

Performance Analysis of Maximum Likelihood Estimation for Multiple Transmitter Locations

Xiaoli Hu^{1,2}, Pin-Han Ho³, and Limei Peng^{4,*}

¹Nanfeng College, Sun Yat-Sen University, Guangzhou 510900, China

²College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

³Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada

⁴School of Computer Science and Engineering, Kyungpook National University, South Korea

* Corresponding author

The paper considers the consistence condition of Maximum Likelihood (ML) estimation for multiple transmitter locations in a wireless network with cooperative receiver nodes. It is found that the location set of receiver nodes should not locate (or asymptotically in some sense) merely in an algebraic curve of order $2M - 1$ if there are totally M transmitters. A sufficient condition for consistence of the ML estimation for M transmitters is that the limit set of locations contains a subset, comprised of $(2M^2 - M + 2)$ points, which is non- $C-2M$ -co-curved, a notion given by Definition IV-B. This condition can be compared to the persistent excitation condition used to guarantee the convergence of least squares algorithm. Numerical experiments are designed to demonstrate the theoretical discoveries in both positive and negative aspects.

Index Terms—Multiple transmitter locations, maximum likelihood estimation, consistence condition.

I. INTRODUCTION

THE locations of transmitters in a wireless network are precious information that can be used to determine the spatial and temporal resources of a sensing node [1]. This has been identified as one of the ultimate tasks in various emerging wireless networking scenarios such as Cognitive Radios (CR). A CR system is designed to operate concurrently with the primary radio system by sharing some unused spectrum resources, referred to as *spectral holes*, while minimizing the interference with the operation of the primary radio system. Note that a spectral hole includes three aspects: frequency, space, and time [2]. Thus, a key step to implement a CR network is to estimate the locations of transmitters in a region of interest. For instance, a scenario is shown in [2] that accurate information of locations of transmitters would help to avoid disruption among the transmission between two cognitive nodes.

Localization of a single transmitter based on received signal strength observations has been extensively studied [1], [3], [4], [5], [6], [7]. A more challenging issue for the localization of multiple transmitters is considered in [2], [8], [9]. For example, authors in [8] addressed the issue of detection and localization of multiple intruders present simultaneously and proposed group-based algorithm, which works by extracting observations across sensors for certain small sets of transmitters. Authors in [9] proposed a measurement clustering method to estimate the locations of multiple radio transmitters based only on the received signal strength and proposed a performance analysis approach based on a hypothesis testing framework. Besides the methods based on clustering or grouping, the estimation based on Maximum Likelihood idea is another natural choice [7], and thus both of the problems

can be formulated into an optimization process. Though some practical numerical algorithms have been reported to solve the optimization issues, it appears that the relevant theoretical performance analysis of the ML estimation is rarely studied. We consider this issue for a single transmitter in [7], and find that the receiver (or observation) locations should be rich enough to guarantee the consistence of ML estimation.

The main purpose of this paper is to discover the consistence condition for ML estimation of multiple transmitters. The novelty of our work can be found in how we formulate the model as follows. Consider a wireless network consisting of M primary nodes, or transmitters, as well as a set of receiver nodes that are listening to the signal launched by the primary nodes. Assume that the primary nodes transmit at a constant power level during the observation period, and the receiver nodes, with their locations known in prior, can exchange their respective received power information with each other [6], [10]. The problem is to estimate the locations of the primary nodes based on a lognormal shadowing model [11], [6].

Denote the unknown two dimensional locations of M transmitters as $\theta = [\theta_1, \dots, \theta_M]^T \in R^{2M}$, where $\theta_i = (a_i, b_i)$ is the location of the i -th transmitter. Assume that M is known as the exact number of transmitters, and that all transmitters have the same power as s_p without loss of generality. The received power at the j -th receiver from the i -th transmitter is given by a lognormal shadowing model ([6], [11]) as

$$s_{ij} = s_p - 10\epsilon \log_{10} d_{ij} + w_{ij}, \quad (1)$$

where $d_{ij} = \sqrt{(x_j - a_i)^2 + (y_j - b_i)^2}$, ϵ is known to represent the path loss exponent, and iid noise $w_{ij} \sim N(0, \sigma^2)$. Thus, the actually observed power at the j -th receiver is given by

$$s_j = \sum_{i=1}^M s_{ij}. \quad (2)$$

The problem is to estimate location vector θ based on the available data set $\{s_j, x_j, y_j, j = 1, 2, \dots, n\}$. In view of the fact that the observation data may be collected from a finite set of monitor locations, it is reasonable to assume that the monitor location (x_i, y_i) is a deterministic value. The main subsequent target is to discover the condition of guaranteeing the consistence of Maximum Likelihood estimation for θ . It will be discovered that the receiver location set $\{(x_i, y_i)\}$ should not be located in a single $(2M - 1)$ -order algebraic curve, i.e., it should be rich enough in some sense (see Theorem IV-B for detail).

The rest of this paper is organized as follows. The ML estimation is formulated for multiple transmitter locations in Section II. In Section III, two numerical experiments are designed to detect the performance of the ML estimate algorithm as the data volume tends to infinity, by the aid of a practical Matlab code `fminsearch` in a recursive form with respect to data volume n . A sufficient condition to guarantee the consistence of ML estimation for the M transmitter locations is proposed in Section IV by the aid of theoretical preliminary in Appendix. Then, the technical mechanisms beneath the numerical experiments are demonstrated and justified therein. Some concluding remarks of the paper are given in Section V.

II. MAXIMUM LIKELIHOOD ESTIMATION FOR MULTIPLE LOCATIONS

In this section, we formulate the ML estimation for multiple locations. The estimation issue turns out to be a nonlinear optimization problem.

As aforementioned, because the locations (x_i, y_i) are taken as deterministic quantities, the random variable s_{ij} has a Gaussian distribution $\mathcal{N}(s_p - 10\epsilon \log_{10} d_{ij}, \sigma^2)$. By noticing the independence of $\{w_{ij}\}$, the density of s_j is given by

$$f(s_j; x_j, y_j, P) = \frac{1}{\sqrt{2\pi M\sigma}} \exp \left\{ -\frac{(s_j - Ms_p + 10\epsilon \sum_{i=1}^M \log_{10} d_{ij})^2}{2M\sigma^2} \right\} \quad (3)$$

where $P = \{P_i(a_i, b_i), i = 1, 2, \dots, M\}$ denotes the location set of transmitters, or primary nodes. It is worth pointing out that though s_j is a sequence of independent random variables, each s_j has different distribution since (x_j, y_j) is different for different j .

The ML method is a wise choice to estimate the parameters θ in equation (2) based on available observation data $\mathcal{O}(n)\{s_j, x_j, y_j, j = 1, 2, \dots, n\}$.

The corresponding ML function is

$$\mathcal{L}(\theta) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi M\sigma}} \exp \left\{ -\frac{(s_j - Ms_p + 10\epsilon \sum_{i=1}^M \log_{10} d_{ij})^2}{2M\sigma^2} \right\} \quad (4)$$

so, the log ML function follows as

$$\log \mathcal{L}(\theta) = -\frac{n}{2} \log(2\pi M) - n \log \sigma - \frac{1}{2M\sigma^2} \sum_{j=1}^n \left(s_j - Ms_p + 10\epsilon \sum_{i=1}^M \log_{10} d_{ij} \right)^2 \quad (5)$$

which is equivalent to minimize

$$M_n(\theta) \sum_{j=1}^n \left(s_j - Ms_p + 10\epsilon \sum_{i=1}^M \log_{10} d_{ij} \right)^2 \quad (6)$$

This turns to be a nonlinear optimization problem regarding $\theta = [a_1, b_1, a_2, b_2, \dots, a_M, b_M]^T$, which can be solved by many mathematical software, e.g., Matlab's optimization toolbox, or by designing special program to solve it. By noticing that the function $M_n(\theta)$ is not necessarily convex, a straightforward search for the minimum may tend to some local minimum rather than global minimum.

III. NUMERICAL EXPERIMENTS

We develop a recursive numerical algorithm to minimize $M_n(\theta)$ in M_n by using Matlab code `fminsearch`. When the n_i th ML estimate for θ is numerically founded to be θ_n , then the next minima θ_{n+1} for $M_{n+1}(\theta)$ is numerically searched based on θ_n as:

$$\theta_{n+1} = \text{fminsearch}(@M_{n+1}, \theta_n) \quad (7)$$

This means the starting point of the search is θ_n , rather than any other value by random guess. So, the information obtained by the former step has been used for the next search, which may increase the effectiveness.

Below we design two numerical experiments to demonstrate the asymptotical performance of ML estimation for θ in model 2 as the number of observation locations tends to infinity. The first one is an example of failure in the localization (referred to as the *failed example*), and then the successful example follows. The essential reason for failure is that the likelihood function 4, or the equivalent function $M_n(\theta)$ in M_n , has multiple roots when some conditions are absent. The technical reason will be explained in subsection IV-C after general analysis of the consistence condition. Note that even in the failed examples, the ML estimation may still work since the true value is also a root of the likelihood function. It depends on the numerical algorithm to find the roots of the likelihood function.

Let $\epsilon = 1$, $s_p = 80$, $M = 2$, $P_1(-20, 0)$, $P_2(20, 0)$ in 12, $w_i \sim \mathcal{N}(0, 0.5^2)$, and all locations (x_i, y_i) are distributed in the following way: generate x_i uniformly distributed in the interval $[-20, 20]$, and then y_i is calculated from the implicit function $40y_i^3 + 400y_i^2 + 40x_i^2y_i - 1200x_i^2 + 160000 = 0$, whose graph is shown in Fig. 3. The whole data to be used is $N = 100$. The performance of the ML estimates has been shown in Fig. 1, where θ_n with $n = 20, 40, 60, 80, 100$ is plotted by "plus" and "cross" signs for the estimates of P_1 and P_2 , respectively. Clearly the estimates for P_1 tend to $(0, 20)$, while the estimates for P_2 tend to $(0, 0)$ in the Fig. 1. The consistence of ML estimators fails in Fig. 1.

With the same setting of Example III, it is expected that all locations (x_i, y_i) are in the set $\{(11,0), (-12,0), (0,9), (0,-8), (10,11), (10,-10), (-10,10), (-10,-10)\}$ successively. The whole data to be used is $N = 100$. The performance of the ML estimates has been shown in Fig. 2, where θ_n with $n = 20, 40, 60, 80, 100$ is plotted by "plus" and "cross" signs for the estimates of P_1 and P_2 , respectively. The ML estimators converge to all true values as shown in Fig. 2.

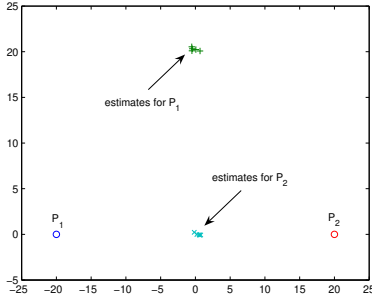


Fig. 1 ML estimation

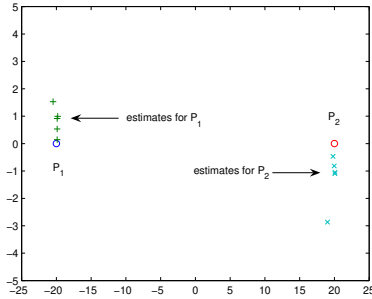


Fig. 2 True values of ML estimation

IV. CONSISTENCE CONDITION

In this section, we will describe the condition of guaranteeing the ML estimate θ_n as $n \rightarrow \infty$, which in turn explains the success and failure of the numerical experiments in the former section. To make the condition easy to be understood, a special case to justify unique solution of a system algebraic equations regarding model 2 without noise is studied in subsection IV-A. The consistence condition is strictly defined and a criterion is given and proved in subsection IV-B. Finally, subsection IV-C serves as a explanation for numerical experiments.

A. Unique Solution Condition for A System of Equations

Let us consider a simple yet important issue regarding to the uniqueness of the solution as a starting point for the ML estimation performance. The uniqueness of the solutions relies on whether the available data set $\{(x_i, y_i)\}_1^m$ serves as a unique solution of a system of equations as:

$$s_j = Ms_p - 10\epsilon \sum_{i=1}^M \log_{10} \sqrt{(x_j - a_i)^2 + (y_j - b_i)^2}, \quad j = 1, 2, \dots, m. \quad (8)$$

Exactly, when it is solvable, one may want to ensure that the solution is unique from the above m equations. This means

that there is no other solution $\theta' = [a'_1, b'_1, \dots, a'_M, b'_M]^T \neq \theta$ (i.e., the two sets of points are different, where $\{(a_i, b_i), i = 1, \dots, M\} \neq \{(a'_i, b'_i), i = 1, \dots, M\}$) to make the data set $\{(a_i, b_i)\}_1^m$ satisfy the following m equations:

$$s_j = Ms_p - 10\epsilon \sum_{i=1}^M \log_{10} \sqrt{(x_j - a'_i)^2 + (y_j - b'_i)^2}, \quad j = 1, 2, \dots, m. \quad (9)$$

Combining 8 and 9, we get:

$$\prod_{i=1}^M \sqrt{(x_j - a_i)^2 + (y_j - b_i)^2} = \prod_{i=1}^M \sqrt{(x_j - a'_i)^2 + (y_j - b'_i)^2}, \quad j = 1, 2, \dots, m. \quad (10)$$

Thus, the solution of 8 is not unique if the data set $\{(x_i, y_i)\}_1^m$ satisfies 10 with a certain θ' . Conversely, if the solution is unique, the data set should never satisfy 10 with any given θ' .

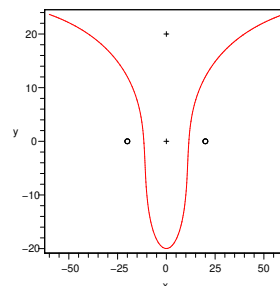
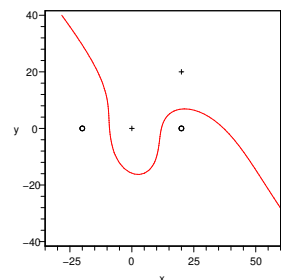
Replacing x_j, y_j as x, y in 10, let us define a curve by equation of x, y as:

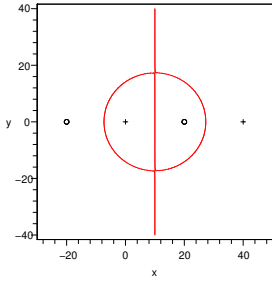
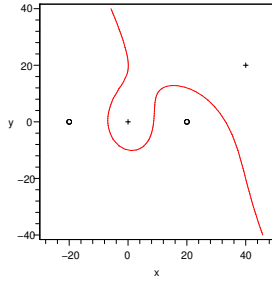
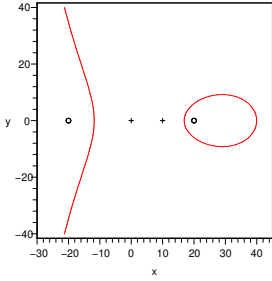
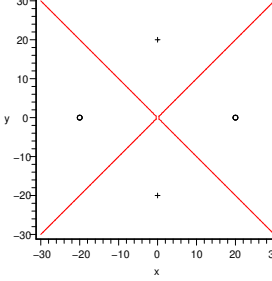
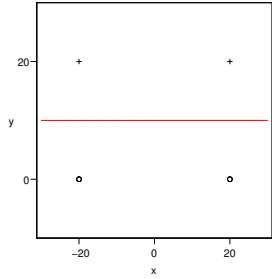
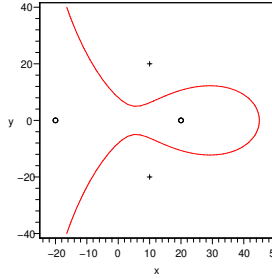
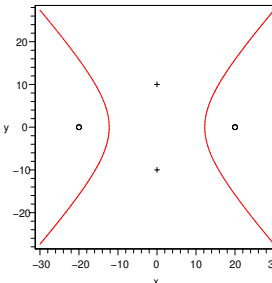
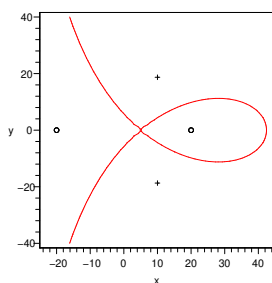
$$\prod_{i=1}^M \sqrt{(x - a_i)^2 + (y - b_i)^2} = \prod_{i=1}^M \sqrt{(x - a'_i)^2 + (y - b'_i)^2} \quad (11)$$

with parameters θ and θ' . By denoting P_i as the point (a_i, b_i) in the plane, and P'_i as (a'_i, b'_i) , $i = 1, 2, \dots, M$, the point $P(x, y)$ in the above curve satisfies $\prod_{i=1}^M |PP_i| = \prod_{i=1}^M |PP'_i|$.

Here the curve defined by 11 is denoted by $C(P_1, \dots, P_M; P'_1, \dots, P'_M)$, or $C-2M$ curve for brief.

We use Maple code `/implicitplot/` to plot some graphs of the $C-2 \times 2$ curve defined by 11 below with $P_1(-20, 0)$, $P_2(20, 0)$, and P'_1, P'_2 are given by the options in Fig. 3-12. From a topological point of view, Figures 3,4,6,9 and 10 have the same topological structure. It seems that each curve has a branch tending to infinity, and that each curve separates the two point sets $\{P_1, P_2\}$ and $\{P'_1, P'_2\}$. These two properties have been stated under general settings for the $C-2M$ case in the following propositions.


 Fig. 3 $P'_1(0,0), P'_2(0,20)$

 Fig. 4 $P'_1(0,0), P'_2(20,20)$


 Fig. 5 $P'_1(0, 0), P'_2(40, 0)$

 Fig. 6 $P'_1(0, 0), P'_2(40, 20)$

 Fig. 7 $P'_1(0, 0), P'_2(10, 0)$

 Fig. 8 $P'_1(0, -20), P'_2(0, 20)$

 Fig. 9 $P'_1(-20, 20), P'_2(20, 20)$

 Fig. 10 $P'_1(10, -20), P'_2(10, 20)$

 Fig. 11 $P'_1(0, -10), P'_2(0, 10)$

 Fig. 12 $P'_1(10, -18.7), P'_2(10, 18.7)$

Each $C(P_1, \dots, P_M; P'_1, \dots, P'_M)$ curve defined by 11 has at least a branch extending to infinity.

Proof. Define a function of $(x, y) \in R^2$ and $\mu \geq 0$ as $G(x, y, \mu) = \frac{\prod_{i=1}^M |PP_i|^2 - \mu \prod_{i=1}^M |PP'_i|^2}{\prod_{i=1}^M [(x - a_i)^2 + (y - b_i)^2] - \mu \prod_{i=1}^M [(x - a'_i)^2 + (y - b'_i)^2]}$. Clearly, the equation $G(x, y, 1) = 0$ defines the curve given by 11. Obviously, the curve is different for different $\mu \geq 0$.

By the geometrical meaning of

$$\frac{\prod_{i=1}^M |PP_i|}{\prod_{i=1}^M |PP'_i|}, \quad (12)$$

it is clear that the curve $G(x, y, \mu) = 0$ is continuous with respect to μ .

Now let us come back to the assertion to be proved. It is sufficient to show that the set defined by equation $G(x, y, 1) = 0$ is unbounded. To prove the assertion, we argue it in a reverse way by firstly assuming the point set of $G(x, y, 1) = 0$ is bounded, while later the contradiction is shown.

By continuity of the curve to μ , there exist $\mu_1^- < 1 < \mu_1^+$ such that the union of curves $G(x, y, \mu) = 0$, denoted by S_1 , is a bounded set for any $\mu \in [\mu_1^-, \mu_1^+]$. Similarly, because the curve $G(x, y, 0) = 0$ is the set of points P_1, \dots, P_M and $G(x, y, \infty) = 0$ the set of points P'_1, \dots, P'_M , there exist $0 < \mu_0 < \mu_1^-$ and $\mu_2 > \mu_1^+$ such that the union of curves $G(x, y, \mu) = 0$, denoted by S_2 , is a bounded set for any $\mu \in [0, \mu_0] \cup [\mu_2, \infty]$. By denoting the union of curves $G(x, y, \mu) = 0$ for $\mu \in [\mu_0, \mu_1^-] \cap [\mu_1^+, \mu_2]$ as S_3 , S_3 is a bounded set by noticing the following fact:

$$G(x, y, \mu) = (1 - \mu)(x^2 + y^2)^{2M} + \text{low order of } (x^2 + y^2)^{2M}. \quad (13)$$

If $\mu \neq 1$, then the right hand side in the above formula tends to ∞ as $(x^2 + y^2) \rightarrow \infty$. Thus, the aforementioned set is bounded.

Now we derive that $S_1 \cap S_2 \cap S_3$ is a bounded set. However, the union should cover the whole plane which is not a bounded set. This contradiction deduces the desired assertion.

Each $C(P_1, \dots, P_M; P'_1, \dots, P'_M)$ curve defined by 11 separates the two point sets $\{P_1, \dots, P_M\}$ and $\{P'_1, \dots, P'_M\}$.

Proof. We will prove that the curve given by 11 divides the two point sets into different connected regions. For this, we need only point out that any curve segment $C-(P_i, P'_j)$ connecting both P_i and P'_j does not belong to a single connected region determined by curve $C(P_1, \dots, P_M; P'_1, \dots, P'_M)$ for any $i, j = 1, 2, \dots, M$.

Consider function

$$G_1(x, y) = \frac{\prod_{i=1}^M |PP_i|}{\prod_{i=1}^M |PP'_i|} \quad (14)$$

along the curve segment $C-(P_i, P'_j)$ connecting P_i and P'_j . Obviously, $G_1(x, y) = 0$ if $(x, y) = P_i$, and $G_1(x, y) \rightarrow \infty$ if $(x, y) = P'_j$. Thus, by the continuity of function $G_1(x, y)$ over $C-(P_i, P'_j)$, there exists a point P in curve segment $C-(P_i, P'_j)$ separating the two points P_i and P'_j . This completes the proof.

Another proof based on function IV-A seems more clear:

$$G(x, y, 1)|_{P_i} < 0, \quad G(x, y, 1)|_{P'_j} > 0, \quad (15)$$

which means points P_i and P'_j is separated by the curve $G(x, y, 1) = 0$.

When calculating the curve equation 11 for $M = 2$, we get

$$a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3 + a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{10}x + a_{01}y + a_{00} = 0, \quad (16)$$

where $a_{30} = a_{12} = 2(a'_1 + a'_2 - a_1 - a_2)$, $a_{03} = a_{21} = 2(b'_1 + b'_2 - b_1 - b_2)$, $a_{20} = r_1^2 + r_2^2 + 4a_1a_2 - r_1'^2 - r_2'^2 - 4a'_1a'_2$, $a_{02} = r_1^2 + r_2^2 + 4b_1b_2 - r_1'^2 - r_2'^2 - 4b'_1b'_2$, $a_{11} = 4(a_1b_2 + a_2b_1 - a'_1b'_2 - a'_2b'_1)$, $a_{10} = 2(a_2r_1'^2 + a'_1r_2'^2 - a_2r_1'^2 - a_1r_2'^2)$,

$a_{01} = 2(b'_2 r_1'^2 + b'_1 r_2'^2 - b_2 r_1^2 - b_1 r_2^2)$,
 $a_{00} = r_1^2 r_2^2 - r_1'^2 r_2'^2$ with $r_1^2 = a_1^2 + b_1^2$, $r_2^2 = a_2^2 + b_2^2$, and r_1'
 and r_2' similarly.

If $a_{30} = a_{03} = 0$, the curve 16 turns to be a hyperbola or a
 combination of two straight lines, as shown in Fig. 6 and 9.

For a given $C(P_1, P_2, P'_1, P'_2)$ curve defined by 11 with
 $P_1(a_1, b_1)$, $P_2(a_2, b_2)$, $P'_1(a'_1, b'_1)$, $P'_2(a'_2, b'_2)$, if $a_1 + a_2 =$
 $a'_1 + a'_2$, $b_1 + b_2 = b'_1 + b'_2$, then the curve is a hyperbola or
 a combination of two straight lines.

Proof. By IV-A, we have

$$r_1^2 + r_2^2 - r_1'^2 - r_2'^2 = 2(a'_1 a'_2 + b'_1 b'_2 - a_1 a_2 - b_1 b_2). \quad (17)$$

By further help of 16, we derive

$$a_{20} = 2(-a'_1 a'_2 + b'_1 b'_2 + a_1 a_2 - b_1 b_2) = -a_{02}. \quad (18)$$

Hence, by the standard discriminant

$$\Delta a_{11}^2 - 4a_{20}a_{02} = a_{11}^2 + 4a_{20}^2 \geq 0, \quad (19)$$

we know that the curve $a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{10}x + a_{01}y +$
 $a_{00} = 0$ is a hyperbola or a combination of two straight lines,
 depending on $\Delta > 0$ or $\Delta = 0$ respectively.

An interesting problem is how to determine whether a set
 of planar points does not belong to any curve described by
 equation 16. Note that the equation 16 has 10 coefficients (the
 freedom degree is actually 7 in view of the fact that $a_{30} = a_{12}$
 and $a_{21} = a_{03}$), the number of planar points should at least
 be greater than 7 to avoid locating in a curve defined by 16.

Define a vector function of $P(x, y) \in R^2 R^8$:
 $V(P) = V(x, y)[x^3 + xy^2, x^2y + y^3, x^2, xy, y^2, x, y, 1]^T$, and a 8×8
 matrix comprised by 8 planar points set $Q = \{Q_i, i =$
 $1, 2, \dots, 8\}$ as: $D(Q)[V(Q_1), V(Q_2), V(Q_3), \dots, V(Q_8)]$.

For a set of 8 points: $Q = \{Q_i, i = 1, 2, \dots, 8\}$, if the
 determinant $|D(Q)| \neq 0$, where $D(Q)$ is given by IV-A, then
 the 8 points do not locate in any curve defined by equation
 16.

Proof. Assume that the 8 points locate in a curve defined by
 equation 16. Thus the 8 corresponding equations with respect
 to coefficients are as following:

$$D(Q)[a_{30}, a_{30}, a_{20}, a_{11}, a_{02}, a_{10}, a_{01}, a_{00}]^T = 0. \quad (20)$$

By the fact $|D(Q)| \neq 0$, we have

$$[a_{30}, a_{30}, a_{20}, a_{11}, a_{02}, a_{10}, a_{01}, a_{00}]^T = 0, \quad (21)$$

which proves the assertion.

A general problem is how to determine whether a set of
 planar points does not belong to any curve described by
 equation 11. When simplifying the equation 11, we know
 that the algebraic expression has $M(2M + 1) = 2M^2 + M$
 coefficients (the freedom degree is actually $2M^2 + M - 1$) by
 the fact that the orders are from $2M - 1$ to 0. By noticing
 that the coefficients of $x^{2M-1}, x^{2M-3}y^2, \dots, xy^{2M-2}$ are all
 $2(a'_1 + \dots + a'_M - a_1 - \dots - a_M)$, and the coefficients
 of $x^{2M-2}y, x^{2M-4}y^3, \dots, y^{2M-1}$ are all $2(b'_1 + \dots + b'_M -$
 $b_1 - \dots - b_M)$ in the expression of 11. Hence, the number
 of planar points belonging to the set should at least be
 $2M^2 + M - 2(M - 1) = 2M^2 - M + 2$ to avoid locating in
 a curve defined by 11.

Define a vector function of $P(x, y) \in R^2 R^{2M^2 - M + 2}$:

$$V_M(P) = V_M(x, y)[\alpha(x, y), \beta(x, y), x^{2M-2}, \dots, y^{2M-2}, \dots, x, y, 1]^T, \quad (22)$$

where $\alpha(x, y) = x^{2M-1} + x^{2M-3}y^2 + \dots + xy^{2M-2}$,
 $\beta(x, y) = x^{2M-2}y + x^{2M-4}y^3 + \dots + y^{2M-1}$, and a $(2M^2 -$
 $M + 2) \times (2M^2 - M + 2)$ matrix comprised by $2M^2 - M + 2$
 planar points set $Q = \{Q_i, i = 1, 2, \dots, 2M^2 - M + 2\}$ as:
 $D_M(Q)[V_M(Q_1), V_M(Q_2), V_M(Q_3), \dots, V_M(Q_{2M^2 - M + 2})]$.

A similar criterion compared to Proposition IV-A is as
 following. The proof is omitted since it is nearly the same
 of the $M = 2$ case (Proposition IV-A).

For a set of $2M^2 - M + 2$ points: $Q = \{Q_i, i =$
 $1, 2, \dots, 2M^2 - M + 2\}$, if the determinant $|D_M(Q)| \neq 0$,
 where $D_M(Q)$ is given by IV-A, then the $2M^2 - M + 2$ points
 do not locate in any curve defined by equation 11.

It is an important issue to find $2M^2 - M + 2$ planar points
 satisfying the restriction $|D_M(Q)| \neq 0$. By observing a class
 of linearly independent polynomial functions, the restriction
 $|D_M(Q)| \neq 0$ holds for almost all the $2M^2 - M + 2$ points.
 In other words, the set of points satisfying $|D_M(Q)| = 0$ in
 $R^{2M^2 - M + 2}$ is a 0 Lebesgue measure set. Therefore, to check
 the restriction, we can randomly select $(2M^2 - M + 2)$ points
 in the plane R^2 , and then check the corresponding $|D_M(Q)|$
 equaling 0 or not. If not, we can slightly change some of them,
 and the set of points probably shows up.

B. Consistence Condition

The consistence condition is defined as the requirement for
 richness of the locations $\{(x_i, y_i)\}$, and a generation of the
 notion introduced in the former subsection. These notions are
 somewhat similar to 'persistence of excitation condition' in
 Least Squares algorithm to guarantee the convergence.

A set of position coordinates $\{(x_i, y_i), i = 1, \dots, t\}$ is
 called C - $2M$ -co-curved with respect to $\{P_i(a_i, b_i), i =$
 $1, 2, \dots, M\}$, if there exist another M positions
 $\{P'_i(a'_i, b'_i), i = 1, 2, \dots, M\}$ such that (x_i, y_i) satisfies
 the equation:

$$\prod_{i=1}^M \sqrt{(x - a_i)^2 + (y - b_i)^2} = \prod_{i=1}^M \sqrt{(x - a'_i)^2 + (y - b'_i)^2} \quad (23)$$

for any $i = 1, \dots, t$. Otherwise, $\{(x_i, y_i), i = 1, \dots, t\}$ is
 called non- C - $2M$ -co-curved.

Now we are in the position of obtaining a consistence result
 in an ML estimation for multiple transmitter locations.

Assume that the location set $\{(x_i, y_i)\}_1^\infty$ is in a bounded
 region, and its limit set in the plane is denoted as Λ . If Λ
 contains a subset Λ_0 consisting of $(2M^2 - M + 2)$ points,
 the points are non- C - $2M$ -co-curved, and a series of noninter-
 secting and successive index sets exist for the whole natural
 numbers as $\{A_j, j = 1, 2, \dots\}$ with $|A_j|$ less than a constant
 positive integer such that $\{(x_i, y_i) : i \in A_j\} \cap B(P, \gamma_j) \neq \emptyset$
 for any ball $B(P, \gamma_j)$ centered at P with radius $\gamma_j [j \infty] 0$,
 $\forall P \in \Lambda_0$. Then, the ML estimation for model 2 with the
 likelihood function given by 4 is in strong consistency, i.e.,

the estimate converges to the true value with probability 1 as the data number tends to infinity.

Proof. Theorem V is taken in the proof. Since the model is assumed to be located in a local region, we need only to check the two conditions therein. By the fact that the likelihood function given by 4 and density for s_j given by 3 are all differentiable and continuous, the condition (ii) of Theorem V is obviously satisfied.

To show V in Theorem V for density 3, by the fact that the function log is invertible and continuous, it is sufficient to show that $\max_{j \in A_i} |\log f(s_j; x_j, y_j, P') - \log f(s_j; x_j, y_j, P)| \geq \delta_0 > 0 \quad \forall i$, hold for $P' \neq P$ with all s_j belong to a fixed compact set. By noticing all parameters are bounded, it is sufficient to require that

$$\max_{j \in A_i} |\log f(s_0; x_j, y_j, P') - \log f(s_0; x_j, y_j, P)| \neq 0 \quad (24)$$

hold for certain fixed s_0 . By further the fact described by IV-B, it is sufficient to show that

$$\max_{(x,y) \in \Lambda_0} |\log f(s_0; x, y, P') - \log f(s_0; x, y, P)| \neq 0, \quad (25)$$

which is equivalent to require that

$$\begin{aligned} \max_{(x,y) \in \Lambda_0} & \left| (s_0 - Ms_p + 10\epsilon \sum_{i=1}^M \log_{10} d_{ij})^2 \right. \\ & \left. - (s_0 - Ms_p + 10\epsilon \sum_{i=1}^M \log_{10} d'_{ij})^2 \right| \neq 0 \end{aligned} \quad (26)$$

by substituting 3 into IV-B. For 26, it is sufficient to require that

$$\max_{(x,y) \in \Lambda_0} \left| \sum_{i=1}^M \log_{10} d_{ij} - \sum_{i=1}^M \log_{10} d'_{ij} \right| \neq 0, \quad (27)$$

which is further sufficient to require that

$$\begin{aligned} \max_{(x,y) \in \Lambda_0} & \left| \prod_{i=1}^M \sqrt{(x - a_i)^2 + (y - b_i)^2} - \right. \\ & \left. \prod_{i=1}^M \sqrt{(x - a'_i)^2 + (y - b'_i)^2} \right| \neq 0. \end{aligned} \quad (28)$$

This is the so-called non- C - $2M$ -co-curved for Λ_0 , and thus the proof of the corresponding condition (i) of Theorem V is finished. Hence, the assertion of the theorem follows.

Based on Theorem IV-B, the ML estimation for M transmitter locations in 2 may fail even if the number of observation tends to infinity in the event that the so-called non- C - $2M$ -co-curved condition fails. For example, the whole location set $\{x_i, y_i\}_1^\infty$ belongs to a single C - $2M$ -curve. As a matter of fact, even in the failure case, it is still possible to see that the ML estimators converge to the true value, since the true value is also the solution of the asymptotic likelihood function.

C. Explanations for Numerical Experiments

Now, by the help of Theorem IV-B, it is easy to explain directly the numerical experiments in the former section.

1) The location set $\{(x_i, y_i), i = 1, 2, \dots\}$ in Example III is C - $2M$ -co-curved, since they are all located in a cubic curve described by following equation:

$$40y^3 + 400y^2 + 40x^2y - 1200x^2 + 160000 = 0, \quad (29)$$

which is equivalent to

$$\sqrt{(x+20)^2 + y^2} \cdot \sqrt{(x-20)^2 + y^2} = \sqrt{x^2 + y^2} \cdot \sqrt{x^2 + (y-20)^2}. \quad (30)$$

This means the corresponding $P'_1(0, 0)$, $P'_2(0, 20)$ in 11 with $P_1(-20, 0)$, $P_2(20, 0)$. In this case, the estimates for the two locations P_1, P_2 may tend to P'_1, P'_2 , which has been discovered by the experiment.

2) Example III is a positive one. The corresponding location set Λ_0 in Theorem IV-B is the following 8 points set:

$$\begin{aligned} \Lambda_0 = \{ & (11, 0), (-12, 0), (0, 9), (0, -8), (10, 11), \\ & (10, -10), (-10, 10), (-10, -10)\}. \end{aligned}$$

By calculating $D_2(\Lambda_0)$ in Proposition IV-A, We find that

$$D_2(\Lambda_0) = 7148845009152000 \neq 0. \quad (31)$$

So, the so-called non- C - $2M$ -co-curved condition hold for Λ_0 . Hence, consistence holds, i.e., the ML estimates for $\theta = [a_1, b_1, a_2, b_2]^T$ tend to true values as data volume tends to infinity, which coincides with the numerical experiments.

V. CONCLUSION

We have analyzed the consistence condition of Maximum Likelihood (ML) estimation for multiple transmitter locations in a wireless network with cooperative receiver nodes. It is found that the location set of receiver nodes should not locate (or asymptotically in some sense) merely in an algebraic curve of order $2M - 1$, given by equation 11, if there are totally M transmitters. The consistence of the ML estimation for M transmitters holds when the limit set of locations contains a subset, comprised of $(2M^2 - M + 2)$ points in a non- C - $2M$ -co-curved manner, which is a notion given by Definition IV-B. We refer to Theorem IV-B for an exact description for the condition. This condition can be compared to the persistent excitation condition used to guarantee the convergence of least squares algorithm [13].

Two numerical experiments were designed to demonstrate the theoretical discoveries in both positive and negative aspects. All this experiments have been analyzed and explained after main theoretical results established.

For future considerations, it is desired to find more efficient numerical algorithm for the involved optimization issue to minimize the function $M_n(\theta)$ given by 6. Because M_n is generally not a convex function, there are many local minimum points among the interested region, which causes some trouble to find the global minimum. It is also of interest to detect the consistence condition for more general and complicated case, e.g., s_p or even ϵ in 1 is unknown.

APPENDIX

Here we copy Theorem 5.2 in [7] as a preliminary for theoretical analysis. Mainly the same notations are introduced below as in [12].

Let $(\mathcal{X}, \mathcal{U}, P_\theta, \Theta)$ be a statistical experiment, which means $(\mathcal{X}, \mathcal{U}, P_\theta)$ is a probability space for any $\theta \in \Theta$. Let P_θ be absolutely continuous with respect to measure μ on \mathcal{U} and thus $\frac{dP_\theta}{d\mu} = p(X, \theta)$, i.e., something like density function. Let X be the observation data generated by certain statistical experiment, then, the function $p(X, \theta)$ is called likelihood function corresponding to the experiment and observation X . The statistic $\hat{\theta}$ defined by

$$\hat{\theta} = \arg \sup_{\theta \in \Theta} p(X, \theta) \quad (32)$$

is called the maximum likelihood estimator for the parameter θ based on the observation X . In the case of independent identically distributed (iid) observations X_1, \dots, X_n , where X_i possesses the density $f(x, \theta)$ with respect to measure μ . The ML estimator $\hat{\theta}_n$ has a simple form as $\hat{\theta}_n = \arg \sup_{\theta \in \Theta} \prod_{i=1}^n f(X_i, \theta)$.

When observations X_1, \dots, X_n are independent yet not identically distributed, the ML estimator $\hat{\theta}_n$ is generally as $\hat{\theta}_n = \arg \sup_{\theta \in \Theta} \prod_{i=1}^n f(X_i, \bar{\theta}_i, \theta)$, where $\bar{\theta}_i$ is a deterministic quantity (or vector). The issue to be faced now is to estimate parameter θ by available observations and the varying data $\{\bar{\theta}_i\}$. Naturally, some restrictions upon the extra data set $\{\bar{\theta}_i\}$ are required to guarantee the consistence.

In aforementioned iid case with likelihood function V , under mild conditions the ML estimator $\hat{\theta}_n$ is strong consistent (see Theorem 4.3 in chapter 1 of book [12]), i.e., $\hat{\theta}_n \theta$ with probability 1 as $n \rightarrow \infty$. Below is a similar result established in [7] for the case with likelihood function as V .

Let Θ be a bounded closed set in \mathcal{R}^k , $f(x, \bar{\theta}_i, \theta)$ be a continuous function of $\theta \in \Theta$ for almost all $x \in \mathcal{X}$, $i = 1, 2, \dots$, $\{\bar{\theta}_i\}$ belongs to a compact set, and let the following conditions be fulfilled:

(i) For any different $\theta, \theta' \in \Theta$, there exist fixed x_0 and $\varepsilon_0 > 0$ s.t. $\inf_j \max_{i \in A_j} \left| \sqrt{f(x_0, \bar{\theta}_i, \theta)} - \sqrt{f(x_0, \bar{\theta}_i, \theta')} \right| > \varepsilon_0$ for a series of nonintersecting and successive index sets for the whole natural numbers as $\{A_j, j = 1, 2, \dots\}$ with the volume of A_j less than or equaling to a constant positive integer.

(ii) For all $\theta \in \Theta$, the derivative of $\sqrt{f(x, \bar{\theta}_i, \theta)}$ satisfies $\sup_i \int_{\mathcal{X}} \sup_{\theta \in \Theta} \left| \frac{\partial \sqrt{f(x, \bar{\theta}_i, \theta)}}{\partial \theta} \right| d\mu < \infty$.

Then, for any fixed $\theta \in \Theta$ the ML estimator $\hat{\theta}_n$ given by V tends to θ as $n \rightarrow \infty$ with probability 1.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (Grant No.: 2020R111A3072688).

REFERENCES

[1] O. I. Khalaf and B. M. Sabbar, An overview on wireless sensor networks and finding optimal location of nodes, *Periodicals of Engineering and Natural Sciences (PEN)* vol. 7, no. 3, pp. 1096-1101, 2019.

[2] J. Liang, M. Huan, X. Deng, T. Bao and G. Wang, Optimal transmitter and receiver placement for localizing 2D interested-region target with constrained sensor regions. *Signal Processing*, vol. 183, pp. 108032, 2021.

[3] J. M. Castro-Arvizu, J. Vilà-Valls, A. Moragrega, P. Closas, and J. A. Fernandez-Rubio, Received signal strength-based indoor localization using a robust interacting multiple model-extended Kalman filter algorithm, *International Journal of Distributed Sensor Networks* vol.13, no. 8 pp. 155014771772215, 2017.

[4] D. A. Ando, R. K. Miranda, J. P. C. L. da Costa and M. T. de Oliveira, A novel direction of arrival estimation algorithm via received signal strength of directional antennas, 2018 Workshop on Communication Networks and Power Systems (WCNPS), November 2018.

[5] R. Niu, A. Vempaty and P. K. Varshney, Received-signal-strength-based localization in wireless sensor networks, *Proceedings of the IEEE*, vol. 106, no. 7, pp. 1166-1182, 2018.

[6] Y. Miao, H. Wu and L. Zhang, The accurate location estimation of sensor node using received signal strength measurements in large-scale farmland. *Journal of Sensors*, vol. 2018, 2018.

[7] X.-L. Hu, P.-H. Ho and L. Peng, Performance Analysis of Maximum Likelihood Estimation for Transmit Power Based on Signal Strength Model, *Journal of Sensor and Actuator Networks*, vol.7, no. 3, pp. 38, 2018.

[8] M. Ghaderibaneh, M. Dasari and H. Gupta, Multiple transmitter localization under time-skewed observations, 2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), November 2019.

[9] M. Saadati and J. K. Nelson, Multiple transmitter localization using clustering by likelihood of transmitter proximity, 2017 51st Asilomar Conference on Signals, Systems, and Computers. IEEE, 2017: 1769-1773.

[10] Wu C M, Lo C P. Distributed MAC protocol for multichannel cognitive radio ad hoc networks based on power control[J]. *Computer Communications*, November 2017.

[11] V. Gupta and B. Singh, Study of range free centroid based localization algorithm and its improvement using particle swarm optimization for wireless sensor networks under log normal shadowing[J]. *International Journal of Information Technology*, vol. 12, no.3, pp. 975-981, 2020.

[12] I. A. Ibragimov, R. Z. Has'minskii, *Statistical estimation, asymptotic theory. Applications of Mathematics*, vol. 16, Springer-Verlag, New York, 1981.

[13] M. Mehrabi, A. Tchamkerten and M. Yousefi, Bounds on the approximation power of feedforward neural networks *International Conference on Machine Learning*, June 2018.