Article

Performance Analysis of Maximum Likelihood Estimation for Transmit Power Based on Signal Strength Model

Xiao-Li Hu [1,2] , Pin-Han Ho [3] and Limei Peng [4,*]

1   Nanfang College, Sun Yat-Sen Univeristy, Guangzhou 510900, China; hutby@yeah.net
2   College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China
3   Department of Electrical and Computer Engineering, University of Waterloo,
     Waterloo, ON N2L3G1, Canada; p4ho@uwaterloo.ca
4   School of Computer Science and Engineering, Kyungpook National University, Daegu 41566, Korea
*   Correspondence: auroraplm@knu.ac.kr; Tel.: +82-053-950-5553

check for
updates

Abstract: We study theoretical performance of Maximum Likelihood (ML) estimation for transmit power of a primary node in a wireless network with cooperative receiver nodes. The condition that the consistence of an ML estimation via cooperative sensing can be guaranteed is firstly defined. Theoretical analysis is conducted on the feasibility of the consistence condition regarding an ML function generated by independent yet not identically distributed random variables. Numerical experiments justify our theoretical discoveries.

Keywords: maximum likelihood; wireless network; cooperative receiver; cooperative sensing

1. Introduction

It has been an important task in the estimation of the transmit power and location of a node based on a set of cooperative receiver (or monitor) nodes in wireless networks. Such an estimation is particularly essential for cooperative sensing based cognitive radio (CR) [1] networks to enable various sensing mechanisms, such as the estimation for the maximum interference-free transmit power of a frequency-agile radio in an Opportunistic Spectrum Sharing system [2]. As pointed out in [3], transmit power estimation is also a fundamental functional block for the detection of adverse behaviors, such as signal jamming attacks [4] and channel capturing [5].

As a special case of transmit power estimation, node positioning based on received signal strength observations has been extensively studied in [6–8], where the transmit power itself is assumed known. In [9], transmit power estimation was studied by using an ad hoc optimization method. A simple yet rough estimator was proposed by a geometrical approach with a deterministic model equation in [3]. However, this method is not an exact way for estimation and may cause a large deviation under some circumstances. In [2], an maximum likelihood (ML) estimation method was introduced to calculate the concerned maximum interference-free transmit power. However, the performance of ML estimation lacks relevant and comprehensive analysis. In [10], deep theoretical results regarding the performance of ML estimation for transmit power have been established by assuming that the observation locations are random variables. The popular ML method has also been applied to [11,12] recently.

Specifically, we concern a wireless network consisting of a primary node, or transmit node, as well as a set of receiver nodes that are listening to the signal launched by the primary node. We assume that the primary node transmits at a constant power level during the observation period, and the receiver nodes, with their locations known previously, can exchange their respective received power

information with each other [2,3]. The problem is estimating the transmit power and location of the primary node based on a lognormal shadowing model [2,13].

Mathematically, we consider an estimation issue regarding the observation model given by a lognormal shadowing model ([2,13]) as

$$s_i = s_p - 10\epsilon \log_{10} \sqrt{(x_i - x_p)^2 + (y_i - y_p)^2} + w_i, \tag{1}$$

where ϵ is the path-loss exponent and w_i is typically a zero mean Gaussian noise $\mathcal{N}(0, \sigma^2)$, $i = 1, 2, \ldots, n$. There are four constant parameters in total in the model (1), i.e., transmit power s_p, path-loss exponent ϵ, and transmitter location (x_p, y_p). The other varying variables except noise are the received power level s_i at the i-th monitor node and the monitor location (x_i, y_i), $i = 1, 2, \ldots, n$. The estimation issue considered in the above references, e.g., [10], are all with respect to (s_p, x_p, y_p), i.e, path-loss exponent ϵ is assumed known. In this paper, we consider a more general estimation issue with respect to the four parameter vector $\theta = (\epsilon, s_p, x_p, y_p)$ for model (1), based on an available data set $\{s_i, x_i, y_i,\ i = 1, 2, \ldots, n\}$. In light of the fact that the observation data may be collected from a finite number of monitor locations, it is reasonably assumed that the monitor location (x_i, y_i) is deterministic, rather than random variable as in [10].

Because of the nonlinearity of the developed models with respect to the parameters to be identified, most classic methods for linear model in [14] are not applicable. However, the ML method is an exception, which is effective for any kind of parametric model if the distribution of noise is known. By assuming that the noise is an independent identically distributed sequence obeying $\mathcal{N}(0, \sigma^2)$, the derived optimization objective function is a summation of the squared difference between the two sides of model Equation (1) (discarding the noise term on the right-hand side). The optimization problem is calculated by Matlab optimization toolbox (2010b). In an iterative process, the former estimation is taken as an initial starting point for the next calculation. In this contribution, we are motivated to analyze the performance of ML-based estimation for parameter θ, since the issue to find the root of the ML function is a standard optimization problem that can be solved by many mathematical or engineering software programs.

Although in practice an ML-based estimation method is usually implemented by a finite data set, there could be a concern regarding the asymptotical performance of the developed ML estimation model as the data volume n tends to infinity. If the ML estimation is biased even when n tends to infinity, how can one expect for the effectiveness with a fixed n? Therefore, the major effort of this paper is to investigate the asymptotical performance of the ML estimate as $n \to \infty$, i.e., when it succeeds and when it fails. It turns out that the monitor location set $\{(x_i, y_i), i = 1, 2, \ldots\}$ should be rich enough in some sense to guarantee the n-th ML estimate, denoted as θ_n, tends to θ as $n \to \infty$. To our knowledge, a special case of this issue has been considered in [10], where the performance of ML estimation for parameters (s_p, x_p, y_p) has been analyzed by assuming that $\{(x_i, y_i)\}$ is a sequence of iid random variables. However, as mentioned above, it is more natural to view the receiver node locations $\{(x_i, y_i)\}$ as deterministic values. Hence, in this paper, we will consider the performance of ML estimation for $\theta = (\epsilon, s_p, x_p, y_p)$ as the number of receiver nodes n tends to infinity by information from deterministic receive nodes.

The main contributions of this paper are as follows.

- We identify the condition that the consistence of ML estimation for transmit power can be guaranteed; i.e., $\theta_n \to \theta$ as $n \to \infty$. Exactly as shown in Remark 1, to achieve a correct observation, the receiver nodes' location set $\{(x_i, y_i), i = 1, 2, \ldots\}$ should not be merely located in a criterion curve (a term given by Definition 2). The variety of the location set is to guarantee the persistently excitation condition, which is used for the convergence of the Least Squares algorithm [14].
- Theoretically, we consider the consistence issue for a ML function generated by independent yet not identically distributed random variables. A simple criterion for checking the consistence is given in Theorem 3.

The rest of this paper is organized as follows. The proposed model and ML estimation for transmit power are presented in Section 2. An extreme case related to consistence condition for ML estimation of transmitter power is studied in Section 3, which may help to establish some intuitive sense for the main results of this paper. Then, as the main practical contribution of the paper, the technical mechanisms beneath the numerical experiments are demonstrated and justified in Section 4. In Section 5, some numerical experiments are designed to detect the performance of the ML estimate algorithm, and explained by relevant theoretical results. The main theoretical contribution is in Section 6, where some theoretical preliminaries to be used in verification of the theoretical results in the former section are prepared. Concluding remarks of the paper are given in Section 7.

2. Maximum Likelihood Estimation for Transmit Power in Localized Signal Strength Model

As aforementioned, we limit our discussion to Signal-Strength (SS) based localization of a single primary transmitter in the geographic coverage area [2]. Let $L = (x_p, y_p)$ denote the location of the primary transmitter. Suppose that a sequence of uncorrelated observed SS measurements, $s_1, \ldots, s_n$, along with the corresponding position coordinates $L_1, \ldots, L_n$, where $L_i = (x_i, y_i)$, $i = 1, \ldots, n$, are available. The set of observations

$$\mathcal{O}(n) \overset{\Delta}{=} \{(s_i, L_i) : i = 1, \ldots, n\},$$

may be obtained in different ways. For example, consider a scenario in which n receiver nodes, located at positions $L_1, \ldots, L_n$, collect the signal strength observations $s_1, \ldots, s_n$ at a given time. These receiver nodes exchange their data among each other, such that at least one of these nodes receives the entire set $\mathcal{O}$. Naturally, the observation set $\mathcal{O}$ may also be obtained by measurements from a single receiver node at n different points in time along a trajectory as the node moves in the coverage area. In general, a given observation (S_i, L_i) may be obtained either from a measurement taken by the receiver node itself in the past, or from a measurement at another receiver node that shares the information between the two nodes. Actually, we will find below that there are a set of requirements on the number of receiver nodes and their locations that can satisfy the consistence condition of ML estimation of transmit power.

The observation equation is given by a lognormal shadowing model ([2,13]) as

$$s_i = s_p - 10\epsilon \log_{10} d_{ip} + w_i, \tag{2}$$

where noise w_i follows a Gaussian distribution $\mathcal{N}(0, \sigma^2)$, and

$$d_{ip} = \sqrt{(x_i - x_p)^2 + (y_i - y_p)^2}. \tag{3}$$

Let us formulate below the ML estimate for θ, or parameters ϵ, s_p, x_p, y_p, based on the given observation data $\mathcal{O}(n)$. Mathematically, when the locations (x_i, y_i) are viewed as deterministic quantity, the random variable s_i has a Gaussian distribution $\mathcal{N}(s_p - 10\epsilon \log_{10} d_{ip}, \sigma^2)$, i.e., its density is

$$f(s_i; x_i, y_i, \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(s_i - s_p + 10\epsilon \log_{10} d_{ip})^2}{2\sigma^2} \right\}. \tag{4}$$

Note that, though $\{s_i\}$ is a sequence of independent random variables, they have different distributions since the locations (x_i, y_i) are probably different with different i.

Clearly, the corresponding ML function is

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(s_i - s_p + 10\epsilon \log_{10} d_{ip})^2}{2\sigma^2} \right\}, \tag{5}$$

and the log ML function can be written as

$$\log \mathcal{L}(\theta) = -\frac{n}{2}\log(2\pi) - n\log\sigma$$

$$-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(s_i - s_p + 10\epsilon\log_{10} d_{ip})^2,$$

which is equivalent to minimize

$$M_n(\theta) \triangleq \sum_{i=1}^{n}(s_i - s_p + 10\epsilon\log_{10} d_{ip})^2. \tag{6}$$

This turns to be a nonlinear optimization problem regarding ϵ, s_p, x_p, y_p, which can be solved by many mathematical software packages such as GAMs (General Algebraic Modeling System) and Matlab toolbox, or by designing a special program. As proposed in the Introduction, we are motivated to investigate the asymptotical performance of the ML estimate as n increases. Precisely, we are interested to know when the ML-based n-th estimate for $\theta = (\epsilon, s_p, x_p, y_p)$ tends to the true value as n tends to infinity.

3. Unique Solution Condition for a System of Observation Equations without Noise

In this section, we consider an extreme case related to a consistence condition for ML estimation of transmitter power, which may help to establish some intuitive sense for the main results of this paper. A similar idea can be found in the proof of Theorem 3 in [10]. By (2) and (6), we have

$$M_n(\theta) = \sum_{i=1}^{n} w_i^2. \tag{7}$$

Thus, $M_n(\theta) = 0$ if there is no noise in the model Equation (2). This further means that each term in the summation (6) equals zero, i.e.,

$$s_i = s_p - 10\epsilon\log_{10}\sqrt{(x_i - x_p)^2 + (y_i - y_p)^2}, \quad i = 1, 2, \ldots, n. \tag{8}$$

We naturally want these n equations regarding $\theta = (\epsilon, s_p, x_p, y_p)$ to have a unique solution. Mathematically, the solution is not unique if we find another $\theta' = (\epsilon', s_p', x_p', y_p') \neq \theta$ such that the data set $\{(x_i, y_i)\}_1^n$ synchronously satisfies the following n equations:

$$s_i = s_p' - 10\epsilon'\log_{10}\sqrt{(x_i - x_p')^2 + (y_i - y_p')^2}, \quad i = 1, 2, \ldots, n. \tag{9}$$

Combining (8) and (9), one gets

$$\frac{d_{ip}^{\epsilon}}{d_{ip'}^{\epsilon'}} = 10^{\frac{1}{10}(s_p - s_p')} \triangleq \lambda(s_p, s_p') \tag{10}$$

for $i = 1, 2, \ldots, n$, where $d_{ip'} = \sqrt{(x_i - x_p')^2 + (y_i - y_p')^2}$ and d_{ip} given by (3). Thus, the solution of n equations given by (8) is not unique if the data set $\{(x_i, y_i)\}_1^n$ satisfies (10) with certain $\theta' = (\epsilon', s_p', x_p', y_p')$. Conversely, if the solution is unique, the whole location set should never satisfy (10) for any given θ'.

After replacing x_i, y_i as x, y in (10), a curve is defined by equation of x, y as

$$\left[(x - x_p)^2 + (y - y_p)^2\right]^{\epsilon/2} = \lambda\left[(x - x_p')^2 + (y - y_p')^2\right]^{\epsilon'/2} \tag{11}$$

with given parameters $\lambda = 10^{\frac{1}{10}(s_p - s'_p)} > 0$, (ϵ, x_p, y_p) and (ϵ', x'_p, y'_p). For convenience, let us call the curve given by (11) a criterion curve with respect to (ϵ, x_p, y_p), (ϵ', x'_p, y'_p), and λ. The class of these criterion curves can be used to to determine whether the solution of a system of n equations given by (8) is unique. Precisely, the condition guaranteeing uniqueness can be presented as: the data set $\{(x_i, y_i)\}_1^n$ should not be located in any single criterion curve.

Below, we will analyze some qualities of criterion curves in detail under general setting and a special case, respectively.

3.1. Qualities of the Criterion Curve

Let us first point out a qualitative property of criterion curve given by (11). If $\epsilon \neq \epsilon'$, say $\epsilon < \epsilon'$, we find that

$$\lim_{x,y \to \infty} \frac{d_{ip}^{\epsilon}}{d_{ip'}^{\epsilon'}} = 0. \tag{12}$$

This means, for any given $\lambda > 0$, the criterion curve defined by (11) is impossible to extend to infinity. However, if $\epsilon = \epsilon'$, we know that the curve turns to be a circle unless $\lambda = 1$. Another immediate discovery is that the curve separates the two points (x_p, y_p) and (x'_p, y'_p), due to the fact that the left-hand side of (11) is less than its right-hand side when substituting $x = x_p$ and $y = y_p$ into (11), and the inequality reversed by substituting $x = x'_p$ and $y = y'_p$. Let us summarize these facts as a proposition below.

Proposition 1. *The criterion curve defined by (11) is a straight line if $\epsilon = \epsilon'$ and $\lambda = 1$. Otherwise, it is a bounded curve. In both cases, the curve separates the two points (x_p, y_p) and (x'_p, y'_p).*

By Equation (11), a criterion curve is symmetrical to the straight line passing through points (x_p, y_p) and (x'_p, y'_p). To know how many points a criterion curve intersects the mentioned straight line, let us introduce a parametrical presentation of the straight line as

$$x = x_p + t\alpha, \quad y = y_p + t\beta, \tag{13}$$

where $\alpha = x'_p - x_p$ and $\beta = y'_p - y_p$. By substituting (13) into (11), we have

$$t^{2\epsilon} = \delta(t-1)^{2\epsilon'}, \tag{14}$$

where $\delta = \lambda^2(\alpha^2 + \beta^2)^{\epsilon' - \epsilon}$. It is easy to see that the Equation (14) (with respect to t) has at least two solutions and at most four solutions for $\epsilon \neq \epsilon'$. If $\epsilon = \epsilon'$ and $\lambda = 1$, it has just one solution $t = \frac{1}{2}$; and it has two solutions for $\lambda \neq 1$. Hence, there are three different structures for the criterion curves if $\epsilon \neq \epsilon'$; and two different structures if $\epsilon = \epsilon'$. Let us check it by numerical experiments. By setting $(x_p, y_p) = (-2, 0)$, $(x'_p, y'_p) = (2, 0)$, $\epsilon = 1.5$, $\epsilon' = 2.5$, we plot graphs of the criterion curves for $\lambda = 0.05, 0.046487, 0.045$ respectively in Figures 1–3. The reason why λ selects these three values is that, under the given parameter setting, $\lambda = 0.046487$ is found numerically at a critical point corresponding to the critical curve among the other two kinds of curves. Thus, the other two values 0.05 and 0.045 from both sides of the critical value are chosen to show the bifurcation phenomenon.

The latter case will be analyzed extensively in the next subsection, so we just summarize the case $\epsilon \neq \epsilon'$ below.
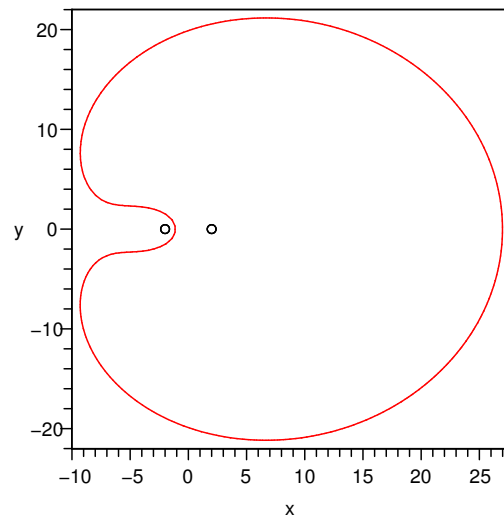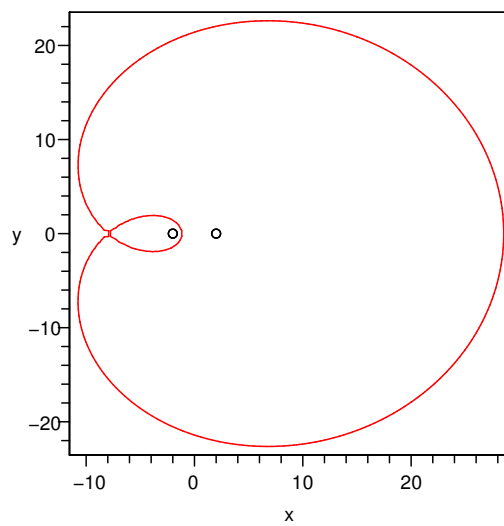
Figure 1. $\lambda = 0.05$.
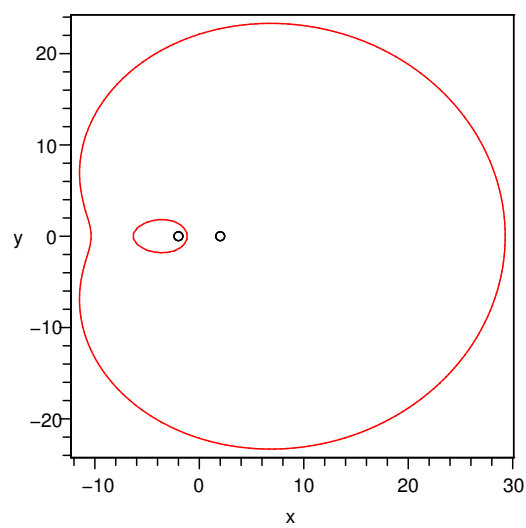

Figure 2. $\lambda = 0.046487$.

Figure 3. $\lambda = 0.045$.

Proposition 2. *For $\epsilon \neq \epsilon'$, the criterion curves defined by Equation (11) can be divided into 3 kinds of curves according to the number of intersecting points of a criterion curve and the straight line determined by the two points (x_p, y_p) and (x'_p, y'_p), i.e., the number of roots of Equation (14). Specifically, the three cases are as follows:*

(a) If the number is 2, the curve is a closed connected curve. A typical graph of this case is shown in Figure 1;
(b) If the number is 3, the curve is a closed connected curve. A typical graph of this case is shown in Figure 2;
(c) If the number is 4, the curve is two separated curves. A typical graph of this case is shown in Figure 3.

3.2. Qualities of the Criterion Curve if ϵ Is Known

If the path-loss exponent ϵ is known, i.e., in this case, the ML estimation is proposed for (s_p, x_p, y_p) as in [10], by a similar reasoning process, we have the criterion curve as

$$\sqrt{(x - x_p)^2 + (y - y_p)^2} = \lambda \sqrt{(x - x'_p)^2 + (y - y'_p)^2} \tag{15}$$

with given parameters $\lambda = 10^{\frac{1}{10\epsilon}(s_p - s'_p)} > 0$, (x_p, y_p) and (x'_p, y'_p).

Clearly, the curve is a straight line when $\lambda = 1$, and a circle when $\lambda \neq 1$. Though these basic facts have been discovered in [10], we describe some more general and deeper qualities in the following propositions.

Proposition 3. *When $\lambda \neq 1$, the circle defined by Label (15) is centered at*

$$\left(\frac{x_p - \lambda^2 x'_p}{1 - \lambda^2}, \frac{y_p - \lambda^2 y'_p}{1 - \lambda^2} \right) \tag{16}$$

with radius $\frac{\lambda}{|1-\lambda^2|} \sqrt{(x_p - x'_p)^2 + (y_p - y'_p)^2}$, i.e., $\frac{\lambda}{|1-\lambda^2|}$ times the distance between (x_p, y_p) and (x'_p, y'_p). While $\lambda = 1$, the straight line defined by (15) is perpendicular to the line connecting (x_p, y_p) and (x'_p, y'_p) and containing their midpoint.

In the circle case, the radius tends to infinity when λ tends to 1, which also means the circle turns out to be a straight line.

With s_p and $P \overset{\Delta}{=} (x_p, y_p)$ fixed, an important topic is checking whether a given circle is a criterion curve, and, if it is, how we can find the corresponding $P' \overset{\Delta}{=} (x'_p, y'_p)$. As shown in Figure 4, given the point P and a circle centered at C with radius r, it is important to figure out whether the circle is a criterion curve, and what P' is at this moment.

Let us change the Formula (16) into a vector form:

$$\vec{C} = \frac{\vec{P} - \lambda^2 \vec{P'}}{1 - \lambda^2}. \tag{17}$$

Thus, we have

$$\vec{P'} = \frac{\vec{P} + (\lambda^2 - 1)\vec{C}}{\lambda^2}. \tag{18}$$

Hence, P', shown in Figure 4, can be calculated from (18) directly if λ, P and C are known.

Let us deduce a formula to calculate λ from given information: points P and C, and circle radius r. In the case of $\lambda > 1$, shown in Figure 4, by the aforementioned fact, it follows that

$$r = \frac{\lambda}{|1 - \lambda^2|} |PP'| \tag{19}$$

and

$$|PP'| = \frac{|\lambda^2 - 1|}{\lambda^2}|PC|, \tag{20}$$

derived by (17). Thus, from (19) and (20), we derive

$$r = \frac{1}{\lambda}|PC|, \tag{21}$$

and thus

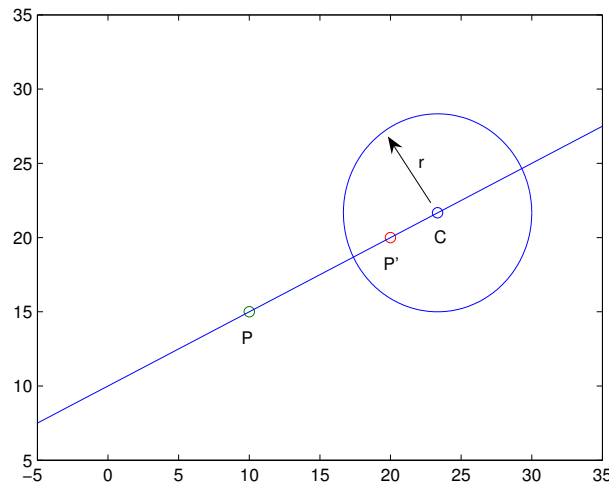$$\lambda = \frac{|PC|}{r}. \tag{22}$$

Figure 4. Calculation of P' based on P and a circle.

Similar results can be derived in the case $0 < \lambda < 1$.

Therefore, now we can calculate P' by Formulas (18) to (22) when $\lambda \neq 1$, i.e., $r \neq |PC|$. In the case $\lambda = 1$, or $r = |PC|$, we know that the corresponding curve should be a straight line rather a circle. Thus, if a given circle contains the point P, it is impossible to be a criterion curve. Let us summarize these facts as a proposition below.

Proposition 4. *Assuming points $P = (x_p, y_p)$ and $C = (x_c, y_c)$, and a circle centered at C with radius $r > 0$ are given, then the following assertions hold:*

(i) If the circle contains point P, i.e., $r = |PC|$, then the circle is not a criterion curve. Moreover, a straight line traversing through P is not a criterion curve.

(ii) If $r \neq |PC|$, then the circle is a criterion curve, and the corresponding $P' = (x'_p, y'_p)$ in (15) can be calculated as:

$$\vec{P'} = \frac{\vec{P} + (\lambda^2 - 1)\vec{C}}{\lambda^2} \tag{23}$$

with $\lambda = |PC|/r$.

Another obvious fact is that, under case (ii) of Proposition 4, the circle contains either one of P and P', and the other is out of the circle, as shown in Figure 4. This also serves as an explanation as to why it fails to be a criterion curve if a circle contains P itself. Moreover, three noncollinear points are not located in any criterion curve, if the circle determined by the three points contains P.

4. Consistence Condition for ML Estimation of Transmitter Power

In this section, we deduce the consistence condition to guarantee convergence of the ML estimate for θ, i.e., $\theta_n \to \theta$ as $n \to \infty$. Roughly speaking, the consistence condition is that the location set $\{(x_i, y_i)\}$ should be rich enough in some sense. A criterion is given and proved.

Let us introduce the definition for consistence condition, which is a requirement for richness of the locations $\{(x_i, y_i)\}$, as introduced in the former section. The requirement is somewhat similar to 'persistence of excitation condition' in a least squares algorithm to guarantee the convergence.

Let us introduce a limit set of a sequence of point sets in terms of standard Euclidean distance.

Definition 1. *Denote a series of nonintersecting and successive index sets for the whole natural numbers as $\{A_j, j = 1, 2, \ldots\}$ with $|A_j| \leq m$. Let $S_k = \{P_i \in R^n, i \in A_k\}$ is a point set in R^n. A point $P \in R^n$ is called a limit point of the sequence of point sets $\{S_k\}$, if for $\forall \varepsilon > 0$, there exists $j_0 > 0$ and $i_j \in A_j$ such that*

$$|P_{i_j} - P| < \varepsilon$$

holds for $j > j_0$. The set comprised of all such limit points is called the limit set of the set sequence $\{S_k\}$, denoted as $Lim(S_k)$.

For example, letting

$$S_k = \left\{ \frac{1}{k}, -1 - \frac{1}{k^2}, 1 + e^{-k} \right\},$$

we have its limit set to $Lim(S_k) = \{0, -1, 1\}$.

Below we define the criterion curve strictly, which actually has been introduced in Section 3.

Definition 2. *A criterion curve is defined by equation of x, y as*

$$\left[(x - x_p)^2 + (y - y_p)^2 \right]^{\epsilon/2} = \lambda \left[(x - x_p')^2 + (y - y_p')^2 \right]^{\epsilon'/2} \tag{24}$$

with given parameters $\lambda = 10^{\frac{1}{10}(s_p - s_p')} > 0$, (ϵ, x_p, y_p) and (ϵ', x_p', y_p').

Remark 1. *When $\epsilon = \epsilon'$, a criterion curve is a circle for if $\lambda \neq 1$ and a line if $\lambda = 1$, as already shown in Section 3. Thus, under this setting, four noncollinear and nonconcyclic points are impossible to locate in any single criterion curve.*

Now, we are in a position to state a consistence result for ML estimation for transmitter power model.

Theorem 1. *Suppose that the location set $\{x_i, y_i\}_1^\infty$ is in a bounded region, and denote*

$$S_k \triangleq \{(x_{km+i}, y_{km+i}), i = 1, \ldots, m\}, \tag{25}$$

where $m \geq 3$ is a positive integer. If the limit set $Lim(S_k)$ does not locate in any single criterion curve given by (24), then the ML estimation for model (2) with likelihood function given by (5) is very consistent, i.e., the estimates converge to the true vale with probability 1 as the data number tends to infinity.

Proof. Theorem 3 is used to prove the above theorem. Since the model is assumed to be located in a local region, we need only to check the two conditions therein. By the likelihood function given by (5) and density for s_i given by (4), we have

$$\frac{\partial \sqrt{f}}{\partial s_p} = \frac{\sqrt{f}(s_p - s_i - 10\epsilon \log_{10} d_{ip})}{2\sigma^2},$$

$$\frac{\partial \sqrt{f}}{\partial \epsilon} = -\frac{\partial \sqrt{f}}{\partial s_p} \cdot 10 \log_{10} d_{ip},$$

$$\frac{\partial \sqrt{f}}{\partial x_p} = -\frac{\partial \sqrt{f}}{\partial s_p} \cdot \frac{10\epsilon(x_p - x_i)}{d_{ip}^2 \log 10},$$

$$\frac{\partial \sqrt{f}}{\partial y_p} = -\frac{\partial \sqrt{f}}{\partial s_p} \cdot \frac{10\epsilon(y_p - x_i)}{d_{ip}^2 \log 10},$$

where $\sqrt{f}$ used on the right-hand side is used to make the expressions brief. By the fact that all parameters belong to a compact set, the condition (ii) of Theorem 3 is satisfied.

The continuity of $\sqrt{f}$ with respect to all variables is obvious. By Remark 5, we need only to verify (53). To show (53) for density (4), it is sufficient to show that, for any different θ and θ', there exist s and $(x, y) \in \mathrm{Lim}(S_k)$ such that

$$f(s; x, y, \theta) \neq f(s; x, y, \theta'). \tag{26}$$

By substituting (4) to (26), we have

$$\left[(x - x_p)^2 + (y - y_p)^2\right]^{\epsilon/2} \neq \lambda \left[(x - x_p')^2 + (y - y_p')^2\right]^{\epsilon'/2}, \tag{27}$$

which means that there is no criterion curve containing all points of the limit set $\mathrm{Lim}(S_k)$. In other words, this is equivalent to say that the limit set $\mathrm{Lim}(S_k)$ is not located in any single criterion curve given by (24), which finishes the proof of the theorem. □

Based on Theorem 1, the ML estimate for transmitter power may fail even if the number of observation tends to infinity if the whole location set $\{x_i, y_i\}_1^\infty$ belongs to a single criterion curve. Below are two typical cases when ϵ is known.

(i). The set $\{x_i, y_i\}_1^\infty$ is located in a straight line, which means $\lambda = 1$ and $\epsilon = \epsilon'$ in (24). In this case, the ML estimator for location (x_p, y_p) may tend to its symmetrical point with respect to the line yet the estimate for s_p still works, as in Example 2. Furthermore, if the set $\{x_i, y_i\}_1^\infty$ is asymptotically located in a straight line, the same phenomenon happens, as shown in Example 4.

(ii). The set $\{x_i, y_i\}_1^\infty$ is located in a circle, which means $\lambda \neq 1$ and $\epsilon = \epsilon'$ in (24). In this case, the ML estimate may tend to another location $L_p' = (x_p', y_p')$ given by (23), and estimate for s_p tends to s_p', as in Example 3.

However, even in the two cases above, it is still possible to see that the ML estimators converge to the true value because the true value are also the solution of the likelihood function.

5. Numerical Algorithm and Experiments

Let us first develop a recursive numerical algorithm to minimize $M_n(\theta)$ in (6) by using Matlab code/fminsearch/. When the n-th ML estimate for θ is numerically found to be θ_n, the next minima θ_{n+1} for $M_{n+1}(\theta)$ is numerically searched based on θ_n as:

$$\theta_{n+1} = \mathrm{fminsearch}(@M_{n+1}, \theta_n).$$

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

For simplicity, the numerical experiments below are designed to investigate the performance of ML estimation for parameters s_p, x_p, y_p, instead of $\theta = (\epsilon, s_p, x_p, y_p)$. Naturally, similar phenomena

exist for ML estimation of θ. Four numerical experiments below are designed to investigate the asymptotical performances of the ML estimation for (s_p, x_p, y_p) in model (2) as the observation number tends to infinity. The first one achieved a successful observation while the rest three failed. The essential reason for the failure is that the likelihood function (5), or the equivalent function $M_n(\cdot, \cdot, \cdot)$ in (6), has multiple roots when some conditions are not met. Note that, even in the three examples of failure, 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.

Example 1. *Let $\epsilon = 1$, $x_p = 20$, $y_p = 0$, $s_p = 80$ in (2) and (3), $w_i \sim \mathcal{N}(0, 0.5^2)$, and assume all locations (x_i, y_i) are all located in the set $\{(0,0), (40,0), (0,-20), (0,20)\}$ successively infinite times. The performance has been shown in Figure 5.*
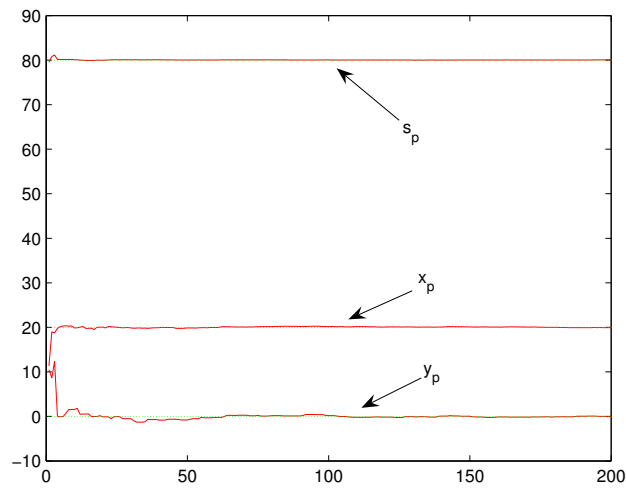

Figure 5. Example 1.

Clearly, the corresponding location set S_k in Theorem 1 is:

$$S_k = \{(0,0), (40,0), (0,-20), (0,20)\}.$$

Hence, $\mathrm{Lim}(S_k)$ is still comprised of these four noncollinear and nonconcyclic points. By Theorem 1 and Remark 1, the ML estimation for (s_p, x_p, y_p) will definitely converge to all true values as shown in Figure 5, which coincides with the numerical experiments.

The set S_k does not need to own four noncollinear and nonconcyclic points. As a matter of fact, A_k consists of three noncollinear points (or three collinear points) may still work.

Example 2. *Under the same setting of Example 1, expect that all locations (x_i, y_i) are located in the set $\{(x,y):x = 0, y \in [-100,100]\}$. The performance has been shown in Figure 6.*

The location set $\{(x_i, y_i), i = 1, 2, \ldots\}$ is obviously located in a criterion curve, since they are all located in the straight line $x = 0$. The corresponding $P' = (x'_p, y'_p) = (-20, 0)$ in (15) by the fact $P = (20, 0)$. In addition, the corresponding $\lambda = 1$ under this case, so, by (10), $s'_p = s_p$. This means that the ML estimates for s_p and y_p have no bias, while the ML estimate for x_p may tend to x'_p, which has been discovered by the experiment. By observing that the symmetric point of $(20, 0)$ with respect to the line $x = 0$ is $(-20, 0)$, the ML estimators for x_p may converge to -20, yet the estimators for y_p and s_p still work, as shown in Figure 6.
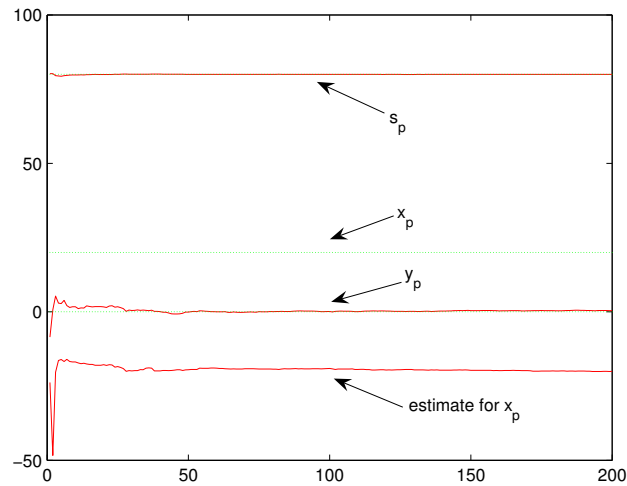
Figure 6. Example 2.

Example 3. *Under the same setting of Example 1, expect that all locations (x_i, y_i) are located in the set*

$$\{(x, y): \sqrt{(x - 20)^2 + y^2} = \lambda\sqrt{(x + 20)^2 + y^2}\}$$

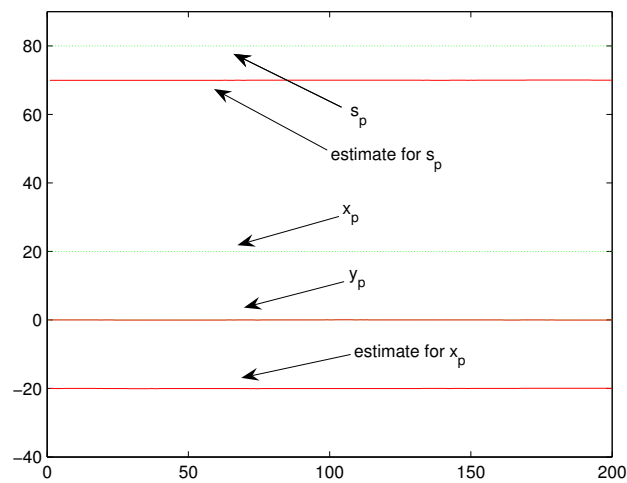*with $\lambda = 10$. The performance has been shown in Figure 7.*

Figure 7. Example 3.

The location set $\{(x_i, y_i), i = 1, 2, \ldots\}$ in Example 3 is also located in a criterion curve, since they are all located in the same circle with center $(-2020/99, 0)$ with radius $r = 400/99$, calculated by Proposition 4. The corresponding $P' = (-20, 0)$ in (15) and the corresponding $\lambda = 10$, so, by (10), $s'_p = 70$. This means that ML estimate for y_p have no bias, while ML estimates for x_p and s_p may tend to x'_p and s'_p respectively, which has been discovered by the experiment, as shown in Figure 7.

Example 4. *Under the same setting of Example 1, expect all locations $\{(x_i, y_i): x_i = \frac{1}{i + 10000}, y_i \in [-100, 100]\}$. The performance has been shown in Figure 8.*

Note that the corresponding location set S_k in Theorem 1 is:

$$S_k = \left\{\left(\frac{1}{10000 + 4k + i}, y\right), y \in [-100, 100], i = 1, 2, 3, 4\right\}.$$

Clearly, $\mathrm{Lim}(S_k)$ is located on line $x = 0$. Thus, it is located on a single criterion curve. A similar phenomenon of Example 2 happens. By observing that the symmetric point of $(20, 0)$ with respect to the line $x = 0$ is $(-20, 0)$, the ML estimators for x_p may converge to -20 yet the estimators for y_p and s_p still work, as shown in Figure 8.

It is similar to design a set of numerical experiments to make the ML estimations for $\theta = (\epsilon, s_p, x_p, y_p)$ all fail. The four simple examples above are just used to demonstrate some basic features of the proposed ML estimate algorithm.
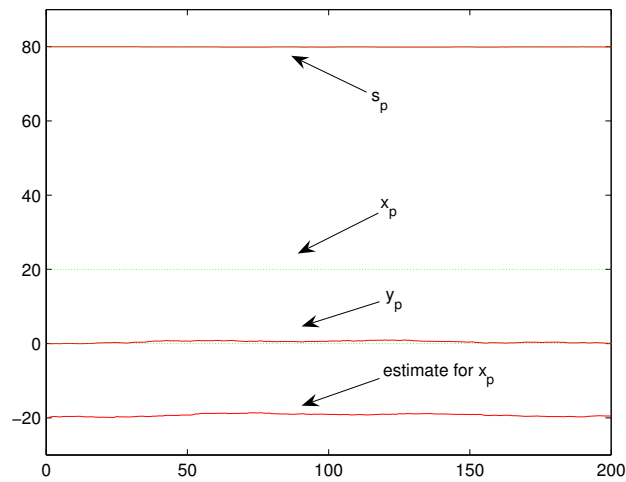

Figure 8. Example 4.

6. Consistence Analysis of Maximum Likelihood Estimation

This section serves as theoretical preliminaries for the former mathematical analysis, which contains general consistence analysis for the proposed ML estimation algorithm based on independent observations with different distributions. As for independent identically distributed observation cases, we refer to Theorem 1.3.3 in Chapter 1 of book [15]. Hence, Theorem 2 here can be taken as an extension of the counterpart in [15]. Furthermore, we propose an easy way of checking the criterion in the following Theorem 3. On the other hand, the readers who are more concerned with practical applications are suggested to pay more attention to Theorem 3.

Although the ML method, suggested by R. A. Fisher, has attracted the interest of mathematicians in theoretical performance analysis for a long time, there are still remain open quite basic issues such as the one investigated in this paper. Thus, this section investigates the basic strong consistence for theoretical preparation.

Let $(\mathcal{X}, \mathcal{U}, P_\theta, \Theta)$ be a statistical experiment, which means that $(\mathcal{X}, \mathcal{U}, P_\theta)$ 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 a certain statistical experiment. Thus, the function $p(X, \theta)$ can be taken as a likelihood function corresponding to the experiment and observation X. The statistic $\hat{\theta}$ defined by

$$\hat{\theta} = \arg\sup_{\theta \in \Theta} p(X, \theta) \tag{28}$$

is called the ML estimator for the parameter θ based on the observation X. In the case of independent identically distributed (iid) observations, $X_1, \ldots, X_n$, 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

$$\hat{\theta}_n = \arg\sup_{\theta \in \Theta} \prod_{i=1}^{n} f(X_i, \theta). \tag{29}$$

Moreover, in many applications, the likelihood function depends on some other varying quantities, e.g., the likelihood function for transmitter power model by (5) involves s_i and (x_i, y_i). This means that the ML estimator $\hat{\theta}_n$ is generally

$$\hat{\theta}_n = \arg \sup_{\theta \in \Theta} \prod_{i=1}^{n} f(X_i, \bar{\theta}_i, \theta), \tag{30}$$

where $\bar{\theta}_i$ is a deterministic quantity (or vector), e.g., $\bar{\theta}_i = [s_i, x_i, y_i]^T$ in (5). The issue to be faced now is to estimate parameter θ via the 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 the aforementioned iid case with likelihood function (29), the ML estimator $\hat{\theta}_n$ is strongly consistent (see Theorem 1.3.3 in Chapter 1 of book [15]), i.e., $\hat{\theta}_n \to \theta$ with probability 1 as $n \to \infty$. Below, we intend to obtain a similar result by using the case with likelihood function as (30).

Theorem 2. *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 $x \in \mathcal{X}$, $i = 1, 2, \ldots$. Let the following conditions be fulfilled with a real number $q \geq 2$:*

(i) Denote a series of nonintersecting and successive index sets for the whole natural numbers as $\{A_j, j = 1, 2, \ldots\}$ with the volume of A_j less than or equal to a constant positive integer m. For all $\theta \in \Theta$ and all $\gamma > 0$, there exists a positive number $\kappa_\theta(\gamma)$ (may depend on θ, q and γ) such that

$$\inf_{j} \max_{i \in A_j} \inf_{|\theta - \theta'| > \gamma} \int_{\mathcal{X}} \left| f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta) - f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta') \right|^q d\mu$$

$$\geq \kappa_\theta(\gamma) > 0. \tag{31}$$

(ii) For any $\bar{\theta}_i$, $\delta > 0$ and $\theta \in \Theta$, there exists a positive number $\omega_\theta(\delta)$ (may depend on θ, q and δ) such that

$$\left(\int_{\mathcal{X}} \sup_{|h| \leq \delta} \left| f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta) - f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + h) \right|^q d\mu \right)^{\frac{1}{q}}$$

$$\leq \omega_\theta(\delta) \xrightarrow[\delta \to 0]{} 0. \tag{32}$$

Then, for any fixed $\theta \in \Theta$, the ML estimator $\hat{\theta}_n$ given by (30) tends to θ as $n \to \infty$ with probability 1.

Let us explain the two conditions briefly here. The condition (i) requires that in any index set A_j there is at least a density function with a different value as θ is changed. This means the data set $\{\bar{\theta}_i, i \in A_j\}$ should be rich enough to make density function $f(x, \bar{\theta}_i, \theta)$ distinguishable according to different values of θ. On the other hand, condition (ii) desires that the density function has certain continuity with respect to θ in some sense.

Remark 2. *This theorem is a twofold extension of Theorem 4.3 in Chapter 1 of book [15]. Firstly, the theorem in book [15] deals with an iid likelihood function as (29), which is just a special case of (30) without $\bar{\theta}_i$ in the likelihood function. Secondly, a parameter q is introduced here in the conditions (i) and (ii) to include the counterpart conditions therein as a special case of $q = 2$, although, $q = 2$ is a wise choice.*

Remark 3. *When we consider a fixed true value θ_0, the condition (i) can be reduced to $\kappa_{\theta_0}(\gamma) > 0$. This means only the difference between (any) θ' and (fixed) θ_0 is required with $|\theta' - \theta_0| > \gamma$.*

Before prove the theorem, let us introduce an extension of Young's inequality [16] below and refer its proof to [17].

Lemma 1. *Suppose that $a \geq 0, b \geq 0$, and $\frac{1}{p} + \frac{1}{q} = 1$ with $1 < p \leq 2 \leq q$. Then,*

$$\frac{1}{q}\left|a^{\frac{1}{q}} - b^{\frac{1}{q}}\right|^q \leq \frac{1}{p}a + \frac{1}{q}b - a^{\frac{1}{p}}b^{\frac{1}{q}} \leq \frac{1}{p}\left|a^{\frac{1}{p}} - b^{\frac{1}{p}}\right|^p, \tag{33}$$

where the equalities hold if $p = q = 2$ or $a = b$.

Remark 4. *This lemma develops an extension of the classic Young's inequality; see, e.g., [16],*

$$a^{\frac{1}{p}}b^{\frac{1}{q}} \leq \frac{1}{p}a + \frac{1}{q}b$$

for $a \geq 0, b \geq 0$, and $\frac{1}{p} + \frac{1}{q} = 1$ with $p > 0, q > 0$, which is a source of lots of important inequalities, e.g., Holder's inequality and Minkowski's inequality. A corollary from the proof is

$$\left|a^{\frac{1}{q}} - b^{\frac{1}{q}}\right|^q \leq \left|a^{\frac{1}{p}} - b^{\frac{1}{p}}\right|^p. \tag{34}$$

The below proof for Theorem 2 shares the main ideas of the counterpart for Theorem 4.3 in [15].

Proof of Theorem 2. Denote

$$\Pi_n(u) = \prod_{j=1}^{n} \prod_{i \in A_j} \frac{f(x_i, \bar{\theta}_i, \theta + u)}{f(x_i, \bar{\theta}_i, \theta)}. \tag{35}$$

Notice that the maximum operator is taking over a finite set A_j in (31) of condition (i), thus there exists an index $i_j \in A_j$ such that

$$\inf_{|\theta - \theta'| > \gamma} \int_{\mathcal{X}} \left| f^{\frac{1}{q}}(x, \bar{\theta}_{i_j}, \theta) - f^{\frac{1}{q}}(x, \bar{\theta}_{i_j}, \theta') \right|^q d\mu \geq \kappa_\theta(\gamma) > 0. \tag{36}$$

Let Γ be a sphere of a small radius δ located entirely in the region $|u| > \frac{1}{2}\gamma$. We shall estimate first $E \sup_\Gamma \Pi_n(u)^{\frac{1}{q}}$. Supposing that u_0 is the center of Γ, then

$$\begin{aligned}
\sup_{u \in \Gamma} \Pi_n(u)^{\frac{1}{q}} &= \sup_{u \in \Gamma} \prod_{j=1}^{n} \prod_{i \in A_j} \left(\frac{f(x_i, \bar{\theta}_i, \theta + u)}{f(x_i, \bar{\theta}_i, \theta)} \right)^{\frac{1}{q}} \\
&\leq \prod_{j=1}^{n} \prod_{i \in A_j} f^{-\frac{1}{q}}(x_i, \bar{\theta}_i, \theta) \left[f^{\frac{1}{q}}(x_i, \bar{\theta}_i, \theta + u_0) \right. \\
&\quad \left. + \sup_{|h| \leq \delta} \left| f^{\frac{1}{q}}(x_i, \bar{\theta}_i, \theta + u_0 + h) - f^{\frac{1}{q}}(x_i, \bar{\theta}_i, \theta + u_0) \right| \right] \\
&\stackrel{\Delta}{=} \prod_{j=1}^{n} \prod_{i \in A_j} \bar{f}_i.
\end{aligned} \tag{37}$$

We now intend to find a suitable upper bound for the expectation of the left-hand side of (37). This can be done by considering the right-hand side in two separated nonintersecting index subsets of i, denoted by $A' \stackrel{\Delta}{=} \{i_j, j = 1, 2, \ldots, n\}$ and $A'' \stackrel{\Delta}{=} \bigcup_{j=1}^{n} A_j - A'$. By taking expectation over index set A' for (37), we have

$$E \prod_{i \in A'} \bar{f}_i \le \Big(\int_{\mathcal{X}} f^{\frac{1}{p}}(x, \bar{\theta}_i, \theta) f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0) d\mu$$
$$+ \int_{\mathcal{X}} \sup_{|h| \le \delta} \left| f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0 + h) - f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0) \right|$$
$$\cdot f^{\frac{1}{p}}(x, \bar{\theta}_i, \theta) d\mu \Big)^n. \tag{38}$$

Here and hereafter, p satisfies $\frac{1}{p} + \frac{1}{q} = 1$. By Lemma 1 and (36), the first term of (38) derives

$$\int_{\mathcal{X}} f^{\frac{1}{p}}(x, \bar{\theta}_i, \theta) f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0) d\mu$$
$$\le \int_{\mathcal{X}} \Big[\frac{1}{p} f(x, \bar{\theta}_i, \theta) + \frac{1}{q} f(x, \bar{\theta}_i, \theta + u_0)$$
$$- \frac{1}{q} \left| f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta) - f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0) \right|^q \Big] d\mu$$
$$\le 1 - \frac{1}{q} \kappa_\theta \left(\frac{\gamma}{2} \right). \tag{39}$$

By Hölder's inequality, the second term of (38) derives

$$\int_{\mathcal{X}} \sup_{|h| \le \delta} \left| f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0 + h) - f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0) \right| f^{\frac{1}{p}}(x, \bar{\theta}_i, \theta) d\mu$$
$$\le \omega_{\theta + u_0}(\delta). \tag{40}$$

By the inequality $1 + x \le e^x$ and (38)–(40), we obtain

$$E \prod_{i \in A'} \bar{f}_i \le \exp \left\{ -n \left(\frac{1}{q} \kappa_\theta \left(\frac{\gamma}{2} \right) - \omega_{\theta + u_0}(\delta) \right) \right\}. \tag{41}$$

By taking expectation over index set A'' for (37), we have

$$E \prod_{i \in A''} \bar{f}_i \le \Big(\int_{\mathcal{X}} f^{\frac{1}{p}}(x, \bar{\theta}_i, \theta) f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0) d\mu$$
$$+ \int_{\mathcal{X}} \sup_{|h| \le \delta} \left| f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0 + h) - f^{\frac{1}{q}}(x, \bar{\theta}_i, \theta + u_0) \right|$$
$$\cdot f^{\frac{1}{p}}(x, \bar{\theta}_i, \theta) d\mu \Big)^{n(m-1)}. \tag{42}$$

By Hölder's inequality, the first term of (42) is not greater than 1, and the second term of (42) satisfies inequality (40). Similar to (41), we obtain

$$E \prod_{i \in A''} \bar{f}_i \le \exp \left\{ n(m-1) \omega_{\theta + u_0}(\delta) \right\}. \tag{43}$$

Finally, by (41) and (43), we have

$$E \sup_{u \in \Gamma} \Pi_n^{\frac{1}{q}}(u) \le \exp \left\{ -n \left(\frac{1}{q} \kappa_\theta \left(\frac{\gamma}{2} \right) - m \omega_{\theta + u_0}(\delta) \right) \right\}. \tag{44}$$

By definition of $\omega_\theta(\delta)$ in (32), it is easy to show that, e.g., see [15],

$$\lim_{\delta \to 0} \sup_{\theta \in \Theta} \omega_\theta(\delta) = 0. \tag{45}$$

Noticing the fact m is a constant integer, this means

$$\frac{1}{q}\kappa_\theta\left(\frac{\gamma}{2}\right) - m\omega_{\theta+u_0}(\delta) > 0 \tag{46}$$

can be held if we select sufficiently small δ. Thus, the right-hand side of (44) tends exponentially to 0 as $n \to \infty$.

Fix $\gamma > 0$ and cover the exterior of the sphere $|u| \le \gamma$ by M spheres Γ_j, $j = 1,\ldots,M$, of radius δ with centers u_j. The number δ is chosen to be sufficiently small so that all the spheres will be located in the region $|u| > \frac{1}{2}\gamma$ and so that, for all j, the inequality $m\omega_{\theta+u_j}(\delta) \le \frac{1}{2q}\kappa_\theta(\frac{1}{2}\gamma)$ satisfied. Then, by definition of Π_n and (44),

$$
\begin{aligned}
P\{|\hat{\theta}_n - \theta| > \gamma\} &\le \sum_{i=1}^{M} P\{|\hat{\theta}_n - \theta| \in \Gamma_i\} \\
&\le \sum_{i=1}^{M} \exp\left\{-\frac{n}{m}\left(\frac{1}{q}\kappa_\theta\left(\frac{\gamma}{2}\right) - m\omega_{\theta+u_0}(\delta)\right)\right\} \\
&\le M\exp\left\{-\frac{n}{2qm}\kappa_\theta\left(\frac{\gamma}{2}\right)\right\}.
\end{aligned}
\tag{47}
$$

Therefore,

$$
\begin{aligned}
P\left\{\bigcup_{l=n}^{\infty}[|\hat{\theta}_l - \theta| > \gamma]\right\} &\le M\sum_{l=n}^{\infty}\exp\left\{-\frac{l}{2qm}\kappa_\theta\left(\frac{\gamma}{2}\right)\right\} \\
&= \frac{M\exp\left\{-\frac{n}{2qm}\kappa_\theta\left(\frac{\gamma}{2}\right)\right\}}{1 - \exp\left\{-\frac{1}{2qm}\kappa_\theta\left(\frac{\gamma}{2}\right)\right\}} \xrightarrow[n\to\infty]{} 0,
\end{aligned}
\tag{48}
$$

which finishes the proof. □

Once the two conditions of Theorem 2 are justified, then strong consistence of ML estimation algorithm follows. However, the two conditions are inconvenient to use in applications. Below is an attempt to find a simple and applicable criterion by using the two conditions (i) and (ii) with $q = 2$ in Theorem 2.

Theorem 3. *Let Θ be a bounded closed set in $\mathcal{R}^k$, $\{\bar{\theta}_i\}$ belong to a compact set, $f(x, \bar{\theta}_i, \theta)$ be a continuous function of $\theta \in \Theta$, $x \in \mathcal{X}$, and $\{\bar{\theta}_i\}$, and A_j is the same notation of Theorem 2. Furthermore, if*
(i') For any different $\theta, \theta' \in \Theta$ and $\forall j$, there exist fixed x_0 and $\varepsilon_0 > 0$ and $i_0 \in A_j$ such that

$$\left|\sqrt{f(x_0, \bar{\theta}_{i_0}, \theta)} - \sqrt{f(x_0, \bar{\theta}_{i_0}, \theta')}\right| > \varepsilon_0; \tag{49}$$

(ii') For all $\theta \in \Theta$, the derivative of $\sqrt{f(x,,\bar{\theta}_i, \theta)}$ satisfies

$$\sup_i \int_{\mathcal{X}} \sup_{\theta\in\Theta} \left|\frac{\partial\sqrt{f(x,\bar{\theta}_i,\theta)}}{\partial\theta}\right| d\mu < \infty. \tag{50}$$

Then, for any fixed $\theta \in \Theta$, the ML estimator $\hat{\theta}_n$ given by (30) tends to θ as $n \to \infty$ with probability 1.

Proof. We need only to justify that the two conditions of Theorem 2 are satisfied with $q = 2$. Clearly, the condition (ii') is stronger than (ii) of Theorem 2 in view of standard Differential Mean Value Theorem with respect to θ. Thus, we need only to guarantee that (i') is a special case of (i).

By (i′), continuity of $\sqrt{f}$, and the fact that Θ is compact, for any given $\gamma > 0$, there exists a number $\varepsilon_0(\gamma) > 0$ such that

$$\inf_j \max_{i \in A_j} \inf_{|\theta - \theta'| \geq \gamma} \left| \sqrt{f(x_0, \bar{\theta}_i, \theta)} - \sqrt{f(x_0, \bar{\theta}_i, \theta')} \right| > \varepsilon_0. \tag{51}$$

Again, by the fact that $\{\bar{\theta}_i\}$ belongs to a compact set, there exists a closed ball (neighborhood) of x_0, denoted by $B(x_0, \delta)$ with $\delta > 0$, such that

$$\inf_j \max_{i \in A_j} \inf_{|\theta - \theta'| \geq \gamma} \left| \sqrt{f(x, \bar{\theta}_i, \theta)} - \sqrt{f(x, \bar{\theta}_i, \theta')} \right| > \varepsilon_0 / 2 \tag{52}$$

holds for any $x \in B(x_0, \delta)$. Clearly, this immediately implies (31) with

$$\kappa_\theta(\gamma) = \varepsilon_0 / 2 |B(x_0, \delta)| > 0,$$

which finishes the proof. □

By the notion introduced by Definition 1, we simplify further the condition (i′) of Theorem 3 for application.

Remark 5. *By Definition 1 and continuity of density $f(\cdot, \cdot, \cdot)$, we find a simple sufficient condition to guarantee the condition (i′) of Theorem 3 as:*
(i″) For any different $\theta, \theta' \in \Theta$, there exist fixed x_0 and $\bar{\theta}_{i_0} \in Lim(\{\bar{\theta}_i, i \in A_j\})$ such that

$$f(x_0, \bar{\theta}_{i_0}, \theta) \neq f(x_0, \bar{\theta}_{i_0}, \theta'). \tag{53}$$

7. Conclusions

We have investigated the consistence condition of Maximum Likelihood (ML) estimation for transmit power of a primary node in a wireless network with cooperative receiver nodes. By relevant theoretical analysis, the condition is that the location set of receiver nodes should not be (or asymptotically in some sense) merely in a criterion curve (which is a term given by Definition 2). In other words, this means the locations of receiver nodes should be rich enough to guarantee consistence of ML estimation. This condition can be compared with the persistent excitation condition used to guarantee the convergence of a least squares algorithm [14].

Moreover, we have also established theoretical analysis for the consistence condition of ML function, which is generated by independent yet not identically distributed random variables. A simple criterion is proposed for application.

Some relevant positive and negative numerical experiments were designed and conducted in order to verify the theoretical discoveries. All of these experiments have been analyzed and explained by relevant theoretical results established.

As our future study, more general scenarios in transmit power estimation are of great interest—for example, the case where there are more unknown parameters in the lognormal shadowing model, as well as the case in which more primary nodes are to be estimated synchronously in the wireless network as that stated in [8]. It is believed that the strict theoretical results developed in the paper will still be useful and will be able to serve as a basis when the more general scenarios of ML estimation are handled.

Author Contributions: X.-L.H. contributed to most of the formulations and simulations, and prepared the original proposed framework under the supervision of P.-H.H. P.-H.H. and L.P. gave advices on the overall work of the paper and helped revise the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Min, A.W.; Kim, K.-H.; Singh, J.P.; Shin, K.G. Opportunistic spectrum access for mobile cognitive radios. In Proceedings of the IEEE INFOCOM 2011, Shanghai, China, 10–15 April 2011.
2. Xue, D.; Ekici, E.; Vuran, M.C. Cooperative Spectrum Sensing in Cognitive Radio Networks Using Multidimensional Correlations. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 1832–1843. [CrossRef]
3. Ho, I.W.-H.; Ko, B.J.; Zafer, M.; Bisdikian, C.; Leung, K.K. Cooperative Transmit-Power Estimation in MANETs. In Proceedings of the 2008 Wireless Communications and Networking Conference (WCNC), Las Vegas, NV, USA, 31 March–3 April 2008.
4. Xu, W.; Trappe, W.; Zhang, Y.; Wood, T. The feasibility of launching and detecting jamming attacks in wireless networks. In Proceedings of the 2005 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing, Urbana-Champaign, IL, USA, 25–28 May 2005.
5. Ho, I.W.-H.; Liew, S.C. Impact of power control on performance of IEEE 802.11 wireless networks. *IEEE Trans. Mob. Comput.* **2007**, *6*, 1245–1258. [CrossRef]
6. Lin, L.; So, H.C.; Chan, Y.T. Received Signal Strength based Positioning for Multiple Nodes in Wireless Sensor Networks. *Digit. Signal Process.* **2014**, *25*, 41–50. [CrossRef]
7. Salman, N.; Ghogho, M.; Kemp, A.H. Optimized Low Complexity Sensor Node Positioning in Wireless Sensor Networks. *IEEE Sens. J.* **2014**, *14*, 39–46. [CrossRef]
8. Nelson, J.K.; Gupta, M.R.; Almodovar, J.; Mortensen, W.H. A Quasi EM Method for Estimating Multiple Transmitter Locations, IEEE Signal Processing Letters. *IEEE Signal Process. Lett.* **2009**, *16*, 354–357. [CrossRef]
9. Kim, S.; Jeon, H.; Lee, H.; Ma, J.S. Robust transmission power and position estimation in cognitive radio. In Proceedings of the 2007 International Conference on Information Networking (ICOIN), Estoril, Portugal, 23–25 January 2007.
10. Zafer, M.; Ko, B.J.; Ho, I. Cooperative Transmit-Power Estimation under Wireless Fading. In Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (ACM Mobihoc), Hong Kong, China, 27–30 May 2008.
11. Jiao, J.; Venkat, K.; Han, Y.; Weissman, T. Maximum likelihood estimation of functionals of discrete distributions. *IEEE Trans. Inf. Theory* **2017**, *63*, 6774–6798. [CrossRef]
12. Imani, M.; Braga-Neto, U.M. Maximum-likelihood adaptive filter for partially observed boolean dynamical systems. *IEEE Trans. Signal Process.* **2017**, *65*, 359–371. [CrossRef]
13. Gudmundson, M. Correlation model for shadow fading in mobile radio systems. *Electron. Lett.* **1991**, *47*, 2145–2146. [CrossRef]
14. Schon, T.B.; Wills, A.; Ninness, B. System Identification of Nonlinear State-space Models. *Automatica* **2011**, *47*, 39–49. [CrossRef]
15. Jiao, J. *Asymptotic Theory of Statistical Estimation*; Standford University: Stanford, CA, USA, 2014.
16. Kuang, J. *Applied Inequalities*, 3rd ed.; Shangdong Technology Press: Jinan, China, 2004. (In Chinese)
17. Hu, X.-L. An extension of Young's inequality and its application. *Appl. Math. Comput.* **2013**, *219*, 6393–6399. [CrossRef]