Article

A Fast ML-Based Single-Step Localization Method Using EM Algorithm Based on Time Delay and Doppler Shift for a Far-Field Scenario

Tianzhu Qin, Lin Li *, Bin Ba and Daming Wang

National Digital Switching System Engineering and Technology Research Center, Zhengzhou 450002, China; skypillar@outlook.com (T.Q.); xidianbabin@163.com (B.B.); wdm_wangdaming@163.com (D.W.)

* Correspondence: llbnerfg@163.com; Tel.: +86-188-3896-7979

check for
updates

Abstract: This study discusses the localization problem based on time delay and Doppler shift for a far-field scenario. The conventional location methods employ two steps that first extract intermediate parameters from the received signals and then determine the source position from the measured parameters. As opposed to the traditional two-step methods, the direct position determination (DPD) methods accomplish the localization in a single step without computing intermediate parameters. However, the DPD cost function often remains non-convex, thereby it will cost a high amount of computational resources to find the estimated position through traversal search. Weiss proposed a DPD estimator to mitigate the computational complexity via eigenvalue decomposition. Unfortunately, when the computational resources are rather limited, Weiss's method fails to satisfy the timeliness. To solve this problem, this paper develops a DPD estimator using expectation maximization (EM) algorithm based on time delay and Doppler shift. The proposed method starts from choosing the transmitter-receiver range vector as the hidden variable. Then, the cost function is separated and simplified via the hidden variable, accomplishing the transformation from the high dimensional nonlinear search problem into a few one dimensional search subproblems. Finally, the expressions of EM repetition are obtained through Laplace approximation. In addition, we derive the Cramér–Rao bound to evaluate the best localization performance in this paper. Simulation results confirm that, on the basis of guaranteeing high accuracy, the proposed algorithm makes a good compromise in localization performance and computational complexity.

Keywords: direct position determination (DPD); maximum likelihood (ML); expectation maximization (EM); Doppler shift; Laplace approximation; Cramér–Rao bound (CRB)

1. Introduction

The transmitter localization is a classic problem in wireless communication systems. It is well known that the conventional location approaches are composed of two separate steps: (1) The intermediate parameters are estimated through the signals. (2) The source position is determined from the measured parameters. In the past three decades, a great deal of work has been done in this field. Generally, the intermediate parameters (e.g., angle of arrival (AOA) [1], time of arrival (TOA) [2], time difference of arrival (TDOA) [3], Doppler shift [4–6] and frequency difference of arrival (FDOA) [7]) are usually estimated by the maximum likelihood (ML)-based method [8] or the subspace data fusion (SDF)-based method [5,9], and the position is mainly determined by an iterative algorithm [10] or a closed-form algorithm [11]. It is worth mentioning that Doppler information is a key parameter in the location problem based on the moving receiver. Gajewski [5] uses SDF criterion based on the Doppler effect

to locate multiple emission sources. Since multipath propagation and Doppler effect are dominant factors to deteriorate localization performance, Kelner [6] uses the signal Doppler frequency method to resist non-line-of-sight (NLOS) conditions. As for the conventional approaches, it must be emphasized that the intermediate parameters are extracted by ignoring the constraint that all measurements must correspond to the same geolocation of the emitter, and more errors will be introduced in two separate processes of the conventional methods [12]. As a result, the conventional two-step methods cannot yield satisfactory location accuracy. For this reason, direct position determination (DPD) techniques that exploit the intrinsic constraint and determine the source position from the received signals directly were developed.

The DPD algorithms have been intensively investigated in recent years. An important classification of localization methods is related to propagation conditions, and there are two general cases: (1) multipath propagations with LOS and NLOS; (2) an additive white Gaussian noise (AWGN). In the first case, the methods [13–16] for LOS or NLOS environments have been deeply developed. Note that Du [16] proposed a DPD estimator for a novel localization architecture in multipath propagations, called the "Multiple Transponders and Multiple Receivers for Multiple Emitters Positioning System". The second case is the most common condition, and the following discussion is also conducted under this assumption. Weiss [17] first proposed an SDF-based DPD algorithm with the utilization of orthogonality between signal subspace and noise subspace to estimate the emitter position. Chen [18] developed a multi-target DPD approach using subspace based on compressive sensing, reaching a high probability of locating the emitter without knowing the number of targets. Even if low computational complexity appears in the SDF-based DPD method, localization performance is deteriorated in low signal-to-noise ratios (SNRs), failing to reach the corresponding Cramér–Rao bound (CRB). To enhance the performance, ML-based DPD approaches were developed. Since the DPD cost function is often non-convex, traversal search is required to find the extremum. However, nuisance parameters (e.g., unknown transmission time [19] and timing errors [20]) will result in high dimensional search, which is impractical in a real-time application. To mitigate the computational load of traversal search, Weiss [21] proposed an ML-based approach via finding the maximum eigenvalue of the Hermitian matrix associated with the cost function, which also exhibits high localization accuracy. It must be emphasized that Weiss's method is an ideal solution to account for localization accuracy and computation complexity in existing DPD algorithms. Moreover, the location performance can be further enhanced by utilizing signal properties. These DPD algorithms [22–24] can obtain superior localization precision by exploiting the constant modulus, orthogonal frequency division multiplexing, and the cyclostationary properties of signals. Unfortunately, the above DPD methods are not fast enough in the presence of limited computing resources, which is often a reality in moving receivers (i.e., airplanes or unmanned aerial vehicles (UAVs)). This limitation results in non-timeliness performance for moving emitter localization. On the other hand, DPD methods based on the ML criterion guarantee superior localization accuracy. Therefore, an ML-based DPD method with rapidity needs to be studied.

The expectation maximization (EM) algorithm is an attractive method of estimating the ML result when data can be divided into "incomplete data" and "complete data" in the model. In the past three decades, the EM algorithm has provided an excellent way to solve machine learning problems (i.e., speech processing and recognition [25] and image identification and restoration [26]). Via choosing the appropriate hidden variable, the EM algorithm decouples a high-dimensional search problem into a few subproblems with much lower computational complexity. In sight of the advantages of the EM algorithm, many scholars have applied it to the parameter estimation domain. Mada [27] uses the EM algorithm to solve the multi-source localization problem, leading to a soft computational load. Moreover, Lu estimates the source position via the EM algorithm for spatially nonwhite noise [28] and nonuniform noise [29], respectively, further demonstrating the effectiveness of the EM model in harsh scenarios. However, the DPD methods have not received the same treatment on the EM application, and only Tzoreff [30] discussed a DPD method based on the EM algorithm. Unfortunately, this method

cannot suit for the moving-receiver scenario. As a result, there is a great demand for developing a DPD method using the EM algorithm for moving receivers.

This paper proposes a fast ML-based DPD algorithm using the EM algortihm based on time delay and Doppler shift, and the main processing steps are exhibited as follows:

(1) The transmitter-receiver range vector is selected as the hidden variable, successfully leading the separation and simplification of the ML cost function.
(2) With the help of Laplace approximation, the high-dimensional multi-parameter search of the prescribed ML estimator is decoupled into a closed form of the transmitter position and a line search of transmitter-receiver distance as well as transmitted time. Therefore, the expressions of EM repetition is determined.
(3) Iteration of the EM expressions, which alternately updates parameters in E-step and M-step, is required until the norm of the difference between the adjacent estimated position converges to a user's predefined number.

In summary, the main contribution of this paper is the improvement of the prescribed ML DPD estimator via the EM algorithm in moving-receiver application for a far-field scenario, leading to a high effectiveness to find the global extreme. In addition, we also have derived the CRB to exhibit the best localization performance in theory. The rest of this paper is organized as follows. Section 2 lists the notations used in this paper. Section 3 constructs the signal model and formulates the problem. Section 4 discusses the DPD methods, including the previous method and the proposed method, and then makes a computational complexity comparison among different methods. Section 5 presents several numerical simulations to verify the theoretical analysis. Finally, Section 6 draws the conclusions.

2. Notations

In this section, some mathematical notation explanations that will be used through this paper are listed in Table 1.

Table 1. Mathematical notation explanation.

Notation	Explanation
$[\bullet]^T$	transpose
$[\bullet]^H$	conjugate transpose
$\bullet^{(R)}$	the real part
$\bullet^{(I)}$	the imaginary part
$\text{diag}\{\bullet\}$	a diagonal matrix with diagonal entries
$\otimes$	Kronecker product
$\|\bullet\|$	Euclidean norm of the matrix
$|\bullet|$	determinant of the matrix
$p_{\mathbf{X},\mathbf{Y}}\{\bullet\}$	the joint distribution of $\mathbf{X}$ and $\mathbf{Y}$
$p_{\mathbf{X}/\mathbf{Y}}\{\bullet\}$	the conditional distribution of $\mathbf{X}$ given $\mathbf{Y}$
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbf{0}_N$	$N \times N$ matrix with zero

3. Signal Model

This paper considers that L moving receivers intercepts the transmitted signal, and the signal is partitioned into K short intervals. Note that the antennas of the emitter and receivers are omnidirectional, and the positions of the emitter and receivers are determined in a two-dimensional

coordinate system. In order to fully introduce the DPD model, three assumptions are included as follows:

Assumption 1. *The receivers are moving. Let $\mathbf{p}_{l,k}$ and $\mathbf{v}_{l,k}$ denote the position and velocity of the lth receivers at the kth interception interval, which are precisely known to us. They are constant at each interception interval.*

Assumption 2. *The stationary emitter, denoted by $\mathbf{p} \triangleq (x, y)$, locates in the far-field of the moving receivers.*

Assumption 3. *The propagation channel is an AWGN channel, and time delay as well as Doppler information are used in the signal model.*

Based on the above assumptions, the complex signals $y_{l,k}(t)$ observed by the lth receiver at the kth interception interval at time t is expressed as [31]

$$y_{l,k}(t) = b_{l,k} s_k \left(t - \tau_{l,k}(\mathbf{p}) - t_0 \right) e^{j2\pi f_{l,k} t} + n_{l,k}(t) \ \ 1 \leqslant l \leqslant L; 1 \leqslant k \leqslant K \tag{1}$$

where t_0 is the signal transmission time, and during the kth interception interval, $s_k(t)$ is the complex envelope of the emitter, $b_{l,k}$ denotes the channel attenuation, $n_{l,k}(t)$ is the white Gaussian, $\tau_{l,k}(\mathbf{p})$ is the propagation time between the emitter and the lth receiver, and Doppler frequency $f_{l,k}$ observed between the emitter and the lth receiver is given by [21]

$$f_{l,k} \triangleq f_c + f_c \mu_{l,k}(\mathbf{p}) \tag{2}$$

where f_c is the nominal frequency of the transmitted signal, assumed known, with

$$\mu_{l,k}(\mathbf{p}) \triangleq \frac{1}{c} \frac{\mathbf{v}_{l,k}^{\mathrm{T}} \left[\mathbf{p} - \mathbf{p}_{l,k} \right]}{\|\mathbf{p} - \mathbf{p}_{l,k}\|} \tag{3}$$

where c is the radio wave propagation speed. It is assume that each receiver performs a down conversion of the intercepted signal by the nominal frequency. Thus, after down conversion the Doppler frequency can be replaced by

$$f_{l,k} \cong f_c \mu_{l,k}(\mathbf{p}). \tag{4}$$

After sampling at $t = nT_s$, the received signal can be shown as

$$\tilde{y}_{l,k}(n) = b_{l,k} \tilde{s}_k \left(nT_s - \tau_{l,k}(\mathbf{p}) - t_k \right) e^{j2\pi f_{l,k} nT_s} + \tilde{n}_{l,k}(n) \ \ n = 1, 2, \ldots, N \tag{5}$$

for $l = 1, \cdots, L$, and $k = 1, \cdots, K$, where N denotes the number of sample points at each interval. Then, Equation (5) can be expressed by a vector form as

$$\tilde{\mathbf{y}}_{l,k} = b_{l,k} \mathbf{A}_{l,k} \tilde{\mathbf{s}}_{\tau_{l,k}} + \tilde{\mathbf{n}}_{l,k} \tag{6}$$

where

$$\begin{aligned}
\tilde{\mathbf{y}}_{l,k} &\triangleq \left[\tilde{y}_{l,k}(1), \tilde{y}_{l,k}(2) \cdots, \tilde{y}_{l,k}(N) \right]^{\mathrm{T}} \\
\tilde{\mathbf{n}}_{l,k} &\triangleq \left[\tilde{n}_{l,k}(1), \tilde{n}_{l,k}(2) \cdots, \tilde{n}_{l,k}(N) \right]^{\mathrm{T}} \\
\mathbf{A}_{l,k} &= \mathrm{diag} \left\{ \exp(-j2\pi f_{l,k} \tilde{\mathbf{N}} T_s) \right\} \\
\tilde{\mathbf{s}}_{\tau_{l,k}} &= \tilde{s}_k \left(\tilde{\mathbf{N}} T_s - \tau_{l,k}(\mathbf{p}) - t_0 \right)^{\mathrm{T}},
\end{aligned} \tag{7}$$

with $\tilde{\mathbf{N}} = [1, 2, \ldots, N]^{\mathrm{T}}$. We define $\tilde{\mathbf{s}}_k = [\tilde{s}_k(T_s), \tilde{s}_k(2T_s), \cdots, \tilde{s}_k(NT_s)]^{\mathrm{T}}$, and the Fourier coefficients of $\tilde{\mathbf{s}}_{\tau_{l,k}}$ and $\tilde{\mathbf{s}}_k$ satisfy

$$\tilde{\mathbf{s}}_{\tau_{l,k}} = \mathbf{F}^{\mathrm{H}} \mathbf{D}_{\tau_{l,k}} \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{D}_{t_0} \mathbf{F} \tilde{\mathbf{s}}_k \tag{8}$$

where

$$
\begin{aligned}
\mathbf{F} &= \frac{1}{\sqrt{N}} \exp\left(-j\frac{2\pi}{N}\tilde{\mathbf{N}}\tilde{\mathbf{N}}^{\mathrm{T}}\right) \\
\mathbf{D}_{\tau_{l,k}} &= \mathrm{diag}\left\{\exp\left(-j\frac{2\pi\tau_{l,k}}{NT_s}\tilde{\mathbf{N}}\right)\right\} \\
\mathbf{D}_{t_0} &= \mathrm{diag}\left\{\exp\left(-j\frac{2\pi t_0}{NT_s}\tilde{\mathbf{N}}\right)\right\}.
\end{aligned}
\tag{9}
$$

$\mathbf{F}$ is the discrete Fourier transform operator. Therefore, Equation (6) is rewritten as

$$
\tilde{\mathbf{y}}_{l,k} = b_{l,k}\mathbf{C}_{l,k}\tilde{\mathbf{s}}_k + \tilde{\mathbf{n}}_{l,k}
\tag{10}
$$

where $\mathbf{C}_{l,k} = \mathbf{A}_{l,k}\mathbf{F}^{\mathrm{H}}\mathbf{D}_{\tau_{l,k}}\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{D}_{t_0}\mathbf{F}$.

4. Direct Position Determination Methods

This section discuss the DPD methods, which locate the emitter directly through the raw data. We first discuss the ML-based DPD method requiring traversal search, which can receive the best location accuracy. Weiss [19] proposed a fast ML-based DPD method based on eigenvalue decomposition. Unfortunately, the above methods cannot work effectively in the presence of limited computational sources. Thus, in light of the EM idea in estimation theory [27], a fast DPD method using EM algorithm will be proposed.

4.1. Previous Work

The DPD optimization based on time delay and Doppler shift is established, which was first introduced by [21]. The received signal $\tilde{\mathbf{y}}_{l,k}$ is a complex Gaussian random vector. Hence, the likelihood function for $\tilde{\mathbf{y}}$ can be formulated by [21]

$$
l(\boldsymbol{\theta}) = \frac{1}{(\pi\sigma^2)^{LKN}} \exp\left(-\frac{1}{\sigma^2}\sum_{k=1}^{K}\sum_{l=1}^{L}\left\|\tilde{\mathbf{y}}_{l,k} - b_{l,k}\mathbf{C}_{l,k}\tilde{\mathbf{s}}_k\right\|^2\right)
\tag{11}
$$

where σ^2 denotes the noise power, and $\boldsymbol{\theta} = \left[t_0, \mathbf{b}^{\mathrm{T}}, \mathbf{p}^{\mathrm{T}}\right]^{\mathrm{T}}$ denotes all unknown parameters, with $\mathbf{b} = \left[\mathbf{b}_1^{\mathrm{T}}, \mathbf{b}_2^{\mathrm{T}}, \cdots, \mathbf{b}_K^{\mathrm{T}}\right]^{\mathrm{T}}$ and $\mathbf{b}_k = \left[\mathbf{b}_{1,k}^{\mathrm{T}}, \mathbf{b}_{2,k}^{\mathrm{T}}, \cdots, \mathbf{b}_{L,k}^{\mathrm{T}}\right]^{\mathrm{T}}$. The associated logarithmic likelihood function can be written as

$$
L(\boldsymbol{\theta}) = -LKN\ln\pi\sigma^2 - \frac{1}{\sigma^2}\sum_{k=1}^{K}\sum_{l=1}^{L}\left\|\tilde{\mathbf{y}}_{l,k} - b_{l,k}\mathbf{C}_{l,k}\tilde{\mathbf{s}}_k\right\|^2.
\tag{12}
$$

Therefore, the estimation of noise power σ^2 is

$$
\hat{\sigma^2} = \frac{1}{LKN}\sum_{k=1}^{K}\sum_{l=1}^{L}\left\|\tilde{\mathbf{y}}_{l,k} - b_{l,k}\mathbf{C}_{l,k}\tilde{\mathbf{s}}_k\right\|^2.
\tag{13}
$$

By substituting Equation (11) into Equation (10), the estimation of parameter $\boldsymbol{\theta}$ can be determined by

$$
\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} f(\boldsymbol{\theta}),
\tag{14}
$$

with

$$
f(\boldsymbol{\theta}) = \sum_{k=1}^{K}\sum_{l=1}^{L}\left\|\tilde{\mathbf{y}}_{l,k} - b_{l,k}\mathbf{C}_{l,k}\tilde{\mathbf{s}}_k\right\|^2.
\tag{15}
$$

Since a high-dimensional nonlinear problem appears in Equation (14), it is difficult to compute the closed-form solution of $\hat{\theta}$. Thus, traversal searches are required among these stray parameters to obtain the best performance. However, this technique will take a long time to find the extreme corresponding to the emitter position, which is not efficient in practical application.

4.2. The Proposed Method

4.2.1. EM Algorithm Review

The EM algorithm is an approach to iterative computation of the ML problem when the observations are regarded as incomplete data. In each iteration, it includes an expectation step and a maximization step. The meaning "incomplete data" reveals that there are two kinds of data: the incomplete data **Y** and the complete data **X**. More specifically, **Y** is the observed data, and **X** is the corresponding hidden data, (not observed directly). We denote the estimated parameter θ and an expression $\mathbf{Y} = f(\mathbf{X})$, which shows a many–one mapping from **X** to **Y**. Via the Bayesian rule, we have [30]

$$L(\theta) \triangleq \log p_{\mathbf{Y}}(\mathbf{Y};\theta) = \log p_{\mathbf{X},\mathbf{Y}}(\mathbf{X},\mathbf{Y};\theta) - \log p_{\mathbf{X}/\mathbf{Y}}(\mathbf{X}/\mathbf{Y};\theta) \tag{16}$$

where $L(\theta)$ is the logarithmic likelihood function.

Since $p_{\mathbf{Y}}(\mathbf{Y};\theta)$ is independent to **X**, the conditional expectation operation with associated with $p_{\mathbf{X}/\mathbf{Y}}(\mathbf{X}/\mathbf{Y};\theta')$ will not make change to $L(\theta)$. The conditional expectation of Equation (14) can be expressed by

$$L(\theta) = E\{\log p_{\mathbf{X},\mathbf{Y}}(\mathbf{X},\mathbf{Y};\theta)\} - E\left\{\log p_{\mathbf{X}/\mathbf{Y}}(\mathbf{X}/\mathbf{Y};\theta)\right\} \tag{17}$$

where $E \equiv E_{(\mathbf{X}/\mathbf{Y};\theta')}\{\bullet\}$ and θ' denotes an arbitrary value of θ.

We define

$$\begin{aligned}
Q\left(\theta,\theta'\right) &= E\{\log p_{\mathbf{X},\mathbf{Y}}(\mathbf{X},\mathbf{Y};\theta)\} \\
W\left(\theta,\theta'\right) &= -E\left\{\log p_{\mathbf{X}/\mathbf{Y}}(\mathbf{X}/\mathbf{Y};\theta)\right\}.
\end{aligned} \tag{18}$$

Thus, Equation (17) can be rewritten as

$$L(\theta) = Q\left(\theta,\theta'\right) + W\left(\theta,\theta'\right). \tag{19}$$

Based on the ML criterion, we can maximize $L(\theta)$ to estimate the unknown parameters. Since $p_{\mathbf{X}/\mathbf{Y}}(\mathbf{X}/\mathbf{Y};\theta)$ is generally unknown to us, we need to approximate $W\left(\theta,\theta'\right)$.

Both sides of Equation (19) subtract $L\left(\theta'\right)$ yielding

$$L(\theta) - L\left(\theta'\right) = Q\left(\theta,\theta'\right) - Q\left(\theta',\theta'\right) + W\left(\theta,\theta'\right) - W\left(\theta',\theta'\right). \tag{20}$$

With the help of Gibbs inequality, we find $W\left(\theta,\theta'\right) \geqslant W\left(\theta',\theta'\right)$. Thus, we obtain

$$L(\theta) - L\left(\theta'\right) \geqslant Q\left(\theta,\theta'\right) - Q\left(\theta',\theta'\right). \tag{21}$$

The above expression shows that an increment of Q associated with θ also ensures an increment of L. Therefore, the maximization problem of L is transformed into the maximization problem of Q. The EM algorithm contains two steps, and each iteration process can be given by

$$\begin{aligned}
&E - step : Calculate\ Q\left(\theta,\theta^{(i)}\right), \\
&M - step : \theta^{(i+1)} = \arg\max_{\theta} Q\left(\theta,\theta^{(i)}\right).
\end{aligned} \tag{22}$$

An iteration cycle exists in Equation (20). We obtain $\theta^{(i+1)}$ in a current iteration step, and it will be the initial value to repeat the EM operation of Equation (20) in the next iteration step. When Q converges, the iteration stops.

4.2.2. Derivation of the EM-DPD Algorithm

After the above analysis, we choose the received signals $\tilde{\mathbf{Y}} = \left[\tilde{\mathbf{y}}_{1,1}^{\mathrm{T}}, \cdots, \tilde{\mathbf{y}}_{l,k}^{\mathrm{T}}, \cdots, \tilde{\mathbf{y}}_{L,K}^{\mathrm{T}} \right]^{\mathrm{T}}$ as the observed data and the transmitter-receiver range vectors $\tilde{\mathbf{X}} = \left[\tilde{\mathbf{x}}_{1,1}^{\mathrm{T}}, \cdots, \tilde{\mathbf{x}}_{l,k}^{\mathrm{T}}, \cdots, \tilde{\mathbf{x}}_{L,K}^{\mathrm{T}} \right]^{\mathrm{T}}$ as the hidden data, respectively. We assume that $\tilde{\mathbf{x}}_{l,k}$ is a Gaussian vector, and the probability distribution function (PDF) of $\tilde{\mathbf{X}}$ can be shown as

$$p_{\tilde{\mathbf{X}}} \left(\tilde{\mathbf{X}}; \mathbf{p} \right) = \prod_{k=1}^{K} \prod_{l=1}^{L} p_{\tilde{\mathbf{x}}} \left(\tilde{\mathbf{x}}_{l,k}; \mathbf{p} \right) = \prod_{k=1}^{K} \prod_{l=1}^{L} N \left(\varepsilon_{l,k} \left(\mathbf{p} \right), \sigma_x^2 \mathbf{I}_2 \right) \tag{23}$$

where $\varepsilon_{l,k} \left(\mathbf{p} \right) = \mathbf{p} - \mathbf{p}_{l,k}$ and σ_x^2 is the variance of x. We define $\mathbf{G}_{l,k} \triangleq \mathbf{C}_{l,k} \tilde{\mathbf{s}}_k$, and Equation (6) is rewritten as

$$\tilde{\mathbf{y}}_{l,k} = b_{l,k} \mathbf{G}_{l,k} + \tilde{\mathbf{n}}_{l,k}. \tag{24}$$

We assume that $\tilde{\mathbf{n}}_{l,k}$ are complex Gaussian vectors with mean $\mathbf{0}_N$ and covariance $\sigma^2 \mathbf{I}_N$. Therefore, the pdf of $\tilde{\mathbf{y}}_{l,k}$ is given by

$$p_{\tilde{\mathbf{y}}} \left(\tilde{\mathbf{y}}_{l,k}; \mathbf{p}, t_0 \right) = \frac{1}{\left(\pi \sigma^2 \right)^N} \exp \left\{ -\frac{\left\| \tilde{\mathbf{y}}_{l,k} - b_{l,k} \mathbf{G}_{l,k} \right\|^2}{\sigma^2} \right\}. \tag{25}$$

Assuming the signals observed by different receivers and different observations intervals are independent, we have

$$p_{\tilde{\mathbf{Y}}} \left(\tilde{\mathbf{Y}}; \mathbf{p}, t_0 \right) = \prod_{k=1}^{K} \prod_{l=1}^{L} p_{\tilde{\mathbf{y}}} \left(\tilde{\mathbf{y}}_{l,k}; \mathbf{p}, t_0 \right) \propto \prod_{k=1}^{K} \prod_{l=1}^{L} \exp \left\{ -\sigma^2 \left\| \tilde{\mathbf{y}}_{l,k} - b_{l,k} \mathbf{G}_{l,k} \right\|^2 \right\}. \tag{26}$$

Note that the source coordinates $\mathbf{p}$ is the direct relation parameter to the observations $\tilde{\mathbf{Y}}$. We need to separate this variable to the subsequent derivation. By introducing $\tilde{\mathbf{X}}$ in Equation (23) and defining $\boldsymbol{\phi} = \left[t_0, \mathbf{b}^{\mathrm{T}} \right]^{\mathrm{T}}$, Equation (26) can be rewritten as

$$p_{\tilde{\mathbf{Y}}} \left(\tilde{\mathbf{Y}}/\tilde{\mathbf{X}}; \boldsymbol{\phi} \right) = \prod_{k=1}^{K} \prod_{l=1}^{L} p_{\tilde{\mathbf{y}}/\tilde{\mathbf{x}}} \left(\tilde{\mathbf{y}}_{l,k}/\tilde{\mathbf{x}}_{l,k}, b_{l,k}, t_0 \right) \propto \prod_{k=1}^{K} \prod_{l=1}^{L} \exp \left\{ -\sigma^2 \left\| \tilde{\mathbf{y}}_{l,k} - b_{l,k} \mathbf{G}_{l,k} \right\|^2 \right\} \tag{27}$$

where $\mathbf{G}_{l,k} \equiv \mathbf{G}_{l,k} \left(\tilde{\mathbf{x}}_{l,k}, t_0 \right)$ for brevity.

From the analysis in Section 4.2.1, we can estimate the emitter position via maximizing the auxiliary function $Q \left(\boldsymbol{\theta}, \boldsymbol{\theta}' \right)$. Next we introduce the key procedure of the derivation.

Proposition 1. *The auxiliary function* $Q \left(\boldsymbol{\theta}, \boldsymbol{\theta}' \right)$ *is separated into*

$$Q \left(\boldsymbol{\theta}, \boldsymbol{\theta}' \right) = Q_1 \left(\boldsymbol{\phi}, \boldsymbol{\phi}' \right) + Q_2 \left(\mathbf{p}, \mathbf{p}' \right). \tag{28}$$

Proof of Proposition. . Via the Bayesian rule, the joint probability $p_{\tilde{\mathbf{X}},\tilde{\mathbf{Y}}}\left(\tilde{\mathbf{X}},\tilde{\mathbf{Y}};\theta\right)$ is expressed by the product of Equations (23) and (26), which depend only on $\mathbf{p}$ and only on ϕ, respectively. Since the $\log(\bullet)$ of $p_{\tilde{\mathbf{X}},\tilde{\mathbf{Y}}}\left(\tilde{\mathbf{X}},\tilde{\mathbf{Y}};\theta\right)$ exists in Equation (18), the separation of $Q\left(\theta,\theta'\right)$ can be shown as

$$Q_1\left(\phi,\phi'\right) = -\sigma^{-2}\sum_{k=1}^{K}\sum_{l=1}^{L}E\left\{\left\|\tilde{\mathbf{y}}_{l,k}-b_{l,k}\mathbf{G}_{l,k}\right\|^2\right\} \tag{29}$$

$$Q_2\left(\mathbf{p},\mathbf{p}'\right) = -\sigma_x^{-2}\sum_{k=1}^{K}\sum_{l=1}^{L}E\left\{\left\|\tilde{\mathbf{x}}_{l,k}-\varepsilon_{l,k}\left(\mathbf{p}\right)\right\|^2\right\}. \tag{30}$$

By maximizing Equation (29), we obtain

$$\hat{b}_{l,k} = \frac{\bar{\mathbf{G}}_{l,k}^{\mathrm{H}}\tilde{\mathbf{y}}_{l,k}}{E_s}\ \forall\ l=1,\cdots,L;\ k=1,\cdots,K \tag{31}$$

where $\bar{\mathbf{G}}_{l,k}\equiv E\{\mathbf{G}_{l,k}\}$ and $E_s\equiv\left\|\mathbf{G}_{l,k}\right\|^2$. Substituting Equation (31) into Equation (29) yields

$$Q_1\left(t_0,t_0'\right) = -\sigma^{-2}\sum_{k=1}^{K}\sum_{l=1}^{L}\tilde{\mathbf{y}}_{l,k}^{\mathrm{H}}\mathbf{\Psi}_{l,k}\tilde{\mathbf{y}}_{l,k} \tag{32}$$

where $\mathbf{\Psi}_{l,k}=\mathbf{I}_N-E_s^{-1}\bar{\mathbf{G}}_{l,k}\bar{\mathbf{G}}_{l,k}^{\mathrm{H}}$. We continue to maximize the above expression, yielding

$$\hat{t}_0 = \arg\max_{t_0}\ \sum_{k=1}^{K}\sum_{l=1}^{L}\left|\tilde{\mathbf{y}}_{l,k}^{\mathrm{H}}\bar{\mathbf{G}}_{l,k}\left(t_0\right)\right|^2. \tag{33}$$

Similarly, maximizing Equation (28) with respect to $\mathbf{p}$ yields

$$\begin{aligned}\hat{\mathbf{p}} &= \arg\min_{\mathbf{p}}\ \sum_{k=1}^{K}\sum_{l=1}^{L}\left\|\mathbf{p}_{l,k}-\mathbf{p}\right\|^2-2\mathbf{p}^{\mathrm{T}}\bar{\mathbf{x}}_{l,k}\\ &= \frac{1}{KL}\sum_{k=1}^{K}\sum_{l=1}^{L}\left(\mathbf{p}_{l,k}-\bar{\mathbf{x}}_{l,k}\right).\end{aligned} \tag{34}$$

After eliminating $\mathbf{b}$ in the optimization process, we obtain the ML estimated expression of t_0 and the closed-form solution of $\mathbf{p}$ through the EM procedure. Note that $\bar{\mathbf{G}}_{l,k}$ in Equation (31) and $\bar{\mathbf{x}}_{l,k}$ in Equation (32) are need to estimate. This work is promoted by Laplace approximation, which is shown in Appendix A. We define $r_{l,k}\triangleq\left\|\mathbf{x}_{l,k}\right\|$, $R_{l,k}\triangleq\left\|\varepsilon_{l,k}\right\|$, and $\lambda=\frac{\sigma_x^2}{\sigma^2}$, the expression of the EM-DPD algorithm is written as follows:

　　E-step:

$$r_{l,k}^* = \arg\max_{r_{l,k}}\ -\frac{r_{l,k}^2}{2}+R_{l,k}r_{l,k}-\lambda\left\|\tilde{\mathbf{y}}_{l,k}-\hat{b}_{l,k}\mathbf{G}_{l,k}\left(r_{l,k}\right)\right\|^2 \tag{35}$$

$$\forall l=1,\cdots,L; k=1,\cdots,K.$$

　　M-step:

$$\mathbf{p}^{(i+1)} = \frac{1}{KL}\sum_{k=1}^{K}\sum_{l=1}^{L}\left[\mathbf{p}_{l,k}\left(1-\frac{r_{l,k}^*}{R_{l,k}}\right)+\frac{r_{l,k}^*}{R_{l,k}}\mathbf{p}^{(i)}\right]. \tag{36}$$

$$t_0^{(i+1)} = \arg\max_{t_0}\ \sum_{k=1}^{K}\sum_{l=1}^{L}\left|\tilde{\mathbf{y}}_{l,k}^{\mathrm{H}}\bar{\mathbf{G}}_{l,k}\left(r_{l,k}^*,t_0^{(i)}\right)\right|^2. \tag{37}$$

□

For easy understanding, the main steps of the proposed method are exhibited in Algorithm 1.

Algorithm 1: The main steps of the proposed method.

Input:
The observed data: $\tilde{\mathbf{y}}_{l,k}$, the position, and the velocity of receiver: $\mathbf{p}_{l,k}$ and $\mathbf{v}_{l,k}$,
　$(l = 1, \cdots, L, k = 1, \cdots, K)$;
　1. Choose a small positive number $\varepsilon > 0$, and set the iteration counter to $i = 1$;
　2. Set i=0, initialize $\mathbf{p}^{(i)}, t_0^{(i)}$;
　3. Calculate $r_{l,k}^*(l = 1, \cdots, L, k = 1, \cdots, K)$ via Equation (35) in E-step;
　4. Substitute $r_{l,k}^*(l = 1, \cdots, L, k = 1, \cdots, K)$ into Equations (36) and (37) to obtain $\mathbf{p}^{(i+1)}$ and
　$t_0^{(i+1)}$ through M-step, respectively. And then set $i = i + 1$;
　5. Calculate $\Delta = \left\| \mathbf{p}^{(i)} - \mathbf{p}^{(i-1)} \right\|$. If $\Delta \leqslant \varepsilon$, stop the iterations; Otherwise, set $i = i + 1$, repeat
　steps 3–4;
Output: The estimated position of target $\hat{\mathbf{p}} = \mathbf{p}^{(i)}$.

Remark 1. *By choosing the hidden data* **X**, *the high-dimensional nonlinear problem in Equation* (14) *is simplified as a few subproblems with soft computational load. The final operations are the line search of $r_{l,k}$ in the E-step and the closed-form solution of* **p** *and the line search of t_0 in the M-step.*

4.3. Computational Complexity Analysis

Based on the above analysis, we have achieved the transformation from high-dimensional multi-parameter nonlinear problem into several optimization subproblems, whose computational complexity is substantially reduced. Nonetheless, the computational complexity differs among the traversal search method, Weiss's method [21], and the proposed method. The traversal search method requires a three-dimensional search to find the extreme of the cost function corresponding to the emitter position. Weiss's method mainly has a two-dimensional search of the maximum eigenvalue of a Hermitian matrix. The proposed method is dominated by the line searches of $r_{l,k}^*(l = 1, \cdots, L, k = 1, \cdots, K)$ in Equation (35) and t_0 in Equation (37).

To better exhibit the computational complexity, Table 2 lists the computational complexity of the traversal search method, Weiss's method, and the proposed method. Note that N_p is the number of grid search points with respect to a line search, and I_{iter} is the number of iterations of the proposed method. It must be emphasized that the key contributor to computational complexity is N_p. Compared with the two- or three-dimensional searches in other methods, the proposed method only has a one-dimensional search, which reduces the computational complexity significantly.

Table 2. Computational complexity.

Algorithm	Amount of Computation
Traversal search method	$O\left(N_p^3 KL \left(8N^3 + 7N^2 + N\right)\right)$
Weiss's method	$O\left(N_p^2 \left(K \left((5L + 3)N^3 + 3N^2 + 2L^3\right)\right)\right)$
Proposed method	$O\left(I_{iter}\left(N_p KL \left(14N^3 + 8N^2 + 3N\right) + KL\right)\right)$

5. Numerical Experiments

In this section, several numerical experiments are reported to corroborate the theoretical analysis. All simulations results are based on 200 Monte Carlo trials. In this scenario, the receivers equipped with only one sensor are included, and the source emits a Gaussian random signal with a center frequency of $f_c = 200$ MHz. The channel attenuation coefficient amplitude obeys a normal distribution with a mean of 1 and a standard deviation of 0.1, and the channel phase obeys a uniform distribution over $[-\pi, \pi]$. Unless otherwise specified, we collect $N = 32$ sample points in each interval at a sampling

rate of f_s = 400 kHz, use L = 3 receivers, perform a total of K = 5 observations, and set the velocity of receiver v = 300 m/s.

To examine the performance comparisons, we take simulations works with the following four algorithms

1. the proposed method in this study;
2. the traversal search method;
3. Weiss's method;
4. the TOA/Doppler two-step algorithm.

The details of the comparison algorithms are shown in Table 3.

Table 3. Algorithms.

Algorithm	Method
TOA/Doppler two-step method	Method in [2] to estimate TOA ; Method in [8] to estimate Doppler; Least square (LS) location using the TOA and Doppler estimations.
Traversal search method	Method in [17] using a three-dimensional search
Weiss's method	Method in [21] using eigenvalue decomposition

Additionally, the CRB presented in Appendix B is also included in this part, providing a theoretical best performance benchmark. Moreover, root mean square error (RMSE) and cumulative distribution function (CDF) are adopted to evaluate localization accuracy in this paper.

We now examine the localization performance of the proposed algorithm for three different scenarios. The target locates at [5, 4] km, and the receivers move along different trajectories, which can be found in Figure 1. As shown in Figure 2, the performance of our algorithm is not sensitive to the receiver trajectories because there is no significant difference between the RMSE curves of our algorithm for different scenarios. Additionally, as expected, the proposed method for Scenario (a) has better localization precision in low SNRs. On the other hand, although the RMSE of our algorithm is far from that of the corresponding CRB in low SNRs, it has a substantial advantage in computation time (see the descriptions for Table 4).

(a)

(b)

(c)
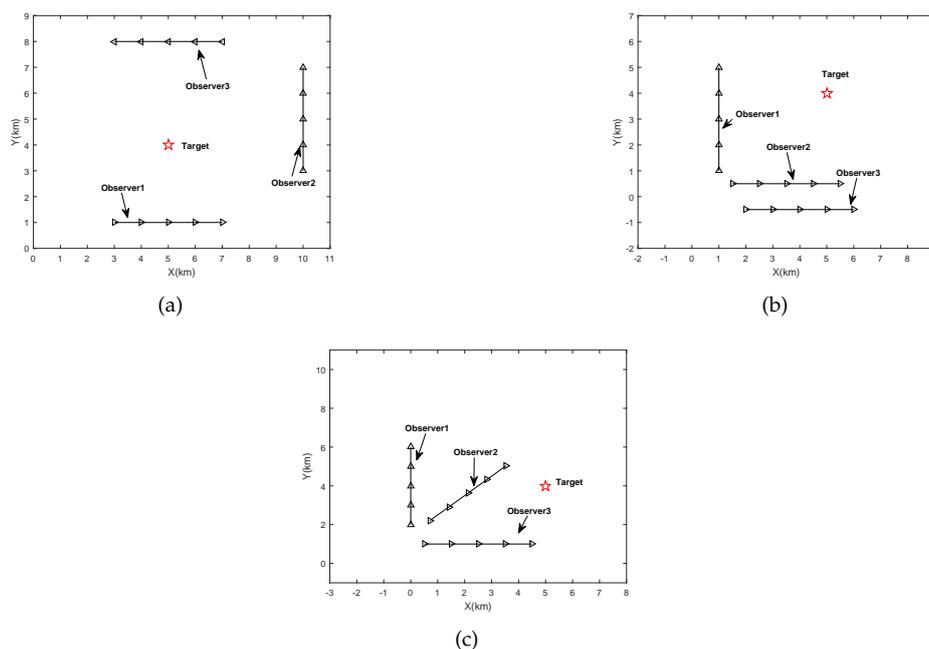
Figure 1. The simulated localization scenarios: (**a**) scenario 1, (**b**) scenario 2, (**c**) scenario 3.
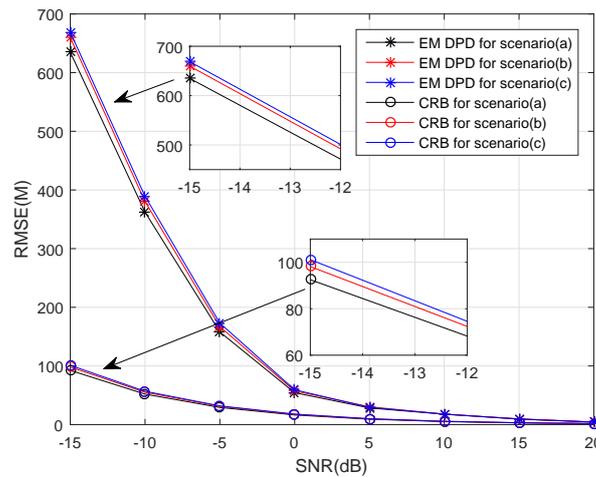
Figure 2. The RMSE and Cramér–Rao bound (CRB) versus SNRs for different scenarios.

It must be emphasized that the next simulations are all based on Scenario (a) in Figure 1. Firstly, we continue to conduct the simulation for algorithm performance comparison versus SNRs. The RMSEs of all algorithms can be easily found in Figure 3. Unsurprisingly, all DPD methods outperform the two-step method, and the traversal search method shows the best localization performance especially in low SNRs. Although the performance of the proposed method is slightly inferior than that of Weiss's method in low SNRs, it is more efficient. However, the other two DPD methods will cost a high amount of time in the presence of limited computing resources, losing its timeliness to moving target. The detailed runtime information can be seen in Table 4.
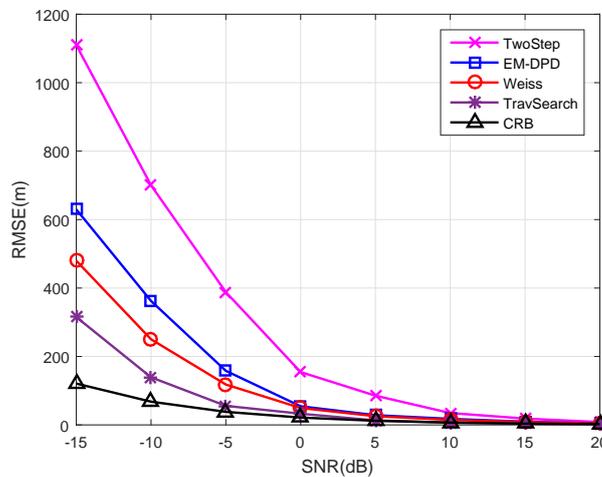
Figure 3. The RMSEs of different algorithms versus SNRs.

Secondly, to better exhibit the computational complexity of each algorithm, the runtime is investigated. From the results in Table 4, we observe that the two-step method requires the least runtime, but its localization performance is deteriorated in low SNRs (see Figure 3). Due to the high-dimensional search, the traversal search method is more computationally expensive than other DPD methods. Obviously, Weiss's method has a great improvement in the running time. Compared with Weiss's method, the run time of our method is further reduced (approximately 40%). However, the reduction of computational complexity of our method does not deteriorate localization performance significantly (also see Figure 3). As indicated by these finding, the proposed approach receives acceptable localization performance with a low computation resource cost.

Table 4. Mean runtime of different methods.

Algorithm	Runtime (s)
Two-step method	0.8
Traversal search method	67
Weiss's method	7.9
Proposed method	4.7

Thirdly, the CDF curves of all methods at the nominal constellation are depicted in Figure 4a,b at the SNR level of 5 dB and −10 dB, respectively. It can be readily observed that the traversal search curve is highest at a different localization error level. The proposed method curve is associated with Weiss's DPD curve, which indicates it can receive high localization accuracy with high SNRs. From Figure 4, the two-step method curve deviates from the DPD methods. The proposed method curve is slightly apart from the curve of Weiss's method, which is acceptable in terms of low complexity.

(a)

(b)
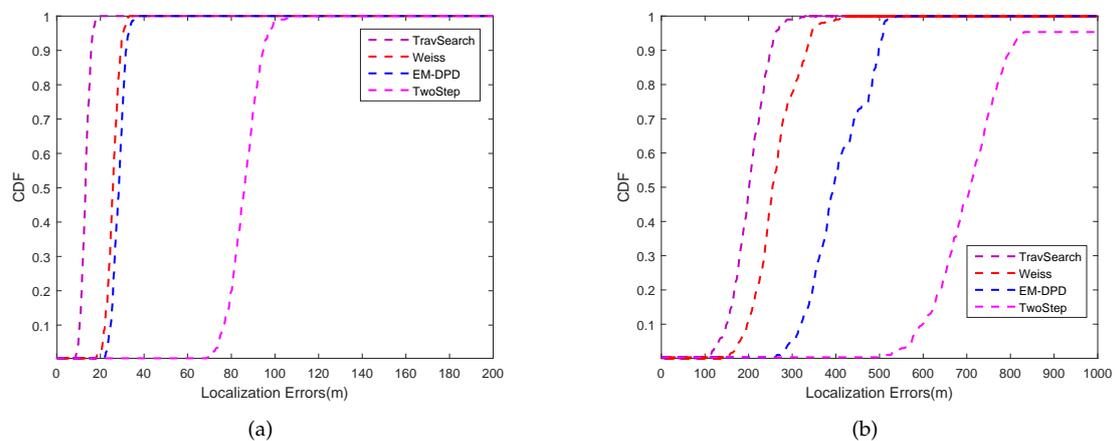
Figure 4. The CDF of different algorithms versus localization errors: (**a**) SNR = 5 dB; (**b**) SNR = −10 dB.

Finally, to further examine the influence of system parameter values on localization performance, we set SNR = −5 dB and exhibit the RMSE curves for *K*,*N*,*L*, and *v* in Figure 5, respectively. Additionally, the simulation conditions are *K* ranging from 2 to 10, *N* ranging from 50 to 200, *L* ranging from 1 to 6, and *v* ranging from 100 to 1000. It can be readily found in Figure 5a–c that all algorithms perform better in terms of localization accuracy, as more system parameter information exists. These results imply that the increase of these system parameters can enhance the localization performance. However, more parameter requirements mean greater system overhead and computational pressure. Consequently, our method can play an excellent role in the presence of limited computational resources. Furthermore, Figure 5d indicates that *v* does not have much impact on position accuracy.
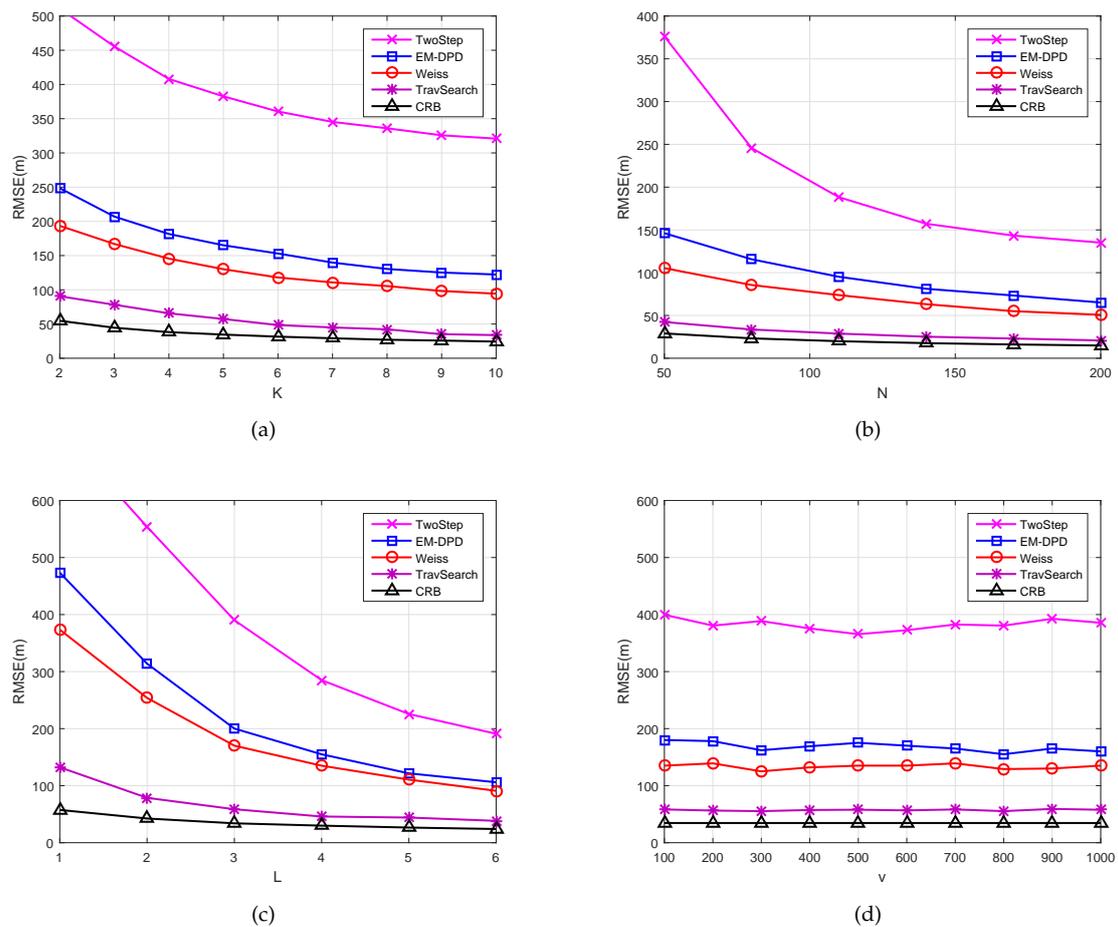
Figure 5. The RMSEs of different algorithms versus system parameters: (**a**) the value of K; (**b**) the value of N; (**c**) the value of L; (**d**) the value of v.

6. Conclusions

The traditional two-step methods have high computational efficiency but low localization performance. To enhance positioning accuracy, DPD approaches have been developed. Since the DPD technology locates the target from the signal directly, computational complexity of this estimator is high. To solve this problem, this paper proposes a fast ML-based direct localization method using an EM algorithm based on time delay and Doppler shift. The EM scheme was developed to solve the ML problem in the DPD model. We simplify this ML problem via the theories of Gibbs inequality and Laplace approximation, transforming the high-dimensional nonlinear search into a few one-dimensional searches. Simulation results show that the proposed algorithm has operates faster than other DPD methods. Furthermore, when limited computing resources appear, our method leads a more efficient result on the basis of guaranteeing high localization accuracy. Therefore, compared with other localization approaches, the proposed method becomes a good way to balance localization accuracy and computational complexity.

Author Contributions: T.Q. derived the proposed method. T.Q. and L.L. conceived and designed the experiments. T.Q. and B.B. performed the simulations. D.W. analyzed the results. T.Q. wrote the paper. L.L. reviewed the paper.

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

Appendix A. Evaluation of $\bar{x}_{l,k}$ and $\bar{G}_{l,k}$ via Laplace Method

We consider the multivariate functions $h(\mathbf{z}) : R^n \to R^m$ and $f(\mathbf{z}) : R^n \to R$. An integral is expressed by

$$I_f = \int_{\mathbf{z} \in \mathbf{Z}} h(\mathbf{z}) e^{f(\mathbf{z})} d\mathbf{z}. \tag{A1}$$

Assume that $f(\mathbf{z})$ has a global maximum within the interval. According to [27], the Laplace approximation is written as

$$I_f \approx (2\pi)^{n/2} |-F(\mathbf{z}^*)|^{-1/2} h(\mathbf{z}^*) e^{f(\mathbf{z}^*)} \tag{A2}$$

where $F(\mathbf{z}^*)$ is the Hessian of $f(\mathbf{z})$ at $\mathbf{z}^*$. Next we approximate the integral of $\bar{x}_{l,k}$ and $\bar{G}_{l,k}$. For the sake of brevity, we omit l, k in the subsequent derivation.

The expression of the conditional expectation is

$$\begin{aligned}
\bar{h}(\tilde{\mathbf{x}}) &\triangleq \int h(\tilde{\mathbf{x}}) p_{\tilde{\mathbf{x}}/\tilde{\mathbf{y}}} d\tilde{\mathbf{x}} = \int h(\tilde{\mathbf{x}}) \frac{p_{\tilde{\mathbf{x}},\tilde{\mathbf{y}}}}{p_{\tilde{\mathbf{y}}}} d\tilde{\mathbf{x}} \\
&\propto p_{\tilde{\mathbf{y}}}^{-1} \int h(\tilde{\mathbf{x}}) e^{-\frac{\|\tilde{\mathbf{x}} - \varepsilon(\mathbf{p})\|^2}{\sigma_{\tilde{x}}^2} - \frac{\|\tilde{\mathbf{y}} - \hat{b} \mathbf{G}(\tilde{\mathbf{x}})\|^2}{\sigma^2}} d\tilde{\mathbf{x}}.
\end{aligned} \tag{A3}$$

The polar coordinates can be written as

$$\tilde{\mathbf{x}} = r \begin{bmatrix} \cos \xi \\ \sin \xi \end{bmatrix}, \quad \varepsilon = R \begin{bmatrix} \cos \eta \\ \sin \eta \end{bmatrix}. \tag{A4}$$

Substituting (A4) into (A3) yields

$$\bar{h}(r, \xi) \propto p_{\tilde{\mathbf{y}}}^{-1} \int_0^{\infty} r e^{f(r)} dr \int_{-\pi}^{\pi} h(r, \xi) e^{\alpha r \cos(\xi - \eta)} d\xi \tag{A5}$$

where $f(r) = -\sigma_x^{-2} \left(\frac{r^2}{2} - \lambda \left\| \tilde{\mathbf{y}} - \hat{b} \mathbf{G}(r) \right\|^2 \right)$, $\lambda = \frac{\sigma_x^2}{\sigma^2}$ and $\alpha = R \sigma_x^{-2}$. We assume $r \gg 1$ and have [28]

$$\begin{aligned}
I_1(r) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos \xi e^{r \cos \xi} d\xi \approx \frac{e^r}{\sqrt{2\pi r}}, \\
I_0(r) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{r \cos \xi} d\xi \approx \frac{e^r}{\sqrt{2\pi r}}, \\
0 &= \int_{-\pi}^{\pi} \sin \xi e^{r \cos \xi} d\xi
\end{aligned} \tag{A6}$$

where $I_1(\bullet)$, $I_0(\bullet)$ denote the first kind of modified Bessel function. We substitute Equation (A6) into Equation (A5), yielding

$$\bar{\tilde{\mathbf{x}}} \propto \begin{bmatrix} \cos \eta \\ \sin \eta \end{bmatrix} \int_0^{\infty} r^{3/2} e^{f(r) + \alpha r} dr. \tag{A7}$$

Recalling the expression in Equation (A2), we can search for the global extreme r^* through

$$r^* = \arg \max_r f(r) + \alpha r. \tag{A8}$$

Therefore, we have

$$\bar{\tilde{\mathbf{x}}} \approx \begin{bmatrix} \cos \eta \\ \sin \eta \end{bmatrix} r^* = \frac{\varepsilon}{\|\varepsilon\|} r^*. \tag{A9}$$

Similarly, we have $\bar{\mathbf{G}} \approx \mathbf{G}\left(r^{*}\right)$.

Appendix B. Derivation of the Cramér–Rao Bound

It is well known that the CRB provides a benchmark for the best localization accuracy for any unbiased estimator. This section derives the compact CRB formula for source position with known signal waveforms and unknown transmitted time.

Let $\mathbf{D}_{l,k}\left(\boldsymbol{\theta}\right) = b_{l,k}\mathbf{C}_{l,k}\tilde{\mathbf{s}}_{k}$. According to [22], the q,jth entry of the fisher information matrix of unknown parameter $\boldsymbol{\theta}$ is shown as

$$[\mathbf{J}]_{q,j} = \frac{2}{\sigma^2} \sum_{k=1}^{K} \sum_{l=1}^{L} \left(\frac{\partial \mathbf{D}_{l,k}\left(\boldsymbol{\theta}\right)}{\partial [\boldsymbol{\theta}]_q^{\mathrm{T}}}\right)^{\mathrm{H}} \left(\frac{\partial \mathbf{D}_{l,k}\left(\boldsymbol{\theta}\right)}{\partial [\boldsymbol{\theta}]_j^{\mathrm{T}}}\right). \tag{A10}$$

For easy derivation, $\boldsymbol{\theta}$ can be redefined by

$$\boldsymbol{\theta} = \left[\mathbf{b}^{(R)\mathrm{T}}, \mathbf{b}^{(I)\mathrm{T}}, t_0, \mathbf{p}^{\mathrm{T}}\right]^{\mathrm{T}} \tag{A11}$$

where $\mathbf{b} = \left[\mathbf{b}_1^{\mathrm{T}}, \mathbf{b}_2^{\mathrm{T}}, \cdots, \mathbf{b}_K^{\mathrm{T}}\right]^{\mathrm{T}}$ and $\mathbf{b}_k = \left[\mathbf{b}_{1,k}^{\mathrm{T}}, \mathbf{b}_{2,k}^{\mathrm{T}}, \cdots, \mathbf{b}_{L,k}^{\mathrm{T}}\right]^{\mathrm{T}}$. Thus, $\mathbf{J}$ can be expressed by

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\mathbf{pp}} & \mathbf{J}_{\mathbf{p}t_0} & \mathbf{J}_{\mathbf{pb}^{(R)}} & \mathbf{J}_{\mathbf{pb}^{(I)}} \\ \mathbf{J}_{t_0\mathbf{p}} & \mathbf{J}_{t_0t_0} & \mathbf{J}_{t_0\mathbf{b}^{(R)}} & \mathbf{J}_{t_0\mathbf{b}^{(I)}} \\ \mathbf{J}_{\mathbf{b}^{(R)}\mathbf{p}} & \mathbf{J}_{\mathbf{b}^{(R)}t_0} & \mathbf{J}_{\mathbf{b}^{(R)}\mathbf{b}^{(R)}} & \mathbf{J}_{\mathbf{b}^{(R)}\mathbf{b}^{(I)}} \\ \mathbf{J}_{\mathbf{b}^{(I)}\mathbf{p}} & \mathbf{J}_{\mathbf{b}^{(I)}t_0} & \mathbf{J}_{\mathbf{b}^{(I)}\mathbf{b}^{(R)}} & \mathbf{J}_{\mathbf{b}^{(I)}\mathbf{b}^{(I)}} \end{bmatrix}. \tag{A12}$$

Next we calculate the subblocks of Equation (A12).

Appendix B.1. The Partial of $D_{l,k}\left(\theta\right)$ with Respect to b

Note that $\mathbf{b}$ is a complex variable, so we obtain

$$\begin{aligned} \left[\frac{\partial \mathbf{D}_{l,k}\left(\boldsymbol{\theta}\right)}{\partial \mathbf{b}^{(R)\mathrm{T}}}\right]_{m,n} &= \begin{cases} \mathbf{G}_{l,k} & m = l, n = k \\ \mathbf{0} & m \neq l, n \neq k \end{cases}, \\ \left[\frac{\partial \mathbf{D}_{l,k}\left(\boldsymbol{\theta}\right)}{\partial \mathbf{b}^{(I)\mathrm{T}}}\right]_{m,n} &= \begin{cases} j\mathbf{G}_{l,k} & m = l, n = k \\ \mathbf{0} & m \neq l, n \neq k \end{cases}. \end{aligned} \tag{A13}$$

Appendix B.2. The Partial of $D_{l,k}\left(\theta\right)$ with Respect to t_0

We obtain

$$\frac{\partial \mathbf{D}_{l,k}\left(\boldsymbol{\theta}\right)}{\partial t_0} = \mathbf{H}_{l,k}, \tag{A14}$$

where $\mathbf{H}_{l,k} = b_{l,k}\mathbf{A}_{l,k}\mathbf{F}^{\mathrm{H}}\mathbf{D}_{\tau_{l,k}}\mathbf{F}\mathbf{F}^{\mathrm{H}}\dot{\mathbf{D}}_{t_0}\mathbf{F}\tilde{\mathbf{s}}_k$, with $\dot{\mathbf{D}}_{t_0} = \mathbf{D}_{t_0} \cdot \mathrm{diag}\left\{-j2\pi\frac{\tilde{\mathbf{N}}}{NT_s}\right\}$.

Appendix B.3. The Partial of $D_{l,k}\left(\theta\right)$ with Respect to p

We define $\boldsymbol{\Gamma}_{l,k} = \frac{\partial \mathbf{A}_{l,k}\mathbf{F}^{\mathrm{H}}\mathbf{D}_{\mathsf{T}_{l,k}}\mathbf{F}}{\partial \mathbf{p}^{\mathrm{T}}}\left(\mathbf{I}_2 \otimes \tilde{\mathbf{s}}_k^{t_0}b_{l,k}\right)$ with $\tilde{\mathbf{s}}_k^{t_0} = \left[\tilde{s}_k\left(T_s - t_0\right), \tilde{s}_k\left(2T_s - t_0\right), \cdots, \tilde{s}_k\left(NT_s - t_0\right)\right]^{\mathrm{T}}$. The derivative with respect to $\mathbf{p}$ can be expressed by

$$\frac{\partial \mathbf{D}_{l,k}\left(\boldsymbol{\theta}\right)}{\partial \mathbf{p}^{\mathrm{T}}} = \boldsymbol{\Gamma}_{l,k} \tag{A15}$$

where

$$\frac{\partial \mathbf{A}_{l,k}\mathbf{F}^H\mathbf{D}_{T_{l,k}}\mathbf{F}}{\partial \mathbf{p}^T} = \frac{\partial \mathbf{A}_{l,k}}{\partial \mathbf{p}^T}\left(\mathbf{I}_2 \otimes \mathbf{F}^H\mathbf{D}_{\tau_{l,k}}\mathbf{F}\right) + \mathbf{A}_{l,k}\mathbf{F}^H\frac{\partial \mathbf{D}_{\tau_{l,k}}}{\partial \mathbf{p}^T}\left(\mathbf{I}_2 \otimes \mathbf{F}\right), \tag{A16}$$

with

$$\frac{\partial \mathbf{A}_{l,k}}{\partial \mathbf{p}^T} = \mathbf{A}_{l,k} \cdot \mathrm{diag}\left\{j2\pi\frac{\partial f_{l,k}}{\partial \mathbf{p}^T}T_s, j2\pi\frac{\partial f_{l,k}}{\partial \mathbf{p}^T}2T_s, \ldots, j2\pi\frac{\partial f_{l,k}}{\partial \mathbf{p}^T}NT_s\right\} \tag{A17}$$

$$\frac{\partial f_{l,k}}{\partial \mathbf{p}^T} = \frac{f_c}{c}\left(\frac{\mathbf{v}_{l,k}^T}{\|\mathbf{p}-\mathbf{p}_{l,k}\|} - \frac{\mathbf{v}_{l,k}^T\left(\mathbf{p}-\mathbf{p}_{l,k}\right)\left(\mathbf{p}-\mathbf{p}_{l,k}\right)^T}{\|\mathbf{p}-\mathbf{p}_{,k}\|^3}\right) \tag{A18}$$

$$\frac{\partial \mathbf{D}_{\tau_{l,k}}}{\partial \mathbf{p}^T} = \mathbf{D}_{\tau_{l,k}} \cdot \mathrm{diag}\left\{-j2\pi\frac{1}{NT_s}\frac{\partial \tau_{l,k}}{\partial \mathbf{p}^T}, -j2\pi\frac{2}{NT_s}\frac{\partial \tau_{l,k}}{\partial \mathbf{p}^T}_s, \ldots, -j2\pi\frac{N}{NT_s}\frac{\partial \tau_{l,k}}{\partial \mathbf{p}^T}_s\right\} \tag{A19}$$

$$\frac{\partial \tau_{l,k}}{\partial \mathbf{p}^T} = \frac{1}{c}\frac{\left(\mathbf{p}-\mathbf{p}_{l,k}\right)^T}{\|\mathbf{p}-\mathbf{p}_{l,k}\|}. \tag{A20}$$

By substituting Equations (A13)–(A15) into Equation (A10), we can obtain the subblocks of $\mathbf{J}_{\xi\xi}$

$$\mathbf{J}_{\mathbf{b}_{m_1,n_1}^{(R)}\mathbf{b}_{m_2,n_2}^{(R)}} = \begin{cases} \frac{2}{\sigma^2}\,\mathrm{Re}\left\{\mathbf{G}_{m_1,n_1}^H\mathbf{G}_{m_2,n_2}\right\} & m_1=m_2 \cap n_1=n_2 \\ \mathbf{0} & m_1\neq m_2 \cup n_1\neq n_2 \end{cases}. \tag{A21}$$

$$\mathbf{J}_{\mathbf{b}_{m_1,n_1}^{(R)}\mathbf{b}_{m_2,n_2}^{(I)}} = \begin{cases} -\frac{2}{\sigma^2}\,\mathrm{Im}\left\{\mathbf{G}_{m_1,n_1}^H\mathbf{G}_{m_2,n_2}\right\} & m_1=m_2 \cap n_1=n_2 \\ \mathbf{0} & m_1\neq m_2 \cup n_1\neq n_2 \end{cases}. \tag{A22}$$

$$\mathbf{J}_{\mathbf{b}_{m,n}^{(R)}t_0} = \frac{2}{\sigma^2}\,\mathrm{Re}\left\{\mathbf{G}_{m,n}^H\mathbf{H}_{m,n}\right\}. \tag{A23}$$

$$\mathbf{J}_{\mathbf{b}_{m,n}^{(R)}\mathbf{p}} = \frac{2}{\sigma^2}\,\mathrm{Re}\left\{\mathbf{G}_{m,n}^H\mathbf{\Gamma}_{m,n}\right\}. \tag{A24}$$

$$\mathbf{J}_{\mathbf{b}_{m_1,n_1}^{(I)}\mathbf{b}_{m_2,n_2}^{(I)}} = \begin{cases} \frac{2}{\sigma^2}\,\mathrm{Re}\left\{\mathbf{G}_{m_1,n_1}^H\mathbf{G}_{m_2,n_2}\right\} & m_1=m_2 \cap n_1=n_2 \\ \mathbf{0} & m_1\neq m_2 \cup n_1\neq n_2 \end{cases}. \tag{A25}$$

$$\mathbf{J}_{\mathbf{b}_{m,n}^{(I)}t_0} = -\frac{2}{\sigma^2}\,\mathrm{Im}\left\{\mathbf{G}_{m,n}^H\mathbf{H}_{m,n}\right\}. \tag{A26}$$

$$\mathbf{J}_{\mathbf{b}_{m,n}^{(I)}\mathbf{p}} = -\frac{2}{\sigma^2}\,\mathrm{Im}\left\{\mathbf{G}_{m,n}^H\mathbf{\Gamma}_{m,n}\right\}. \tag{A27}$$

$$\mathbf{J}_{t_0 t_0} = \frac{2}{\sigma^2}\sum_{l=1}^{L}\sum_{k=1}^{K}\mathrm{Re}\left\{\mathbf{H}_{l,k}^H\mathbf{H}_{l,k}\right\}. \tag{A28}$$

$$\mathbf{J}_{t_0\mathbf{p}} = \frac{2}{\sigma^2}\sum_{k=1}^{K}\sum_{l=1}^{L}\mathrm{Re}\left\{\mathbf{H}_{l,k}^H\mathbf{\Gamma}_{l,k}\right\}. \tag{A29}$$

$$\mathbf{J}_{\mathbf{pp}} = \frac{2}{\sigma^2}\sum_{l=1}^{L}\sum_{k=1}^{K}\mathrm{Re}\left\{\mathbf{\Gamma}_{l,k}^H\mathbf{\Gamma}_{l,k}\right\}. \tag{A30}$$

Since submatrices of Equation (A12) with respect to diagonal symmetry are transposed matrices of each other, we can easily obtain the remaining submatrices.

Thus, the inverse of the CRB of position can be obtained by

$$
(\mathbf{CRB_p})^{-1} = \mathbf{J_{pp}} - \begin{bmatrix} \mathbf{J_{pb}}^{(R)} & \mathbf{J_{pb}}^{(I)} & \mathbf{J_p}t_0 \end{bmatrix} \begin{bmatrix} \mathbf{J_{b}}^{(R)}\mathbf{b}^{(R)} & \mathbf{J_{b}}^{(R)}\mathbf{b}^{(I)} & \mathbf{J_{b}}^{(R)}t_0 \\ \mathbf{J_{b}}^{(I)}\mathbf{b}^{(R)} & \mathbf{J_{b}}^{(I)}\mathbf{b}^{(I)} & \mathbf{J_{b}}^{(I)}t_0 \\ \mathbf{J_{t_0}}\mathbf{b}^{(R)} & \mathbf{J_{t_0}}\mathbf{b}^{(I)} & \mathbf{J_{t_0 t_0}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{J_{pb}}^{(R)} & \mathbf{J_{pb}}^{(I)} & \mathbf{J_p}t_0 \end{bmatrix}^{\mathrm{T}}. \quad \text{(A31)}
$$

This concludes the derivation.

References

1. Zhang, Y.; Xu, X.; Sheikh, Y.A.; Ye, Z. A rank-reduction based 2-D DOA estimation algorithm for three parallel uniform linear arrays. *Signal Process.* **2016**, *120*, 305–310. [CrossRef]
2. Oh, D.; Kim, S.; Yoon, S.H.; Chong, J.W. Two-dimensional ESPRIT-like shift-invariant TOA estimation algorithm using multi-band chirp signals robust to carrier frequency Offset. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 3130–3139. [CrossRef]
3. Cao, H.; Chan, Y.T.; So, H.C. Maximum likelihood TDOA estimation from compressed sensing samples without reconstruction. *IEEE Signal Process. Lett.* **2017**, *24*, 564–568. [CrossRef]
4. Tahat, A.; Kaddoum, G.; Yousefi, S.; Valaee, S.; Gagnon, F. A look at the recent wireless positioning techniques with a focus on algorithms for moving receivers. *IEEE Access.* **2016**, *4*, 6652–6680. [CrossRef]
5. Gajewski, P.; Ziolkowski, C.; Kelner, J.M. Using SDF method for simultaneous location of multiple radio transmitters. In Proceedings of the 2012 19th International Conference on Microwave Radar and Wireless Communications (MIKON), Warsaw, Poland, 21–23 May 2012; IEEE: Warsaw, Poland, 2012; Volume 2, pp. 634–637.
6. Kelner, J.M.; Ziolkowski, C.; Nowosielski, L.; Wnuk, M. Localization of emission source in urban environment based on the Doppler effect. In Proceedings of the 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), Nanjing, China, 15–18 May 2016; pp. 1–5.
7. Yeredor, A.; Angel, E. Joint TDOA and FDOA Estimation: A Conditional Bound and Its Use for Optimally Weighted Localization. *IEEE Trans. Signal Process.* **2011**, *59*, 1612–1623. [CrossRef]
8. Stein, A. Differential delay/Doppler ML estimation with unknown signals. *IEEE Trans. Signal Process.* **1993**, *41*, 2717–2719. [CrossRef]
9. Viberg, M.; Ottersten, B. Sensor array processing based on subspace fitting. *IEEE Trans. Signal Process.* **1991**, *39*, 1110–1121. [CrossRef]
10. Wang, Y.L.; Wu, Y. An efficient semidefinite relaxation algorithm for moving source localization using TDOA and FDOA measurements. *IEEE Commun. Lett.* **2017**, *21*, 80–83. [CrossRef]
11. Park, C.H.; Chang, J.H. Closed-Form Localization for Distributed MIMO Radar Systems Using Time Delay Measurements. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 1480–1490. [CrossRef]
12. Xu, N.W.; Tang, C.N.; Wu, S.H.; Li, G.L.; Yang, J.Y. Optimal design of microseismic monitoring networking and rrror analysis of seismic source location for rock slope. *Adv. Mater. Res.* **2011**, *163–167*, 2991–2999.
13. Papakonstantinou, K.; Slock, D. Direct location estimation using single-bounce NLOS time-varying channel models. In Proceedings of the 2008 IEEE 68th Vehicular Technology Conference, Calgary, BC, Canada, 21–24 September 2008; pp. 1–5.
14. Demissie, B. Direct localization and detection of multiple sources in multi-path environment. In Proceedings of the 14th International Conference on Information Fusion, Chicago, IL, USA, 5–8 July 2011; pp. 1–8.
15. Yin, J.X.; Wang, D.; Wu, Y.; Liu, R.R. A decoupled direct position determination algorithm for multiple targets in mixed LOS/NLOS environments. *Acta Aeronaut. Astronaut. Sin.* **2018**, *39*, 321338–321338.
16. Du, J.; Wang, D.; Yu, W.; Yu, H.; Du, J.; Wang, D.; Yu, W.; Yu, H. Direct position determination of unknown signals in the presence of multipath propagation. *Sensors* **2018**, *18*, 892.
17. Weiss, A.J.; Amar, A. Direct position determination of multiple radio signals. *EURASIP J. Appl. Signal Process.* **2005**, *1*, 1–13. [CrossRef]
18. Chen, L.; Zhang, H.; Sun, H. Multi-target direct position determination using subspace based compressive sensing. *J. Signal Process.* **2015**, *31*, 1272–1278.
19. Weiss, A.J. Direct position determination of narrowband radio frequency transmitters. *IEEE Signal Process. Lett.* **2004**, *11*, 513–516. [CrossRef]

20. Tzoreff, E.; Bobrovsky, B.Z.; Weiss, A.J. Single receiver emitter geolocation based on signal periodicity with oscillator instability. *IEEE Trans. Signal Process*. **2014**, *62*, 1377–1385. [CrossRef]

21. Weiss, A.J. Direct geolocation of wideband emitters based on delay and Doppler. *IEEE Trans. Signal Process*. **2011**, *59*, 2513–2521. [CrossRef]

22. Wang, D.; Zhang, G.; Shen, C.; Zhang, J. A direct position determination algorithm for constant modulus signals with single moving observer. *Acta Aeronaut. Astronaut. Sin*. **2016**, *37*, 1622–1633.

23. Lu, Z.; Wang, J.; Ba, B.; Wang, D. A novel direct position determination algorithm for orthogonal frequency division multiplexing signals based on the time and angle of arrival. *IEEE Access*. **2017**, *5*, 25312–25321. [CrossRef]

24. Reuven, A.M.; Weiss, A.J. Direct position determination of cyclostationary signals. *Signal Process*. **2009**, *89*, 2448–2464. [CrossRef]

25. Huda, S.; Yearwood, J.; Togneri, R. A stochastic version of Expectation Maximization algorithm for better estimation of Hidden Markov Model. *Pattern Recogn. Lett*. **2009**, *30*, 1301–1309. [CrossRef]

26. Katsaggelos, A.K. Image identification and restoration based on the expectation-maximization algorithm. *Opt Eng*. **1990**, *29*, 436. [CrossRef]

27. Mada, K.K.; Wu, H.-C.; Iyengar, S.S. Efficient and robust EM algorithm for multiple wideband source localization. *IEEE Trans. Veh. Technol*. **2009**, *58*, 3071–3075. [CrossRef]

28. Lu, L.; Wu, H.-C. Robust expectation-maximization direction-of-arrival estimation algorithm for wideband source signals. *IEEE Trans. Veh. Technol*. **2011**, *60*, 2395–2400. [CrossRef]

29. Lu, L.; Wu, H.-C.; Yan, K.; Iyengar, S.S. Robust expectation-maximization algorithm for multiple wideband acoustic source localization in the presence of nonuniform noise variances. *IEEE Sens. J*. **2011**, *11*, 536–544. [CrossRef]

30. Tzoreff, E.; Weiss, A.J. Expectation-maximization algorithm for direct position determination. *Signal Process*. **2017**, *133*, 32–39. [CrossRef]

31. Yin, J.X.; Wang, D.; Wu, Y.; Tang, T. Single-step localization using multiple moving arrays in the presence of observer location errors. *Signal Process*. **2018**, *152*, 382–410. [CrossRef]