}{\sigma_{kl} f_{v_{kl}}(v_{kl})} \right), \quad \frac{\partial l_{kl}}{\partial y_l} = -\frac{y_k - y_l}{d_{kl}} \left( \frac{v_{kl} - \mu}{\sigma_{kl}^2} - \frac{g_{kl}(v_{kl})}{\sigma_{kl} f_{v_{kl}}(v_{kl})} \right) \end{aligned} \quad (20)$$

where

$$g_{kl}(v_{kl}) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left(-\frac{(v_{ij} - \mu)^2}{2\sigma_{ij}^2}\right) \sum_{s=1}^{\infty} \sigma^s \sum_{\{k_m\}} (s + 2r) A_s H_{s+2r-1}(\hat{x}) \quad (21)$$

The elements of the Fisher Information are similar with the GK except for the coefficient:

$$A_{kl} = \mathbb{E} \left[ \left( \frac{v_{kl} - \mu}{\sigma_{kl}^2} - \frac{g_{kl}(v_{kl})}{\sigma_{kl} f_{v_{kl}}(v_{kl})} \right)^2 \right] = \int_{-\infty}^{\infty} \left( \frac{v_{kl} - \mu}{\sigma_{kl}^2} - \frac{g_{kl}(v_{kl})}{\sigma_{kl} f_{v_{kl}}(v_{kl})} \right)^2 f_{v_{kl}}(v_{kl}) dv_{kl} \quad (22)$$

## 5. PERFORMANCE EVALUATION

Moving forward from the various theoretical analyses presented in this paper, we can state that the EE methods can approximate the statistics of ranging errors (using Nakagami Distribution), with lesser samples and more accuracy with respect to the GK as such we now consider real network topologies to further illustrate the performances of both methods. Therefore, a region of  $10m \times 10m$  is employed, where three ( $a_n = 3$ ) anchors are placed to form a triangular shape and three ( $n = 3$ ) blindfolded devices (targets), not connected together, are randomly placed within the convex of the anchors. The two error bounds - the CRLB and the PEB - will be utilized to evaluate the performance of the two estimators - EE and GK. The average CRLB for any network topology can be computed using

$$\bar{\varepsilon} = \frac{1}{n} \{ \mathbf{J}^{-1} \} \quad (23)$$

while the PEB can be illustrated by the 95% Confidence Interval  $C_i = 0.95$ , whose mathematical formulation is shown. The Fisher Ellipse parameters of the  $i$ -th target  $\theta_i$  are estimated from the covariance matrix  $\mathbf{\Omega}_{\theta_i}$ , which is a combination of the error variance  $\sigma_{i:x}^2$  and  $\sigma_{i:y}^2$  on the “x” and “y” dimensions, respectively and the cross-term  $\sigma_{i:xy}$ , given as

$$\mathbf{\Omega}_{\theta_i} \triangleq \begin{bmatrix} \sigma_{i:x}^2 & \sigma_{i:xy} \\ \sigma_{i:xy} & \sigma_{i:y}^2 \end{bmatrix} \quad (24)$$

The directions of the scattering in the space for the vector  $\theta_i$  are known to be directly proportional to the eigenvalues associated to  $\mathbf{\Omega}_{\theta_i}$  up to a factor of  $\kappa_i$  [25, 11, 12]. In particular, the axis direction of the ellipse which describes this scattering in the space is  $2\sqrt{\kappa_i \lambda_{i:1}}$ ,  $2\sqrt{\kappa_i \lambda_{i:2}}$ , where

$$\lambda_{i:1} \triangleq \frac{1}{2} \left[ \sigma_{i:x}^2 + \sigma_{i:y}^2 + \sqrt{(\sigma_{i:x}^2 - \sigma_{i:y}^2)^2 + 4\sigma_{i:xy}^2} \right], \quad \lambda_{i:2} \triangleq \frac{1}{2} \left[ \sigma_{i:x}^2 + \sigma_{i:y}^2 - \sqrt{(\sigma_{i:x}^2 - \sigma_{i:y}^2)^2 + 4\sigma_{i:xy}^2} \right] \quad (25)$$

If  $\sigma_{i:y} > \sigma_{i:x}$ , then in (25) the orders of  $\lambda_{i:1}$  and  $\lambda_{i:2}$  are swapped. The proportionality factor  $\kappa_i$  can be related to the confidence interval  $C_i$  in that the target  $\theta_i$  is enclosed in an ellipse, as such  $\kappa_i = -2 \ln(1 - C_i)$

It follows that the Fisher Ellipse for the  $i$ -th target  $\theta_i$  is described through the following in [25]

$$\frac{[(x - p_{i:x}) \cos \gamma_i + (y - p_{i:y}) \sin \gamma_i]^2}{\kappa_i \cdot \lambda_{i:1}} + \frac{[(x - p_{i:x}) \sin \gamma_i - (y - p_{i:y}) \cos \gamma_i]^2}{\kappa_i \cdot \lambda_{i:2}} = 1 \quad (26)$$

where the *rotation angle*  $\gamma_i$  describes the offset between the principal axis for the ellipse and reference axis and it is defined as  $\gamma_i \triangleq \frac{1}{2} \arctan \left( \frac{2\sigma_{i:xy}}{\sigma_{i:x}^2 - \sigma_{i:y}^2} \right)$  - Note that for  $\sigma_{i:x}^2 = \sigma_{i:y}^2$ , then  $\gamma_i = 0$ .

Figure 4 depicts the level of accuracy and efficiency of the approximated CRLBs  $\bar{\varepsilon}$ , as a function of the sample number  $P$  using Nakagami distributed random variables with  $b_{i,j}$  uniformly selected between 0 - 4 for NLOS and  $\sigma = 0.5$  for both LOS and NLOS ranging errors within each target-to-anchor links.

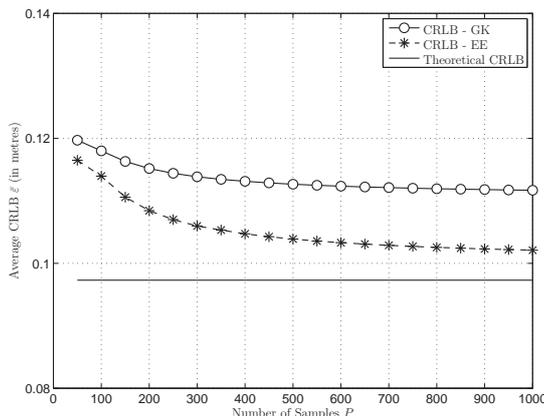


Figure 4. Average CRLB as a function of samples

To compare both estimators, a line representing the theoretical CRLB, i.e. computed with perfect knowledge of the statistic of the propagation channel has been added to the plots. Clearly the CRLB of reconstructed EE is much closer to the theoretical one than the GK: the EE performs better for any samples. From the above derived CRLBs, the minimum number of samples  $P$  required for obtaining accurate results are analyzed. As shown in Figure 3, it is seen that the non-parametric estimators converges to the true PDF with sufficient sample size, therefore the estimated CRLBs converges quickly to a stable value as  $P$  increases.

Furthermore, the two estimators are now represented by their respective Fisher ellipses (theoretical and reconstructed from the two methods) for  $P = 50$  samples as seen in Figure 5. Clearly, the samples reconstructed with the EE estimator almost perfectly match the theoretical one, where they vary only in the axis orientation. As the sample number increases to  $P = 250$  samples, the Fisher ellipses of the two estimators have almost or matching axis orientation to the theoretical PEB with the EE method much closer than the GK method. To clearly and better capture the differences between the two estimators with respect to the theoretical

PEB, Figure 6 depicts, as a function of the number of samples, the inner product  $PEB \Delta \triangleq \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\langle \hat{A}_i \cdot A_i \rangle}{\sqrt{\hat{A}_i A_i}}$ ,

where  $\langle \cdot \rangle$  denotes the inner product,  $A$  is the area of the theoretical Fisher ellipse and  $\hat{A}$  is the area of the reconstructed methods (EE or GK method). From the discussion in this section, it can be clearly seen that Edgeworth Expansion method performs far better than the Gaussian Kernel method, which therein implies that the approximated Fisher Ellipses of Edgeworth Expansion are much closer in size and orientation to the theoretical Fisher Ellipses.

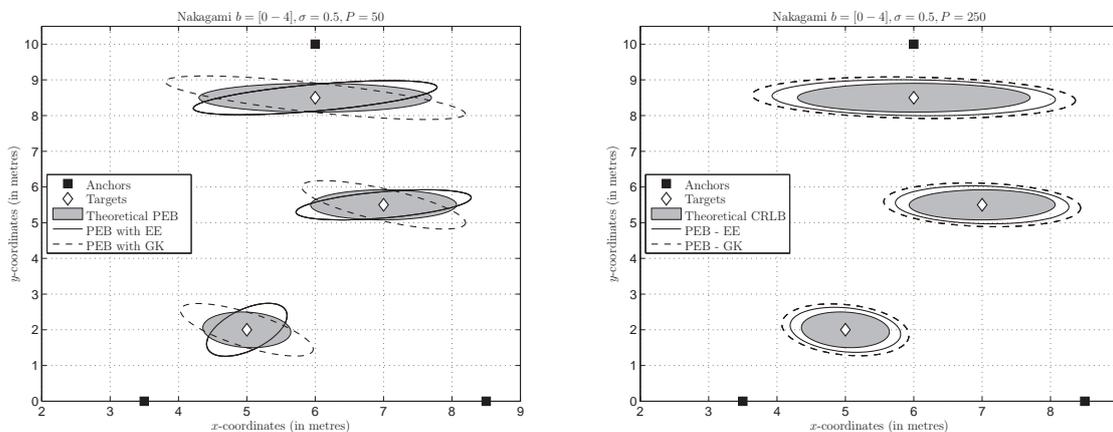


Figure 5. The 95% Fisher ellipses, theoretical, and estimated with  $P = 50$  &  $250$  samples collected per link

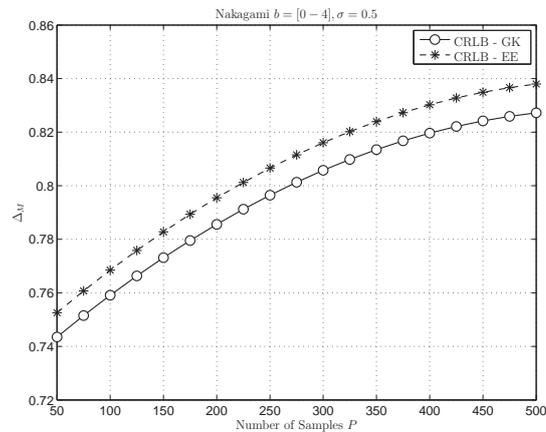


Figure 6.  $\Delta$  as a function of the number of samples

## 6. CONCLUSIONS

This article clearly focuses on the error analyses of approximating and reconstructing the statistics of the ranging measurements without a priori knowledge of the wireless channel. A popular and non-parametric estimator, the Gaussian kernel was first described and utilized for the approximation and reconstruction of the error distributions from samples and the corresponding error bound was derived. Furthermore, an Edgeworth Expansion method was therein, to reconstruct the error distribution statistics from samples of the ranging measurements by exploiting the efficiency and effectiveness of moment convergence. This approach was clearly shown and proven to be valid for Non-Line-of-Sight conditions, where it is impossible to estimate the statistics of the ranging errors a priori. Results and figures illustrated showed that the Edgeworth expansion technique is a far more efficient and accurate technique than the Gaussian kernel method, requiring lesser sample size to reach the similar level accuracy.

## REFERENCES

- [1] D. Macagnano, G. Destino, and G. Abreu, "A comprehensive tutorial on localization: algorithms and performance analysis tools," *International Journal of Wireless Information Networks*, vol. 19, no. 4, pp. 290-314, 2012.
- [2] G. Yanying, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 13-32, 2009.
- [3] L. Hui, H. Darabi, P. Banerjee, and L. Jing, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1067-1080, 2007.
- [4] N. Patwari, J. Ash, S. Kyperountas, et al, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54-69, 2005.
- [5] K.-T. Feng, C.-L. Chen, and C.-H. Chen, "Gale: An enhanced geometry-assisted location estimation algorithm for nlos environments," *IEEE Trans. on Mobile Computing*, vol. 7, no. 2, pp. 199-213, 2008.
- [6] Y. Qi and H. Kobayashi, "On geolocation accuracy with prior information in non-line-of-sight environment," *IEEE Proceedings on the 56th Vehicular Technology Conference-Fall*, vol. 1, pp. 285-288, 2002.
- [7] H. V. Poor, "An introduction to signal detection and estimation 2nd Ed," *Springer-Verlag New York*, 1994.
- [8] H. Van Trees, "Detection, Estimation, and Modulation Theory," *Wiley*, 2004.
- [9] S. Kay, "Fundamentals Statistical Processg," *Prentice Hall*, 2001.
- [10] T. Wang, "Cramer-rao bound for localization with a priori knowledge on biased range measurements," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 48, no. 1, pp. 468-476, 2012.
- [11] S. Severi, G. Abreu, G. Destino, and D. Dardari, "Efficient and Accurate Localization in Multihop Networks," *Proc. Asilomar Conf. Signals, Systems, and Computers*, 2009.
- [12] J. Saloranta, S. Severi, D. Macagnano, and G. Abreu, "Algebraic confidence for sensor localization," *Proc. Asilomar Conf. Signals, Systems, and Computers*, 2012.

- [13] Eric Weisstein, "Nakagami distribution," [Online], Available: <http://mathworld.wolfram.com/NakaDistribution.html>, 2020.
- [14] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137-2148, 2003.
- [15] H. Miao, K. Yu, and M. J. Juntti, "Positioning for nlos propagation: Algorithm derivations and cramer-rao bounds," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 2568-2580, 2007.
- [16] P. H. Tseng and K. T. Feng, "Geometry-assisted localization algorithms for wireless networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 4, pp. 774-789, 2013.
- [17] J. Prieto, A. Bahillo, et al, "Nlos mitigation based on range estimation error characterization in an rtt-based IEEE 802.11, indoor location," *IEEE International Symposium on Intelligent Signal Processing*, pp. 61-66, 2009.
- [18] J. Huang and Q. Wan, "The crlb for wsns location estimation in nlos environments," *International Conference on Communications, Circuits and Systems*, pp. 83-86, 2010.
- [19] S. Blinnikov and R. Moessner, "Expansions for nearly gaussian distributions," *Astronomy and Astrophysics Supplement Series*, vol. 130, pp. 193-205, 1998.
- [20] D. Jourdan, J. Deyst, J.J., M. Win, and N. Roy, "Monte carlo localization in dense multipath environments using uwb ranging," *IEEE International Conference on Ultra-Wideband*, pp. 314-319, 2005.
- [21] G.L. Stüber, "Principles of Mobile Communication," *Springer Publishing Company, Incorporated*, 2011.
- [22] J. D. Parson, "The Mobile Radio Propagation Channel," *Wiley*, 2000.
- [23] C. Bourin and P. Bondon, "Efficiency of high-order moment estimates," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 255-258, 1998.
- [24] G. de Abreu, "Cth16-4: Accurate simulation of piecewise continuous arbitrary nakagami-m phasor processes," *IEEE Global Telecommunications Conference*, pp. 1-6, 2006.
- [25] D. Torrieri, "Statistical theory of passive location systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 2, pp. 183-198, 1984.

## BIOGRAPHIES OF AUTHORS



**Omotayo Oshiga** received his B.Tech degree in Physics and Telecommunications from the Federal University of Technology, Minna, Nigeria in 2010 and the M.Sc degree in Wireless Communication Systems from Brunel University, London, UK in 2011. He received his Ph.D degree in Electrical Engineering from Jacobs University, Bremen, Germany in February, 2015. He is currently a Senior Lecturer in the department of Electrical/Electronics Engineering at the Nile University of Nigeria, in Abuja, Nigeria. His research interest include various topics in wireless communication and applied mathematics such as estimation theory, signal processing, wireless communication systems, positioning, tracking and navigation, wireless sensing etc.



**Ali Nyangwarimam Obadiah** received his Bachelor of Engineering (B.Eng.) in Electrical and Computer Engineering from the Federal university of Technology Minna, Niger state, Nigeria in 2010 and a Master of Engineering (M.Eng.) in Electrical-Electronics and Telecommunication Engineering from the University of Technology Malaysia in 2014. He received a PhD in Electrical Engineering from University of Technology Malaysia in 2018. He was with the Advanced RF and Microwave Research Group. He is now a Lecturer with the Department of Computer Engineering, Faculty of Engineering, Nile University of Nigeria, Abuja, Nigeria. His research interest includes Radio frequency devices, reconfigurable antennas, filters and filter-antennas.