IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 7, 2020, accepted January 11, 2021, date of publication January 25, 2021, date of current version February 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3053994

Cramer-Rao Lower Bound Analysis of Data Fusion for Fingerprinting Localization in Non-Line-of-Sight Environments

JUN LI[1], (Graduate Student Member, IEEE), I-TAI LU[1], (Senior Member, IEEE), AND JONATHAN LU[2], (Member, IEEE)

[1]Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, New York, NY 11201, USA
[2]ZaiNar, Redwood City, CA 94063, USA

Corresponding author: Jun Li (jl7333@nyu.edu)

ABSTRACT This paper proposes a novel framework for analyzing the localization accuracy of data fusion for fingerprinting approaches in non-line-of-sight (NLOS) environments. Using simulation data generated for two very different NLOS environments (a suburban area of $3.3km \times 3.3km$ in Santa Clara, California, and a mountainous area of $11.4km \times 11.4km$ in the Caspian region), we establish novel channel models for measurement differences of three data types (received signal strength indicator (RSS), time of arrival (TOA) and direction of arrival (DOA)) at K neighboring nodes of an arbitrary node. The crucial point is that the modeling errors for each of the three data types are shown to be jointly Gaussian distributed. Based on these measurement difference models, Cramer-Rao Lower Bound (CRLB) is used as a benchmark to evaluate K-nearest neighbor (KNN) and Weighted K-Nearest Neighbor (WKNN). It is shown that the proposed CRLB analyses can be employed to evaluate fingerprinting systems with various designs (such as different data types and fusion options) and different configurations (such as densities of reference nodes and numbers of anchor nodes) in diverse NLOS environments(such as suburban and mountainous regions).

INDEX TERMS Localization, NLOS, 5G, weighted K-nearest neighborhood, RSS, TOA, DOA.

I. INTRODUCTION

With rapid increases in mobile data volume in millimeter-wave (mm-wave) communications and fast development of Internet of things (IoT), many location-based services such as advertisements, smart cities, social networking, and robot navigation, become prevalent [1]–[3]. Since location information is required in IoT and various mobile systems for applications such as device tracking and device to device resource allocation [4], accurate and low-cost localization approaches become of high demand nowadays.

The most well-known localization technology is the global positioning system (GPS) [5], [6], which enables users to find their positions almost everywhere in the world. However, installing a global positioning system (GPS) in each device will undoubtedly increase the manufacturing and networking costs [7]. Moreover, its localization accuracy is inferior in situations where the line-of-sight (LOS) ray does not exist. Note that the LOS ray will be blocked entirely or significantly attenuated as long as there are obstructions (e.g., walls, vegetation, buildings, and mountains) between the transmitter and the receiver. Under such situations, non-line-of-sight (NLOS) propagation becomes dominant, and those localization approaches depending mainly on LOS propagation (such as GPS) are no longer applicable. It is then necessary to use other localization approaches that take into consideration of NLOS propagation.

One of the most frequently applied localization approaches in wireless networks under NLOS conditions is the fingerprinting-based approach [8]–[12], which is GPS-free and easy hardware-implementation. However, it is expensive and time consuming to collect fingerprinting data in large areas, especially the outdoors. With the development of the high-performance computing technologies such as parallel computing on GPUs, the offline environment-driven measurements can be generated by numerical simulations [13], making fingerprinting-based localization approaches more competitive.

Note that different types of measurement data, such as received signal strength indicator (RSS), time of

The associate editor coordinating the review of this manuscript and approving it for publication was Irfan Ahmed.

arrival (TOA), or direction of arrival (DOA), can be employed for fingerprinting [14]. It is known that the RSS-based fingerprinting approaches are low cost and can have excellent performance in indoor localization. Nevertheless, its estimation error could become very large due to wireless channel fading in the outdoors [15], [16]. In [17]–[20], TOA-based fingerprinting approaches are shown to provide better accuracy in specific environments but require a more extensive system bandwidth and additional sophisticated hardware [15]. Notably, as shown in [20], [21], TOA information can be used to improve RSS-based positioning accuracy drastically. In [14], [22], [23], fingerprinting-based approaches using DOA measurements are proposed. It is shown in [23] that DOA is similar to TOA, which is position sensitive and suitable for localization. However, accurate DOA can only be obtained using large receive antenna arrays, which is usually unavailable for civilian applications in the past. As radio communications move to include the mm-Wave band in 5G networks, antenna dimensions become very small and large antenna arrays become possible [24]. Furthermore, the raw time resolution [25] in mmWave systems is much improved due to the much higher available bandwidths [26], [27]. Thus, accurate DOA measurements will soon be easily obtained.

As RSS, TOA, and DOA have complementary properties, combining two or three of these different data types (data fusion) to improve localization accuracy is desirable in future 5G fingerprinting-based positioning systems. However, since the system using more features will cost more, it will be beneficial to know the Cramer-Rao Lower Bound (CRLB) [28] of localization accuracy for each data type alone (without fusion) and each possible combination of these data types (fusion option) before system deployment. It turns out that some researchers have foreseen this need and have been working toward this goal in recent years. However, most researches have been done assuming the LOS scenario. For examples, the authors in [3], [29], [30] have presented CRLB analyses for RSS, TOA, and DOA individually and discuss their results; the authors of [31]–[33] performed CRLB analyses for data fusion of TOA and DOA; the authors of [34]–[36] performed CRLB analyses for data fusion of RSS and TOA; the authors of [22], [37], [38] have performed CRLB analyses for data fusion of RSS and DOA; and the authors of [39] have performed CRLB analyses for data fusion of RSS, DOA, and TOA. However, these analyses cannot be employed for the NLOS scenario since the propagation mechanisms and models for the NLOS scenario are very different from those for the LOS scenario.

The CRLB of localization accuracy analyses have also been carried out for the NLOS scenario by various authors. Nevertheless, few papers are on data fusion. The authors in the very well-cited papers [40]–[42] have shown CRLB analyses for RSS, TOA, and DOA individually and discuss their results. As for data fusion, the authors of [43], [44] have performed CRLB analyses for data fusion of TOA and DOA; and the authors of [45], [46] have performed CRLB analyses for data fusion of RSS and TOA. However, we have not found papers on CRLB analyses for data fusion of RSS and DOA or data fusion of RSS, DOA, and TOA.

Note that some of NLOS approaches are like those for LOS. For example, the authors in [43] conducted the localization accuracy analysis using AOD/AOA/TOA measurements in a NLOS indoor environment. However, the analysis assumes single-bounce reflection in every NLOS propagation path. By image theory, it is known that a single-bounce reflection is equivalent to a LOS ray from the image source. Another example is in [47], where the authors conduct the CRLB analysis in a mixed LOS/NLOS environment, but require identifying the number of NLOS paths in advance. In this paper, NLOS paths are not desired. They are sources of error of a LOS-based localization approach. Unlike [43], [47], the approach in [48] does not require to identify the NLOS paths for indoor fingerprinting. However, it uses the Gaussian Process to interpolate 1-D magnetic data and RSS data, which cannot be employed to model the TOA and DOA data in large NLOS environments, especially the outdoors.

Unlike [48], the authors in [40]–[42], [46], [49]–[52] use sophisticated statistical models for RSS, TOA, and DOA because simple Gaussian model cannot fit the measurement data well in complex NLOS environments. The authors of these papers use maximum likelihood estimation (MLE) or maximum a posteriori probability (MAP) approaches for localization and use CRLB to find the best achievable positioning accuracy. Considering TOA as an example, the error term is typically modeled as Exponential, Rayleigh, or Gamma distributed in these papers. These sophisticated non-Gaussian models are environment-dependent and cannot be generalized easily to different environments.

The CRLBs derived in the papers mentioned above for the NLOS scenario are not suitable to characterize the localization accuracy of fingerprinting approaches. The systems in these papers consist of multiple transmitting anchor nodes and the unknown receiving node(s). There are no reference nodes. However, the systems in fingerprinting approaches typically consist of multiple transmitting anchor nodes, many receiving reference nodes, and the unknown receiving node(s). Similarities of measurement data at the unknown receiving node and those at its neighboring reference nodes play crucial roles in fingerprinting approaches. To the best of our knowledge, no one has provided a practical localization accuracy analysis for fingerprinting approaches in NLOS environments.

In this paper, we propose a novel framework for CRLB analysis to characterize the localization accuracy of fingerprinting approaches in different NLOS environments. Since fingerprinting localization depends on differences of measurement data (such as ROA, TOA, and DOA) between the unknown node and its neighboring nodes, we do not need sophisticated models for measurement data at each receiving node explicitly. Thus, instead of modeling the measurement data directly, we propose to model the measurement differences.

Note that the authors in [53] analyze the CRLB of localization accuracy using RSS difference. However, their RSS difference is utterly different from our RSS difference here. The RSS difference in [53] is the difference between two RSS measurements at the unknown node due to two far apart transmitting anchors. In this manner, it still needs the sophisticated measurement data model for each transmitter-receiver pair explicitly because there is no correlation between the two anchors. Nevertheless, our RSS difference is the difference between two RSS measurements at two neighboring nodes due to the same transmitting anchor. In this manner, we do not need the sophisticated measurement data model for each transmitter-receiver pair explicitly because there is a strong correlation between the two neighboring nodes, and the RSS difference will remove the correlated portions of NLOS propagation.

There are two significant differences between our proposed CRLB approaches for fingerprinting localization and the aforementioned approaches for MLE and MAP localization in the NLOS scenario. Firstly, authors in the papers mentioned above model the measurement data at all possible receiver locations in a vast region directly, but we model the measurement differences of neighboring nodes located in a small confined region. Although the small confined region can be anywhere in the vast region of interest, the proposed modeling process is simple because the derived models are Gaussian distributed, which can be generalized to different NLOS environments. However, the modeling process for the existing approaches is complicated because their derived models are usually non-Gaussian distributed, which cannot be generalized to different NLOS environments. Secondly, the above-mentioned approaches do not take advantage of the similarities in measurement data at the neighboring nodes and, therefore, miss modeling the correlations between measurements at neighboring nodes, which are essential to fingerprinting approaches. However, the proposed fingerprinting approaches model the measurement differences jointly, which account for the correlation between neighboring nodes' measurements. Thus, the proposed approaches can characterize the localization accuracy of fingerprinting approaches better.

Our proposed models are based on physical insights of wave propagation principles. Note that the dominant propagation mechanism is most likely NLOS in complex environments when the separation distance between the transmitter and the receiver is large. However, as neighboring nodes usually see the same dominant NLOS ray, the ray trajectory of this dominant NLOS ray in between these neighboring nodes is linear (i.e., LOS). Thus, the relations between these neighboring nodes' measurements are governed by the LOS plane wave propagation mechanism. Motivated by this physical insight, we hypothesize that the measured data difference between a given node and one of its nearby neighbors can be modeled as a linear function of a deterministic factor (determined by the LOS propagation mechanism) plus a random modeling error term. Moreover, concerning the given node, the modeling errors of measurement differences at the K nearest neighbors for the same data type (RSS, TOA, or DOA) are assumed jointly Gaussian distributed, but the modeling errors for different data types are assumed uncorrelated. Here, the number K is to be determined for each specific problem. With these statistical models of measurement differences at hand, the CRLBs for localization accuracy of fingerprinting approaches using various data types and different fusion options can then be easily derived.

To validate the proposed models for measurement differences, we consider two very different complex environments: a suburban region and a mountainous region. The suburban region is an area of $3.3km \times 3.3km$ in Santa Clara, California. Note that the environment is very complex, and predominant NLOS rays are from street corner diffraction at faraway receivers. Our previously developed ray launcher [13] is employed to generate the RSS, TOA, and DOA data. The ray launcher is fully discrete in order to takes advantage of environment-driven parallel processing so as to efficiently trace the launched rays undergoing both specular and diffuse interactions. The mountainous region is a $11.4km \times 11.4km$ area in the Caspian land. The RSS, TOA, and DOA data are generated using our previously developed ray tracer [54]. Accounting for the mountainous terrain and its undulations in the transmitter and receiver's vicinity, this simulator predicts the propagation characteristics of rays traveling in the vertical plane containing the transmitter and receiver and terrain scattered rays traveling outside of this plane.

In this paper, system formulation, measurement differences modeling, and CRLB derivation for fingerprinting approaches are shown in Section II. The ray launcher and ray tracer for generating measurement data (RSS, TOA, and DOA) in suburban and mountainous environments are described in Section III. In Section IV, the proposed models for RSS, TOA, and DOA measurement differences at K nearest neighbors of a given receiving node are established and validated by linear regression using simulation data for both Santa Clara suburban and Caspian mountainous regions. In Section V, CRLB is used as a benchmark to evaluate two popular fingerprinting approaches (K-nearest neighbor (KNN) and Weighted K-Nearest Neighbor (WKNN) [12]) in both suburban and mountainous regions for various data types and fusion options. In Section VI, it is concluded that, with measurement difference modeling, CRLB accuracy analyses for fingerprinting-based localization approaches using various data types and different fusion options can be conducted for distinct system configurations in diverse NLOS environments.

II. CRLB ANALYSIS FOR FINGERPRINTING LOCALIZATION IN NLOS ENVIRONMENTS

Consider a system with N_T anchors, N_R reference nodes and N_U unknown nodes. Each receiver (representing a reference node or an unknown node) will receive sounding signals from the transmitters (representing anchor nodes) via LOS or NLOS propagation paths. The propagation environments of interest are assumed to be very complex, where a lot of main

arrivals are not LOS rays. Furthermore, it is assumed that we have no prior knowledge to determine whether a particular measured RSS, TOA, and DOA is via a LOS path or not. Therefore, approaches depending on LOS rays or needing to distinguish LOS and NLOS rays are not applicable here.

Fortunately, there exist features, such as RSS, TOA, and DOA, that are functions of locations. Moreover, in local regions, autocorrelations of these features at two different locations decrease when the separation distance between the two locations decreases. Thus, each or a combination of these features can be used for localization via fingerprinting approaches.

A. FINGERPRINTING LOCALIZATION APPROACHES

Let T_i be the positions of the anchors with $i \in \{1, \cdots, N_T\}$. The positions of receivers are denoted by R_j with $j \in \{1, \cdots, N_R + N_U\}$ where indexes $\{1, \cdots, N_R\}$ denote the reference nodes and indexes $\{1 + N_R, \cdots, N_R + N_U\}$ denote the unknown nodes. Here, T_i and R_j are $D \times 1$ vectors, where D is the dimension of the coordinates of the environment of interest, typically $D = 2$ or 3. We consider a two-dimensional localization problem and, therefore, $D = 2$ in this paper. All our results can be extended to $D = 3$ by following the same procedure presented in this paper.

Denote the RSS, TOA, and DOA measurements at the j^{th} receiver from the i^{th} anchor as $r_{j|i}$, $t_{j|i}$, and $(\hat{a}_{j|i}, \check{a}_{j|i})$, respectively. For convenience, TOA is normalized with $1/c$ where c is the speed of light. Thus, $t_{j|i}$ has the unit of *meter*. RSS $r_{j|i}$ has the unit of *dBm*. Note that the DOA measurement is two-dimensional in a two-dimensional problem. This is due to the fact that the angle in a two-dimensional plane has a mod 2π property and cannot be described by a single number. Here, we have defined $\hat{a}_{j|i}$ and $\check{a}_{j|i}$ as real and imaginary parts of $e^{j\theta}$, respectively, where θ is the angle of arrival. Thus, both $\hat{a}_{j|i}$ and $\check{a}_{j|i}$ are dimensionless.

Let the $N_T \times 1$ vectors r_j, t_j and $(\hat{a}_j, \check{a}_j)$ represent the RSS, TOA, and DOA measurements at the j^{th} node from all N_T anchors, respectively. In this paper, the measurements r_j, t_j and $(\hat{a}_j, \check{a}_j)$ can be used individually or fused together for localization. For convenience, use s_j to represent all possible measurement data vectors to be adopted for localization. Thus, if we fuse all three types of measurements, $s_j = [r_j^T \quad t_j^T \quad \hat{a}_j^T \quad \check{a}_j^T]^T$ is a $4N_T \times 1$ vector. Similarly, if we fuse only RSS and TOA measurements, $s_j = [r_j^T \quad t_j^T]^T$ is a $2N_T \times 1$ vector. However, if we fuse only RSS and DOA measurements, $s_j = [r_j^T \quad \hat{a}_j^T \quad \check{a}_j^T]^T$ is a $3N_T \times 1$ vector. Likewise, if we adopt only RSS measurements, $s_j = r_j$ is a $N_T \times 1$ vector. But, if we adopt only DOA measurements, $s_j = [\hat{a}_j^T \quad \check{a}_j^T]^T$ is a $2N_T \times 1$ vector.

Fingerprinting localization approaches consist of two steps: the offline step and the online step. During the offline step, we need to find a model that can best describe the relationship between the position R_k and the measurement s_k at the k^{th} unknown node, $k \in \{N_R + 1, \cdots, N_R + N_U\}$. For the WKNN approach, the positions and measurements of K nearest reference-node neighbors of the unknown node are included in determining the position of the unknown node. Then, this model is written as a vector function as shown below:

$$R_k \approx f_1(s_k; s_k^{(1)}, s_k^{(2)}, \ldots, s_k^{(K)}; \bar{R}_k^{(1)}, \bar{R}_k^{(2)}, \ldots, \bar{R}_k^{(K)}) \tag{1}$$

where $s_k^{(m)}$ and $\bar{R}_k^{(m)}$ are the measurement and position of the m^{th} "nearest" reference-node neighbor of the k^{th} unknown node, withe $m = 1, 2, \ldots, K$. Here, "nearest" is not measured according to the physical location distances between the unknown node and the reference nodes since we do not know the exact location, R_k, of the k^{th} unknown node. Once $f_1(\cdot)$ is determined, the offline step is done. During the online step, R_k can simply be estimated using (1).

Note that, in order to find the vector function $f_1(\cdot)$ in (1), we need to determine the K "nearest" neighbors first. In this paper, the K "nearest" neighbors of the k^{th} unknown node are the K reference nodes which have the measurement vectors most similar to the measurement vector of the unknown node. Define the weighted Euclidean distance as

$$\epsilon_k^j = \| \Lambda(s_k - s_j) \|_2 \tag{2}$$

with $j = 1, \ldots, N_R$. Here, Λ is a diagonal matrix whose diagonal elements are the weights put on different data types when these data types are fused for fingerprinting localization. Specific weights can be found by trial and error. Using (2), the K "nearest" neighbors of the k^{th} unknown node are the K reference nodes with the K least weighted Euclidean distances $\{\epsilon_k^j\}$. For convenience, we re-index the reference nodes according to the weighted Euclidean distances with respect to the k^{th} unknown node and let $(s_k^{(m)}, \bar{R}_k^{(m)})$ denote the measurement and position information of the m^{th} "nearest" neighbor, $m \in \{1, 2, \cdots, K\}$, for the k^{th} unknown node. Also, define $\epsilon_k^{(m)}$ as the weighted Euclidean distance between the measurement vector of the k^{th} unknown node and the measurement vector of its m^{th} "nearest" reference node:

$$\epsilon_k^{(m)} = \| \Lambda(s_k - s_k^{(m)}) \|_2, \quad \epsilon_k^{(1)} \le \epsilon_k^{(2)} \cdots \le \epsilon_k^{(K)}. \tag{3}$$

Using the weighted Euclidean distances in (3), the vector function $f_1(\cdot)$ in (1) can be simplified as:

$$R_k \approx f_2(\epsilon_k^{(1)}, \epsilon_k^{(2)}, \ldots, \epsilon_k^{(K)}; \bar{R}_k^{(1)}, \bar{R}_k^{(2)}, \ldots, \bar{R}_k^{(K)}) \tag{4}$$

for the well-known WKNN fingerprinting localization approach. Specifically, (4) is written as

$$R_k \approx f_2 = \sum_{m=1}^{K} w(\epsilon_k^{(m)}) \bar{R}_k^{(m)} \tag{5}$$

where $\{w(\cdot)\}$ are the weights used for averaging the locations of K "nearest" neighbors. The WKNN approach is reduced to the KNN approach if the weights w in (5) is $\frac{1}{K}$ for every m. Note that the number K in (5) is to be determined in practical applications. In our numerical procedure, we use a cross-validation data set to find the best parameter K, the optimum weight factors Λ in (2) and $\{w(\cdot)\}$ in (5). Representing typical fingerprinting approaches, numerical results of KNN

and WKNN will be compared with the CRLB proposed in the next section.

B. CRAMER-RAO LOWER BOUND (CRLB) ANALYSIS

To derive the CRLB for a localization approach in a given environment, one needs to have a workable statistical model for the measurements of that environment. However, in complex NLOS environments, no simple models exist. Therefore, scarce work has been done on CRLB analysis in NLOS environments and, for those existing ones, their models are very complicated and only for exceptional cases.

To tackle this challenging problem, we observe from (4) that the fingerprinting localization approaches use only the measurement differences between the unknown node and its ''nearest'' neighboring reference nodes. As the measurements at the unknown node and those at the reference nodes do not appear in (4) independently, we realize that we do *not* need to model the measurements directly, which is very difficult. For CRLB analysis for fingerprinting approaches, we just need to model the measurement differences, which is feasible, flexible, and general (can be extended for cases with various data types and in various environments).

For convenience, let $v_{j|i}$ represent any of the following data: $r_{j|i}$, $t_{j|i}$, $\hat{a}_{j|i}$ and $\breve{a}_{j|i}$ (which are the RSS, TOA, and DOA measurements at the j^{th} receiver due to the transmission from the i^{th} anchor). Also define $\delta v_{k|i}^{(m)} = v_{k|i} - v_{k|i}^{(m)}$ as the difference between the measurement $v_{k|i}$ of the k^{th} unknown node at R_k and the measurement $v_{k|i}^{(m)}$ of its true m^{th} nearest reference node at $R_k^{(m)}$ due to the transmission from the i^{th} anchor. Note that the K true nearest neighbors are determined according to their physical distances with respect to R_k of the node of interest,

$$\| R_k - R_k^{(1)} \|_2 \leq \| R_k - R_k^{(2)} \|_2 \leq \cdots \leq \| R_k - R_k^{(K)} \|_2 \qquad (6)$$

Here, using (6) (instead of (3)) to determine the K nearest neighbors in the CRLB formulation is to obtain a smaller CRLB.

Typically, $\delta v_{k|i}^{(m)}$ and $\delta v_{k|i}^{(\hat{m})}$ are correlated because both m^{th} and $\hat{m}^{th}$ neighbors of the k^{th} node are in the same local region. Contrarily, $\delta v_{k|i}^{(m)}$ and $\delta v_{k|\hat{i}}^{(m)}$ are not correlated as the i^{th} and $\hat{i}^{th}$ anchors are far apart. Note that ''local region'' is used to classify reference nodes (receivers) here. As we will use the K nearest neighbors to build the difference models of measurement data for computing CRLB, the definition of the local region is the region occupied by the K nearest neighbors. Those reference nodes which are not in the local region are not employed and are not of concerned. Contrarily, ''far apart'' is used to describe anchor nodes (transmitters). In our examples, the separation distance between any pair of anchor nodes is much larger than that between any pair of reference nodes in a local region.

Furthermore, we assume that different data types are not correlated because they represent different physical mechanisms that are affected by the environment differently. Then, the novel statistical model of the measurement difference is proposed to be:

$$\boldsymbol{\delta} v_{k|i} = \boldsymbol{\mu}_{k|i}^{v}(R_k) + q_{k|i}^{v} \qquad (7)$$

where $\boldsymbol{\delta} v_{k|i} = \left[\delta v_{k|i}^{(1)},\ \delta v_{k|i}^{(2)},\ \cdots,\ \delta v_{k|i}^{(K)}\right]^T$, $q_{k|i}^{v}$ is a random vector and $\boldsymbol{\mu}_{k|i}^{v}(\cdot)$ is a deterministic vector function of the unknown node's position R_k. Note that the m^{th} element of $\boldsymbol{\mu}_{k|i}^{v}(\cdot)$ is also a function of the location of the m^{th} nearest neighboring reference node: $R_k^{(m)}$. Here, v and v denote the type of data measurement. Ideally, $\boldsymbol{\mu}_{k|i}^{v}(\cdot)$ has to be chosen such that $q_{k|i}^{v}$ follows the zero-mean multivariate normal distribution (i.e. $q_{k|i}^{v} \sim N(0, C_{k|i}^{v})$). Here, we would also like to minimize the diagonal elements of the covariance matrix $C_{k|i}^{v}$.

As the measurement differences are correlated with respect to different nearest neighbor indexes $\{m\}$'s, but are not correlated with respect to different anchor indexes i's or to different data types, it is much more convenient and informative to rewrite (1) in terms of $\delta v_{k|i}$ for every i (instead of s_k and $s_k^{(m)}$ for every m as shown in (1)) for the following derivations. Thus, from (1), we have:

$$R_k \approx f_3(\{\boldsymbol{\delta} v_{k|i}\}_{i;v}; R_k^{(1)}, R_k^{(2)}, \ldots, R_k^{(K)}) \qquad (8)$$

where $\{\delta v_{k|i}\}_{i;v}$ in (8) represents all data types $\{v\}$ and all anchor nodes $\{i\}$ used in localization. For convenience and readability, we do not specifically list all data types since many data fusion options are considered in this paper. Note that, in (8), we have also used $\{R_k^{(1)}, R_k^{(2)}, \ldots, R_k^{(K)}\}$ of the K true nearest neighbors instead of using $\{\bar{R}_k^{(1)}, \bar{R}_k^{(2)}, \ldots, \bar{R}_k^{(K)}\}$ of the K estimated ''nearest'' neighbors in (1).

Suppose we use minimum mean squared error (MMSE) as our evaluation criterion, it is lower bounded by the well-known CRLB:

$$\min \sqrt{E\{||R_k - f_3||_2^2\}} \geq \sqrt{tr\{[I_k(R_k)]^{-1}\}} \qquad (9)$$

where

$$I_k(R_k) = \sum_{v} \sum_{i=1}^{N_T} I_{k|i}^{v}(R_k). \qquad (10)$$

Since the fingerprinting-based approaches are asymptotically unbiased [55]–[57], the use of CRLB as a benchmark is appropriate. Here, $I_{k|i}^{v}(R_k)$ is the Fisher Information Matrix (FIM) for data type v and anchor i:

$$I_{k|i}^{v}(R_k) = \begin{pmatrix} [I_{k|i}^{v}(R_k)]_{xx} & [I_{k|i}^{v}(R_k)]_{xy} \\ [I_{k|i}^{v}(R_k)]_{yx} & [I_{k|i}^{v}(R_k)]_{yy} \end{pmatrix}. \qquad (11)$$

Derivations of FIM are given in Appendix A. Let ζ (and ξ) denote the spatial coordinates x or y. Then, from (11), we have

$$[I_{k|i}^{v}(R_k)]_{\zeta\xi} = \left[\frac{\partial \boldsymbol{\mu}_{k|i}^{v}(R_k)}{\partial \zeta}\right]^T \{C^v\}^{-1} \left[\frac{\partial \boldsymbol{\mu}_{k|i}^{v}(R_k)}{\partial \xi}\right]. \qquad (12)$$

C. MODELS FOR MEASUREMENT DIFFERENCES

Even though the dominant propagation mechanism between a transmitter and faraway receivers in complex environments is most likely NLOS, the relations between the measurements at neighboring receivers may be governed by the LOS propagation mechanism if these neighbors are nearby one another. Therefore, our models for (7) are based on the hypothesis of dominant LOS propagation between the K nearest neighbors. The models for RSS, TOA and DOA measurement differences at these neighbors will be established and validated using simulation data at a mountainous environment and a suburban environment in Section IV.

In order to reduce the number of models required in a given CRLB analysis, the measurement difference models in (7) are made to depend only on the differences of node location vectors. Thus, (7) can be rewritten as:

$$\boldsymbol{\delta} v = \boldsymbol{\mu}^v(\{T-R\}, \{T-R^{(m)}\}_{m=1,2,\ldots,K}) + q^v \qquad (13)$$

where the subscript $k|i$ in (7) and the subscripts k and i for the location vectors are removed. Here, T denotes the location of an anchor node, R denotes the location of an unknown node, and $R^{(m)}$ denotes the location of the m^{th} nearest neighbor of the unknown node for every $m = 1, 2, \ldots, K$. Specifically, the m^{th} row in the vector equation (13) is corresponding to the m^{th} nearest neighbor and can be written as

$$\delta v_m = \mu_m^v(\{T-R\},\ \{T-R^{(m)}\}) + q_m^v. \qquad (14)$$

For each data type v, (14) shows that the same deterministic function μ_m^v and random variable q_m^v are employed to model for any unknown node k and any anchor node i in consideration. However, (14) needs to be re-established if the environment or the system configuration changes.

For simplicity, we use linear regression to model μ_m^v, which will be shown to be reasonable in Section IV via suburban and mountainous data. So the μ_m^v in (14) is chosen to be

$$\mu_m^v = \beta^r \eta_m^v + \alpha_m^r \qquad (15)$$

where, for different data types, η_m^v in (15) are chosen to be

$$\begin{aligned}
RSS:\ \eta_m^r &= 10\log_{10}\left(\frac{||R-T||_2}{||R^{(m)}-T||_2}\right)\\
TOA:\ \eta_m^t &= (||R-T||_2 - ||R^{(m)}-T||_2)\\
DOA_x:\ \eta_m^{\hat{a}} &= \left(\frac{[R-T]_x}{||R-T||_2} - \frac{[R^{(m)}-T]_x}{||R^{(m)}-T||_2}\right)\\
DOA_y:\ \eta_m^{\check{a}} &= \left(\frac{[R-T]_y}{||R-T||_2} - \frac{[R^{(m)}-T]_y}{||R^{(m)}-T||_2}\right).
\end{aligned} \qquad (16)$$

Here, $[\cdot]_x$ and $[\cdot]_y$ are the x and y components of the location vector $[\cdot]$, respectively. Note that $\{\beta^v\}$, with $v = r,\ t,\ \hat{a}$ or $\check{a}$, in (16) are chosen to be independent of the nearest neighbor index m. Ideally, q^v in (13) or $[q_1^v, q_2^v, \ldots, q_K^v]^T$ in (14) should follow the zero-mean multivariate normal distribution (i.e. $q^v \sim N(0, C^v)$).

III. RAY-BASED PROPAGATION SIMULATORS IN SUBURBAN AND MOUNTAINOUS ENVIRONMENTS

Ideally, one would like to use data obtained through rigorous experiments to establish the models for measurement differences in (13). However, it is costly and time consuming to conduct meaningful and reliable experiments in any large region of practical interests. Therefore, people usually use simulated data to conduct preliminary studies on propagation modelings, algorithm validations, system developments, etc., for practical localization applications. In this paper, to show that the models for measurement differences proposed in the previous section are very general and flexible, we will use our previously developed ray-based propagation simulators to generate RSS, TOA and DOA data in two very different environments: a suburban region and a mountainous region.

Note that ray-based algorithms can be divided in two main groups: Ray Tracing and Ray Launching. In ray tracing algorithms, the exact positions of both the transmitter and the receiver are taken into account from the beginning. Thus rays (called eigenrays) satisfying geometrical optics rules for those specific locations are searched using various techniques. Searching for these eigenrays becomes very difficult and computational expensive when the environment is complex. Fortunately, for special cases where rays predominantly lie in the vertical plane containing the source and receiver, the efforts of tracing the ray can be greatly reduced. In ray launching algorithms, rays start from the transmitter with a pre-determined angular separation and are traced regardless of the receivers' position. These rays are ideally propagated by the algorithm along their trajectory until they encounter an obstacle, where they are reflected, diffracted, transmitted, or scattered. The subsequent rays are propagated following geometrical optics rules. Here, a space discretization is assumed that, in principle, limits field prediction accuracy. In summary, ray launching is more efficient, although theoretically less accurate, than ray tracing to perform wave propagation characterization over vast areas. However, when the vertical plane propagation is the dominant physical mechanism, ray tracing usually becomes more efficient than ray launching.

Here, a ray launching algorithm will be used for a suburban environment since the environment is very complicated and a ray tracing algorithm will be used for a mountainous environment since, in some cases, predominant rays lie in the vertical plane containing the transmitter and receiver.

A. RAY LAUNCHER FOR SUBURBAN ENVIRONMENTS

Consider an area of $3.3km \times 3.3km$ in a suburban region in Santa Clara, CA, as shown in Fig. 1. As the environment is very complex and predominant NLOS rays are from street corner diffraction, we will use our previously developed ray launcher [13] to generate the RSS, TOA and DOA data. The ray launcher is fully discrete in order to take advantage of environment preprocessing to efficiently trace rays undergoing both specular and diffuse interactions. Explicitly, the

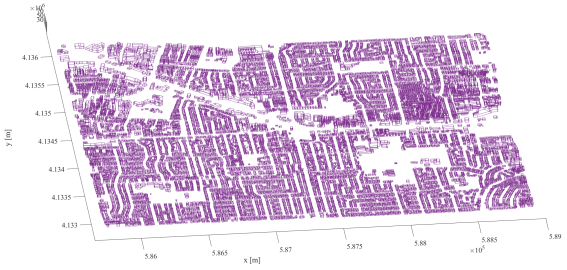

FIGURE 1. 3D view of a Suburban Environment in Santa Clara.

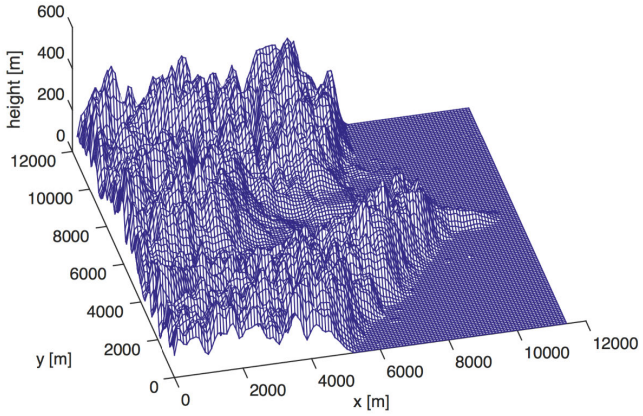

FIGURE 2. 3D view of a Mountainous Environment in Caspian Region.

environment is discretized into simple regular shapes to facilitate faster geometric computations. In addition to seamless space tessellation, other advanced features have been implemented to achieve a very high accuracy while drastically reducing computation time, as listed below.

1) Environment discretization. Environment surfaces (such as walls and streets) are discretized into ''tiles'' to simplify the ray launching algorithm and perform basic operations such as beam cross-section cutting across obstacles' edges without complex polygon clipping algorithms.

2) Environment-driven ray launching. Efficiency has been improved by launching rays on the base of the geometrical distribution of the obstacles (tiles) present in the environment, instead of using a constant angular discretization, and is particularly useful for outdoor application where obstacles are sparse or there are large open sky sectors where launching many rays would be useless.

3) Visibility preprocessing. Since the environment is simplified into a set of tiles, the potential existence of a propagation path between a generic pair of tiles (i.e. a ''visibility relation'') can be pre-computed for a given environment and saved in a file (the ''visibility matrix''), thus greatly simplifying and speeding up the computation of multiple-bounce rays.

4) Parallelization on a Graphical Processing Unit (GPU). Since the ray launching approach is inherently fit to parallel computing, the whole algorithm, including visibility preprocessing, has been parallelized, thus further reducing computation time with respect to traditional implementations.

B. RAY TRACER FOR MOUNTAINOUS ENVIRONMENTS

Consider a $11.4km \times 11.4km$ area in the Caspian region terrain database (see Fig. 2). The RSS, TOA, and DOA data are generated using our previously developed ray tracer [54]. For any arbitrary transmitter and receiver pair, this simulator predicts the propagation characteristics of rays traveling in the vertical plane containing the transmitter and receiver and terrain scattered rays traveling outside of this plane.

To predict the rays traveling between a transmitter and receiver pair, the simulator accounts for the terrain and its undulations in the vicinity of the transmitter and receiver. The terrain information is inputted to the simulator in the form of a terrain database. This terrain database is conceptually a matrix of terrain elevation samples over a uniformly spaced rectangular grid. The grid spacing between samples varies depending on the terrain data source.

For rays traveling in the vertical plane between the transmitter and receiver, the simulator feeds the vertical plane terrain profile to the Terrain-Integrated Rough Earth Model (TIREM) [58] to predict the propagation characteristics of the vertical plane rays. Note that depending on the terrain geometry, the TIREM model will use different prediction methods. For example, when the transmitter has LOS to the receiver, the TIREM model simply accounts for the direct and ground reflected rays. While, when the transmitter has NLOS to the receiver, TIREM uses either the Bullington method or the Epstein-Peterson multiple diffraction formulas to predict the vertical plane ray's propagation characteristics. In both LOS and NLOS cases, rays scattered from the troposphere and surface waves are also accounted for. Parameters used for TIREM in our simulations are: Earth's surface conductivity $= 0.01 S/m$; Surface humidity near the antennas $= 7.5\ g/m^3$; Surface refractivity $= 289$; Relative permittivity of earth's surface $= 15$; Polarization is VV (vertical to vertical polarized).

For rays that travel outside of the vertical plane containing the transmitter and receiver, the rural channel simulator accounts for terrain scattered rays. These rays are especially significant when the transmitter has NLOS to the receiver in mountainous or hilly environments. In these cases, the vertical plane rays are expected to be very weak after diffracting over the mountains or hills between the transmitter and receiver. The simulator is able to account for single bounce terrain scattered rays. Multiple terrain scattered rays are not accounted for because they are assumed to be relatively weak due to the high loss nature of scattering. To find the single bounced rays, the surface of the terrain is first partitioned into triangles. The triangles which are co-visible to both the transmitter and receiver are then used for single bounce scattering. The amplitude of the scattered ray is computed using the bi-static scattering equation with a Lambertian scattering coefficient.

IV. MODELS FOR MEASUREMENT DIFFERENCES IN SUBURBAN AND MOUNTAINOUS ENVIRONMENTS

In this section, we will use the RSS, TOA and DOA data generated using ray-based propagation simulators discussed in Section III to establish the models for measurement differences (see (13)-(16)) in a Santa Clara suburban region and a Caspian mountainous region. Specifically, we will use these data to validate the models proposed in (13) by finding the parameters β^v and α_m^v in (15) and the covariance matrix C^v of q^v in (13).

A. MODELING A SANTA CLARA SUBURBAN ENVIRONMENT

The 2D view of the Santa Clara environment in Fig. 1 is shown in Fig. 3 where 6 anchors and 3000 receivers are also shown. In each simulation, we randomly choose 600 receivers as nodes of interest (where the remaining 2400 receivers serve as reference nodes) to develop our models.

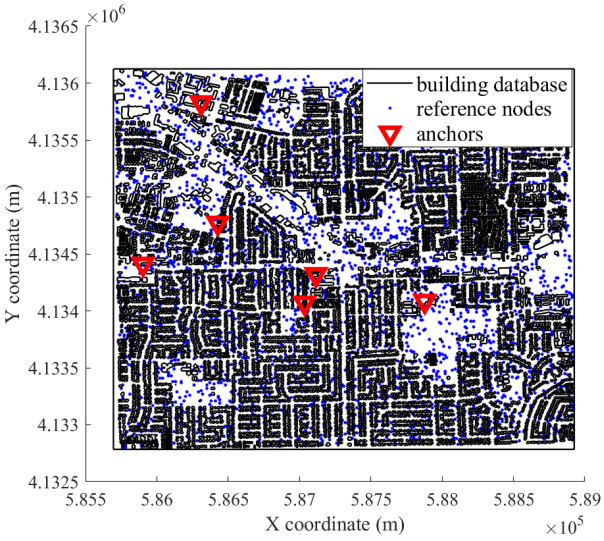

FIGURE 3. 2D view of the Santa Clara suburban environment in Fig. 1.

Let $K = 4$. One typical chosen node of interest at R and its K true nearest neighbors at $R^{(m)}$, $m = 1, 2, 3, 4$, are shown in the enlarged figure (see Fig. 4) as an example. On average, the distance between the first true nearest neighbor and the chosen node is $29.1m$. The distance between the second true nearest neighbor and the chosen node is $43.6m$. The distance between the third true nearest neighbor and the chosen node is $54.2m$. Finally, the distance between the forth true nearest neighbor and the chosen node is $63.9m$.

With given β^v and α_m^v for a data type v, a realization of the random variable q_m^v, the modeling error in (14), for a receiver at R, its m^{th} nearest neighbor at $R^{(m)}$, and a transmitter at T is re-written as

$$q_m^v(T, R) = \delta v_m(T, R) - \mu_m^v(T, R). \tag{17}$$

Note that all factors in (17) depend on locations T, R, and $R^{(m)}$. However, only the dependence on the locations T and R is explicitly indicated where the dependence on $R^{(m)}$ is hidden, because $R^{(m)}$ is determined when T and R are specified.

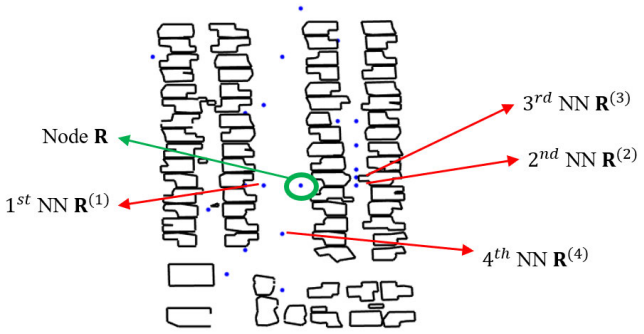

FIGURE 4. Illustration of 4 true nearest neighbors at $R^{(m)}$, $m = 1, 2, 3, 4$, of a given unknown node at R.

The optimum β^v and α_m^v in (15) can be found by minimizing the sum of squared modeling error over all possible $\{T, R\}$ pairs:

$$(\beta^v, \alpha_m^v) = \arg\min \sum_R \sum_T |q_m^v(T, R)|^2 \tag{18}$$

Equation (18) can be solved by the least square curve fitting. Recall that μ_m^v in (15) is a linear function of η_m^v. Thus, the solution in (18) can be obtained by linear regression which fits a line (μ_m^v as a linear function of η_m^v) to the simulated data (δv_m as a random function of η_m^v). This process will be demonstrated below for various data types $\{v\}$.

1) RSS DIFFERENCE MODEL

For RSS, as the data type, $v = r$. With $m = 1$ as an example, the curve fitting result using the optimum β^r and α_m^r is illustrated in Fig. 5. Note that the linear regression model is typically employed for modeling path loss. Therefore, RSS difference is modeled by linear regression here.

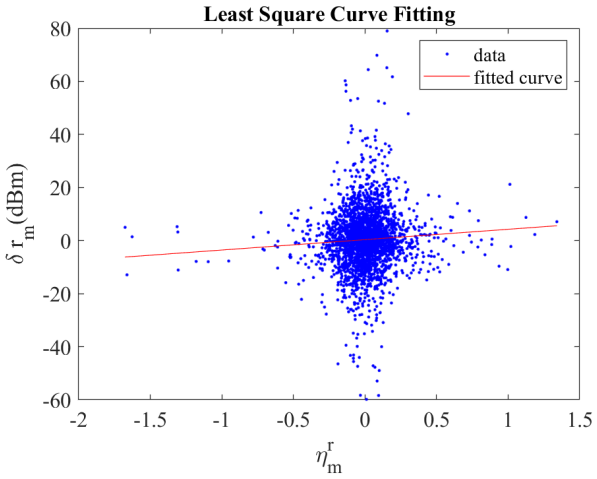

FIGURE 5. Linear regression to find optimum β^r and α_m^r for $m = 1$.

Following (18), the optimum β^r and α_m^r are given as:

$$\beta^r = -3.1$$
$$[\alpha_1^r, \alpha_2^r, \alpha_3^r, \alpha_4^r] = [0.88,\ 0.35, -0.18,\ 0.27]$$

With the optimum β^r and α_m^r given above, the realizations of the random variable q_m^r in (17) with $m \in \{1, \cdots, 4\}$ can be obtained and the covariance matrix of q^r can then be found:

$$C^r = \begin{bmatrix} 5.3 & 6.9 & 6.3 & 8.0 \\ 6.9 & 41.0 & 20.3 & 23.1 \\ 6.3 & 20.3 & 50.4 & 27.1 \\ 8.0 & 23.1 & 27.1 & 75.7 \end{bmatrix}$$

Fig. 6 shows the histograms of q_m^r in (17) with $m \in \{1, \cdots, 4\}$. where each q_m^r can be approximately considered as a Gaussian random variable. Note that with the introduction of α_m^r, each q_m^r is made to be zero mean.

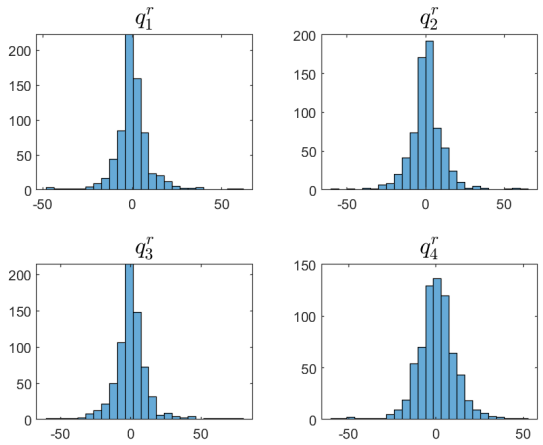

FIGURE 6. Histograms of q_m^r for $m = 1, 2, 3, 4$.

Compared with the widely used RSS path loss model, the variances for the RSS difference model shown above are smaller, since RSS differences are computed between closely located nodes.

2) TOA DIFFERENCE MODEL

For TOA as the data type, $v = t$. With $m = 1$ as an example, the curve fitting result using the optimum β^t and α_m^t is illustrated in Figure.7.

Following (18), the optimum β^t and α_m^t are given as:

$$\beta^t = 2$$
$$[\alpha_1^t, \alpha_2^t, \alpha_3^t, \alpha_4^t] = [27.2,\ 11.7,\ 20,\ 30]$$

With the optimum β^t and α_m^t given above, the realizations of the random variable q_m^t in (17) with $m \in \{1, \cdots, 4\}$ can be obtained and the covariance matrix of q^t can then be found:

$$C^t = \begin{bmatrix} 6.2 & 4.6 & 4.2 & 3.2 \\ 4.6 & 27 & 13 & 10 \\ 4.2 & 13 & 33 & 12 \\ 3.2 & 10 & 12 & 30 \end{bmatrix} \times 10^3$$

Fig. 8 shows the histograms of q_m^t in (17) with $m \in \{1, \cdots, 4\}$. where each q_m^t can be approximately considered as a Gaussian random variable. Note that with the introduction of α_m^t, each q_m^t is made to be zero mean.

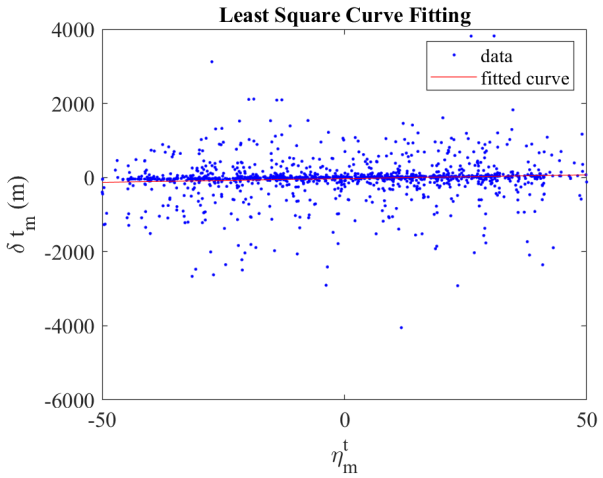

FIGURE 7. Linear regression to find optimum β^t and α_m^t for $m = 1$.

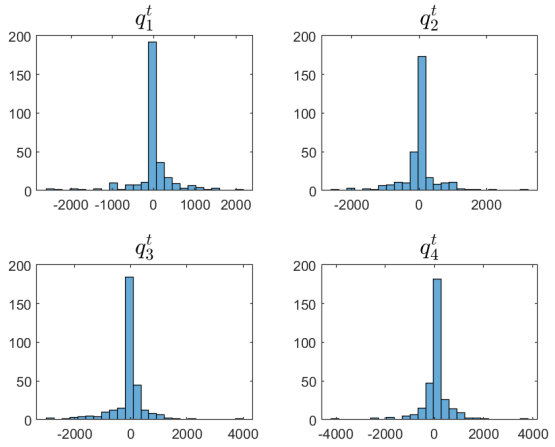

FIGURE 8. Histograms of q_m^t for $m = 1, 2, 3, 4$.

3) DOA DIFFERENCE MODEL

For 2-D cases, DOA difference Model is built in two dimensions. Thus, for DOA as the data type, both $v = \hat{a}$ and $v = \check{a}$ need to be considered. With $m = 1$ as an example, the curve fitting results using optimum ($\beta^{\hat{a}}$,$\alpha_m^{\hat{a}}$) and ($\beta^{\check{a}}$,$\alpha_m^{\check{a}}$) are illustrated in Fig.9. For simplicity, we use linear regression where the resulting remainders $q_m^{\hat{a}}$ and $q_m^{\check{a}}$ in (17) can be approximated as Gaussian random variables. This linear model of DOA difference is based on the fact that LOS propagation typically occurs between nearest neighbors, even though NLOS propagation typically characterizes the transmission from the transmitter to receivers.

Following (18), the optimum ($\beta^{\hat{a}}$,$\alpha_m^{\hat{a}}$) and ($\beta^{\check{a}}$,$\alpha_m^{\check{a}}$) are given as:

$$\beta^{\hat{a}} = 0.59,\ \beta^{\check{a}} = 0.6$$
$$[\hat{\alpha}_1^{\hat{a}}, \hat{\alpha}_2^{\hat{a}}, \hat{\alpha}_3^{\hat{a}}, \hat{\alpha}_4^{\hat{a}}] = [5,\ 3, -6,\ 5] \times 10^{-3}$$
$$[\hat{\alpha}_m^{\hat{a}}, \hat{\alpha}_m^{\hat{a}}, \hat{\alpha}_m^{\hat{a}}, \hat{\alpha}_m^{\hat{a}}] = [9,\ 7,\ 3, -25] \times 10^{-3}$$

With the optimum ($\beta^{\hat{a}}$,$\alpha_m^{\hat{a}}$) and ($\beta^{\check{a}}$,$\alpha_m^{\check{a}}$) given above, the realizations of the random variables $q_m^{\hat{a}}$ and $q_m^{\check{a}}$ in (17) with

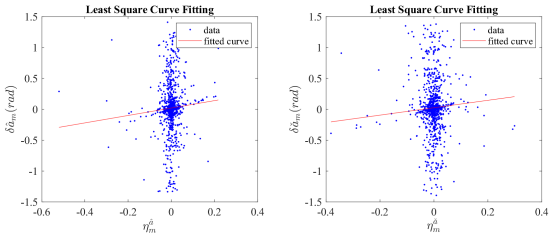

FIGURE 9. Linear regression to find optimum $(\beta^{\hat{a}}, \alpha_m^{\hat{a}})$ and $(\beta^{\check{a}}, \alpha_m^{\check{a}})$ for $m = 1$.

$m \in \{1, \cdots, 4\}$ can be obtained and their covariance matrices are

$$C^{\hat{a}} = \begin{bmatrix} 4.2 & 2.6 & 1.9 & 3.4 \\ 2.6 & 16.1 & 5.7 & 7.6 \\ 1.9 & 5.7 & 17.4 & 6.4 \\ 3.4 & 7.6 & 6.4 & 27.6 \end{bmatrix} \times 10^{-3}$$

and

$$C^{\check{a}} = \begin{bmatrix} 15.6 & 7.3 & 7.0 & 9.2 \\ 7.3 & 15.9 & 7.5 & 7.2 \\ 7.0 & 7.5 & 20.2 & 10.3 \\ 9.2 & 7.2 & 10.3 & 29.9 \end{bmatrix} \times 10^{-3}$$

Fig. 10 shows the histograms of $q_m^{\hat{a}}$ and $q_m^{\check{a}}$ in (17) with $m \in \{1, \cdots, 4\}$. Here each $q_m^{\hat{a}}$ or $q_m^{\check{a}}$ can be approximately considered as a Gaussian random variable. Note that with the introduction of $\alpha_m^{\hat{a}}$ and $\alpha_m^{\check{a}}$, $q_m^{\hat{a}}$ and $q_m^{\check{a}}$ are made to be zero mean.

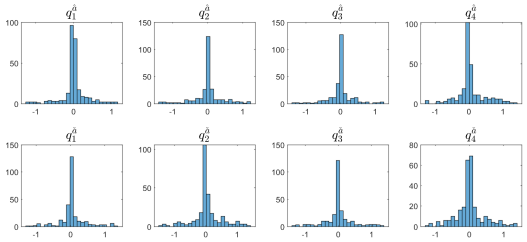

FIGURE 10. Histograms of $q_m^{\hat{a}}$ and $q_m^{\check{a}}$ for $m = 1, 2, 3, 4$.

B. MODELING A CASPIAN MOUNTAINOUS ENVIRONMENT

The 2D view of the Caspian region in Fig. 2 is shown in Fig. 11, where 6 anchors and 2000 receivers are also shown. We randomly choose 400 receivers as nodes of interest (where the remaining 1600 receivers serve as reference nodes) to develop our models.

Let $K = 4$. One typical chosen node of interest at R and its K true nearest neighbors at $R^{(m)}$, $m = 1, 2, 3, 4$, are shown in the enlarged figure (see Fig. 12) as an example. On average, the distance between the first true nearest neighbor and the chosen node is $102m$. The distance between the second true nearest neighbor and the chosen node is $164m$. The distance between the third true nearest neighbor and the chosen node

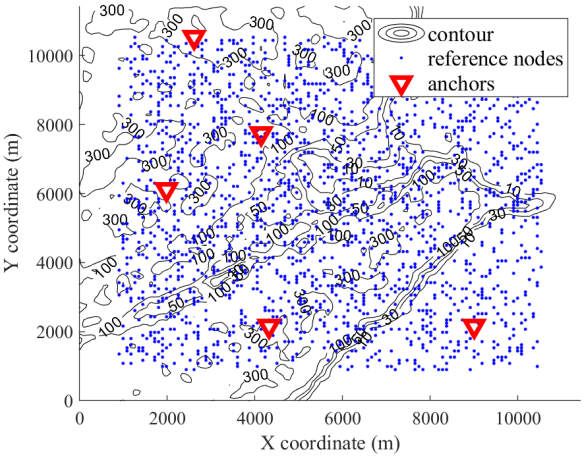

FIGURE 11. 2D view of the Caspian mountainous environment in Fig.2.

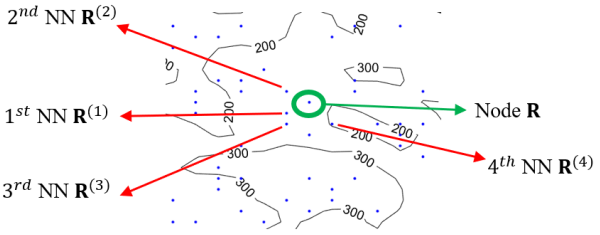

FIGURE 12. Illustration of $K = 4$ nearest neighbors of a given node at R.

is $203m$. Finally, the distance between the forth true nearest neighbor and the chosen node is $239m$.

Following the same procedure as shown in the previous section, we can obtain our measurement difference models in this Caspian mountainous environment. As the linear regression process employed for the Caspian mountainous environment is exactly the same as that for the Santa Clara suburban region, we will omit the showing of least square curve fitting and histogram figures here. Instead, optimum model parameters β^v and α_m^v and covariance matrix of q^v for each data type will be listed here. These numbers will be used in generating CRLBs in the Section V.

1) RSS DIFFERENCE MODEL

Following (18), the optimum β^r and α_m^r are given as:

$$\beta^r = -4.5$$
$$[\alpha_1^r, \alpha_2^r, \alpha_3^r, \alpha_4^r] = [0.88, -0.32, -0.24, 0.31]$$

With the optimum β^r and α_m^r given above, the realizations of the random variable q_m^r in (17) with $m \in \{1, \cdots, 4\}$ can be obtained and the covariance matrix of q^r can then be found:

$$C^r = \begin{bmatrix} 5.3 & 4.2 & 5.6 & 4.5 \\ 4.2 & 18.5 & 11.4 & 13.4 \\ 5.6 & 11.4 & 35.9 & 19.1 \\ 4.5 & 13.4 & 19.1 & 42.3 \end{bmatrix}$$

2) TOA DIFFERENCE MODEL

Following (18), the optimum β^t and α_m^t are given as:

$$\beta^t = 1.02$$
$$[\alpha_1^t, \alpha_2^t, \alpha_3^t, \alpha_4^t] = [13.6, 23.8, 19.2, 26.4]$$

With the optimum β^t and α_m^t given above, the realizations of the random variable q_m^t in (17) with $m \in \{1, \cdots, 4\}$ can be obtained and the covariance matrix of q^t can then be found:

$$C^t = \begin{bmatrix} 13.6 & 6.4 & 6.6 & 5.9 \\ 6.4 & 14.5 & 3.8 & 6.0 \\ 6.6 & 3.8 & 31.8 & 10.2 \\ 5.9 & 6.0 & 10.2 & 17.5 \end{bmatrix} \times 10^3$$

3) DOA DIFFERENCE MODEL

Following (18), the optimum ($\beta^{\hat{a}}$,$\alpha_m^{\hat{a}}$) and ($\beta^{\check{a}}$,$\alpha_m^{\check{a}}$) are given as:

$$\beta^{\hat{a}} = 1.01, \beta^{\check{a}} = 0.991$$
$$[\alpha_1^{\hat{a}}, \alpha_2^{\hat{a}}, \alpha_3^{\hat{a}}, \alpha_4^{\hat{a}}] = [8.7, 8, 7.6, 4.2] \times 10^{-3}$$
$$[\alpha_m^{\check{a}}, \alpha_m^{\check{a}}, \alpha_m^{\check{a}}, \alpha_m^{\check{a}}] = [0.4, 7, 10, 4] \times 10^{-3}$$

With the optimum ($\beta^{\hat{a}}$,$\alpha_m^{\hat{a}}$) and ($\beta^{\check{a}}$,$\alpha_m^{\check{a}}$) given above, the realizations of the random variables $q_m^{\hat{a}}$ and $q_m^{\check{a}}$ in (17) with $m \in \{1, \cdots, 4\}$ can be obtained and their covariance matrices are

$$C^{\hat{a}} = \begin{bmatrix} 4.0 & 1.9 & 2.5 & 1.8 \\ 1.9 & 3.2 & 3.0 & 2.2 \\ 2.5 & 3.0 & 6.6 & 2.9 \\ 1.8 & 2.2 & 2.9 & 3.7 \end{bmatrix} \times 10^{-4}$$

and

$$C^{\check{a}} = \begin{bmatrix} 16 & 8.9 & 8.7 & 8.8 \\ 8.9 & 16 & 12 & 9.6 \\ 8.7 & 12 & 22 & 11 \\ 8.8 & 9.6 & 11 & 20 \end{bmatrix} \times 10^{-4}$$

V. NUMERICAL RESULTS OF FINGERPRINTING LOCALIZATION

We will use the CRLB in (9) as a benchmark to evaluate two fingerprinting localization approaches, KNN and WKNN, with different data fusion options of RSS, TOA, and DOA in both Santa Clara and Caspian regions, shown in Fig. 1 and Fig. 2, respectively.

A. SANTA CLARA SUBURBAN REGION

We use the simulation setup in Section IV.A to perform 10 fingerprinting localization simulations here. Consider the six anchor nodes and 3000 receivers as shown in Fig. 3. In each simulation, we randomly choose 600 receivers as unknown nodes and treat the remaining 2400 receivers as reference nodes. The estimated locations of the 400 unknown nodes can be obtained by KNN and WKKN using (5) with their corresponding weights, respectively. Then, the root mean square errors (RMSEs) of the location estimations of these 400 unknown nodes can be calculated for both KNN and WKNN approaches, respectively. The means, over 10 simulations, of the RMSEs for various data types and fusion options using KNN and WKNN, respectively, are listed in Table 1. Also listed in Table 1 are the mean CRLBs which are obtained by averaging firstly over the 400 unknown nodes in each simulation and then over the 10 simulations for various single-data types and fusion options.

TABLE 1. Fingerprinting localization results (Santa Clara Suburban Region).

Measurement	KNN(m)	WKNN(m)	CRLB(m)
RSS	654	654	262.4
TOA	273	243	45.0
DOA	731	689	125.2
RSS/TOA	272	240	43.1
RSS/DOA	486	475	70.2
TOA/DOA	262	240	31.8
RSS/TOA/DOA	255	232	30.8

Listed below are a few key observations made from Table 1. At first, let us focus on performance evaluations of the KNN and WKNN localization approaches. In this example, we observe that KNN and WKNN have similar performances. For same data type or fusion option, WKNN result is just a little bit better than the corresponding KNN result. However, both KNN and WKNN estimation errors are much higher than the corresponding CRLB for each single-data type or fusion option. Because some of the ''nearest'' neighbors determined using (3) are, in fact, very faraway from the unknown node. (Note that the average distance between an unknown node to its K actual nearest neighbors is around 48 meters.) So the average or weighted average of the locations of ''nearest'' neighbors can be very different from the actual location of the unknown node. As the gaps between KNN/WKNN results and CRLBs are huge, there are many rooms for improvement on the KNN/WKNN fingerprinting approaches. In addition, significant improvements should come from a better strategy on choosing ''nearest'' neighbors rather than a better strategy on weight design, because the performance of KNN/WKNN can be greatly improved if one can detect and remove fake ''nearest'' neighbors,

Secondly, we discuss the results using different single-data types (i.e., without data fusion). Note that RSS results (including KNN, WKNN, and CRLB) are inferior due to strong spatial fading in the complex NLOS Santa Clara suburban environment. However, TOA results (including KNN, WKNN and CRLB) are significantly better than the RSS results even in the same NLOS suburban environments. Because TOA information is primarily determined by the dominant ray, which has little fading effect. Although DOA information is also primarily determined by the dominant ray, which has little fading effect, it is not surprising that the KNN/WKNN DOA results are more unsatisfactory than the KNN/WKNN RSS results. Because the dominant rays are mainly propagating along streets. Thus, the street orientation where an unknown node is located will dictate the DOA information of that node. Since the number of street orientations

is very small in the environment considered (see Fig. 3), the variations of DOAs for different nodes are limited. Therefore, it is challenging to find correct ''nearest'' neighbors using DOA in this kind of environments. However, contrary to the KNN/WKNN results, the CRLB of DOA is much better (i.e., smaller) than the CRLB of RSS. Because we use true nearest neighbors in formulating the CRLBs in Section IV and the dominant ray has little fading effect. One last remark is that the CRLB of DOA is larger than that of TOA. Because DOAs have less variations than TOAs as mentioned before.

Thirdly, we examine the results using different fusion options. It is clear from (9) and (10) that more data types employed should improve the localization performance. Thus, RSS/TOA/DOA has the best performance, RSS/TOA is better than RSS or TOA, RSS/DOA is better than RSS or TOA, and TOA/DOA is better than TOA or DOA. When the performances of two options are comparable, the fusion of these two will significantly improve the performance. This can be seen by comparing RSS/DOA with RSS or DOA. When the performance of two options are very different, the fusion of these two will improve little. This can be seen by comparing RSS/TOA with TOA. It can also be seen by comparing TOA/DOA with TOA or comparing RSS/TOA/DOA with TOA/DOA. In summary, the fusion of three different measurements (RSS/TOA/DOA) provides the best localization result which shows the advantage of data fusion. However, comparing RSS/TOA/DOA with TOA, TOA/DOA or RSS/TOA, the improvement gained by data fusion is not significant because the position information obtained from RSS and DOA measurements is limited. Thus, in practical applications, there should be a trade-off between the accuracy improvement and equipment complexity.

Table 1 shows only the mean results. Sometimes, it is more informative showing the cumulative distribution function (CDF) of the localization error than just showing the mean. Thus, we use one of the 10 simulations and define the CDF of the location errors of 400 unknown nodes as:

$$\mathrm{CDF}(\phi) = \mathrm{Prob}(||R_k - f_2||_2 \leq \phi) \tag{19}$$

Fig. 13 shows the CDF results with different single-data types and fusion options for KNN/WKNN. All discussions about Table 1 are applicable here.

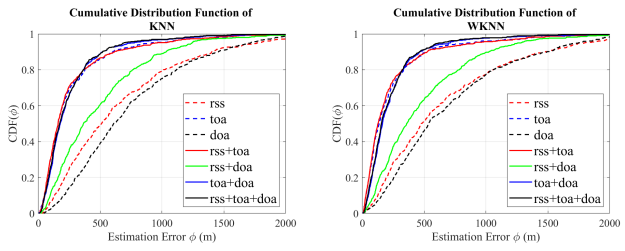

FIGURE 13. CDFs of position estimation errors (Santa Clara Suburban Region).

B. CASPIAN MOUNTAINOUS REGION

We use the simulation setup in Section IV.B to perform 10 fingerprinting localization simulations here. Consider the six anchor nodes and 2000 receivers as shown in Fig. 11. In each simulation, we randomly choose 400 receivers as unknown nodes and treat the remaining 1600 receivers as reference nodes. The estimated locations of the 400 unknown nodes can be obtained by KNN and WKKN using (5) with their corresponding weights, respectively. Then, the root mean square errors (RMSEs) of the location estimations of these 400 unknown nodes can be calculated for both KNN and WKNN approaches, respectively. The means, over 10 simulations, of the RMSEs for various data types and fusion options using KNN and WKNN, respectively, are listed in Table 2. Also listed in Table 2 are the mean CRLBs which are obtained by averaging firstly over the 400 unknown nodes in each simulation and then over the 10 simulations for various single-data types and fusion options.

TABLE 2. Fingerprinting localization results (Caspian Mountainous Region).

Measurement	KNN(m)	WKNN(m)	CRLB(m)
RSS	963	647	433.3
TOA	276	120	79.6
DOA	354	172	80.2
RSS/TOA	276	120	76.6
RSS/DOA	343	171	71.8
TOA/DOA	238	119	49.1
RSS/TOA/DOA	238	118	47.3

Results in Table 2 and those in Table 1 are somewhat different because the dominant ray for many receivers in the Caspian region is primarily due to diffraction over the mountain top while the dominant ray for many receivers in the Santa Clara region is primarily due to diffraction around the street corners. Thus, the DOA information of NLOS rays in the Caspian region is more like the DOA information of the LOS ray in 2D plane, which is very useful for fingerprinting localization. However, the DOA information of NLOS rays in Santa Clara region is more like the orientations of the streets, which is not very useful for fingerprinting localization. Listed below are a few key observations made from Table 2.

At first, let us focus on performance evaluations of the KNN and WKNN localization approaches. Different from Table 1, WKNN is significantly better than KNN because of the LOS-like propagation mechanism in this mountainous region. Similar to (but not as bad as) Table 1, both KNN and WKNN estimation errors are higher than the corresponding CRLB for each single-data type or fusion option. This is mainly because some of the ''nearest'' neighbors determined using (3) can be far from the unknown node. Nevertheless, this is not as bad as that in Santa Clara suburban region because of the over-mountain-top diffraction feature of the dominant ray mentioned above. (Note that the average distance between an unknown node to its K actual nearest neighbors is around 177 meters.)

Secondly, we discuss the results using different single-data types (i.e., without data fusion). Same as Table 1, RSS results are deteriorated due to spatial fading, but TOA results are significantly better than the RSS results (due to that TOA is primarily a feature of the dominant ray). Different from Table 1, DOA results are much better than RSS results because of the over-the-mountain-top diffraction feature of the dominant ray mentioned above. Furthermore, the KNN/WKNN DOA results are still poorer than the KNN/WKNN TOA results, which is mainly due to the mode-2π feature of the DOAs. However, the DOA CRLB is almost as good as the TOA CRLB which is also due to the over-the-mountain-top diffraction feature of the dominant ray.

Thirdly, we examine the results using different fusion options. It is clear from (9) and (10) that more data types employed should improve the localization performance. This indeed can be seen in the CRLB results, but is not profound in KNN/WKNN results. Again, this shows that there are rooms for improvement for KNN/WKNN.

For different single-data types and fusion options, Fig. 14 shows the CDF (see (19)) of the location errors of 400 unknown nodes in one of the 10 simulation using KNN/WKNN for the Caspian mountain region. All discussions about Table 2 are applicable here.

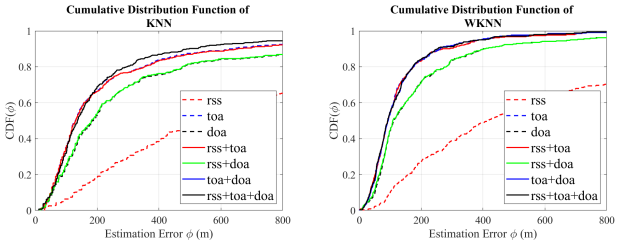

FIGURE 14. CDFs of position estimation errors (Caspian Mountainous Region).

C. EFFECTS OF REFERENCE NODE DENSITY AND MEASUREMENT ERROR ON CRLB

Using the Caspian mountainous region shown in Fig. 11 as an example, we will at first demonstrate how to improve the accuracy of fingerprinting approaches by increasing the density of reference nodes. Secondly, we will demonstrate the effects of measurement errors on CRLB.

In Fig. 15, we compare the average CRLBs for all data types and fusion options for different node densities. The average distance between two adjacent reference nodes is also plotted in Fig. 15. As the number of reference nodes increases from 500 to 5000 in the region, the average CRLBs decreases. However, the decreasing rates of the average CRLBs become less steep as the number of reference nodes approaches 5000. This shows that the achievable average CRLBs are limited by measurement errors when the node density is high. Note that the "measurement errors" here are generated by the limited accuracy of our channel simulator.

In Fig. 16, we show the CRLB contour plot with 500 reference nodes without additionally added

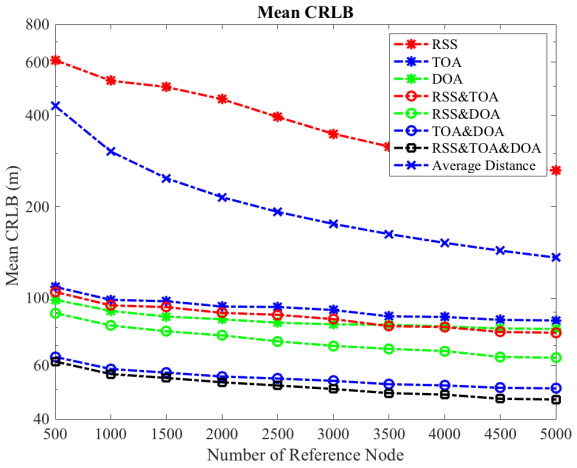

FIGURE 15. Mean CRLB with different numbers of reference nodes.

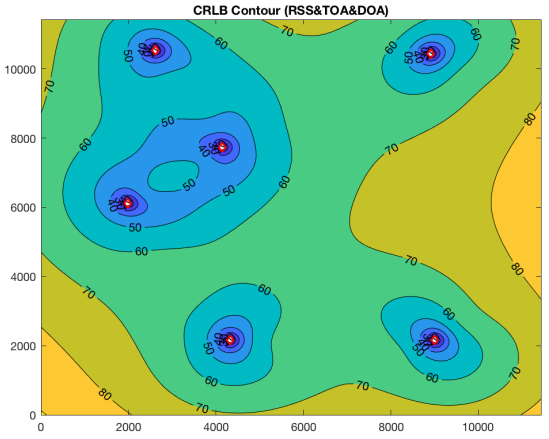

FIGURE 16. CRLB contour plot with 500 reference nodes (without additional measurement errors).

measurement errors. However, as mentioned before, the simulation data does contain simulation errors due to the finite grid size defined in our simulator. With additionally added measurement errors, we show the CRLB contour plot with 500 reference nodes in Fig. 17 and the CRLB contour plot with 5000 reference nodes in Fig. 18. Specifically, in Figs. 17 and 18, the additional RSS measurement error is modeled as a zero mean Gaussian random variable where the standard deviation is $5.8 \times 10^{-12} mW$; the additional TOA measurement error is modeled as a zero mean Gaussian random variable where the standard deviation is $5.7ns$; and the additional DOA measurement error is modeled as a uniform distributed random variable, ranging from $-5°$ to $5°$. For convenience, we show the fusion results of RSS, TOA and DOA only. Similar conclusions can be made from other single data-type and fusion options.

From Figs. 16, 17 and 18, one can see that the areas closer to any of the six anchors show smaller CRLBs. The areas surrounded by multiple anchors have smaller CRLBs comparing to areas near the four edges or four corners of the region map. Comparing Fig. 16 with Fig. 17, CRLBs increase

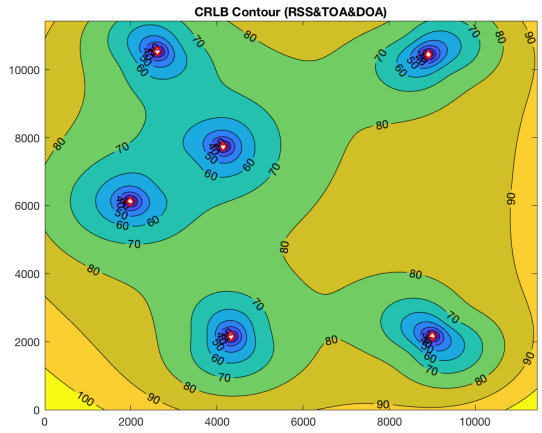

FIGURE 17. CRLB contour plot with 500 reference nodes (with additional measurement errors).

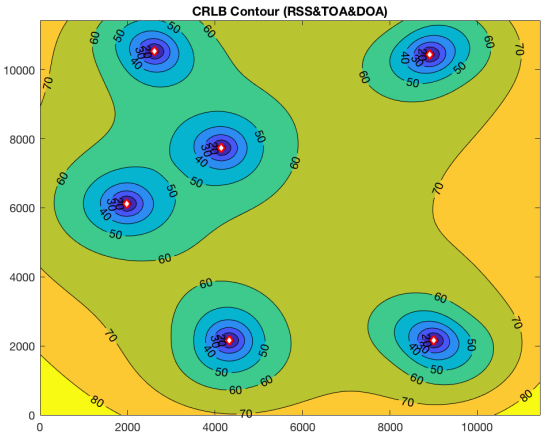

FIGURE 18. CRLB contour plot with 5000 reference nodes (with additional measurement errors).

with the measurement errors. However, the CRLBs are less affected by the measurement errors in the areas near any of the six anchors. Comparing Fig. 17 with Fig. 18, CRLBs decrease as the node density increases. It is remarkable that the CRLB contours in Fig. 16 are similar to those in Fig. 18. Higher node densities show diversity-like gains to mitigate the effects of measurement errors.

D. EFFECTS OF CELLULAR ID ON FINGERPRINTING LOCALIZATION ACCURACY

Using the Santa Clara suburban region shown in Fig. 3 as an example, we will demonstrate how to improve the accuracy of fingerprinting approaches by increasing the number of anchors in this section. As mentioned before, a better strategy for choosing "nearest" neighbors is an effective way to improve performance. Here, we will show that better estimates of the "nearest" neighbors can be achieved with more anchors. Consequently, as seen from (5), the location estimation errors will be reduced.

Consider 200 anchors and 3000 receivers randomly distributed in the Santa Clara suburban region considered in Fig. 3. Among the 3000 receivers, we randomly choose 600 as unknown nodes and treat the remaining 2400 as reference nodes. Based on the 200 anchors' locations, we divide the region in Fig. 3 into 200 cells to form a cellular system. In this case, the average cell radius is around 220 meters. Assuming that each unknown node is given a cell ID specifying which cell it is located, we can then use the cell ID to do an initial classification of the reference nodes before applying KNN/WKNN to find the location of an unknown node. Specifically, when we search for the "nearest" neighbors of this unknown node, only the reference nodes in the given cell and its 6 neighboring cells (see Fig. 19) can be potential candidates. In this manner, the risk to find a faraway fake "nearest" neighbor is removed.

To estimate the location of this unknown node, it is obvious that one would normally use the anchors in the given cell and its neighboring cells (see Fig. 19) for better results. Thus, we use the anchors in the six neighboring cells to perform the fingerprinting operation in this section. To derive the CRLB in (9), we have to use the new anchors and the K neighbors of each of the 600 unknown nodes to derive the measurement difference models in (15)-(17) first. To reduce the paper length, the modeling results will not be presented here as the modeling process is exactly the same as that employed in Section IV.

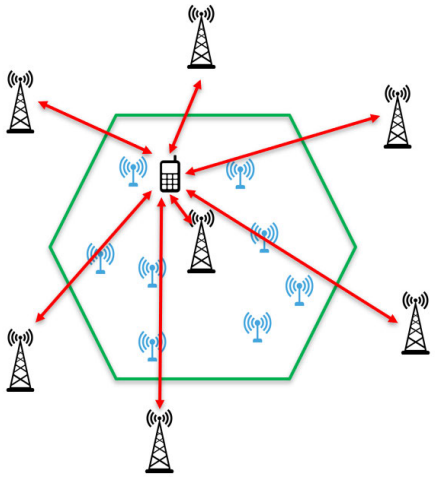

FIGURE 19. Localization with Cell ID.

As in Section V.A, 10 simulations (where the locations of 600 unknown nodes are estimated in each simulation) are performed to compute the CRLB and the mean RMSEs for various single-data types and fusion options. Table 3 shows the localization results and CRLB results when Cell ID is employed for the initial classification of the reference nodes. It can be seen that all values in Table 3 are much smaller than their corresponding values in Table 1. This is because the faraway fake "nearest" neighbors can no longer exist when cell ID is employed.

For different single-data types and fusion options, Fig. 20 shows the CDF (see (19)) of the location errors

TABLE 3. Fingerprinting localization results (with Cell ID, Santa Clara Suburban Region).

Measurement	KNN(m)	WKNN(m)	CRLB(m)
RSS	52.7	43.3	38.0
TOA	54.1	43.2	34.1
DOA	63.3	54.6	17.3
RSS/TOA	50.8	40.2	23.4
RSS/DOA	49.5	40.4	12.6
TOA/DOA	52.5	42.9	12.1
RSS/TOA/DOA	47.7	37.9	10.6

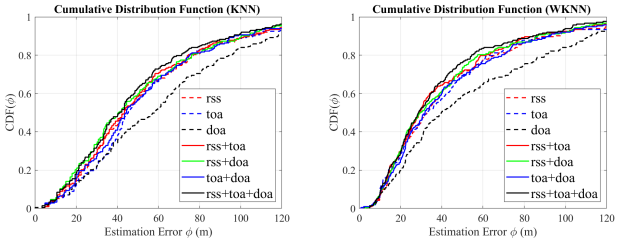

FIGURE 20. The CDFs of position estimation errors in Santa Clara Suburban Region with Cell ID.

of 600 unknown nodes in one simulation using KNN/WKNN. All discussions about Table 3 are applicable here.

VI. CONCLUSION

Since measurement data for reference nodes are provided for fingerprinting approaches, the true optimum location estimation and CRLB calculation should consider joint distributions of measurement data of the unknown node and its neighboring reference nodes. This poses a challenging problem in NLOS environments because joint statistics of measurement data at neighboring locations are usually non-Gaussian and unavailable.

Based on physical insights of NLOS propagation, we propose in this paper to use joint statistics of measurement differences (instead of joint statistics of measurement data directly) to circumvent the difficulty. As the ray trajectory of the dominant NLOS ray in between neighboring nodes is usually LOS, joint statistics of measurement differences are approximated as Gaussian. To show the robustness of our approaches, we use two very different NLOS environments (a suburban area of $3.3km \times 3.3km$ in Santa Clara, California, and a mountainous area of $11.4km \times 11.4km$ in the Caspian region) to validate our proposed models for RSS, TOA and DOA. As the two environments are very complicated, we use our previously developed ray launcher and ray tracers to generate propagation data.

With the measurement difference models at hand, CRLB is used as a benchmark to evaluate KNN and WKNN in the Santa Clara suburban and Caspian mountainous regions for different data types and various fusion options. Numerical results show that there is room for improvement for the KNN or WKNN approaches. One way to improve the localization performance is to take advantage of the geometrical relations of the neighboring nodes revealed by the measured TOAs and DOAs. By exploiting physical insights of propagation mechanisms, more intelligent feature fusion strategies for multi-feature fingerprinting localization will be developed in the future.

The proposed framework for CRLB analyses for NLOS environments are very general and flexible. With the aid of ray launcher and tracer to generate simulation data, the proposed CRLB analyses can be employed to evaluate fingerprinting systems with various designs (such as data type and fusion option) and different configurations (such as densities of reference nodes and numbers of anchor nodes) in diverse NLOS environments (such as suburban and mountainous regions).

Note that different environments have different propagation characteristics. For examples, the dominant NLOS propagator is over-the-mountain-top diffraction in mountainous regions. However, diffraction around street corners is the dominant NLOS propagator in dense suburban regions even though over-the-roof-top rays sometimes play important roles as well. Thus, the DOA data is very useful for localization in mountainous regions but is almost useless in suburban regions. These important points are concluded from the presented CRLB results. This shows that CRLB analyses can be very useful system design tools.

The hypothesis of jointly Gaussian of the measurement differences between a node of interest and its nearest neighbors is the key to simplify the CRLB derivation. This hypothesis has been verified using numerical simulations in this paper. But rigorous validations of the proposed models using comprehensive measured data need to be done in the future.

This is especially true for DOA because it may not be easy to satisfy the differential similarity in complicated environments. As we observe in our numerical simulations, differential similarity is difficult to be satisfied well in the suburban scenario due to its diffraction-by-street-corner and over-the-roof-top propagation characteristics. This is because that the DOA of diffraction-by-street-corner ray is determined by the street orientation while the DOA of over-the-roof-top ray is very close to the DOA of the LOS ray assuming no blockage between the transmitter and the receiver. When diffraction-by-street-corner rays are dominant, the far-apart receivers with the same street orientation will have differential similarity. This is not desirable. Furthermore, when two neighboring nodes have two different dominant propagation mechanisms, their DOAs may not satisfy the differential similarity. This is also not desirable. However, we find that differential similarity is satisfied well in the mountainous scenario due to its over-the-mountain-top propagation characteristics. This is because the DOA of over-the-mountain-top ray is very close to the DOA of the LOS ray assuming no blockage between the transmitter and the receiver. We will verify our models with measured data in the future.

Lastly, we have assumed that RSS and TOA are uncorrelated for simplicity. Since the two NLOS environments

considered in this paper are very complex, the correlation of RSS and TOA is indeed not significant. But this needs to be verified with measured data. Generally speaking, TOA is correlated to RSS in LOS environments. Specifically, a short TOA generally associates with a large RSS. But this is not always true in complex NLOS environments. For example, if the dominant ray at a receiver has to go through two or more diffractions, the RSS at this receiver will be very weak. It could be much weaker than the RSS at the other receiver with a longer TOA if the dominant ray of the second receiver goes through only one or no diffraction. Further studies on the impact on CRLB by the correlation between TOA and RSS will be carried out in the future.

APPENDIX A
DERIVATIONS OF FIM

Consider the fusion of RSS, TOA and DOA as an example. Following the general format in (7), the differences between the measurements at a given receiver and its k^{th} nearest neighboring receiver, with respect to the i^{th} transmitter, can be expressed as:

$$\begin{aligned}
\delta r_{k|i} &= \boldsymbol{\mu}^r_{k|i}(R_k) + q^r_{k|i} \\
\delta t_{k|i} &= \boldsymbol{\mu}^t_{k|i}(R_k) + q^t_{k|i} \\
\delta \hat{a}_{k|i} &= \boldsymbol{\mu}^{\hat{a}}_{k|i}(R_k) + q^{\hat{a}}_{k|i} \\
\delta \check{a}_{k|i} &= \boldsymbol{\mu}^{\check{a}}_{k|i}(R_k) + q^{\check{a}}_{k|i}
\end{aligned} \tag{20}$$

And the conditional probability density distributions of the measurement differences are:

$$\begin{aligned}
f^r_{k|i}(\delta r_{k|i}) &= \frac{e^{\left[-\frac{1}{2}(\delta r_{k|i}-\boldsymbol{\mu}^r_{k|i}(R_k))^T (C^r)^{-1}(\delta r_{k|i}-\boldsymbol{\mu}^r_{k|i}(R_k))\right]}}{\sqrt{(2\pi)^K |C^r|}} \\
f^t_{k|i}(\delta t_{k|i}) &= \frac{e^{\left[-\frac{1}{2}(\delta t_{k|i}-\boldsymbol{\mu}^t_{k|i}(R_k))^T (C^t)^{-1}(\delta t_{k|i}-\boldsymbol{\mu}^t_{k|i}(R_k))\right]}}{\sqrt{(2\pi)^K |C^t|}} \\
f^{\hat{a}}_{k|i}(\delta \hat{a}_{k|i}) &= \frac{e^{\left[-\frac{1}{2}(\delta \hat{a}_{k|i}-\boldsymbol{\mu}^{\hat{a}}_{k|i}(R_k))^T (C^{\hat{a}})^{-1}(\delta \hat{a}_{k|i}-\boldsymbol{\mu}^{\hat{a}}_{k|i}(R_k))\right]}}{\sqrt{(2\pi)^K |C^{\hat{a}}|}} \\
f^{\check{a}}_{k|i}(\delta \check{a}_{k|i}) &= \frac{e^{\left[-\frac{1}{2}(\delta \check{a}_{k|i}-\boldsymbol{\mu}^{\check{a}}_{k|i}(R_k))^T (C^{\check{a}})^{-1}(\delta \check{a}_{k|i}-\boldsymbol{\mu}^{\check{a}}_{k|i}(R_k))\right]}}{\sqrt{(2\pi)^K |C^{\check{a}}|}}
\end{aligned} \tag{21}$$

Here, we use C^v instead of $C^v_{k|i}$ because we assume the correlation matrix does not change with the receiver index k or transmitter index i. Assume different data types are collected independently and the measurements received from different transmitters are also independent. Then,

$$\begin{aligned}
&f_k(\delta r_k, \delta t_k, \delta \hat{a}_k, \delta \check{a}_k) \\
&\quad = \prod_{i=1}^{N_T} f_{k|i}(\delta r_{k|i}, \delta t_{k|i}, \delta \hat{a}_{k|i}, \delta \check{a}_{k|i}) \\
&\quad = \prod_{i=1}^{N_T} \{f^r_{k|i}(\delta r_{k|i}) f^t_{k|i}(\delta t_{k|i}) f^{\hat{a}}_{k|i}(\delta \hat{a}_{k|i}) f^{\check{a}}_{k|i}(\delta \check{a}_{k|i})\}
\end{aligned} \tag{22}$$

Finally we can get the xx element of the fisher information matrix in (11) for feature fusion:

$$\begin{aligned}
&[I_k(R_k)]_{xx} \\
&\quad = -E\{\frac{\partial^2 \ln f_k(\delta r_k, \delta t_k, \delta \hat{a}_k, \delta \check{a}_k)}{\partial x^2}\} \\
&\quad = \sum_{i=1}^{N_T} \left\{ \left[\frac{\partial \boldsymbol{\mu}^r_{k,i}(R_k)}{\partial x}\right]^T (C^r)^{-1} \left[\frac{\partial \boldsymbol{\mu}^r_{k|i}(R_k)}{\partial x}\right] \right. \\
&\qquad + \left[\frac{\partial \boldsymbol{\mu}^t_{k|i}(R_k)}{\partial x}\right]^T (C^t)^{-1} \left[\frac{\partial \boldsymbol{\mu}^t_{k|i}(R_k)}{\partial x}\right] \\
&\qquad + \left[\frac{\partial \boldsymbol{\mu}^{\hat{a}}_{k|i}(R_k)}{\partial x}\right]^T (C^{\hat{a}})^{-1} \left[\frac{\partial \boldsymbol{\mu}^{\hat{a}}_{k|i}(R_k)}{\partial x}\right] \\
&\qquad \left. + \left[\frac{\partial \boldsymbol{\mu}^{\check{a}}_{k|i}(R_k)}{\partial x}\right]^T (C^{\check{a}})^{-1} \left[\frac{\partial \boldsymbol{\mu}^{\check{a}}_{k|i}(R_k)}{\partial x}\right] \right\}
\end{aligned} \tag{23}$$

Similarly, we can get the $[I_k(R_k)]_{yy}$ and $[I_k(R_k)]_{xy}$.

ACKNOWLEDGMENT

The authors are grateful for the precious time and comments offered by the editor and anonymous peer reviewers. They would like to thank Polaris Wireless, USA, for use of their ray launcher in simulating Santa Clara measurements.

REFERENCES

[1] J. H. Reed, K. J. Krizman, B. D. Woerner, and T. S. Rappaport, "An overview of the challenges and progress in meeting the E-911 requirement for location service," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 30–37, Apr. 1998.
[2] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005.
[3] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.
[4] L. Zhang, J. Liu, and H. Jiang, "Energy-efficient location tracking with smartphones for IoT," in *Proc. IEEE Sensors*, Oct. 2012, pp. 1–4.
[5] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low-cost outdoor localization for very small devices," *IEEE Pers. Commun.*, vol. 7, no. 5, pp. 28–34, Oct. 2000.
[6] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice*. Springer, 2012.
[7] S. Tomic, M. Beko, and R. Dinis, "3-D target localization in wireless sensor networks using RSS and AoA measurements," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3197–3210, Apr. 2017.
[8] K. Li, P. Jiang, E. L. Bodanese, and J. Bigham, "Outdoor location estimation using received signal strength feedback," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 978–981, Jul. 2012.
[9] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM. Conf. Comput. Commun. 19th Annu. Joint Conf. IEEE Comput. Commun. Societies*, vol. 2, Mar. 2000, pp. 775–784.
[10] X. Liu, S. Zhang, J. Quan, and X. Lin, "The experimental analysis of outdoor positioning system based on fingerprint approach," in *Proc. IEEE 12th Int. Conf. Commun. Technol.*, Nov. 2010, pp. 369–372.
[11] A. Arya, P. Godlewski, and P. Melle, "Performance analysis of outdoor localization systems based on RSS fingerprinting," in *Proc. 6th Int. Symp. Wireless Commun. Syst.*, Sep. 2009, pp. 378–382.
[12] H. Koyuncu and S. H. Yang, "A 2D positioning system using WSNs in indoor environment," *Int. J. Elect. Comput. Sci.*, vol. 11, no. 3, pp. 70–77, 2011.

[13] J. S. Lu, E. M. Vitucci, V. Degli-Esposti, F. Fuschini, M. Barbiroli, J. A. Blaha, and H. L. Bertoni, "A discrete environment-driven GPU-based ray launching algorithm," *IEEE Trans. Antennas Propag.*, vol. 67, no. 2, pp. 1180–1192, Feb. 2019.

[14] K. Pahlavan, X. Li, and J. P. Makela, "Indoor geolocation science and technology," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 112–118, Feb. 2002.

[15] R. S. S. Singh and Y. Singh, "Localization techniques in wireless sensor networks," *Int. J. Comput. Sci. Inf. Technol.*, vol. 6, no. 1, pp. 840–850, 2015.

[16] Y. Tsuda, Q. Kong, and T. Maekawa, "Detecting and correcting WiFi positioning errors," in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput.* New York, NY, USA: Association Computing Machinery, Sep. 2013, pp. 777–786.

[17] M. Kanaan and K. Pahlavan, "CN-TOAG: A new algorithm for indoor geolocation," in *Proc. IEEE 15th Int. Symp. Pers., Indoor Mobile Radio Commun.*, vol. 3, Sep. 2004, pp. 1906–1910.

[18] A. Hatami and K. Pahlavan, "Performance comparison of RSS and TOA indoor geolocation based on UWB measurement of channel characteristics," in *Proc. IEEE 17th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2006, pp. 1–6.

[19] J. Sun and C. Li, "Tunnel personnel positioning method based on TOA and modified location-fingerprint positioning," *Int. J. Mining Sci. Technol.*, vol. 26, no. 3, pp. 429–436, May 2016.

[20] J. Li, I.-T. Lu, J. S. Lu, and L. Zhang, "Robust kernel-based machine learning localization using NLOS TOAs or TDOAs," in *Proc. IEEE Long Island Syst., Appl. Technol. Conf. (LISAT)*, May 2017, pp. 1–6.

[21] L. Yu, M. Laaraiedh, S. Avrillon, and B. Uguen, "Fingerprinting localization based on neural networks and ultra-wideband signals," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2011, pp. 184–189.

[22] Z. Wei, Y. Zhao, X. Liu, and Z. Feng, "DoA-LF: A location fingerprint positioning algorithm with millimeter-wave," *IEEE Access*, vol. 5, pp. 22678–22688, 2017.

[23] J. Li, I.-T. Lu, and J. S. Lu, "DOA-based localization algorithms under NLOS conditions," in *Proc. IEEE Long Island Syst., Appl. Technol. Conf. (LISAT)*, May 2018, pp. 1–6.

[24] M. Li and Y. Lu, "Angle-of-arrival estimation for localization and communication in wireless networks," in *Proc. 16th Eur. Signal Process. Conf.*, Aug. 2008, pp. 1–5.

[25] F. Lemic, J. Martin, C. Yarp, D. Chan, V. Handziski, R. Brodersen, G. Fettweis, A. Wolisz, and J. Wawrzynek, "Localization as a feature of mmWave communication," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Sep. 2016, pp. 1033–1038.

[26] T. Rappaport, R. Heath, R. Daniels, and J. Murdock, *Millimeter Wave Wireless Communications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2015, pp. 585–651.

[27] O. Kanhere and T. S. Rappaport, "Position locationing for millimeter wave systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, U.A.E., Dec. 2018, pp. 1–6.

[28] S. M. Kay, *Fundamentals of Statistical Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[29] K. Cheung, H. So, W.-K. Ma, and Y. Chan, "A constrained least squares approach to mobile positioning: Algorithms and optimality," *EURASIP J. Adv. Signal Process.*, vol. 2006, no. 1, Dec. 2006, Art. no. 020858, doi: 10.1155/ASP/2006/20858.

[30] S. Pino-Povedano and F.-J. Gonzalez-Serrano, "On the use of modified Cramér-Rao bound in sensor deployment," *IEEE Sensors J.*, vol. 13, no. 11, pp. 4163–4171, Nov. 2013.

[31] Y. Fu and Z. Tian, "Cramer–Rao bounds for hybrid TOA/DOA-based location estimation in sensor networks," *IEEE Signal Process. Lett.*, vol. 16, no. 8, pp. 655–658, Aug. 2009.

[32] C. Botteron, A. Host-Madsen, and M. Fattouche, "Cramer-Rao bound for location estimation of a mobile in asynchronous DS-CDMA systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, May 2001, pp. 2221–2224.

[33] M. Wax and A. Leshem, "Joint estimation of time delays and directions of arrival of multiple reflections of a known signal," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. Conf.*, vol. 5. Washington, DC, USA: IEEE Computer Society, May 1996, pp. 2622–2625.

[34] A. Catovic and Z. Sahinoglu, "The Cramer-Rao bounds of hybrid TOA/RSS and TDOA/RSS location estimation schemes," *IEEE Commun. Lett.*, vol. 8, no. 10, pp. 626–628, Oct. 2004.

[35] Z. Sahinoglu and A. Catovic, "A hybrid location estimation scheme (H-LES) for partially synchronized wireless sensor networks," in *Proc. IEEE Int. Conf. Commun.*, vol. 7, Jun. 2004, pp. 3797–3801.

[36] Y. Wang, G. Leus, and A.-J. van der Veen, "Cramér-Rao bound for range estimation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 3301–3304.

[37] J. Wang, J. Chen, and D. Cabric, "Cramer-Rao bounds for joint RSS/DoA-based primary-user localization in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1363–1375, Mar. 2013.

[38] J. Werner, J. Wang, A. Hakkarainen, D. Cabric, and M. Valkama, "Performance and Cramer–Rao bounds for DoA/RSS estimation and transmitter localization using sectorized antennas," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3255–3270, May 2016.

[39] Q. Jiang, F. Qiu, M. Zhou, and Z. Tian, "Benefits and impact of joint metric of AOA/RSS/TOF on indoor localization error," *Appl. Sci.*, vol. 6, no. 10, p. 296, Oct. 2016.

[40] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2137–2148, Aug. 2003.

[41] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: Possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 41–53, Jul. 2005.

[42] S. Gezici, "A survey on wireless position estimation," *Wireless Pers. Commun.*, vol. 44, no. 3, pp. 263–282, Feb. 2008.

[43] X. Wei, N. Palleit, and T. Weber, "AOD/AOA/TOA-based 3D positioning in NLOS multipath environments," in *Proc. IEEE 22nd Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2011, pp. 1289–1293.

[44] C. Botteron, M. Fattouche, and A. Host-Madsen, "Statistical theory of the effects of radio location system design parameters on the positioning performance," in *Proc. IEEE 56th Veh. Technol. Conf.*, vol. 2, Sep. 2002, pp. 1187–1191.

[45] J. Prieto, S. Mazuelas, A. Bahillo, P. Fernandez, R. M. Lorenzo, and E. J. Abril, "Adaptive data fusion for wireless localization in harsh environments," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1585–1596, Apr. 2012.

[46] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 672–681, Mar. 2006, doi: 10.1109/TWC.2006.1611097.

[47] B. T. Sieskul, F. Zheng, and T. Kaiser, "On the effect of shadow fading on wireless geolocation in mixed LoS/NLoS environments," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4196–4208, Nov. 2009.

[48] A. Nikitin, C. Laoudias, G. Chatzimilioudis, P. Karras, and D. Zeinalipour-Yazti, "Indoor localization accuracy estimation from fingerprint data," in *Proc. 18th IEEE Int. Conf. Mobile Data Manage. (MDM)*, May 2017, pp. 196–205.

[49] Y. Qi, "Wireless geolocation in non-line-of-sight environment," Ph.D. dissertation, Dept. Elect. Eng., Princeton Univ., Princeton, NJ, USA, 2003.

[50] Y. Qi and H. Kobayashi, "On geolocation accuracy with prior information in non-line-of-sight environment," in *Proc. IEEE 56th Veh. Technol. Conf.*, vol. 1, Sep. 2002, pp. 285–288.

[51] Y. Wang, K. Gu, Y. Wu, W. Dai, and Y. Shen, "NLOS effect mitigation via spatial geometry exploitation in cooperative localization," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6037–6049, Sep. 2020.

[52] Y. Wang, Y. Wu, and Y. Shen, "Multipath effect mitigation by joint spatiotemporal separation in large-scale array localization," in *Proc. GLOBECOM-IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.

[53] A. K. M. M. Hossain and W.-S. Soh, "Cramer-Rao bound analysis of localization using signal strength difference as location fingerprint," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[54] J. S. Lu, X. Han, and H. L. Bertoni, "The influence of terrain scattering on radio links in hilly/mountainous regions," *IEEE Trans. Antennas Propag.*, vol. 61, no. 3, pp. 1385–1395, Mar. 2013.

[55] R. Shibata, "A note on asymptotic unbiasedness of estimates," *Kodai Math. J.*, vol. 2, no. 1, pp. 38–42, 1979, doi: 10.2996/kmj/1138035964.

[56] K.-C. Li, "Consistency for cross-validated nearest neighbor estimates in nonparametric regression," *Ann. Statist.*, vol. 12, no. 1, pp. 230–240, Mar. 1984, doi: 10.1214/aos/1176346403.

[57] G. S. Watson, "Smooth regression analysis," *Sankhya, Indian J. Statist.*, vol. 26, no. 4, pp. 359–372, 1964.

[58] D. Eppink and W. Kuebler, "Tirem/sem handbook," Electromagn. Compat. Anal. Center, Annapolis, MD, USA, Tech. Rep. ECAC-HDBK-86-076, 1994.

JUN LI (Graduate Student Member, IEEE) received the M.S. degree in electrical engineering from the Tandon School of Engineering, New York University (NYU), Brooklyn, NY, USA, in 2017, where he is currently pursuing the Ph.D. degree in electrical engineering with the NYU WIRELESS Research Center.

His research interests include wireless positioning system design, interference mitigation algorithms design, automotive radar signal processing, sparse signal recovery, and Bayesian optimization. He was a recipient of the School of Engineering (SoE) Fellowship from the NYU Tandon School of Engineering. He received the Myron M. Rosenthal Award for Best M.S. Academic Achievement in Electrical and Computer Engineering, in 2017.

I-TAI LU (Senior Member, IEEE) received the Ph.D. degree in E.E. from Polytechnic University. He is currently a Professor with the Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University. He has published more than 200 journals and proceeding papers and patents. He has given many invited lectures and short courses on wireless communications in various high-tech companies and universities. His current research interests include wireless communications, in which he has made contributions to the development and standardization of various wireless LAN systems, the third generation cellular communications, the fourth generation (3GPP LTE-A), and future generations (including 5G) of cellular communications systems.

JONATHAN LU (Member, IEEE) received the B.S. and M.S. degrees in electrical engineering and the Ph.D. degree from New York University (NYU), New York, NY, USA, in 2010 and 2014, respectively.

From 2008 to 2014, he was a Research Fellow with NYU. From 2014 to 2019, he was a Senior Research Engineer and a Senior Staff Research Engineer with Polaris Wireless, Mountain View, CA, USA. Since 2019, he has been a Principal Engineer with Zainar, Redwood City, CA. He is the author of over 25 technical articles. He has contributed to 3GPP and ATIS standards for emergency location and 5G positioning. He is also an Inventor of 15 issued patents in the positioning and geolocation space. His research interests include wave propagation modeling, and positioning algorithms and system design.