RESEARCH   **Open Access**

CrossMark

Robust direct position determination methods in the presence of array model errors

Ding Wang[1,2], Jiexin Yin[1,2*] (iD), Ruirui Liu[1,2], Zhidong Wu[1,2] and Tao Tang[1,2]

Abstract

Direct position determination (DPD) methods are known to have many advantages over the traditional two-step localization method, especially for low signal-to-noise ratios (SNR) and/or short data records. However, similar to conventional direction-of-arrival (DOA) estimation methods, the performance of DPD estimators can be dramatically degraded by inaccuracies in the array model. In this paper, we present robust DPD methods that can mitigate the effects of these uncertainties in the array manifold. The proposed technique is related to the classical auto-calibration procedure under the assumption that prior knowledge of the array response errors is available. Localization is considered for the cases of both unknown and a priori known transmitted signals. The corresponding *maximum a posteriori* (MAP) estimators for these two cases are formulated, and two alternating minimization algorithms are derived to determine the source location directly from the observed signals. The Cramér-Rao bounds (CRBs) for position estimation are derived for both unknown and known signal waveforms. Simulation results demonstrate that the proposed algorithms are asymptotically efficient and very robust to array model errors.

Keywords: Array signal processing, Source localization, Direct position determination (DPD), Array model errors, Alternating minimization algorithm, Cramér-Rao bound (CRB)

1 Introduction

Emitter localization using direction-of-arrival (DOA) measurements [1–3] has received significant attention because of its importance in fields such as radar, sonar, seismology, vehicle navigation, and wireless communications. In this type of localization system, a single moving observer or multiple stationary observers are used to determine the emitter position. Generally, each base station is equipped with an antenna array that can be used to estimate the angle of arrival of the transmitted signals. With multiple DOA estimates from different observer locations, it is possible to locate the source [4–6]. This emitter localization procedure is called two-step location. In the first step, the signal parameters (e.g., DOA [1–6], time difference of arrival (TDOA) [7, 8],

time of arrival (TOA) [9, 10], frequency difference of arrival (FDOA) [11, 12], frequency of arrival (FOA) [13], received signal strength (RSS) [14], gain ratios of arrival (GROA) [15]) are obtained independently from the intercepted signals by spatially separated sensors. In the second step, the measurements of all sensors are transferred to a central unit and the transmitter location is estimated. Note that this two-step procedure is also known as the decentralized approach [16].

It is worth mentioning that, although the two-step approach is extensively used in modern localization systems, it is only a suboptimal position determination technique. This is because the signal metrics extracted from the waveforms ignore the constraint that all measurements must correspond to the same transmitter. Consequently, information loss between the two steps is inevitable. Indeed, the extended invariance principle (EXIP) [17] can be used to show that the two-step procedure can also provide an asymptotically efficient estimate under certain

* Correspondence: Cindyin0807@163.com
[1]National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450002, People's Republic of China
[2]Zhengzhou Information Science and Technology Institute, Zhengzhou 450002, Henan, People's Republic of China

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 2 of 28

conditions. In practice, however, these requirements are not easily satisfied.

To improve the localization accuracy of two-step location methods, the direct position determination (DPD) technique has been proposed and extensively developed. DPD is a centralized and single-step location technique in which the estimator uses exactly the same signal samples as in two-step methods, but estimates the source location directly, skipping the intermediate (first) step. This can be viewed as searching for the emitter location that best explains the collected data. In general, the DPD method is superior to the classical two-step methods for low signal-to-noise ratios (SNR) and/or short data records. Moreover, DPD can be applied to many wireless positioning systems. In particular, DPD methods for locating a narrowband radio emitter based on Doppler shift [18, 19] and for locating a wideband source using a time delay metric [20–22] have been presented. Additionally, DPD estimators using both the Doppler effect and the relative delay have been developed [23–26]. In the DPD methods mentioned above, several platforms each carrying single-antenna receivers are used for source location; hence, the DOA information of the impinging signals cannot be exploited. To use such signal metrics, DPD methods based on multiple static stations, each equipped with an antenna array, have been proposed [27]. This single-step location technique models the array response as a function of the source position, requiring only a two-dimensional search for planar geometry and a three-dimensional search for the general case. Following the methods in [27], other DPD estimators for special localization scenarios have been reported in the literature. Specifically, DPD methods for multiple-source scenarios are studied in [28, 29], and some high-resolution DPD methods are proposed in [30–32]. Additionally, DPD methods tailored to special signal structures (e.g., orthogonal frequency division multiplexing signals, cyclostationary signals, noncircular signals, and intermittent emissions) have been developed [33–37]. Note that the experimental results in [18–37] demonstrate that the single-step approach outperforms the two-step method in scenarios with low SNR and/ or relatively few data records.

High-resolution DOA estimation methods are known to be sensitive to errors in the array response model [38–42]. This is because these approaches require exact knowledge of the array manifold. Note that the DPD methods using an antenna array also rely on the accuracy of the array model. As a result, it is reasonable to expect their localization accuracy to deteriorate significantly in the event of an array

model mismatch. In [43–45], for the case of a known signal waveform, the closed-form expression for the mean square error (MSE) of the DPD estimator was derived for the case of array model errors. However, in practice, the signal waveform is rarely known. Consequently, an alternative statistical analysis of the DPD method in the presence of array model errors was performed [46] under the assumption that the signal waveform is unknown. All the theoretical and experimental results in [43–46] demonstrate that the accuracy loss caused by array model errors is considerable. Hence, a new DPD technique that accounts for array model errors is needed to improve the emitter location accuracy. However, to the best of our knowledge, very few studies have considered this topic. This paper presents robust DPD methods that can reduce the impact of uncertainties in the array manifold. The proposed technique is similar to the traditional auto-calibration procedure in the field of array signal processing and assumes that certain prior knowledge of the array response errors is available. This is a reasonable assumption in most applications, and it allows for more general perturbation models. We consider two different localization cases: the case of a priori known signal waveforms and the more realistic case where the transmitted signals are unknown to the location systems. The corresponding *maximum* a posteriori (MAP) estimators for the two cases are formulated and two alternating minimization algorithms are developed to estimate the emitter position directly from the received signal samples. Hence, the proposed methods follow the Bayesian framework [47–50]. Additionally, to verify the asymptotic efficiency of the new methods, the Cramér-Rao bounds (CRBs) for position estimation are derived for both unknown and known signal waveforms.

The remainder of this paper is organized as follows. Section 2 describes the method and experimental used in this paper. In Section 3, the signal model for DPD is formulated and the array error model is also discussed. Section 4 presents a robust DPD method in the presence of array model errors, when signal waveform is unknown. In Section 5, an alternative robust DPD method in the presence of uncertainties in array manifold is proposed for the case of known transmitted signals. In Section 6, the CRB expressions for the position estimation are derived for both unknown and known signal waveforms. Simulation results are reported in Section 7. Conclusions are drawn in Section 8. The proofs of the main results are given in the Appendix 1, 2, 3, 4, and 5.

Wang et al. EURASIP Journal on Advances in Signal Processing (2018) 2018:38

Page 3 of 28

2 Methods/experimental

This study refers to the field of emitter localization. The aim of this study is to propose the robust DPD methods that can mitigate the effects of these uncertainties in the array manifold. The methods are designed based on MAP criterion. We consider two different localization cases: the case of a priori known signal waveforms and the more realistic case where the transmitted signals are unknown to the location systems. The numerical optimization technology is applied and two alternating minimization algorithms to identify the unknown parameters are presented. To verify the asymptotic efficiency of the new methods, we derive the CRBs on position estimation for both unknown and known signal waveforms.

We perform a set of Monte Carlo simulations to examine the behavior of the proposed robust DPD methods. The root mean square error (RMSE) of position estimate is employed to assess and compare the performance. All the simulation results are averaged over 2000 independent runs. We compare our algorithms with the algorithms in [27], and the traditional two-step localization algorithms, as well as the CRB for unknown and known signal waveforms. Besides, all the experiments are conducted for the 3D localization.

All data and procedures performed in the paper are in accordance with the ethical standards of research community. This paper does not contain any studies with human participants or animals performed by any of the authors.

3 Problem formulation

3.1 Signal model for direct position determination

3.1.1 Time-domain signal model

Consider a stationary radio emitter and N base stations that can intercept the transmitted signal. Each base station is equipped with an antenna array composed of M sensors. The transmitter's position is denoted by an $L \times 1$ vector of coordinates $\mathbf{p}$. In this paper, we consider the three-dimensional (3D) scenario, in which $\mathbf{p} = [p_x \quad p_y \quad p_z]^T$ and $L = 3$. The complex envelopes of the signal observed by the nth base station are modeled by

$$\mathbf{x}_n(t) = \beta_n \mathbf{a}_n(\mathbf{p}, \boldsymbol{\mu}_n) s(t - \tau_n(\mathbf{p}) - t_0) + \boldsymbol{\varepsilon}_n(t) \quad (1 \le n \le N) \tag{1}$$

where

- $\mathbf{a}_n(\mathbf{p})$ is the array response of the nth station to a signal transmitted from position $\mathbf{p}$,

- $s(t - \tau_n(\mathbf{p}) - t_0)$ is the signal waveform transmitted at unknown time t_0 and delayed by $\tau_n(\mathbf{p})$,
- $\tau_n(\mathbf{p})$ is the signal propagation time from the emitter to the nth base station (i.e., distance divided by signal propagation speed),
- β_n is an unknown complex scalar representing the channel attenuation between the emitter and the nth base station,
- $\boldsymbol{\varepsilon}_n(t)$ denotes temporally white, circularly symmetric complex Gaussian noise with zero mean and covariance matrix $\sigma_\varepsilon^2 \mathbf{I}_M$,
- $\boldsymbol{\mu}_n$ represents the perturbed array parameters and is used to model the uncertainty in the array steering vector.

We assume that the observation vector $\mathbf{x}_n(t)$ is sampled with period T. The kth data sample can then be expressed as

$$\mathbf{x}_{n,k} = \beta_n \mathbf{a}_n(\mathbf{p}, \boldsymbol{\mu}_n) s(kT - \tau_n(\mathbf{p}) - t_0) + \boldsymbol{\varepsilon}_{n,k} \quad (1 \le k \le K) \tag{2}$$

where K is the number of snapshots.

3.1.2 Frequency-domain signal model

In order to determine the emitter position directly from the received signal samples, it is desirable to separate the propagation delay $\tau_n(\mathbf{p})$ and transmit time t_0 from the signal waveform. This is easy when using the frequency-domain representation of the problem. Taking the discrete Fourier transform (DFT) of (2), we get

$$\bar{\mathbf{x}}_{n,k} = \beta_n \mathbf{a}_n(\mathbf{p}, \boldsymbol{\mu}_n) \bar{s}_k \cdot \exp\{-j\omega_k(\tau_n(\mathbf{p}) + t_0)\} + \bar{\boldsymbol{\varepsilon}}_{n,k}$$
$$(1 \le n \le N \; ; \; 1 \le k \le K) \tag{3}$$

where

- $\omega_k = 2\pi(k-1)/(KT)$ is the kth known discrete frequency point,
- $\bar{s}_k$ is the kth Fourier coefficient of the unknown signal corresponding to frequency ω_k,
- $\bar{\boldsymbol{\varepsilon}}_{n,k}$ is the kth Fourier coefficient of the random noise corresponding to frequency ω_k.

As the DFT is an orthogonal linear transformation, the probability distribution of the noise vector $\bar{\boldsymbol{\varepsilon}}_{n,k}$ is the same as that of $\boldsymbol{\varepsilon}_{n,\ k}$, with first- and second-order moments given by

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 4 of 28

$$\begin{cases} E\left[\bar{\boldsymbol{\varepsilon}}_{n,k}\right] = \mathbf{O}_{M \times 1} \\ E\left[\bar{\boldsymbol{\varepsilon}}_{n,k} \bar{\boldsymbol{\varepsilon}}_{n,k}^{\mathrm{T}}\right] = \mathbf{O}_{M \times M}, \quad E\left[\bar{\boldsymbol{\varepsilon}}_{n,k} \bar{\boldsymbol{\varepsilon}}_{n,k}^{\mathrm{H}}\right] = \sigma_{\varepsilon}^2 \mathbf{I}_M \quad (1 \leq k, \ l \leq K; \ k \neq l) \\ E\left[\bar{\boldsymbol{\varepsilon}}_{n,k} \bar{\boldsymbol{\varepsilon}}_{n,l}^{\mathrm{T}}\right] = E\left[\bar{\boldsymbol{\varepsilon}}_{n,k} \bar{\boldsymbol{\varepsilon}}_{n,l}^{\mathrm{H}}\right] = \mathbf{O}_{M \times M} \end{cases}$$

$$(4)$$

Note that the robust DPD methods proposed in this paper are derived from the frequency-domain signal model (3). For notational convenience, we introduce the following three vectors

$$\begin{cases} \bar{\mathbf{x}}_n = \left[\bar{\mathbf{x}}_{n,1}^{\mathrm{H}} \quad \bar{\mathbf{x}}_{n,2}^{\mathrm{H}} \quad \cdots \quad \bar{\mathbf{x}}_{n,K}^{\mathrm{H}}\right]^{\mathrm{H}}, \quad \boldsymbol{\beta} = \left[\boldsymbol{\beta}_1 \quad \boldsymbol{\beta}_2 \quad \cdots \quad \boldsymbol{\beta}_N\right]^{\mathrm{T}} \\ \bar{\mathbf{s}} = \left[\bar{s}_1 \cdot \exp\{-j\omega_1 t_0\} \quad \bar{s}_2 \cdot \exp\{-j\omega_2 t_0\} \quad \cdots \quad \bar{s}_K \cdot \exp\{-j\omega_K t_0\}\right]^{\mathrm{T}} \end{cases}$$

$$(5)$$

3.2 Array error model

This subsection describes the array error model. According to the discussion in [47, 48, 50], $\boldsymbol{\mu}_n$ can be modeled by considering it as a real random vector and is composed of the array parameters that are subject to perturbations. A nominal value of $\boldsymbol{\mu}_n$, denoted by $\boldsymbol{\mu}_{n,\,0}$, is known, resulting in a nominal array manifold $\mathbf{a}_n(\mathbf{p}, \boldsymbol{\mu}_{n,\,0})$. As a consequence, the actual array response may be different. We denote the length of $\boldsymbol{\mu}_n$ as Q_n.

It is known that $\boldsymbol{\mu}_n$ may include either structured parameters, such as the sensor gain, phase, position, and/or mutual coupling, or unstructured parameters. Generally, the perturbation parameter vector $\boldsymbol{\mu}_n$ can be modeled as a Gaussian random variable with mean $E[\boldsymbol{\mu}_n] = \boldsymbol{\mu}_{n,\,0}$ and covariance

$$E\left[\left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)\left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)^{\mathrm{T}}\right] = \boldsymbol{\Omega}_n \quad (1 \leq n \leq N) \quad (6)$$

Moreover, $\boldsymbol{\mu}_{n_1}$ and $\boldsymbol{\mu}_{n_2}$ are statistically independent for $n_1 \neq n_2$. Note that the covariance matrices $\{\boldsymbol{\Omega}_n\}_{1 \leq n \leq N}$ are also assumed to be known. These matrices can be determined, for example, using sample statistics from a number of independent, identical calibration experiments or using tolerance data specified by the sensor manufacturer [51].

In the next sections, the a priori information in the form of a probability distribution is used to develop two robust DPD methods with regard to array model mismatch. The first DPD method covers the common case of signals with unknown waveforms, and the second one is applicable to the less common case of signals with known waveforms. For notational simplicity, we define the following two vectors

$$\begin{cases} \boldsymbol{\mu}_0 = \left[\boldsymbol{\mu}_{1,0}^{\mathrm{H}} \quad \boldsymbol{\mu}_{2,0}^{\mathrm{H}} \quad \cdots \quad \boldsymbol{\mu}_{N,0}^{\mathrm{H}}\right]^{\mathrm{H}} \\ \boldsymbol{\mu} = \left[\boldsymbol{\mu}_1^{\mathrm{H}} \quad \boldsymbol{\mu}_2^{\mathrm{H}} \quad \cdots \quad \boldsymbol{\mu}_N^{\mathrm{H}}\right]^{\mathrm{H}} \end{cases} \quad (7)$$

4 Robust direct position determination method for case of unknown signal waveform

In this section, we propose a new DPD method that is robust to array model errors and assumes that the signal waveforms are unknown. The optimal MAP estimator for the problem at hand is formulated, and a computationally efficient numerical algorithm is derived.

4.1 Optimization criterion for maximum a posteriori estimator

If the transmitted signals are unknown, the parameters to be determined involve $\mathbf{p}$, $\{\bar{s}_k\}_{1 \leq k \leq K}$, t_0, $\boldsymbol{\beta}$, and $\boldsymbol{\mu}$. From (5), both $\{\bar{s}_k\}_{1 \leq k \leq K}$ and t_0 are contained in the vector $\bar{\mathbf{s}}$. Hence, the estimation of $\{\bar{s}_k\}_{1 \leq k \leq K}$ and t_0 can be replaced by the estimation of $\bar{\mathbf{s}}$.

When deriving the MAP estimator, the a priori distribution of $\{\boldsymbol{\mu}_n\}_{1 \leq n \leq N}$ must be exploited. Following [47–50], the joint estimation of $\mathbf{p}$, $\bar{\mathbf{s}}$, $\boldsymbol{\beta}$, and $\boldsymbol{\mu}$ is then obtained as

$$\min_{\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu}} J_{\text{U-MAP}}(\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu}) = \min_{\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu}} \{J_{\text{U-ML}}(\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu}) \quad (8)$$

$$+ \frac{1}{2} \cdot \sum_{n=1}^{N} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)^{\mathrm{T}} \boldsymbol{\Omega}_n^{-1} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)\}$$

where

$$J_{\text{U-MAP}}(\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu}) = J_{\text{U-ML}}(\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu}) \quad (9)$$

$$+ \frac{1}{2} \cdot \sum_{n=1}^{N} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)^{\mathrm{T}} \boldsymbol{\Omega}_n^{-1} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)$$

Here, $J_{\text{U-ML}}(\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu})$ is the negative log-likelihood function, which is given by

$$J_{\text{U-ML}}(\mathbf{p}, \bar{\mathbf{s}}, \boldsymbol{\beta}, \boldsymbol{\mu}) = \frac{1}{\sigma_{\varepsilon}^2} \cdot \sum_{n=1}^{N} \sum_{k=1}^{K} \|\bar{\mathbf{x}}_{n,k}$$
$$- \beta_n \mathbf{a}_n(\mathbf{p}, \boldsymbol{\mu}_n) \bar{s}_k \cdot \exp\{-j\omega_k(\tau_n(\mathbf{p})$$
$$+ t_0)\}\|_2^2$$
$$= \frac{1}{\sigma_{\varepsilon}^2} \cdot \sum_{n=1}^{N} \left\|\bar{\mathbf{x}}_n - ((\bar{\mathbf{s}} \odot \boldsymbol{\gamma}_n(\mathbf{p})) \otimes \mathbf{a}_n(\mathbf{p}, \boldsymbol{\mu}_n)) \beta_n\right\|_2^2$$

$$(10)$$

where $\otimes$ denotes Kronecker product, and $\odot$ represents Schur product (element by element multiplication), and

$$\boldsymbol{\gamma}_n(\mathbf{p}) = [\ \exp\{-j\omega_1 \tau_n(\mathbf{p})\} \quad \exp\{-j\omega_2 \tau_n(\mathbf{p})\} \quad \cdots$$
$$\exp\{-j\omega_K \tau_n(\mathbf{p})\}]^{\mathrm{T}}$$

$$(11)$$

Setting $\boldsymbol{\Omega}_{n,0} = \boldsymbol{\Omega}_n / \sigma_{\varepsilon}^2$, the optimization criterion can be finally formulated as

Wang et al. EURASIP Journal on Advances in Signal Processing (2018) 2018:38

Page 5 of 28

$$\min_{\mathbf{p},\bar{\mathbf{s}},\boldsymbol{\beta},\boldsymbol{\mu}}\Big\{\sum_{n=1}^{N}\big\|\bar{\mathbf{x}}_n-((\bar{\mathbf{s}}\odot\boldsymbol{\gamma}_n(\mathbf{p}))\otimes\mathbf{a}_n(\mathbf{p},\boldsymbol{\mu}_n))\beta_n\big\|_2^2 \quad (12)$$
$$+\frac{1}{2}\cdot\sum_{n=1}^{N}\big(\boldsymbol{\mu}_n-\boldsymbol{\mu}_{n,0}\big)^{\mathrm{T}}\boldsymbol{\Omega}_{n,0}^{-1}\big(\boldsymbol{\mu}_n-\boldsymbol{\mu}_{n,0}\big)\Big\}$$

where we assume, without loss of generality, that the first element in $\bar{\mathbf{s}}$ is equal to one so that the solution is unique and unambiguous.

Obviously, (12) is a multidimensional nonlinear minimization problem, and no closed-form solution is available. In the next subsection, we present an efficient numerical algorithm to solve (12). For this purpose, the following matrices are introduced:

$$\begin{cases}\mathbf{A}(\mathbf{p},\boldsymbol{\mu})=\big[(\mathbf{A}_1(\mathbf{p},\boldsymbol{\mu}_1))^{\mathrm{H}} \quad (\mathbf{A}_2(\mathbf{p},\boldsymbol{\mu}_2))^{\mathrm{H}} \quad \cdots \quad (\mathbf{A}_N(\mathbf{p},\boldsymbol{\mu}_N))^{\mathrm{H}}\big]^{\mathrm{H}}\\ \mathbf{A}_n(\mathbf{p},\boldsymbol{\mu}_n)=\mathrm{blkdiag}\big[\mathbf{a}'_{n,1}(\mathbf{p},\boldsymbol{\mu}_n) \quad \mathbf{a}'_{n,2}(\mathbf{p},\boldsymbol{\mu}_n) \quad \cdots \quad \mathbf{a}'_{n,K}(\mathbf{p},\boldsymbol{\mu}_n)\big]\end{cases}$$
$$(1\le n\le N)$$
$$(13)$$

where

$$\mathbf{a}'_{n,k}(\mathbf{p},\boldsymbol{\mu}_n)=\mathbf{a}_n(\mathbf{p},\boldsymbol{\mu}_n)\cdot\exp\{-\mathrm{j}\omega_k\tau_n(\mathbf{p})\} \quad (14)$$

4.2 Numerical algorithm

Recalling (12), it is apparent that the unknown parameter vectors of $\mathbf{p}$, $\bar{\mathbf{s}}$, $\boldsymbol{\beta}$, and $\boldsymbol{\mu}$ cannot be completely separated. Hence, the straightforward minimization of (12) with respect to $\mathbf{p}$, $\bar{\mathbf{s}}$, $\boldsymbol{\beta}$, and $\boldsymbol{\mu}$ is rarely feasible. Because of the diversity of the unknowns, we derive an alternating minimization algorithm to identify the unknown parameters in an iterative manner. This numerical method has been successfully applied to a variety of signal processing issues [52, 53].

After careful analysis of the present problem, the unknown parameters can be grouped into two categories: one is composed of $\mathbf{p}$ and $\bar{\mathbf{s}}$, and the other comprises $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$. The two sets of variables are alternately optimized in succession. First, the minimization is performed with respect to $\mathbf{p}$ and $\bar{\mathbf{s}}$, while $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ are kept unchanged. Subsequently, we solve the problem for $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ while keeping $\mathbf{p}$ and $\bar{\mathbf{s}}$ constant. This procedure is repeated until the convergence criterion is satisfied. A detailed description of the two steps in each iteration is given in the following subsection.

4.2.1 Joint optimization of p and s̄

First, consider the joint optimization of $\mathbf{p}$ and $\bar{\mathbf{s}}$, with $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ given by $\hat{\boldsymbol{\beta}}^{(\mathrm{a})}$ and $\hat{\boldsymbol{\mu}}^{(\mathrm{a})}$, respectively. Fortunately, when $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ are fixed, the estimation of $\mathbf{p}$ and $\bar{\mathbf{s}}$ can be decoupled.

Using the result in [27], we have the following optimization problem for finding the position vector $\mathbf{p}$.

$$\max_{\mathbf{p}}f_1(\mathbf{p})=\max_{\mathbf{p}}\lambda_{\max}\Big\{\big(\mathbf{B}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)\big)^{\mathrm{H}}\mathbf{B}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)\Big\}$$
$$=\max_{\mathbf{p}}\lambda_{\max}\Big\{\mathbf{C}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)\Big\}$$
$$(15)$$

where $\lambda_{\max}\{\cdot\}$ denotes the maximal eigenvalue of its input matrix and

$$\begin{cases}\mathbf{C}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)=\big(\mathbf{B}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)\big)^{\mathrm{H}}\mathbf{B}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)\\ \mathbf{B}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)=\big(\mathbf{A}\big(\mathbf{p},\hat{\boldsymbol{\mu}}^{(\mathrm{a})}\big)\big)^{\mathrm{H}}\overline{\mathbf{X}}\end{cases} \quad (16)$$

with $\overline{\mathbf{X}}=\mathrm{blkdiag}[\bar{\mathbf{x}}_1 \quad \bar{\mathbf{x}}_2 \quad \cdots \quad \bar{\mathbf{x}}_N]$. Note that the cost function $f_1(\mathbf{p})$ in (15) is expressed as the maximal eigenvalue of some positive semidefinite matrix and, hence, the functional form of $f_1(\mathbf{p})$ with respect to $\mathbf{p}$ is not explicit. The most straightforward method for solving (15) may be a grid search, as recommended in [27]. However, this method is very computationally expensive when the area of interest is large and the grid step size is small. To avoid a multidimensional search, we derive an efficient Gauss-Newton algorithm, which has much faster convergence speed than the steepest ascent and steepest descent methods. For this purpose, we first introduce the following proposition, which is associated with the matrix eigenvalue perturbation result.

Proposition 1: *Let $\mathbf{Z}\in\mathbf{C}^{n\times n}$ be a positive semidefinite matrix with eigenvalues $\lambda_1\le\lambda_2\le\cdots\le\lambda_n$, associated with unit eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$, respectively. Moreover, λ_j differs from the other eigenvalues. Assume $\mathbf{Z}$ is corrupted by a Hermitian error matrix $\boldsymbol{\delta}\mathbf{Z}\in\mathbf{C}^{n\times n}$, and the corresponding perturbed matrix is denoted as $\hat{\mathbf{Z}}$, i.e., $\hat{\mathbf{Z}}=\mathbf{Z}+\boldsymbol{\delta}\mathbf{Z}\in\mathbf{C}^{n\times n}$. If the eigenvalues of matrix $\hat{\mathbf{Z}}$ are defined by $\hat{\lambda}_1\le\hat{\lambda}_2\le\cdots\le\hat{\lambda}_n$, then the relationship between $\hat{\lambda}_j$ and λ_j can be described by*

$$\lambda_j=\lambda_j+\mathbf{v}_j^{\mathrm{H}}\cdot\boldsymbol{\delta}\mathbf{Z}\cdot\mathbf{v}_j+\mathbf{v}_j^{\mathrm{H}}\cdot\boldsymbol{\delta}\mathbf{Z}\cdot\mathbf{V}_j\cdot\boldsymbol{\delta}\mathbf{Z}\cdot\mathbf{v}_j+o\big(\|\boldsymbol{\delta}\mathbf{Z}\|_{\mathrm{F}}^2\big)$$
$$(17)$$

where

$$\mathbf{V}_j=\sum_{\substack{i=1\\i\ne j}}^{n}\frac{\mathbf{v}_i\mathbf{v}_i^{\mathrm{H}}}{\lambda_j-\lambda_i} \quad (18)$$

The proof of Proposition 1 is given in Appendix 1. The result in Proposition 1 can be used to obtain

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 6 of 28

the second-order Taylor series expansion of the cost function $f_1(\mathbf{p})$, from which the gradient and Hessian matrix of $f_1(\mathbf{p})$ can be obtained.

Assume that vector $\hat{\mathbf{p}}$ belongs to some neighborhood of the true value $\mathbf{p}$. The second-order Taylor series expansion of $\mathbf{C}(\hat{\mathbf{p}}, \hat{\boldsymbol{\mu}}^{(a)})$ around $\mathbf{p}$ leads to

$$
\begin{aligned}
\mathbf{C}\left(\hat{\mathbf{p}}, \hat{\boldsymbol{\mu}}^{(a)}\right) = & \mathbf{C}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) + \sum_{l=1}^{L} <\boldsymbol{\delta}\mathbf{p}>_l \cdot \dot{\mathbf{C}}_l\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \frac{1}{2} \cdot \sum_{l_1=1}^{L}\sum_{l_2=1}^{L} <\boldsymbol{\delta}\mathbf{p}>_{l_1} \cdot <\boldsymbol{\delta}\mathbf{p}>_{l_2} \cdot \ddot{\mathbf{C}}_{l_1 l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + o\left(\|\boldsymbol{\delta}\mathbf{p}\|_2^2\right)
\end{aligned}
\tag{19}
$$

where

$$
\begin{aligned}
\dot{\mathbf{C}}_l\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) = & \frac{\partial \mathbf{C}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)}{\partial <\mathbf{p}>_l} = \left(\dot{\mathbf{B}}_l\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \mathbf{B}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \left(\mathbf{B}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \dot{\mathbf{B}}_l\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
\ddot{\mathbf{C}}_{l_1 l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) = & \frac{\partial^2 \mathbf{C}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)}{\partial <\mathbf{p}>_{l_1}\partial <\mathbf{p}>_{l_2}} = \left(\ddot{\mathbf{B}}_{l_1 l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \mathbf{B}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \left(\mathbf{B}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \ddot{\mathbf{B}}_{l_1 l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \left(\dot{\mathbf{B}}_{l_1}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \dot{\mathbf{B}}_{l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \left(\dot{\mathbf{B}}_{l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \dot{\mathbf{B}}_{l_1}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)
\end{aligned}
\tag{21}
$$

with $\dot{\mathbf{B}}_l(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}) = \frac{\partial \mathbf{B}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})}{\partial <\mathbf{p}>_l}$ and $\ddot{\mathbf{B}}_{l_1 l_2}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}) = \frac{\partial^2 \mathbf{B}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})}{\partial <\mathbf{p}>_{l_1}\partial <\mathbf{p}>_{l_2}}$. Let $\lambda_1(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}) \le \lambda_2(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}) \le \cdots \le \lambda_N(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})$ and $\mathbf{v}_1(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})$, $\mathbf{v}_2(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})$, $\cdots$, $\mathbf{v}_N(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})$ be the eigenvalues and relevant unit eigenvectors of matrix $\mathbf{C}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})$, respectively. Combining (15), (17), and (19) leads to

$$
\begin{aligned}
f_1(\hat{\mathbf{p}}) = & \lambda_{\max}\left\{\mathbf{C}\left(\hat{\mathbf{p}}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right\} = \lambda_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \sum_{l=1}^{L} <\boldsymbol{\delta}\mathbf{p}>_l \cdot \left(\mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \dot{\mathbf{C}}_l\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \frac{1}{2} \cdot \sum_{l_1=1}^{L}\sum_{l_2=1}^{L} <\boldsymbol{\delta}\mathbf{p}>_{l_1} \cdot <\boldsymbol{\delta}\mathbf{p}>_{l_2} \\
& \cdot \left(\mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} (2\dot{\mathbf{C}}_{l_1}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{V}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \dot{\mathbf{C}}_{l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& + \ddot{\mathbf{C}}_{l_1 l_2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) + o\left(\|\boldsymbol{\delta}\mathbf{p}\|_2^2\right)
\end{aligned}
\tag{22}
$$

where

$$
\mathbf{V}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) = \sum_{n=1}^{N-1} \frac{\mathbf{v}_n\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\left(\mathbf{v}_n\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}}}{\lambda_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) - \lambda_n\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)}
\tag{23}
$$

Define the following vector and matrices

$$
\begin{cases}
\mathbf{h}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) = \begin{bmatrix} \left(\mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \dot{\mathbf{C}}_1\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\ \left(\mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \dot{\mathbf{C}}_2\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\ \vdots \\ \left(\mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}} \dot{\mathbf{C}}_L\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \end{bmatrix} \\
\mathbf{H}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) = \mathbf{H}_1\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) + \mathbf{H}_2\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)
\end{cases}
\tag{24}
$$

where

$$
\begin{aligned}
\mathbf{H}_1\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) = & \left(\mathbf{I}_L \otimes \left(\mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{\mathrm{H}}\right) \begin{bmatrix} \ddot{\mathbf{C}}_{11}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) & \ddot{\mathbf{C}}_{12}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) & \cdots & \ddot{\mathbf{C}}_{1L}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\ \ddot{\mathbf{C}}_{21}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) & \ddot{\mathbf{C}}_{22}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) & \cdots & \ddot{\mathbf{C}}_{2L}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \ddot{\mathbf{C}}_{L1}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) & \ddot{\mathbf{C}}_{L2}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) & \cdots & \ddot{\mathbf{C}}_{LL}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \end{bmatrix} \cdot \left(\mathbf{I}_L \otimes \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right) \\
\mathbf{H}_2\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) = & 2 \cdot \left[\dot{\mathbf{C}}_1\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mid \dot{\mathbf{C}}_2\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mid \cdots \mid \dot{\mathbf{C}}_L\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right]^{\mathrm{H}} \cdot \mathbf{V}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \\
& \times \left[\dot{\mathbf{C}}_1\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mid \dot{\mathbf{C}}_2\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mid \cdots \mid \dot{\mathbf{C}}_L\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \mathbf{v}_N\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right]
\end{aligned}
\tag{25}
$$

Then, $f_1(\hat{\mathbf{p}})$ in (22) can be rephrased as

$$
\begin{aligned}
f_1(\hat{\mathbf{p}}) = & f_1(\mathbf{p}) + \boldsymbol{\delta}\mathbf{p}^{\mathrm{T}} \cdot \mathbf{h}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) + \frac{1}{2} \cdot \boldsymbol{\delta}\mathbf{p}^{\mathrm{T}} \cdot \mathbf{H}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right) \cdot \boldsymbol{\delta}\mathbf{p} \\
& + o\left(\|\boldsymbol{\delta}\mathbf{p}\|_2^2\right) = f_1(\mathbf{p}) + \boldsymbol{\delta}\mathbf{p}^{\mathrm{T}} \cdot \mathrm{Re}\left\{\mathbf{h}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right\} \\
& + \frac{1}{2} \cdot \boldsymbol{\delta}\mathbf{p}^{\mathrm{T}} \cdot \mathrm{Re}\left\{\mathbf{H}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right\} \cdot \boldsymbol{\delta}\mathbf{p} + o\left(\|\boldsymbol{\delta}\mathbf{p}\|_2^2\right)
\end{aligned}
\tag{26}
$$

where the second equality follows from the fact that $f_1(\cdot)$ is a real function and $\boldsymbol{\delta}\mathbf{p}$ is a real vector. From (26), $\mathrm{Re}\{\mathbf{h}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})\}$ and $\mathrm{Re}\{\mathbf{H}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})\}$ are the gradient and Hessian matrix of $f_1(\mathbf{p})$, respectively. Consequently, the Gauss-Newton algorithm for solving (15) is given by

$$
\begin{aligned}
\mathbf{p}^{(i+1)} = & \hat{\mathbf{p}}^i - \alpha^i \left(\mathrm{Re}\left\{\mathbf{H}\left(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right\}\right)^{-1} \\
& \cdot \mathrm{Re}\left\{\mathbf{h}\left(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right\}
\end{aligned}
\tag{27}
$$

where $\alpha \in (0, 1)$ denotes the step size, and superscript i indexes the ith iteration.

Before proceeding, some remarks are concluded.

Remark 1: Clearly, (15) is a nonlinear least-squares optimization problem, so it can be effectively solved by the Gauss-Newton algorithm.

Remark 2: Note that the imaginary parts of $\mathbf{h}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})$ and $\mathbf{H}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})$ are ignored in (26) and (27). However,

this does not lead to any information loss or performance degradation. The vanishing term in the second equality in (26) is $\boldsymbol{\delta p}^T \cdot \text{Im}\{\mathbf{h}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})\} + \boldsymbol{\delta p}^T \cdot \text{Im}\{\mathbf{H}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})\} \cdot \boldsymbol{\delta p}/2$, which is equal to zero because $f_1(\hat{\mathbf{p}})$, $f_1(\mathbf{p})$, and $\boldsymbol{\delta p}$ are real scalars or vectors. As a result, $\text{Re}\{\mathbf{h}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})\}$ and $\text{Re}\{\mathbf{H}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(a)})\}$ are the true gradient and Hessian matrix of $f_1(\mathbf{p})$, respectively, and no approximation is necessary. Hence, (27) can be viewed as a standard Gauss-Newton algorithm.

Remark 3: It is worth noting that apart from Gauss-Newton algorithm, there are many other local search algorithms, such as Newton algorithm, steepest descent algorithm, Levenberg-Marquardt algorithm, etc. However, Newton algorithm is much more computationally demanding because it requires computing the second-order derivative of the cost function; steepest descent algorithm has slower convergence rate than the other algorithms; Levenberg-Marquardt algorithm needs to introduce a reasonable damping factor, which is difficult to determine. For these reasons, we use Gauss-Newton algorithm to solve (15). In our simulation performed in Section 7, this algorithm can provide satisfactory results. Besides, it must be emphasized that all these algorithms can only guarantee local convergence; that is to say, it is not guaranteed to obtain the global optimal solution for these algorithms.

Remark 4: In the simulations described in Section 7, the step size α is set to 0.85. A number of simulation results indicate that this value for α provides good estimates of the source position.

Assuming that the convergence result of (27) is $\hat{\mathbf{p}}^{(a)}$, the derivation in [27] implies that the optimal solution of $\overline{\mathbf{s}}$ is given by

$$\widehat{\overline{\mathbf{s}}}^{(a)} = \mathbf{v}_{\max}\left\{ \mathbf{B}\left(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}^{(a)}\right)\left(\mathbf{B}\left(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}^{(a)}\right)\right)^{H} \right\} \tag{28}$$

where $\mathbf{v}_{\max}\{\cdot\}$ denotes the eigenvector (with first element equal to one) corresponding to the largest eigenvalue of its matrix argument.

4.2.2 Joint optimization of β and μ
This subsection describes the minimization of the criterion in (12) with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$, while $\mathbf{p}$ and $\overline{\mathbf{s}}$ are fixed at $\hat{\mathbf{p}}^{(a)}$ and $\widehat{\overline{\mathbf{s}}}^{(a)}$, respectively.

Likewise, the estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ can also be decoupled. First, the channel attenuation scalar $\{\beta_n\}_{1 \le n \le N}$ that minimizes (12) is given by

$$\hat{\beta}_{n,\text{opt}} = \frac{1}{\left\| \mathbf{a}_n\left(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n\right) \right\|_2^2 \cdot \left\| \widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(a)}\right) \right\|_2^2} \\ \cdot \left(\left(\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(a)}\right)\right) \otimes \mathbf{a}_n\left(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n\right)\right)^{H} \overline{\mathbf{x}}_n \quad (29) \\ (1 \le n \le N)$$

Inserting (29) back into (12) we get the following concentrated problem

$$\min_{\boldsymbol{\mu}} f_2(\boldsymbol{\mu}) = \min_{\boldsymbol{\mu}}\left\{ \sum_{n=1}^{N}\left\| \overline{\mathbf{x}}_n - \frac{\left(\left(\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})\right) \otimes \mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n)\right)\left(\left(\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})\right) \otimes \mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n)\right)^{H}}{\|\mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n)\|_2^2 \cdot \|\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})\|_2^2} \cdot \overline{\mathbf{x}}_n \right\|_2^2 \right. \\ \left. + \frac{1}{2}\sum_{n=1}^{N}(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0})^T \boldsymbol{\Omega}_{n,0}^{-1}(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}) \right\} \\ = \min_{\boldsymbol{\mu}}\left\{ \sum_{n=1}^{N}\left\| \begin{bmatrix} \overline{\mathbf{x}}_n \\ \frac{1}{\sqrt{2}}\cdot\boldsymbol{\Omega}_{n,0}^{-1/2}\boldsymbol{\mu}_{n,0} \end{bmatrix} - \begin{bmatrix} \frac{\left(\left(\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})\right) \otimes \mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n)\right)\left(\left(\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})\right) \otimes \mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n)\right)^{H}}{\|\mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n)\|_2^2 \cdot \|\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})\|_2^2} \cdot \overline{\mathbf{x}}_n \\ \frac{1}{\sqrt{2}}\cdot\boldsymbol{\Omega}_{n,0}^{-1/2}\boldsymbol{\mu}_n \end{bmatrix} \right\|_2^2 \right\} \tag{30}$$

As each term in the sum only depends on the parameter of one array (i.e., $\boldsymbol{\mu}_n$), the minimization in (30) can be performed by solving the following N subproblems:

$$\min_{\boldsymbol{\mu}_n} f_{2,n}(\boldsymbol{\mu}_n) = \min_{\boldsymbol{\mu}_n}\left\{ \left\| \begin{bmatrix} \overline{\mathbf{x}}_n \\ \frac{1}{\sqrt{2}}\cdot\boldsymbol{\Omega}_{n,0}^{-1/2}\boldsymbol{\mu}_{n,0} \end{bmatrix} - \mathbf{g}_n(\boldsymbol{\mu}_n) \right\|_2^2 \right\} \quad (1 \le n \le N) \tag{31}$$

where

$$\mathbf{g}_n(\boldsymbol{\mu}_n) = \begin{bmatrix} \frac{((\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})) \otimes \mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n))((\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})) \otimes \mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n))^{H}}{\|\mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \boldsymbol{\mu}_n)\|_2^2 \cdot \|\widehat{\overline{\mathbf{s}}}^{(a)} \odot \boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})\|_2^2} \cdot \overline{\mathbf{x}}_n \\ \frac{1}{\sqrt{2}}\cdot\boldsymbol{\Omega}_{n,0}^{1/2}\boldsymbol{\mu}_n \end{bmatrix} \tag{32}$$

Obviously, (31) is a nonlinear least-squares optimization problem, which can be efficiently solved by the Gauss-Newton algorithm. The corresponding iterative formula is given by

$$\hat{\boldsymbol{\mu}}_n^{(i+1)} = \hat{\boldsymbol{\mu}}_n^{(i)} + (\text{Re}\{(\mathbf{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)}))^H \mathbf{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\})^{-1} \cdot \text{Re}\left\{ (\mathbf{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)}))^H \left(\begin{bmatrix} \overline{\mathbf{x}}_n \\ \frac{1}{\sqrt{2}}\cdot\boldsymbol{\Omega}_{n,0}^{-1/2}\boldsymbol{\mu}_{n,0} \end{bmatrix} - \mathbf{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)}) \right) \right\} \tag{33}$$

where i is the iteration number and $\mathbf{G}_n(\boldsymbol{\mu}_n)$ is the Jacobian matrix of $\mathbf{g}_n(\boldsymbol{\mu}_n)$, i.e.,

$$\mathbf{G}_n(\boldsymbol{\mu}_n) = \frac{\partial \mathbf{g}_n(\boldsymbol{\mu}_n)}{\partial \boldsymbol{\mu}_n^T} = \begin{bmatrix} \mathbf{G}_{n,1}(\boldsymbol{\mu}_n) \\ \mathbf{G}_{n,2}(\boldsymbol{\mu}_n) \end{bmatrix} \tag{34}$$

From (32), it can be checked that

$$
\begin{aligned}
\mathbf{G}_{n,1}(\boldsymbol{\mu}_n) =& \frac{\left(\left(\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right)\otimes\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right)^{\mathrm{H}}\overline{\mathbf{x}}_n}{\left\|\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right\|_2^2\cdot\left\|\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right\|_2^2} \\
&\cdot\left(\left(\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right)\otimes\frac{\partial\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)}{\partial\boldsymbol{\mu}_n^{\mathrm{T}}}\right) \\
&+\frac{\left(\left(\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right)\otimes\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right)\overline{\mathbf{x}}_n^{\mathrm{T}}}{\left\|\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right\|_2^2\cdot\left\|\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right\|_2^2} \\
&\cdot\left(\left(\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right)^*\otimes\frac{\partial\left(\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right)^*}{\partial\boldsymbol{\mu}_n^{\mathrm{T}}}\right) \\
&-\frac{2\left(\left(\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right)\otimes\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right)^{\mathrm{H}}\overline{\mathbf{x}}_n}{\left\|\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right\|_2^4\cdot\left\|\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right\|_2^2} \\
&\cdot\left(\left(\widehat{\overline{\mathbf{s}}}^{(\mathrm{a})}\odot\boldsymbol{\gamma}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})}\right)\right)\otimes\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right)\left(\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)\right)^{\mathrm{H}} \\
&\cdot\frac{\partial\mathbf{a}_n\left(\hat{\mathbf{p}}^{(\mathrm{a})},\boldsymbol{\mu}_n\right)}{\partial\boldsymbol{\mu}_n^{\mathrm{T}}}
\end{aligned}
\tag{35}
$$

$$
\mathbf{G}_{n,2}(\boldsymbol{\mu}_n)=\frac{1}{\sqrt{2}}\cdot\boldsymbol{\Omega}_{n,0}^{-1/2}
\tag{36}
$$

We conclude some remarks as follows.

Remark 5: Similarly, as (31) is also a nonlinear least-squares optimization problem, we can use the Gauss-Newton algorithm to find its optimal solution.

Remark 6: The imaginary parts of $\left(\mathbf{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\right)^{\mathrm{H}}\mathbf{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})$ and $\left(\mathbf{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\right)^{\mathrm{H}}\left(\left[\begin{matrix}\overline{\mathbf{x}}_n\\ \frac{1}{\sqrt{2}}\cdot\boldsymbol{\Omega}_{n,0}^{-1/2}\boldsymbol{\mu}_{n,0}\end{matrix}\right]-\mathbf{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\right)$ are neglected in (33). Nevertheless, this does not mean that there are any approximations in (33), because $\boldsymbol{\mu}_n$ is a real vector. The detailed derivation of (33) is given in Appendix 2.

Assuming that the convergence result of (33) is $\hat{\boldsymbol{\mu}}_n^{(\mathrm{a})}$, substituting this back into (29) leads to the estimation of $\boldsymbol{\beta}$, which is denoted by $\hat{\boldsymbol{\beta}}^{(\mathrm{a})}$.

4.2.3 Summary of the alternating minimization algorithm

The ingredients of the previous two subsections can be combined to form the proposed alternating minimization algorithm, which is summarized as follows.

Proposed alternating minimization algorithm I

Step 1: Define a convergence threshold $\delta > 0$ and choose the initial values $\hat{\mathbf{p}}^{(0)}$, $\hat{\mathbf{s}}^{(0)}$, $\hat{\boldsymbol{\mu}}^{(0)}$ and $\hat{\boldsymbol{\beta}}^{(0)}$.

Step 2: Set the iteration counter $m = 0$ and compute the cost function $J_{\mathrm{U}}^{(m)}$ using (12).

Step 3: Calculate $\hat{\mathbf{p}}^{(m)}$ using the Gauss-Newton algorithm given in (27).

Step 4: Compute $\hat{\mathbf{s}}^{(m)}$ according to (28).

Step 5: Calculate $\hat{\boldsymbol{\mu}}^{(m)}$ from (33).

Step 6: Compute $\hat{\boldsymbol{\beta}}^{(m)}$ via (29).

Step 7: Increment the iteration counter $m = m + 1$ and compute the cost function $J_{\mathrm{U}}^{(m)}$ using (12). If $|J_{\mathrm{U}}^{(m)}-J_{\mathrm{U}}^{(m-1)}|\le\delta$, stop the procedure; otherwise go to Step 3.

The following remarks concern the alternating minimization algorithm described above.

Remark 7: The algorithm iterates until the convergence criterion is satisfied. Similar to the analysis in [54], it follows that, at each step, the cost function reduces so that $J_{\mathrm{U}}^{(1)} > J_{\mathrm{U}}^{(2)} > \cdots > J_{\mathrm{U}}^{(m)}\ge 0$. Therefore, $\{J_{\mathrm{U}}^{(m)}\}_{m=1}^{+\infty}$ is a convergent series and the convergence is ensured. Based on our experimental results, 20 iterations are sufficient to guarantee the convergence.

Remark 8: As in many iterative algorithms, convergence to the global optimum of the cost function depends on the initial estimate. If the initial values are properly chosen, the effects of these initial errors can largely be removed. For the problem at hand, the initial value of $\boldsymbol{\mu}$ can be chosen as its nominal value $\boldsymbol{\mu}_0$, and the starting point of $\mathbf{p}$ can be obtained using the traditional two-step localization method, where the DOA parameter can be estimated by subspace-based methods [1–3]. If the unknowns to be found are DOAs, the array manifold in (1)-(3) should be modeled as a function of direction, as in [1–3]. This is not difficult, because the position vector $\mathbf{p}$ is a function of direction. Once the DOAs relative to all arrays have been estimated, the initial source position can be easily determined using the closed-form method [4]. In addition, $\overline{\mathbf{s}}$ and $\boldsymbol{\beta}$ can be initialized from (28) and (29), respectively. The simulation results in Section 7 demonstrate that using these initial estimates gives satisfactory performance.

4.2.4 Complexity analysis

Here, the computational complexity of the proposed DPD method is studied in terms of the number of multiplication. Table 1 summarizes the numerical complexity.

5 Direct position determination method for case of known signal waveform

The aim of this section is to propose an alternative robust DPD method that accounts for the uncertainties in the

Wang et al. EURASIP Journal on Advances in Signal Processing (2018) 2018:38

Page 9 of 28

Table 1 Complexity of proposed method

Calculation module	Computational unit	Computational complexity of each unit	Total computational complexity
Gauss-Newton algorithm in (27)	$\mathbf{C}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)})$	$MN^2K^2 + N^2K$	$I_N \left(\begin{array}{l} MN^2K^2 + N^2K + N^2(N-1) \\ +N^2(L^2+2L) + N(L^2+3L) \\ +O(N^3) + O(L^3) \end{array} \right)$
	Eigen decomposition of $\mathbf{C}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)})$	$O(N^3)$	
	$\mathbf{V}_N(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)})$	$N^2(N-1)$	$+I_{G-N}N \left(\begin{array}{l} Q_n^2(MK+3) \\ +Q_n(M^2(K^2+K)+2MK) \\ +M^2(3K^2+K)+M(2K+1) \\ +2K+O(Q_n^3) \end{array} \right)$
	$\mathbf{H}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)})$	$N^2(L^2+L)$ $+N(L^2+2L)$	$+M(N^2K^2+N(2K+1))$ $+N(K^2+K)+O(K^3)$
	$\mathbf{h}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)})$	N^2L+NL	(where I_N, I_{G-N} and I_{Iter} denote the iteration numbers for Gauss-Newton algorithm and alternating minimization algorithm, respectively)
	$\alpha \left(\left(\mathrm{Re}\{\mathbf{H}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)})\}\right)^{-1} \cdot \mathrm{Re}\{\mathbf{h}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(a)})\}\right.$	$O(L^3)$	
Estimation of $\bar{\mathbf{s}}$	$\mathbf{B}(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}^{(a)})$	MN^2K^2	
	$\mathbf{B}(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}^{(a)})\left(\mathbf{B}(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}^{(a)})\right)^{\mathsf{H}}$	NK^2	
	Eigen decomposition of $\mathbf{B}(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}^{(a)})\left(\mathbf{B}(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}^{(a)})\right)^{\mathsf{H}}$	$O(K^3)$	
Gauss-Newton algorithm in (33)	$\mathbf{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)})$	$Q_n^2+2M^2K^2$ $+M(K+1)+2K$	
	$\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})$	$Q_n \left(\begin{array}{l} M^2(K^2+K) \\ +MK \end{array} \right)$	
	$\left(\mathrm{Re}\{(\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)}))^{\mathsf{H}} \boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\}\right)^{-1}$	$+M^2(K^2+K)+MK$	
	$\times \mathrm{Re}\left\{(\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)}))^{\mathsf{H}} \left(\left[\dfrac{1}{\sqrt{2}} \cdot \dfrac{\bar{\boldsymbol{x}}_n}{\boldsymbol{\Omega}_{n,0}^{-1/2}\,\boldsymbol{\mu}_{n,0}}\right] - \boldsymbol{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\right)\right\}$	$Q_n^2(MK+2)$ $+Q_nMK+O(Q_n^3)$	
Estimation of $\boldsymbol{\beta}$	$\dfrac{\left(\left(\tilde{\bar{\mathbf{s}}}^{(a)} \odot \gamma_n(\hat{\mathbf{p}}^{(a)})\right)\otimes \mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}_n^{(a)})\right)^{\mathsf{H}} \bar{\mathbf{x}}_n}{\|\mathbf{a}_n(\hat{\mathbf{p}}^{(a)}, \hat{\boldsymbol{\mu}}_n^{(a)})\|_2^2 \cdot \|\tilde{\bar{\mathbf{s}}}^{(a)} \odot \gamma_n(\hat{\mathbf{p}}^{(a)})\|_2}$	$M(2K+1)+K$	

Wang et al. EURASIP Journal on Advances in Signal Processing (2018) 2018:38

Page 10 of 28

array manifold for the case of known signal waveforms. Applications of this study can be found in, for example, wireless communications, where some known preamble sequences are often transmitted for training or synchronization purposes [55, 56]. Similar to Section 4, the robust DPD algorithm developed here is also based on an optimal MAP estimator for the current problem.

5.1 Optimization criterion for maximum a posteriori estimator

If the transmitted signals are a priori known, the sole unknown parameter embedded in $\bar{\mathbf{s}}$ is t_0. Consequently, we can consider the estimation of t_0 instead of $\bar{\mathbf{s}}$. Using the a priori distribution of $\{\boldsymbol{\mu}_n\}_{1 \le n \le N}$, the MAP estimator with respect to $\mathbf{p}$, t_0, $\boldsymbol{\beta}$, and $\boldsymbol{\mu}$ can be written as

$$\min_{\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu}} J_{\text{K-MAP}}(\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu}) = \min_{\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu}} \{J_{\text{K-ML}}(\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu})$$
$$+ \frac{1}{2} \cdot \sum_{n=1}^{N} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)^{\mathrm{T}} \boldsymbol{\Omega}_n^{-1} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)\}$$

(37)

where

$$J_{\text{K-MAP}}(\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu}) = J_{\text{K-ML}}(\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu}) + \frac{1}{2}$$
$$\cdot \sum_{n=1}^{N} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)^{\mathrm{T}} \boldsymbol{\Omega}_n^{-1} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)$$

(38)

Here, $J_{\text{K-ML}}(\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu})$ is the negative log-likelihood function for the case of known waveform, which can be formulated as

$$J_{\text{K-ML}}(\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu}) = \frac{1}{\sigma_\varepsilon^2} \cdot \sum_{n=1}^{N} \sum_{k=1}^{K} \|\bar{\mathbf{x}}_{n,k} - \beta_n \mathbf{a}_n(\mathbf{p},\boldsymbol{\mu}_n)\bar{s}_k$$
$$\cdot \exp\{-j\omega_k(\tau_n(\mathbf{p}) + t_0)\}\|_2^2$$
$$= \frac{1}{\sigma_\varepsilon^2} \cdot \sum_{n=1}^{N} \|\bar{\mathbf{x}}_n - ((\bar{\mathbf{s}}' \odot \boldsymbol{\phi}_n(\mathbf{p},t_0)) \otimes \mathbf{a}_n(\mathbf{p},\boldsymbol{\mu}_n))\beta_n\|_2^2$$

(39)

where

$$\begin{cases} \bar{\mathbf{s}}' = [\bar{s}_1 & \bar{s}_2 & \cdots & \bar{s}_K]^{\mathrm{T}} \\ \boldsymbol{\phi}_n(\mathbf{p},t_0) = [\exp\{-j\omega_1(\tau_n(\mathbf{p})+t_0)\} & \exp\{-j\omega_2(\tau_n(\mathbf{p})+t_0)\} \\ & \cdots & \exp\{-j\omega_K(\tau_n(\mathbf{p})+t_0)\}]^{\mathrm{T}} \end{cases}$$

(40)

Note that the vector $\bar{\mathbf{s}}'$ is known here. Combining (37)–(39), we obtain the following optimization problem:

$$\min_{\mathbf{p},t_0,\boldsymbol{\beta},\boldsymbol{\mu}} \left\{ \sum_{n=1}^{N} \|\bar{\mathbf{x}}_n - ((\bar{\mathbf{s}}' \odot \boldsymbol{\phi}_n(\mathbf{p},t_0)) \otimes \mathbf{a}_n(\mathbf{p},\boldsymbol{\mu}_n))\beta_n\|_2^2 \right.$$
$$\left. + \frac{1}{2} \cdot \sum_{n=1}^{N} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right)^{\mathrm{T}} \boldsymbol{\Omega}_{n,0}^{-1} \left(\boldsymbol{\mu}_n - \boldsymbol{\mu}_{n,0}\right) \right\}$$

(41)

It does not seem possible to solve (41) in closed form because of its nonlinear nature. In the next subsection, an efficient numerical algorithm for solving this problem is presented.

5.2 Numerical algorithm

The algorithm proposed in this subsection is similar to that in Subsection 4.2, and is also implemented by adopting the alternating minimization algorithm. After a close inspection of (41), we can divide the unknown parameters into two categories: one is composed of $\mathbf{p}$ and t_0, and the other comprises $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$. The two sets of variables are alternately updated by separate optimization procedures. A detailed description of the proposed algorithm is given in the following subsections.

5.2.1 Joint optimization of p and t_0

We perform joint optimization with respect to $\mathbf{p}$ and t_0, with $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ fixed to $\hat{\boldsymbol{\beta}}^{(b)}$ and $\hat{\boldsymbol{\mu}}^{(b)}$, respectively.

According to the derivation in [27], the joint estimates of $\mathbf{p}$ and t_0 can be obtained by solving the following optimization problem:

$$\max_{\mathbf{p},t_0} f_3(\mathbf{p},t_0) = \max_{\mathbf{p},t_0} \left\| (\mathbf{B}(\mathbf{p},\hat{\boldsymbol{\mu}}^{(b)}))^{\mathrm{H}} \bar{\mathbf{S}}' \mathbf{d}(t_0) \right\|_2^2 \quad (42)$$

where

$$\begin{cases} \bar{\mathbf{S}}' = \text{diag}[\bar{\mathbf{s}}'] = \text{diag}[\bar{s}_1 & \bar{s}_2 & \cdots & \bar{s}_K] \\ \mathbf{d}(t_0) = [\exp\{-j\omega_1 t_0\} & \exp\{-j\omega_2 t_0\} & \cdots & \exp\{-j\omega_K t_0\}]^{\mathrm{T}} \end{cases}$$

(43)

We define a permutation matrix $\boldsymbol{\Pi}$ subject to

$$\text{vec}\left(\left(\mathbf{B}\left(\mathbf{p},\hat{\boldsymbol{\mu}}^{(b)}\right)\right)^{\mathrm{H}}\right) = \boldsymbol{\Pi} \cdot \text{vec}\left(\mathbf{B}\left(\mathbf{p},\hat{\boldsymbol{\mu}}^{(b)}\right)\right) \quad (44)$$

Using the identity $\text{vec}(\mathbf{XYZ}) = (\mathbf{Z}^{\mathrm{T}} \otimes \mathbf{X}) \cdot \text{vec}(\mathbf{Y})$, (42) can then be written as

$$\max_{\mathbf{p},t_0} f_3(\mathbf{p},t_0) = \max_{\mathbf{p},t_0} \left\| \left(\left(\bar{\mathbf{S}}' \mathbf{d}(t_0)\right)^{\mathrm{T}} \otimes \mathbf{I}_N \right) \boldsymbol{\Pi} \mathbf{b}(\mathbf{p},\hat{\boldsymbol{\mu}}^{(b)}) \right\|_2^2$$

(45)

where $\mathbf{b}(\mathbf{p},\hat{\boldsymbol{\mu}}^{(b)}) = \text{vec}(\mathbf{B}(\mathbf{p},\hat{\boldsymbol{\mu}}^{(b)}))$. Note that, instead of finding $\mathbf{p}$ and t_0 jointly, we can estimate them separately in sequence to reduce the computational complexity.

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 11 of 28

First, we consider t_0 as a nuisance parameter and rewrite (45) as an optimization problem for the position vector $\mathbf{p}$ as follows:

$$\max_{\mathbf{p}} f_4(\mathbf{p}) = \max_{\mathbf{p}} \left\{ \max_{t_0} \left\| \left(\left(\overline{\mathbf{S}}' \mathbf{d}(t_0) \right)^{\mathrm{T}} \otimes \mathbf{I}_N \right) \mathbf{\Pi} \mathbf{b}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}) \right\|_2^2 \right\}$$
(46)

where

$$f_4(\mathbf{p}) = \max_{t_0} \left\| \left(\left(\overline{\mathbf{S}}' \mathbf{d}(t_0) \right)^{\mathrm{T}} \otimes \mathbf{I}_N \right) \mathbf{\Pi} \mathbf{b}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}) \right\|_2^2$$
(47)

In (47), the functional form of $f_4(\mathbf{p})$ is not explicit. Therefore, the most straightforward method for solving (46) may be a grid search. However, this is computationally intensive, as stated previously. Thus, we develop an iterative algorithm for solving (44) based on the Gauss-Newton algorithm.

To obtain an iterative procedure, we need to derive the gradient and Hessian matrix of the cost function $f_4(\mathbf{p})$. For this purpose, a preliminary mathematical result is presented as follows.

Proposition 2: *Consider a twice continuously differentiable function $f(\mathbf{z}, y)$ that depends on the independent variables $y \in \mathbf{R}$ and $\mathbf{z} \in \mathbf{R}^{m \times 1}$. Define a vector function $f_{\mathrm{m}}(\mathbf{z}) = \max_{y}\{f(\mathbf{z}, y)\}$ and assume that, for a given $\mathbf{z}$, the extreme point that maximizes $f(\mathbf{z}, y)$ with respect to y is denoted as $y_{\mathrm{m}}(\mathbf{z})$. The gradient and Hessian matrix of $f_{\mathrm{m}}(\mathbf{z})$ are then given by*

$$\begin{cases} \mathbf{f}_{\mathrm{m}}(\mathbf{z}) = \dfrac{\partial f_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} = \dot{\mathbf{f}}_1(\mathbf{z}, y_{\mathrm{m}}(\mathbf{z})) \\[2mm] \mathbf{F}_{\mathrm{m}}(\mathbf{z}) = \dfrac{\partial^2 f_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z} \partial \mathbf{z}^{\mathrm{T}}} = \ddot{\mathbf{F}}_{11}(\mathbf{z}, y_{\mathrm{m}}(\mathbf{z})) - \dfrac{\ddot{\mathbf{f}}_{21}(\mathbf{z}, y_{\mathrm{m}}(\mathbf{z})) \ddot{\mathbf{f}}_{21}^{\mathrm{T}}(\mathbf{z}, y_{\mathrm{m}}(\mathbf{z}))}{\ddot{f}_{22}(\mathbf{z}, y_{\mathrm{m}}(\mathbf{z}))} \end{cases}$$
(48)

where

$$\begin{cases} \dot{f}_2(\mathbf{z}, y) = \dfrac{\partial f(\mathbf{z}, y)}{\partial y} \quad , \quad \ddot{f}_{22}(\mathbf{z}, y) = \dfrac{\partial^2 f(\mathbf{z}, y)}{\partial y^2} \\[2mm] \dot{\mathbf{f}}_1(\mathbf{z}, y) = \dfrac{\partial f(\mathbf{z}, y)}{\partial \mathbf{z}} \quad , \quad \ddot{\mathbf{f}}_{21}(\mathbf{z}, y) = \dfrac{\partial \dot{f}_2(\mathbf{z}, y)}{\partial \mathbf{z}} = \dfrac{\partial^2 f(\mathbf{z}, y)}{\partial y \partial \mathbf{z}} \\[2mm] \ddot{\mathbf{F}}_{11}(\mathbf{z}, y) = \dfrac{\partial^2 f(\mathbf{z}, y)}{\partial \mathbf{z} \partial \mathbf{z}^{\mathrm{T}}} \end{cases}$$
(49)

The proof of Proposition 2 is provided in [Appendix 3](). Note that the gradient and Hessian matrix of the criterion function $f_4(\mathbf{p})$ can be obtained from this result.

We first assume that, for a given $\mathbf{p}$, the point that maximizes $f_3(\mathbf{p}, t_0)$ with respect to t_0 is denoted as $t_{0,\mathrm{m}}(\mathbf{p})$. This can be obtained by the fast Fourier transform (FFT) [27]. Using (45) and the first equality in (48), the gradient of $f_4(\mathbf{p})$ can be expressed as

$$\begin{aligned} \mathbf{r}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) &= \frac{\partial f_4(\mathbf{p})}{\partial \mathbf{p}} = \left. \frac{\partial f_3(\mathbf{p}, t_0)}{\partial \mathbf{p}} \right|_{t_0 = t_{0,m}(\mathbf{p})} \\ &= 2 \cdot \mathrm{Re}\left\{ \left(\frac{\partial \mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right)}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{\Pi}^{\mathrm{T}} \right. \\ &\quad ((\overline{\mathbf{S}}'^{*}(\mathbf{d}(t_{0,m}(\mathbf{p})))) \\ &\quad {}^{*}(\mathbf{d}(t_{0,m}(\mathbf{p})))^{\mathrm{T}} \overline{\mathbf{S}}'^{\mathrm{T}}) \otimes \mathbf{I}_N) \mathbf{\Pi} \mathbf{b}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}) \} \end{aligned}$$
(50)

Furthermore, using the second equality in (48), the Hessian matrix of $f_4(\mathbf{p})$ can be formulated as

$$\begin{aligned} \mathbf{R}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) &= \frac{\partial^2 f_4(\mathbf{p})}{\partial \mathbf{p} \partial \mathbf{p}^{\mathrm{T}}} \\ &= \mathbf{R}_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) - \frac{1}{r_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right)} \\ &\quad \cdot \mathbf{r}_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) \left(\mathbf{r}_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) \right)^{\mathrm{T}} \end{aligned}$$
(51)

where

$$\begin{aligned} \mathbf{R}_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) &= \left. \frac{\partial^2 f_3(\mathbf{p}, t_0)}{\partial \mathbf{p} \partial \mathbf{p}^{\mathrm{T}}} \right|_{t_0 = t_{0,m}(\mathbf{p})} \quad , \quad \mathbf{r}_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) \\ &= \left. \frac{\partial^2 f_3(\mathbf{p}, t_0)}{\partial \mathbf{p} \partial t_0} \right|_{t_0 = t_{0,m}(\mathbf{p})} \quad , \quad r_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) \\ &= \left. \frac{\partial^2 f_3(\mathbf{p}, t_0)}{\partial t_0^2} \right|_{t_0 = t_{0,m}(\mathbf{p})} \end{aligned}$$
(52)

Combining (45) and (50) leads to

$$\begin{aligned} \mathbf{R}_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right) &= 2 \cdot \mathrm{Re}\left\{ \left(\frac{\partial \mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right)}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{\Pi}^{\mathrm{T}} ((\overline{\mathbf{S}}'^{*}(\mathbf{d}(t_{0,m}(\mathbf{p})))) \right. \\ &\quad {}^{*}(\mathbf{d}(t_{0,m}(\mathbf{p})))^{\mathrm{T}} \overline{\mathbf{S}}'^{\mathrm{T}}) \otimes \mathbf{I}_N) \mathbf{\Pi} \cdot \frac{\partial \mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right)}{\partial \mathbf{p}^{\mathrm{T}}} \} + 2 \\ &\quad \cdot \mathrm{Re}\{((\mathbf{\Pi}^{\mathrm{T}}((\overline{\mathbf{S}}'^{*}(\mathbf{d}(t_{0,m}(\mathbf{p})))) \\ &\quad {}^{*}(\mathbf{d}(t_{0,m}(\mathbf{p})))^{\mathrm{T}} \overline{\mathbf{S}}'^{\mathrm{T}}) \otimes \mathbf{I}_N) \mathbf{\Pi} \mathbf{b}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}))^{\mathrm{T}} \otimes \mathbf{I}_L) \\ &\quad \cdot \frac{\partial}{\partial \mathbf{p}^{\mathrm{T}}} \mathrm{vec}\left(\left(\frac{\partial \mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(\mathrm{b})}\right)}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \right) \} \end{aligned}$$
(53)

$$
\begin{aligned}
\mathbf{r}_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right) = {} & 2\cdot \operatorname{Re}\left\{ \left(\frac{\partial \mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right)}{\partial \mathbf{p}^{\mathrm{T}}}\right)^{\mathrm{H}} \mathbf{\Pi}^{\mathrm{T}}\left(\left(\overline{\mathbf{S}}^{\prime *}\left(\dot{\mathbf{d}}\left(t_{0,m}(\mathbf{p})\right)\right)^{*}\left(\mathbf{d}\left(t_{0,m}(\mathbf{p})\right)\right)^{\mathrm{T}} \overline{\mathbf{S}}^{\prime \mathrm{T}}\right) \otimes \mathbf{I}_N\right) \mathbf{\Pi}\mathbf{b}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}) \right\} \\
& + 2\cdot \operatorname{Re}\left\{ \left(\frac{\partial \mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right)}{\partial \mathbf{p}^{\mathrm{T}}}\right)^{\mathrm{H}} \mathbf{\Pi}^{\mathrm{T}}\left(\left(\overline{\mathbf{S}}^{\prime *}\left(\mathbf{d}\left(t_{0,m}(\mathbf{p})\right)\right)^{*}\left(\dot{\mathbf{d}}\left(t_{0,m}(\mathbf{p})\right)\right)^{\mathrm{T}} \overline{\mathbf{S}}^{\prime \mathrm{T}}\right) \otimes \mathbf{I}_N\right) \mathbf{\Pi}\mathbf{b}(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}) \right\}
\end{aligned}
\tag{54}
$$

$$
\begin{aligned}
r_0\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right) = {} & 2\cdot \operatorname{Re}\left\{ \left(\mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right)\right)^{\mathrm{H}} \mathbf{\Pi}^{\mathrm{T}}\left(\left(\overline{\mathbf{S}}^{\prime *}\left(\ddot{\mathbf{d}}\left(t_{0,m}(\mathbf{p})\right)\right)^{*}\left(\mathbf{d}\left(t_{0,m}(\mathbf{p})\right)\right)^{\mathrm{T}} \overline{\mathbf{S}}^{\prime \mathrm{T}}\right) \otimes \mathbf{I}_N\right) \mathbf{\Pi}\mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right) \right\} \\
& + 2\cdot \operatorname{Re}\left\{ \left(\mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right)\right)^{\mathrm{H}} \mathbf{\Pi}^{\mathrm{T}}\left(\left(\overline{\mathbf{S}}^{\prime *}\left(\dot{\mathbf{d}}\left(t_{0,m}(\mathbf{p})\right)\right)^{*}\left(\dot{\mathbf{d}}\left(t_{0,m}(\mathbf{p})\right)\right)^{\mathrm{T}} \overline{\mathbf{S}}^{\prime \mathrm{T}}\right) \otimes \mathbf{I}_N\right) \mathbf{\Pi}\mathbf{b}\left(\mathbf{p}, \hat{\boldsymbol{\mu}}^{(b)}\right) \right\}
\end{aligned}
\tag{55}
$$

At this point, the Gauss-Newton algorithm for solving (46) can be formulated as

$$
\hat{\mathbf{p}}^{(i+1)} = \hat{\mathbf{p}}^{(i)} - \alpha^i \left(\mathbf{R}\left(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(b)}\right)\right)^{-1} \mathbf{r}\left(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(b)}\right)
\tag{56}
$$

where $\alpha \in (0, 1)$ is the step size, and superscript i represents the iteration number.

Let the convergence result of (56) be $\hat{\mathbf{p}}^{(b)}$. Then, t_0 can be obtained from the following minimization problem:

$$
\max_{t_0} f_5(t_0) = \max_{t_0} \left\| (\mathbf{B}(\hat{\mathbf{p}}^{(b)}, \hat{\boldsymbol{\mu}}^{(b)}))^{\mathrm{H}} \overline{\mathbf{S}}' \mathbf{d}(t_0) \right\|_2^2
\tag{57}
$$

Note that the cost function in (57) can be minimized via FFT algorithm, which has been discussed in [27].

5.2.2 Joint optimization of β and μ

The objective function in (41) is now minimized with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$, with $\mathbf{p}$ and t_0 given by $\hat{\mathbf{p}}^{(b)}$ and $\hat{t}_0^{(b)}$, respectively. Note that the algorithm presented in Section 4.2.2 can still be applied here, but $\hat{\mathbf{p}}^{(a)}$, $\boldsymbol{\gamma}_n(\hat{\mathbf{p}}^{(a)})$, and $\hat{\overline{\mathbf{s}}}^{(a)}$ must be replaced with $\hat{\mathbf{p}}^{(b)}$, $\boldsymbol{\phi}_n(\hat{\mathbf{p}}^{(b)}, \hat{t}_0^{(b)})$, and

$$
\begin{aligned}
\hat{\overline{\mathbf{s}}}^{(b)} = \big[& \bar{s}_1 \cdot \exp\left\{-j\omega_1 \hat{t}_0^{(b)}\right\} \quad \bar{s}_2 \\
& \cdot \exp\left\{-j\omega_2 \hat{t}_0^{(b)}\right\} \quad \cdots \quad \bar{s}_K \cdot \exp\left\{-j\omega_K \hat{t}_0^{(b)}\right\} \big]^{\mathrm{T}},
\end{aligned}
\tag{58}
$$

respectively.

5.2.3 Summary of the alternating minimization algorithm
A possible implementation of the proposed alternating minimization algorithm is outlined as follows.

Proposed alternating minimization algorithm II

Step 1: Define a convergence threshold $\delta > 0$ and choose the initial values $\hat{\mathbf{p}}^{(0)}$, $\hat{t}_0^{(0)}$, $\hat{\boldsymbol{\mu}}^{(0)}$ and $\hat{\boldsymbol{\beta}}^{(0)}$.

Step 2: Set the iteration counter $m = 0$ and compute the cost function $J_K^{(m)}$ using (41).

Step 3: Calculate $\hat{\mathbf{p}}^{(m)}$ using the Gauss-Newton algorithm given in (56).

Step 4: Compute $\hat{t}_0^{(m)}$ from (57) by FFT algorithm.

Step 5: Calculate $\hat{\boldsymbol{\mu}}^{(m)}$ from (33) and (58).

Step 6: Compute $\hat{\boldsymbol{\beta}}^{(m)}$ via (29) and (58).

Step 7: Increment the iteration counter $m = m + 1$ and compute the cost function $J_K^{(m)}$ using (41). If $| J_K^{(m)} - J_K^{(m-1)} | \leq \delta$, stop the procedure; otherwise go to Step 3.

In the following, we make two remarks concerning the alternating minimization algorithm described above.

Remark 9: Similar to sequence $\{J_U^{(m)}\}_{m=1}^{+\infty}$, $\{J_K^{(m)}\}_{m=1}^{+\infty}$ is also monotonically decreasing and, hence, convergence is guaranteed. From our simulation results, it can be observed that 20 iterations are generally enough to satisfy the convergence criterion.

Remark 10: The selection of the initial values is again important. For the current problem, the initial value of $\boldsymbol{\mu}$ can be chosen as its nominal value $\boldsymbol{\mu}_0$ and the starting point of $\mathbf{p}$ can be obtained using the two-step localization method, where the DOA parameter can be estimated by the algorithm proposed for the scenario of known waveforms [55, 56]. As stated in Remark 8, when the unknowns to be determined are DOAs, we need to express the array manifold in (1)–(3) as a function of direction, as in [1–3]. This can easily be achieved, because the position vector $\mathbf{p}$ is a function of DOA. Once the DOAs corresponding to all the arrays have been estimated, the estimated initial position is given by the closed-form method

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 13 of 28

[4]. Further, from (57), we have the initial solution of t_0 from the FFT algorithm, and the initial guess of $\boldsymbol{\beta}$ can be given by (29) and (58). From the results in Section 7, it is clear that these initial estimates allow the proposed algorithm to provide satisfactory estimation accuracy.

5.2.4 Complexity analysis

We now address the computational complexity of the proposed DPD method in terms of the number of multiplications. Table 2 summarizes the numerical complexity.

6 Cramér-Rao bound on covariance matrix of localization errors

The CRB gives a lower bound on the asymptotic covariance matrix of any asymptotically unbiased estimator. The bound can be obtained by the inverse of the Fisher information matrix (FIM). Relating the parameter estimates to this bound can provide a nature measure of performance. The purpose of this section is to derive the CRBs for the estimates of the emitter's position. We consider the cases when the signal waveform is unknown and when the signal waveform is known.

6.1 Cramér-Rao bound on position estimate for case of unknown signal waveform

This subsection is devoted to the derivation of the CRB for localization under the assumption that the signal waveform is unknown. For this case, the full parameter set contains both the deterministic parameters σ_ε^2, $\mathbf{p}$, $\boldsymbol{\beta}$, $\bar{\mathbf{s}}$

, and the stochastic parameter $\boldsymbol{\mu}$. Hence, the CRB derivation should follow the Bayesian theory framework [47–50]. Note that the CRB derivation can also be used for stochastic parameters, as in [47–50, 57, 58]. To this end, a novel parameter vector that comprises all the deterministic and stochastic unknowns is introduced as

$$
\begin{aligned}
\boldsymbol{\eta}^{(a)} &= \begin{bmatrix} \sigma_\varepsilon^2 & \mathbf{p}^T & (\,\mathrm{Re}\{\boldsymbol{\beta}\})^T & (\,\mathrm{Im}\{\boldsymbol{\beta}\})^T & (\,\mathrm{Re}\{\bar{\mathbf{s}}\})^T & (\,\mathrm{Im}\{\bar{\mathbf{s}}\})^T & \boldsymbol{\mu} \end{bmatrix}^T \\
&= \begin{bmatrix} \sigma_\varepsilon^2 & \boldsymbol{\eta}'^{(a)} \end{bmatrix}^T
\end{aligned}
$$

$$(59)$$

where

$$
\boldsymbol{\eta}'^{(a)} = \begin{bmatrix} \mathbf{p}^T & (\,\mathrm{Re}\{\boldsymbol{\beta}\})^T & (\,\mathrm{Im}\{\boldsymbol{\beta}\})^T & (\,\mathrm{Re}\{\bar{\mathbf{s}}\})^T & (\,\mathrm{Im}\{\bar{\mathbf{s}}\})^T & \boldsymbol{\mu} \end{bmatrix}^T
$$

$$(60)$$

We proceed to define a data vector as

$$
\begin{aligned}
\bar{\mathbf{x}} &= \begin{bmatrix} \bar{\mathbf{x}}_1^H & \bar{\mathbf{x}}_2^H & \cdots & \bar{\mathbf{x}}_N^H \end{bmatrix}^H \\
&= \begin{bmatrix} \bar{\mathbf{x}}_{1,1}^H & \bar{\mathbf{x}}_{1,2}^H & \cdots & \bar{\mathbf{x}}_{1,K}^H & \bar{\mathbf{x}}_{2,1}^H & \bar{\mathbf{x}}_{2,2}^H \\
& \cdots & \bar{\mathbf{x}}_{2,K}^H & \cdots & \cdots & \bar{\mathbf{x}}_{N,1}^H & \bar{\mathbf{x}}_{N,2}^H \\
& \cdots & \bar{\mathbf{x}}_{N,K}^H \end{bmatrix}^H
\end{aligned}
$$

$$(61)$$

The mean vector of $\bar{\mathbf{x}}$ is given by

$$
\begin{aligned}
\bar{\mathbf{x}}_0 = \mathrm{E}[\bar{\mathbf{x}}] = \begin{bmatrix} \bar{\mathbf{x}}_{1,1,0}^H & \bar{\mathbf{x}}_{1,2,0}^H & \cdots & \bar{\mathbf{x}}_{1,K,0}^H & \bar{\mathbf{x}}_{2,1,0}^H & \bar{\mathbf{x}}_{2,2,0}^H \\
& \cdots & \bar{\mathbf{x}}_{2,K,0}^H & \cdots \\
& \cdots & \bar{\mathbf{x}}_{N,1,0}^H & \bar{\mathbf{x}}_{N,2,0}^H \\
& \cdots & \bar{\mathbf{x}}_{N,K,0}^H \end{bmatrix}^H
\end{aligned}
$$

$$(62)$$

where

$$
\begin{aligned}
\bar{\mathbf{x}}_{n,k,0} = \mathrm{E}\left[\bar{\mathbf{x}}_{n,k}\right] &= \beta_n \mathbf{a}_n(\mathbf{p}, \boldsymbol{\mu}_n) \bar{s}_k \\
&\cdot \exp\{-\mathrm{j}\omega_k(\tau_n(\mathbf{p}) + t_0)\}
\end{aligned}
$$

$$(63)$$

When the deterministic and stochastic parameters coexist, we should use the hybrid CRB. As a consequence, the FIM for vector $\boldsymbol{\eta}^{(a)}$ is given by [47–50, 57, 58].

$$
\begin{aligned}
< \mathbf{FIM}\left(\boldsymbol{\eta}^{(a)}\right) >_{ij} &= < \mathbf{FIM}_1\left(\boldsymbol{\eta}^{(a)}\right) >_{ij} + < \mathbf{FIM}_2\left(\boldsymbol{\eta}^{(a)}\right) >_{ij} \\
&= \mathrm{E}_{\bar{\mathbf{x}}, \boldsymbol{\mu}}\left[\frac{\partial^2 f_{\mathrm{ML}}\left(\boldsymbol{\eta}^{(a)} | \bar{\mathbf{x}}\right)}{\partial < \boldsymbol{\eta}^{(a)} >_i \partial < \boldsymbol{\eta}^{(a)} >_j} \right] \\
&+ \mathrm{E}_{\boldsymbol{\mu}}\left[\frac{1}{2} \cdot \frac{\partial^2 (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \boldsymbol{\Omega}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)}{\partial < \boldsymbol{\eta}^{(a)} >_i \partial < \boldsymbol{\eta}^{(a)} >_j} \right]
\end{aligned}
$$

$$(64)$$

where

$$
\begin{cases}
< \mathbf{FIM}_1\left(\boldsymbol{\eta}^{(a)}\right) >_{ij} = \mathrm{E}_{\bar{\mathbf{x}}, \boldsymbol{\mu}}\left[\dfrac{\partial^2 f_{\mathrm{ML}}\left(\boldsymbol{\eta}^{(a)} | \bar{\mathbf{x}}\right)}{\partial < \boldsymbol{\eta}^{(a)} >_i \partial < \boldsymbol{\eta}^{(a)} >_j} \right] \\
< \mathbf{FIM}_2\left(\boldsymbol{\eta}^{(a)}\right) >_{ij} = \mathrm{E}_{\boldsymbol{\mu}}\left[\dfrac{1}{2} \cdot \dfrac{\partial^2 (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \boldsymbol{\Omega}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)}{\partial < \boldsymbol{\eta}^{(a)} >_i \partial < \boldsymbol{\eta}^{(a)} >_j} \right]
\end{cases}
$$

$$(65)$$

in which $f_{\mathrm{ML}}(\boldsymbol{\eta}^{(a)} | \bar{\mathbf{x}})$ is the maximum likelihood (ML) function of the compound data vector $\bar{\mathbf{x}}$. Note that the first term, $\mathbf{FIM}_1(\boldsymbol{\eta}^{(a)})$, resembles the standard expression for the FIM in the unperturbed model augmented by the perturbation parameter $\boldsymbol{\mu}$. However, $\mathbf{FIM}_1(\boldsymbol{\eta}^{(a)})$ is not easily calculated, because the $\boldsymbol{\mu}$-parameter generally behaves in a nonlinear fashion, making the expectation with respect to $\boldsymbol{\mu}$ difficult to compute. In [57], the above expression is examined for some special cases, but it is hard to generalize this result. Indeed, a more common approach is to ignore the expectation with respect to $\boldsymbol{\mu}$ and calculate $\mathbf{FIM}_1(\boldsymbol{\eta}^{(a)})$ at $\boldsymbol{\mu}_0$, as in [47–50, 58]. Following the analysis in [47–50, 58], it can be shown that this approximation is $O(1)$, which implies that

Table 2 Complexity of proposed method

Calculation module	Computational unit	Computational complexity of each unit	Total computational complexity
Gauss-Newton algorithm in (56)	Calculation of $t_{0,m}(\hat{\mathbf{p}}^{(i)})$	$MN^2K^2 + (N+2)K^2 + 0.5NK \cdot \log_2 K$	$I_{\text{iter}}\left\{ I_N\left(\begin{array}{l} MN^2K^2 + N^2(3K^2L + 2K^2) \\ +N\left(\begin{array}{l} KL^3 + KL^2 + 2KL \\ +K^2 + K + 0.5K \cdot \log_2 K \\ +6K^2 + 4K + 2L^2 + O(L^3) \end{array} \right) \end{array} \right) \right.$ $\left. +I_{\text{G-N}}N\left(\begin{array}{l} Q_n^2(MK+3) \\ +Q_n(M^2(K^2+K)+2MK) \\ +M^2(3K^2+K)+M(2K+1) \\ +2K+O(Q_n^3) \end{array} \right) \right.$ $\left. +M(N^2K^2+2NK+N)+(N+2)K^2 +0.5NK\cdot\log_2 K+NK \right\}$
	$\mathbf{R}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(b)})$	$N^2(2K^2L + 2K^2) + N\left(\begin{array}{l} KL^3 + KL^2 \\ +KL + K \end{array} \right) + 3K^2 + 3K + 2L^2$	(where I_N, $I_{\text{G-N}}$
	$\mathbf{r}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(b)})$	$N^2K^2L + NKL + K^2 + K$	and I_{iter} denote the iteration numbers for Gauss-Newton algorithm and alternating minimization algorithm, respectively)
	$\alpha^i(\mathbf{R}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(b)}))^{-1} \mathbf{r}(\hat{\mathbf{p}}^{(i)}, \hat{\boldsymbol{\mu}}^{(b)})$	$O(L^3)$	
Estimation of t_0	$\mathbf{B}(\hat{\mathbf{p}}^{(b)}, \hat{\boldsymbol{\mu}}^{(b)})$	MN^2K^2	
	$\bar{\mathbf{S}}^{H} \mathbf{B}(\hat{\mathbf{p}}^{(b)}, \hat{\boldsymbol{\mu}}^{(b)})(\mathbf{B}(\hat{\mathbf{p}}^{(b)}, \hat{\boldsymbol{\mu}}^{(b)}))^{H} \bar{\mathbf{S}}'$	$(N+2)K^2$	
	FFT algorithm	$0.5NK \cdot \log_2 K$	
Gauss-Newton algorithm in (33)	$\mathbf{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)})$	$Q_n^2 + 2M^2K^2 + M(K+1) + 2K$	
	$\mathbf{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})$	$Q_n\left(\begin{array}{l} M^2(K^2+K) \\ +MK \end{array} \right) + M^2(K^2+K) + MK$	
	$(\text{Re}\{(G_n(\hat{\mu}_n^{(i)}))^H G_n(\hat{\mu}_n^{(i)})\})^{-1}$ $\times \text{Re}\left\{ (G_n(\hat{\mu}_n^{(i)}))^H \left[\frac{1}{\sqrt{2}} \cdot \boldsymbol{\Omega}_{n,0}^{-1/2} \boldsymbol{\mu}_{n,0} - \frac{\bar{x}_n}{\sqrt{2}} \cdot \boldsymbol{\Omega}_{n,0}^{-1/2} \boldsymbol{\mu}_{n,0} - g_n(\mu_n^{(i)}) \right] \right\}$	$Q_n^2(MK+2) + Q_n MK + O(Q_n^3)$	
Estimation of $\boldsymbol{\beta}$	$\dfrac{((\widehat{\bar{\mathbf{s}}}^{(b)} \odot \boldsymbol{\varphi}_{t_0}(\hat{\mathbf{p}}^{(b)}, \hat{t}_0^{(b)})) \otimes \mathbf{a}_r(\hat{\mathbf{p}}^{(b)}, \hat{\boldsymbol{\mu}}_n^{(b)}))^{H} \bar{\mathbf{x}}_n}{\|\mathbf{a}_r(\hat{\mathbf{p}}^{(b)}, \hat{\boldsymbol{\mu}}_n^{(b)})\|_2^2 \cdot \|\widehat{\bar{\mathbf{s}}}^{(b)} \odot \boldsymbol{\varphi}_{t_0}(\hat{\mathbf{p}}^{(b)}, \hat{t}_0^{(b)})\|_2}$	$M(2K+1)+K$	

$$
E_{\overline{\mathbf{x}},\boldsymbol{\mu}} \left[\frac{\partial^2 f_{\mathrm{ML}}\left(\boldsymbol{\eta}^{(\mathrm{a})}|\overline{\mathbf{x}}\right)}{\partial <\boldsymbol{\eta}^{(\mathrm{a})}>_i \partial <\boldsymbol{\eta}^{(\mathrm{a})}>_j} \right]
$$

$$
= E_{\overline{\mathbf{x}}} \left[\frac{\partial^2 f_{\mathrm{ML}}\left(\boldsymbol{\eta}^{(\mathrm{a})}|\overline{\mathbf{x}}\right)}{\partial <\boldsymbol{\eta}^{(\mathrm{a})}>_i \partial <\boldsymbol{\eta}^{(\mathrm{a})}>_j} \Big|_{\boldsymbol{\mu}_0} \right] + O(1) \qquad (66)
$$

Additionally, note that $E_{\overline{\mathbf{x}},\boldsymbol{\mu}}[\frac{\partial^2 f_{\mathrm{ML}}(\boldsymbol{\eta}^{(\mathrm{a})}|\overline{\mathbf{x}})}{\partial <\boldsymbol{\eta}^{(\mathrm{a})}>_i \partial <\boldsymbol{\eta}^{(\mathrm{a})}>_j}] = O(NK)$, so we can still obtain an asymptotically valid CRB. Another reason for ignoring this approximation is that the distribution of $\boldsymbol{\mu}$ is symmetric and, in this paper, the error $\boldsymbol{\mu} - \boldsymbol{\mu}_0$ is assumed to be sufficiently small.

Combining (64) and the above discussion, the FIM for vector $\boldsymbol{\eta}^{'(\mathrm{a})}$ can be approximately expressed as

$$
< \mathbf{FIM}\left(\boldsymbol{\eta}^{'(\mathrm{a})}\right)>_{ij} = \frac{2}{\sigma_\varepsilon^2} \cdot
$$
$$
< \mathrm{Re}\left\{\mathbf{T}_{\boldsymbol{\eta}^{'(\mathrm{a})}}^{\mathrm{H}} \mathbf{T}_{\boldsymbol{\eta}^{'(\mathrm{a})}}\right\}>_{ij} +
$$
$$
< \boldsymbol{\Omega}^{-1}>_{ij} \cdot \delta(i,j) \qquad (67)
$$

where $\delta(i,j)$ is an indicator function such that $\delta(i,j) = 1$ if both i and j correspond to the element in $\boldsymbol{\mu}$, $\delta(i,j) = 0$ otherwise, $\boldsymbol{\Omega} = \mathrm{blkdiag}[\boldsymbol{\Omega}_1 \quad \boldsymbol{\Omega}_2 \quad \cdots \quad \boldsymbol{\Omega}_N]$, and

$$
\mathbf{T}_{\boldsymbol{\eta}^{'(\mathrm{a})}} = \frac{\partial \overline{\mathbf{x}}_0}{\partial \boldsymbol{\eta}^{'(\mathrm{a})\mathrm{T}}} = [\frac{\partial \overline{\mathbf{x}}_0}{\partial \mathbf{p}^{\mathrm{T}}} \quad \frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Re}\{\boldsymbol{\beta}\})^{\mathrm{T}}} \quad \frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Im}\{\boldsymbol{\beta}\})^{\mathrm{T}}} \quad \frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Re}\{\overline{\mathbf{s}}\})^{\mathrm{T}}}
$$

$$
\frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Im}\{\overline{\mathbf{s}}\})^{\mathrm{T}}} \quad \frac{\partial \overline{\mathbf{x}}_0}{\partial \boldsymbol{\mu}^{\mathrm{T}}}]
$$

$$
= \begin{bmatrix} \mathbf{T}_{\mathbf{p}} & \mathbf{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}} & \mathbf{T}_{\mathrm{Im}\{\boldsymbol{\beta}\}} & \mathbf{T}_{\mathrm{Re}\{\overline{\mathbf{s}}\}} & \mathbf{T}_{\mathrm{Im}\{\overline{\mathbf{s}}\}} & \mathbf{T}_{\boldsymbol{\mu}} \end{bmatrix} \qquad (68)
$$

It follows from (67) that

$$
\mathbf{FIM}\left(\boldsymbol{\eta}^{'(\mathrm{a})}\right) = \frac{2}{\sigma_\varepsilon^2} \cdot \mathrm{Re}\left\{\mathbf{T}_{\boldsymbol{\eta}^{'(\mathrm{a})}}^{\mathrm{H}} \mathbf{T}_{\boldsymbol{\eta}^{'(\mathrm{a})}}\right\} + \begin{bmatrix} \mathbf{OO} \\ \mathbf{O}\boldsymbol{\Omega}^{-1} \end{bmatrix} \qquad (69)
$$

Using (62) and (63), and after some algebraic manipulations, the sub-matrices in (68) can be written as

$$
\begin{cases}
\mathbf{T}_{\mathbf{p}} = \frac{\partial \overline{\mathbf{x}}_0}{\partial \mathbf{p}^{\mathrm{T}}} = \mathrm{diag}[\boldsymbol{\beta} \otimes \mathbf{1}_{MK\times 1}] \cdot \mathrm{diag}[\mathbf{1}_{N\times 1} \otimes \overline{\mathbf{s}} \otimes \mathbf{1}_{M\times 1}] \cdot \frac{\partial(\mathbf{A}(\mathbf{p},\boldsymbol{\mu}_0)\mathbf{1}_{K\times 1})}{\partial \mathbf{p}^{\mathrm{T}}} \\
\mathbf{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}} = \frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Re}\{\boldsymbol{\beta}\})^{\mathrm{T}}} = \mathrm{diag}[\mathbf{1}_{N\times 1} \otimes \overline{\mathbf{s}} \otimes \mathbf{1}_{M\times 1}] \cdot \mathrm{blkdiag}\left[\mathbf{A}_1\left(\mathbf{p},\boldsymbol{\mu}_{1,0}\right)\mathbf{1}_{K\times 1} \quad \mathbf{A}_2\left(\mathbf{p},\boldsymbol{\mu}_{2,0}\right)\mathbf{1}_{K\times 1} \quad \cdots \quad \mathbf{A}_N\left(\mathbf{p},\boldsymbol{\mu}_{N,0}\right)\mathbf{1}_{K\times 1}\right] \\
\mathbf{T}_{\mathrm{Im}\{\boldsymbol{\beta}\}} = \frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Im}\{\boldsymbol{\beta}\})^{\mathrm{T}}} = \mathrm{j}\mathbf{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}} = \mathrm{j} \cdot \mathrm{diag}[\mathbf{1}_{N\times 1} \otimes \overline{\mathbf{s}} \otimes \mathbf{1}_{M\times 1}] \cdot \mathrm{blkdiag}\left[\mathbf{A}_1\left(\mathbf{p},\boldsymbol{\mu}_{1,0}\right)\mathbf{1}_{K\times 1} \quad \mathbf{A}_2\left(\mathbf{p},\boldsymbol{\mu}_{2,0}\right)\mathbf{1}_{K\times 1} \quad \cdots \quad \mathbf{A}_N\left(\mathbf{p},\boldsymbol{\mu}_{N,0}\right)\mathbf{1}_{K\times 1}\right] \\
\mathbf{T}_{\mathrm{Re}\{\overline{\mathbf{s}}\}} = \frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Re}\{\overline{\mathbf{s}}\})^{\mathrm{T}}} = \mathrm{diag}[\boldsymbol{\beta} \otimes \mathbf{1}_{MK\times 1}] \cdot \mathbf{A}(\mathbf{p},\boldsymbol{\mu}_0) \\
\mathbf{T}_{\mathrm{Im}\{\overline{\mathbf{s}}\}} = \frac{\partial \overline{\mathbf{x}}_0}{\partial (\mathrm{Im}\{\overline{\mathbf{s}}\})^{\mathrm{T}}} = \mathrm{j}\mathbf{T}_{\mathrm{Re}\{\overline{\mathbf{s}}\}} = \mathrm{j} \cdot \mathrm{diag}[\boldsymbol{\beta} \otimes \mathbf{1}_{MK\times 1}] \cdot \mathbf{A}(\mathbf{p},\boldsymbol{\mu}_0) \\
\mathbf{T}_{\boldsymbol{\mu}} = \frac{\partial \overline{\mathbf{x}}_0}{\partial \boldsymbol{\mu}^{\mathrm{T}}} = \mathrm{diag}[\boldsymbol{\beta} \otimes \mathbf{1}_{MK\times 1}] \cdot \mathrm{diag}[\mathbf{1}_{N\times 1} \otimes \overline{\mathbf{s}} \otimes \mathbf{1}_{M\times 1}] \cdot \frac{\partial(\mathbf{A}(\mathbf{p},\boldsymbol{\mu}_0)\mathbf{1}_{K\times 1})}{\partial \boldsymbol{\mu}_0^{\mathrm{T}}}
\end{cases} \qquad (70)
$$

Substituting (68) and (70) into (69) leads to

$$
\mathbf{FIM}(\boldsymbol{\eta}^{'(\mathrm{b})}) = \frac{2}{\sigma_\varepsilon^2} \mathrm{Re} \begin{bmatrix} T_p^{\mathrm{H}} T_p & T_p^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}} & T_p^{\mathrm{H}} T_{\mathrm{Im}\{\boldsymbol{\beta}\}} & T_p^{\mathrm{H}} T_{\overline{s}_r} & T_p^{\mathrm{H}} T_\mu \\ T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_p & T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}} & T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\mathrm{Im}\{\boldsymbol{\beta}\}} & T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\overline{s}_r} & T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_\mu \\ T_{\mathrm{Im}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_p & T_{\mathrm{Im}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}} & T_{\mathrm{Im}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\mathrm{Im}\{\boldsymbol{\beta}\}} & T_{\mathrm{Im}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\overline{s}_r} & T_{\mathrm{Im}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_\mu \\ T_{\overline{s}_r}^{\mathrm{H}} T_p & T_{\overline{s}_r}^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}} & T_{\overline{s}_r}^{\mathrm{H}} T_{\mathrm{Im}\{\boldsymbol{\beta}\}} & T_{\overline{s}_r}^{\mathrm{H}} T_{\overline{s}_r} & T_{\overline{s}_r}^{\mathrm{H}} T_\mu \\ T_\mu^{\mathrm{H}} T_p & T_\mu^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}} & T_\mu^{\mathrm{H}} T_{\mathrm{Im}\{\boldsymbol{\beta}\}} & T_\mu^{\mathrm{H}} T_{\overline{s}_r} & T_\mu^{\mathrm{H}} T_\mu \end{bmatrix} + \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Omega}^{-1} \end{bmatrix}
$$
$$
= \frac{2}{\sigma_\varepsilon^2} \mathrm{Re} \begin{bmatrix} Z_1^{(\mathrm{b})} & Z_2^{(\mathrm{b})} \\ Z_2^{(\mathrm{b})\mathrm{H}} & Z_3^{(\mathrm{b})} \end{bmatrix} + \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Omega}^{-1} \end{bmatrix} \qquad (79)
$$

where

$$
\begin{bmatrix} \mathbf{Z}_1^{(\mathrm{b})} = T_p^{\mathrm{H}} T_p, \ \mathbf{Z}_2^{(\mathrm{b})} = [[1 \ \mathrm{j}] \otimes (T_p^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}}) \ \vdots \ T_p^{\mathrm{H}} T_{\overline{s}_r} \ \vdots \ T_p^{\mathrm{H}} T_\mu] \\ \mathbf{Z}_3^{(\mathrm{b})} = \begin{bmatrix} \begin{bmatrix} 1 & \mathrm{j} \\ -\mathrm{j} & 1 \end{bmatrix} \otimes (T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}}) & \begin{bmatrix} 1 \\ -\mathrm{j} \end{bmatrix} \otimes (T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_{\overline{s}_r}) & \begin{bmatrix} 1 \\ -\mathrm{j} \end{bmatrix} \otimes (T_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} T_\mu) \\ [1 \ \mathrm{j}] \otimes (T_{\overline{s}_r}^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}}) & T_{\overline{s}_r}^{\mathrm{H}} T_{\overline{s}_r} & T_{\overline{s}_r}^{\mathrm{H}} T_\mu \\ [1 \ \mathrm{j}] \otimes (T_\mu^{\mathrm{H}} T_{\mathrm{Re}\{\boldsymbol{\beta}\}}) & T_\mu^{\mathrm{H}} T_{\overline{s}_r} & T_\mu^{\mathrm{H}} T_\mu \end{bmatrix} \end{bmatrix} \qquad (80)
$$

The details of calculating the matrices in (72) are shown in Appendix 4.

Note that, in practice, only the $\mathbf{p}$ corner of the CRB matrix is of interest. Invoking the partitioned matrix inversion formula, the CRB matrix for position vector $\mathbf{p}$ can be written as

$$
\mathbf{CRB}^{(\mathrm{a})}(\mathbf{p}) = \frac{\sigma_\varepsilon^2}{2} \cdot \left(\begin{aligned} &\left(\mathrm{Re}\left\{\mathbf{Z}_1^{(\mathrm{a})}\right\}\right)^{-1} + \left(\mathrm{Re}\left\{\mathbf{Z}_1^{(\mathrm{a})}\right\}\right)^{-1} \\ &\cdot \mathrm{Re}\left\{\mathbf{Z}_2^{(\mathrm{a})}\right\} \cdot (\mathrm{Re}\left\{\mathbf{Z}_3^{(\mathrm{a})}\right\} \\ &- \mathrm{Re}\left\{\mathbf{Z}_2^{(\mathrm{a})\mathrm{H}}\right\} \cdot \left(\mathrm{Re}\left\{\mathbf{Z}_1^{(\mathrm{a})}\right\}\right)^{-1} \\ &\cdot \mathrm{Re}\left\{\mathbf{Z}_2^{(\mathrm{a})}\right\} + \begin{bmatrix} \mathbf{OO} \\ \mathbf{O}\sigma_\varepsilon^2\boldsymbol{\Omega}^{-1}/2 \end{bmatrix}\right)^{-1} \\ &\times \mathrm{Re}\left\{\mathbf{Z}_2^{(\mathrm{a})\mathrm{H}}\right\} \cdot \left(\mathrm{Re}\left\{\mathbf{Z}_1^{(\mathrm{a})}\right\}\right)^{-1}) \end{aligned} \right)
$$
$$
\qquad (73)
$$

Note that the superscript "a" in (73) indicates that this CRB corresponds to the case of an unknown signal waveform. In addition, the trace of this CRB matrix is the minimum achievable localization MSE.

6.2 Cramér-Rao bound on position estimate for case of known signal waveform

In this subsection, the CRB on the covariance matrix of location estimation is deduced for the case where the transmitted signals are a priori known. Accordingly, the complete parameter set comprises the deterministic parameters σ_ε^2, $\mathbf{p}$, $\boldsymbol{\beta}$, and t_0, as well as the random parameter $\boldsymbol{\mu}$. Hence, the hybrid CRB should also be considered. To proceed, we collect all these parameters into the following vector:

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 16 of 28

$$
\begin{aligned}
\boldsymbol{\eta}^{(\mathrm{b})} &= \begin{bmatrix} \sigma_\varepsilon^2 & \mathbf{p}^{\mathrm{T}} & (\ \mathrm{Re}\{\boldsymbol{\beta}\})^{\mathrm{T}} & (\ \mathrm{Im}\{\boldsymbol{\beta}\})^{\mathrm{T}} & t_0 & \boldsymbol{\mu} \end{bmatrix}^{\mathrm{T}} \\
&= \begin{bmatrix} \sigma_\varepsilon^2 & \boldsymbol{\eta}'^{(\mathrm{b})} \end{bmatrix}^{\mathrm{T}}
\end{aligned}
\tag{74}
$$

where

$$
\boldsymbol{\eta}'^{(\mathrm{b})} = \begin{bmatrix} \mathbf{p}^{\mathrm{T}} & (\ \mathrm{Re}\{\boldsymbol{\beta}\})^{\mathrm{T}} & (\ \mathrm{Im}\{\boldsymbol{\beta}\})^{\mathrm{T}} & t_0 & \boldsymbol{\mu} \end{bmatrix}^{\mathrm{T}}
\tag{75}
$$

Likewise, the FIM for vector $\boldsymbol{\eta}'^{(\mathrm{b})}$ can be approximately expressed as

$$
\begin{aligned}
\mathbf{FIM}\left(\boldsymbol{\eta}'^{(\mathrm{b})}\right) &= \frac{2}{\sigma_\varepsilon^2} \cdot \ \mathrm{Re}\left\{ \mathbf{T}_{\boldsymbol{\eta}'^{(\mathrm{b})}}^{\mathrm{H}} \mathbf{T}_{\boldsymbol{\eta}'^{(\mathrm{b})}} \right\} \\
&\quad + \begin{bmatrix} \mathbf{OO} \\ \mathbf{O}\boldsymbol{\Omega}^{-1} \end{bmatrix}
\end{aligned}
\tag{76}
$$

where

$$
\begin{aligned}
\mathbf{T}_{\boldsymbol{\eta}'^{(\mathrm{b})}} &= \frac{\partial \overline{\mathbf{x}}_0}{\partial \boldsymbol{\eta}'^{(\mathrm{b})\mathrm{T}}} = \begin{bmatrix} \dfrac{\partial \overline{\mathbf{x}}_0}{\partial \mathbf{p}^{\mathrm{T}}} & \dfrac{\partial \overline{\mathbf{x}}_0}{\partial (\ \mathrm{Re}\{\boldsymbol{\beta}\})^{\mathrm{T}}} & \dfrac{\partial \overline{\mathbf{x}}_0}{\partial (\ \mathrm{Im}\{\boldsymbol{\beta}\})^{\mathrm{T}}} & \dfrac{\partial \overline{\mathbf{x}}_0}{\partial t_0} & \dfrac{\partial \overline{\mathbf{x}}_0}{\partial \boldsymbol{\mu}^{\mathrm{T}}} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{T_p} & \mathbf{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}} & \mathbf{T}_{\mathrm{Im}\{\boldsymbol{\beta}\}} & \mathbf{T}_{t_0} & \mathbf{T}_{\boldsymbol{\mu}} \end{bmatrix}
\end{aligned}
\tag{77}
$$

Note that matrix $\mathbf{FIM}(\boldsymbol{\eta}'^{(\mathrm{b})})$ is also computed at the nominal value $\boldsymbol{\mu}_0$. According to the analysis in Subsection 6.1, we can still obtain an asymptotically valid CRB. In addition, all the sub-matrices in (77) are given explicitly in (70), except for $\mathbf{T}_{t_0}$. Therefore, we need only derive an expression for $\mathbf{T}_{t_0}$. From (62) and (63), it can be checked that

$$
\begin{aligned}
\mathbf{T}_{t_0} &= \frac{\partial \overline{\mathbf{x}}_0}{\partial t_0^{\mathrm{T}}} = -\mathrm{j} \cdot \ \mathrm{diag}[\boldsymbol{\beta} \otimes \mathbf{1}_{MK \times 1}] \\
&\quad \cdot \ \mathrm{diag}[\mathbf{1}_{N \times 1} \otimes (\overline{\mathbf{s}} \odot \boldsymbol{\omega}) \otimes \mathbf{1}_{M \times 1}] \\
&\quad \cdot (\mathbf{A}(\mathbf{p}, \boldsymbol{\mu}_0) \mathbf{1}_{K \times 1})
\end{aligned}
\tag{78}
$$

where $\boldsymbol{\omega} = \begin{bmatrix} \omega_1 & \omega_2 & \cdots & \omega_K \end{bmatrix}^{\mathrm{T}}$. Inserting (70), (77), and (78) into (76) produces

where

$$
\begin{cases}
\boldsymbol{Z}_1^{(\mathrm{b})} = \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_p \ , \ \boldsymbol{Z}_2^{(\mathrm{b})} = \left[[1 \ \ \mathrm{j}] \otimes (\boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}) \ \middle| \ \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_{t_0} \ \middle| \ \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_{\boldsymbol{\mu}} \right] \\[4pt]
\boldsymbol{Z}_3^{(\mathrm{b})} = \begin{bmatrix} \begin{bmatrix} 1 & \mathrm{j} \\ -\mathrm{j} & 1 \end{bmatrix} \otimes (\boldsymbol{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}) & \begin{bmatrix} 1 \\ -\mathrm{j} \end{bmatrix} \otimes (\boldsymbol{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} \boldsymbol{T}_{t_0}) & \begin{bmatrix} 1 \\ -\mathrm{j} \end{bmatrix} \otimes (\boldsymbol{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}} \boldsymbol{T}_{\boldsymbol{\mu}}) \\ [1 \ \mathrm{j}] \otimes (\boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}) & \boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_{t_0} & \boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_{\boldsymbol{\mu}} \\ [1 \ \mathrm{j}] \otimes (\boldsymbol{T}_{\boldsymbol{\mu}}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}) & \boldsymbol{T}_{\boldsymbol{\mu}}^{\mathrm{H}} \boldsymbol{T}_{t_0} & \boldsymbol{T}_{\boldsymbol{\mu}}^{\mathrm{H}} \boldsymbol{T}_{\boldsymbol{\mu}} \end{bmatrix}
\end{cases}
\tag{80}
$$

In Appendix 5, the details of calculating the matrices in (80) are provided.

With the application of the partitioned matrix inversion formula, the CRB matrix for position vector $\mathbf{p}$ can be expressed as

$$
\begin{aligned}
\mathbf{CRB}^{(\mathrm{b})}(\mathbf{p}) = \frac{\sigma_\varepsilon^2}{2} \cdot \Bigg(& \left(\mathrm{Re}\left\{ \mathbf{Z}_1^{(\mathrm{b})} \right\} \right)^{-1} + \left(\mathrm{Re}\left\{ \mathbf{Z}_1^{(\mathrm{b})} \right\} \right)^{-1} \cdot \ \mathrm{Re}\left\{ \mathbf{Z}_2^{(\mathrm{b})} \right\} \\
& \cdot \left(\mathrm{Re}\left\{ \mathbf{Z}_3^{(\mathrm{b})} \right\} - \mathrm{Re}\left\{ \mathbf{Z}_2^{(\mathrm{b})\mathrm{H}} \right\} \right. \\
& \left. \cdot \left(\mathrm{Re}\left\{ \mathbf{Z}_1^{(\mathrm{b})} \right\} \right)^{-1} \cdot \ \mathrm{Re}\left\{ \mathbf{Z}_2^{(\mathrm{b})} \right\} \right. \\
& \left. + \begin{bmatrix} \mathbf{OO} \\ \mathbf{O}\sigma_\varepsilon^2 \boldsymbol{\Omega}^{-1}/2 \end{bmatrix} \right)^{-1} \\
& \times \ \mathrm{Re}\left\{ \mathbf{Z}_2^{(\mathrm{b})\mathrm{H}} \right\} \cdot \left(\mathrm{Re}\left\{ \mathbf{Z}_1^{(\mathrm{b})} \right\} \right)^{-1} \Bigg)
\end{aligned}
\tag{81}
$$

where the superscript "b" indicates the case of a known signal waveform, and the trace of $\mathbf{CRB}^{(\mathrm{b})}(\mathbf{p})$ is the minimum achievable MSE for any unbiased localization method.

7 Simulation results

This section presents a set of Monte Carlo simulations to examine the behavior of the proposed robust DPD methods. The RMSE of position estimate is employed to assess and compare the performance. All the simulation results are averaged over 2000 independent runs. We compare our algorithms with the algorithms in [27], and the traditional two-step localization algorithms, as well as the CRB for unknown and known signal waveforms. Besides, all the experiments are conducted for the 3-D localization.

$$
\begin{aligned}
\mathbf{FIM}(\boldsymbol{\eta}'^{(\mathrm{b})}) &= \frac{2}{\sigma_\varepsilon^2} \cdot \mathrm{Re}\left\{ \begin{bmatrix} \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_p & \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\beta\}} & \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Im}\{\beta\}} & \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_{t_0} & \boldsymbol{T}_p^{\mathrm{H}} \boldsymbol{T}_{\mu} \\ \boldsymbol{T}_{\mathrm{Re}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_p & \boldsymbol{T}_{\mathrm{Re}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\beta\}} & \boldsymbol{T}_{\mathrm{Re}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Im}\{\beta\}} & \boldsymbol{T}_{\mathrm{Re}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{t_0} & \boldsymbol{T}_{\mathrm{Re}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{\mu} \\ \boldsymbol{T}_{\mathrm{Im}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_p & \boldsymbol{T}_{\mathrm{Im}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\beta\}} & \boldsymbol{T}_{\mathrm{Im}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Im}\{\beta\}} & \boldsymbol{T}_{\mathrm{Im}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{t_0} & \boldsymbol{T}_{\mathrm{Im}\{\beta\}}^{\mathrm{H}} \boldsymbol{T}_{\mu} \\ \boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_p & \boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\beta\}} & \boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Im}\{\beta\}} & \boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_{t_0} & \boldsymbol{T}_{t_0}^{\mathrm{H}} \boldsymbol{T}_{\mu} \\ \boldsymbol{T}_{\mu}^{\mathrm{H}} \boldsymbol{T}_p & \boldsymbol{T}_{\mu}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Re}\{\beta\}} & \boldsymbol{T}_{\mu}^{\mathrm{H}} \boldsymbol{T}_{\mathrm{Im}\{\beta\}} & \boldsymbol{T}_{\mu}^{\mathrm{H}} \boldsymbol{T}_{t_0} & \boldsymbol{T}_{\mu}^{\mathrm{H}} \boldsymbol{T}_{\mu} \end{bmatrix} \right\} + \begin{bmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{\Omega}^{-1} \end{bmatrix} \\
&= \frac{2}{\sigma_\varepsilon^2} \cdot \mathrm{Re}\left\{ \begin{bmatrix} \boldsymbol{Z}_1^{(\mathrm{b})} & \boldsymbol{Z}_2^{(\mathrm{b})} \\ \boldsymbol{Z}_2^{(\mathrm{b})\mathrm{H}} & \boldsymbol{Z}_3^{(\mathrm{b})} \end{bmatrix} \right\} + \begin{bmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{\Omega}^{-1} \end{bmatrix}
\end{aligned}
\tag{79}
$$

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 17 of 28

7.1 Simulation results for case of unknown signal waveform

In this subsection, the algorithm presented in Section 4 is evaluated and its estimation accuracy is compared with that of the first algorithm in [27], which assumes the signal waveform is unknown. Both of these algorithms are denoted as algorithm I in the figures below. In addition, the two-step localization algorithm used for comparison here is realized using the robust Bayesian algorithm [48] to estimate the DOAs in the first step and exploiting the Taylor series (TS) algorithm [59] to locate the source in the second phase. If the unknowns to be found are DOAs, the array manifold in (1)–(3) should be modeled as a function of direction, as in [1–3]. This is not difficult to achieve because position vector $\mathbf{p}$ is a function of direction. The array model mismatch considered here is caused by sensor location perturbations. The sensor position errors are zero-mean Gaussian random variables in both the x and y directions, and are independent and identically distributed (IID) from sensor to sensor and array to array. The standard deviation of the position errors is denoted by σ_L.

In the first set of experiments, we consider three base stations placed at [0, 2500, 0] m, [0, 0, 0] m, and [0, – 2500, 0] m and a single emitter located at [1500, 2000, 2000] m. The transmitted waveforms are realizations of a narrowband Gaussian random process, and are unknown to the receivers. Each base station is equipped with a uniform circular array (UCA). The channel attenuation magnitude is fixed at 1, and the channel phase is selected at random from a uniform distribution over [−π, π). Additionally, unless stated otherwise, we use the following settings: (1) $K = 128$ samples; (2) SNR = 10 dB; (3) $M = 5$ sensors; (4) $\sigma_L = 0.03\lambda$ (where λ is the wavelength of the carrier signal); and (5) array radius equal to the wavelength. Figures 1, 2, 3, 4, and 5 display the RMSEs of the localization methods as functions of the SNR of the emitter signal, number of snapshots K, standard deviation of the sensor position errors σ_L, number of array elements M, and ratio of array radius to wavelength, respectively.

It can be observed from Figs. 1, 2, 3, 4, and 5 that algorithm I in this paper outperforms algorithm I in [27] in terms of the estimation accuracy. The performance gap is especially noticeable for high SNRs and large sensor position errors. This is because, in these cases, the array model errors dominate the performance. Our technique accounts for these errors and incorporates the prior statistics of the array perturbations for source localization. The results in Figs. 1, 2, 3, 4, and 5 show that, when the SNR decreases and the array aperture increases, the proposed algorithm gives only a minor performance improvement. This is because, for lower SNRs and large array apertures, the measurement noise is the dominant error source, and the two algorithms give similar performance. In addition, the estimation accuracy of our algorithm achieves the CRB provided by (73) at moderate noise and error levels. Finally, it is worth noting that the DPD estimators perform significantly better than the two-step localization method. This improvement has been explained in the literature.
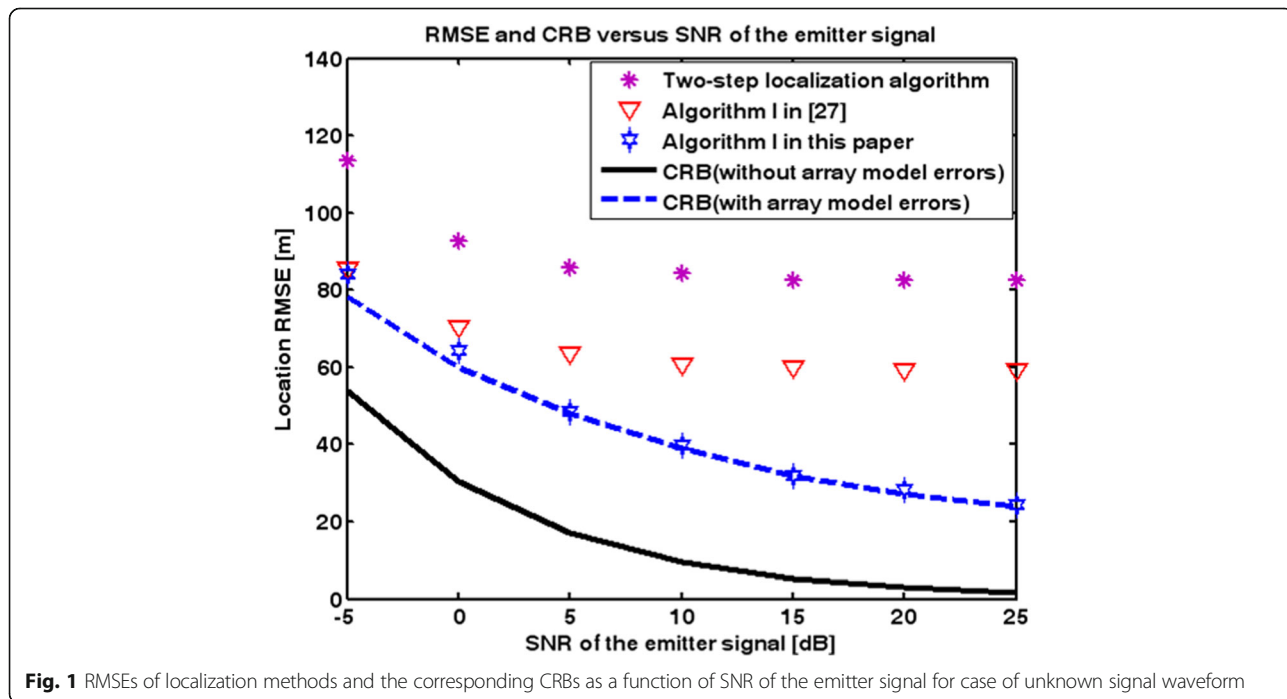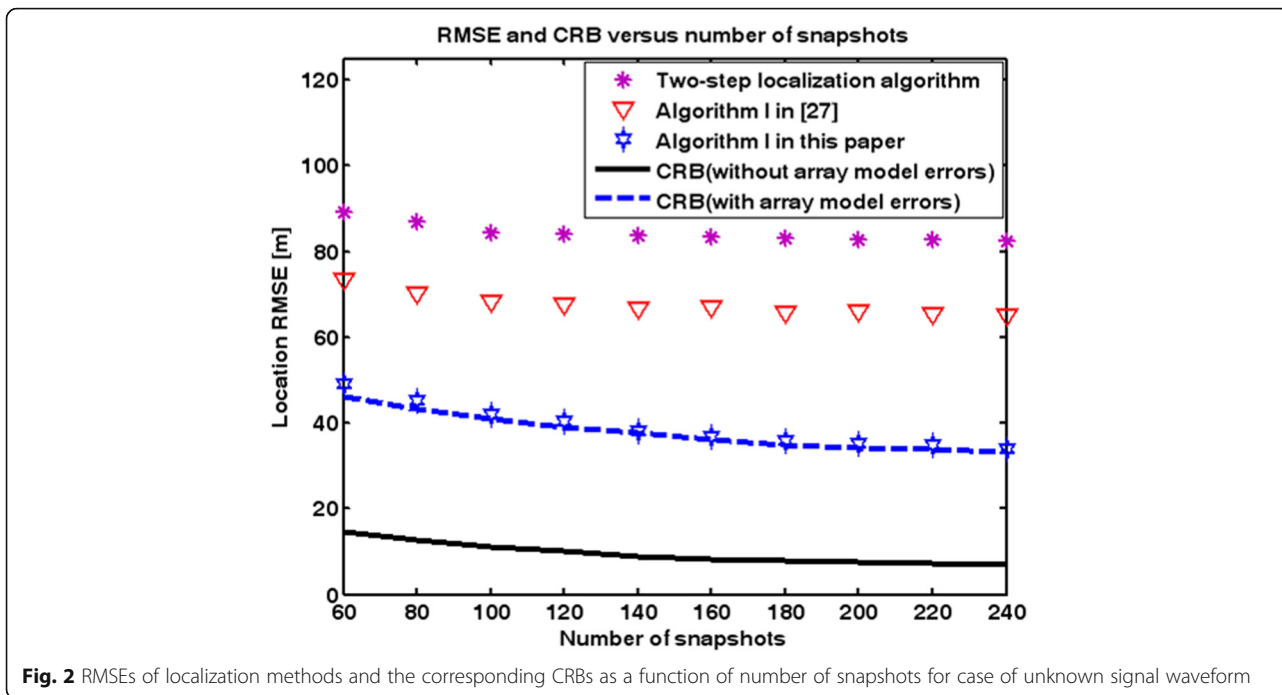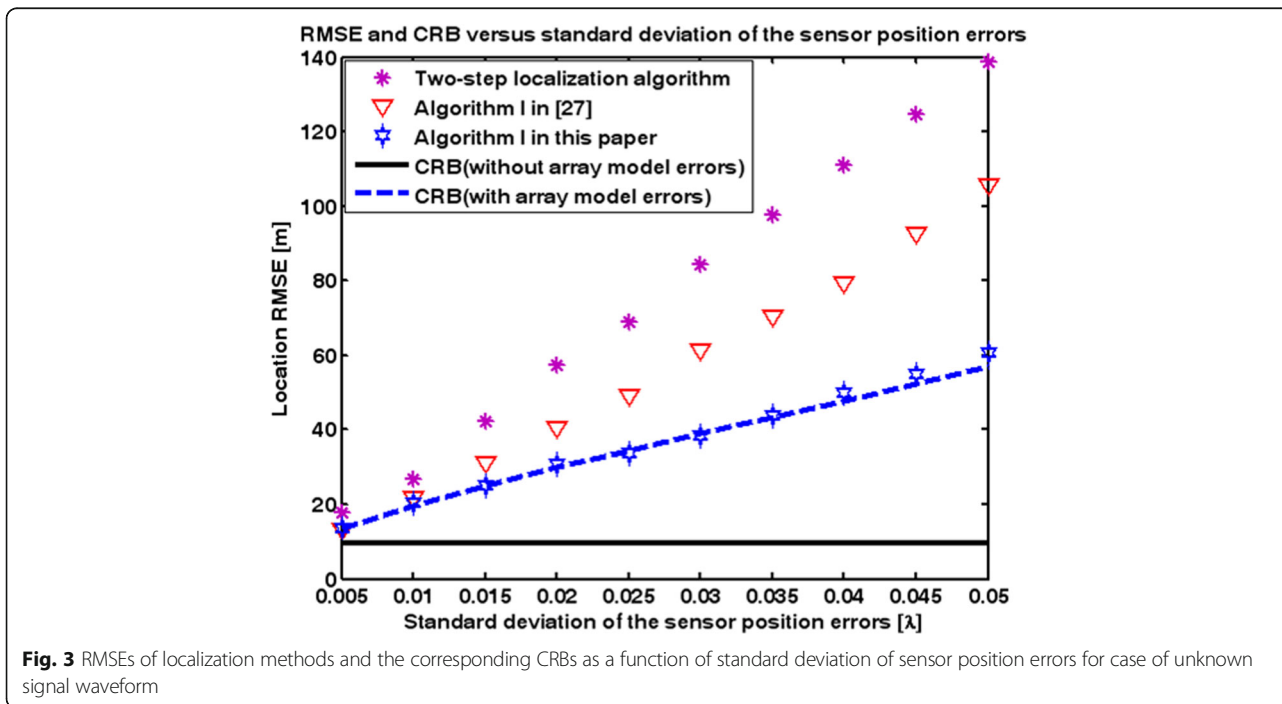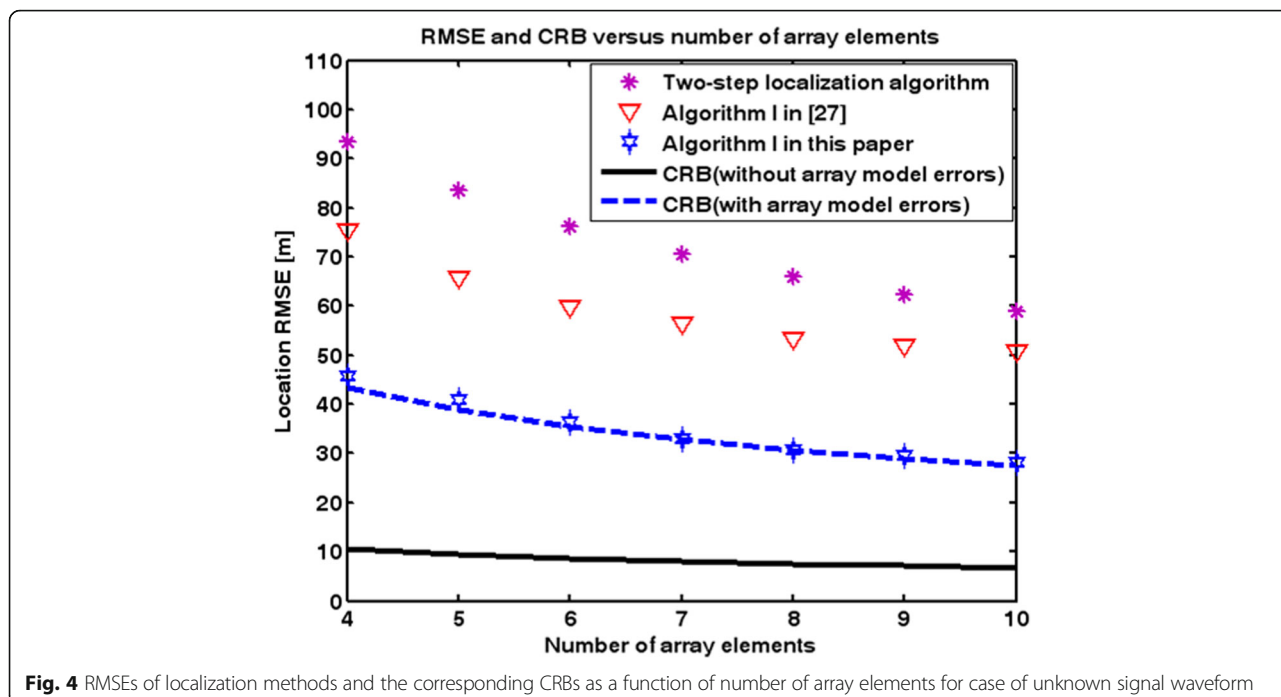

Fig. 1 RMSEs of localization methods and the corresponding CRBs as a function of SNR of the emitter signal for case of unknown signal waveform

Fig. 2 RMSEs of localization methods and the corresponding CRBs as a function of number of snapshots for case of unknown signal waveform

In the second experiment, we use the same simulation settings with a varying emitter location. The source coordinate is set as [1500, 2000, 2000] + $\alpha \cdot$ [200, 200, 200] m, where α ranges from 0 to 10. The distance between the emitter and the receivers increases as α increases. Figure 6 illustrates the RMSEs of the localization methods with respect to α.

From Fig. 6, we can draw similar conclusions similar to those for Figs. 1, 2, 3, 4, and 5; they are not repeated here for reasons of brevity. We only emphasize that the RMSE improvement from algorithm I in this paper over algorithm I in [27] is more significant as α increases. Hence, the gain in localization accuracy of our algorithm increases as the source moves farther away from the receivers.

Fig. 3 RMSEs of localization methods and the corresponding CRBs as a function of standard deviation of sensor position errors for case of unknown signal waveform

Fig. 4 RMSEs of localization methods and the corresponding CRBs as a function of number of array elements for case of unknown signal waveform

7.2 Simulation results for case of known signal waveform

In this subsection, the positioning accuracy of the algorithm given in Section 5 is compared with that of the second algorithm given in [27], which assumes the signal waveform is known. These are referred to as algorithm II in the following figures. Additionally, for a fair comparison, in the simulated two-step localization

algorithm, the DOAs are estimated by the algorithm presented in [56], which assumes a known waveform. As stated above, when the unknowns to be determined are DOAs, we need to express the array manifold in (1)–(3) as a function of direction, following [1–3]. This can be easily achieved because position vector **p** is a function of DOA. The array model error results from sensor gain

Fig. 5 RMSEs of localization methods and the corresponding CRBs as a function of ratio of array radius to wavelength for case of unknown signal waveform

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 20 of 28
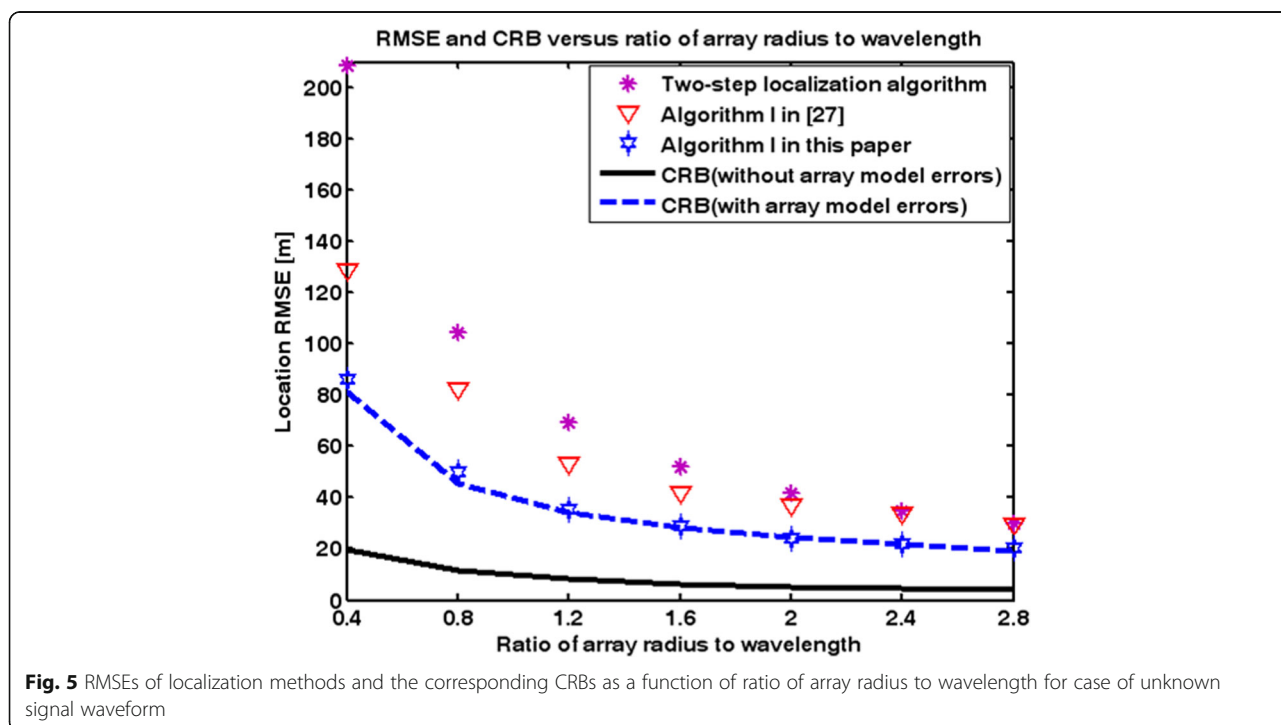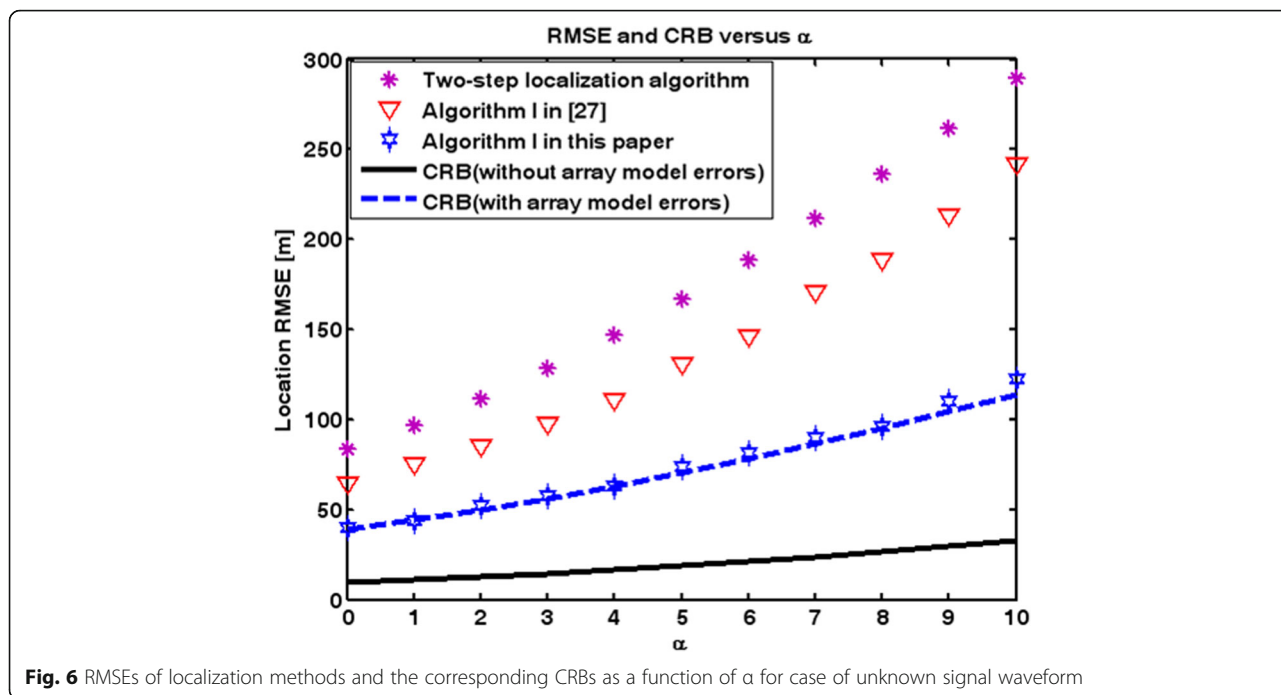
Fig. 6 RMSEs of localization methods and the corresponding CRBs as a function of α for case of unknown signal waveform

and phase uncertainties, corresponding to case 2 in Section 4 of [39]. Moreover, both gain and phase perturbations are independent and identically distributed zero-mean Gaussian random variables. The standard deviations of the gain and phase errors are denoted as σ_G and σ_P, respectively. We assume $\sigma_G = 0.01\sigma_P$ hereafter; thus, if σ_P changes, σ_G alters accordingly.

In the first set of experiments, the location system consists of three base stations located at [500, 2500, 0] m, [0, 0, 0] m, and [500, − 2500, 0] m. The emitter is positioned at [1500, 500, 2500] m. The signal waveforms are generated in the same way as described in Subsection 7.1, but they are known to the receivers. Each base station is equipped with a UCA. The channel attenuation magnitude is fixed at 1, and the channel phase is selected at random from a uniform distribution over [−π, π). Additionally, unless otherwise specified, we adopt the following simulation parameters: (1) $K = 128$ samples; (2) SNR = 10 dB; (3) $M = 6$ sensors; (4) $\sigma_P = 0.1$rad; and (5) array radius equal to wavelength. Figures 7, 8, 9, 10, and 11 plot the RMSEs of the localization methods against the SNR of the emitter signal, number of snapshots K, standard deviation of the sensor phase errors σ_P, number of array elements M, and ratio of array radius to wavelength, respectively.

In Figures 7, 8, 9, 10, and 11, the RMSEs clearly demonstrate the superior performance of algorithm II in this paper over algorithm II in [27] and the two-step localization algorithm. Moreover, the performance improvement is more pronounced as the SNR and sensor model errors increase and the array aperture decreases.

The impact of array model mismatch is effectively mitigated by our algorithm. Additionally, the proposed algorithm yields a solution that attains the CRB accuracy provided by (81) at moderate noise and error levels.
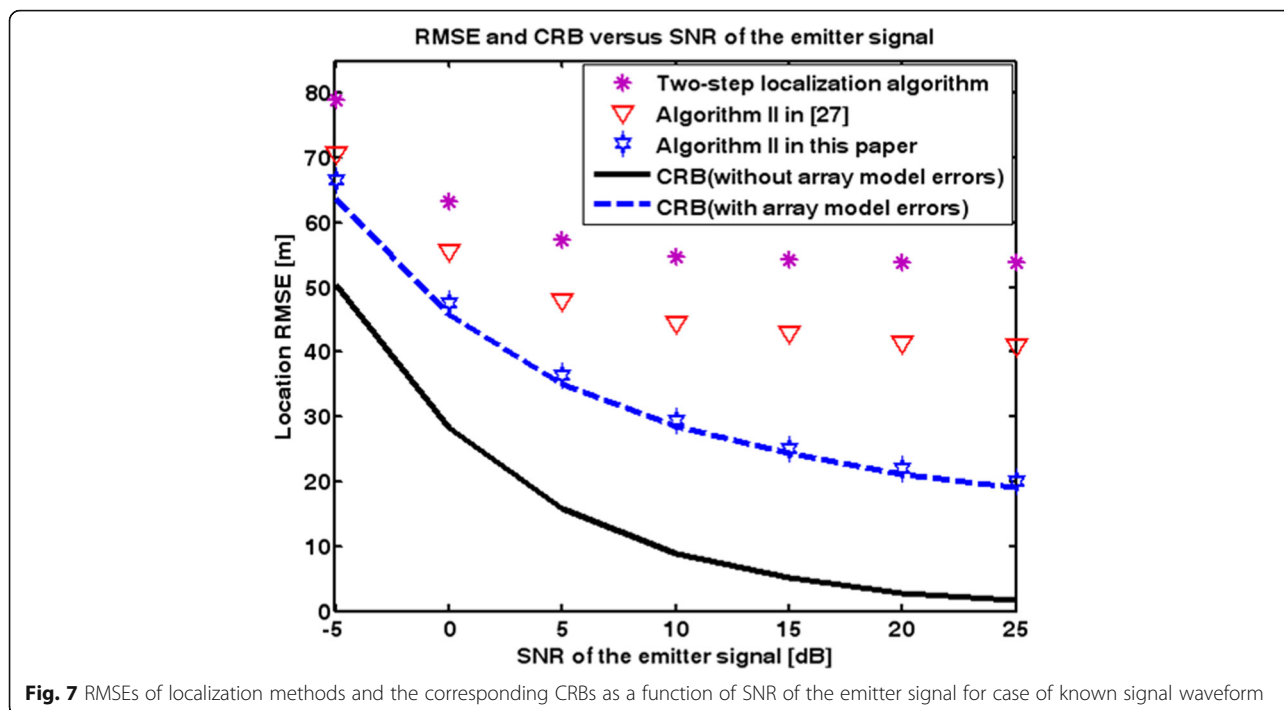
In the second experiment, the same simulation parameters are used, but the source location changes. The emitter position coordinates are set as [1500, 500, 2500] + α·[200, 200, 200] m, where α varies from 0 to 10. Again, the source becomes farther away from the receivers as α increases. Figure 12 shows the RMSEs of the localization methods with respect to α.

The performance gain of the proposed algorithm over the other algorithms is corroborated by the results in Fig. 12. Moreover, the RMSE improvement of the new algorithm over algorithm II in [27] becomes stronger as the source moves farther away from the receivers. This observation is consistent with the results in Fig. 6.

7.3 Running time of the localization methods
In this subsection, the runtime of all the localization methods considered in the experiments is compared. All the simulations were implemented using MATLAB R2016b on a ThinkPad laptop equipped with Intel Core i7–7500 CPU and 8 GB RAM.

First, we adopt the simulation settings used to produce Fig. 2. Figure 13 depicts the average runtime of the considered localization methods as a function of snapshot number. Second, the simulation parameters used to produce Fig. 8 were applied. In Fig. 14, the average runtime versus the snapshot number is compared for the considered localization methods.

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 21 of 28

Fig. 7 RMSEs of localization methods and the corresponding CRBs as a function of SNR of the emitter signal for case of known signal waveform
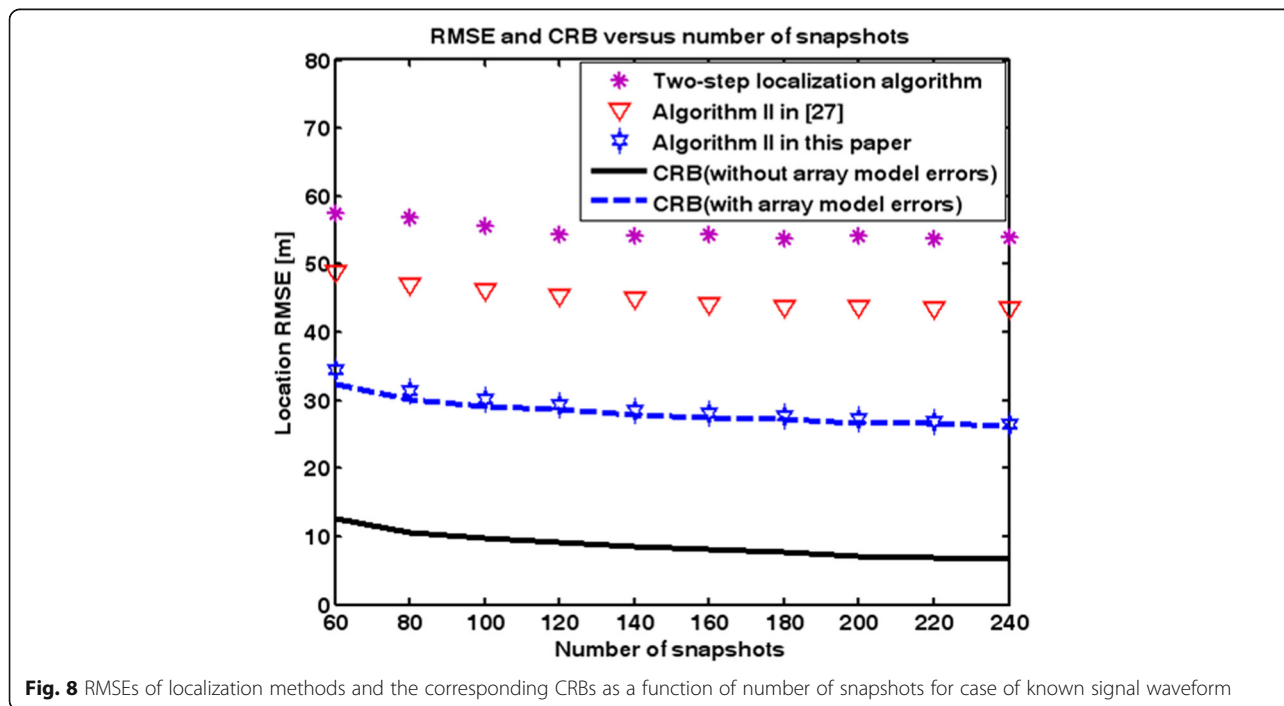
It is easily observed from Figs. 13 and 14 that the runtime of the proposed DPD methods is significantly higher than that of the other methods. This is because the model errors are considered and there are too many nuisance variables to be optimized besides the source position vector. The two-step localization method has the fastest runtime among the compared methods, mainly because it is a decentralized method.

8 Discussion

From the simulation results described above, we can observe that the proposed algorithms can effectively mitigate the effects of array model mismatch, and they
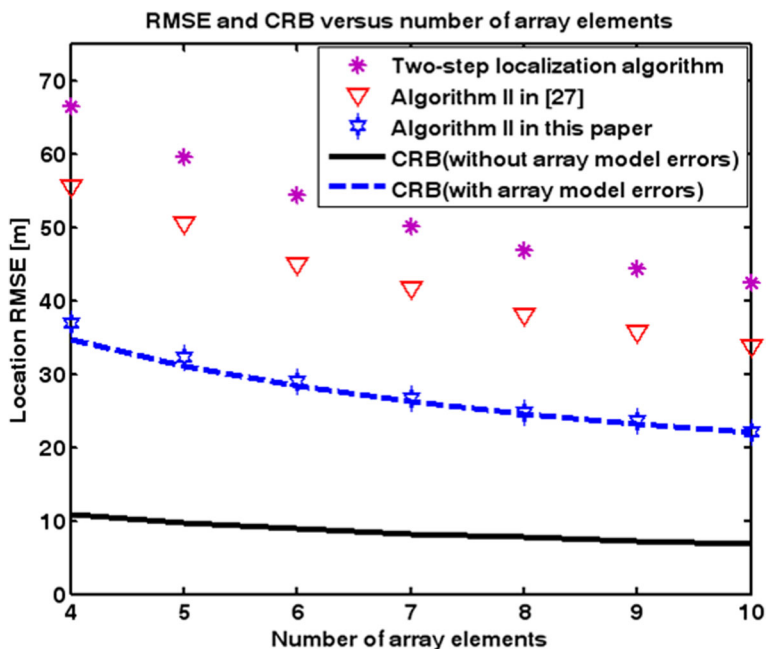

Fig. 8 RMSEs of localization methods and the corresponding CRBs as a function of number of snapshots for case of known signal waveform

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 22 of 28
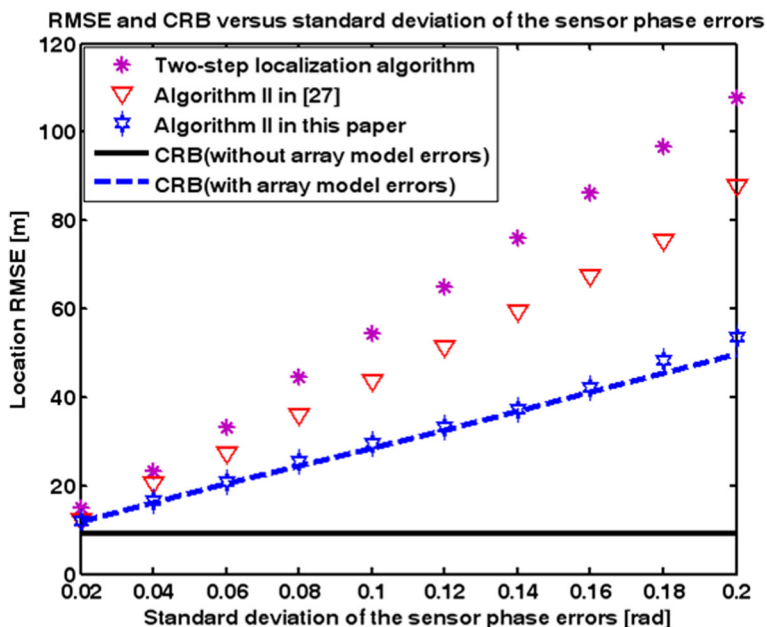
Fig. 9 RMSEs of localization methods and the corresponding CRBs as a function of standard deviation of sensor phase errors for case of known signal waveform

outperform the algorithms in [27] obviously for high SNRs and large sensor position errors. Moreover, the gain in localization accuracy of our algorithms increases as the source moves farther away from the receivers. The estimation performance of the developed algorithms can attain the corresponding CRB at moderate noise and error levels. However, the complexity of the proposed algorithms are higher than that of the algorithms in [27] because they account for the array model errors and require more computational procedure. Finally, it needs to be mentioned

Fig. 10 RMSEs of localization methods and the corresponding CRBs as a function of number of array elements for case of known signal waveform

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 23 of 28
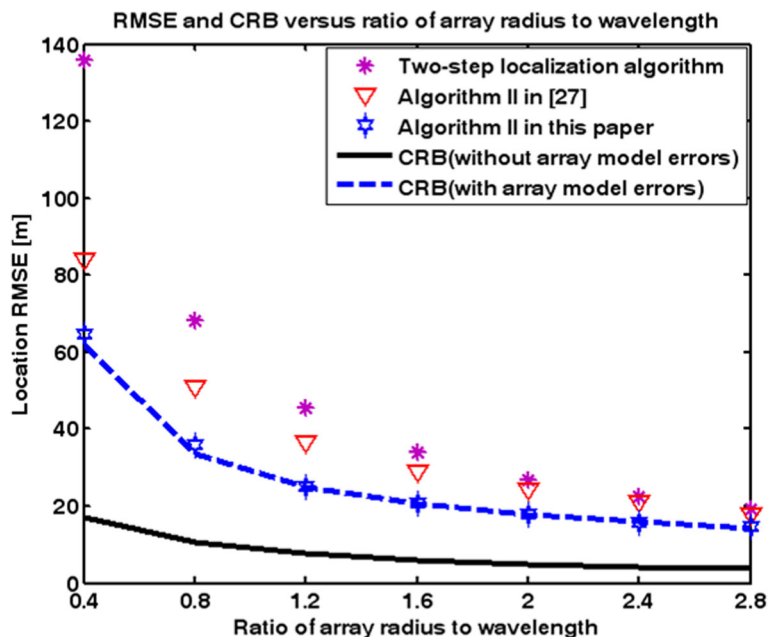
Fig. 11 RMSEs of localization methods and the corresponding CRBs as a function of ratio of array radius to wavelength for case of known signal waveform

that the present study only deals with the single-source situation. In our future work, we intend to extend the proposed methods to the multiple-source scenario.

9 Conclusions

This paper presents robust DPD methods that can reduce the negative effect of array model mismatch.

The idea behind the newly proposed technique is similar to that of the array error auto-calibration methods, which assume that certain prior statistical distribution of the array model errors is available. We consider two different localization cases, the case of a priori known signal waveform and the more realistic case where transmitted signals are unknown to the
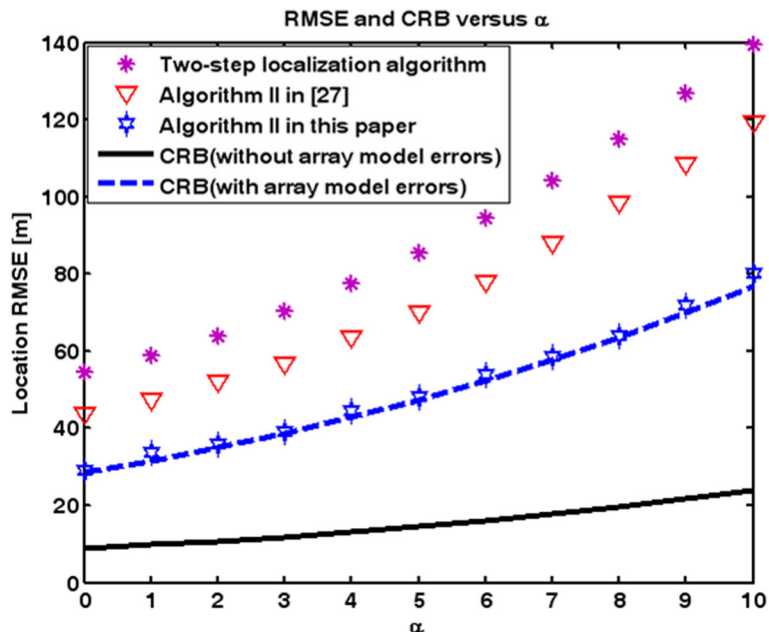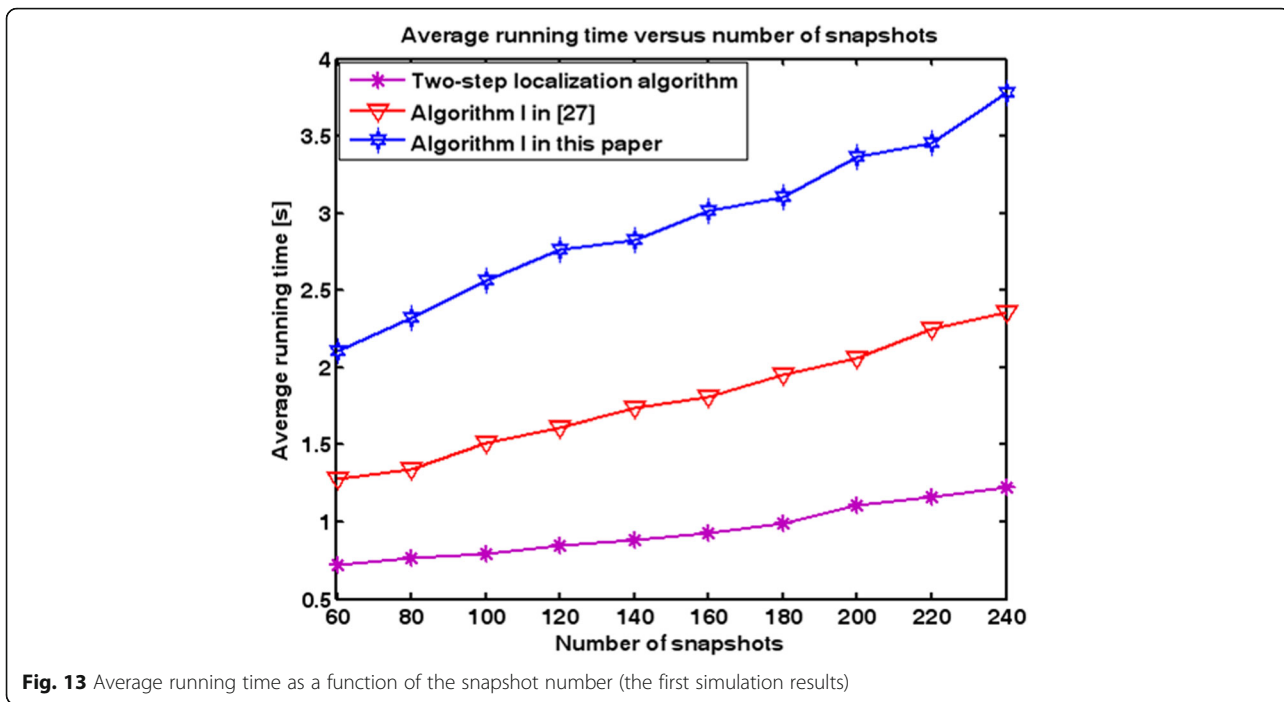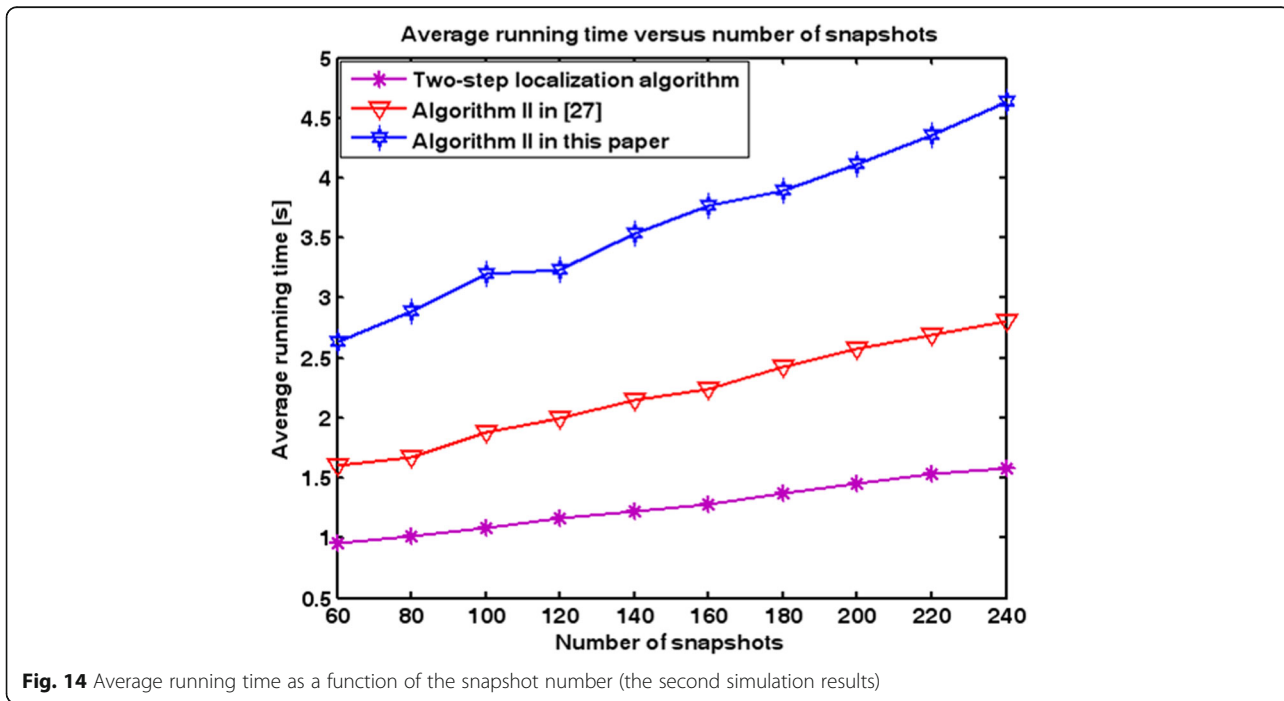

Fig. 12 RMSEs of localization methods and the corresponding CRBs as a function of α for case of known signal waveform

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 24 of 28

Fig. 13 Average running time as a function of the snapshot number (the first simulation results)

location systems. The corresponding MAP estimators for the two cases are formulated and two effective alternating minimization algorithms are developed to locate the source directly from the signals captured at several antenna arrays simultaneously. The proposed methods follow the Bayesian framework given in [47–50]. Besides, for the purpose of verifying the

asymptotic efficiency of the new methods, the CRB expressions for position estimation are also deduced for both unknown and known signal waveforms. Simulation results confirm that the proposed algorithms are able to provide the CRB accuracy and the effect of array model errors can be decreased considerably.

Fig. 14 Average running time as a function of the snapshot number (the second simulation results)

Wang et al. EURASIP Journal on Advances in Signal Processing (2018) 2018:38

Page 25 of 28

10 Appendix 1

10.1 Proof of Proposition 1

Assume the normalized eigenvectors of matrix $\hat{\mathbf{Z}}$ are $\hat{\mathbf{v}}_1$, $\hat{\mathbf{v}}_2$, $\cdots$, $\hat{\mathbf{v}}_n$. Applying the Hermitian matrix eigen-perturbation theory [60], we have

$$\begin{cases} \hat{\lambda}_j = \lambda_j + \delta\lambda_j^{(1)} + \delta\lambda_j^{(2)} + \cdots \\ \hat{\mathbf{v}}_j = \left(1 - \frac{1}{2} \cdot \sum_{\substack{i=1 \\ i \neq j}}^{n} \left| \xi_{ji}^{(1)} \right|^2 \right) \mathbf{v}_j + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \mathbf{v}_i + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(2)} \mathbf{v}_i + \cdots \end{cases} \tag{82}$$

where $\delta\lambda_j^{(1)}$ and $\xi_{ji}^{(1)}$ are the first-order perturbation terms, i.e., $\delta\lambda_j^{(1)} = O(\|\boldsymbol{\delta Z}\|_2)$ and $\xi_{ji}^{(1)} = O(\|\boldsymbol{\delta Z}\|_2)$, and $\delta\lambda_j^{(2)}$ and $\xi_{ji}^{(2)}$ are the second-order perturbation terms, i.e., $\delta\lambda_j^{(2)} = O(\|\boldsymbol{\delta Z}\|_2^2)$ and $\xi_{ji}^{(2)} = O(\|\boldsymbol{\delta Z}\|_2^2)$. It follows from the matrix eigen-equation that

$$\hat{\mathbf{Z}}\hat{\mathbf{v}}_j = \hat{\lambda}_j \hat{\mathbf{v}}_j \Leftrightarrow (\mathbf{Z} + \boldsymbol{\delta Z}) \left(\left(1 - \frac{1}{2} \cdot \sum_{\substack{i=1 \\ i \neq j}}^{n} \left| \xi_{ji}^{(1)} \right|^2 \right) \mathbf{v}_j \right.$$
$$\left. + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \mathbf{v}_i + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(2)} \mathbf{v}_i + \cdots \right)$$
$$= \left(\lambda_j + \delta\lambda_j^{(1)} + \delta\lambda_j^{(2)} + \cdots \right) \left(\left(1 - \frac{1}{2} \cdot \sum_{\substack{i=1 \\ i \neq j}}^{n} \left| \xi_{ji}^{(1)} \right|^2 \right) \mathbf{v}_j \right.$$
$$\left. + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \mathbf{v}_i + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(2)} \mathbf{v}_i + \cdots \right) \tag{83}$$

Comparing the first-order perturbation terms between both sides of (83), it can be observed that

$$\sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \lambda_i \mathbf{v}_i + \boldsymbol{\delta Z} \cdot \mathbf{v}_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \lambda_j \mathbf{v}_i + \mathbf{v}_j \cdot \delta\lambda_j^{(1)} \tag{84}$$

Premultiplying $\mathbf{v}_j^{\mathrm{H}}$ and $\mathbf{v}_i^{\mathrm{H}} (i \neq j)$ on both sides of (84) leads to

$$\begin{cases} \delta\lambda_j^{(1)} = \mathbf{v}_j^{\mathrm{H}} \cdot \boldsymbol{\delta Z} \cdot \mathbf{v}_j \\ \xi_{ji}^{(1)} = \dfrac{\mathbf{v}_i^{\mathrm{H}} \cdot \boldsymbol{\delta Z} \cdot \mathbf{v}_j}{\lambda_j - \lambda_i} \end{cases} \tag{85}$$

Furthermore, comparing the second-order perturbation terms between both sides of (83), it follows that

$$\sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(2)} \lambda_i \mathbf{v}_i + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \cdot \boldsymbol{\delta Z} \cdot \mathbf{v}_i$$
$$= \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(2)} \lambda_j \mathbf{v}_i + \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \cdot \delta\lambda_j^{(1)} \cdot \mathbf{v}_i + \delta\lambda_j^{(2)} \cdot \mathbf{v}_j \tag{86}$$

Premultiplying $\mathbf{v}_j^{\mathrm{H}}$ on both sides of (86) and using (85) yields

$$\delta\lambda_j^{(2)} = \sum_{\substack{i=1 \\ i \neq j}}^{n} \xi_{ji}^{(1)} \mathbf{v}_j^{\mathrm{H}} \cdot \boldsymbol{\delta Z} \cdot \mathbf{v}_i$$
$$= \mathbf{v}_j^{\mathrm{H}} \cdot \boldsymbol{\delta Z} \cdot \left(\sum_{\substack{i=1 \\ i \neq j}}^{n} \frac{\mathbf{v}_i \mathbf{v}_i^{\mathrm{H}}}{\lambda_j - \lambda_i} \right) \cdot \boldsymbol{\delta Z} \cdot \mathbf{v}_j$$
$$= \mathbf{v}_j^{\mathrm{H}} \cdot \boldsymbol{\delta Z} \cdot \mathbf{V}_j \cdot \boldsymbol{\delta Z} \cdot \mathbf{v}_j \tag{87}$$

Combining (82), (85) and (87) completes the proof.

11 Appendix 2

11.1 Proof of (33)

Assume that $\hat{\boldsymbol{\mu}}_n^{(i)}$ is the estimation of $\boldsymbol{\mu}_n$ at the ith iteration step. In order to obtain the updated estimate of $\boldsymbol{\mu}_n$ at the $(i+1)$th iteration, we need to substitute the first-order Taylor series expansion of $\mathbf{g}_n(\boldsymbol{\mu}_n)$ round $\hat{\boldsymbol{\mu}}_n^{(i)}$ into (31), which yields

$$\hat{\boldsymbol{\mu}}_n^{(i+1)} = \arg\min_{\boldsymbol{\mu}_n} \left\{ \left\| \begin{bmatrix} \overline{\boldsymbol{x}}_n \\ \frac{1}{\sqrt{2}} \cdot \boldsymbol{\Omega}_{n,0}^{-1/2} \boldsymbol{\mu}_{n,0} \end{bmatrix} - \boldsymbol{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)}) - \boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})(\boldsymbol{\mu}_n - \hat{\boldsymbol{\mu}}_n^{(i)}) \right\|_2^2 \right\} \tag{88}$$

Noting that $\boldsymbol{\mu}_n - \hat{\boldsymbol{\mu}}_n^{(i)}$ is a real vector, we can reformulate (88) as

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 26 of 28

$$
\hat{\boldsymbol{\mu}}_n^{(i+1)} = \arg\min_{\mu_n} \left\| \left[\begin{array}{c} \text{Re}\left\{ \left[\dfrac{\overline{x}_n}{\frac{1}{\sqrt{2}}\cdot \boldsymbol{\Omega}_{n,0}^{-1/2}\,\boldsymbol{\mu}_{n,0}} \right] - \boldsymbol{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)}) \right\} \\ \hline \text{Im}\left\{ \left[\dfrac{\overline{x}_n}{\frac{1}{\sqrt{2}}\cdot \boldsymbol{\Omega}_{n,0}^{-1/2}\,\boldsymbol{\mu}_{n,0}} \right] - \boldsymbol{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)}) \right\} \end{array} \right] - \left[\begin{array}{c} \text{Re}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \\ \hline \text{Im}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \end{array} \right] \cdot (\boldsymbol{\mu}_n - \hat{\boldsymbol{\mu}}_n^{(i)}) \right\|_2^2
\tag{89}
$$

It is obvious that (89) is a linear least-squares estimator, and its optimal solution satisfies

$$
\hat{\boldsymbol{\mu}}_n^{(i+1)} - \hat{\boldsymbol{\mu}}_n^{(i)} = \left(\left[\begin{array}{c} \text{Re}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \\ \text{Im}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \end{array} \right]^{\mathrm{T}} \cdot \left[\begin{array}{c} \text{Re}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \\ \text{Im}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \end{array} \right] \right)^{-1} \cdot \left[\begin{array}{c} \text{Re}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \\ \text{Im}\{\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\} \end{array} \right]^{\mathrm{T}} \cdot \left[\begin{array}{c} \text{Re}\left\{ \left[\frac{\overline{x}_n}{\frac{1}{\sqrt{2}}\cdot \boldsymbol{\Omega}_{n,0}^{-1/2}\,\boldsymbol{\mu}_{n,0}} \right] - \boldsymbol{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)}) \right\} \\ \hline \text{Im}\left\{ \left[\frac{\overline{x}_n}{\frac{1}{\sqrt{2}}\cdot \boldsymbol{\Omega}_{n,0}^{-1/2}\,\boldsymbol{\mu}_{n,0}} \right] - \boldsymbol{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)}) \right\} \end{array} \right]
$$

$$
= (\text{Re}\{(\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)}))^{\mathrm{H}}\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)})\})^{-1} \cdot \text{Re}\left\{ (\boldsymbol{G}_n(\hat{\boldsymbol{\mu}}_n^{(i)}))^{\mathrm{H}} \left(\left[\frac{\overline{x}_n}{\frac{1}{\sqrt{2}}\cdot \boldsymbol{\Omega}_{n,0}^{-1/2}\,\boldsymbol{\mu}_{n,0}} \right] - \boldsymbol{g}_n(\hat{\boldsymbol{\mu}}_n^{(i)}) \right) \right\}
\tag{90}
$$

which indicates that (33) holds true.

12 Appendix 3
12.1 Proof of Proposition 2
From the definition of $y_{\mathrm{m}}(\mathbf{z})$, it follows that

$$
f_{\mathrm{m}}(\mathbf{z}) = \max_y \{f(\mathbf{z},y)\} = f(\mathbf{z},y_{\mathrm{m}}(\mathbf{z}))
\tag{91}
$$

Then, the gradient of $f_{\mathrm{m}}(\mathbf{z})$ can be written as

$$
\mathbf{f}_{\mathrm{m}}(\mathbf{z}) = \frac{\partial f_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} = \dot{\mathbf{f}}_1(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) + \dot{f}_2(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) \cdot \frac{\partial y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}}
\tag{92}
$$

From (92), we can obtain the Hessian matrix of $f_{\mathrm{m}}(\mathbf{z})$ as

$$
\mathbf{F}_{\mathrm{m}}(\mathbf{z}) = \frac{\partial^2 f_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}\partial \mathbf{z}^{\mathrm{T}}} = \frac{\partial \mathbf{f}_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}^{\mathrm{T}}} = \ddot{\mathbf{F}}_{11}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) + \ddot{\mathbf{f}}_{21}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z}))
$$
$$
\cdot \left(\frac{\partial y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} \right)^{\mathrm{T}} + \frac{\partial y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} \cdot \ddot{\mathbf{f}}_{21}^{\mathrm{T}}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) + \ddot{f}_{22}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z}))
$$
$$
\cdot \frac{\partial y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} \cdot \left(\frac{\partial y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} \right)^{\mathrm{T}} + \dot{f}_2(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) \cdot \frac{\partial^2 y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}\partial \mathbf{z}^{\mathrm{T}}}
\tag{93}
$$

On the other hand, using the definition of $y_{\mathrm{m}}(\mathbf{z})$ again, we have

$$
\dot{f}_2(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) = \left. \frac{\partial f(\mathbf{z},y)}{\partial y} \right|_{y=y_{\mathrm{m}}(\mathbf{z})} = 0
\tag{94}
$$

Substituting (94) into (92) proves the first equality in (48). Taking the derivative with respect to $\mathbf{z}$ on both sides of (94) yields

$$
\ddot{\mathbf{f}}_{21}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) + \ddot{f}_{22}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z})) \cdot \frac{\partial y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} = \mathbf{O}_{m\times 1}
\tag{95}
$$

which implies

$$
\frac{\partial y_{\mathrm{m}}(\mathbf{z})}{\partial \mathbf{z}} = -\frac{\ddot{\mathbf{f}}_{21}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z}))}{\ddot{f}_{22}(\mathbf{z},y_{\mathrm{m}}(\mathbf{z}))}
\tag{96}
$$

The second equality in (48) can be proved by inserting (94) and (96) into (93). At this point, the proof of Proposition 2 is completed.

13 Appendix 4
13.1 Detailed derivation of matrices in (72)
It can be seen from (72) that, in order to calculate matrices $\mathbf{Z}_1^{(\mathrm{a})}$, $\mathbf{Z}_2^{(\mathrm{a})}$, and $\mathbf{Z}_3^{(\mathrm{a})}$, we must derive the explicit expressions for matrices $\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\mathbf{p}}$, $\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}$, $\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\overline{\mathbf{s}}\}}$, $\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\boldsymbol{\mu}}$, $\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}$, $\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\overline{\mathbf{s}}\}}$, $\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}}\mathbf{T}_{\boldsymbol{\mu}}$, $\mathbf{T}_{\text{Re}\{\overline{\mathbf{s}}\}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\overline{\mathbf{s}}\}}$, $\mathbf{T}_{\text{Re}\{\overline{\mathbf{s}}\}}^{\mathrm{H}}\mathbf{T}_{\boldsymbol{\mu}}$, and $\mathbf{T}_{\boldsymbol{\mu}}^{\mathrm{H}}\mathbf{T}_{\boldsymbol{\mu}}$.

Using (70) and after some algebraic manipulation, we have

$$
\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\mathbf{p}} = \sum_{n=1}^{N}\sum_{k=1}^{K} |\beta_n|^2 \cdot |\overline{s}_k|^2
$$
$$
\cdot \left(\frac{\partial \mathbf{a}_{n,k}'\left(\mathbf{p},\boldsymbol{\mu}_{n,0}\right)}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \cdot \frac{\partial \mathbf{a}_{n,k}'\left(\mathbf{p},\boldsymbol{\mu}_{n,0}\right)}{\partial \mathbf{p}^{\mathrm{T}}}
\tag{97}
$$

$$
\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}} = \sum_{k=1}^{K} |\overline{s}_k|^2 \left[\left(\frac{\partial \mathbf{a}_{1,k}'(\mathbf{p},\boldsymbol{\mu}_{1,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{a}_{1,k}'(\mathbf{p},\boldsymbol{\mu}_{1,0}) \left| \left(\frac{\partial \mathbf{a}_{2,k}'(\mathbf{p},\boldsymbol{\mu}_{2,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{a}_{2,k}'(\mathbf{p},\boldsymbol{\mu}_{2,0}) \right| \cdots \left| \left(\frac{\partial \mathbf{a}_{N,k}'(\mathbf{p},\boldsymbol{\mu}_{N,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{a}_{N,k}'(\mathbf{p},\boldsymbol{\mu}_{N,0}) \right]
$$
$$
\times \text{diag}[\boldsymbol{\beta}^*]
\tag{98}
$$

$$
\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\overline{s}\}} = \sum_{n=1}^{N} |\beta_n|^2 \left[\left(\frac{\partial \mathbf{a}_{n,1}'(\mathbf{p},\boldsymbol{\mu}_{n,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{a}_{n,1}'(\mathbf{p},\boldsymbol{\mu}_{n,0}) \left| \left(\frac{\partial \mathbf{a}_{n,2}'(\mathbf{p},\boldsymbol{\mu}_{n,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{a}_{n,2}'(\mathbf{p},\boldsymbol{\mu}_{n,0}) \right| \cdots \left| \left(\frac{\partial \mathbf{a}_{n,K}'(\mathbf{p},\boldsymbol{\mu}_{n,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \mathbf{a}_{n,K}'(\mathbf{p},\boldsymbol{\mu}_{n,0}) \right]
$$
$$
\times \text{diag}[\overline{\mathbf{s}}^*]
\tag{99}
$$

$$
\mathbf{T}_{\mathbf{p}}^{\mathrm{H}}\mathbf{T}_{\boldsymbol{\mu}} = \sum_{k=1}^{K} |\overline{s}_k|^2 \left[\left(\frac{\partial \mathbf{a}_{1,k}'(\mathbf{p},\boldsymbol{\mu}_{1,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \cdot \frac{\partial \mathbf{a}_{1,k}'(\mathbf{p},\boldsymbol{\mu}_{1,0})}{\partial \boldsymbol{\mu}_{1,0}^{\mathrm{T}}} \left| \left(\frac{\partial \mathbf{a}_{2,k}'(\mathbf{p},\boldsymbol{\mu}_{2,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \cdot \frac{\partial \mathbf{a}_{2,k}'(\mathbf{p},\boldsymbol{\mu}_{2,0})}{\partial \boldsymbol{\mu}_{2,0}^{\mathrm{T}}} \right| \cdots \left| \left(\frac{\partial \mathbf{a}_{N,k}'(\mathbf{p},\boldsymbol{\mu}_{N,0})}{\partial \mathbf{p}^{\mathrm{T}}} \right)^{\mathrm{H}} \cdot \frac{\partial \mathbf{a}_{N,k}'(\mathbf{p},\boldsymbol{\mu}_{N,0})}{\partial \boldsymbol{\mu}_{N,0}^{\mathrm{T}}} \right]
$$
$$
\times \text{diag}[|\boldsymbol{\beta}|^2]
\tag{100}
$$

$$
\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}} = \sum_{k=1}^{K} |\overline{s}_k|^2 \cdot \text{diag}[\| \mathbf{a}_{1,k}'(\mathbf{p},\boldsymbol{\mu}_{1,0})\|_2^2 \,|\, \| \mathbf{a}_{2,k}'(\mathbf{p},\boldsymbol{\mu}_{2,0})\|_2^2 \,|\, \cdots \,|\, \| \mathbf{a}_{N,k}'(\mathbf{p},\boldsymbol{\mu}_{N,0})\|_2^2]
\tag{101}
$$

$$
\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}}\mathbf{T}_{\text{Re}\{\overline{s}\}} = \text{diag}[\boldsymbol{\beta}]\cdot[\mathbf{A}_1^{\mathrm{H}}(\mathbf{p},\boldsymbol{\mu}_{1,0})\mathbf{A}_1(\mathbf{p},\boldsymbol{\mu}_{1,0})\overline{\mathbf{s}} \,|\, \mathbf{A}_2^{\mathrm{H}}(\mathbf{p},\boldsymbol{\mu}_{2,0})\mathbf{A}_2(\mathbf{p},\boldsymbol{\mu}_{2,0})\overline{\mathbf{s}} \,|\, \cdots \,|\, \mathbf{A}_N^{\mathrm{H}}(\mathbf{p},\boldsymbol{\mu}_{N,0})\mathbf{A}_N(\mathbf{p},\boldsymbol{\mu}_{N,0})\overline{\mathbf{s}}]^{\mathrm{H}}
\tag{102}
$$

$$
\mathbf{T}_{\text{Re}\{\boldsymbol{\beta}\}}^{\mathrm{H}}\mathbf{T}_{\boldsymbol{\mu}} = \sum_{k=1}^{K} |\overline{s}_k|^2 \cdot \text{blkdiag}[\beta_1 \mathbf{a}_{1,k}'^{\mathrm{H}}(\mathbf{p},\boldsymbol{\mu}_{1,0})\cdot \frac{\partial \mathbf{a}_{1,k}'(\mathbf{p},\boldsymbol{\mu}_{1,0})}{\partial \boldsymbol{\mu}_{1,0}^{\mathrm{T}}} \,|\, \beta_2 \mathbf{a}_{2,k}'^{\mathrm{H}}(\mathbf{p},\boldsymbol{\mu}_{2,0})\cdot \frac{\partial \mathbf{a}_{2,k}'(\mathbf{p},\boldsymbol{\mu}_{2,0})}{\partial \boldsymbol{\mu}_{2,0}^{\mathrm{T}}} \,|\, \cdots \,|\, \beta_N \mathbf{a}_{N,k}'^{\mathrm{H}}(\mathbf{p},\boldsymbol{\mu}_{N,0})\cdot \frac{\partial \mathbf{a}_{N,k}'(\mathbf{p},\boldsymbol{\mu}_{N,0})}{\partial \boldsymbol{\mu}_{N,0}^{\mathrm{T}}}]
\tag{103}
$$

$$\mathbf{T}^{\mathrm{H}}_{\mathrm{Re}\{\overline{\mathbf{s}}\}} \mathbf{T}_{\mathrm{Re}\{\overline{\mathbf{s}}\}} = \sum_{n=1}^{N} |\beta_n|^2$$
$$\cdot \operatorname{diag}\left[\left\| \mathbf{a}'_{n,1}(\mathbf{p}, \boldsymbol{\mu}_{1,0}) \right\|_2^2 \quad \left\| \mathbf{a}'_{n,2}(\mathbf{p}, \boldsymbol{\mu}_{2,0}) \right\|_2^2 \right.$$
$$\left. \cdots \quad \left\| \mathbf{a}'_{n,K}(\mathbf{p}, \boldsymbol{\mu}_{N,0}) \right\|_2^2 \right]$$

$$(104)$$

$$\mathbf{T}^{\mathrm{H}}_{\mathrm{Re}\{\overline{s}\}}\mathbf{T}_{\boldsymbol{\mu}} = \operatorname{diag}[\overline{s}] \cdot \left[|\beta_1|^2 \cdot (A_1(p,\mu_{1,0}))^{\mathrm{H}} \cdot \frac{\partial (A_1(p,\mu_{1,0})I_{K\times1})}{\partial \mu_{1,0}^{\mathrm{T}}} \; \vdots \; |\beta_2|^2 \cdot (A_2(p,\mu_{2,0}))^{\mathrm{H}} \cdot \frac{\partial (A_2(p,\mu_{2,0})I_{K\times1})}{\partial \mu_{2,0}^{\mathrm{T}}} \; \vdots \; \cdots \right.$$
$$\left. \cdots \; \vdots \; |\beta_N|^2 \cdot (A_N(p,\mu_{N,0}))^{\mathrm{H}} \cdot \frac{\partial (A_N(p,\mu_{N,0})I_{K\times1})}{\partial \mu_{N,0}^{\mathrm{T}}} \right]$$

$$(105)$$

$$\mathbf{T}^{\mathrm{H}}_{\boldsymbol{\mu}}\mathbf{T}_{\boldsymbol{\mu}} = \sum_{k=1}^{K} |\overline{s}_k|^2 \cdot \operatorname{blkdiag} \left[|\beta_1|^2 \left(\frac{\partial a'_{1,k}(p,\mu_{1,0})}{\partial \mu_{1,0}} \right)^{\mathrm{H}} \; \vdots \; |\beta_2|^2 \left(\frac{\partial a'_{2,k}(p,\mu_{2,0})}{\partial \mu_{2,0}} \right)^{\mathrm{H}} \; \vdots \; \cdots \; \vdots \; |\beta_N|^2 \left(\frac{\partial a'_{N,k}(p,\mu_{N,0})}{\partial \mu_{N,0}} \right)^{\mathrm{H}} \right.$$
$$\left. \times \frac{\partial a'_{1,k}(p,\mu_{1,0})}{\partial \mu_{1,0}^{\mathrm{T}}} \qquad \times \frac{\partial a'_{2,k}(p,\mu_{2,0})}{\partial \mu_{2,0}^{\mathrm{T}}} \qquad \times \frac{\partial a'_{N,k}(p,\mu_{N,0})}{\partial \mu_{N,0}^{\mathrm{T}}} \right]$$

$$(106)$$

14 Appendix 5

14.1 Detailed derivation of matrices in (80)

From (80) and the results in Appendix 4, in order to calculate matrices $\mathbf{Z}_1^{(\mathrm{b})}$, $\mathbf{Z}_2^{(\mathrm{b})}$, and $\mathbf{Z}_3^{(\mathrm{b})}$, we need to derive the closed-form expressions for matrices $\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{t_0}$, $\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{\mathbf{p}}$, $\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}}$, and $\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{\boldsymbol{\mu}}$.

Combining (70) and (78) and after some algebraic manipulation, we get

$$\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{t_0} = \sum_{n=1}^{N} \sum_{k=1}^{K} \omega_k^2 \cdot |\beta_n|^2 \cdot |\overline{s}_k|^2 \cdot \left\| \mathbf{a}'_{n,k}(\mathbf{p}, \boldsymbol{\mu}_{n,0}) \right\|_2^2 \tag{107}$$

$$\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{\mathbf{p}} = \sum_{n=1}^{N} \sum_{k=1}^{K} \mathrm{j}\omega_k \cdot |\beta_n|^2 \cdot |\overline{s}_k|^2 \cdot \left(\mathbf{a}'_{n,k}(\mathbf{p}, \boldsymbol{\mu}_{n,0}) \right)^{\mathrm{H}} \cdot \frac{\partial \mathbf{a}'_{n,k}(\mathbf{p}, \boldsymbol{\mu}_{n,0})}{\partial \mathbf{p}^{\mathrm{T}}} \tag{108}$$

$$\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{\mathrm{Re}\{\boldsymbol{\beta}\}} = \sum_{k=1}^{K} \mathrm{j}\omega_k \cdot |\overline{s}_k|^2 \cdot \{ \| a'_{1,k}(p,\mu_{1,0}) \|_2^2 \; \vdots \; \| a'_{2,k}(p,\mu_{2,0}) \|_2^2 \; \vdots \; \cdots \; \vdots \; \| a'_{N,k}(p,\mu_{N,0}) \|_2^2 \} \cdot \operatorname{diag}[\boldsymbol{\beta}^*]$$

$$(109)$$

$$\mathbf{T}^{\mathrm{H}}_{t_0}\mathbf{T}_{\boldsymbol{\mu}} = \sum_{k=1}^{K} \mathrm{j}\omega_k \cdot |\overline{s}_k|^2 \cdot \left[(a'_{1,k}(p,\mu_{1,0}))^{\mathrm{H}} \cdot \frac{\partial a'_{1,k}(p,\mu_{1,0})}{\partial \mu_{1,0}^{\mathrm{T}}} \; \vdots \; (a'_{2,k}(p,\mu_{2,0}))^{\mathrm{H}} \cdot \frac{\partial a'_{2,k}(p,\mu_{2,0})}{\partial \mu_{2,0}^{\mathrm{T}}} \; \vdots \; \cdots \; \vdots \; (a'_{N,k}(p,\mu_{N,0}))^{\mathrm{H}} \cdot \frac{\partial a'_{N,k}(p,\mu_{N,0})}{\partial \mu_{N,0}^{\mathrm{T}}} \right]$$
$$\times \operatorname{diag}[|\boldsymbol{\beta}|^2]$$

$$(110)$$

Abbreviations
3D: Three-dimensional; CRB: Cramér-Rao bound; DFT: Discrete Fourier transform; DOA: Direction-of-arrival; DPD: Direct position determination; EXIP: Extended invariance principle; FDOA: Frequency difference of arrival; FFT: Fast Fourier transform; FIM: Fisher information matrix; FOA: Frequency of arrival; GROA: Gain ratios of arrival; IID: Independent and identically distributed; MAP: Maximum a posteriori; ML: Maximum likelihood; MSE: Mean square error; RMSE: Root-mean-square-error; RSS: Received signal strength; SNR: Signal-to-

Authors' contributions
DW and JY wrote the manuscript and RL helped with writing and theoretical derivation. ZW and TT were in charge of the experiment and its results. All authors had significant contribution on the development of early ideas and design of the final algorithms. All authors read and approved the final manuscript.

Ethics approval and consent to participate
All data and procedures performed in paper were in accordance with the ethical standards of research community. This paper does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication
Informed consent was obtained from all authors included in the study.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. RO Schmidt, Multiple emitter location and signal parameter estimation[J]. IEEE Trans. Antennas Propag. **34**(3), 267–280 (1986).
2. P Stoica, A Nehorai, MUSIC, maximum likelihood, and Cramér-Rao bound[J]. IEEE Trans. Acoust. Speech Signal Process. **37**(5), 720–741 (1989).
3. M Viberg, B Ottersten, Sensor array processing based on subspace fitting[J]. IEEE Trans. Signal Process. **39**(5), 1110–1121 (1991).
4. SC Nardone, ML Graham, A closed-form solution to bearings-only target motion analysis[J]. IEEE J. Ocean. Eng. **22**(1), 168–178 (1997).
5. D Kutluyil, Bearings-only target localization using total least squares[J]. Signal Process. **85**(9), 1695–1710 (2005).
6. Z Lin, T Han, R Zheng, M Fu, Distributed localization for 2-D sensor networks with bearing-only measurements under switching topologies[J]. IEEE Trans. Signal Process. **64**(23), 6345–6359 (2016).
7. L Yang, KC Ho, An approximately efficient TDOA localization algorithm in closed-form for locating multiple disjoint sources with erroneous sensor positions[J]. IEEE Trans. Signal Process. **57**(12), 4598–4615 (2009).
8. W Jiang, C Xu, L Pei, W Yu, Multidimensional scaling-based TDOA localization scheme using an auxiliary line[J]. IEEE Signal Process Lett. **23**(4), 546–550 (2016).
9. KW Cheung, HC So, YT Chan, Least squares algorithms for time-of-arrival-based mobile location[J]. IEEE Trans. Signal Process. **52**(4), 1121–1128 (2004).
10. H Shen, Z Ding, S Dasgupta, C Zhao, Multiple source localization in wireless sensor networks based on time of arrival measurement[J]. IEEE Trans. Signal Process. **62**(8), 1938–1949 (2014).

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:38

Page 28 of 28

11. KC Ho, X Lu, L Kovavisaruch, Source localization using TDOA and FDOA measurements in the presence of receiver location errors: analysis and solution[J]. IEEE Trans. Signal Process. **55**(2), 684–696 (2007).

12. G Wang, Y Li, N Ansari, A semidefinite relaxation method for source localization using TDOA and FDOA measurements[J]. IEEE Trans. Veh. Technol. **62**(2), 853–862 (2013).

13. J Mason, Algebraic two-satellite TOA/FOA position solution on an ellipsoidal earth[J]. IEEE Trans. Aerosp. Electron. Syst. **40**(7), 1087–1092 (2004).

14. KC Ho, M Sun, An accurate algebraic closed-form solution for energy-based source localization[J]. IEEE Trans. Audio Speech Lang. Process. **15**(8), 2542–2550 (2007).

15. KC Ho, M Sun, Passive source localization using time differences of arrival and gain ratios of arrival[J]. IEEE Trans. Signal Process. **56**(2), 464–477 (2008).

16. M Wax, T Kailath, Decentralized processing in sensor arrays[J]. IEEE Trans. Signal Process. **33**(4), 1123–1129 (1985).

17. P Stoica, On reparametrization of loss functions used in estimation and the invariance principle[J]. Signal Process. **17**(4), 383–387 (1989).

18. A Amar, AJ Weiss, Localization of narrowband radio emitters based on Doppler frequency shifts[J]. IEEE Trans. Signal Process. **56**(11), 5500–5508 (2008).

19. D Wang, Y Wu, Statistical performance analysis of direct position determination method based on doppler shifts in presence of model errors[J]. Multidim. Syst. Sign. Process. **28**(1), 149–182 (2017).

20. E Tzoreff, AJ Weiss, Expectation-maximization algorithm for direct position determination. Signal Process. **97**(4), 32–39 (2017).

21. N Vankayalapati, S Kay, Q Ding, TDOA based direct positioning maximum likelihood estimator and the Cramer-Rao bound[J]. IEEE Trans. Aerosp. Electron. Syst. **50**(3), 1616–1634 (2014).

22. W Xia, W Liu, LF Zhu, Distributed adaptive direct position determination based on diffusion framework[J]. J. Syst. Eng. Electron. **27**(1), 28–38 (2016).

23. AJ Weiss, Direct geolocation of wideband emitters based on delay and Doppler[J]. IEEE Trans. Signal Process. **59**(6), 2513–2520 (2011).

24. M Pourhomayoun, ML Fowler, Distributed computation for direct position determination emitter location[J]. IEEE Trans. Aerosp. Electron. Syst. **50**(4), 2878–2889 (2014).

25. O Bar-Shalom, AJ Weiss, Emitter geolocation using single moving receiver[J]. Signal Process. **94**(12), 70–83 (2014).

26. JZ Li, L Yang, FC Guo, WL Jiang, Coherent summation of multiple short-time signals for direct positioning of a wideband source based on delay and Doppler[J]. Digital Signal Process. **48**(1), 58–70 (2016).

27. AJ Weiss, Direct position determination of narrowband radio frequency transmitters[J]. IEEE Signal Process Lett. **11**(5), 513–516 (2004).

28. AJ Weiss, A Amar, Direct position determination of multiple radio signals[J]. EURASIP J Appl. Signal Process. **2005**(1), 37–49 (2005).

29. A Amar, AJ Weiss, A decoupled algorithm for geolocation of multiple emitters[J]. Signal Process. **87**(10), 2348–2359 (2007).

30. T Tirer, AJ Weiss, High resolution direct position determination of radio frequency sources[J]. IEEE Signal Process Lett. **23**(2), 192–196 (2016).

31. L Tzafri, AJ Weiss, High-resolution direct position determination using MVDR[J]. IEEE Trans. Wireless Commun. **15**(9), 6449–6461 (2016).

32. T Tirer, AJ Weiss, Performance analysis of a high-resolution direct position determination method[J]. IEEE Trans. Signal Process. **65**(3), 544–554 (2017).

33. O Bar-Shalom, AJ Weiss, Efficient direct position determination of orthogonal frequency division multiplexing signals[J]. IET Radar Sonar Navig. **3**(2), 101–111 (2009).

34. AM Reuven, AJ Weiss, Direct position determination of cyclostationary signals[J]. Signal Process. **89**(12), 2448–2464 (2009).

35. JX Yin, W Ying, W Ding, Direct position determination of multiple noncircular sources with a moving array[J]. Circuits Syst. Signal Process. 2017;36(10):4050-76.

36. JX Yin, W Ding, W Ying, XW Yao, ML-based single-step estimation of the locations of strictly noncircular sources[J]. Digital Signal Process. **69**(10), 224–236 (2017).

37. M Oispuu, U Nickel, Direct detection and position determination of multiple sources with intermittent emission[J]. Signal Process. **90**(12), 3056–3064 (2010).

38. B Friedlander, A sensitivity analysis of the MUSIC algorithm[J]. IEEE Trans. Acoust. Speech Signal Process. **38**(11), 1740–1751 (1990).

39. A Swindlehurst, T Kailath, A performance analysis of subspace-based methods in the presence of model errors, part I: The MUSIC algorithm[J]. IEEE Trans. Signal Process. **40**(7), 1758–1773 (1992).

40. A Ferréol, P Larzabal, M Viberg, Statistical analysis of the MUSIC algorithm in the presence of modeling errors, taking into account the resolution probability[J]. IEEE Trans. Signal Process. **58**(8), 4156–4166 (2010).

41. M Khodja, A Belouchrani, K Abed-Meraim, Performance analysis for time-frequency MUSIC algorithm in presence of both additive noise and array calibration errors[J]. EURASIP J. Adv. Signal Process. **2012**(1), 94–104 (2012).

42. V Inghelbrecht, J Verhaevert, T van Hecke, H Rogier, The influence of random element displacement on DOA estimates obtained with (Khatri-Rao-) root-MUSIC[J]. Sensors **14**(11), 21258–21280 (2014).

43. A Amar, AJ Weiss, *Analysis of the direct position determination approach in the presence of model errors[A]. Proceedings of the IEEE Convention on Electrical and Electronics Engineers[C]* (IEEE Press, Tel Aviv, 2004), pp. 408–411.

44. A Amar, AJ Weiss, *Analysis of direct position determination approach in the presence of model errors[A]. Proceedings of the IEEE Workshop on Statistical Signal Processing[C]* (IEEE Press, Novosibirsk, 2005), pp. 521–524.

45. A Amar, AJ Weiss, Direct position determination in the presence of model errors—known waveforms[J]. Digital Signal Process. **16**(1), 52–83 (2006).

46. W Ding, HY Yu, ZD Wu, C Wang, Performance analysis of the direct position determination method in the presence of array model errors[J]. Sensors **17**(7), 1550–1590 (2017).

47. B Wahlberg, B Ottersten, M Viberg, *Robust signal parameter estimation in the presence of array perturbations[A]. Proceedings of the IEEE International conference on acoustic, speech and signal processing[C]*, vol 5 (IEEE Press, Toronto, 1991), pp. 3277–3280.

48. M Viberg, AL Swindlehurst, A Bayesian approach to auto-calibration for parametric array signal processing[J]. IEEE Trans. Signal Process. **42**(12), 3495–3507 (1994).

49. A Flieller, A Ferréol, P Larzabal, H Clergeot, *Robust bearing estimation in the presence of direction-dependent modeling errors: identifiability and treatment[A]. Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing[C]* (IEEE Press, Detroit, 1995), pp. 1884–1887.

50. M Jansson, AL Swindlehurst, B Ottersten, Weighted subspace fitting for general array error models[J]. IEEE Trans. Signal Process. **46**(9), 2484–2498 (1998).

51. EK Hung, Matrix-construction calibration method for antenna arrays[J]. IEEE Trans. Aerosp. Electron. Syst. **36**(3), 819–828 (2000).

52. QM Bao, C KO C, WJ Zhi, DOA estimation under unknown mutual coupling and multipath[J]. IEEE Trans. Aerosp. Electron. Syst. **41**(2), 565–573 (2005).

53. J Zheng, YC Wu, Joint time synchronization and localization of an unknown node in wireless sensor networks[J]. IEEE Trans. Signal Process. **58**(3), 1309–1320 (2010).

54. B Friedlander, AJ Weiss, Direction finding in the presence of mutual coupling[J]. IEEE Trans. Antennas Propag. **39**(3), 273–284 (1991).

55. J Li, RT Compton, Maximum likelihood angle estimation for signals with known waveforms[J]. IEEE Trans. Signal Process. **41**(9), 2850–2862 (1993).

56. J Li, B Halder, P Stoica, M Viberg, Computationally efficient angle estimation for signals with known waveforms[J]. IEEE Trans. Signal Process. **43**(9), 2154–2163 (1995).

57. Y Rockah, Schultheiss, Array shape calibration using sources in unknown locations—part I: far-field sources[J]. IEEE Trans. Acoust. Speech Signal Process. **35**(3), 286–299 (1987).

58. JX Zhu, H Wang, *Effects of sensor position and pattern perturbations on CRLB for direction finding of multiple narrowband sources[A]. Proceedings of the Fourth IEEE Workshop on Spectrum Estimation and Modeling[C]* (IEEE Press, Minneapolis, 1988), pp. 98–102.

59. XN Lu, KC Ho, *Taylor-series technique for source localization using AOAs in the presence of sensor location errors[A]. Proceedings of the Fourth IEEE Workshop on Sensor Array and Multichannel Processing[C]* (IEEE Press, Waltham, 2006), pp. 190–194.

60. M Kaveh, AJ Barabell, The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise[J]. IEEE Trans. Acoust. Speech Signal Process. **34**(2), 331–341 (1986).