

System error iterative identification for underwater positioning based on spectral clustering

LU Yu, WANG Jiongqi, HE Zhangming^{*}, ZHOU Haiyin, XING Yao, and ZHOU Xuanying

College of Sciences, National University of Defense Technology, Changsha 410073, China

Abstract: The observation error model of the underwater acoustic positioning system is an important factor to influence the positioning accuracy of the underwater target. For the position inconsistency error caused by considering the underwater target as a mass point, as well as the observation system error, the traditional error model best estimation trajectory (EMBT) with little observed data and too many parameters can lead to the ill-condition of the parameter model. In this paper, a multi-station fusion system error model based on the optimal polynomial constraint is constructed, and the corresponding observation system error identification based on improved spectral clustering is designed. Firstly, the reduced parameter unified modeling for the underwater target position parameters and the system error is achieved through the polynomial optimization. Then a multi-station non-oriented graph network is established, which can address the problem of the inaccurate identification for the system errors. Moreover, the similarity matrix of the spectral clustering is improved, and the iterative identification for the system errors based on the improved spectral clustering is proposed. Finally, the comprehensive measured data of long baseline lake test and sea test show that the proposed method can accurately identify the system errors, and moreover can improve the positioning accuracy for the underwater target positioning.

Keywords: acoustic positioning, reduced parameter, system error identification, improved spectral clustering, accuracy analysis.

DOI: [10.23919/JSEE.2024.000069](https://doi.org/10.23919/JSEE.2024.000069)

1. Introduction

The high accuracy measuring and positioning for the underwater targets has become an important issue in the underwater resource and environment detection, the underwater measurement and control, etc. For the under-

water target, the strong absorption of the radio waves limits the application of global navigation satellite system (GNSS). The underwater acoustic navigation and positioning technology has become an important technical mean for the underwater target positioning because of the good propagation characteristics of the acoustic waves in water [1–4].

With the development of the underwater acoustic navigation and positioning technology, the representative systems such as long baseline (LBL) system, short baseline (SBL) system and ultra SBL (USBL) system have emerged [5,6]. Among them, both the LBL and SBL systems obtain the positions of underwater targets by measuring the propagation time of hydroacoustic signals, which can improve the positioning accuracy for the underwater target by increasing the number of acoustic base stations. The USBL system determines the target position by measuring the propagation time and the azimuth angle [7,8].

The acoustic positioning accuracy of the underwater targets depends on many factors, which include the measurement accuracy of acoustic stations, the geometric dilution of precision (GDOP) of underwater positioning systems (i.e., station layout and configuration design for underwater positioning systems), and the observation system error such as the position errors of underwater stations, the sound velocity error and the time delay error. The uncertainty of the observation system error model caused by the complex underwater environment is the key factor to restrict the high accuracy acoustic navigation and positioning [9].

The observation error can be divided into random error, gross error, and system error. The system error has obvious deterministic and regularity, which has obvious negative effects in the practical navigation and positioning applications. Effective identification and compensation for the system error is an effective way to restrict the negative impacts [10,11].

Manuscript received August 11, 2022.

^{*}Corresponding author.

This work was supported by the National Natural Science Foundation of China (61903086;61903366;62001115), the Natural Science Foundation of Hunan Province (2019JJ50745;2020JJ4280;2021JJ40133), and the Fundamentals and Basic of Applications Research Foundation of Guangdong Province (2019A1515110136).

Taking the underwater acoustic positioning system based on the distance intersection as an example, there are two main aspects for the system error. (i) The position inconsistency error caused by ignoring the structure of the target and regarding the target as a particle. For example, in the process of cylindrical target positioning, ignoring the position inconsistency between the beacon and the target center will lead to systematic deviation. (ii) The measure system error, such as acoustic time delay, position error (i.e., installation and calibration deviation), and sound velocity correlated error, which can be effectively weakened or suppressed by fine calibration and model improvement [12,13].

At present, the identification and compensation for the underwater acoustic positioning system errors mainly include differential elimination, system error modeling, and unified estimation. Xin constructed the single difference and the double difference positioning models to eliminate the influence of the system error [14]. The main forms of system error modeling are error fitting, observation constraints, and parameter solving. Yang et al. used the relationship between the range error and the time for the system error modeling, which can eliminate the influence of the representative errors in sound velocity profiles [15]. Other effective ways to restrict the representative error of sound velocity is to construct a three-dimensional acoustic field model or add some constraints [16]. In addition, in the case of the accurate sound velocity profile, sound tracking is a direct method to eliminate the influence of refractive effect [17]. Xin et al. proposed an underwater positioning method based on the equivalent sound velocity [18]. In [18], the high accuracy sound tracking and positioning solution without the sound velocity profile was realized by iteratively solving the equivalent sound velocity gradient and the position of the target as unknown parameters. In terms of the unified estimation, the most commonly used method is the error model best estimation trajectory (EMBET). The EMBET method based on the spline restraint on the basis of the traditional EMBET method was studied in [19], and the purposes to estimate trajectory, calibrate the system error coefficients and identify the dynamic performance indexes of the equipment were realized.

However, the traditional EMBET method only combines the observation equation and the system error model, which does not make full use of the trajectory kinematics characteristics. The EMBET method based on the spline restraint or polynomial restraint needs to clarify the form of system error. The polynomial order in the traditional EMBET model is predetermined. Adding too many system error parameters increases the complexity of the model, and may also lead to the ill-condition of the

parameter model. Therefore, an optimization criterion is established to adaptively select the polynomial order, and the optimal polynomial model is used to represent the target trajectory. The position errors between the target center and the beacons and the measurement error are unified to decrease the parameters.

After the model is established, the system error of each station is related to the measurement accuracy and the distance between the station and the target. Therefore, the selection of the number of the system errors can be transformed into the classification of the stations. The traditional partitioning-based clustering methods such as the K-means algorithm, the K-means++ algorithm and the model-based clustering such as Gaussian mixture model (GMM) clustering algorithm need to know the number of clusters in advance [20]. However, in the underwater LBL fusion positioning, it is difficult to determine the appropriate identification number of the system error without prior knowledge. Too few or too many system error parameters will make the identification results inaccurate [21]. Therefore, this paper improves the similarity matrix of the spectral clustering, and applies it to system errors classification and identification in proposed model.

The main contributions of this paper are as follows.

(i) Unified modeling. Considering that too many parameters will lead to the ill-condition for the traditional EMBET model, the optimal selection criterion of the fusion model is proposed for the underwater positioning. A multi-station fusion system error model based on the optimal polynomial constraint is constructed. The unified modeling with the target trajectory parameters and the system error parameters is realized by multi-station, multi-beacon, and multi-time fusion processing.

(ii) System errors classification and identification. In view of the problem that the system errors identification is not accurate or overfitting in the process of the underwater cylindrical target positioning, the classification of the system errors for the stations only according to the relative distances between the stations and the target fails to consider the connection of the system errors between the stations. Therefore, this paper establishes a multi-station non-oriented graph network. Then the similarity matrix of the spectral clustering is improved by comprehensively considering the measurement accuracy, the rotation angle of the target and the system errors. On this basis, the iterative identification method of the system errors based on the improved spectral clustering is put forward.

(iii) The model and method proposed in this paper are proved to improve the target positioning accuracy through the comprehensive measured data test of the LBL lake test and the sea test.

2. Multi-station fusion system error model based on optimal polynomial constraints

The EMBET method is widely used in fusion estimation [22]. Using the measurement data of multiple stations, the joint equations between the target coordinate and the measurement data are established, and the target position parameters and the measurement system errors are calculated by the least square method. However, too many parameters will lead to the ill-condition of the model since the traditional EMBET method does not make full use of the matching information. The target trajectory can be accurately expressed by the time function, such as the polynomial and spline function. However, the traditional polynomial model cannot adaptively match the trajectory of the target because of its order determination. In this section, the principle of the LBL positioning is briefly analyzed. Then, combined with the traditional EMBET model, the optimal model selection criterion is used to adaptively select the order of the polynomial function to represent the motion model for the underwater target, which realizes the target position and velocity matching. In addition, considering the position errors between the target center and the beacons, the measurement element errors which include the time delay error, the sound velocity correlated error and the station position error are equivalent to the constant system error of each acoustic station. Further, the unified modeling of reduced parameters for the target trajectory and the system error is realized.

2.1 Underwater LBL positioning

The underwater LBL system consists of the acoustic stations laid on the seafloor, the acoustic beacons mounted on the target, and a calibration vessel. As shown in Fig. 1, more than three acoustic stations are laid on the seafloor to form a seafloor positioning baseline array with a certain geometry. The positions of the acoustic stations are pre-calibrated by the calibration vessel. The acoustic stations are used to receive the signal emitted by the beacons and measure the propagation time of the signal. Assume that the target is a cylinder within the baseline array, and the target center position is $\mathbf{X} = (x, y, z)^T$. Suppose that the positions of m acoustic stations are $\mathbf{X}_i = (x_i, y_i, z_i)^T (i = 1, 2, \dots, m)$. Then the measurement equation of the i th acoustic station is as follows:

$$R_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} + \Delta_i + \varepsilon_i \quad (1)$$

where R_i is the observation of the i th acoustic station, Δ_i is the observation system error of the i th acoustic station caused by the clock error, and ε_i is the measurement random error.

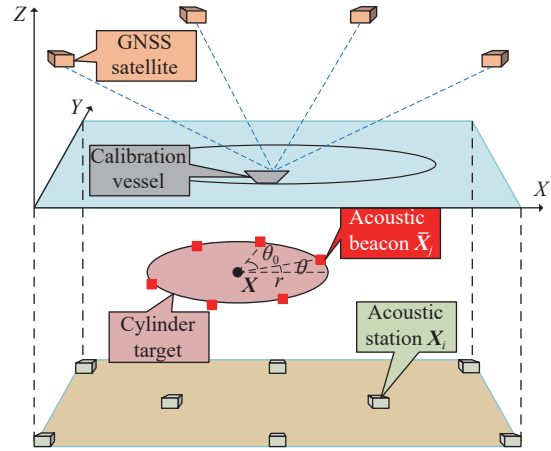


Fig. 1 Underwater LBL positioning system

However, with the increasing complexity of the underwater maneuvering target motion and the improvement of the underwater observation accuracy, the method of describing the underwater target as a particle cannot satisfy the needs of high-accuracy positioning [23]. Suppose that n beacons are uniformly mounted on the surface of the cylinder, as shown in Fig. 1, the positions of the beacons are $\bar{\mathbf{X}}_j = (\bar{x}_j, \bar{y}_j, \bar{z}_j)^T (j = 1, 2, \dots, n)$. Assume that the cylinder rotates during the motion with the initial phase θ . Then the conversion relationship between the position of the j th beacon and the cylinder center position is described as follows:

$$\begin{cases} \bar{x}_j = x + r \cos((j-1)\theta_0 + \theta) \\ \bar{y}_j = y + r \sin((j-1)\theta_0 + \theta) \\ \bar{z}_j = z \end{cases} \quad (2)$$

where $j = 1, 2, \dots, n$, r is the radius of the cylinder, and θ_0 is the angle of the two lines connecting the two adjacent beacons and the cylinder center.

It is reasonable to assume that the signal received by the i th acoustic station at a certain time is from the j th beacon, where j is random. Then, the measurement equation will be

$$R_{ij} = \sqrt{(x_i - \bar{x}_j)^2 + (y_i - \bar{y}_j)^2 + (z_i - \bar{z}_j)^2} + \Delta_i + \varepsilon_i. \quad (3)$$

Substitute (2) into (3), and rewrite it in functional form as follows:

$$R_{ij} = h_i(x, y, z, \theta, \Delta_i, \varepsilon_i). \quad (4)$$

2.2 Multi-station fusion system error model

Combined with the measurement equations of m acoustic stations at time t , the measurement equations of the target center position, the initial phase and the system errors at time t can be obtained.

$$\begin{pmatrix} R_{1j_1} \\ R_{2j_2} \\ \vdots \\ R_{mj_m} \end{pmatrix} = \begin{pmatrix} h_1^{(i)}(x, y, z, \theta, \Delta_1, \varepsilon_1) \\ h_2^{(i)}(x, y, z, \theta, \Delta_2, \varepsilon_2) \\ \vdots \\ h_m^{(i)}(x, y, z, \theta, \Delta_m, \varepsilon_m) \end{pmatrix} \quad (5)$$

where j_1, j_2, \dots, j_m represent the serial numbers of the beacons received by m acoustic stations respectively at time t . Let

$$\mathbf{Y}_e = (R_{1j_1}, R_{2j_2}, \dots, R_{mj_m})^T. \quad (6)$$

And combine N sampling points to obtain the EMBET model as follows:

$$\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_N \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1(x, y, z, \theta, \Delta_1, \Delta_2, \dots, \Delta_m, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_m) \\ \mathbf{F}_2(x, y, z, \theta, \Delta_1, \Delta_2, \dots, \Delta_m, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_m) \\ \vdots \\ \mathbf{F}_N(x, y, z, \theta, \Delta_1, \Delta_2, \dots, \Delta_m, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_m) \end{pmatrix} = \begin{pmatrix} (h_1^{(1)}, h_2^{(1)}, \dots, h_m^{(1)})^T \\ (h_1^{(2)}, h_2^{(2)}, \dots, h_m^{(2)})^T \\ \vdots \\ (h_1^{(N)}, h_2^{(N)}, \dots, h_m^{(N)})^T \end{pmatrix}. \quad (7)$$

The trajectory kinematics characteristics and the matching information are not involved in the traditional EMBET model shown in (7). Little observed data and too many parameters will lead to the ill-condition of the model. The target trajectory constrained by the polynomial function based on the optimal selection criterion is given below, and then the optimal polynomial coefficients, the target rotation angle and the system errors are unified modeling to decrease the parameter numbers.

The target moves fast and short in the shallow sea area. Therefore, the target trajectory approximately satisfies the polynomial constraints, that is, the target position $\mathbf{X} = (x, y, z)^T$ satisfies the following equation:

$$\begin{cases} x = f_1(a_1, b_1, \dots, t) \\ y = f_2(a_2, b_2, \dots, t) \\ z = f_3(a_3, b_3, \dots, t) \end{cases} \quad (8)$$

where f_1, f_2, f_3 are the polynomial functions with indefinite order, a and b are the polynomial parameters, t is the time. The polynomial orders are determined by the optimal model selection criterion. Substitute the target center position $\mathbf{X} = (x, y, z)^T$ which satisfies (8) into (3). Equating the delay measurement error, the station position error and the sound velocity correlated error to a constant error, the measurement equation will be rewritten as follows:

$$\begin{cases} R_{ij} = \sqrt{(x_i - \bar{x}_j)^2 + (y_i - \bar{y}_j)^2 + (z_i - \bar{z}_j)^2} + \Delta R_i \\ \bar{x}_j = x + r \cos((j-1)\theta_0 + \theta) \\ \bar{y}_j = y + r \sin((j-1)\theta_0 + \theta) \\ \bar{z}_j = z \\ i = 1, 2, \dots, m; j = 1, 2, \dots, n \end{cases} \quad (9)$$

where R_{ij} is the observation of the i th acoustic station, ΔR_i is the constant range error equivalent to the time delay measurement error, the station position error and the sound velocity correlated error of the i th acoustic station.

Similarly, the fusion model of the polynomial parameters, the initial phase and the system errors can be obtained as follows:

$$\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_N \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1(a_1, b_1, \dots, \theta, \Delta R_1, \Delta R_2, \dots, \Delta R_m) \\ \mathbf{F}_2(a_1, b_1, \dots, \theta, \Delta R_1, \Delta R_2, \dots, \Delta R_m) \\ \vdots \\ \mathbf{F}_N(a_1, b_1, \dots, \theta, \Delta R_1, \Delta R_2, \dots, \Delta R_m) \end{pmatrix} = \begin{pmatrix} (g_1^{(1)}, g_2^{(1)}, \dots, g_m^{(1)})^T \\ (g_1^{(2)}, g_2^{(2)}, \dots, g_m^{(2)})^T \\ \vdots \\ (g_1^{(N)}, g_2^{(N)}, \dots, g_m^{(N)})^T \end{pmatrix}. \quad (10)$$

The above formula is the multi-station fusion system error model with the polynomial constraints, where the order of the indefinite polynomial is determined by the optimal model selection criterion.

For the linear models, assume that the true distance is $\mathbf{R} = [R_1, R_2, \dots, R_{mN}]^T$ and the measurement is $\mathbf{D} = [D_1, D_2, \dots, D_{mN}]^T$ in the LBL fusion positioning, the measurement data model is as follows:

$$\begin{cases} \mathbf{D} = \mathbf{R} + \mathbf{e} \\ \mathbf{e} \sim \mathbf{N}(\mathbf{0}, \delta^2 \mathbf{I}) \end{cases} \quad (11)$$

where \mathbf{e} is a random error vector with the mean $\mathbf{0}$ and the covariance matrix $\delta^2 \mathbf{I}$.

It is assumed that the estimated parameter of the model is $\hat{\boldsymbol{\beta}} = (a_1, b_1, \dots, \theta, \Delta R_1, \Delta R_2, \dots, \Delta R_m)^T$ and the design matrix is \mathbf{H} after the polynomial order is determined.

Let $\mathbf{H}_X = \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$. Then the above equation can be written as follows:

$$\mathbb{E} \|\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{R}\|^2 = \|(\mathbf{I} - \mathbf{H}_X)\mathbf{b}\|^2 + n_\beta \delta^2 \quad (12)$$

where $\hat{\boldsymbol{\beta}}$ is the estimation value of the parameter $\boldsymbol{\beta}$, and n_β is the number of the parameters. According to the residual sum of squares (RSS) formula, the RSS is calculated as follows:

$$\text{RSS} = \mathbb{E} \|\mathbf{D} - \mathbf{H}\hat{\boldsymbol{\beta}}\|^2 = \|(\mathbf{I} - \mathbf{H}_X)\mathbf{R}\|^2 + (mN - n_\beta)\delta^2. \quad (13)$$

Then, we can get

$$E\|\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{R}\|^2 = \text{RSS} + (2n_\beta - mN)\delta^2. \quad (14)$$

The above formula is the optimal test statistic of the polynomial system error model, denoted as

$$S = \text{RSS} + (2n_\beta - mN)\delta^2. \quad (15)$$

It can be seen from the above expression that the test statistic S does not depend on whether the model is linear or not. Therefore, the above statistic is still available for the nonlinear multi-beacon measurement system. It is required to minimize the RSS and the number of parameters for the target trajectory model with the optimal polynomial constraints. The model with the smallest S is selected as the multi-station fusion system errors model constrained by the optimal polynomial based on the optimal selection criterion.

Theoretically, unlike the space-segment ballistics, the complex underwater environment leads to a larger uncertainty of the motion model. And the trajectory approximately conforms to the polynomial model during the small range and the short time motion of the underwater target. Therefore, the polynomial order is preferred by using the preference criterion, which considers the RSS of the model and the number of parameters. The criterion optimizes the polynomial order in terms of both the positioning accuracy and the model complexity. The match between the motion model and the actual motion state is improved. Under the condition that the accuracy of the positioning device is determined, the selection of the positioning model parameters has a great influence on the positioning accuracy.

Remark 1 If δ^2 is unknown and the statistic $Q = mN - n_\beta$ is large, the RSS of the range can be used to give an estimation of δ^2 since the first part of (13) is relatively small. It can be ignored in the estimation. And the result is equivalent to the optimal result with δ^2 . It is denoted that the RSS corresponding to the system errors models with q polynomial constraints of different orders are $\text{RSS}_1, \text{RSS}_2, \dots, \text{RSS}_q$, and the corresponding statistics Q are Q_1, Q_2, \dots, Q_q . Then the estimation δ_*^2 of δ^2 is

$$\delta_*^2 = \min \left\{ \frac{\text{RSS}_1}{Q_1}, \frac{\text{RSS}_2}{Q_2}, \dots, \frac{\text{RSS}_q}{Q_q} \right\}. \quad (16)$$

After obtaining the multi-station fusion system error model constrained by the optimal polynomial, the traditional identification method is nonlinear parameter estimation.

When the observation error is independent and obeys the Gaussian distribution with the mean 0 and the variance σ^2 , the stable optimal estimation can be obtained according to the following Gauss-Newton iterative formula:

$$\hat{\boldsymbol{\beta}}_{k+1} = \hat{\boldsymbol{\beta}}_k + (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T (\mathbf{Y} - \mathbf{F}(\boldsymbol{\beta}_k, \mathbf{X}_i)) \quad (17)$$

where $\hat{\boldsymbol{\beta}}_k$ is the parameter estimation in the k th iteration and \mathbf{J} is the Jacobian matrix.

However, (12) shows that the estimation error of the traditional estimation method is

$$E\|\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{H}\boldsymbol{\beta}\|^2 = n_\beta \delta^2. \quad (18)$$

It shows that when $\mathbf{H}\hat{\boldsymbol{\beta}}$ is used as the estimation of $\mathbf{H}\boldsymbol{\beta}$, the error is proportional to n_β . The more the parameters, the worse the estimation.

After the polynomial is optimized by the model optimal selection criterion and the system errors are modeled uniformly by the reduced parameters, the inaccurate system error parameter number will lead to the low positioning accuracy based on the traditional identification method while the identification of redundant parameters is inaccurate. Therefore, it is necessary to select the appropriate number of the system errors. It is an effective method to classify and identify the system errors with multiple stations.

3. Iterative identification for system error based on improved spectral clustering

In the multi-station fusion system errors model (10) constrained by the optimal polynomial, whether the system errors $\Delta R_1, \Delta R_2, \dots, \Delta R_m$ can be accurately identified has a great influence on the positioning accuracy. Assume all the stations be the same error parameters, that will lead to the inaccurate calculation and the low positioning accuracy. However, too many error parameters will also lead to overfitting in the model, therefore, it is an effective way to improve the positioning accuracy to classify and identify the multi-station system errors, and to select the appropriate number of the system errors.

The spectral method transforms the data clustering problem into the graph cut problem with the similarity matrix. The predefined truncation function is optimized by calculating the eigenvalues and eigenvectors of Laplacian matrix, and the relaxation solution of the minimum cut problem is obtained. Finally, the cluster structure of the data is obtained by mapping to the original problem [24]. However, the similarity matrix established based on the distance generally reflects the distance similarity between data, and still cannot reflect the similarity of the system errors between the stations. Therefore, the similarity matrix of the spectral clustering is improved by the similarity of the system errors between the stations. The feedback and adjustment are carried out according to the results of the measurement element accuracy, the target rotation angle and the system errors identification. An iterative identification algorithm for the system error based on the improved spectral clustering is formed.

3.1 Spectral method

The purpose of the spectral clustering is to cluster the given data into different sets according to the similarity matrix, so that the data of the same cluster are as similar as possible, and the data between different clusters are as different as possible. For a given dataset and the similarity graph, it is illustrated in the graph theory to find a cut that divides a graph into several subgraphs with small connection weights between the subgraphs and the large connection weights.

Assume that the similarity matrix between data is $\mathbf{W} = (w_{ij})_{i,j=1,2,\dots,n}$ for the given dataset $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$, where w_{ij} denotes the similarity between \mathbf{x}_i and \mathbf{x}_j . A calculation method is Gaussian similarity function as follows:

$$w_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (19)$$

where the hyperparameter σ is used to control the amplitude of the adjacent relationship. A non-oriented graph $G(V, E)$ is generated by treating each data as a vertex according to the similarity matrix, where V and E represent the vertex set and the edge set, respectively. Then, the weighted adjacency matrix of the graph is the similarity matrix \mathbf{W} . For the vertex set V , the degree of each vertex is defined as follows:

$$d_i = \sum_{j=1}^n w_{ij}. \quad (20)$$

The degree matrix is defined as $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$. Then, the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$ and $\tilde{\mathbf{L}} = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2}$ are obtained. Obviously, its solution of the relaxed linear programming \mathbf{T}^* is composed of the eigenvectors t_1, t_2, \dots, t_k corresponding to the first k minimal eigenvalues of $\tilde{\mathbf{L}}$, namely $\mathbf{T}^* = (t_1, t_2, \dots, t_k)$.

To map the solution of the relaxed linear programming to the original problem, let $\mathbf{H}^* = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)^\top$, where $\mathbf{u}_i \in \mathbf{R}^k$ is the line i of \mathbf{H}^* , i.e., it is a dimension reduction. Cluster $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ by K-means to get the label set (C_1, C_2, \dots, C_k) . Then the final clustering result A_1, A_2, \dots, A_k of the original problem satisfies the following relationship:

$$A_i = \{x_j | u_j \in C_i\}. \quad (21)$$

3.2 System error identification based on improved spectral clustering

The system errors of different stations are not only related to the relative distances between the stations and the target, but also related to the measurement accuracy. Therefore, the similarity matrix based on the station position cannot fully reflect the similarity of the system errors

between different stations. Therefore, the similarity matrix of the spectral method is improved.

The target positioning accuracy is related to the geometry of the positioning system and the inverse measurement accuracy of the parameter estimation. An effective way to improve the target positioning accuracy is to improve the inverse measurement accuracy and the target rotation angle accuracy when the station positions are fixed. Therefore, the inverse measurement accuracy, the rotation angle and the system errors are considered to improve the similarity matrix. Obviously, the difference between the system errors identification results of the two stations as the same type and the real system errors is smaller, the system error similarity of the two stations is higher. Therefore, in this case that all the stations are involved in the positioning, the system error similarity of any two stations is defined as

$$s_{ij} = \exp\left(-\frac{\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\|^2}{2\sigma^2}\right) \quad (22)$$

where $\hat{\boldsymbol{\alpha}} = (\hat{\sigma}, \hat{\theta}, \Delta\hat{R}_1, \Delta\hat{R}_2, \dots, \Delta\hat{R}_m)$, and $\hat{\sigma}$ is the accuracy of the inverse measurement element, $\hat{\theta}$ is the estimation of the target rotation angle, $\Delta\hat{R}_1, \Delta\hat{R}_2, \dots, \Delta\hat{R}_m$ are the identification results of the station system errors when station i and station j are the same type. $\boldsymbol{\alpha} = (0, \theta, \Delta R_1, \Delta R_2, \dots, \Delta R_m)$, θ and $\Delta R_1, \Delta R_2, \dots, \Delta R_m$ are the corresponding true value, respectively. According to (22), the similarity matrix is continuously iteratively updated to obtain the accurate classification results of the system errors, and then to improve the positioning accuracy of the underwater targets.

In the underwater LBL positioning system, it is assumed that m acoustic stations are placed on the seabed-based platform for ranging the cylinder, and n beacons are uniformly installed on the cylinder. The initial phase of the cylinder is θ . Then, the system errors iterative identification algorithm flow for the underwater LBL positioning based on the improved spectral clustering is shown in Fig. 2. The specific process is as follows:

Step 1 Optimize the polynomial order to construct the kinematic equation of the cylinder centroid based on the optimal selection criterion. The relative position relationship between the beacons and the cylinder is determined, as shown in (9) and (2), where

$$\theta_0 = \frac{2\pi}{n}. \quad (23)$$

Step 2 Construct the weighted graph according to the stations' positions. The initial similarity matrix is set to

$$w_{ij}^{(0)} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (24)$$

where \mathbf{x}_i is the position of the i th station.

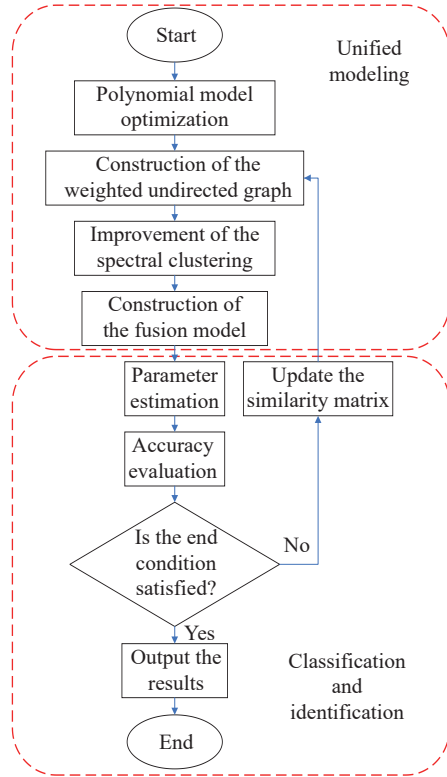


Fig. 2 Algorithm process

Step 3 Establish the degree matrix \mathbf{D} , Laplacian matrix \mathbf{L} and $\tilde{\mathbf{L}}$. Carry out the feature decomposition of $\tilde{\mathbf{L}}$. The classification number k is determined according to the distribution of eigenvalues. Calculate the feature vectors t_1, t_2, \dots, t_k corresponding to the first k minimal eigenvalues and $\mathbf{H}^* = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)^T$. The K-means method is used to cluster n data vectors $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$, and the label set (C_1, C_2, \dots, C_k) is obtained. Finally, the clustering results of the original problem are obtained:

$$A_i = \{\mathbf{X}_j | \mathbf{u}_j \in C_i\}. \quad (25)$$

Step 4 Give different system errors to each cluster according to the clustering results, so as to establish the corresponding measurement equation at a certain time:

$$\begin{pmatrix} R_{1j_1} \\ R_{2j_2} \\ \vdots \\ R_{mj_m} \end{pmatrix} = \begin{pmatrix} g_1(a_1, b_1, \dots, \theta, \Delta R_{A_{k_1}}) \\ g_2(a_1, b_1, \dots, \theta, \Delta R_{A_{k_2}}) \\ \vdots \\ g_m(a_1, b_1, \dots, \theta, \Delta R_{A_{k_m}}) \end{pmatrix} \quad (26)$$

where $\Delta R_{A_{k_i}}$ is the system error corresponding to the cluster A_{k_i} and $A_{k_i} \in \{A_1, A_2, \dots, A_k\}$.

Step 5 Construct the fusion model as shown in (10), and the parameter estimation is iteratively obtained according to (17):

$$\hat{\boldsymbol{\beta}} = (\hat{a}_1, \hat{b}_1, \dots, \hat{\theta}, \Delta \hat{R}_{A_{k_1}}, \dots, \Delta \hat{R}_{A_{k_m}}). \quad (27)$$

Step 6 Substitute the parameter estimation (27) into the measurement equation, the accuracy of the inverse measurement element at time t can be obtained:

$$\sigma_t = \sqrt{\frac{\|\mathbf{Y}_t - \hat{\mathbf{Y}}_t(\hat{\boldsymbol{\beta}})\|^2}{m-1}} \quad (28)$$

where m is the equation number at time t , \mathbf{Y}_t and $\hat{\mathbf{Y}}_t(\hat{\boldsymbol{\beta}})$ represent the measurement and the inverse measurement at time t , respectively.

Step 7 Substitute the parameter estimation (27) into (2), the estimation of the target trajectory and the GDOP at time t can be obtained:

$$\text{GDOP}_t = \sqrt{\text{trace}(\mathbf{J}_t^T \mathbf{J}_t)^{-1}} \quad (29)$$

where \mathbf{J}_t is the direction cosine matrix of the target at time t .

Step 8 The estimation accuracy of the target position can be obtained by integrating the inverse element accuracy and the GDOP for the whole target trajectory:

$$\eta = \overline{\text{GDOP}_t} \cdot \overline{\sigma_t} \quad (30)$$

where $\overline{\text{GDOP}_t}$ and $\overline{\sigma_t}$ represent the average value of N sampling points for the target trajectory.

Step 9 If the maximum iteration number is reached or the clustering results of the stations converge, the clustering results A_1, A_2, \dots, A_k and the accuracy η are output. Otherwise, the similarity matrix is updated according to the existing clustering results:

$$w_{ij} = \begin{cases} \exp\left(-\frac{\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\|^2}{2\sigma^2}\right), & A_{k_i} = A_{k_j} \\ \exp\left(-\frac{\|\hat{\boldsymbol{\alpha}}' - \boldsymbol{\alpha}\|^2}{2\sigma^2}\right), & A_{k_i} \neq A_{k_j} \end{cases} \quad (31)$$

where $A_{k_i} = A_{k_j}$ and $A_{k_i} \neq A_{k_j}$ represent the same cluster and different clusters of the two stations. $\hat{\boldsymbol{\alpha}} = (\overline{\sigma_e}, \hat{\theta}, \Delta \hat{R}_{A_{k_1}}, \Delta \hat{R}_{A_{k_2}}, \dots, \Delta \hat{R}_{A_{k_m}})$. $\overline{\sigma_e}$ is the average inverse measurement accuracy. $\hat{\boldsymbol{\alpha}}' = (\overline{\sigma_e}', \hat{\theta}', \Delta \hat{R}'_{A_{k_1}}, \Delta \hat{R}'_{A_{k_2}}, \dots, \Delta \hat{R}'_{A_{k_m}})$ represents the parameter estimation results of taking the station i and the station j as the same new class. Then, repeat Steps 3–9 based on the updated similarity matrix.

4. Test validation

In order to verify the effectiveness of the proposed identification method for the underwater LBL positioning based on the improved spectral clustering, it is applied to the LBL lake test of the underwater maneuvering target trajectory fusion solution, and the experimental analysis is carried out using the LBL station data and the target trajectory data in Songhua Lake. On the other hand, com-

pared with the conditions of the lake test, the working condition of the sea test is more complex. Therefore, the proposed method is also used for the experimental analysis in the shallow sea area. To further verify the performance of the method, the calculation results are compared with the solution method without considering the target rotation angle (Method 1), the solution method without system error identification (Method 2), the solution method with full system error identification (Method 3) and the system error identification method directly based on K-means clustering algorithm (Method 4). The full system error identification refers to taking each station as a cluster, that is, the number of the system errors is equal to the station number.

4.1 Test scenario

The equipment used in the LBL lake test include the acoustic stations, the beacons, and the target cylinder. The number of acoustic stations $m = 11$, 10 of which are placed within $1 \text{ km} \times 2 \text{ km}$ of the bottom and another mounted on a surface boat. A crane is used to control the cylinder on a large surface ship. The distance between the two ships is 50–100 m during the experiment. The beacons are mounted on the target cylinder. The stations' positions are calibrated by the survey ship before the test. The cylinder is put into the water by the crane after the beginning of the test. After the cylinder is placed in water, the acoustic beacons began to emit signals, and the acoustic stations began to receive signals and collect data. The cylinder stops when it reaches a certain depth. And the cylinder is pulled to the water surface at a uniform speed by the crane after the signal is emitted. During the period, both the large ship and the small ship drift, and there is displacement in the horizontal direction. After the cylinder reaches the water surface, stop emitting signals, the cylinder is out of the water, and the single group test ends. A global positioning system (GPS) antenna locates directly above the cylinder point and a depth gauge mounted on the cylinder provides horizontal and vertical position reference during the process from the time the cylinder enters the water to the time it exits the water.

The scenario of the LBL sea test is the similarity with the lake test, but the complex environment of the sea test results in only eight out of ten acoustic stations valid at the bottom.

4.2 Test result analysis

Firstly, the optimal selection criterion of the model is used to adaptively select the polynomial order. After the unified modeling of the multi-station system errors with the reduced parameters, the proposed iterative identification for the system error based on the improved spectral clustering is used for the parameter identification, the

cylinder trajectory estimation and the accuracy evaluation. In the two test scenarios, the RSS and the statistics for different polynomial models are shown in Table 1.

Table 1 Comparison of different polynomial models

Test scenario	Polynomial order	RSS	n_β	Q	S
Lake test ($m=11, N=9999$)	1	1423.70	18	109971	311.28
	2	1129.02	21	109968	16.66
	3	1112.54	24	109965	0.24
	4	1125.68	27	109962	13.45
Sea test ($m=9, N=8460$)	1	1456.43	16	76124	618.73
	2	1402.20	19	76121	564.57
	3	837.81	22	76118	0.24
	4	845.77	25	76115	8.26

The statistics S decreases with the increase of the order when the polynomial order is less than 3. Then, S begins to increase when the polynomial order is more than 3. Therefore, the cylinder trajectory model is optimized to the form of cubic polynomial. In the experiment, the reference trajectory of the cylinder and the overall layout of the acoustic stations after the data preprocessing of the measured data are shown in Fig. 3.

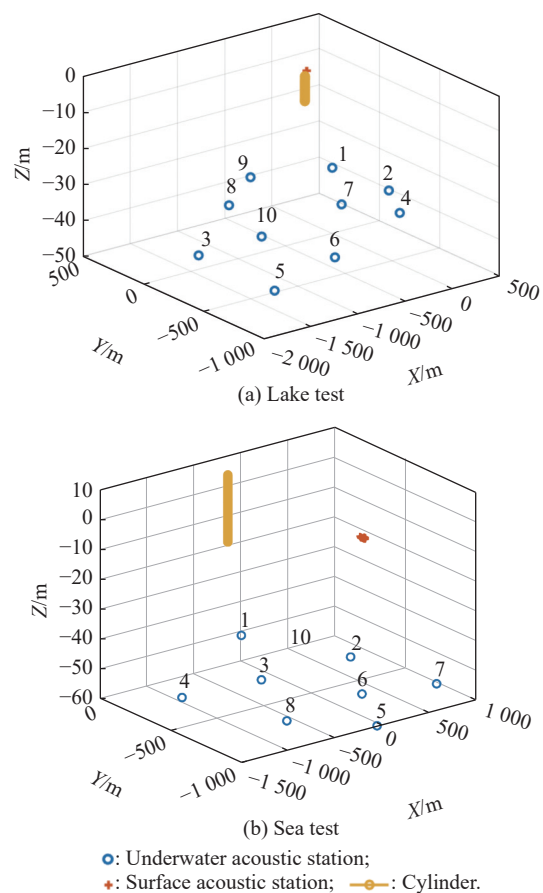


Fig. 3 Overall layout of stations and the cylinder

The iterative identification based on the improved spectral clustering is used to identify the system error of the surface acoustic station and the underwater acoustic stations. The initial similarity matrix is generated by the stations' positions, as shown in (19). In the lake test and the sea test, the underwater acoustic stations are divided into six categories and four categories after three and two iterations, respectively. When the classification number no longer changes, the system error and the rotation angle can be accurately identified. The variation of the eigenvalues' distribution of \tilde{L} during iterations is shown in Fig. 4 and Fig. 5.

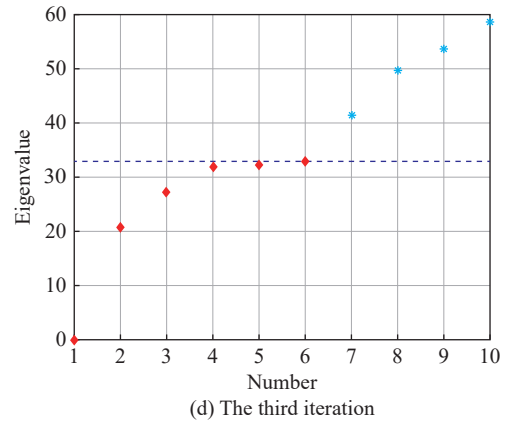
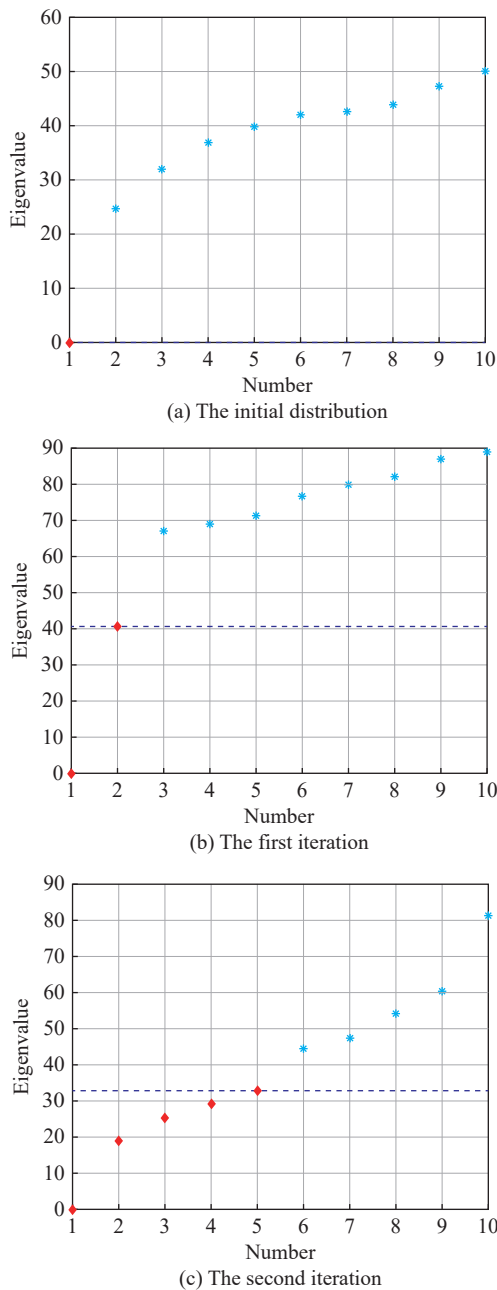


Fig. 4 Eigenvalues of \tilde{L} in the lake test

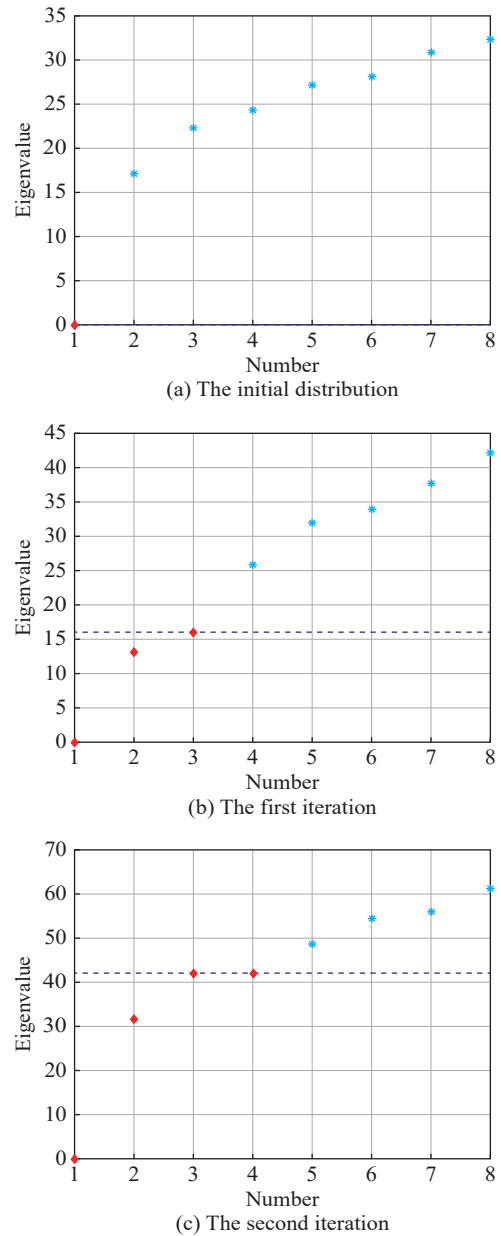


Fig. 5 Eigenvalues of \tilde{L} in the sea test

Select the cluster number of the stations and the identification number of the system errors adaptively changes according to the distribution of the eigenvalues during iterations. It can be seen from Fig. 4 that in the lake test scenario, the initial cluster number is selected as 1 according to the eigenvalues' distribution of the initial similarity matrix. The cluster number of the stations finally converges to 6 and does not change after three iterations. The clustering results are shown in Fig. 6. Fig. 5 shows that in the sea test scenario, the initial cluster number is selected to be 1 according to the eigenvalues' distribution of the initial similarity matrix. The cluster number of the stations converges to 4 and does not change after two iterations. The clustering results are shown in Fig. 7. It can be seen that the system error similarity between the stations is not only determined by the distances between the stations and the target.

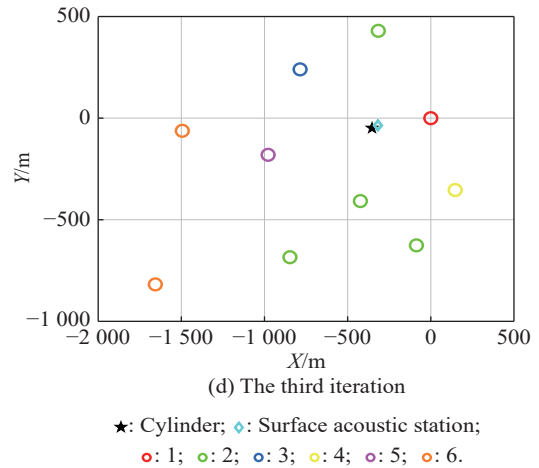
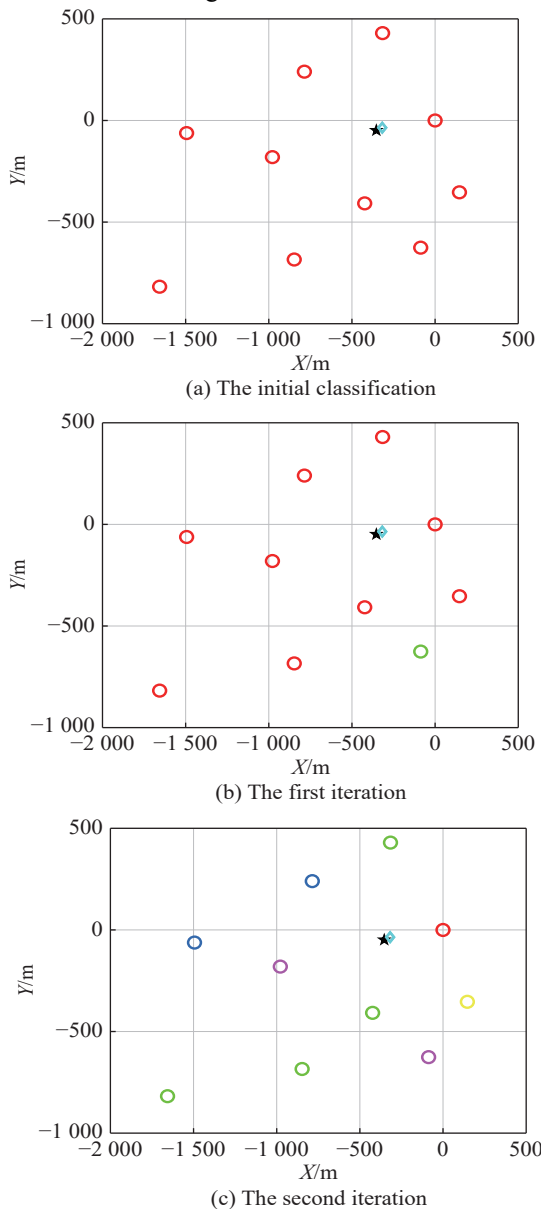


Fig. 6 Classification results of the lake test

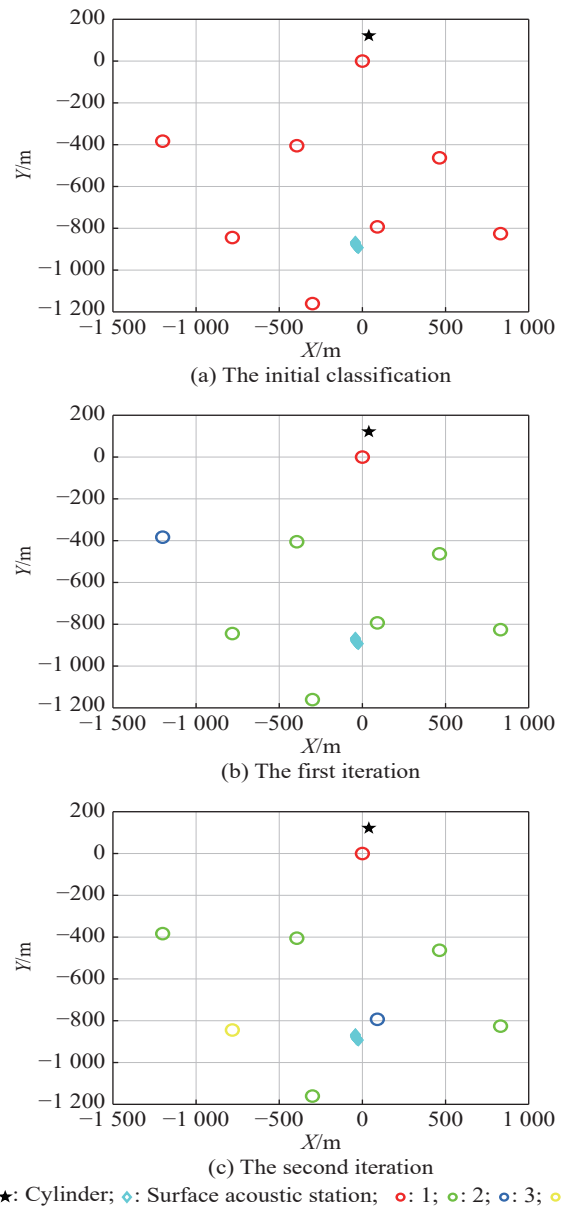


Fig. 7 Classification results of the sea test

The system error parameters are assigned according to the final classification results of the stations, and the target position parameters are solved by the model. Taking the cylinder position provided by the GPS antenna as a reference, the cylinder trajectory obtained by the system

error iterative identification with the improved spectral clustering (SEI-ISC), the target rotation angle and the system error proposed are compared with the trajectory results obtained by other methods. The comparison of the RSS is shown in Fig. 8 and Fig. 9.

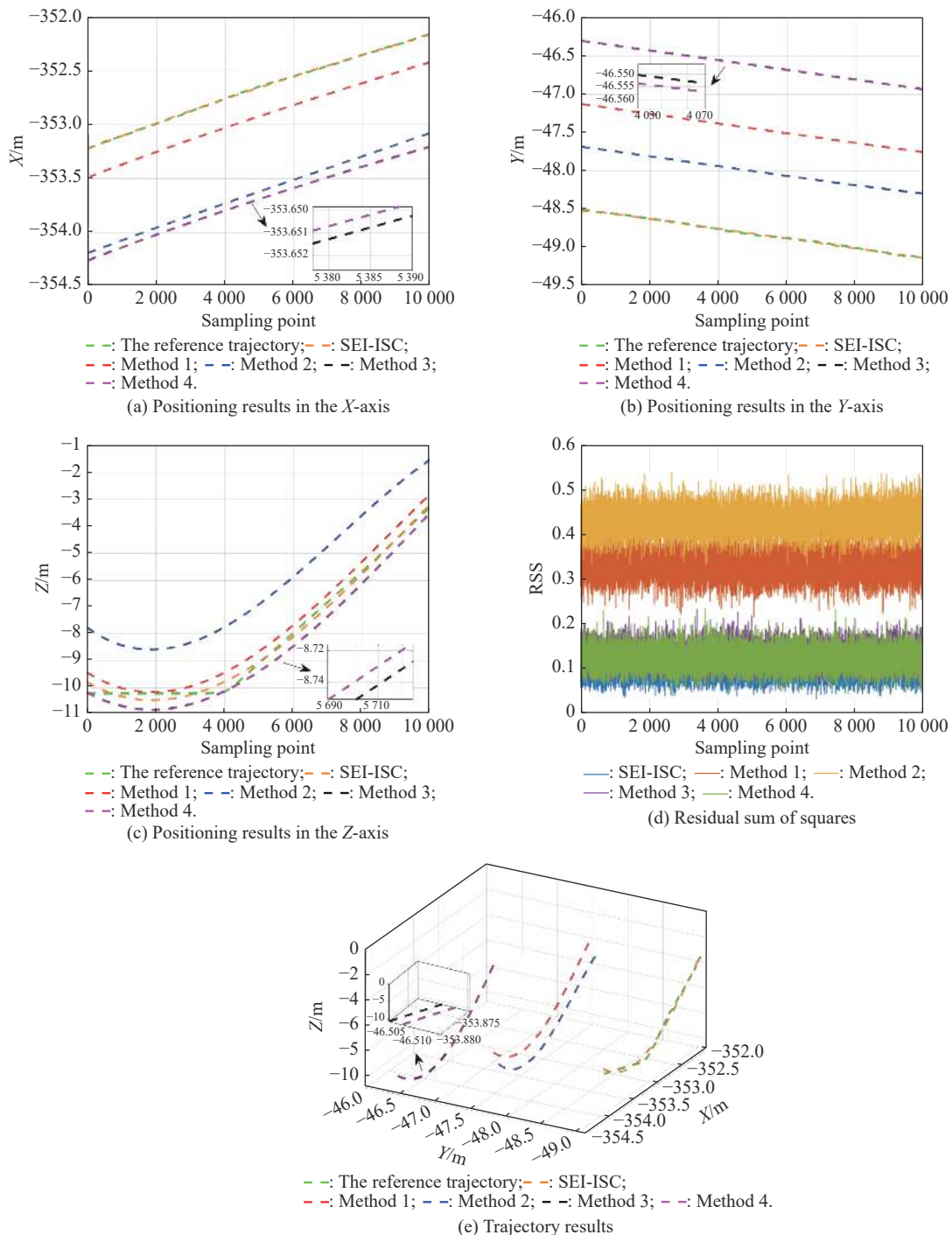


Fig. 8 Results of various methods in the lake test

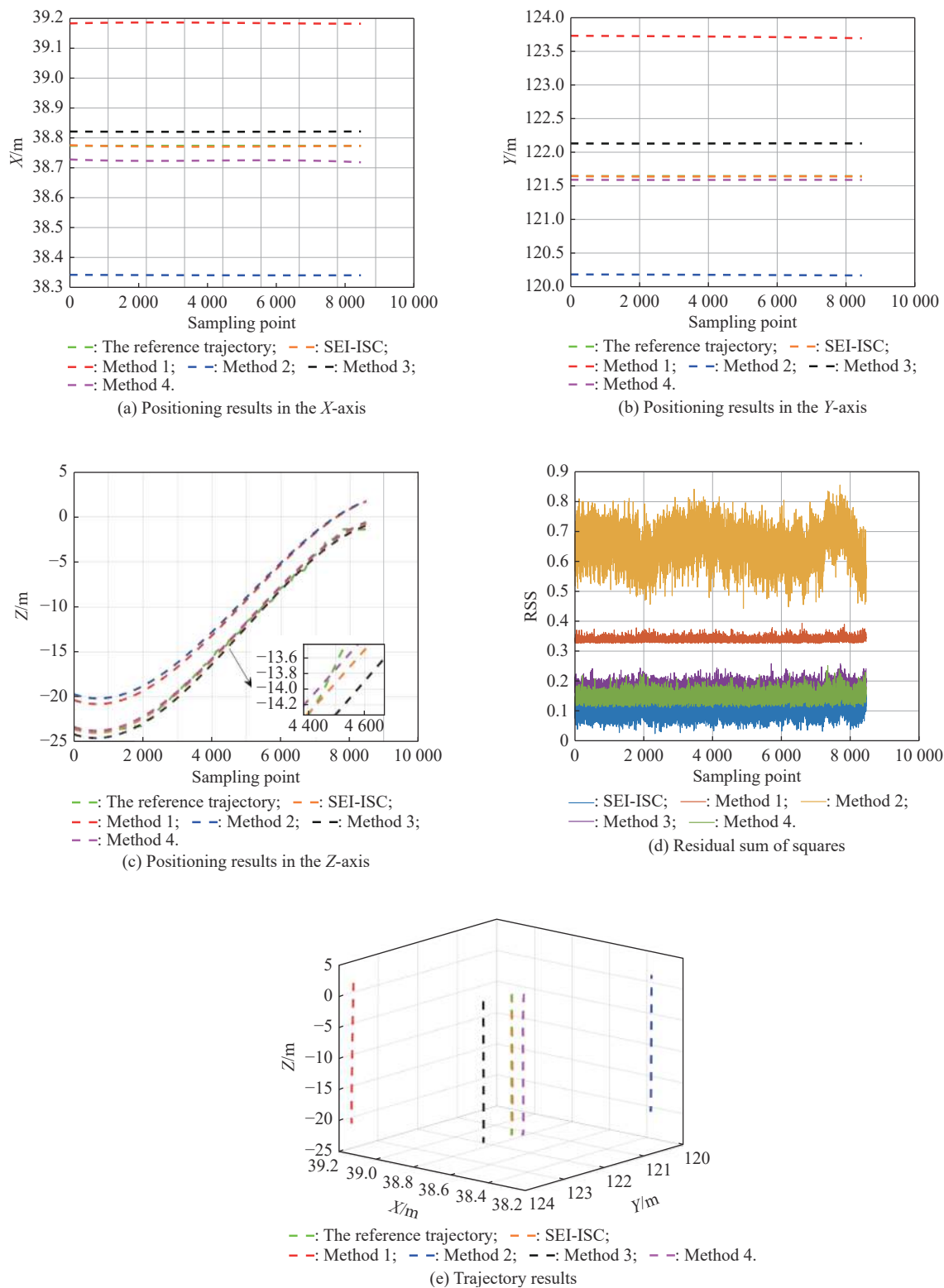


Fig. 9 Results of various methods in the sea test

The comparison between the measurement accuracy $\hat{\sigma}$, the cylinder rotation angle $\hat{\theta}$, the GDOP of the target, and

the accuracy of target position η calculated by each method are shown in Table 2.

Table 2 Results of various methods

Test scenario	Method	$\hat{\sigma}/\text{m}$	$\hat{\theta}/(^{\circ})$	GDOP	η/m
Lake test	SEI-ISC	0.0982	179.97	4.8129	0.4727
	Method 1	0.3293	–	4.9212	1.6207
	Method 2	0.4327	74.57	5.2131	2.2560
	Method 3	0.1304	55.80	4.7909	0.6248
	Method 4	0.1232	55.90	4.7904	0.5904
Sea test	SEI-ISC	0.1017	180.24	4.4101	0.4487
	Method 1	0.3397	–	4.2362	1.4392
	Method 2	0.6546	55.59	4.0927	2.6792
	Method 3	0.1816	158.27	4.4894	0.8154
	Method 4	0.1480	152.39	4.3844	0.6490

The true cylinder rotation angle is 180° . It can be concluded as follows:

(i) Regardless of the rotation angle of the target or the system errors of the stations will lead to the inaccurate calculation result and the low accuracy. And the calculation trajectory deviates from the reference trajectory.

(ii) For the solution method with full system error identification or the system error identification method directly based on the K-means clustering algorithm, the calculation accuracy is slightly improved. However, the target rotation angle and the system errors are still inaccurate, which results in that the calculation trajectory deviates from the reference trajectory.

(iii) When the iterative identification for the underwater positioning system error based on the improved spectral clustering is adopted, the calculation of the target rotation angle and the system errors of stations is more accurate. The measurement accuracy is about 0.1 m, and the positioning accuracy is the highest, better than 0.5 m. And the calculated trajectory is almost coincident with the reference trajectory.

(iv) The positioning accuracy of the sea test is lower than that of the lake test. This is because that the working condition of the sea test is more complex and the measurement uncertainty factors are more numerous.

5. Conclusions

Taking the target tracking of the underwater positioning system and the problem of the system error identification as the background, the iterative identification method for the underwater system error is given. Due to the ill-condition of the traditional EMBET model with too many parameters to be estimated, a multi-station fusion system error model based on the optimal polynomial constraint is proposed for the underwater positioning in the small area and the short time. The unified modeling of the reduced parameters for the trajectory parameters and the system

error coefficients is realized. In view of the problem of the inaccurate system error identification or the over-fitting in the process of the target positioning, the classification of the system errors of the stations only according to the distances between the stations and the target fails to consider the connection of the system error between the stations. Therefore, a multi-station non-oriented graph network is established, and an improved spectral clustering iterative identification method based on the measurement accuracy, the rotation angle and the system error is proposed. The system error of the stations is accurately identified, and the accurate estimation of the target position parameters is realized. Finally, it is applied to the LBL lake test and the sea test scenarios. And the system error of the stations is accurately identified, and the target positioning accuracy is improved.

References

- [1] ZHANG J C, SHI C H, SUN D J, et al. High-precision, limited-beacon-aided AUV localization algorithm. *Ocean Engineering*, 2018, 149(1): 106–112.
- [2] ZHANG T W, HAN G J, LIN C, et al. Integration of communication, positioning, navigation and timing for deep-sea vehicles. *IEEE Network*, 2020, 34(2): 121–127.
- [3] ZHANG T W, TANG J L, QIN S J, et al. Review of navigation and positioning of deep-sea manned submersibles. *The Journal of Navigation*, 2019, 72(4): 1021–1034.
- [4] WU Y H, TA X X, XIAO R C, et al. Survey of underwater robot positioning navigation. *Applied Ocean Research*, 2019, 90(4): 101845.
- [5] HUANG Y, WU L H, YU Q. Underwater square-root cubature attitude estimator by use of quaternion-vector switching and geomagnetic field tensor. *Journal of Systems Engineering and Electronics*, 2020, 31(4): 804–814.
- [6] CHEN H H. The estimation of angular misalignments for ultra short baseline navigation systems. Part II: experimental results. *The Journal of Navigation*, 2013, 66(5): 773–787.
- [7] HAN Y F, ZHENG C E, SUN D J. An optimized estimation method in long baseline acoustic positioning systems. *Acta Acustica*, 2017, 42(1): 14–20. (in Chinese)
- [8] DESSET S, DAMUS R, MORASH J, et al. Use of GIBs in AUVs for underwater archaeology. *Sea Technology*, 2003, 44(12): 22–27.
- [9] HAN Y R, WANG B, DENG Z H, et al. Point mass filter based matching algorithm in gravity aided underwater navigation. *Journal of Systems Engineering and Electronics*, 2018, 29(1): 152–159.
- [10] BIANCO M, GERSTOFT P. Dictionary learning of sound speed profiles. *The Journal of the Acoustical Society of America*, 2017, 141(3): 1749–1758.
- [11] CREAMER D B. On using polynomial chaos for modeling uncertainty in acoustic propagation. *The Journal of the Acoustical Society of America*, 2006, 119(4): 1979–1994.
- [12] TONG J W, XU X S, ZHANG T, et al. A misalignment angle error calibration method of underwater acoustic array in strapdown inertial navigation system/ultrashort baseline integrated navigation system based on single transponder mode. *Review of Scientific Instruments*, 2019, 90(8): 085001.
- [13] CHEN G X, LIU Y, LIU Y X, et al. Adjustment of transceiver lever arm offset and sound speed bias for GNSS-acoustic positioning. *Remote Sensing*, 2019, 11(13): 1606.
- [14] XIN M Z. Research on key technologies of GNSS-an under-

water positioning and navigation. Ji'nan: Shandong University of Science and Technology, 2020. (in Chinese)

- [15] YANG F L, LU X S, LI J B, et al. Precise positioning of underwater static objects without sound speed profile. *Marine Geodesy*, 2011, 34(2): 138–151.
- [16] XU P L, ANDO M, TADOKORO K. Precise, three-dimensional seafloor geodetic deformation measurements using difference techniques. *Earth, Planets and Space*, 2005, 57(9): 795–808.
- [17] GENG X Y, ZIELINSKI A. Precise multibeam acoustic bathymetry. *Marine Geodesy*, 1999, 22(3): 157–167.
- [18] XIN M Z, YANG F L, WANG F X, et al. A TOA/AOA underwater acoustic positioning system based on the equivalent sound speed. *The Journal of Navigation*, 2018, 71(6): 1431–1440.
- [19] GE T Z. Research on EMBET method for range task evaluation. Harbin: Harbin Institute of Technology, 2017. (in Chinese)
- [20] YANG J, WANG J. Tag clustering algorithm LMMSK: improved K-means algorithm based on latent semantic analysis. *Journal of Systems Engineering and Electronics*, 2017, 28(2): 374–384.
- [21] SHEN X L, SONG Z Y, FAN H Q, et al. Fast density peak-based clustering algorithm for multiple extended target tracking. *Journal of Systems Engineering and Electronics*, 2019, 30(3): 435–447.
- [22] QIAN K C, WAN Y, FAN Y C, et al. A method of measuring data fusion based on EMBET. Proc. of the IEEE 6th International Conference on Computer and Communication Systems, 2021: 33–37.
- [23] LI X H, ZHAO C X, YU J, et al. Underwater bearing-only and bearing-Doppler target tracking based on square root unscented Kalman filter. *Entropy*, 2019, 21(8): 740.
- [24] BIANCHI F M, GRATAROLA D, ALIPPI C. Spectral clustering with graph neural networks for graph pooling. Proc. of the International Conference on Machine Learning, 2020: 874–883.

Biographies



LU Yu was born in 1997. He received his B.S. degree in information and computational science from Shandong University, Ji'nan, China, in 2020. Currently, he is an M.S. candidate at the College of Sciences, National University of Defense Technology. His main research interests include data fusion and underwater acoustic positioning.

E-mail: luyu_gfkd@163.com



WANG Jiongqi was born in 1979. He received his B.S. degree in applied mathematics from Zhejiang University, Hangzhou, China, in 2002 and M.S. and Ph.D. degrees in applied mathematics from National University of Defense Technology, in 2004 and 2008, respectively. He is a professor in the College of Sciences, National University of Defense Technology, Changsha, China. His main

research interests include measurement data analysis, parameter estimation, system identification, and space target state filter and its applications.

E-mail: wjq_gfkd@163.com



HE Zhangming was born in 1985. He received his B.S. and M.S. degrees in applied mathematics from National University of Defense Technology, Changsha, China, in 2008 and 2010, respectively. From 2013 to 2014, he was a visiting scholar in Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Duisburg, Germany. He received his Ph.D. degree in system

science from National University of Defense Technology, in 2015. He is an associate professor in the College of Sciences, National University of Defense Technology, Changsha, China. His main research interests include fault diagnosis and prognosis, signal processing, and system identification.

E-mail: hzmnudt@163.com



ZHOU Haiyin was born in 1965. He received his B.S. degree in applied mathematics from Wuhan University, Wuhan, China, in 1986, M.S. degree from Hunan University, Changsha, China in 1989 and Ph.D. degree in systems engineering from National University of Defense Technology, Changsha, China, in 2003. Since 2009, he has been a professor in the College of Sciences,

National University of Defense Technology. His main research interests include data driven diagnosis approaches, power system dynamics and control, advanced signal processing, and data information fusion and its application.

E-mail: gfkd_zhy@sina.com



XING Yao was born in 1997. He received his B.S. degree in applied mathematics from National University of Defense Technology, Changsha, China, in 2020. Currently, he is an M.S. candidate at the College of Sciences, National University of Defense Technology. His main research interests are data fusion, systematic error identification, and target localization using underwater

sound.

E-mail: y_1326@163.com



ZHOU Xuanying was born in 1991. She received her B.S., M.S., and Ph.D. degrees in applied mathematics respectively in 2013, 2016, and 2019 from National University of Defense Technology, Changsha, China. She is a lecturer at the College of Sciences, National University of Defense Technology, Changsha, China. Her main research interests include system modelling, missiles, and

signal process and its applications.

E-mail: Julia_chow07@163.com