Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Structural total least squares algorithm for locating multiple disjoint sources based on AOA/TOA/FOA in the presence of system error*

Xin CHEN[1,2], Ding WANG[†‡1,2], Rui-rui LIU[1,2], Jie-xin YIN[1,2], Ying WU[1,2]

[1]*National Digital Switching System Engineering and Technology Research Center, Zhengzhou 450002, China*

[2]*Zhengzhou Institute of Information Science and Technology, Zhengzhou 450002, China*

[†]E-mail: wang_ding814@aliyun.com

Received Nov. 8, 2017; Revision accepted Feb. 5, 2018; Crosschecked July 12, 2018

Abstract: Single-station passive localization technology avoids the complex time synchronization and information exchange between multiple observatories, and is increasingly important in electronic warfare. Based on a single moving station localization system, a new method with high localization precision and numerical stability is proposed when the measurements from multiple disjoint sources are subject to the same station position and velocity displacement. According to the available measurements including the angle-of-arrival (AOA), time-of-arrival (TOA), and frequency-of-arrival (FOA), the corresponding pseudo linear equations are deduced. Based on this, a structural total least squares (STLS) optimization model is developed and the inverse iteration algorithm is used to obtain the stationary target location. The localization performance of the STLS localization algorithm is derived, and it is strictly proved that the theoretical performance of the STLS method is consistent with that of the constrained total least squares method under first-order error analysis, both of which can achieve the Cramér-Rao lower bound accuracy. Simulation results show the validity of the theoretical derivation and superiority of the new algorithm.

Key words: Single-station; Structural total least squares; Inverse iteration; Angle-of-arrival (AOA); Time-of-arrival (TOA); Frequency-of-arrival (FOA); Disjoint sources

https://doi.org/10.1631/FITEE.1700735 **CLC number:** TN911.7

1 Introduction

Radiation source location technology can be divided into active location technology and passive location technology according to the nature of the radiation source. According to the number of observation stations, it can be divided into single-station location (Bar-Shalom and Weiss, 2014; Li et al., 2014) and multi-station location (Ho et al., 2007; Chalise et al., 2014). Compared with active location technology, passive location technology has the advantages of stronger concealment and longer detection distance, which can greatly improve the survivability and combat capability of the system in electronic warfare (Torrieri, 1984). Single-station passive location technology not only avoids time synchronization and data fusion between multiple observation stations, but also has the characteristics of flexibility, mobility, and simplicity. Thus, it has been of wide concern and has been studied intensely in recent decades. A maneuverable single station receives source emissions and extracts positioning parameters at different measurement points to determine source

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 61201381, 61401513, and 61772548), the China Postdoctoral Science Foundation (No. 2016M592989), the Self-Topic Foundation of Information Engineering University, China (No. 2016600701), and the Outstanding Youth Foundation of Information Engineering University, China (No. 2016603201)

ORCID: Xin CHEN, http://orcid.org/0000-0002-0562-0319

positions. Generally, the positioning parameters can be obtained, including the angle of arrival (AOA) (Knapp and Carter, 1976; Li and Lu, 2008; Yin et al., 2014; Giacometti et al., 2016), time of arrival (TOA) (Bamler, 1991; Belouchrani and Aouada, 2003; Antreich et al., 2011), and frequency of arrival (FOA) (Wang et al., 2001; Liu et al., 2017). TOA can be converted to the distance between the target and observation station. FOA is caused by the Doppler effect, and is equivalent to the range rate. Parameter estimation is an important research issue in the field of array signal processing. There are two main approaches for estimating these parameters. The first is to estimate individually each parameter of interest in the signal (Knapp and Carter, 1976; Bamler, 1991; Li and Lu, 2008; Yin et al., 2014), and then the results of parameter estimation are paired using a matching algorithm. This estimation method belongs to the traditional one-dimensional parameter estimation category, and does not make full use of the relationship between the parameters. There is a multi-parameter matching threshold problem. To obtain higher estimation accuracy, joint parameter estimation was proposed as an alternative (Antreich et al., 2011). This method can make full use of the relationship between the parameters to improve the estimation accuracy and achieve multi-parameter automatic matching. Because the joint estimation problem of three parameters is difficult to solve, multi-parameter joint estimation includes mainly the joint angle and delay estimation (Wang et al., 2001; Belouchrani and Aouada, 2003) and joint angle and Doppler frequency estimation (Lemma et al., 2003; Liu et al., 2010). Generally, the estimated parameters are algebraic equations with respect to target position, and the solution of the passive location is to determine the target position from these algebraic equations.

Clearly, existing measurements are non-linear functions about the target position, and we can transform them into pseudo linear equations according to the algebraic characteristics of the functions. During the past few decades, a number of least squares localization algorithms have been proposed, and most are for a single source location. The Taylor series method presented by Foy (1976) provides a least-sum-squared-error solution for a set of simultaneously linearized algebraic equations. Although the method can converge to the correct solution under certain conditions, it is very sensitive to the initial value. Thus, location robustness needs to be enhanced further. Doğançay (2005) and Sun et al. (2008) proposed a total least squares (TLS) positioning approach based on different measurements, which achieves zero bias and coincides approximately with the Cramér-Rao lower bound (CRLB) under small measurement noise. The TLS algorithm can achieve an asymptotic optimal performance when the noise component is independent and identically distributed in the observation matrix. Nevertheless, after pseudo linearization is used, noise components in the observation matrix are structured, which will lead to a sharp drop in the performance of the TLS algorithm. There is no doubt that the TLS method performs poorly with high noise levels, and cannot approach the CRLB. To overcome the drawback of the TLS algorithm, Abatzoglou et al. (1991), Yang et al. (2010), and Yu et al. (2012) proposed the constrained total least squares (CTLS) localization algorithm, which considers fully the algebraic structure of noise distribution in the observation matrix, and thus has high positioning accuracy. Wang et al. (2007) proposed CTLS methods based only on AOA measurements. Although the corresponding performance curve attains the CRLB and possesses a higher threshold compared with the closed-form method, the Newton iterative algorithm used in the CTLS method has high computational complexity. Wang et al. (2011) extended the CTLS method to localization scenarios based on wireless sensor networks. To suppress the influence of coefficient matrix disturbance in the linear equation, de Moor (1994) and Lemmerling et al. (1996) proposed another method to solve the parameters: the structural total least squares (STLS) algorithm. This method also needs to accurately model the perturbation term of the coefficient matrix, and a new optimization model and numerical iterative algorithm have been designed from the perspective of rank loss. However, in the literature, there are only few studies in which the STLS method was applied to the passive location, and only Wang et al. (2009) and Wu et al. (2015) proposed the STLS method based on angle information. However, the algorithm was deduced under a single measurement and did not consider the errors of the station position and velocity.

In this study, we consider the problem of localizing multiple disjoint sources (Carevic, 2007; Yang

and Ho, 2009; Sun and Ho, 2011; Hao et al., 2012; Shen et al., 2014; Griffin et al., 2015; Li et al., 2015; Jia et al., 2017) whose emissions do not interfere with each other, and the disjointedness can be achieved by limiting the source signal to different frequency bands and/or time intervals (Rockah and Schultheiss, 1987a; Raykar et al., 2005). In practice, multiple disjoint sources exist in some specific location scenes. In addition to practical consideration, the measurements from different sources contain the same number of displacements in the station position and velocity, which helps improve the positioning accuracy, and a lower CRLB can be obtained. This approach is also called self-calibration in direction of arrival (DOA) estimation (Rockah and Schultheiss, 1987b; Weiss and Friedlander, 1989; Raykar et al., 2005; Mir et al., 2007).

In this study, we develop an STLS localization algorithm for locating multiple disjoint sources when the station position and velocity are not known accurately. First, we deduce the pseudo linear equations according to the available spatial-, time-, and frequency-domain measurements. Then, we construct an STLS optimization model and apply the inverse iteration method to solve this problem and obtain the source position. Moreover, we develop a unified theoretical framework in which the STLS localization algorithm performance is derived. It is strictly proved that the theoretical performances of the STLS and CTLS methods are consistent under first-order error analysis and coincides with the CRLB. Finally, simulation results verify the validity of theoretical derivation, higher location accuracy than the general single-source location, and lower computational complexity compared with the CTLS and Taylor series method.

Notations used in this paper are listed in Table 1.

2 Measurement model and pseudo linearization process

Consider a single moving station localization system placed on an airborne or carrier platform in a three-dimensional (3D) scenario. Assume that M measurements are obtained during its movement. The true position and speed of the station at the i^{th} measurement are denoted by $\boldsymbol{s}_i^{\text{o}}=[x_{s_i}^{\text{o}},y_{s_i}^{\text{o}},z_{s_i}^{\text{o}}]^{\text{T}}$ and $\dot{\boldsymbol{s}}_i^{\text{o}}=[\dot{x}_{s_i}^{\text{o}},\dot{y}_{s_i}^{\text{o}},\dot{z}_{s_i}^{\text{o}}]^{\text{T}}$, i=1, 2, …, M, respectively, and are assumed to be known. Let the unknown fixed source position be $\boldsymbol{u}_j^{\text{o}}=[x_{\boldsymbol{u}_j}^{\text{o}},y_{\boldsymbol{u}_j}^{\text{o}},z_{\boldsymbol{u}_j}^{\text{o}}]^{\text{T}}$, j=1, 2, …, N.

Table 1 Mathematical notation

Notation	Description
$\boldsymbol{O}_{N\times n}$	$N\times n$ matrix with all-zero entries
$\boldsymbol{i}_M^{(l)}$	$M\times 1$ vector of zeros except for the l^{th} element, which has a value of 1
$\boldsymbol{I}_N$	N-dimensional identity matrix
$\otimes$	Kronecker product
diag( )	Diagonal matrix equipped with the elements of a vector
blkdiag( )	Block diagonal matrix consisting of the matrix or vector
$<>_n$	n^{th} element of a vector
$\|\|\ \|\|$	Euclidean norm of a vector

In the absence of the measurement noise and system error, the AOA measurement at the i^{th} measurement for the j^{th} source can be modeled as

$$\begin{cases}\theta_{i,j}^{\text{o}}=\arctan\dfrac{y_{\boldsymbol{u}_j}^{\text{o}}-y_{s_i}^{\text{o}}}{x_{\boldsymbol{u}_j}^{\text{o}}-x_{s_i}^{\text{o}}},\\ \beta_{i,j}^{\text{o}}=\arctan\dfrac{z_{\boldsymbol{u}_j}^{\text{o}}-z_{s_i}^{\text{o}}}{\sqrt{(x_{\boldsymbol{u}_j}^{\text{o}}-x_{s_i}^{\text{o}})^2+(y_{\boldsymbol{u}_j}^{\text{o}}-y_{s_i}^{\text{o}})^2}},\end{cases}\tag{1}$$

where θ_{ij}^{o} and β_{ij}^{o} represent the true azimuth and elevation for i=1, 2, …, M, j=1, 2, …, N, respectively. The range and range rate at the i^{th} measurement for the j^{th} source are

$$\begin{cases}r_{i,j}^{\text{o}}=\|\,u_j^{\text{o}}-s_i^{\text{o}}\,\|,\\ \dot{r}_{i,j}^{\text{o}}=\dfrac{-\dot{s}_i^{\text{oT}}(u_j^{\text{o}}-s_i^{\text{o}})}{\|\,u_j^{\text{o}}-s_i^{\text{o}}\,\|},\end{cases}\tag{2}$$

where i=1, 2, …, M and j=1, 2, …, N. The TOA and FOA measurements are then denoted by

$$\begin{cases}\tau_{i,j}^{\text{o}}=\dfrac{r_{i,j}^{\text{o}}}{c},\\ f_{i,j}^{\text{o}}=\dfrac{f_{\text{c}j}}{c}\dot{r}_{i,j}^{\text{o}},\end{cases}\tag{3}$$

where c is the signal propagation speed, and f_{cj} is the carrier frequency for the j^{th} source, i=1, 2, …, M, and j=1, 2, …, N. Note that f_{cj} is assumed to be known because it can be estimated or measured in practical applications (Luise and Reggiannini, 1995). For convenience, we replace the TOA and FOA measurements with the range r_{ij}^{o} and range rate $\dot{r}_{i,j}^{o}$, respectively, in the following localization problem.

Collect the measurements from the same source together given by $\boldsymbol{\theta}_j^{o}=[\theta_{1,j}^{o},\theta_{2,j}^{o},\cdots,\theta_{M,j}^{o}]^{T}$, $\boldsymbol{\beta}_j^{o}=[\beta_{1,j}^{o},\beta_{2,j}^{o},\cdots,\beta_{M,j}^{o}]^{T}$, $\boldsymbol{r}_j^{o}=[r_{1,j}^{o},r_{2,j}^{o},\cdots,r_{M,j}^{o}]^{T}$, and $\dot{\boldsymbol{r}}_j^{o}=[\dot{r}_{1,j}^{o},\dot{r}_{2,j}^{o},\cdots,\dot{r}_{M,j}^{o}]^{T}$. Then all measurements are collected in the $4M\times1$ vectors $\boldsymbol{\xi}_j^{o}=[\boldsymbol{\theta}_j^{oT},\boldsymbol{\beta}_j^{oT},\boldsymbol{r}_j^{oT},\dot{\boldsymbol{r}}_j^{oT}]^{T}$.

To make full use of the information with respect to the target location to improve the positioning accuracy, we construct a joint positioning mode based on the AOA/TOA/FOA. Note that some measurement equations can be transformed into pseudo linear equations (So and Lin, 2011) directly, and other measurement equations can be transformed into pseudo linear equations by combining other measurement equations.

AOA measurement equations can be formatted into pseudo linear equations directly:

$$\begin{cases}(x_{\boldsymbol{u}_j}^{o}-x_{\boldsymbol{s}_i}^{o})\sin\theta_{\boldsymbol{s}_i}^{o}=(y_{\boldsymbol{u}_j}^{o}-y_{\boldsymbol{s}_i}^{o})\cos\theta_{i,j}^{o},\\(z_{\boldsymbol{u}_j}^{o}-z_{\boldsymbol{s}_i}^{o})\cos\beta_{\boldsymbol{s}_i}^{o}=[(x_{\boldsymbol{u}_j}^{o}-x_{\boldsymbol{s}_i}^{o})\cos\theta_{i,j}^{o}\\\qquad\qquad+(y_{\boldsymbol{u}_j}^{o}-y_{\boldsymbol{s}_i}^{o})\sin\theta_{i,j}^{o}]\sin\beta_{i,j}^{o}.\end{cases}\tag{4}$$

Rewrite Eq. (4) into vector format:

$$\begin{cases}\boldsymbol{\omega}_1^{T}(\theta_{i,j}^{o})\boldsymbol{u}_j^{o}=\boldsymbol{\omega}_1^{T}(\theta_{i,j}^{o})\boldsymbol{s}_i^{o},\\\boldsymbol{\omega}_2^{T}(\theta_{i,j}^{o},\beta_{i,j}^{o})\boldsymbol{u}_j^{o}=\boldsymbol{\omega}_2^{T}(\theta_{i,j}^{o},\beta_{i,j}^{o})\boldsymbol{s}_i^{o},\end{cases}\tag{5}$$

where

$$\begin{cases}\boldsymbol{\omega}_1(\theta_{i,j}^{o})=[\sin\theta_{i,j}^{o},-\cos\theta_{i,j}^{o},0]^{T}\\\boldsymbol{\omega}_2(\theta_{i,j}^{o},\beta_{i,j}^{o})=[\cos\theta_{i,j}^{o}\sin\beta_{i,j}^{o},\\\qquad\qquad\sin\theta_{i,j}^{o}\sin\beta_{i,j}^{o},-\cos\beta_{i,j}^{o}]^{T}.\end{cases}\tag{6}$$

Different from AOA equations, TOA equations cannot be transformed into pseudo linear equations directly. However, this problem can be solved effectively by combining angle measurements. Resort to the geometry relationship of targets and observation stations as follows:

$$\|\boldsymbol{u}_j^{o}-\boldsymbol{s}_i^{o}\|=\boldsymbol{\omega}_3^{T}(\theta_{i,j}^{o},\beta_{i,j}^{o})(\boldsymbol{u}_j^{o}-\boldsymbol{s}_i^{o}),\tag{7}$$

where

$$\begin{aligned}\boldsymbol{\omega}_3(\theta_{i,j}^{o},\beta_{i,j}^{o})=[&\cos\theta_{i,j}^{o}\cos\beta_{i,j}^{o},\\&\sin\theta_{i,j}^{o}\cos\beta_{i,j}^{o},\sin\beta_{i,j}^{o}]^{T}.\end{aligned}\tag{8}$$

Substitute Eq. (7) into TOA measurement equations and obtain the TOA equations:

$$r_{i,j}^{o}+\boldsymbol{\omega}_3^{T}(\theta_{i,j}^{o},\beta_{i,j}^{o})\boldsymbol{s}_i^{o}=\boldsymbol{\omega}_3^{T}(\theta_{i,j}^{o},\beta_{i,j}^{o})\boldsymbol{u}_j^{o}.\tag{9}$$

FOA measurement equations can also be transformed into pseudo linear equations using Eq. (8):

$$(\dot{r}_{i,j}^{o}\boldsymbol{\omega}_3^{T}(\theta_{i,j}^{o},\beta_{i,j}^{o})+\dot{\boldsymbol{s}}_i^{oT})\boldsymbol{u}_j^{o}=\dot{r}_{i,j}^{o}\boldsymbol{\omega}_3^{T}(\theta_{i,j}^{o},\beta_{i,j}^{o})\boldsymbol{s}_i^{o}+\dot{\boldsymbol{s}}_i^{oT}\boldsymbol{s}_i^{o}.\tag{10}$$

We define the measurement vector $\boldsymbol{\xi}^{o}=[\xi_1^{o},\xi_2^{o},\cdots,\xi_N^{o}]^{T}$ (let its dimension be p=4MN), observation station position and velocity vector $\boldsymbol{s}^{o}=[\boldsymbol{s}_1^{oT},\dot{\boldsymbol{s}}_1^{oT},\boldsymbol{s}_2^{oT},\dot{\boldsymbol{s}}_2^{oT},\ldots,\boldsymbol{s}_M^{oT},\dot{\boldsymbol{s}}_M^{oT}]^{T}$ (let its dimension be r=6M), and unknown source position vector $\boldsymbol{u}^{o}=[\boldsymbol{u}_1^{oT},\boldsymbol{u}_2^{oT},\ldots,\boldsymbol{u}_N^{oT}]^{T}$ (let its dimension be q=3N). Then the simultaneous equations of Eqs. (5), (9), and (10) give the joint localization equation set based on the AOA, TOA, and FOA measurements:

$$\boldsymbol{A}(\boldsymbol{\xi}^{o},\boldsymbol{s}^{o})\boldsymbol{u}^{o}=\boldsymbol{b}(\boldsymbol{\xi}^{o},\boldsymbol{s}^{o}),\tag{11}$$

where $\boldsymbol{A}(\boldsymbol{\xi}^{o},\boldsymbol{s}^{o})$ represents the coefficient matrix and $\boldsymbol{b}(\boldsymbol{\xi}^{o},\boldsymbol{s}^{o})$ the pseudo linear measurement vector. Their composition is described in Appendix A.

In practice, the AOA, TOA, and FOA measurements available for processing, denoted by $\boldsymbol{\xi}$, have random noise. The measurement noise in $\boldsymbol{\xi}$ is

$$\boldsymbol{e}_1=\boldsymbol{\xi}-\boldsymbol{\xi}^{o},\tag{12}$$

where $\boldsymbol{e}_1=[\boldsymbol{e}_{1,1}^{T},\boldsymbol{e}_{1,2}^{T},\cdots,\boldsymbol{e}_{1,N}^{T}]^{T}=[e_{1,1},\ e_{1,2},\ \ldots,\ e_{1,4MN}]^{T}$

follows a zero mean Gaussian distribution with a covariance matrix $Q_1=E[e_1e_1^{\mathrm{T}}]$. Specifically, $\boldsymbol{e}_{1,j}=[e_{1,4M(j-1)+1},\ e_{1,4M(j-2)+1},\ \ldots,\ e_{1,4Mj}]^{\mathrm{T}}$ represents the measurement noise for the j^{th} source, where the first $2M$ elements, $2M+1$ to $3M$ elements, and the latter M elements represent the measurement noise of the AOA, TOA, and FOA, respectively. We define the covariance matrices of these measurement noises:

$$\begin{cases}\boldsymbol{Q}_{1,j}^{\mathrm{AOA}}=\sigma_{\mathrm{AOA}}^2\boldsymbol{I}_{2M},\\ \boldsymbol{Q}_{1,j}^{\mathrm{TOA}}=\sigma_{\mathrm{TOA}}^2\boldsymbol{I}_{M},\\ \boldsymbol{Q}_{1,j}^{\mathrm{FOA}}=\sigma_{\mathrm{FOA}}^2\boldsymbol{I}_{M},\end{cases}\tag{13}$$

where σ_{AOA}, σ_{TOA}, and σ_{FOA} represent the noise power of the AOA, TOA, and FOA measurements for $j=1, 2, \ldots, N$, respectively. Assume that different measurements are independent of each other. Therefore, we have $\boldsymbol{Q}_{1,j}=\mathrm{blkdiag}\,(\boldsymbol{Q}_{1,j}^{\mathrm{AOA}},\boldsymbol{Q}_{1,j}^{\mathrm{TOA}},\boldsymbol{Q}_{1,j}^{\mathrm{FOA}})$ and $\boldsymbol{Q}_1=\mathrm{blkdiag}(\boldsymbol{Q}_{1,1}, \boldsymbol{Q}_{1,2}, \ldots, \boldsymbol{Q}_{1,N})$. Let the AOA measurement for the j^{th} source be $\boldsymbol{\xi}_{j,\mathrm{AOA}}=[\boldsymbol{\theta}_j^{\mathrm{T}},\boldsymbol{\beta}_j^{\mathrm{T}}]^{\mathrm{T}}$, and probability density functions (PDF) of the AOA, TOA, and FOA are given by

$$\begin{cases}f(\boldsymbol{\xi}_{j,\mathrm{AOA}})=\dfrac{1}{(2\pi)^{M/2}\,|\boldsymbol{Q}_{1,j}^{\mathrm{AOA}}|^{1/2}}\\ \cdot\exp[-(\boldsymbol{\xi}_{j,\mathrm{AOA}}-\boldsymbol{\xi}_{j,\mathrm{AOA}}^{\mathrm{o}})^{\mathrm{T}}(\boldsymbol{Q}_{1,j}^{\mathrm{AOA}})^{-1}(\boldsymbol{\xi}_{j,\mathrm{AOA}}-\boldsymbol{\xi}_{j,\mathrm{AOA}}^{\mathrm{o}})/2],\\ f(\boldsymbol{r}_j)=\dfrac{1}{(2\pi)^{M/2}\,|\boldsymbol{Q}_{1,j}^{\mathrm{TOA}}|^{1/2}}\\ \cdot\exp[-(\boldsymbol{r}_j-\boldsymbol{r}_j^{\mathrm{o}})^{\mathrm{T}}(\boldsymbol{Q}_{1,j}^{\mathrm{TOA}})^{-1}(\boldsymbol{r}_j-\boldsymbol{r}_j^{\mathrm{o}})/2],\\ f(\dot{\boldsymbol{r}}_j)=\dfrac{1}{(2\pi)^{M/2}\,|\boldsymbol{Q}_{1,j}^{\mathrm{FOA}}|^{1/2}}\\ \cdot\exp[-(\dot{\boldsymbol{r}}_j-\dot{\boldsymbol{r}}_j^{\mathrm{o}})^{\mathrm{T}}(\boldsymbol{Q}_{1,j}^{\mathrm{FOA}})^{-1}(\dot{\boldsymbol{r}}_j-\dot{\boldsymbol{r}}_j^{\mathrm{o}})/2].\end{cases}\tag{14}$$

Further, the AOA, TOA, and FOA measurements are located in the following intervals with a probability of 99.7%:

$$\begin{cases}\langle\boldsymbol{\xi}_{j,\mathrm{AOA}}^{\mathrm{o}}\rangle_i-2.58\sigma_{\mathrm{AOA}}\le\langle\boldsymbol{\xi}_{j,\mathrm{AOA}}\rangle_i\le\langle\boldsymbol{\xi}_{j,\mathrm{AOA}}^{\mathrm{o}}\rangle_i+2.58\sigma_{\mathrm{AOA}},\\ \langle\boldsymbol{r}_j^{\mathrm{o}}\rangle_i-2.58\sigma_{\mathrm{AOA}}\le\langle\boldsymbol{r}_j\rangle_i\le\langle\boldsymbol{r}_j^{\mathrm{o}}\rangle_i+2.58\sigma_{\mathrm{AOA}},\\ \langle\dot{\boldsymbol{r}}_j^{\mathrm{o}}\rangle_i-2.58\sigma_{\mathrm{AOA}}\le\langle\dot{\boldsymbol{r}}_j\rangle_i\le\langle\dot{\boldsymbol{r}}_j^{\mathrm{o}}\rangle_i+2.58\sigma_{\mathrm{AOA}}.\end{cases}\tag{15}$$

By contrast, although $\boldsymbol{s}^{\mathrm{o}}$ is easy to access through the actual measurement, it may inevitably be disturbed by noise, such as sensors fixed on the moving airborne and marine platforms, or random positions. Then, we assume that the system error is

$$\boldsymbol{e}_2=\boldsymbol{s}-\boldsymbol{s}^{\mathrm{o}},\tag{16}$$

where $\boldsymbol{s}$ is the station position and velocity vector available and $\boldsymbol{e}_2=[e_{2,1}, e_{2,2}, \ldots, e_{2,6M}]^{\mathrm{T}}$ is the zero mean distributed with a covariance matrix $\boldsymbol{Q}_2=E[\boldsymbol{e}_2\boldsymbol{e}_2^{\mathrm{T}}]$.

Because the measurement noise $\boldsymbol{e}_1$ and system error $\boldsymbol{e}_2$ are independent of each other for different originations, we can gather the two types of noise as vector $\boldsymbol{e}=[\boldsymbol{e}_1^{\mathrm{T}},\boldsymbol{e}_2^{\mathrm{T}}]^{\mathrm{T}}$, and the covariance matrix can be expressed as $\boldsymbol{Q}=\mathbf{cov}(\boldsymbol{e})=E[\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}]=\mathrm{blkdiag}(\boldsymbol{Q}_1, \boldsymbol{Q}_2)$.

3 STLS localization algorithm

3.1 Constructing the STLS localization model

According to Doğançay (2005), Markovsky and van Huffel (2007), and Sun et al. (2008), the measurement noise can be effectively suppressed using the TLS positioning algorithm when the coefficient matrix is subject to noise. However, the TLS method can achieve optimal statistical performance when all disturbance components are independent and identically distributed in the coefficient matrix and observation vector. Considering that the disturbance in $\boldsymbol{A}(\boldsymbol{\xi}, \boldsymbol{s})$ and $\boldsymbol{b}(\boldsymbol{\xi}, \boldsymbol{s})$ must be structuralized for $\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})$ and $\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})$, which are the function matrices with respect to $\boldsymbol{\xi}^{\mathrm{o}}$ and $\boldsymbol{s}^{\mathrm{o}}$, collectively, the TLS localization algorithm cannot achieve the optimal statistical performance. Through more detailed modeling analysis, the STLS method can overcome the shortcomings of the TLS method, so that a more accurate positioning solution can be obtained.

To establish the STLS positioning optimization model, we first expand $\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})$ and $\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})$ with respect to $\boldsymbol{\xi}$ and $\boldsymbol{s}$ using the Taylor series, and ignore the second- and higher-order terms:

$$\begin{cases} \boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})=\boldsymbol{A}(\boldsymbol{\xi},\boldsymbol{s})+\sum\limits_{k=1}^{p}\dot{\boldsymbol{A}}_{1k}(\boldsymbol{\xi},\boldsymbol{s})\langle\boldsymbol{\xi}^{\mathrm{o}}-\boldsymbol{\xi}\rangle_k \\ \qquad\qquad +\sum\limits_{k=1}^{r}\dot{\boldsymbol{A}}_{2k}(\boldsymbol{\xi},\boldsymbol{s})\langle\boldsymbol{s}^{\mathrm{o}}-\boldsymbol{s}\rangle_k \\ \qquad\qquad =\boldsymbol{A}(\boldsymbol{\xi},\boldsymbol{s})-\sum\limits_{k=1}^{p}\langle\boldsymbol{e}_1\rangle_k\dot{\boldsymbol{A}}_{1k}(\boldsymbol{\xi},\boldsymbol{s}) \\ \qquad\qquad -\sum\limits_{k=1}^{r}\langle\boldsymbol{e}_2\rangle_k\dot{\boldsymbol{A}}_{2k}(\boldsymbol{\xi},\boldsymbol{s}), \\ \boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})=\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s})+\boldsymbol{B}_1(\boldsymbol{\xi},\boldsymbol{s})(\boldsymbol{\xi}^{\mathrm{o}}-\boldsymbol{\xi}) \\ \qquad\qquad +\boldsymbol{B}_2(\boldsymbol{\xi},\boldsymbol{s})(\boldsymbol{s}^{\mathrm{o}}-\boldsymbol{s}) \\ \qquad\qquad =\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s})-\boldsymbol{B}_1(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{e}_1-\boldsymbol{B}_2(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{e}_2 \\ \qquad\qquad =\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s})-\sum\limits_{k=1}^{p}\langle\boldsymbol{e}_1\rangle_k\boldsymbol{B}_1(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_p^{(k)} \\ \qquad\qquad -\sum\limits_{k=1}^{r}\langle\boldsymbol{e}_2\rangle_k\boldsymbol{B}_2(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_r^{(k)}, \end{cases} \tag{17}$$

where

$$\begin{cases} \dot{\boldsymbol{A}}_{1k}(\boldsymbol{\xi},\boldsymbol{s})=\left.\dfrac{\partial\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})}{\partial\langle\boldsymbol{\xi}^{\mathrm{o}}\rangle_k}\right|_{\boldsymbol{\xi}^{\mathrm{o}}=\boldsymbol{\xi},\boldsymbol{s}^{\mathrm{o}}=\boldsymbol{s}}, \\ \dot{\boldsymbol{A}}_{2k}(\boldsymbol{\xi},\boldsymbol{s})=\left.\dfrac{\partial\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})}{\partial\langle\boldsymbol{s}^{\mathrm{o}}\rangle_k}\right|_{\boldsymbol{\xi}^{\mathrm{o}}=\boldsymbol{\xi},\boldsymbol{s}^{\mathrm{o}}=\boldsymbol{s}}, \end{cases} \tag{18}$$

and

$$\begin{cases} \boldsymbol{B}_1(\boldsymbol{\xi},\boldsymbol{s})=\left.\dfrac{\partial\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})}{\partial\boldsymbol{\xi}^{\mathrm{o}\mathrm{T}}}\right|_{\boldsymbol{\xi}^{\mathrm{o}}=\boldsymbol{\xi},\boldsymbol{s}^{\mathrm{o}}=\boldsymbol{s}}, \\ \boldsymbol{B}_2(\boldsymbol{\xi},\boldsymbol{s})=\left.\dfrac{\partial\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})}{\partial\boldsymbol{s}^{\mathrm{o}\mathrm{T}}}\right|_{\boldsymbol{\xi}^{\mathrm{o}}=\boldsymbol{\xi},\boldsymbol{s}^{\mathrm{o}}=\boldsymbol{s}}. \end{cases} \tag{19}$$

Substituting Eq. (17) into Eq. (11) yields the following approximate equation constraint:

$$\begin{aligned} &[-\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s})\big|\boldsymbol{A}(\boldsymbol{\xi},\boldsymbol{s})]\cdot\begin{bmatrix}1\\ \boldsymbol{u}\end{bmatrix} \\ &+\begin{pmatrix}\sum\limits_{k=1}^{p}\langle\boldsymbol{e}_1\rangle_k[\boldsymbol{B}_1(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_p^{(k)}\,|-\dot{\boldsymbol{A}}_{1k}(\boldsymbol{\xi},\boldsymbol{s})] \\ +\sum\limits_{k=1}^{r}\langle\boldsymbol{e}_2\rangle_k[\boldsymbol{B}_2(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_r^{(k)}\,|-\dot{\boldsymbol{A}}_{2k}(\boldsymbol{\xi},\boldsymbol{s})]\end{pmatrix}\cdot\begin{bmatrix}1\\ \boldsymbol{u}\end{bmatrix}=\boldsymbol{0}_{p\times1}. \end{aligned} \tag{20}$$

According to Eq. (20) and the second-order statistical properties of measurement noise $\boldsymbol{e}_1$ and system error $\boldsymbol{e}_2$, we can establish the following STLS optimization model:

$$\begin{aligned} &\min_{\substack{\boldsymbol{r}\in\mathrm{R}^{(q+1)\times1}\\ \boldsymbol{e}\in\mathrm{R}^{(p+r)\times1}}}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{e}, \\ &\text{s.t.}\begin{pmatrix}[-\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s})\big|\boldsymbol{A}(\boldsymbol{\xi},\boldsymbol{s})] \\ +\sum\limits_{k=1}^{p}\langle\boldsymbol{e}_1\rangle_k[\boldsymbol{B}_1(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_p^{(k)}\,|-\dot{\boldsymbol{A}}_{1k}(\boldsymbol{\xi},\boldsymbol{s})] \\ +\sum\limits_{k=1}^{r}\langle\boldsymbol{e}_2\rangle_k[\boldsymbol{B}_2(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_r^{(k)}\,|-\dot{\boldsymbol{A}}_{2k}(\boldsymbol{\xi},\boldsymbol{s})]\end{pmatrix}\cdot\boldsymbol{r}=\boldsymbol{0}_{p\times1}, \\ &\boldsymbol{i}_{q+1}^{(1)\mathrm{T}}\boldsymbol{r}=1. \end{aligned} \tag{21}$$

Define the vector

$$\begin{cases} \boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})=[-\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s})\big|\boldsymbol{A}(\boldsymbol{\xi},\boldsymbol{s})], \\ \boldsymbol{R}_{1k}(\boldsymbol{\xi},\boldsymbol{s})=[\boldsymbol{B}_1(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_p^{(k)}\,|-\dot{\boldsymbol{A}}_{1k}(\boldsymbol{\xi},\boldsymbol{s})], \\ \boldsymbol{R}_{2k}(\boldsymbol{\xi},\boldsymbol{s})=[\boldsymbol{B}_2(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{i}_r^{(k)}\,|-\dot{\boldsymbol{A}}_{2k}(\boldsymbol{\xi},\boldsymbol{s})], \\ \boldsymbol{R}_k(\boldsymbol{\xi},\boldsymbol{s})=\boldsymbol{R}_{1k}(\boldsymbol{\xi},\boldsymbol{s}),\ 1\le k\le p, \\ \boldsymbol{R}_{p+k}(\boldsymbol{\xi},\boldsymbol{s})=\boldsymbol{R}_{2k}(\boldsymbol{\xi},\boldsymbol{s}),\ 1\le k\le r. \end{cases} \tag{22}$$

Furthermore, Eq. (21) can be simplified as

$$\begin{aligned} &\min_{\substack{\boldsymbol{r}\in\mathrm{R}^{(q+1)\times1}\\ \boldsymbol{e}\in\mathrm{R}^{(p+r)\times1}}}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{e} \\ &\text{s.t.}\left(\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})+\sum_{k=1}^{p+r}\langle\boldsymbol{e}\rangle_k\boldsymbol{R}_k(\boldsymbol{\xi},\boldsymbol{s})\right)\cdot\boldsymbol{r}=\boldsymbol{0}_{p\times1},\ \boldsymbol{i}_{q+1}^{(1)\mathrm{T}}\boldsymbol{r}=1. \end{aligned} \tag{23}$$

To obtain the standard form of the STLS optimization model, define $\overline{\boldsymbol{e}}=\boldsymbol{Q}^{-1/2}\boldsymbol{e}$. Then we have

$$\langle\boldsymbol{e}\rangle_k=\langle\boldsymbol{Q}^{1/2}\overline{\boldsymbol{e}}\rangle_k=\sum_{m=1}^{p+r}\langle\boldsymbol{Q}^{1/2}\rangle_{km}\langle\overline{\boldsymbol{e}}\rangle_m,\ 1\le k\le p+r. \tag{24}$$

Therefore, Eq. (23) can be reformulated as

$$\begin{aligned} &\min_{\substack{\boldsymbol{r}\in\mathrm{R}^{(q+1)\times1}\\ \overline{\boldsymbol{e}}\in\mathrm{R}^{(p+r)\times1}}}\|\overline{\boldsymbol{e}}\|_2^2 \\ &\text{s.t.}\begin{pmatrix}\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s}) \\ +\sum\limits_{k=1}^{p+r}\langle\overline{\boldsymbol{e}}\rangle_k\left(\sum\limits_{m=1}^{p+r}\langle\boldsymbol{Q}^{1/2}\rangle_{km}\boldsymbol{R}_m(\boldsymbol{\xi},\boldsymbol{s})\right)\end{pmatrix}\cdot\boldsymbol{r}=\boldsymbol{0}_{p\times1}, \\ &\boldsymbol{i}_{q+1}^{(1)\mathrm{T}}\boldsymbol{r}=1. \end{aligned} \tag{25}$$

To proceed, we define vector

$$\bar{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})=\sum_{m=1}^{p+r}\langle \boldsymbol{Q}^{1/2}\rangle_{km}\boldsymbol{R}_m(\boldsymbol{\xi},\boldsymbol{s}). \tag{26}$$

As a result, Eq. (25) can be further simplified as

$$\begin{aligned}&\min_{\substack{\boldsymbol{r}\in\mathrm{R}^{(q+1)\times 1}\\ \bar{\boldsymbol{e}}\in\mathrm{R}^{(p+r)\times 1}}} \|\bar{\boldsymbol{e}}\|_2^2\\ &\text{s.t.}\ \left(\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})+\sum_{k=1}^{p+r}\langle\bar{\boldsymbol{e}}\rangle_k\bar{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\right)\cdot\boldsymbol{r}=\boldsymbol{0}_{p\times 1},\ \boldsymbol{i}_{q+1}^{(1)\mathrm{T}}\boldsymbol{r}=1.\end{aligned} \tag{27}$$

Because vector $\boldsymbol{r}$ is in the null space of $\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})+\sum_{k=1}^{p+r}\langle\bar{\boldsymbol{e}}\rangle_k\bar{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})$, where the first element is one, the change of the constraint on $\boldsymbol{r}$ has no direct influence on the solution to the location problem. Then, we need only to let $\hat{\boldsymbol{u}}=\hat{\boldsymbol{r}}(2:q+1)/\hat{\boldsymbol{r}}(1)$ be the solution of the location problem. Thus, we transform the linear constraint in Eq. (27) into the secondary constraint $\|\boldsymbol{r}\|_2=1$. The standard form of the STLS optimization model is given by

$$\begin{aligned}&\min_{\substack{\boldsymbol{r}\in\mathrm{R}^{(q+1)\times 1}\\ \bar{\boldsymbol{e}}\in\mathrm{R}^{(p+r)\times 1}}} \|\bar{\boldsymbol{e}}\|_2^2\\ &\text{s.t.}\ \left(\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})+\sum_{k=1}^{p+r}\langle\bar{\boldsymbol{e}}\rangle_k\bar{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\right)\cdot\boldsymbol{r}=\boldsymbol{0}_{p\times 1},\ \|\boldsymbol{r}\|_2=1.\end{aligned} \tag{28}$$

3.2 Inverse iteration algorithm to solve the STLS localization model

Eq. (28) is a standard form of the STLS optimization model; thus, there is a unified numerical algorithm. de Moor (1994) provided the basic concept of solving the STLS problem. However, it is not an executable numerical algorithm. In this subsection, we propose a specific algorithm and steps. The essence of the solution is to determine the triad ($\boldsymbol{\alpha}$, τ, $\boldsymbol{\beta}$) that minimizes $|\tau|$, in which $\boldsymbol{\alpha}\in\mathbb{R}^{p\times 1}$, $\boldsymbol{\beta}\in\mathbb{R}^{(q+1)\times 1}$ and $\tau\in\mathbb{R}$, and they meet the following constraints:

$$\begin{cases}\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{\beta}=\tau\boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{\alpha},\ \ \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{\alpha}=1,\\ \boldsymbol{R}_{\mathrm{o}}^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{\alpha}=\tau\boldsymbol{S}_2(\boldsymbol{\alpha})\boldsymbol{\beta},\ \ \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{S}_2(\boldsymbol{\alpha})\boldsymbol{\beta}=1,\end{cases} \tag{29}$$

where

$$\begin{cases}\boldsymbol{S}_1(\boldsymbol{\beta})=\sum_{k=1}^{p+r}\bar{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}\bar{\boldsymbol{R}}_k^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s}),\\ \boldsymbol{S}_2(\boldsymbol{\alpha})=\sum_{k=1}^{p+r}\bar{\boldsymbol{R}}_k^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\bar{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s}).\end{cases} \tag{30}$$

Eq. (30) is also called 'Riemannian singular value decomposition', which can be solved using the inverse iteration algorithm. The main idea is to use the current latest iteration values $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to calculate matrices $\boldsymbol{S}_2(\boldsymbol{\alpha})$ and $\boldsymbol{S}_1(\boldsymbol{\beta})$, respectively, in each iteration. Then, they are substituted into Eq. (29) as the constant matrix to obtain the iteration update values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. This process is repeated until the iteration converges.

The iteration update values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be obtained using QR decomposition of $\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})$. The QR decomposition of $\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})$ is expressed as

$$\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})=[\boldsymbol{P}_1\quad\boldsymbol{P}_2]\cdot\begin{bmatrix}\boldsymbol{R}\\ \boldsymbol{Q}\end{bmatrix}, \tag{31}$$

where $\boldsymbol{P}_1\in\mathbb{R}^{p\times(q+1)}$, $\boldsymbol{P}_2\in\mathbb{R}^{p\times(p-q-1)}$, $\boldsymbol{R}\in\mathbb{R}^{(q+1)\times(q+1)}$, and $\boldsymbol{Q}\in\mathbb{R}^{(p-q-1)\times(q+1)}$. $\boldsymbol{P}=[\boldsymbol{P}_1,\ \boldsymbol{P}_2]$ is an orthogonal matrix; that is, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{I}_p$. $\boldsymbol{R}$ is the upper triangular matrix, and $\boldsymbol{\alpha}$ can be decomposed as

$$\boldsymbol{\alpha}=\boldsymbol{P}_1\boldsymbol{\alpha}_1+\boldsymbol{P}_2\boldsymbol{\alpha}_2, \tag{32}$$

where $\boldsymbol{\alpha}_1\in\mathbb{R}^{(q+1)\times 1}$ and $\boldsymbol{\alpha}_2\in\mathbb{R}^{(p-q-1)\times 1}$. From Eq. (29), the following linear equations can be deduced:

$$\begin{bmatrix}\boldsymbol{R}^{\mathrm{T}} & \boldsymbol{0}_{(q+1)\times(p-q-1)} & \boldsymbol{0}_{(q+1)\times(q+1)}\\ \boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{P}_1 & \boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{P}_2 & \boldsymbol{0}_{(p-q-1)\times(q+1)}\\ \tau\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{P}_1 & \tau\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{P}_2 & -\boldsymbol{R}\end{bmatrix}\cdot\begin{bmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \boldsymbol{\beta}\end{bmatrix}=\begin{bmatrix}\tau\boldsymbol{S}_2(\boldsymbol{\alpha})\boldsymbol{\beta}\\ \boldsymbol{0}_{(p-q-1)\times 1}\\ \boldsymbol{0}_{(q+1)\times 1}\end{bmatrix}. \tag{33}$$

It is clear that the number of unknowns in Eq. (33) is $p+q+1$ and the number of equations is $p+q+1$. Hence, a unique solution can be determined. Because the coefficient matrix on the left-hand side of Eq. (33) is a block of the lower triangular matrix, the solution can be determined using the form of the

recurrence formulas as follows:

$$\begin{cases} \boldsymbol{\alpha}_1 = \tau \boldsymbol{R}^{-\mathrm{T}} \boldsymbol{S}_2(\boldsymbol{\alpha})\boldsymbol{\beta}, \\ \boldsymbol{\alpha}_2 = -(\boldsymbol{P}_2^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{P}_2)^{-1} \boldsymbol{P}_2^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{P}_1 \boldsymbol{\alpha}_1, \\ \boldsymbol{\alpha} = \boldsymbol{P}_1 \boldsymbol{\alpha}_1 + \boldsymbol{P}_2 \boldsymbol{\alpha}_2, \\ \boldsymbol{\beta} = \tau \boldsymbol{R}^{-1} \boldsymbol{P}_1^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta})\boldsymbol{\alpha}. \end{cases} \tag{34}$$

The specific implementation of the inverse iteration algorithm is listed as follows:

Step 1: initialization. Select $\boldsymbol{\alpha}_0$, $\boldsymbol{\beta}_0$, and τ_0, and construct $\boldsymbol{S}_1(\boldsymbol{\beta}_0)$ and $\boldsymbol{S}_2(\boldsymbol{\alpha}_0)$. Then, normalize them to satisfy

$$\boldsymbol{\alpha}_0^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta}_0)\boldsymbol{\alpha}_0 = \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{S}_2(\boldsymbol{\alpha}_0)\boldsymbol{\beta}_0 = 1. \tag{35}$$

Step 2: use QR decomposition of matrix $\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi}, \boldsymbol{s})$ according to Eq. (31) to obtain $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, and $\boldsymbol{R}$.

Step 3: let k:=1 and compute the following equations in turn:

(1) $\boldsymbol{\alpha}_{1k} := \tau_{k-1} \boldsymbol{R}^{-\mathrm{T}} \boldsymbol{S}_2(\boldsymbol{\alpha}_{k-1})\boldsymbol{\beta}_{k-1}$.

(2) $\boldsymbol{\alpha}_{2k} := -(\boldsymbol{P}_2^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta}_{k-1})\boldsymbol{P}_2)^{-1} \boldsymbol{P}_2^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta}_{k-1})\boldsymbol{P}_1 \boldsymbol{\alpha}_{1k}$.

(3) $\boldsymbol{\alpha}_k := \boldsymbol{P}_1 \boldsymbol{\alpha}_{1k} + \boldsymbol{P}_2 \boldsymbol{\alpha}_{2k}$ and $\boldsymbol{S}_2(\boldsymbol{\alpha}_k)$.

(4) $\boldsymbol{\beta}_k := \boldsymbol{R}^{-1} \boldsymbol{P}_1^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta}_{k-1})\boldsymbol{\alpha}_k$.

(5) $\boldsymbol{\beta}_k := \boldsymbol{\beta}_k / \| \boldsymbol{\beta}_k \|_2$ and $\boldsymbol{S}_1(\boldsymbol{\beta}_k)$.

(6) $x_k := (\boldsymbol{\alpha}_k^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta}_k)\boldsymbol{\alpha}_k)^{1/4}$.

(7) $\boldsymbol{\alpha}_k := \boldsymbol{\alpha}_k / x_k$ and $\boldsymbol{\beta}_k := \boldsymbol{\beta}_k / x_k$.

(8) $\boldsymbol{S}_1(\boldsymbol{\beta}_k) := \boldsymbol{S}_1(\boldsymbol{\beta}_k) / x_k^2$ and $\boldsymbol{S}_2(\boldsymbol{\alpha}_k) := \boldsymbol{S}_2(\boldsymbol{\alpha}_k) / x_k^2$.

(9) $\tau_k := \boldsymbol{\alpha}_k^{\mathrm{T}} \boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi}, \boldsymbol{s})\boldsymbol{\beta}_k$.

(10) $\langle \overline{\boldsymbol{e}} \rangle_m := -\tau_k \boldsymbol{\alpha}_k^{\mathrm{T}} \overline{\boldsymbol{R}}_m(\boldsymbol{\xi}, \boldsymbol{s})\boldsymbol{\beta}_k$, $1 \le m \le p + r$.

(11) $\boldsymbol{X}_k = \boldsymbol{R}_o(\boldsymbol{\xi}, \boldsymbol{s}) + \sum_{m=1}^{p+r} \langle \overline{\boldsymbol{e}} \rangle_m \cdot \overline{\boldsymbol{R}}_m(\boldsymbol{\xi}, \boldsymbol{s})$.

Step 4: calculate the maximum singular value $\sigma_{k,\max}$ and minimum singular value $\sigma_{k,\min}$ of $\boldsymbol{X}_k$. Then, assess whether $\sigma_{k,\max}/\sigma_{k,\min} < \varepsilon$ and pause the iteration if true; otherwise, let k:=k+1 and go to step 3.

Finally, if the iterative convergence value of $\boldsymbol{\beta}_k$ in the above iterative algorithm is $\hat{\boldsymbol{\beta}}$, then the solution to the STLS localization problem is $\hat{\boldsymbol{u}} = \hat{\boldsymbol{\beta}}(2: q+1) / \hat{\boldsymbol{\beta}}(1)$.

For the above iteration algorithm, three remarks are provided:

Remark 1 In step 1, the initial values $\boldsymbol{\alpha}_0$, $\boldsymbol{\beta}_0$, and τ_0 can be obtained using singular value decomposition of matrix $\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi}, \boldsymbol{s})$, where τ_0 is the minimum singular value, and $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ are the left singular vector and right singular vector, respectively, corresponding to the smallest singular value.

Remark 2 To normalize $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ according to Eq. (35), let $\boldsymbol{\alpha}_0$:=$\boldsymbol{\alpha}_0/x$ and $\boldsymbol{\beta}_0$:=$\boldsymbol{\beta}_0/y$. Thus, we have

$$\begin{cases} \boldsymbol{\alpha}_0^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta}_0)\boldsymbol{\alpha}_0 := (\boldsymbol{\alpha}_0^{\mathrm{T}} / x)(\boldsymbol{S}_1(\boldsymbol{\beta}_0) / y^2)(\boldsymbol{\alpha}_0 / x) \\ \qquad\qquad = \boldsymbol{\alpha}_0^{\mathrm{T}} \boldsymbol{S}_1(\boldsymbol{\beta}_0)\boldsymbol{\alpha}_0 / (x^2 y^2) = 1, \\ \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{S}_2(\boldsymbol{\alpha}_0)\boldsymbol{\beta}_0 := (\boldsymbol{\beta}_0^{\mathrm{T}} / x)(\boldsymbol{S}_2(\boldsymbol{\alpha}_0) / y^2)(\boldsymbol{\beta}_0 / x) \\ \qquad\qquad = \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{S}_2(\boldsymbol{\alpha}_0)\boldsymbol{\beta}_0 / (x^2 y^2) = 1. \end{cases} \tag{36}$$

For simplicity, let $x = y = (\boldsymbol{\alpha}_0^{\mathrm{T}} \boldsymbol{S}_1 \boldsymbol{\beta}_0 \boldsymbol{\alpha}_0)^{1/4} = (\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{S}_2 \boldsymbol{\alpha}_0 \boldsymbol{\beta}_0)^{1/4}$.

Remark 3 The purpose of normalizing vector $\boldsymbol{\beta}_k$ in Eq. (5) of step 3 is to make the algorithm more robust.

It is noteworthy that although our approach is based on a single moving station localization scenario, it can be easily extended to a multi-station scenario. Measurement model (11) is extensible. For a single-station scenario, $\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}) = \boldsymbol{A}_1(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})$ and $\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}) = \boldsymbol{b}_1(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})$, while for a multi-station scenario, $\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}) = [\boldsymbol{A}_1^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}), \boldsymbol{A}_2^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}), \cdots, \boldsymbol{A}_L^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})]^{\mathrm{T}}$ and $\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}) = [\boldsymbol{b}_1^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}), \boldsymbol{b}_2^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}}), \cdots, \boldsymbol{b}_L^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})]^{\mathrm{T}}$, where L is the number of observation stations. On the other hand, we establish a measurement model of a moving observation station for M measurement moments in Section 2, which is essentially equivalent to the measurement model of M moving stations for one measurement moment.

4 Performance analysis of the STLS localization algorithm

4.1 Cramér-Rao lower bound (CRLB)

Given that the measurement noise vector $\boldsymbol{e}_1$ and system error vector $\boldsymbol{e}_2$ are Gaussian distributed and independent of each other, the logarithm of the joint probability density function of $\boldsymbol{\xi}$ and $\boldsymbol{s}$ parameterized on $\boldsymbol{\rho} = [\boldsymbol{u}^{\mathrm{oT}}, \boldsymbol{s}^{\mathrm{oT}}]^{\mathrm{T}}$ is

$$\begin{aligned}\ln f(\boldsymbol{\xi},\boldsymbol{s};\boldsymbol{\rho}) &= \ln f(\boldsymbol{\xi};\boldsymbol{\rho}) + \ln f(\boldsymbol{s};\boldsymbol{\rho}) \\ &= k_1 - \frac{1}{2}(\boldsymbol{\xi}-\boldsymbol{\xi}^{\mathrm{o}})^{\mathrm{T}}\boldsymbol{Q}_1^{-1}(\boldsymbol{\xi}-\boldsymbol{\xi}^{\mathrm{o}}) \\ &\quad + k_2 - \frac{1}{2}(\boldsymbol{s}-\boldsymbol{s}^{\mathrm{o}})^{\mathrm{T}}\boldsymbol{Q}_2^{-1}(\boldsymbol{s}-\boldsymbol{s}^{\mathrm{o}}),\end{aligned} \tag{37}$$

where $k_1=-\ln[(2\pi)^{2p}|\boldsymbol{Q}_1|]/2$ and $k_2=-\ln[(2\pi)^{2r}|\boldsymbol{Q}_2|]/2$ are constants.

The CRLB of $\boldsymbol{\rho}$ is

$$\begin{aligned}\mathrm{CRLB}(\boldsymbol{\rho}) &= -E\left[\left.\frac{\partial^2 \ln f(\boldsymbol{\xi},\boldsymbol{s};\boldsymbol{\rho})}{\partial\boldsymbol{\rho}\partial\boldsymbol{\rho}^{\mathrm{T}}}\right|_{\substack{\boldsymbol{u}=\boldsymbol{u}^{\mathrm{o}}\\ \boldsymbol{s}=\boldsymbol{s}^{\mathrm{o}}}}\right]^{-1} \\ &= \begin{bmatrix}\boldsymbol{X}_{11} & \boldsymbol{X}_{12}\\ \boldsymbol{X}_{12}^{\mathrm{T}} & \boldsymbol{X}_{22}\end{bmatrix}^{-1},\end{aligned} \tag{38}$$

where

$$\begin{cases}\boldsymbol{X}_{11} = \boldsymbol{\Omega}_1^{\mathrm{T}}\boldsymbol{Q}_1^{-1}\boldsymbol{\Omega}_1,\\ \boldsymbol{X}_{12} = \boldsymbol{\Omega}_1^{\mathrm{T}}\boldsymbol{Q}_1^{-1}\boldsymbol{\Omega}_2,\\ \boldsymbol{X}_{22} = \boldsymbol{\Omega}_2^{\mathrm{T}}\boldsymbol{Q}_1^{-1}\boldsymbol{\Omega}_2 + \boldsymbol{Q}_2^{-1},\\ \boldsymbol{\Omega}_1 = \dfrac{\partial\boldsymbol{\xi}^{\mathrm{o}}}{\partial\boldsymbol{u}^{\mathrm{o}\mathrm{T}}},\\ \boldsymbol{\Omega}_2 = \dfrac{\partial\boldsymbol{\xi}^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{o}\mathrm{T}}}.\end{cases} \tag{39}$$

The detailed derivatives of $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ are shown in Appendix B. Applying the matrix inversion formula (Kay, 1993), the CRLB of $\boldsymbol{u}$ is given by

$$\begin{aligned}\mathrm{CRLB}(\boldsymbol{u}) &= (\boldsymbol{X}_{11} - \boldsymbol{X}_{12}^{\mathrm{T}}\boldsymbol{X}_{22}^{-1}\boldsymbol{X}_{12})^{-1} \\ &= [\boldsymbol{\Omega}_1^{\mathrm{T}}(\boldsymbol{Q}_1 + \boldsymbol{\Omega}_2\boldsymbol{Q}_2\boldsymbol{\Omega}_2^{\mathrm{T}})^{-1}\boldsymbol{\Omega}_1]^{-1}.\end{aligned} \tag{40}$$

The square root of the diagonal element sum in Eq. (40) is the minimum possible source location root mean square error (RMSE) that any linear unbiased estimator can achieve (Eldar, 2006).

4.2 Theoretical performance of the STLS localization algorithm

In this subsection, we demonstrate the optimal estimate performance of the STLS method using first-order perturbation analysis.

Proposition 1 The optimal solution to Eq. (27) satisfies the following constraint optimization problem:

$$\min_{\boldsymbol{r}\in\mathrm{R}^{(q+1)\times 1}} \boldsymbol{r}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{o}}^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})(\boldsymbol{S}_1(\boldsymbol{r}))^{-1}\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{r} \ \text{ s.t. } \boldsymbol{i}_{q+1}^{(1)}\boldsymbol{r}=1. \tag{41}$$

Proof Because Eq. (27) is an optimization problem with equality constraints, we construct the corresponding Lagrange function as follows:

$$\begin{aligned}J(\overline{\boldsymbol{e}},\boldsymbol{r},\boldsymbol{l},\lambda) &= \overline{\boldsymbol{e}}^{\mathrm{T}}\overline{\boldsymbol{e}} + 2\boldsymbol{l}^{\mathrm{T}}\left(\boldsymbol{R}_o(\boldsymbol{\xi},\boldsymbol{s}) + \sum_{k=1}^{p+r}\langle\overline{\boldsymbol{e}}\rangle_k\cdot\overline{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\right)\cdot\boldsymbol{r} \\ &\quad + \lambda(1-\boldsymbol{i}_{q+1}^{(1)\mathrm{T}}\boldsymbol{r}),\end{aligned} \tag{42}$$

where $\boldsymbol{l}$ and λ represent the Lagrange multiplier vector and scalar, respectively. We consider the derivation in Eq. (42) with respect to each component and make the gradient zero:

$$\begin{cases}\langle\overline{\boldsymbol{e}}\rangle_k = -\boldsymbol{l}^{\mathrm{T}}\overline{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{r},\ 1\le k\le p+r,\\ \boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{r} = \left(\sum_{k=1}^{p+r}\overline{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{r}\boldsymbol{r}^{\mathrm{T}}\overline{\boldsymbol{R}}_k^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})\right)\boldsymbol{l} = \boldsymbol{S}_1(\boldsymbol{r})\boldsymbol{l},\\ \boldsymbol{R}_{\mathrm{o}}^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{l} = \left(\sum_{k=1}^{p+r}\overline{\boldsymbol{R}}_k^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}\overline{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\right)\boldsymbol{r} = \boldsymbol{S}_2(\boldsymbol{l})\boldsymbol{r}.\end{cases} \tag{43}$$

Thus, the objective function in Eq. (27) can be expressed as

$$\|\overline{\boldsymbol{e}}\|_2^2 = \overline{\boldsymbol{e}}^{\mathrm{T}}\overline{\boldsymbol{e}} = \boldsymbol{l}^{\mathrm{T}}\left(\sum_{k=1}^{p+r}\overline{\boldsymbol{R}}_k(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{r}\boldsymbol{r}^{\mathrm{T}}\overline{\boldsymbol{R}}_k^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})\right)\boldsymbol{l} = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{r})\boldsymbol{l}. \tag{44}$$

The second equation in Eq. (43) can be rewritten as

$$\boldsymbol{l} = (\boldsymbol{S}_1(\boldsymbol{r}))^{-1}\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{r}. \tag{45}$$

Substituting Eq. (45) into Eq. (44), the cost function can be expressed as

$$\|\overline{\boldsymbol{e}}\|_2^2 = \boldsymbol{l}^{\mathrm{T}}\boldsymbol{S}_1(\boldsymbol{r})\boldsymbol{l} = \boldsymbol{r}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{o}}^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s})(\boldsymbol{S}_1(\boldsymbol{r}))^{-1}\boldsymbol{R}_{\mathrm{o}}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{r}. \tag{46}$$

At this point, the proof is completed.

Proposition 2 The localization performances of the STLS method and CTLS method are consistent, and optimal solutions $\hat{\boldsymbol{u}}_{\mathrm{STLS}}$ and $\hat{\boldsymbol{u}}_{\mathrm{CTLS}}$ both satisfy the following unconstrained optimization problem:

$$\begin{aligned}&\min_{\boldsymbol{u}\in\mathrm{R}^{q\times 1}}(\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s}) - \boldsymbol{A}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u})^{\mathrm{T}}(\boldsymbol{\Phi}(\boldsymbol{\xi},\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\xi},\boldsymbol{s},\boldsymbol{u}))^{-1} \\ &\quad\cdot(\boldsymbol{b}(\boldsymbol{\xi},\boldsymbol{s}) - \boldsymbol{A}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}).\end{aligned} \tag{47}$$

Proof Based on Eqs. (11) and (17), the CLTS optimization model can be formulated as

$$\begin{aligned}&\min \boldsymbol{e}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{e}\\ &\text{s.t.}\ \left(\boldsymbol{A}(\xi,\boldsymbol{s})-\sum_{k=1}^{p}\dot{\boldsymbol{A}}_{1k}(\xi,\boldsymbol{s})\langle\boldsymbol{e}_1\rangle_k-\sum_{k=1}^{r}\dot{\boldsymbol{A}}_{2k}(\xi,\boldsymbol{s})\langle\boldsymbol{e}_2\rangle_k\right)\boldsymbol{u}\\ &\quad=\boldsymbol{b}(\xi,\boldsymbol{s})-\boldsymbol{B}_1(\xi,\boldsymbol{s})\boldsymbol{e}_1-\boldsymbol{B}_2(\xi,\boldsymbol{s})\boldsymbol{e}_2.\end{aligned}\tag{48}$$

We define

$$\begin{cases}\boldsymbol{E}_1(\xi,\boldsymbol{s},\boldsymbol{u})=[\dot{\boldsymbol{A}}_{11}(\xi,\boldsymbol{s})\boldsymbol{u},\dot{\boldsymbol{A}}_{12}(\xi,\boldsymbol{s})\boldsymbol{u},\cdots,\dot{\boldsymbol{A}}_{1p}(\xi,\boldsymbol{s})\boldsymbol{u}],\\ \boldsymbol{E}_2(\xi,\boldsymbol{s},\boldsymbol{u})=[\dot{\boldsymbol{A}}_{21}(\xi,\boldsymbol{s})\boldsymbol{u},\dot{\boldsymbol{A}}_{22}(\xi,\boldsymbol{s})\boldsymbol{u},\cdots,\dot{\boldsymbol{A}}_{2r}(\xi,\boldsymbol{s})\boldsymbol{u}],\\ \boldsymbol{\Phi}(\xi,\boldsymbol{s},\boldsymbol{u})=[\boldsymbol{\Phi}_1(\xi,\boldsymbol{s},\boldsymbol{u}),\boldsymbol{\Phi}_2(\xi,\boldsymbol{s},\boldsymbol{u})]\\ \qquad\qquad=[\boldsymbol{B}_1(\xi,\boldsymbol{s})-\boldsymbol{E}_1(\xi,\boldsymbol{s},\boldsymbol{u}),\boldsymbol{B}_2(\xi,\boldsymbol{s})-\boldsymbol{E}_2(\xi,\boldsymbol{s},\boldsymbol{u})].\end{cases}\tag{49}$$

Then, Eq. (48) can be simplified as

$$\begin{aligned}&\min \boldsymbol{e}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{e}\\ &\text{s.t.}\ \ \boldsymbol{b}(\xi,\boldsymbol{s})-\boldsymbol{A}(\xi,\boldsymbol{s})\boldsymbol{u}=(\boldsymbol{B}_1(\xi,\boldsymbol{s})-\boldsymbol{E}_1(\xi,\boldsymbol{s},\boldsymbol{u}))\boldsymbol{e}_1\\ &\qquad\qquad+(\boldsymbol{B}_2(\xi,\boldsymbol{s})-\boldsymbol{E}_2(\xi,\boldsymbol{s},\boldsymbol{u}))\boldsymbol{e}_2\\ &\qquad\qquad=\boldsymbol{\Phi}(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{e}.\end{aligned}\tag{50}$$

Using the Lagrange multiplier method, Eq. (50) can be transformed into an unconstrained optimization problem (Eq. (47)).

Based on Eqs. (22) and (25), we have

$$\boldsymbol{R}_{\mathrm{o}}(\xi,\boldsymbol{s})\boldsymbol{r}=[-\boldsymbol{b}(\xi,\boldsymbol{s})\,|\,\boldsymbol{A}(\xi,\boldsymbol{s})]\cdot\begin{bmatrix}1\\ \boldsymbol{u}\end{bmatrix}=\boldsymbol{A}(\xi,\boldsymbol{s})\boldsymbol{u}-\boldsymbol{b}(\xi,\boldsymbol{s}),\tag{51}$$

$$\begin{aligned}\bar{\boldsymbol{R}}_k(\xi,\boldsymbol{s})\boldsymbol{r}&=\sum_{m=1}^{p+r}\langle\boldsymbol{Q}^{1/2}\rangle_{km}\cdot\boldsymbol{R}_m(\xi,\boldsymbol{s})\cdot\boldsymbol{r}\\ &=\sum_{m=1}^{p}\langle\boldsymbol{Q}_1^{1/2}\rangle_{km}\cdot[\boldsymbol{B}_1(\xi,\boldsymbol{s})\boldsymbol{i}_p^{(k)}\,|-\dot{\boldsymbol{A}}_{1k}(\xi,\boldsymbol{s})]\cdot\boldsymbol{r}\\ &=(\boldsymbol{B}_1(\xi,\boldsymbol{s})-[\dot{\boldsymbol{A}}_{11}(\xi,\boldsymbol{s})\boldsymbol{u},\dot{\boldsymbol{A}}_{12}(\xi,\boldsymbol{s})\boldsymbol{u},\cdots,\dot{\boldsymbol{A}}_{1p}(\xi,\boldsymbol{s})\boldsymbol{u}])\boldsymbol{Q}_1^{1/2}\boldsymbol{i}_p^{(k)}\\ &=\boldsymbol{\Phi}_1(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}_1^{1/2}\boldsymbol{i}_p^{(k)},\ 1\le k\le p,\end{aligned}\tag{52}$$

$$\begin{aligned}\bar{\boldsymbol{R}}_{p+k}(\xi,\boldsymbol{s})\boldsymbol{r}&=\sum_{m=1}^{p+r}\langle\boldsymbol{Q}^{1/2}\rangle_{p+k,m}\cdot\boldsymbol{R}_m(\xi,\boldsymbol{s})\\ &=\sum_{m=1}^{r}\langle\boldsymbol{Q}_2^{1/2}\rangle_{km}\cdot[\boldsymbol{B}_2(\xi,\boldsymbol{s})\boldsymbol{i}_p^{(k)}\,|-\dot{\boldsymbol{A}}_{2k}(\xi,\boldsymbol{s})]\cdot\boldsymbol{r}\\ &=(\boldsymbol{B}_2(\xi,\boldsymbol{s})-[\dot{\boldsymbol{A}}_{21}(\xi,\boldsymbol{s})\boldsymbol{u},\dot{\boldsymbol{A}}_{22}(\xi,\boldsymbol{s})\boldsymbol{u},\cdots,\dot{\boldsymbol{A}}_{2r}(\xi,\boldsymbol{s})\boldsymbol{u}])\boldsymbol{Q}_2^{1/2}\boldsymbol{i}_r^{(k)}\\ &=\boldsymbol{\Phi}_2(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}_2^{1/2}\boldsymbol{i}_r^{(k)},\ 1\le k\le r.\end{aligned}\tag{53}$$

Applying Eqs. (30), (52), and (53), $\boldsymbol{S}_1(\boldsymbol{r})$ can be reformulated as

$$\begin{aligned}\boldsymbol{S}_1(\boldsymbol{r})&=\sum_{k=1}^{p+r}\bar{\boldsymbol{R}}_k(\xi,\boldsymbol{s})\boldsymbol{r}\boldsymbol{r}^{\mathrm{T}}\bar{\boldsymbol{R}}_k^{\mathrm{T}}(\xi,\boldsymbol{s})\\ &=\sum_{k=1}^{p}\boldsymbol{\Phi}_1(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}_1^{1/2}\boldsymbol{i}_p^{(k)}\boldsymbol{i}_p^{(k)\mathrm{T}}\boldsymbol{Q}_1^{1/2}\boldsymbol{\Phi}_1^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u})\\ &\quad+\sum_{k=1}^{r}\boldsymbol{\Phi}_2(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}_2^{1/2}\boldsymbol{i}_r^{(k)}\boldsymbol{i}_r^{(k)\mathrm{T}}\boldsymbol{Q}_2^{1/2}\boldsymbol{\Phi}_2^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u})\\ &=\boldsymbol{\Phi}_1(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}_1\boldsymbol{\Phi}_1^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u})\\ &\quad+\boldsymbol{\Phi}_2(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}_2\boldsymbol{\Phi}_2^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u})\\ &=(\boldsymbol{\Phi}(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}\boldsymbol{\Phi}^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u}))^{-1}.\end{aligned}\tag{54}$$

Based on Eqs. (51) and (54), we obtain

$$\begin{aligned}&\boldsymbol{r}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{o}}^{\mathrm{T}}(\xi,\boldsymbol{s})(\boldsymbol{S}_1(\boldsymbol{r}))^{-1}\boldsymbol{R}_{\mathrm{o}}(\xi,\boldsymbol{s})\boldsymbol{r}\\ &=(\boldsymbol{A}(\xi,\boldsymbol{s})\boldsymbol{u}-\boldsymbol{b}(\xi,\boldsymbol{s}))^{\mathrm{T}}(\boldsymbol{\Phi}(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}\boldsymbol{\Phi}^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u}))^{-1}\\ &\cdot(\boldsymbol{A}(\xi,\boldsymbol{s})\boldsymbol{u}-\boldsymbol{b}(\xi,\boldsymbol{s})).\end{aligned}\tag{55}$$

At this point, the proof is completed.

The cost function in Eq. (47) can be rewritten as

$$J(\boldsymbol{u})=\boldsymbol{\varphi}^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{\Psi}(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u}),\tag{56}$$

where

$$\begin{cases}\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u})=\boldsymbol{A}(\xi,\boldsymbol{s})\boldsymbol{u}-\boldsymbol{b}(\xi,\boldsymbol{s}),\\ \boldsymbol{\Psi}(\xi,\boldsymbol{s},\boldsymbol{u})=(\boldsymbol{\Phi}(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{Q}\boldsymbol{\Phi}^{\mathrm{T}}(\xi,\boldsymbol{s},\boldsymbol{u}))^{-1}.\end{cases}\tag{57}$$

The derivation of $J(\boldsymbol{u})$ with respect to $\boldsymbol{u}$ yields the gradient vector as

$$\begin{aligned}\boldsymbol{h}(\boldsymbol{u})&=2\boldsymbol{A}^{\mathrm{T}}(\xi,\boldsymbol{s})\boldsymbol{\Psi}(\xi,\boldsymbol{s},\boldsymbol{u})\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u})\\ &\quad+\left(\frac{\partial\mathrm{vec}\left(\boldsymbol{\Psi}(\xi,\boldsymbol{s},\boldsymbol{u})\right)}{\partial\boldsymbol{u}^{\mathrm{T}}}\right)^{\mathrm{T}}\left(\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u})\otimes\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u})\right).\end{aligned}\tag{58}$$

Assume that the solution of the estimator, denoted by $\boldsymbol{u}_{\mathrm{STLS}}$, is the global minimum, which makes the gradient zero. Then, we have

$$\begin{aligned}&2\boldsymbol{A}^{\mathrm{T}}(\xi)\boldsymbol{\Psi}(\xi,\boldsymbol{s},\boldsymbol{u}_{\mathrm{STLS}})\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u}_{\mathrm{STLS}})\\ &+\left(\left.\frac{\partial\mathrm{vec}\left(\boldsymbol{\Psi}(\xi,\boldsymbol{s},\boldsymbol{u})\right)}{\partial\boldsymbol{u}^{\mathrm{T}}}\right|_{\boldsymbol{u}=\boldsymbol{u}_{\mathrm{STLS}}}\right)^{\mathrm{T}}\\ &\cdot(\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u}_{\mathrm{STLS}})\otimes\boldsymbol{\varphi}(\xi,\boldsymbol{s},\boldsymbol{u}_{\mathrm{STLS}}))=\boldsymbol{0}_{3N\times1}.\end{aligned}\tag{59}$$

The estimate can be expressed as the true value plus the error:

$$\boldsymbol{u}_{\mathrm{STLS}}=\boldsymbol{u}^{\mathrm{o}}+\Delta\boldsymbol{u}_{\mathrm{STLS}}. \tag{60}$$

Substituting Eq. (59) into $\boldsymbol{\varphi}(\boldsymbol{\xi}, \boldsymbol{s}, \boldsymbol{u})$ in Eq. (57) and ignoring the second-order error term yield

$$\begin{aligned}
&\boldsymbol{\varphi}(\boldsymbol{\xi},\boldsymbol{s},\boldsymbol{u}_{\mathrm{STLS}})\\
&=\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})\boldsymbol{u}^{\mathrm{o}}-\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})+\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}})\Delta\boldsymbol{u}_{\mathrm{STLS}}\\
&\quad+([\dot{\boldsymbol{A}}_{11}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\dot{\boldsymbol{A}}_{12}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\cdots\quad\dot{\boldsymbol{A}}_{1p}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}]-\boldsymbol{B}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}}))\boldsymbol{e}_1\\
&\quad+([\dot{\boldsymbol{A}}_{21}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\dot{\boldsymbol{A}}_{22}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\cdots\quad\dot{\boldsymbol{A}}_{2r}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}]-\boldsymbol{B}_2(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}}))\boldsymbol{e}_2\\
&=\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})\Delta\boldsymbol{u}_{\mathrm{STLS}}-\boldsymbol{\Phi}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{e}_1-\boldsymbol{\Phi}_2(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{e}_2\\
&=\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})\Delta\boldsymbol{u}_{\mathrm{STLS}}-\boldsymbol{\Phi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{e}.
\end{aligned} \tag{61}$$

Applying Eq. (61) and omitting the second term in Eq. (59), we have

$$\begin{aligned}
\boldsymbol{0}_{3N\times1}\approx{}&\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}})\boldsymbol{\Psi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})(\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}})\Delta\boldsymbol{u}_{\mathrm{STLS}}\\
&-\boldsymbol{\Phi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{e}).
\end{aligned} \tag{62}$$

Consequently, the estimation error is given by

$$\begin{aligned}
\Delta\boldsymbol{u}_{\mathrm{STLS}}\approx{}&(\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}})\boldsymbol{\Psi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}}))^{-1}\\
&\cdot\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}})\boldsymbol{\Psi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{\Phi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{e}.
\end{aligned} \tag{63}$$

Multiplying Eq. (63) by its transpose and taking the expectation yield the covariance matrix as

$$\begin{aligned}
\mathbf{cov}(\Delta\boldsymbol{u}_{\mathrm{STLS}})&=E[\Delta\boldsymbol{u}_{\mathrm{STLS}}\Delta\boldsymbol{u}_{\mathrm{STLS}}^{\mathrm{T}}]\\
&=(\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})\boldsymbol{\Psi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}}))^{-1}\\
&=(\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})(\boldsymbol{\Phi}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{Q}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}}))^{-1}\\
&\quad\cdot\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}}))^{-1}.
\end{aligned} \tag{64}$$

4.3 Asymptotic optimal performance of the STLS localization algorithm

In this subsection, we analyze the asymptotic performance of the STLS method. Substituting $\boldsymbol{\xi}^{\mathrm{o}}=\boldsymbol{\xi}^{\mathrm{o}}(\boldsymbol{s}^{\mathrm{o}}, \boldsymbol{u}^{\mathrm{o}})$ into Eq. (11) yields

$$\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}}(\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}}),\boldsymbol{s}^{\mathrm{o}})\boldsymbol{u}^{\mathrm{o}}=\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}}(\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}}),\boldsymbol{s}^{\mathrm{o}}). \tag{65}$$

Computing the Jacobi matrix (Golub and van Loan, 2012) of both sides of Eq. (65) with respect to $\boldsymbol{u}^{\mathrm{o}}$ leads to

$$\begin{aligned}
&\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})+[\dot{\boldsymbol{A}}_{11}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\dot{\boldsymbol{A}}_{12}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\cdots\quad\dot{\boldsymbol{A}}_{1p}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}]\boldsymbol{\Omega}_1\\
&=\boldsymbol{B}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})\boldsymbol{\Omega}_1.
\end{aligned} \tag{66}$$

Based on Eqs. (49) and (66), $\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}}, \boldsymbol{s}^{\mathrm{o}})$ is given by

$$\begin{aligned}
\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})&=(\boldsymbol{B}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})\\
&\quad-[\dot{\boldsymbol{A}}_{11}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\dot{\boldsymbol{A}}_{12}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\cdots\quad\dot{\boldsymbol{A}}_{1p}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}])\boldsymbol{\Omega}_1\\
&=\boldsymbol{\Phi}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{\Omega}_1.
\end{aligned} \tag{67}$$

Calculating the Jacobi matrix with respect to $\boldsymbol{s}^{\mathrm{o}}$ again, Eq. (65) are rewritten as

$$\begin{aligned}
&[\dot{\boldsymbol{A}}_{11}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\dot{\boldsymbol{A}}_{12}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\cdots\quad\dot{\boldsymbol{A}}_{1p}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}]\boldsymbol{\Omega}_2\\
&+[\dot{\boldsymbol{A}}_{21}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\dot{\boldsymbol{A}}_{22}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}\quad\cdots\quad\dot{\boldsymbol{A}}_{2r}(\boldsymbol{\xi},\boldsymbol{s})\boldsymbol{u}]\\
&=\boldsymbol{B}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})\boldsymbol{\Omega}_2+\boldsymbol{B}_2(\boldsymbol{\xi}^{\mathrm{o}}).
\end{aligned} \tag{68}$$

With some algebraic manipulations, we have the following equality relationship:

$$-\boldsymbol{\Phi}_1^{-1}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{\Phi}_2(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})=\boldsymbol{\Omega}_2. \tag{69}$$

Combining Eqs. (67) and (69), Eq. (64) can be simplified as

$$\begin{aligned}
&\mathbf{cov}(\Delta\boldsymbol{u}_{\mathrm{CTLS}})\\
&=\left(\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}})\begin{pmatrix}\boldsymbol{\Phi}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{Q}_1\boldsymbol{\Phi}_1^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\\+\boldsymbol{\Phi}_2(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{Q}_2\boldsymbol{\Phi}_2^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\end{pmatrix}^{-1}\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}})\right)^{-1}\\
&=\left(\begin{pmatrix}\boldsymbol{\Omega}_1^{\mathrm{T}}\boldsymbol{\Phi}_1^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\begin{pmatrix}\boldsymbol{\Phi}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{Q}_1\boldsymbol{\Phi}_1^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\\+\boldsymbol{\Phi}_2(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{Q}_2\boldsymbol{\Phi}_2^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\end{pmatrix}^{-1}\\\cdot\boldsymbol{\Phi}_1(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{\Omega}_1\end{pmatrix}^{-1}\right)^{-1}\\
&=\left(\boldsymbol{\Omega}_1^{\mathrm{T}}\begin{pmatrix}\boldsymbol{Q}_1+\boldsymbol{\Phi}_1^{-1}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{\Phi}_2(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{Q}_2\\\cdot\boldsymbol{\Phi}_2^{\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\boldsymbol{\Phi}_1^{-\mathrm{T}}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}},\boldsymbol{u}^{\mathrm{o}})\end{pmatrix}^{-1}\boldsymbol{\Omega}_1\right)^{-1}\\
&=\left(\boldsymbol{\Omega}_1^{\mathrm{T}}(\boldsymbol{Q}_1+\boldsymbol{\Omega}_2\boldsymbol{Q}_2\boldsymbol{\Omega}_2^{\mathrm{T}})^{-1}\boldsymbol{\Omega}_1\right)^{-1}\\
&=\mathrm{CRLB}(\boldsymbol{u}).
\end{aligned} \tag{70}$$

In conclusion, the proposed method has asymptotic optimal performance and can achieve the CRLB accuracy for a moderate noise level.

5 Simulations

In the following simulations, the proposed method (labeled 'STLS') is compared with the TLS method (labeled 'TLS'), the CTLS localization algorithm (labeled 'CTLS'), and the Taylor series iterative algorithm (labeled 'Taylor') in the presence of the measurement noise and system error. We illustrate that both measurement noise and system error are Gaussian distributed. To verify the validity of theoretical analysis in the text, the following graphs also show the theoretical performance curve of the STLS localization algorithm and the curve of CRLB. Besides, in the following simulation scenarios. we assume that the single moving station is equipped with a uniform circular array (UCA), which is comprised by omni-directional antennas (such as monopole and dipole antennas).

The upper and lower bounds of TOA and FOA measurements are closely related to the localization scene. For example, in the ultra-short wave band, the communication distance is about 20 km, and thus the upper bound of TOA is about 66.7 μs. In the short wave band, the communication distance is about 300 km, so the upper bound of TOA is about 1 ms. The upper and lower bounds of FOA measurement are closely related to not only the wave band but also the speed of the targets and observation stations. In the following simulation scenario, we assume the upper and lower bounds of AOA, TOA, and FOA as

$$\begin{cases} 0° \le \theta_{i,j}^{\mathrm{o}} \le 360°, \\ 0° \le \beta_{i,j}^{\mathrm{o}} \le 90°, \\ 5 \le \tau_{i,j} \le 20\ (\mu\mathrm{s}), \\ 3 \le f_{i,j} \le 10\ (\mathrm{Hz}). \end{cases} \tag{71}$$

5.1 Comparison of localization performance for different algorithms

We examine the estimation accuracy of the source position in a 3D scenario in which a single observation station is driving. We assume that three measurements are obtained during its movement and its locations and velocities are listed in Table 2. There are two radiation sources that are located at $\boldsymbol{u}_1^{\mathrm{o}}$ =[2500, 2000, 1000]$^{\mathrm{T}}$ m and $\boldsymbol{u}_2^{\mathrm{o}}$ =[1500, 1000, −1500]$^{\mathrm{T}}$ m, and their carrier frequencies are $f_{\mathrm{c}1}$=50 MHz and $f_{\mathrm{c}2}$=100 MHz, respectively, belonging to the ultra-short wave band. The localization geometry is shown in Fig. 1. The open pentagons indicate the locations of the disjoint sources and the open circles indicate the true station positions at three measurements. The noise powers at different measurement moments are unequal with $\boldsymbol{R}_{\mathrm{s}}=\sigma_{\mathrm{s}}^2\cdot\mathrm{blkdiag}(\boldsymbol{I}_3,\ \boldsymbol{0}_{3\times3},\ 3\boldsymbol{I}_3,\ \boldsymbol{0}_{3\times3},\ 5\boldsymbol{I}_3,\ \boldsymbol{0}_{3\times3})$ m and $\boldsymbol{R}_{\dot{\mathrm{s}}}=\sigma_{\dot{\mathrm{s}}}^2\cdot\mathrm{blkdiag}(\boldsymbol{0}_{3\times3},\ 2\boldsymbol{I}_3,\ \boldsymbol{0}_{3\times3},\ 4\boldsymbol{I}_3,\ \boldsymbol{0}_{3\times3},\ \boldsymbol{I}_3)$ m/s, and $Q_2=\boldsymbol{R}_{\mathrm{s}}+\boldsymbol{R}_{\dot{\mathrm{s}}}$. The noise powers of AOA, TOA, and FOA measurements are $\boldsymbol{R}_{\mathrm{a}}=\sigma_{\mathrm{AOA}}^2\cdot\mathrm{blkdiag}(\boldsymbol{I}_6,\ \boldsymbol{0}_{6\times6},\ 2\boldsymbol{I}_6,\ \boldsymbol{0}_{6\times6})$ m, $\boldsymbol{R}_{\mathrm{t}}=\sigma_{\mathrm{TOA}}^2\cdot\mathrm{blkdiag}(\boldsymbol{0}_{6\times6},\ \boldsymbol{I}_3,\ \boldsymbol{0}_{9\times9},\ 3\boldsymbol{I}_3,\ \boldsymbol{0}_{3\times3})$ m and $\boldsymbol{R}_{\mathrm{f}}=\sigma_{\mathrm{FOA}}^2c^2\cdot\mathrm{blkdiag}(\boldsymbol{0}_{9\times9},\ \boldsymbol{I}_3,\ \boldsymbol{0}_{9\times9},\ 4\boldsymbol{I}_3)$ m/s, respectively, and $\boldsymbol{Q}_1=\boldsymbol{R}_{\mathrm{a}}+\boldsymbol{R}_{\mathrm{t}}+\boldsymbol{R}_{\mathrm{f}}$. We first fix the noise power of the station position and velocity as σ_{s}^2 =100 m^2 and $\sigma_{\dot{\mathrm{s}}}^2=0.1$ m^2/s^2, respectively, and let $\sigma_{\mathrm{AOA}}^2=10^{-4}\sigma_{\mathrm{m}}^2$ rad^2, $\sigma_{\mathrm{TOA}}^2c^2=\sigma_{\mathrm{m}}^2$ m^2, and $\sigma_{\mathrm{FOA}}^2c^2=10^{-3}\sigma_{\mathrm{m}}^2$ m^2/s^2, where σ_{m}^2 represents the measurement noise power which varies from −20 dB to 15 dB in 5 dB intervals. We conduct N=500 Monte Carlo (MC) experiments and the estimation accuracy is evaluated in terms of $\mathrm{RMSE}(\boldsymbol{u})=\sqrt{\frac{1}{N}\sum_{n=1}^{N}\|\boldsymbol{u}-\boldsymbol{u}^{\mathrm{o}}\|^2}$, which is shown in log scale as the noise power increases.

Table 2 Location and velocity of the observation station at different measurement time

Parameter	Measurement value		
	1	2	3
x_i^{o} (m)	1200	300	−1500
y_i^{o} (m)	−500	−1100	600
z_i^{o} (m)	150	−350	−250
$\dot{x}_i^{\mathrm{o}}$ (m/s)	−20	10	20
$\dot{y}_i^{\mathrm{o}}$ (m/s)	30	−20	−20
$\dot{z}_i^{\mathrm{o}}$ (m/s)	10	−30	−10

The localization accuracy of the proposed solution is shown in Fig. 2, from which we can observe that the TLS algorithm has the worst position estimation performance and always deviates from CRLB because of the numerical structure mismatch of the coefficient matrix in Eq. (11). Additionally, with the exception of the TLS method, the STLS, CTLS, and

Taylor series methods attain CRLB at a low noise level. As the measurement noise increases, the CTLS and Taylor series methods perform worse, and the thresholding effect generally appears successively. As expected from Proposition 2, under certain conditions, the STLS and CTLS algorithms have the same theoretical localization performance. However, the numerical stability of the STLS method is higher than that of the CTLS method, which can be interpreted as the numerical stability of singular value decomposition being higher than that of the Newton iteration. After $10\log(\sigma_{\mathrm{m}}^2)$ reaches 15 dB, the proposed STLS method also suffers from the thresholding effect.

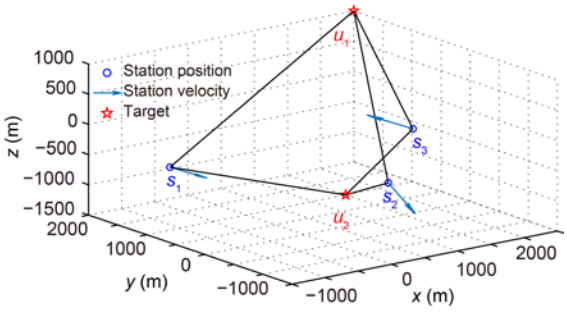

Fig. 1 Localization geometry

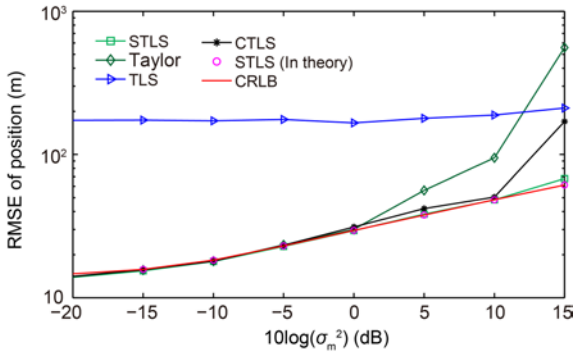

Fig. 2 Curves of source position RMSEs as the measurement noise varies

Next, we assume that σ_{m}^2 =5 dB, so $\sigma_{\mathrm{AOA}}^2=10^{-3.5}$ rad^2, $\sigma_{\mathrm{TOA}}^2 c^2=10^{0.5}$ m^2, and $\sigma_{\mathrm{FOA}}^2 c^2=10^{-2.5}\sigma_{\mathrm{m}}^2$ m^2/s^2. Let $\sigma_{\mathrm{s}}^2=\sigma_0^2$ m^2 and $\sigma_{\mathrm{s}}^2=\sigma_0^2/200$ m^2/s^2, where σ_0^2 represents the system error power which varies from −10 dB to 30 dB in 5 dB intervals. The simulation results are shown in Fig. 3, from which we can obtain a conclusion similar to that drawn from Fig. 2; that is, the three methods can achieve the CRLB under moderate noise.

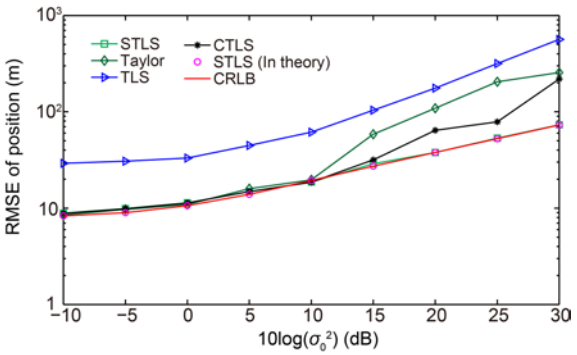

Fig. 3 Curves of source position RMSEs as the system error varies

With the increase of the system error, the Taylor series method suffers from the thresholding effect at $10\log(\sigma_0^2)$ =15 dB, possibly because the initial source positions of the Taylor series method (we randomly select the initial values within a 1000-m radius of the true value) are far from the true locations. Different from Fig. 2, the proposed STLS method never suffers from the thresholding effect within the range of the given system error. This difference demonstrates that the STLS method is more robust to the system error than the CTLS and Taylor series method. It is evident from Figs. 2 and 3 that the STLS method demonstrates superior performance. It can attain the accuracy set by CRLB for both sources before the threshold effect occurs, and its theoretical RMSE coincides with CRLB perfectly. These phenomena verify the former analysis.

5.2 Comparison of localization performance for a single source and multiple sources

In this subsection, the superiority of the proposed method in a multi-source scenario is demonstrated. For simplicity, a two-source scenario is considered in which the two sources are located at $\boldsymbol{u}_1^{\mathrm{o}}$ =[2500, 2000, 1000]$^{\mathrm{T}}$ m and $\boldsymbol{u}_2^{\mathrm{o}}$ =[1500, 1000, 1500]$^{\mathrm{T}}$ m. The simulation conditions are similar to those in the case in Section 5.1, and only one source position is of interest. Figs. 4 and 5 show the results of the proposed solution for the source $\boldsymbol{u}_1^{\mathrm{o}}$. The symbol STLS-a represents the RMSE of $\boldsymbol{u}_1$, which is obtained by combining the measurements of unknown source $\boldsymbol{u}_2$ to estimate the locations of $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. STLS-b represents the RMSE of $\boldsymbol{u}_1$ derived from conventional

single source positioning. The traces of their CRLBs are also shown for comparison.

It is evident that compared with the previously developed single source localization algorithm, the proposed algorithm can achieve a lower CRLB; that is, lower optimal performance is obtained. From Fig. 4, with the increase of the measurement noise σ_{m}^2, the performance of the positioning solution STLS-a is significantly superior to that of STLS-b. The performance gain of STLS-a reaches 17 m over the positioning solution STLS-b when $10\log(\sigma_{\mathrm{m}}^2)$ =10 dB. The significant gain in localization accuracy may be attributed to the effect of the measurements from source $\boldsymbol{u}_2$ with the same position and velocity displacements. However, as measurement noise σ_{m}^2 further increases, the positioning solutions STLS-a and STLS-b start deviating from their respective CRLB and suffering from the thresholding effect at $10\log(\sigma_{\mathrm{m}}^2)$ =15 dB. In Fig. 5, when the measurement noise is certain, as the system error increases, the separation between STLS-a and STLS-b indicates that the greater the influence of the public position and velocity displacements of measurements from different sources, the greater the cooperative location gain. These phenomena demonstrate that for the multi-source scenario, when only one source position is of interest, the proposed algorithm can still explore the measurements from other unknown sources with the same position and velocity displacements, to improve source localization accuracy under moderate measurement noise or a large system error. However, the performance improvement comes at the cost of increasing computational complexity from the joint estimation of multiple sources.

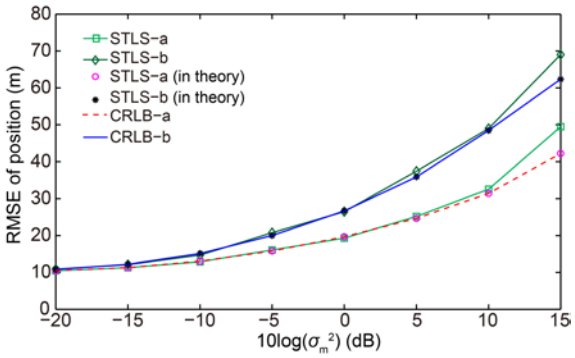

Fig. 4 Curves of source position RMSEs as the measurement noise varies

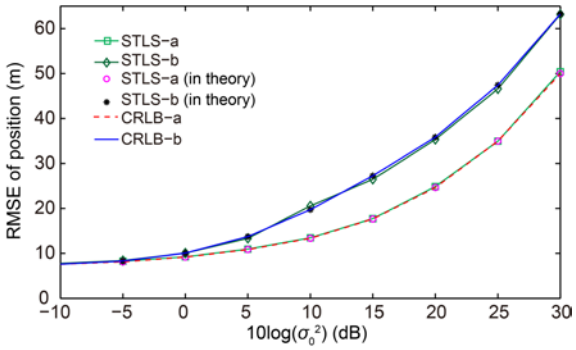

Fig. 5 Curves of source position RMSEs as the system error varies

Next, the performance of the proposed algorithm is examined in a multi-source scenario as the distance from the sources to the station changes (Fig. 6). The simulation conditions are similar to those in the case in Section 5.1, and only the source position $\boldsymbol{u}_1^{\mathrm{o}}$ is of interest. k represents the distance factor which varies from 1 to 20, and the larger the value, the greater the distance from the sources to the station. From Fig. 6 it is observed that both methods attain CRLB at any distance factor. However, the proposed algorithm STLS-a has a much better theoretical localization accuracy than the general single source positioning algorithm STLS-b. Fig. 6 also indicates that the larger the distance from the sources to the station, the greater the cooperative localization gain.

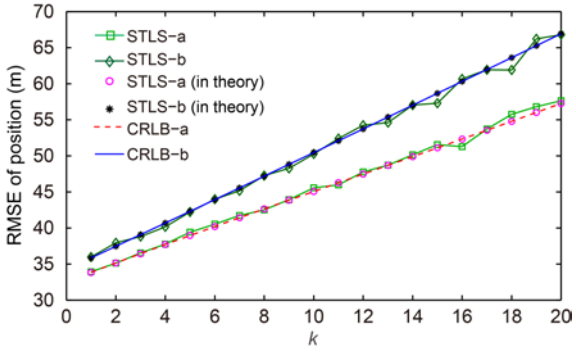

Fig. 6 Curves of source position RMSEs vs. the distance factor

5.3 Comparison of computational complexity

To highlight the superiority of the STLS method, the complexity of the proposed method is compared with those of the Taylor series method and CTLS localization method. Proposition 1 has proved that the STLS method and CTLS method have the same

localization error under certain conditions, and the Taylor series method can also achieve CRLB before the threshold effect occurs when a good initial value is obtained. However, by comparing the computations of the three methods, we determined that each iteration of the STLS method requires a $(p-q-1)$-order matrix inversion (Eq. (2) in step 3). In contrast, each iteration of the CTLS method (the variable-step Newton iteration method is used to solve this problem) requires a number of p-order matrix inversions and a number of q-order matrix inversions. Additionally, each iteration of the Taylor series requires a p-order matrix inversion and q-order matrix inversion. Thus, the computation quantity of the STLS method is less than that of the CTLS and Taylor series methods. Additionally, through a simulation, it is found that under the given simulation conditions, the number of iterations of the STLS method is generally smaller than five, which is smaller than that of the CTLS and Taylor series methods. A 3D scenario is considered with the same simulation conditions as those in the case in Section 5.1. Simultaneously, we set the noise power of station position and velocity as $\sigma_s^2=100$ m^2 and $\sigma_{\dot{s}}^2=0.1$ m^2/s^2, respectively, and let $\sigma_{\mathrm{AOA}}^2=10^{-4}$ rad^2, $\sigma_{\mathrm{TOA}}^2 c^2=1$ m^2, and $\sigma_{\mathrm{FOA}}^2 c^2=10^{-3}$ m^2/s^2. The simulations are based on 500 MC runs. The RMSEs, average iteration numbers, and total running time of three methods are shown in Table 3. Then, the noise power of the station position and velocity are changed to $\sigma_s^2=1000$ m^2 and $\sigma_0^2=1$ m^2/s^2, respectively, and the simulation results are shown in Table 4.

It can be observed from Table 3 that there are no significant differences in the localization errors for the above three algorithms. However, the STLS method has a smaller average iteration number and a shorter total running time than the other two methods, which confirms that the STLS method requires a smaller iteration number and fewer calculations. From Table 4, the RMSEs of the STLS and Taylor series methods are much greater than that of the STLS method. This demonstrates that, with the increase of the system error, the solutions to the CTLS method and Taylor series method are prone to deterioration, and the solution to the STLS method is more stable. So, we can conclude that the STLS method is more robust to system errors than the CTLS and Taylor series methods.

Table 3 Comparison of three algorithms in calculation with $\sigma_s^2=100$ m^2 and $\sigma_0^2=0.1$ m^2/s^2

Method	RMSE of position (m)	Average iteration number	Total running time (s)
STLS	37.7781	1.1700	3.8286
CTLS	37.7245	8.2067	477.7594
Taylor series	37.7098	9.2233	4.8978

Table 4 Comparison of three algorithms in calculation with $\sigma_s^2=1000$ m^2 and $\sigma_0^2=1$ m^2/s^2

Method	RMSE of position (m)	Average iteration number	Total running time (s)
STLS	72.1213	1.9333	5.1219
CTLS	205.0680	9.1833	532.6493
Taylor series	337.7194	9.8600	5.2545

5.4 Localization performance under a multi-station scenario

In this subsection, we aim at verifying that the proposed algorithm can be applied to a multi-station scenario. It is assumed that there are four available array observation stations, which can receive and locate the radiated source signals, and their locations and velocities are listed in Table 5. The positions of two radiation sources are $\boldsymbol{u}_1^{\mathrm{o}}=[2500, 2000, 1000]^{\mathrm{T}}$ m and $\boldsymbol{u}_2^{\mathrm{o}}=[1500, 1000, -1500]^{\mathrm{T}}$ m. The localization geometry is illustrated in Fig. 7. Then, let $\boldsymbol{Q}_1=\mathrm{blkdiag}(\sigma_{\mathrm{AOA}}^2\boldsymbol{I}_8,\ \sigma_{\mathrm{TOA}}^2c^2\boldsymbol{I}_4,\ \sigma_{\mathrm{FOA}}^2c^2\boldsymbol{I}_4,\ 2\sigma_{\mathrm{AOA}}^2\boldsymbol{I}_8,\ 3\sigma_{\mathrm{TOA}}^2c^2\boldsymbol{I}_4,\ 4\sigma_{\mathrm{FOA}}^2c^2\boldsymbol{I}_4)$ and $\boldsymbol{Q}_2=\mathrm{blkdiag}\ (\sigma_s^2\boldsymbol{I}_3,\ 2\sigma_{\dot{s}}^2\boldsymbol{I}_3, 3\sigma_s^2\boldsymbol{I}_3, 4\sigma_{\dot{s}}^2\boldsymbol{I}_3, 5\sigma_s^2\boldsymbol{I}_3, 3\sigma_{\dot{s}}^2\boldsymbol{I}_3, 2\sigma_s^2\boldsymbol{I}_3, \sigma_{\dot{s}}^2\boldsymbol{I}_3)$.

Table 5 Location and velocity of the observation stations at one measurement moment for four different observation stations

Parameter	Measurement value			
	1	2	3	4
x_i^{o} (m)	1200	300	−1500	600
y_i^{o} (m)	−500	−1100	600	3000
z_i^{o} (m)	150	−350	−250	−500
$\dot{x}_i^{\mathrm{o}}$ (m/s)	−20	10	20	−20
$\dot{y}_i^{\mathrm{o}}$ (m/s)	30	−20	−20	−10
$\dot{z}_i^{\mathrm{o}}$ (m/s)	10	−30	−10	30

All other simulation conditions are the same as those described in Section 5.1. The simulation results are shown in Figs. 8 and 9. As expected, the proposed algorithm can achieve CRLB under moderate noise, and suffers from the thresholding effect later than other algorithms. From these figures, we can conclude that, the proposed algorithm is still valid and outperforms the other localization algorithms under the condition of multiple stations.

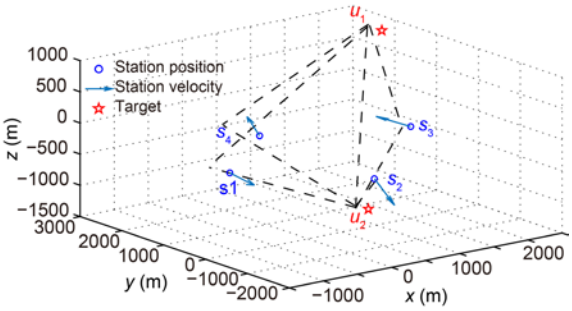

Fig. 7 Localization geometry under the multi-station scenario

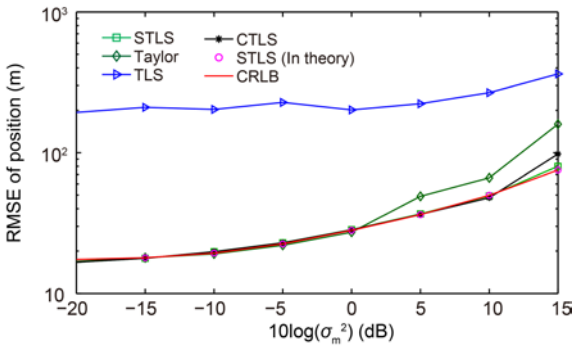

Fig. 8 Curves of source position RMSEs as the measurement noise varies

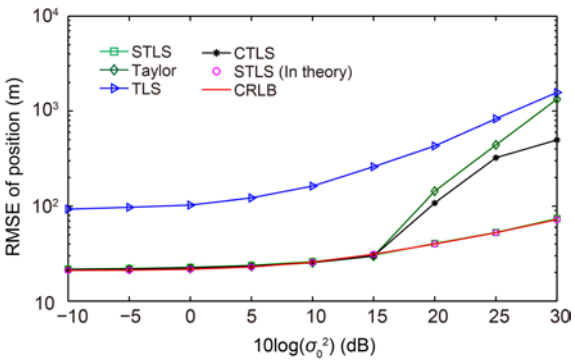

Fig. 9 Curves of source position RMSEs as the system error varies

6 Conclusions

Based on the STLS localization algorithm with only angle information (Wang et al., 2009), in this study, we extended the STLS algorithm to the single-station passive location of multiple disjoint sources, and the observation station position and velocity errors were considered. The study began with the establishment of the measurement model according to the available measurements: AOA, TOA, and FOA. Then, we deduced the corresponding pseudo linear equations, based on which the STLS localization model was formulated and the inverse iteration algorithm was used to solve this problem. Moreover, it was strictly proved that the theoretical performance of the STLS method and CTLS method was consistent under first-order error analysis, and coincided with that of CRLB. Several simulations based on AOA/TOA/FOA measurements were designed to verify the superiority of the new algorithm and the validity of theoretical derivation. For the multi-source scenario, when only one source position is of interest, the proposed algorithm can still exploit the same station position and velocity displacements of measurements from other unknown sources to improve the source localization accuracy. Furthermore, simulation results illustrated that the cooperative positioning gain is proportional to the distance from the sources to the station, and that the proposed localization algorithm requires less computation than the CTLS and Taylor series algorithms. Finally, the proposed algorithm can be easily extended to the multi-station scenario.

References

Abatzoglou TJ, Mendel JM, Harada GA, 1991. The constrained total least squares technique and its applications to harmonic superresolution. *IEEE Trans Signal Process*, 39(5):1070-1087. https://doi.org/10.1109/78.80955

Antreich F, Nossek JA, Seco-Granados G, et al., 2011. The extended invariance principle for signal parameter estimation in an unknown spatial field. *IEEE Trans Signal Process*, 59(7):3213-3225. https://doi.org/10.1109/TSP.2011.2140107

Bamler R, 1991. Doppler frequency estimation and the Cramer-Rao bound. *IEEE Trans Geosci Remote Sens*, 29(3):385-390. https://doi.org/10.1109/36.79429

Bar-Shalom O, Weiss AJ, 2014. Emitter geolocation using single moving receiver. *Signal Process*, 105:70-83. https://doi.org/10.1016/j.sigpro.2014.05.006

Belouchrani A, Aouada S, 2003. Maximum likelihood joint angle and delay estimation in unknown noise fields. IEEE

Int Conf on Acoustics, Speech, and Signal Processing, p.V-265. https://doi.org/10.1109/ICASSP.2003.1199919

Carevic D, 2007. Automatic estimation of multiple target positions and velocities using passive TDOA measurements of transients. *IEEE Trans Signal Process*, 55(2): 424-436. https://doi.org/10.1109/TSP.2006.885745

Chalise BK, Zhang YD, Amin MG, et al., 2014. Target localization in a multi-static passive radar system through convex optimization. *Signal Process*, 102:207-215. https://doi.org/10.1016/j.sigpro.2014.02.023

de Moor B, 1994. Total least squares for affinely structured matrices and the noisy realization problem. *IEEE Trans Signal Process*, 42(11):3104-3113. https://doi.org/10.1109/78.330370

Doğançay K, 2005. Bearings-only target localization using total least squares. *Signal Process*, 85(9):1695-1710. https://doi.org/10.1016/j.sigpro.2005.03.007

Eldar YC, 2006. Uniformly improving the CramÉr-Rao bound and maximum-likelihood estimation. *IEEE Trans Signal Process*, 54(8):2943-2956. https://doi.org/10.1109/TSP.2006.877648

Foy WH, 1976. Position-location solutions by Taylor-series estimation. *IEEE Trans Aerosp Electron Syst*, AES-12(2): 187-194. https://doi.org/10.1109/TAES.1976.308294

Giacometti R, Baussard A, Cornu C, et al., 2016. Accuracy studies for TDOA-AOA localization of emitters with a single sensor. IEEE Radar Conf, p.1-4. https://doi.org/10.1109/RADAR.2016.7485185

Golub GH, van Loan CF, 2012. Matrix Computations (4th Ed.). The Johns Hopkins University Press, Baltimore, USA.

Griffin A, Alexandridis A, Pavlidi D, et al., 2015. Localizing multiple audio sources in a wireless acoustic sensor network. *Signal Process*, 107:54-67. https://doi.org/10.1016/j.sigpro.2014.08.013

Hao BJ, Li Z, Si JB, et al., 2012. Passive multiple disjoint sources localization using TDOAs and GROAs in the presence of sensor location uncertainties. IEEE Int Conf on Communications, p.47-52. https://doi.org/10.1109/ICC.2012.6364164

Ho KC, Lu XN, Kovavisaruch L, 2007. Source localization using TDOA and FDOA measurements in the presence of receiver location errors: analysis and solution. *IEEE Trans Signal Process*, 55(2):684-696. https://doi.org/10.1109/TSP.2006.885744

Jia TY, Wang HY, Shen XH, et al., 2017. Bearing-only multiple sources localization and the spatial spectrum. OCEANS, p.1-5. https://doi.org/10.1109/OCEANSE.2017.8084819

Kay SM, 1993. Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory. Prentice Hall, Englewood Cliffs, New Jersey, USA, p.465-466.

Knapp C, Carter G, 1976. The generalized correlation method for estimation of time delay. *IEEE Trans Acoust Speech Signal Process*, 24(4):320-327. https://doi.org/10.1109/TASSP.1976.1162830

Lemma AN, van der Veen AJ, Deprettere EF, 2003. Analysis of joint angle-frequency estimation using ESPRIT. *IEEE Trans Signal Process*, 51(5):1264-1283. https://doi.org/10.1109/TSP.2003.810306

Lemmerling P, de Moor B, van Huffel S, 1996. On the equivalence of constrained total least squares and structured total least squares. *IEEE Trans Signal Process*, 44(11): 2908-2911. https://doi.org/10.1109/78.542454

Li J, Zhao YJ, Li DH, 2014. Accurate single-observer passive coherent location estimation based on TDOA and DOA. *Chin J Aeronaut*, 27(4):913-923. https://doi.org/10.1016/j.cja.2014.06.004

Li JZ, Pang HW, Guo FC, et al., 2015. Localization of multiple disjoint sources with prior knowledge on source locations in the presence of sensor location errors. *Dig Signal Process*, 40:181-197. https://doi.org/10.1016/j.dsp.2015.02.003

Li MH, Lu YL, 2008. Angle-of-arrival estimation for localization and communication in wireless networks. 16th European Signal Processing Conf, p.1-5.

Liu FL, Wang JK, Du RY, 2010. Unitary-JAFE algorithm for joint angle-frequency estimation based on Frame-Newton method. *Signal Process*, 90(3):809-820. https://doi.org/10.1016/j.sigpro.2009.08.013

Liu RR, Wang YL, Yin JX, et al., 2017. Passive source localization using importance sampling based on TOA and FOA measurements. *Front Inform Technol Electron Eng*, 18(8):1167-1179. https://doi.org/10.1631/FITEE.1601657

Luise M, Reggiannini R, 1995. Carrier frequency recovery in all-digital modems for burst-mode transmissions. *IEEE Trans Commun*, 43(2-4):1169-1178. https://doi.org/10.1109/26.380149

Markovsky I, van Huffel S, 2007. Overview of total least-squares methods. *Signal Process*, 87(10):2283-2302. https://doi.org/10.1016/j.sigpro.2007.04.004

Mir HS, Sahr JD, Hatke GF, et al., 2007. Passive source localization using an airborne sensor array in the presence of manifold perturbations. *IEEE Trans Signal Process*, 55(6):2486-2496. https://doi.org/10.1109/TSP.2007.893936

Raykar VC, Kozintsev IV, Lienhart R, 2005. Position calibration of microphones and loudspeakers in distributed computing platforms. *IEEE Trans Speech Audio Process*, 13(1):70-83. https://doi.org/10.1109/TSA.2004.838540

Rockah Y, Schultheiss PM, 1987a. Array shape calibration using sources in unknown location—Part I: far-field sources. *IEEE Trans Acoust Speech Signal Process*, 35(3): 286-299. https://doi.org/10.1109/TASSP.1987.1165144

Rockah Y, Schultheiss PM, 1987b. Array shape calibration using sources in unknown locations—Part II: near-field sources and estimator implementation. *IEEE Trans Acoust Speech Signal Process*, 35(6):724-735. https://doi.org/10.1109/TASSP.1987.1165222

Shen H, Ding Z, Dasgupta S, et al., 2014. Multiple source localization in wireless sensor networks based on time of arrival measurement. *IEEE Trans Signal Process*, 62(8):

1938-1949. https://doi.org/10.1109/TSP.2014.2304433

So HC, Lin LX, 2011. Linear least squares approach for accurate received signal strength based source localization. *IEEE Trans Signal Process*, 59(8):4035-4040. https://doi.org/10.1109/TSP.2011.2152400

Sun M, Ho KC, 2011. An asymptotically efficient estimator for TDOA and FDOA positioning of multiple disjoint sources in the presence of sensor location uncertainties. *IEEE Trans Signal Process*, 59(7):3434-3440. https://doi.org/10.1109/TSP.2011.2131135

Sun XY, Li JD, Huang PY, et al., 2008. Total least-squares solution of active target localization using TDOA and FDOA measurements in WSN. 22nd Int Conf on Advanced Information Networking and Applications, p.995-999. https://doi.org/10.1109/WAINA.2008.150

Torrieri DJ, 1984. Statistical theory of passive location systems. *IEEE Trans Aerosp Electron Syst*, AES-20(2): 183-198. https://doi.org/10.1109/TAES.1984.310439

Wang D, Zhang L, Wu Y, 2007. Constrained total least squares algorithm for passive location based on bearing-only measurements. *Sci China Ser F Inform Sci*, 50(4): 576-586. https://doi.org/10.1007/s11432-007-0023-8

Wang D, Zhang L, Wu Y, 2009. The structured total least squares algorithm research for passive location based on angle information. *Sci China Ser F Inform Sci*, 52(6): 1043-1054. https://doi.org/10.1007/s11432-009-0114-9

Wang L, Hu AQ, Bai GW, et al., 2011. Enhanced constrained total least squares estimator in TDOA localization for WSNs. Int Conf on Computer Science and Service System, p.389-392. https://doi.org/10.1109/CSSS.2011.5974564

Wang YY, Chen JT, Fang WH, 2001. TST-MUSIC for joint DOA-delay estimation. *IEEE Trans Signal Process*, 49(4):721-729. https://doi.org/10.1109/78.912916

Weiss AJ, Friedlander B, 1989. Array shape calibration using sources in unknown locations—a maximum likelihood approach. *IEEE Trans Acoust Speech Signal Process*, 37(12):1958-1966. https://doi.org/10.1109/29.45542

Wu H, Chen SX, Zhang YH, et al., 2015. Robust structured total least squares algorithm for passive location. *J Syst Eng Electron*, 26(5):946-953. https://doi.org/10.1109/JSEE.2015.00103

Yang K, An JP, Bu XY, et al., 2010. Constrained total least-squares location algorithm using time-difference-of-arrival measurements. *IEEE Trans Veh Technol*, 59(3): 1558-1562. https://doi.org/10.1109/TVT.2009.2037509

Yang L, Ho KC, 2009. An approximately efficient TDOA localization algorithm in closed-form for locating multiple disjoint sources with erroneous sensor positions. *IEEE Trans Signal Process*, 57(12):4598-4615. https://doi.org/10.1109/TSP.2009.2027765

Yin JX, Wu Y, Wang D, 2014. On 2-D direction-of-arrival estimation performance for rank reduction estimator in presence of unexpected modeling errors. *Circ Syst Signal Process*, 33(2):515-547. https://doi.org/10.1007/s00034-013-9654-8

Yu H, Huang G, Gao J, 2012. Constrained total least-squares localisation algorithm using time difference of arrival and frequency difference of arrival measurements with sensor location uncertainties. *IET Radar Sonar Navig*, 6(9): 891-899. https://doi.org/10.1049/iet-rsn.2011.0205

Appendix A: Composition of $\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})$ and $\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})$

$$\boldsymbol{A}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})=\mathrm{blkdiag}\left(\begin{bmatrix}\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{1,1}^{\mathrm{o}})\\ \vdots\\ \boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}})\\ \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{1,1}^{o},\beta_{1,1}^{o})\\ \vdots\\ \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}},\beta_{M,1}^{\mathrm{o}})\\ \boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,1}^{o},\beta_{1,1}^{o})\\ \vdots\\ \boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}},\beta_{M,1}^{\mathrm{o}})\\ \dot{r}_{1,1}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,1}^{\mathrm{o}},\beta_{1,1}^{\mathrm{o}})+\dot{\boldsymbol{s}}_1^{\mathrm{oT}}\\ \vdots\\ \dot{r}_{M,1}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}},\beta_{M,1}^{\mathrm{o}})+\dot{\boldsymbol{s}}_M^{\mathrm{oT}}\end{bmatrix},\begin{bmatrix}\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{1,2}^{\mathrm{o}})\\ \vdots\\ \boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}})\\ \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{1,2}^{\mathrm{o}},\beta_{1,2}^{\mathrm{o}})\\ \vdots\\ \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}},\beta_{M,2}^{\mathrm{o}})\\ \boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,2}^{\mathrm{o}},\beta_{1,2}^{\mathrm{o}})\\ \vdots\\ \boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}},\beta_{M,2}^{\mathrm{o}})\\ \dot{r}_{1,2}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,2}^{\mathrm{o}},\beta_{1,2}^{\mathrm{o}})+\dot{\boldsymbol{s}}_1^{\mathrm{oT}}\\ \vdots\\ \dot{r}_{M,2}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}},\beta_{M,2}^{\mathrm{o}})+\dot{\boldsymbol{s}}_M^{\mathrm{oT}}\end{bmatrix},\cdots,\begin{bmatrix}\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{1,N}^{\mathrm{o}})\\ \vdots\\ \boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}})\\ \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{1,N}^{\mathrm{o}},,\beta_{1,N}^{\mathrm{o}})\\ \vdots\\ \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}},\beta_{M,N}^{\mathrm{o}})\\ \boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,N}^{\mathrm{o}},\beta_{1,N}^{\mathrm{o}})\\ \vdots\\ \boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}},\beta_{M,N}^{\mathrm{o}})\\ \dot{r}_{1,N}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,N}^{\mathrm{o}},\beta_{1,N}^{\mathrm{o}})+\dot{\boldsymbol{s}}_1^{\mathrm{oT}}\\ \vdots\\ \dot{r}_{M,N}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}},\beta_{M,N}^{\mathrm{o}})+\dot{\boldsymbol{s}}_M^{\mathrm{oT}}\end{bmatrix}\right),\quad(\mathrm{A1})$$

$$\boldsymbol{b}(\boldsymbol{\xi}^{\mathrm{o}},\boldsymbol{s}^{\mathrm{o}})=\begin{bmatrix}
\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{1,1}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}} & \boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{1,2}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}} & \cdots & \boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{1,N}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}}\\
\vdots & \vdots & & \vdots\\
\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}} & \boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}} & \cdots & \boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}}\\
\boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{1,1}^{o},\beta_{1,1}^{o})\boldsymbol{s}_1^{\mathrm{o}} & \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{1,2}^{o},\beta_{1,2}^{o})\boldsymbol{s}_1^{\mathrm{o}} & \cdots & \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{1,N}^{o},\beta_{1,N}^{o})\boldsymbol{s}_1^{\mathrm{o}}\\
\vdots & \vdots & & \vdots\\
\boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}},\beta_{M,1}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}} & \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}},\beta_{M,2}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}} & \cdots & \boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}},\beta_{M,N}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}}\\
r_1^{\mathrm{o}}+\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,1}^{\mathrm{o}},\beta_{1,1}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}} & r_1^{\mathrm{o}}+\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,2}^{\mathrm{o}},\beta_{1,2}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}} & \cdots & r_1^{\mathrm{o}}+\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,N}^{\mathrm{o}},\beta_{1,N}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}}\\
\vdots & \vdots & & \vdots\\
r_{M,1}^{\mathrm{o}}+\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}},\beta_{M,1}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}} & r_{M,2}^{\mathrm{o}}+\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}},\beta_{M,2}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}} & \cdots & r_{M,N}^{\mathrm{o}}+\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}},\beta_{M,N}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}}\\
\dot{r}_{1,1}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,1}^{\mathrm{o}},\beta_{1,1}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}}+\dot{\boldsymbol{s}}_1^{\mathrm{oT}}\boldsymbol{s}_1^{\mathrm{o}} & \dot{r}_{1,2}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,2}^{\mathrm{o}},\beta_{1,2}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}}+\dot{\boldsymbol{s}}_1^{\mathrm{oT}}\boldsymbol{s}_1^{\mathrm{o}} & \cdots & \dot{r}_{1,N}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,N}^{\mathrm{o}},\beta_{1,N}^{\mathrm{o}})\boldsymbol{s}_1^{\mathrm{o}}+\dot{\boldsymbol{s}}_1^{\mathrm{oT}}\boldsymbol{s}_1^{\mathrm{o}}\\
\vdots & \vdots & & \vdots\\
\dot{r}_{M,1}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,1}^{\mathrm{o}},\beta_{M,1}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}}+\dot{\boldsymbol{s}}_M^{\mathrm{oT}}\boldsymbol{s}_M^{\mathrm{o}} & \dot{r}_{M,2}^{\mathrm{o}}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,2}^{\mathrm{o}},\beta_{M,2}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}}+\dot{\boldsymbol{s}}_M^{\mathrm{oT}}\boldsymbol{s}_M^{\mathrm{o}} & \cdots & \dot{r}_{M,N}^{o}\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,N}^{\mathrm{o}},\beta_{M,N}^{\mathrm{o}})\boldsymbol{s}_M^{\mathrm{o}}+\dot{\boldsymbol{s}}_M^{\mathrm{oT}}\boldsymbol{s}_M^{\mathrm{o}}
\end{bmatrix}. \tag{A2}$$

Appendix B: Expressions of $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$

The detailed derivatives of $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ are addressed below:

$$\begin{cases}
\boldsymbol{\Omega}_1=\dfrac{\partial\boldsymbol{\xi}^{\mathrm{o}}}{\partial\boldsymbol{u}^{\mathrm{oT}}}=\left[\dfrac{\partial\boldsymbol{\xi}_1^{\mathrm{o}}}{\partial\boldsymbol{u}^{o\mathrm{T}}};\dfrac{\partial\boldsymbol{\xi}_2^{\mathrm{o}}}{\partial\boldsymbol{u}^{o\mathrm{T}}};\ldots;\dfrac{\partial\boldsymbol{\xi}_N^{\mathrm{o}}}{\partial\boldsymbol{u}^{\mathrm{oT}}}\right]=\mathrm{diag}\left(\dfrac{\partial\boldsymbol{\xi}_1^{\mathrm{o}}}{\partial\boldsymbol{u}_1^{\mathrm{oT}}},\dfrac{\partial\boldsymbol{\xi}_2^{\mathrm{o}}}{\partial\boldsymbol{u}_2^{\mathrm{oT}}},\ldots,\dfrac{\partial\boldsymbol{\xi}_N^{\mathrm{o}}}{\partial\boldsymbol{u}_N^{\mathrm{oT}}}\right),\\
\boldsymbol{\Omega}_2=\dfrac{\partial\boldsymbol{\xi}^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}}=\left[\dfrac{\partial\boldsymbol{\xi}_1^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}};\dfrac{\partial\boldsymbol{\xi}_2^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}};\ldots;\dfrac{\partial\boldsymbol{\xi}_N^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}}\right],
\end{cases} \tag{B1}$$

$$\begin{cases}
\dfrac{\partial\boldsymbol{\xi}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}}=\left[\dfrac{\partial\boldsymbol{\theta}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}};\dfrac{\partial\boldsymbol{\beta}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}};\dfrac{\partial\boldsymbol{r}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}};\dfrac{\partial\dot{\boldsymbol{r}}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}}\right],\\
\dfrac{\partial\boldsymbol{\xi}_j^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}}=\left[\dfrac{\partial\boldsymbol{\theta}_j^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}};\dfrac{\partial\boldsymbol{\beta}_j^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}};\dfrac{\partial\boldsymbol{r}_j^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}};\dfrac{\partial\dot{\boldsymbol{r}}_j^{\mathrm{o}}}{\partial\boldsymbol{s}^{\mathrm{oT}}}\right],
\end{cases} \tag{B2}$$

$$\begin{cases}
\dfrac{\partial\boldsymbol{\theta}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}}=\mathrm{diag}(\|[\boldsymbol{I}_2,\boldsymbol{0}_{2\times1}]\cdot(\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_1^{\mathrm{o}})\|^{-1},\|[\boldsymbol{I}_2,\boldsymbol{0}_{2\times1}]\cdot(\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_2^{\mathrm{o}})\|^{-1},\ldots,\|[\boldsymbol{I}_2,\boldsymbol{0}_{2\times1}]\cdot(\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_M^{\mathrm{o}})\|^{-1}]\\
\qquad\cdot[-\sin\boldsymbol{\theta}_j^{\mathrm{o}},\cos\boldsymbol{\theta}_j^{\mathrm{o}},\boldsymbol{0}_{M\times1}],\\
\dfrac{\partial\boldsymbol{\beta}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}}=\mathrm{diag}(\|\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_1^{\mathrm{o}}\|^{-1},\|\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_2^{\mathrm{o}}\|^{-1},\ldots,\|\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_M^{\mathrm{o}}\|^{-1}]\cdot[-\cos\boldsymbol{\theta}_j^{\mathrm{o}}\cdot\sin\boldsymbol{\beta}_j^{\mathrm{o}},-\sin\boldsymbol{\theta}_j^{\mathrm{o}}\cdot\sin\boldsymbol{\beta}_j^{\mathrm{o}},\cos\boldsymbol{\beta}_j^{\mathrm{o}}],\\
\dfrac{\partial\boldsymbol{r}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}}=[\cos\boldsymbol{\theta}_j^{\mathrm{o}}\cdot\cos\boldsymbol{\beta}_j^{\mathrm{o}},\sin\boldsymbol{\theta}_j^{\mathrm{o}}\cdot\cos\boldsymbol{\beta}_j^{\mathrm{o}},\sin\boldsymbol{\beta}_j^{\mathrm{o}}],\\
\dfrac{\partial\dot{\boldsymbol{r}}_j^{\mathrm{o}}}{\partial\boldsymbol{u}_j^{\mathrm{oT}}}=\mathrm{diag}(\|(\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_1^{\mathrm{o}})\|^{-1},\|(\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_2^{\mathrm{o}})\|^{-1},\ldots,\|(\boldsymbol{u}_j^{\mathrm{o}}-\boldsymbol{s}_M^{\mathrm{o}})\|^{-1}]\cdot\begin{pmatrix}-[\dot{\boldsymbol{s}}_1^{\mathrm{o}},\dot{\boldsymbol{s}}_2^{\mathrm{o}},\ldots,\dot{\boldsymbol{s}}_M^{\mathrm{o}}]^{\mathrm{T}}-\mathrm{diag}[\dot{\boldsymbol{r}}_1^{\mathrm{o}},\dot{\boldsymbol{r}}_2^{\mathrm{o}},\ldots,\dot{\boldsymbol{r}}_M^{\mathrm{o}}]\cdot\\ [\cos\boldsymbol{\theta}_j^{\mathrm{o}}\cdot\cos\boldsymbol{\beta}_j^{\mathrm{o}},\sin\boldsymbol{\theta}_j^{\mathrm{o}}\cdot\cos\boldsymbol{\beta}_j^{\mathrm{o}},\sin\boldsymbol{\beta}_j^{\mathrm{o}}]\end{pmatrix},
\end{cases} \tag{B3}$$

$$
\begin{cases}
\dfrac{\partial \boldsymbol{\theta}_j^{\mathrm{o}}}{\partial \boldsymbol{s}^{\mathrm{oT}}} = \mathrm{diag}(\| [\boldsymbol{I}_2, \boldsymbol{0}_{2\times 1}] \cdot (\boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_1^{\mathrm{o}}) \|^{-1}, \| [\boldsymbol{I}_2, \boldsymbol{0}_{2\times 1}] \cdot_2 (\boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_2^{\mathrm{o}}) \|^{-1}, \ldots, \| [\boldsymbol{I}_2, \boldsymbol{0}_{2\times 1}] \cdot_2 (\boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_M^{\mathrm{o}}) \|^{-1}) \\
\qquad \cdot \mathrm{blkdiag}([\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{1,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], [\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{2,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], \ldots, [\boldsymbol{\omega}_1^{\mathrm{T}}(\theta_{M,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}]), \\
\dfrac{\partial \boldsymbol{\beta}_j^{\mathrm{o}}}{\partial \boldsymbol{s}^{\mathrm{oT}}} = \mathrm{diag}(\| \boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_1^{\mathrm{o}} \|^{-1}, \| \boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_2^{\mathrm{o}} \|^{-1}, \ldots, \| \boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_M^{\mathrm{o}} \|^{-1}) \\
\qquad \cdot \mathrm{blkdiag}([\boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{1,j}^{\mathrm{o}}, \beta_{1,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], [\boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{2,j}^{\mathrm{o}}, \beta_{2,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], \ldots, [\boldsymbol{\omega}_2^{\mathrm{T}}(\theta_{M,j}^{\mathrm{o}}, \beta_{M,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}]), \\
\dfrac{\partial \boldsymbol{r}_j^{\mathrm{o}}}{\partial \boldsymbol{s}^{\mathrm{oT}}} = \mathrm{blkdiag}([-\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,j}^{\mathrm{o}}, \beta_{1,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], [-\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{2,j}^{\mathrm{o}}, \beta_{2,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], \ldots, [-\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,j}^{\mathrm{o}}, \beta_{M,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}]), \\
\dfrac{\partial \dot{\boldsymbol{r}}_j^{\mathrm{o}}}{\partial \boldsymbol{s}^{\mathrm{oT}}} = \mathrm{diag}(\| (\boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_1^{\mathrm{o}}) \|^{-1}, \| (\boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_2^{\mathrm{o}}) \|^{-1}, \ldots, \| (\boldsymbol{u}_j^{\mathrm{o}} - \boldsymbol{s}_M^{\mathrm{o}}) \|^{-1}) \\
\qquad \cdot \begin{pmatrix} \mathrm{diag}(\dot{\boldsymbol{r}}_{1,j}^{\mathrm{o}}, \dot{\boldsymbol{r}}_{2,j}^{\mathrm{o}}, \ldots, \dot{\boldsymbol{r}}_{M,j}^{\mathrm{o}}) \cdot [[\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{1,j}^{\mathrm{o}}, \beta_{1,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], [\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{2,j}^{\mathrm{o}}, \beta_{2,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}], \ldots, [\boldsymbol{\omega}_3^{\mathrm{T}}(\theta_{M,j}^{\mathrm{o}}, \beta_{M,j}^{\mathrm{o}}), \boldsymbol{0}_{1\times 3}]] \\ +\mathrm{blkdiag}([\dot{\boldsymbol{s}}_1^{\mathrm{oT}}, \boldsymbol{s}_1^{\mathrm{oT}} - \boldsymbol{u}_j^{\mathrm{oT}}], [\dot{\boldsymbol{s}}_2^{\mathrm{oT}}, \boldsymbol{s}_2^{\mathrm{oT}} - \boldsymbol{u}_j^{\mathrm{oT}}], \ldots, [\dot{\boldsymbol{s}}_M^{\mathrm{oT}}, \boldsymbol{s}_M^{\mathrm{oT}} - \boldsymbol{u}_j^{\mathrm{oT}}]) \end{pmatrix}.
\end{cases}
\tag{B4}
$$