WLS-Based Self-Localization Using Perturbed Anchor Positions and RSSI Measurements

Vikram Kumar, Reza Arablouei, Brano Kusy, Raja Jurdak, and Neil W. Bergmann

Abstract—We consider the problem of self-localization by a resource-constrained node within a network given radio signal strength indicator (RSSI) measurements from a set of anchor nodes where the RSSI measurements as well as the anchor position information are subject to perturbation. In order to achieve a computationally efficient estimate for the unknown position, we minimize a weighted sum-square-distance-error cost function in an iterative fashion utilizing the gradient-descent method. We calculate the weights in the cost function by taking into account perturbations in both RSSI measurements and anchor node position information while assuming normal distribution for the perturbations in the anchor node position information and log-normal distribution for the RSSI-induced distance estimates. The latter assumption is due to considering the log-distance path-loss model with normally-distributed perturbations for the RSSI measurements in the logarithmic scale. We also derive the Cramer-Rao lower bound associated with the considered position estimation problem. We evaluate the performance of the proposed algorithm considering various arbitrary network topologies and compare it with an existing algorithm that is based on a similar approach but only accounts for perturbations in the RSSI measurements. The experimental results show that the proposed algorithm yields significant improvement in localization performance over the existing algorithm while maintaining its computational efficiency. This makes the proposed algorithm suitable for real-world applications where the information available about the positions of anchor nodes often suffer from uncertainty due to observational noise or error and the computational and energy resources of mobile nodes are limited, prohibiting the use of more sophisticated techniques such as those based on semidefinite or second-order cone programming.

Index Terms—Cramer-Rao lower bound, radio signal strength indicator, self-localization, multilateration, weighted least-squares, wireless sensor networks.

I. INTRODUCTION

Location information is important in most sensor network applications such as environmental, wildlife, or mobile asset monitoring. The global positioning system (GPS) has revolutionized location-based services in the past few decades offering an accuracy range of 1-50m with consumer-grade GPS receivers in outdoor environments [1]. However, the GPS is a relatively power-hungry technology and localization using the GPS is not practical in resource-constrained scenarios such as those involving wireless sensors networks (WSNs) with battery-operated mobile sensors [2], [3]. Relatively high cost and size of the currently available GPS receivers is another hindrance to this technology in becoming the primary choice of localization in most resource-constrained applications.

Cooperative localization based on radio signal strength indicator (RSSI) measurements is an effective method for localizing groups of battery-operated mobile sensors [4]–[6]. In cooperative localization, any node interested in estimating its own position, referred to as the “blind node”, receives position and distance information from its neighboring nodes, called the “anchor nodes”, and estimates its position through multilateration. Such localization techniques can be energy/cost efficient by minimizing the use of expensive localization methods such as the GPS. The localization performance of these techniques largely depends on the accuracy of the RSSI measurements and the anchor node position information. In practice, the RSSI measurements are subject to perturbations arising from model inaccuracy, thermal noise, measurement error, etc. The anchor node position information is also corrupted by noise or error, especially when it is the product of a previous estimation process including the GPS.

Substantial research effort has been dedicated to develop localization methods that can compensate for the adverse effects of perturbations in the RSSI measurements [7]–[12] and the anchor node position information [13]–[22]. Most of the proposed methods that can cope with perturbations in both RSSI measurements and anchor node position information rely on optimization techniques such as semidefinite programming (SDP) or second-order cone programming (SOCP). Therefore, high computational complexity of the SDP and SOCP often makes them unsuitable for localization in resource-constrained applications.

In this paper, we propose a computationally efficient algorithm to localize a blind node when perturbations are present in both RSSI measurements and anchor node position information. Our algorithm is of weighted least-squares (WLS) type as it is based on minimizing a weighted sum-square-error cost function using the gradient-descent method. Each term in the cost function is the weighted square of the difference between the distance inferred from an RSSI measurement and the Euclidean distance between the blind node and the anchor node to which the RSSI measurement corresponds. We weight the square-error terms by their variances, which are estimated by taking into account perturbations in both RSSI measurements and anchor node position information. We assume normal distribution for the perturbations in the available information of the anchor node positions. In addition, we adopt the well-known log-normal shadowing path-loss model. Therefore, we assume that the RSSI measurements in the logarithmic (dBm) scale are affected by normally-distributed...
perturbation and consequently the distance estimates inferred from the RSSI measurements have log-normal distribution. We derive the Cramer-Rao lower bound (CRLB) on the variance of any unbiased estimator of the blind node position in the considered problem with the assumed realistic distributions for the perturbations. We carry out numerical simulations with various geometrical arrangements of the blind node and the anchor nodes in arbitrary network topologies to evaluate the performance of the proposed algorithm in comparison with an existing WLS-based algorithm \cite{10} that only takes into account the perturbations in the RSSI measurements and has similar computational requirements. The simulation results show that the proposed algorithm offers significant improvement in the localization performance over the algorithm of \cite{10}.

Our work is motivated by several use-case scenarios relevant to localization in emerging cooperative wireless networks of mobile nodes such as robots, vehicles, and equipment or wildlife trackers deployed in large outdoor areas \cite{23, 24, 25, 26}. Typically, such networks are very large and contain a mixture of GPS-equipped and non-GPS-equipped nodes. GPS-equipped nodes generally act as anchor nodes for the rest of the network. However, poor performance of the GPS in dense forests, urban areas, or even with a cloudy sky results in inaccuracies/uncertainties in the form of perturbations in the anchor node position information.

The remainder of the paper is organized as follows. After describing the notations used in this paper, in Section II, we review the existing works on the problem of self-localization in WSNs and highlight their differences with our work. For the sake of completeness, we also cover some works on the closely-related problem of source/emitter localization in WSNs. In Sections III and IV, we provide a formal statement of the considered problem and the details of the proposed algorithm, respectively. In Section V, we calculate the CRLB. We explain the experimental setup used for performance evaluation of the proposed algorithm in Section VI and provide the results of the experiments in Section VII before concluding the paper in Section VIII.

**Notations:** The symbol $\mathbb{N}^+$ denotes the set of positive integers and $\mathbb{R}$ denotes the set of real numbers. A lower-case letter, e.g., $x$, represents a scalar variable; a lower-case bold letter, e.g., $\mathbf{x}$, represents a vector; and an upper-case bold letter, e.g., $\mathbf{F}$, represents a matrix. The superscript $(\cdot)^T$ denotes the vector/matrix transpose and $(\cdot)^{-1}$ denotes the matrix inverse. A letter with a tilde accent, e.g., $\tilde{x}$, represents the noisy observation of the original variable, $x_i$; a letter with a hat accent, e.g., $\hat{x}$, represents an estimated value; and a letter with an overbar, e.g., $\bar{c}$, represents an approximate value. The operator $\frac{\partial}{\partial x}$ is the partial derivative with respect to the variable $x$, $\text{Cov}(\cdot)$ returns the covariance matrix of its vector argument, $\mathbb{E}[\cdot]$ is the expectation operator, and $\text{Tr}\{\cdot\}$ is the matrix trace operator. The symbol $\mathcal{N}(\mu, \sigma)$ represents the normal (Gaussian) distribution with mean $\mu$ and standard deviation $\sigma$.

## II. RELATED WORK

We divide the related existing work into the following categories:

- **self-localization by a static blind node (single/multiple):** self-localization by a single or multiple static blind nodes in a sensor network with computation load at the blind node itself;
- **target localization (single/multiple):** localization of a single or multiple mobile/static blind nodes by a group of nodes with processing load either on a central facility or distributed among the group of nodes interested in the target localization;
- **joint multiple blind and anchor node localization:** problem of refining the anchor position information in addition to the blind node localization with distributed or central processing;
- **self-localization by a mobile node:** self-localization by a mobile node using information from other mobile nodes within its radio communication range and computation load only at the blind node.

The first category “self-localization by a static blind node” relates to the static node localization in sensor networks such as those used in agriculture and environmental monitoring \cite{27, 28}. Generally, such networks contain a small number of nodes with independent positioning sensors such as GPS receivers. The rest of the nodes use some cooperative localization techniques to estimate their positions. An iterative multilateration technique in which three or more nodes compute their positions and then act as anchors in the subsequent iterations is given in \cite{29}. In \cite{11}, a multidimensional scaling technique that uses node connectivity information for localization is proposed. It presents two methods, one builds a global map directly and the other builds local maps then stitches them together to build a global map. The authors in \cite{30} present a convex position estimation technique in WSN based exclusively on connectivity-induced constraints processed in a centralized resource center. They model the known peer-to-peer communication constraints as a set of geometric constraints on the node position and solve the problem using SDP.

In \cite{15}, a range-based positioning technique is proposed that has a computational complexity lower than the least-square method. Their algorithm is based on the linearized range measurement equations and implementation of a WLS criterion in a computationally efficient way. Similarly, there are other works \cite{7, 14} that focus on providing low-complexity solutions. The work of \cite{13} implements a robust joint localization and time synchronization in WSN with bounded anchor position uncertainties. It implements a robust joint estimator based on minimizing the worst-case mean square error and the solution is obtained by solving a SDP problem. Similarly, there are other RSSI-based node localization methods such as \cite{31, 32, 33}. However, none of these works consider the problem of localizing blind nodes when perturbations are present in both RSSI and anchor position information.

The second category “target localization” has received considerable attention due to its wider application domain...
including target tracking, habitat monitoring, and military tracking [34]–[37]. In [12], the authors proposed an SDP relaxation-based method for the localization in ad-hoc WSNs. The approach is to convert the non-convex quadratic distance constraints into linear constraints by introducing a relaxation to remove the quadratic term from the constraint. In contrast to our problem, it assumes the availability of accurate anchor position information. A closed form solution and corresponding CRLB for localizing a stationary source/Emitter based on time difference of arrival (TDOA) measurements is presented in [22]. The authors assume normally-distributed error in both anchor position information and TDOA measurements. In contrast, we consider RSSI measurements modeled using a log-normal shadowing path loss model. An SOCP-based approach for sensor network localization with anchor position uncertainty is given in [18]. It presents a robust localization approach using maximum-likelihood criteria under an unbounded uncertainty model for the anchor position error. A distributed multidimensional scaling approach for localization in WSNs is proposed in [38]. It weights the range measurements based on their accuracy to account for the communication constraints in a sensor network.

The focus of the third category “joint multiple blind and anchor node localization” is on refining the anchor position information in addition to localizing a blind node. In [19], the authors propose a two-step distributed sensor network localization approach using SCOP. In the first step, the blind nodes compute their locations using available anchor node position information. In the second step, anchor nodes refine their position estimates. Node localization for the underwater and underground networks is considered in [16]. The authors propose an SDP-based localization algorithm in the presence of anchor position error and uncertainties in the signal propagation speed affecting the time of arrival (TOA)-based distance measurement. They also derive the CRLB for the corresponding problem.

In [39]–[41], the authors propose an RSSI-based solution for joint location estimation of a source and multiple anchor nodes. They also provide the theoretical bounds for their solution as well as an interpretation of the theoretical bounds. The nature of the perturbations considered in [39]–[41] is the same as what we consider in this paper. However, their algorithm requires multiple RSSI and anchor position measurements. Nonetheless, our work is focused on energy-efficient outdoor localization of mobile nodes where the GPS is the main source of anchor position information while being responsible for the majority of energy consumed. Therefore, the requirement of multiple measurements of a anchor position is not in line with our goal of energy/resource-efficient localization. In a nutshell, this category is focused on Joint blind node-anchor node localization using complex optimization techniques such as SOCP and SDP, generally considered unsuitable for energy- and resource-constrained scenarios.

The last category is “self-localization by a mobile node”. The example application cases for this category are proximity services, mobile phone users tracking, and animal tracking with standalone tracking devices [42]–[44]. This category has the following features:

- Unlike in target localization, the anchor nodes are not required to be in the communication range of each other.
- Limited resources highlight the need to avoid complex algorithms in performing multi-source or joint localization.
- The uncontrolled mobile nature of the nodes leads to temporary grouping behavior. The grouping interval sometimes may not be enough to receive or share the results of joint localization.

In [45], the authors proposed a localization technique using RSSI based distance from neighboring nodes. They proposed to augment GPS-based positioning with more energy-efficient location sensors to bound position estimate uncertainty while the GPS is off. They used RSSI based distance estimation from the nodes with better position estimate as a measure to reduce self-position uncertainty. In [46], the authors explored the concept of collaborative GPS duty cycling using Wi-Fi ad-hoc connectivity while ensuring application specific error bounds. Whenever a node approaches the error bound, first it requests a better position estimate from its neighborhood and wait for a response within a certain time. In the case of a positive response, the node will update its position estimate. Otherwise, it will go for its own GPS lock followed by a broadcast of new GPS position coordinates.

In [45], [46], the authors present an algorithm for self-localization in resource-constrained environments but their main focus is to minimize the energy consumption. They neither consider the problem of perturbation in distance estimation nor in the neighborhood position information. In [47], the authors focus on improving the Wi-Fi positioning in indoor environments by using the RSSI-based distances among neighbors. They calculate confidence scores as weights in deciding the positions of the neighbors. The confidence of the Wi-Fi-based position is a function of the standard deviation of the multiple Wi-Fi scans for the same point. A lower standard deviation means higher confidence on position and vice-versa. Similarly, the confidence score of the Bluetooth is assigned by RSSI modeling in different settings of the indoor environment. Lastly, game theory is used to determine the final position of the nodes. Similar to others, this work only considered the error in the RSSI measurements. A simple subspace-based algorithm for single mobile node positioning using TOA measurements from three or more anchors with exact position information is given in [48].

In [10], the authors propose two WLS-based algorithms, called hyperbolic and circular, to localize a node in the presence of log-normal perturbations in the RSSI-based distance measurements. The hyperbolic algorithm linearizes the problem and solves it using the WLS method. The circular algorithm minimize the weighted approximation of the original non-linear sum-square-error cost function using the gradient-descent method. The circular algorithm performs better than the hyperbolic algorithm due to minimization of the original cost function by an iterative approach. The proposed algorithms match the low computational requirement of our applications of interest but do not consider perturbations in the anchor node positions. Our proposed algorithm is based on a similar approach of non-linear WLS type but considers...
the error/noise in the anchor positions as well.

In summary, the existing relevant works either assume the perturbations in both RSSI-based distance measurements and anchor positions to have normal distribution, which is not realistic given that the RSSI-based distance measurements follow log-normal distribution [49], or propose solutions that are based on complex optimization techniques such as SDP or SOCP. The applications of our interest fall into the category of “self-localization by a mobile node” where a resource-constrained mobile node has access to perturbed anchor node positions and RSSI measurements and is interested in localizing itself in a resource-efficient way. We consider the more realistic log-normal perturbation for the RSSI-based distance measurements and normal distribution for anchor position perturbations. To meet the resource-efficiency requirement, we find a weighted nonlinear least-squares solution for the considered location estimation problem by minimizing a weighted sum-square error cost function using the gradient-descent method.

III. PROBLEM STATEMENT

We consider the problem of self-localization by a single node, referred to as the blind node, on a two-dimensional Cartesian plane. The blind node is interested in obtaining an estimate of its true position, denoted by \((x_b, y_b)\). There are \(M \geq 3\) nodes, referred to as the anchor nodes, arbitrarily distributed within the communication range of the blind node at locations \((x_i, y_i)\), \(i = 1, ..., M\). The locations of the anchor nodes are known to the blind node only approximately as they are corrupted by random perturbations. The blind node reconciles its distance from the anchor nodes using available RSSI measurements that are also subject to random perturbations. We denote the perturbed knowledge of the anchor node positions at the blind node by \((\tilde{x}_i, \tilde{y}_i)\), \(i = 1, ..., M\), and the corresponding perturbed RSSI measurements by \(\tilde{p}_i\), \(i = 1, ..., M\), in the linear (mW) scale and by \(\tilde{p}_i\) (dBm), \(i = 1, ..., M\), in the logarithmic (dBm) scale.

We adopt the following common assumptions:

A1: The available knowledge of the position of the \(i\)th anchor node on \(x\) and \(y\) axes are corrupted by independent additive zero-mean Gaussian perturbations with standard deviation \(\sigma_a\). The perturbations in the knowledge of different anchor node positions are independent of each other and the values of \(\sigma_a\) may not be the same for different anchor nodes. Therefore, we have

\[
\tilde{x}_i = x_i + n_{x_i}, \quad \tilde{y}_i = y_i + n_{y_i}, \quad n_{x_i}, n_{y_i} \sim \mathcal{N}(0, \sigma_a).
\]

A2: The path-loss model for the radio signal propagation is the log-normal shadowing model. Therefore, the RSSI measurement of the signal transmitted from the \(i\)th anchor node and received at the blind node has a nominal value of \(\tilde{p}_i\) in the logarithmic (dBm) domain. However, the actual measured value is a realization of the nominal value corrupted by a zero-mean Gaussian perturbation with standard deviation \(\sigma_p\), i.e.,

\[
\tilde{p}_i(dBm) = \tilde{p}_i(dBm) + n_{p_i} \quad (1)
\]

According to the shadowing path-loss model, we have

\[
\tilde{p}_i(dBm) = p_0(dBm) - 10\eta \log \frac{d_i}{d_0} \quad (2)
\]

where

\[
d_i = \sqrt{(x_i - x_b)^2 + (y_i - y_b)^2}
\]

is the distance between the blind node and the \(i\)th anchor node.

Furthermore, we assume that the blind node and the anchor nodes have limited computational and energy resources. Hence, at any particular instance of localization, only one RSSI measurement and position estimate from each anchor node is available to the blind node.

IV. PROPOSED ALGORITHM

One can estimate the position of the blind node by minimizing the following sum-square-error cost function

\[
c(x, y) = \sum_{i=1}^{M} e_i^2(x, y)
\]

where

\[
e_i(x, y) = \delta_i - d_i
\]

and

\[
\delta_i = \sqrt{(x_i - x_b)^2 + (y_i - y_b)^2}.
\]

Given an initial estimate, the blind node’s position can be estimated using a gradient-descent method as follows:

\[
\begin{align*}
\dot{x}_b^{(k+1)} &= \dot{x}_b^{(k)} - \alpha \frac{\partial c}{\partial x} |_{x = \tilde{x}_b^{(k)}} \\
\dot{y}_b^{(k+1)} &= \dot{y}_b^{(k)} - \alpha \frac{\partial c}{\partial y} |_{y = \tilde{y}_b^{(k)}}
\end{align*}
\]

where \(\alpha > 0\) is the step-size parameter and \(\dot{x}_b^{(k)}\) and \(\dot{y}_b^{(k)}\) are the estimate of \(x_b\) and \(y_b\) at iteration \(k\), respectively.

However, we do not have access to the unperturbed values \(x_i\), \(y_i\), and \(d_i\). Hence, we replace them with their corresponding perturbed observations \(\tilde{x}_i\), \(\tilde{y}_i\), and \(\tilde{d}_i\) and approximate the \(i\)th error term with

\[
\tilde{e}_i(x, y) = \tilde{\delta}_i - \tilde{d}_i
\]

where

\[
\tilde{\delta}_i = \sqrt{(\tilde{x}_i - x_b)^2 + (\tilde{y}_i - y_b)^2}
\]

Furthermore, we assume that the blind node and the anchor nodes have limited computational and energy resources. Hence, at any particular instance of localization, only one RSSI measurement and position estimate from each anchor node is available to the blind node.
In addition, to factor in the difference in the scale and statistical properties of the values associated with different anchor nodes, we weight each error term with the inverse of its standard deviation. Therefore, we minimize the following WLS cost function

$$\tilde{e}(x, y) = \sum_{i=1}^{M} \frac{e_{i}^2}{\text{Var}(e_{i})}. \quad (4)$$

Since the perturbations of the anchor positions and RSSI-induced distances are independent of among each other, each variance term $\text{Var}(e_{i})$ can be calculated as

$$\text{Var}(e_{i}) = \text{Var}(\tilde{d}_{i}) + \text{Var}(d_{i}). \quad (5)$$

To calculate the first term on the right-hand side of (5), we note that $\delta_{i}$ is the Euclidean distance of the points $(x, y)$ and $(\tilde{x}_i, \tilde{y}_i)$ where $x$ and $y$ are deterministic variables and $\tilde{x}_i$ and $\tilde{y}_i$ are independent stochastic variables that have Gaussian distributions with means $x_i$ and $y_i$, respectively, and the same variance $\sigma^2_{a_i}$. Therefore, $\delta_{i}$ has a Rice distribution \[5\] with its approximate value of $\delta_{i}$ and its standard deviation. Therefore, we minimize the following

In addition, to factor in the difference in the scale and variance term $\text{Var}(\tilde{d}_{i})$ can be calculated as

$$\text{Var}(\tilde{d}_{i}) = \text{Var}(\tilde{\delta}_{i}) + \text{Var}(d_{i}). \quad (6)$$

where

$$\delta_{i} = \sqrt{(x_i - x)^2 + (y_i - y)^2} \quad \text{and} \quad L_{1/2}(\cdot) \text{ is a Laguerre polynomial expressed as}$$

$$L_{1/2}(z) = \exp(z/2) \left[ (1 - z) I_0 \left(-\frac{z}{2}\right) - z I_1 \left(-\frac{z}{2}\right) \right]$$

with $I_0(.)$ and $I_1(.)$ being the modified Bessel functions of the first kind with order zero and one, respectively.

Considering the assumption A2, the second term on the right-hand side of (5) is calculated as \[10\]

$$\text{Var}(d_{i}) = d_{i}^2 \left[ \exp(2\sigma^2_{a_i}) - \exp(\sigma^2_{a_i}) \right] \quad (7)$$

where

$$\sigma_{d_i} = \frac{\ln 10}{10\eta} \sigma_{p_i}.$$

We estimate the blind node position by minimizing the formulated cost function \[4\] in an iterative manner using the gradient-descent method. Since the unperturbed anchor positions $(x_i, y_i)$, $i = 1, \ldots, M$, are unknown, we replace $\tilde{\delta}_{i}$ in \[6\] with its approximate value of

$$\bar{\delta}_{i}^{(k)} = \sqrt{(\tilde{x}_i - \hat{x}_i)^2 + (\tilde{y}_i - \hat{y}_i)^2}$$

where $(\tilde{x}_i^{(k)}, \tilde{y}_i^{(k)})$ is the most recent estimate of the blind node location. In addition, we replace the unknown values of $d_{i}, i = 1, \ldots, M$ in \[7\] with their available approximate values $\tilde{d}_{i}, i = 1, \ldots, M$. Therefore, we calculate the gradients as

$$\frac{\partial \bar{e}}{\partial x}|_{x = \hat{x}_i^{(k)}} = -2 \sum_{i=1}^{M} \frac{\tilde{e}_{i}^{(k)} - \bar{\delta}_{i}^{(k)}}{w_{i}^{(k)} \bar{\delta}_{i}^{(k)}} \hat{x}_i - \hat{x}_i^{(k)}$$

and

$$\frac{\partial \bar{e}}{\partial y}|_{y = \hat{y}_i^{(k)}} = -2 \sum_{i=1}^{M} \frac{\tilde{e}_{i}^{(k)} - \bar{\delta}_{i}^{(k)}}{w_{i}^{(k)} \bar{\delta}_{i}^{(k)}} \hat{y}_i - \hat{y}_i^{(k)}$$

where the weight $w_{i}^{(k)}$ that is an estimate of the variance of the error due to the measurements related to the $i$th anchor node is calculated at the $k$th iteration as

$$w_{i}^{(k)} = \left(\hat{\delta}_{i}^{(k)}\right)^2 + 2\sigma^2_{a_i} - \frac{\pi\sigma^2_{a_i}\bar{L}_{1/2}}{2\sigma^2_{a_i}} \left(\frac{\bar{\delta}_{i}^{(k)}}{2\sigma^2_{a_i}}\right)^2\right) + \bar{d}_{i}^2 \left[ \exp(2\sigma^2_{a_i}) - \exp(\sigma^2_{a_i}) \right].$$

A. Computational Complexity

At each iteration, the proposed algorithm requires $16M$ multiplications and $4M$ divisions as well as $M$ square-root and $4M$ exponentiation operations. The required Bessel function values can be obtained using look-up tables, which only have some extra storage requirement. Therefore, the proposed algorithm has a computational complexity of $O(M)$ where $I$ is the total number of iterations needed for convergence.

V. CRAMER-RAO LOWER BOUND

Given an observation vector $x$ with known distribution that is related to an unknown parameter vector $\theta$, the Cramer-Rao lower bound (CRLB) sets a lower bound on the covariance of any unbiased non-Bayesian estimator of $\theta$. The CRLB is the inverse of the Fisher information matrix (FIM), denoted by $F(\theta)$. Therefore, we have

$$\text{Cov}(\hat{\theta}) \geq F^{-1}(\theta).$$

The FIM represents the information provided by the observation $x$ about the unobserved parameter vector $\theta$ and is calculated as

$$F(\theta) = -\mathbb{E} \left[ \frac{\partial^2 l(\theta | x)}{\partial \theta^2} \right]$$

where $l(\theta | x)$ is the log-likelihood function of $\theta$ given $x$.

In our self-localization problem, the observation vector $x$ contains the perturbed anchor positions $(\tilde{x}_i, \tilde{y}_i)$, $i = 1, \ldots, M$, as well as the perturbed RSSI measurements $\tilde{p}_i$, $i = 1, \ldots, M$, i.e., it can be written as

$$x = [\tilde{p}_1, \tilde{x}_1, \tilde{y}_1, \tilde{p}_2, \tilde{x}_2, \tilde{y}_2, \ldots, \tilde{p}_M, \tilde{x}_M, \tilde{y}_M]^T.$$

Given \[1\] and \[2\], the probability density function (pdf) of $\tilde{p}_i$ is expressed as \[8\]

$$f_{\tilde{p}_i}(\tilde{p}_i) = \frac{10/\ln 10 \exp \left[ -\frac{b_i}{8} \left( \ln \frac{d_{i}^2}{\bar{d}_{i}^2} \right)^2 \right]}{\sqrt{2\pi\sigma^2_{p_i}}}$$

where

$$b_i = \frac{(10\eta)^2}{(\sigma_{p_i} \ln 10)^2}.$$

In view of the assumption A1, the pdfs of the perturbed anchor positions $(\tilde{x}_i, \tilde{y}_i)$, $i = 1, \ldots, M$, are written as

$$f_{\tilde{x}_i}(\tilde{x}_i) = \frac{1}{\sqrt{2\pi\sigma^2_{a_i}}} \exp \left[ -\frac{(\tilde{x}_i - x_i)^2}{2\sigma^2_{a_i}} \right]$$

$$f_{\tilde{y}_i}(\tilde{y}_i) = \frac{1}{\sqrt{2\pi\sigma^2_{a_i}}} \exp \left[ -\frac{(\tilde{y}_i - y_i)^2}{2\sigma^2_{a_i}} \right].$$
Our unknown parameters of interest are the position coordinates of the blind node \((x_b, y_b)\). However, since the RSSI measurements are functions of the unknown unperturbed anchor positions, \((x_i, y_i), i = 1, ..., M\), we include the unperturbed anchor positions as the nuisance parameters. Hence, our unknown parameter vector \(\theta\) is
\[
\theta = [x_b, y_b, x_1, y_1, x_2, y_2, \ldots, x_M, y_M]^T.
\]

The perturbed anchor node positions \((\tilde{x}_i, \tilde{y}_i), i = 1, ..., M\), are statistically independent of each other as well as the RSSI measurements. Therefore, we write the log-likelihood function as
\[
l(\theta | x) = \sum_{i=1}^{M} \ln f_{\tilde{x}_i}(\tilde{x}_i | \theta) + \sum_{i=1}^{M} \ln f_{\tilde{y}_i}(\tilde{y}_i | \theta)
\]
and derive its second-order partial derivatives with respective to the entries of \(\theta\), required for the calculation of the FIM, in the Appendix. Using
\[
E \left[ \ln \left( \frac{d_i^2}{d_j^2} \right) \right] = 0
\]
with the results in the Appendix, we can express the FIM as
\[
F(\theta) = \begin{bmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{bmatrix}
\]
where
\[
F_{11} = \sum_{i=1}^{M} \frac{b_i}{d_i^4} Q_i,
\]
\[
F_{12} = -\left[ \frac{b_1}{d_1^2} Q_1, \ldots, \frac{b_M}{d_M^2} Q_M \right],
\]
\[
F_{22} = \text{blockdiag} \left\{ \frac{b_1}{d_1^2} Q_1 + \frac{1}{\sigma_{a1}^2} I_2, \ldots, \frac{b_M}{d_M^2} Q_M + \frac{1}{\sigma_{aM}^2} I_2 \right\},
\]
\[
Q_i = \begin{bmatrix} (x_i - x_b)^2 & (x_i - x_b)(y_i - y_b) \\ (x_i - x_b)(y_i - y_b) & (y_i - y_b)^2 \end{bmatrix},
\]
and \(I_2\) is the \(2 \times 2\) identity matrix. Consequently, we obtain a lower bound on the localization root-mean-square error (RMSE) of any unbiased estimator of the blind node location as below
\[
\sqrt{E \left[ (\hat{x}_b - x_b)^2 + (\hat{y}_b - y_b)^2 \right]} \geq \sqrt{\text{Tr} \left\{ (F_{11} - F_{12} F_{22}^{-1} F_{12}^T)^{-1} \right\}}.
\]

VI. EXPERIMENTAL SETUP

We consider the scenario of self-localization by a mobile blind node in the presence of six anchor nodes. All the nodes are assumed to be battery-powered and equipped with a processing unit and a low-power radio interface. Only anchor nodes are assumed to have a GPS unit for positioning. The perturbations on the position information of all anchor nodes are assumed to have Gaussian distribution with zero mean and known standard deviations. The radio propagation environment is assumed to be reasonably re-presentable by a log-normal shadowing path loss model with known parameters.

We evaluate the performance of the proposed algorithm in a \(35m \times 35m\) region. We consider both homogeneous and heterogeneous GPS noise scenarios in the region. The homogeneous scenario refers to the case where the standard deviation of the GPS noise is the same all over the region. In the heterogeneous scenario, the standard deviation of the GPS noise varies within the region. In practice, if the GPS operational context, e.g., surroundings, hardware, etc., is the same for all the nodes, we have a homogeneous scenario; otherwise, a heterogeneous one. As an example, the GPS localization performance is a function of the time during which the GPS receiver is active. Therefore, variations in the GPS activity time may lead to a heterogeneous scenario. The heterogeneous scenario may also arise with cooperative or group/cluster-based energy-efficient localization schemes where the GPS activation time is decided based on the available energy budget of the individual nodes. In our experiments, the standard deviations of the perturbations on the anchor node position information, \(\sigma_{a_i}, i = 1, ..., M\), are the same for all the anchor nodes when the GPS noise scenario is homogeneous. However, in the heterogeneous scenario, these standard deviations may differ depending on the position of the anchor nodes within the region.

The performance of the RSSI-based localization highly depends on the network geometry. Keeping this in mind, we further divide our experiments into two categories of “fixed” and “region-based random node placement”. In the fixed category, we evaluate the performance of the proposed algorithm with arbitrarily selected anchor and blind node positions, which we call semi-linear, semi-circle, corner anchor, arbitrary node placements. We also consider several values for the standard deviation of the GPS noise, i.e., 1m, 3m, and 5m. The RSSI measurement errors range from 0dBm to 5dBm in all the experiments. To avoid any bias due to blind node position initialization, we select an arbitrary point in the region as the initial estimate for the blind node. The initialization point is kept considerably far away from the true position of the blind node while being fixed in all trials of each experiment.

The region-based random node placement category is designed to emulate the geometries encountered by mobile nodes. In this category, we assign a region for the random placement of the anchor nodes. Similarly, we assign a region for the true position of the blind node as well as its initial position estimate. To further neutralize the impact of the network geometry on the localization performance, we consider two different anchor node placement arrangements, which we
Fig. 1. Homogeneous GPS error, fixed nodes, semi-circle anchor node arrangement; (a) the network topology: the stars are the anchor nodes, the circle is the true position of the blind node, the triangle is the initial estimate for the blind node position; (b) the RMSE of the existing algorithm for different values of $\sigma_{pi}$ and $\sigma_{ai}$; (c) the RMSE of the proposed algorithm for different values of $\sigma_{pi}$ and $\sigma_{ai}$; (d) the corresponding CRLB values.

Fig. 2. Homogeneous GPS error, fixed nodes, semi-linear anchor node arrangement; (a) the network topology: the stars are the anchor nodes, the circle is the true position of the blind node, the triangle is the initial estimate for the blind node position; (b) the RMSE of the existing algorithm for different values of $\sigma_{pi}$ and $\sigma_{ai}$; (c) the RMSE of the proposed algorithm for different values of $\sigma_{pi}$ and $\sigma_{ai}$; (d) the corresponding CRLB values.

call corner and semi-circle, as well as varying the GPS noise standard-deviation within the range of 1 – 10m.

The values of the path-loss model parameters used in our experiments are $d_0 = 1$m, $p_0$(dBm) = −33.44 and $\eta = 3.567$. These values are based on the results reported in [26]. We tune the step-size parameter of our gradient-descent algorithm to obtain convergence within 300 iterations. We compare the performance of the proposed algorithm with that of the so-called weighted circular algorithm proposed in [10]. This algorithm iteratively produces a WLS solution based on the assumption that only the RSSI measurements are corrupted by noise/error and the anchor node positions are exactly known at the blind node. Therefore, the proposed algorithm can be viewed as an improvement over the algorithm of [10] taking into account the effects of perturbation in anchor node positions. We compare our algorithm only with this algorithm since, to the best of our knowledge, it is the only relevant algorithm in the literature that has a computational complexity comparable to that of the proposed algorithm. We use the same step-sizes and maximum number of iterations in the implementations of both algorithms.

VII. SIMULATION RESULTS

We use the RMSE as the performance measure and calculate it by averaging over 1000 independent trials for each experiment. We also present the corresponding values of the theoretical lower bound on the RMSE, when appropriate, although the estimates produced by the proposed algorithm may be biased in some scenarios.

A. Fixed Node Placement

Fig. 1 presents the results for a homogeneous semi-circle arrangement with various values of $\sigma_{ai}$ and $\sigma_{pi}$. For $\sigma_{ai} = 5$m and $\sigma_{pi}$ in the range of 1dBm to 3dBm, the proposed algorithm performs around 20% better than the algorithm of [10]. However, for other combinations of $\sigma_{ai}$ and $\sigma_{pi}$, both algorithms perform nearly similar except for some benefits of the proposed algorithm when $\sigma_{ai} = 1$m and $\sigma_{p} = 1$, 2dBm. Fig. 2 shows the results for a semi-linear arrangement with a noise environment similar to the above semi-circle arrangement. The proposed algorithm performs around 15% better than the algorithm of [10] for the combination of the higher range of $\sigma_{ai}$ and the middle range of $\sigma_{pi}$. 
Fig. 3. Heterogeneous GPS error, fixed nodes, and arbitrary node placement; (a) the network topology: the stars are the anchor nodes with $\sigma_{a_i} = 6m$, the squares are the anchor nodes with $\sigma_{a_i} = 3m$, the circle is the true position of the blind node, the triangle is the initial estimate for the blind node position; (b) the RMSE of the proposed and existing algorithms as well as the corresponding CRLB for different values of $\sigma_{p_i}$; (c) the localization error histograms of the proposed and existing algorithms when $\sigma_{p_i} = 2$dBm.

Fig. 4. Heterogeneous GPS error, fixed nodes, and arbitrary node placement; (a) the network topology: the stars are the anchor nodes with $\sigma_{a_i} = 4m$, the squares are the anchor nodes with $\sigma_{a_i} = 2m$, the circle is the true position of the blind node, the triangle is the initial estimate for the blind node position; (b) the RMSE of the proposed and existing algorithms as well as the corresponding CRLB for different values of $\sigma_{p_i}$; (c) the localization error histograms of the proposed and existing algorithms when $\sigma_{p_i} = 2$dBm.

Fig. 5. Homogeneous GPS error, region-based random node placement, corner anchor node arrangement; (a) the network topology: the grey rectangles contain the anchor nodes with $\sigma_{a_i} = 5m$, the green rectangle contains the true position of the blind node, the yellow rectangle contains the initial estimate for the blind node position; (b) the RMSE of the proposed and existing algorithms; (c) the localization error histograms of the proposed and existing algorithms when $\sigma_{p_i} = 3$dBm.
Fig. 6. Homogeneous GPS error, region-based random node placement, semi-circle anchor node arrangement; (a) network topology: the grey rectangles contain the anchor nodes with $\sigma_a = 3\text{m}$, the green rectangle contains the true position of the blind node, the yellow rectangle contains the initial estimate for the blind node position; (b) the RMSE of the proposed and existing algorithms; (c) the localization error histograms of the proposed and existing algorithms when $\sigma_p = 3\text{dBm}$.

Fig. 7. Heterogeneous GPS error, region-based random node placement, corner anchor node arrangement; (a) network topology: the grey rectangle contains the anchor nodes with $\sigma_a = 5\text{m}$, the blue rectangle contains the anchor nodes with $\sigma_a = 1\text{m}$, the green rectangle contains the true position of the blind node, the yellow rectangle contains the initial estimate for the blind node position; (b) the RMSE of the proposed and existing algorithms; (c) the localization error histograms of the proposed and existing algorithms when $\sigma_p = 3\text{dBm}$.

Fig. 8. Heterogeneous GPS error, region-based random node placement, semi-circle anchor node arrangement; (a) network topology: the grey rectangles contain the anchor nodes with $\sigma_a = 4\text{m}$, the blue rectangle contains the anchor nodes with $\sigma_a = 1\text{m}$, green rectangle contains the true position of the blind node, the yellow rectangle contains the initial estimate for the blind node position; (b) the RMSE of the proposed and existing algorithms; (c) the localization error histograms of the proposed and existing algorithms when $\sigma_p = 3\text{dBm}$. 
Fig. 3 shows the results for a heterogeneous arbitrary selected node placement. In this experiment, the anchor nodes in the area with high GPS noise have $\sigma_{a1} = 6\text{m}$ and the other ones have $\sigma_{a1} = 3\text{m}$. Overall, the proposed algorithm reduces the RMSE by around 15% to 30%. The proposed algorithm reduces the gap between the algorithm of [10] and the CRLB by around 50%.

The results for another arbitrarily selected node placement with variation in $\sigma_{a1}$ is given in Fig. 4. There are three anchor nodes with $\sigma_{a1} = 4\text{m}$ and three anchor nodes with $\sigma_{a1} = 2\text{m}$. In low-noise radio environments, the RMSE of the proposed algorithm is just over half of that of the algorithm of [10]. The proposed algorithm performs relatively close to the CRLB compared with the algorithm of [10]. The error histogram of the algorithm of [10] is also wider than that of the proposed algorithm for $\sigma_{p1} = 2\text{dBm}$.

**B. Region-Based Random Node Placement**

Fig. 5 presents the results for the corner anchor node placement with homogeneous GPS error where the anchor nodes are randomly placed in two pre-specified regions. The values of $\sigma_{a1}$ are 5m in both regions and each region has three anchor nodes. The RSSI error $\sigma_{p1}$ varies from 1dBm to 5dBm. The proposed scheme reduces the RMSE by around 0.5m for all the variations of RSSI error. Fig. 5 shows a semi-circle arrangement of the nodes with $\sigma_{a1} = 3\text{m}$. In this arrangement, the proposed scheme improves the RMSE around 10 to 20%.

The results for a corner anchor node placement arrangement with $\sigma_{a1} = 1\text{m}$ and $5\text{m}$, respectively, for the low and high GPS error areas are given in Fig. 7. The proposed scheme reduces the RMSE by approximately 25% in the low RSSI error areas. With a high RSSI error, e.g., $\sigma_{p1} = 5\text{dBm}$, the proposed algorithm still shows an improvement of 7%. The error histogram of the proposed algorithm for $\sigma_{p1} = 3\text{dBm}$ is also slightly compacter than that of the algorithm of [10].

Fig. 8 shows the results for the heterogeneous semi-circle arrangement. The standard-deviation of the anchor position error $\sigma_{a1}$ is 4m in the poor GPS performance area and 1m in the high performance area. The proposed scheme performs 20 to 25% better in term of the RMSE with different RSSI error values.

In summary, the proposed algorithm significantly outperforms the algorithm of [10] in the realistic scenarios of heterogeneous GPS error. However, its advantages in the homogeneous GPS error scenarios is less pronounced.

**VIII. CONCLUSION**

We proposed a localization algorithm in the presence of the perturbation in the RSSI-based distance measurements as well as anchor positions information. We also derived the CRLB for the given problem. We evaluated the performance of the proposed algorithm in comparison with a previously-proposed algorithm that only accounts for perturbations in the RSSI measurements considering several arbitrary arrangements of anchor nodes and the blind node. Our simulation results showed that the proposed algorithm can significantly reduce the localization RMSE with respect to the existing algorithm.

**IX. APPENDIX**

The second-order partial derivatives of the log-likelihood function $l (\theta \mid x)$ with respect to the entries of its argument $\theta$, which are required for the calculation of the FIM, are computed as in the following:

$$
\frac{\partial^2 l (\theta \mid x)}{\partial x_i \partial x_j} = \sum_{i=1}^{M} \left \{ -\frac{b_i (x_i - x_b)^2}{d_i^4} + b_i \ln \left ( \frac{d_i^2}{d_i^2} \right ) \left [ \frac{(x_i - x_b)^2}{d_i^4} - \frac{1}{2d_i^2} \right ] \right \}
$$

$$
\frac{\partial^2 l (\theta \mid x)}{\partial y_i \partial y_j} = \sum_{i=1}^{M} \left \{ -\frac{b_i (y_i - y_b)^2}{d_i^4} + b_i \ln \left ( \frac{d_i^2}{d_i^2} \right ) \left [ \frac{(y_i - y_b)^2}{d_i^4} - \frac{1}{2d_i^2} \right ] \right \}
$$

$$
\frac{\partial^2 l (\theta \mid x)}{\partial x_b \partial y_b} = \sum_{i=1}^{M} \left \{ \frac{b_i (x_i - x_b)(y_i - y_b)}{d_i^4} \left [ \ln \left ( \frac{d_i^2}{d_i^2} \right ) - 1 \right ] \right \}
$$

For $i = j$, we have:

$$
\frac{\partial^2 l (\theta \mid x)}{\partial x_i \partial x_j} = \left \{ -\frac{b_i (x_i - x_b)^2}{d_i^4} + \frac{b_i \ln \left ( \frac{d_i^2}{d_i^2} \right )}{d_i^4} \left [ \frac{(x_i - x_b)^2}{d_i^4} - \frac{1}{2d_i^2} \right ] \right \} i = j
$$

For $i \neq j$, we have:

$$
\frac{\partial^2 l (\theta \mid x)}{\partial x_i \partial x_j} = \left \{ -\frac{b_i (x_i - x_b)^2}{d_i^4} + \frac{b_i \ln \left ( \frac{d_i^2}{d_i^2} \right )}{d_i^4} \left [ \frac{(x_i - x_b)^2}{d_i^4} - \frac{1}{2d_i^2} \right ] \right \} i \neq j
$$

For $i = j$, we have:

$$
\frac{\partial^2 l (\theta \mid x)}{\partial y_i \partial y_j} = \left \{ -\frac{b_i (y_i - y_b)^2}{d_i^4} + \frac{b_i \ln \left ( \frac{d_i^2}{d_i^2} \right )}{d_i^4} \left [ \frac{(y_i - y_b)^2}{d_i^4} - \frac{1}{2d_i^2} \right ] \right \} i = j
$$

For $i \neq j$, we have:

$$
\frac{\partial^2 l (\theta \mid x)}{\partial y_i \partial y_j} = \left \{ -\frac{b_i (y_i - y_b)^2}{d_i^4} + \frac{b_i \ln \left ( \frac{d_i^2}{d_i^2} \right )}{d_i^4} \left [ \frac{(y_i - y_b)^2}{d_i^4} - \frac{1}{2d_i^2} \right ] \right \} i \neq j
$$

**REFERENCES**

[1] N. Vallina-Rodriguez and J. Crowcroft, “When assistance becomes dependence: characterizing the costs and inefficiencies of A-GPS,” ACM SIGMOBILE Mobile Computing and Communications Review, vol. 17, no. 4, pp. 3–14, 2013. [Online]. Available: http://dl.acm.org/citation.cfm?id=2557970

[2] F. B. Abdesslem, A. Phillips, and T. Henderson, “Less is more: energy-efficient mobile sensing with senseless,” Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhelds., pp. 61–62, 2009. [Online]. Available: http://dl.acm.org/citation.cfm?id=1592621

[3] I. Constandache, G. Gaonkar, M. Sayler, R. R. Choudhury, and L. Cox, “Enloc: Energy-efficient localization for mobile phones,” in INFOCOM 2009, IEEE, IEEE, 2009, pp. 2716–2720.

[4] R. Jurdak, B. Kusy, and A. Cotillon, “Group-based Motion Detection for Energy-Efficient Localisation,” Journal of Sensor and Actuator Networks, vol. 1, no. 3, pp. 183–216, 2012.

[5] Y. Lee, Y. Ju, C. Min, S. Kang, I. Hwang, and J. Song, “Comon: Cooperative ambiance monitoring platform with continuity and benefit awareness,” in Proceedings of the 10th international conference on Mobile systems, applications, and services. ACM, 2012, pp. 43–56.
[48] H.-C. So and F. K. Chan, “A generalized subspace approach for mobile positioning with time-of-arrival measurements,” *IEEE transactions on Signal Processing*, vol. 55, no. 10, pp. 5103–5107, 2007.

[49] T. S. Rappaport et al., *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[50] S. B. L. Rice, “Mathematical Analysis of Random Noise Part III: statistical properties of random noise currents \_bstj24-1-46.pdf,” pp. 46–156, 1945.

[51] N. H. Nguyen and K. Doğançay, “Optimal geometry analysis for multistatic toa localization,” *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4180–4193, 2016.