# Robust Beamforming for Data Correction via Interference-Noise Covariance Reconstruction and Adaptive Error Compensation

Lili Wang
School of Intelligent Manufacturing, Yancheng Polytechnic College, Yancheng 224000, China
E-mail: liliwang962464@163.com

**Keywords:** matrix reconstruction, robust beamforming algorithm, robustness, data correction, performance optimization

**Received:** April 10, 2025

*Accurate data detection is an important basis for achieving industrial process operation, performance control, and optimization. In response to the problem of poor accuracy in existing data correction methods, a data correction method based on a matrix reconstruction robust beamforming algorithm is proposed. However, this method still has significant correction errors for the data. Therefore, this study optimizes the matrix reconstruction robust beamforming algorithm to optimize the performance of data correction methods. Simulations were conducted on synthetic nonlinear dynamic data with a sampling frequency of 30 and an SNR of 10 dB. In the simulation results, when the incident angle was 30°, the signal power estimates of traditional beamforming algorithms and the proposed algorithm were 27.57 dB and 30.00 dB, respectively. This indicated that the proposed algorithm could effectively solve the problem of signal power underestimation. Under random error, the reaction concentration state value of the proposed algorithm at a time of 10 seconds was 0.152 J/kg·K, which differed from the true state value by 0.008 J/kg·K. Compared to the RCB baseline, this proposed algorithm reduced the average sum of squared errors and total sum of squared errors by 74.60% and 72.66%, respectively. The results indicate that the proposed algorithm has superior data correction performance. This study has contributed to improving the performance and robustness of beamforming algorithms in practical environments.*

*Povzetek: Prispevek predstavi izboljšano metodo robustnega oblikovanja smernosti snopa za korekcijo industrijskih podatkov prek rekonstrukcije kovariančne matrike motenj in prilagodljivega modeliranja napak.*

## 1 Introduction

As computer science and artificial intelligence rapidly develop, the process industry has entered the era of informatization. Accurate data detection is a reliable basis for the operation, real-time monitoring and analysis, performance control and optimization, and maintenance and management of the entire industrial production process [1-2]. The accuracy of detection data is the foundation of the operation of all production control systems. However, in the actual detection process, there may be random errors and significant errors, which make the data results inaccurate and unreliable, thereby affecting production operation and increasing the uncontrollability of system operation [3]. Therefore, it is crucial to explore an accurate and efficient data correction method. The beamforming algorithm has been extensively studied by scholars in recent years due to its ability to effectively reduce errors in data correction [4-5].

Yang et al. proposed a robust adaptive beamforming method based on Covariance Matrix Reconstruction (CMR) and subspace decomposition for steering vector estimation and correction. Therefore, the projection method of the signal subspace was corrected, improving the robustness of vector mismatch. This method has achieved good performance in various mismatch situations [6]. Luo et al. designed an effective orthogonal CMR method to remove unwanted signals, addressing the issue of performance degradation in traditional adaptive beamforming algorithms. This method removed useless signals from the received data by constructing a projection matrix, and its performance was superior to traditional methods [7]. Sun et al. developed a nonlinear beamforming method based on an improved efficiency genetic algorithm, aiming to enhance the Signal-to-Noise Ratio (SNR) and achieve effective nonlinear beamforming algorithms. The modified method could effectively enhance the testing of nonlinear beams and had better performance [8]. Yang and Dong developed a robust adaptive beamforming algorithm for steering vector estimation and subspace orthogonal interference power correction to address the performance degradation of Capon beamformers when expected signals appear. This method had good robustness against common mismatch errors [9]. Yu et al. proposed a low complexity receiver for millimeter wave beamforming, beam tracking, and data symbol detection to address the issue of poor accuracy in signal and data detection, aiming to effectively handle dynamic changes. This method could achieve the best spectral efficiency [10].

Due to the poor robustness of traditional beamforming algorithms and certain deficiencies in data correction and signal power estimation, many experts have optimized the performance of related algorithms. Wang et al. proposed

an improved fast iterative shrinkage threshold algorithm for fast compression beamforming to address the issue of high computational complexity in traditional beamforming methods. The method could reduce the amount of floating-point operations by three orders of magnitude [11]. Yang et al. proposed a robust adaptive beamforming method based on CMR and interference power estimation, aiming to reduce the impact of the signal of interest on traditional Capon beamformers and improve the robustness of the algorithm. This method had strong robustness against mismatch errors [12]. Du et al. proposed a robust adaptive beamforming algorithm based on the covariance matrix. This method decomposed the eigenvalues of the reconstructed matrix, derived the correspondence between eigenvalues and power, calculated weight vectors, and proved the effectiveness of the algorithm [13]. Mohammadzadeh et al. designed a robust adaptive beamforming technique for uniform linear arrays based on uncertain regions. This method introduced power estimation based on interference and noise components, and its performance approached the optimal value over a large SNR range [14]. Ge et al. developed a robust adaptive beamforming method based on sparse Bayesian and CMR. It maximized the marginal likelihood function of the data received through the array to accurately estimate the error of data correction. This method could approach the optimal output SNR and has better performance [15]. Table 1 shows the summary of related works.

In summary, numerous scholars have improved the relevant algorithms of beamforming and achieved certain research results. However, the currently advanced Robust Capon Beamforming (RCB) algorithm suffers from significant array position and amplitude/phase errors. In view of this, this study proposes an Interference-Plus-Noise Covariance Matrix Robust Adaptive Beamforming (IPNCM-RAB) algorithm. The algorithm automatically adjusts the column weights to correct the data. In response to the problem of large data correction errors in this method, this study optimizes the performance of the proposed data correction method. The IPNCM-RAB algorithm automatically adjusts the column weights to correct the data. To further enhance the accuracy and robustness of the IPNCM-RAB algorithm, the Array Amplitude Phase Error Correction (AA-PEC) method is integrated.

The research objective is to propose a data correction method based on subspace matrix reconstruction and Robust Beamforming Algorithm (RBA) to improve the accuracy and robustness of data correction. The second objective is to combine AA-PEC to optimize the performance of the proposed algorithm, further reducing correction errors and enhancing the applicability of the algorithm in practical engineering scenarios. The effectiveness of the proposed algorithm has been verified through simulation and comparison with existing methods under various conditions, including linear and nonlinear dynamic data environments. The innovation of this study lies in reconstructing the covariance matrix of interference and noise, which improves the robustness and effectiveness of beamforming algorithms. Then, the AA-PEC method is used to correct the data, minimizing the correction error and improving the application value of the algorithm in practical engineering.

Table 1: Related work summary table

| Method | Key Technique | Input Conditions | Dataset Type | Metrics | Limitations | References |
|---|---|---|---|---|---|---|
| CMR-SVEC | Subspace decomposition, SV estimation/correction | SV mismatch, 20dB SNR | Synthetic array | SINR, Robustness | Requires accurate subspace estimation | Yang J et al. [6] |
| URGLQ | Orthogonal CMR | Strong interference, 15dB SNR | Simulated array | IRR, MSE | Performance drops in dynamic environments | Luo T et al. [7] |
| GA-Nonlinear | Efficiency-improved genetic algorithm | Nonlinear signals, 10dB SNR | 3D seismic data | SNR enhancement | Computationally intensive | Sun Y et al. [8] |
| Subspace-Ortho | Steering vector estimation via subspace orthogonality | Angle errors, 12dB SNR | Experimental array | SINR, MSE | Limited to small-angle mismatches | Yang H et al. [9] |
| ML-MMSE | Millimeter-wave beamforming/tracking | Dynamic channels, 5dB SNR | mmWave experimental | Spectral efficiency | Complex implementation | Yu JH et al. [10] |
| FISTA-Compressive | Modified fast iterative shrinkage-thresholding | High-dimension data | Acoustic array | Computation speed | Requires sparsity assumption | Wang S et al. [11] |
| CMR-IPE | Covariance reconstruction + interference power estimation | 10-25dB SNR | Uniform linear array | SINR, Robustness | Sensitive to interference DOA errors | Yang H et al. [12] |
| Coprime-IPNCM | Interference-plus-noise matrix reconstruction | Coprime array geometry | Sparse array | Resolution | Restricted to specific array types | Du Y et al. [13] |
| PSR-UR | Power spectral estimation + uncertainty regions | Wide SNR range | Experimental | Near-optimal SINR | Requires region parameter tuning | Mohammadzadeh S et al. [14] |
| SBL-CMR | Sparse Bayesian learning + CMR | Low SNR | Simulated | Output SNR, MSE | High computational complexity | Ge S et al. [15] |

## 2 Methods and materials

This section first proposes a data correction method based on matrix reconstruction RBA, which automatically adjusts the array weights to correct the data. To improve the accuracy of data correction, this study proposes a data correction method based on AA-PEC to optimize its performance.

### 2.1 Data correction method based on subspace matrix reconstruction of RBA

The beamforming algorithm is widely used in the field of array signals due to its ability to accurately correct data errors. In recent years, matrix reconstruction algorithms have developed rapidly due to their strong robustness. However, this algorithm still has certain shortcomings in data correction and source power estimation [16]. Due to the high requirement for prior information of the array in the beamforming algorithm based on matrix reconstruction, it is necessary to provide robustness of the beamforming algorithm under adaptive conditions to achieve accurate data correction. However, existing RCB algorithms suffer from power underestimation and cannot accurately correct data [17]. Therefore, this study designs a data correction method based on the IPNCM-RAB algorithm. This algorithm improves the accuracy of data correction by automatically adjusting the column weights. Figure 1 shows the signal received by a uniform linear array in a subspace matrix.

Figure 1 shows a schematic diagram of a uniform linear array receiving a signal. The figure shows the interaction between the wavefront of a planar wave array and its various elements (Element M-2 to Element 1). Each array element is located on its own normal, which is perpendicular to the reference element of the array. The incident signal reaches the array at an angle θ, resulting in differences in the arrival time of the signal between different elements, where d is the spacing between the elements. This study applies $Q+1$ signals with a center frequency of $f_s$ to a uniform front line, where the uniform front line consists of $M$ elements. At time $t$, the signal received by the array element is as shown in equation (1).

$$x_m(t)=\sum_{q=0}^{Q}s_q(t-\tau_{mq})+n_m(t) \tag{1}$$

In equation (1), $x_m(t)$ is the signal received by the $m$-th element. $s_q(t)$ is the complex envelope vector and $n_m(t)$ is the noise in the element space. $\tau_{mq}$ is the delay between the signal received by the $m$-th element and the reference element. When the distance difference between signals reaching adjacent array elements is $d\sin\theta_q$, the matrix representation of the received signals in the array is as shown in equation (2).

$$\begin{cases}\tau_{mq}=\dfrac{(m-1)d\sin\theta_q}{c}\\ x(t)=\sum_{q=0}^{Q}s_q(t)a_q+n(t)=AS(t)+n(t)\end{cases} \tag{2}$$

In equation (2), $d\sin\theta_q$ is the difference in signal arrival distance. $c$ means the speed of light. $x(t)$ is the matrix of the array received signal model. $a_q$ means the guiding vector of the $q$-th incident source. $A$ is an array manifold matrix. $n(t)$ is a noise vector. $S(t)$ means the complex envelope vector of the incident signal. The beamformer weights and sums the signals received by the array based on the spatial information of different incident signals. The schematic diagram of its beamforming and uniform linear array robust beamformer is shown in Figure 2.

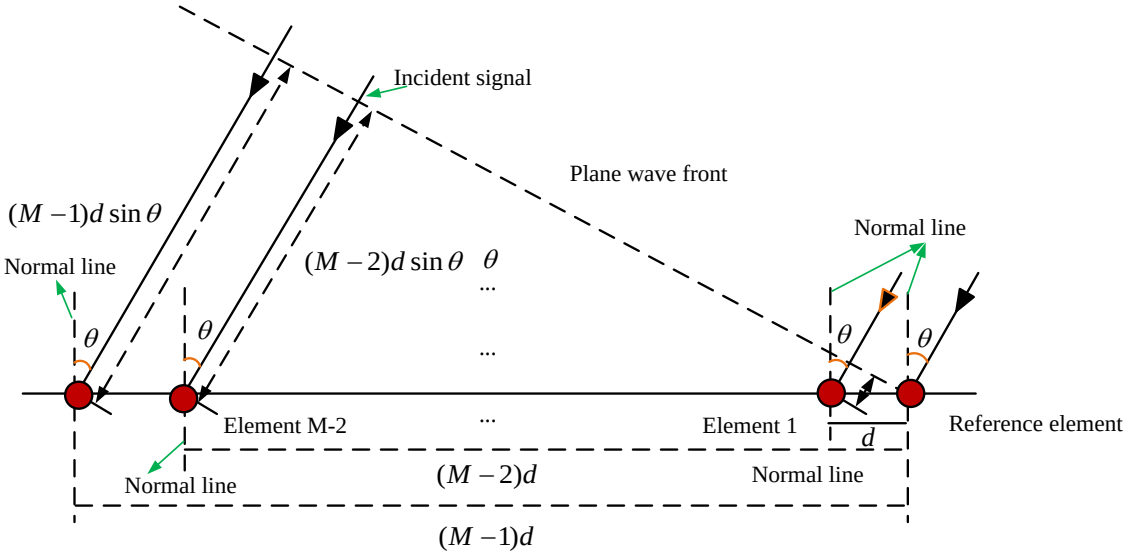
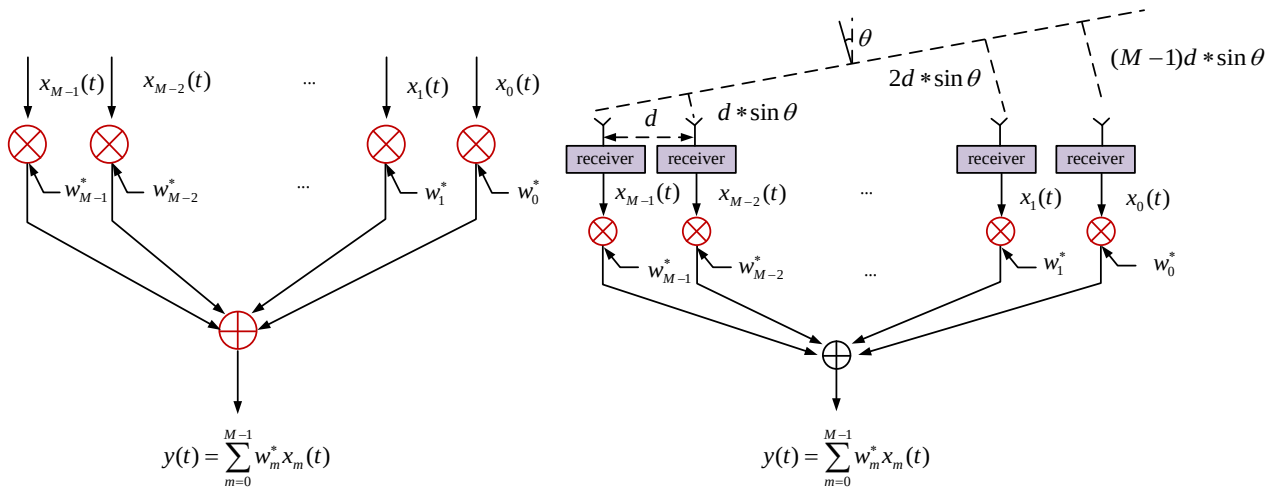



Figure 1: Uniform linear array receiving signal schematic.




(a) Beam alignment diagram

(b) Uniform linear array robust beam synthesizer diagram

Figure 2: Diagram of beamforming and uniform linear array robust beamformers.

In Figure 2 (a), the signals $x_{M-2}(T)$ to $x_0(T)$ received by each element in the array are weighted by their respective weights $w^*_{M-2}$ to $w^*_0$ and summed to form the output signal $y(t)$. In Figure 2 (b), each received signal is also weighted, but to adapt to the incident angle $\theta$, the time difference of signal arrival is also considered. The impact of different path lengths on the signal, such as $d\cdot\sin\theta$ and $2d\cdot\sin\theta$, is marked in the figure, and these differences are compensated for by weight adjustment to optimize the beamforming effect. The robust estimation of beamforming algorithms mainly relies on data regression theory. An unbiased estimation function using data regression theory is designed. This function can effectively complete data correction under the condition of insensitivity to data bias, thereby obtaining coordinated data that is closer to the measured values [18]. The impact function is a core indicator function in robust estimation, which can display the importance and impact of various estimation errors. Its calculation is shown in equation (3).

$$I(\xi_0)=\lim_{t\to 0}\frac{T\left[(1-t)f+t\delta(\xi-\xi_0)\right]-T\left[f\right]}{t} \tag{3}$$

In equation (3), $I(\xi_0)$ is the impact function of robust estimation. $\delta(\xi-\xi_0)$ is the particle distribution function centered around $\xi_0$. $T$ is a robust function. $\{\xi_1,\ldots\xi_n\}$ follows a distribution function. $\xi_0$ is the effect function of robust estimation. The impact function quantifies the degree to which an estimate is affected by outliers or deviations from the expected data distribution. This function is crucial for understanding how robust estimators minimize the impact of outliers. The expression of the influence function at the data point and the expression of the joint weight function are denoted in equation (4).

$$\begin{cases} I(\xi_i)=\dfrac{d\rho}{d\xi_i} \\ W(\xi_i)=\dfrac{I(\xi_i)}{\xi_i} \end{cases} \tag{4}$$

In equation (4), $I(\xi_i)$ is the influence function at point $i$ of the data band. $W(\xi_i)$ is a joint weight function. $\xi_i$ is the measurement error value at data point $i$. $\rho$ is the objective function. The joint weight function combines the deviation of the influence function from the mean, providing a measure of the relative importance of each data point in the estimation process. This helps to understand how each data point contributes to the robustness of the overall estimation. According to the theory of robust estimation, a robust estimation function $\rho(x)$ is designed, which is expressed as equation (5) [19].

$$I(x)=\frac{d\rho(x)}{dx}=\begin{cases}-e^{-x}x^2+2xe^{-x}+c(1-e^{-x})+cxe^{-x}, x\geq 0\\ e^{-x}x^2+2xe^{-x}+c(1-e^{-x})+cx(-e^{x}), x<0\end{cases} \tag{5}$$

In equation (5), $I(x)$ is the influence function of the robust estimation objective function. $x$ is the standard error. $c$ is an adjustable parameter. This function minimizes the sum of robust function values to ensure that the estimation results are less affected by outliers. This is a key step in developing robust estimators that can effectively handle noisy or contaminated data. To investigate the impact of interference power estimation on RBA, this study analyzes the Interference Rejection Ratio (IRR). The calculation of the $i$ -th IRR is given by equation (6).

$$\frac{\sigma_i^2}{\sigma_{io}^2}=\frac{E\{|s_i(k)|^2\}}{E\{|w^H a_i s_i(k)|^2\}}=\frac{E\{|s_i(k)|^2\}}{\mu^2|a_0 R_{i+n}^{-1}a_i|^2 E\{|s_i(k)|^2\}} \tag{6}$$

In equation (6), $\sigma_i^2$ is the output power of interference. $\sigma_{io}^2$ is the output power of the beamforming system. $s_i(k)$ is the complex envelope vector of the $i$ -th interference. $a_0$ is the expected signal. $\mu$ is a constant. $a_i$ is the $i$ -th interference signal. $R_{i+n}^{-1}$ is the corresponding covariance matrix. The higher the IRR value, the better the suppression effect on interference, which is crucial for maintaining the integrity of the expected signal in strong interference environments. Finally, based on the obtained noise power estimation, the

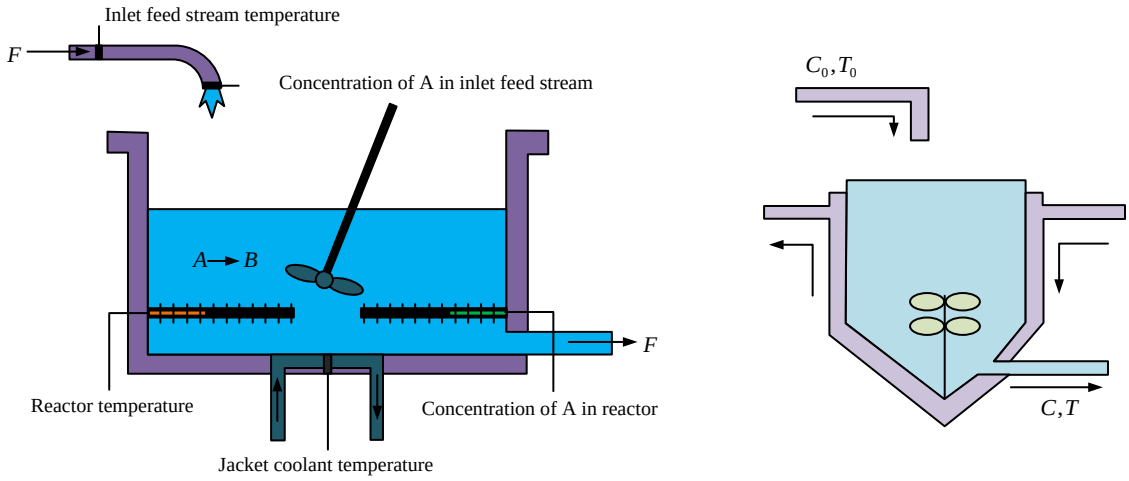
interference plus noise covariance matrix is reconstructed. The expression and weight vector of the beamforming algorithm are shown in equation (7).

$$\begin{cases} \hat{R}_{i+n} = \sum_{q=1}^{Q} \hat{\sigma}_q^2 \hat{a}(\theta_q)\hat{a}(\theta_q)^H + \hat{\sigma}_n^2 I \\ w_{IPNCM} = \dfrac{\hat{R}_{i+n}^{-1}\hat{a}(\theta_0)}{\hat{a}(\theta_0)^H \hat{R}_{i+n}^{-1}\hat{a}(\theta_0)} \end{cases} \tag{7}$$

In equation (7), $\hat{R}_{i+n}^{-1}$ is the inverse matrix of the corresponding covariance matrix. $w_{IPNCM}$ is the weight vector of the beamforming algorithm. $Q$ is the signal that reaches the array. $\sigma_q^2$ means the output power of the $q$-th interference. The weight vector of beamforming algorithm optimizes beamforming performance by minimizing interference and noise while maximizing the signal of interest. This study enhances the robustness of the beamforming algorithm by obtaining weight vectors, thereby improving the accuracy of the algorithm for data correction. To further improve the accuracy and robustness of the IPNCM-RAB algorithm, several variants of the IPNCM algorithm are developed and compared, including IPNCM-cor, IPNCM-sub, IPNCM-qcqp, and IPNCM-ortho. These variants use different techniques to reconstruct the interference plus noise covariance matrix. Among them, IPNCM-cor focuses on using correlation-based methods to correct the covariance matrix of interference and noise. By utilizing the correlation between estimated interference components and noise components, this method improves the accuracy of the reconstructed matrix. This method ensures that the covariance matrix better represents the actual interference and noise conditions, thereby improving the robustness of the beamforming algorithm. IPNCM-sub uses subtraction-based techniques to estimate and remove interference components from the noise covariance matrix. By accurately removing interfering components, the remaining noise covariance matrix becomes more accurate, resulting in better performance in data correction and signal estimation. IPNCM-qcqp utilizes the Quadratic Constrained Quadratic Programming (QCQP) method to optimize the reconstruction of the interference plus noise covariance matrix. The QCQP method ensures that the reconstructed matrix satisfies specific constraints, such as positive definiteness and orthogonality, which are crucial for the robustness and accuracy of beamforming algorithms. IPNCM-ortho introduces an orthogonalization process to ensure that the interference component is orthogonal to the signal component. By enforcing orthogonality, the algorithm can more effectively separate the expected signal from interference, thereby improving the robustness and accuracy of beamforming algorithms. For the correction of dynamic data, this study adopts the Dynamic System Model (DSM) as the constraint equation to achieve data correction. DSM is typically represented using time-varying state variable functions. DSM's data correction utilizes the time redundancy feature to correct real-time read data [20]. In the process of correcting nonlinear dynamic data, this study optimizes the constraint conditions of the dynamic model, as shown in equation (8).

$$\begin{cases} \min \Phi = \sum_{k=0}^{c} \left[\hat{x}(t_k) - \tilde{x}(t_k)\right] \sum^{-1} \left[\hat{x}(t_k) - \tilde{x}(t_k)\right] \\ s.t.\, f\left(\dfrac{d\hat{x}(t)}{dt}, \hat{x}(t)\right) = 0 \\ h(\hat{x}(t)) = 0 \\ g(\hat{x}(t)) < 0 \end{cases} \tag{8}$$

In equation (8), $\min \Phi$ is the estimated minimum deviation of the measured value. $c$ is the current time. $k$ is the sampling time. $\Sigma$ is the covariance matrix. $\hat{x}(t)$ is the estimation function. $\hat{x}(t_k)$ is the value of the estimated function at time $t_k$. $\tilde{x}(t_k)$ is the measurement value at time $t_k$. $f$ is a differential constraint equation. $h$ and $g$ are equality constraints, including upper and lower bounds on variables. The DSM ensures that the corrected data are consistent with the expected behavior of the system over time, which is crucial for accurate real-time data processing. To further validate the effectiveness of the proposed RBA in practical applications, the Continuous Stirred Tank Reactor (CSTR) model is employed as a dynamic system framework to simulate real-time data correction. The CSTR model provides a practical testing platform for evaluating the performance of algorithms in correcting dynamic data, ensuring that the robustness and accuracy proven in theoretical development can be effectively translated into practical scenarios. This study uses a CSTR-DSM to calibrate dynamic data. Figure 3 shows the framework of the model [21]. In Figure 3 (b), $C_0$ and $T_0$ are input variables, where $C_0$ is the system feed concentration and $T_0$ is the system feed temperature. $C$ and $T$ are state variables, representing the system output concentration and temperature, respectively. Based on chemical kinetics and thermodynamics, the mass differential equation and thermal differential equation for the DSM of CSTR are designed. The concentration and temperature are dimensionless relative concentrations and temperatures.



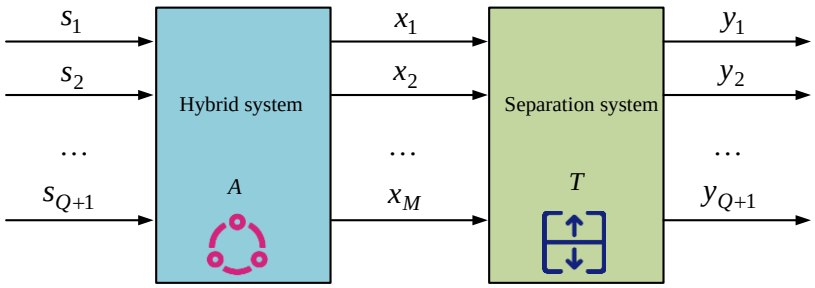

(a) Continuous stirred tank system model diagram (b) CSTR system model structure diagram

Figure 3: Schematic diagram of dynamic model of CSTR

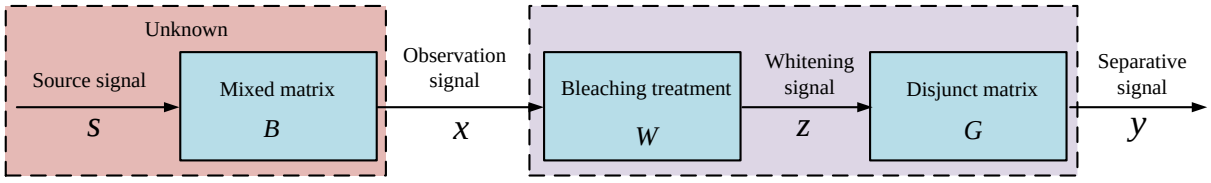
Figure 3 (a) shows the material flow and reaction process inside the reactor, where A and B represent reactants, F represents feed stream, and the reactor temperature and feed stream temperature are indicated separately. The figure also shows the concentration changes of A inside the reactor. Figure 3 (b) shows the external structure of the reactor, including the jacket coolant temperatures Ca, Ta, and the fluid temperatures C,T inside the reactor. These graphs are used to design the mass differential equation and thermal differential equation of CSTR-DSM for simulating and analyzing the behavior of the reactor under different operating conditions. The concentration and temperature are dimensionless relative values.

## 2.2 Performance optimization of data correction method based on AA-PEC

The data correction method based on matrix reconstruction RBA still has defects such as large correction errors and poor accuracy when correcting data. Therefore, it is necessary to provide robustness of beamforming algorithms under adaptive conditions to achieve accurate data correction. The traditional beamforming algorithm suffers from significant errors in the direction of incoming waves, array position, and amplitude and phase, leading to a decline in its performance. Therefore, it is necessary to optimize the performance of data correction methods [22]. The AA-PEC method is mainly used to correct amplitude and phase errors in data. After initial calibration by AA-PEC, DSM serves as the validator for evaluating the accuracy and reliability of the calibration data. DSM models the dynamic behavior of the system using time-varying state variable functions and compares the corrected data with the expected system performance for verification. This two-stage process ensures the accuracy and robustness of data correction. The AA-PEC method aims to optimize the performance of the data correction method for matrix reconstruction beamforming algorithm. This method introduces blind source separation theory into beamforming algorithms to estimate amplitude and phase errors, and finally reconstructs the covariance matrix of interference and noise to improve the accuracy of data correction. The AA-PEC method is used to correct amplitude and phase errors in data, thereby improving the overall accuracy and robustness of the IPNCM-RAB algorithm. Assuming that the mixed system through which source signal $s$ passes is $A$, the array system consists of $M$ elements, and the received replacement measurement data are $x = [x_1, x_2, \ldots, x_M]^T$ [23]. In this study, the separation system $T$ designed using Independent Component Analysis (ICA) is used to process the observed data, and the output data of the source signal are $y = [y_1, y_2, \ldots, y_{Q+1}]^T$ [24]. The number of independent sources is determined based on the complexity of the signal mixture and the specific application requirements. In this study, the number of independent sources is set to n, which is determined through preliminary analysis of the signal data. The convergence criteria for the ICA algorithm are based on the stability of the estimated independent components. The algorithm iterates until the change in the estimated components is below a predefined threshold, which is set to 10−5 in this study. This ensures that the algorithm reaches a stable solution without unnecessary computation. Figure 4 shows the models of the hybrid system and the separation system, as well as the structure of the blind source separation system.




(a) Hybrid system and separate system model diagram




(b) Diagram of blind source separation system

Figure 4: Hybrid and separation system model and blind source separation system diagram.

Figure 4 (a) shows the model diagrams of the hybrid system and the separation system. Multiple source signals $s_1$ to $s_{Q+1}$ enter the hybrid system through the mixing matrix A, generate observation signals $x_1$ to $x_M$, and then obtain separated signals $y_1$ to $y_{Q+1}$ through the separation system T. Figure 4 (b) provides a detailed description of the blind signal separation process, including the unknown signal mixing process and the blind signal separation process. The source signal S generates the observation signal $x$ through the mixing matrix B, undergoes whitening processing W and whitening signal Z, and then obtains the separated signal $y$ through the separation matrix G. In the process of signal mixing and separation, the received signal at time $k$ and the observed data output by the separation system are shown in equation (9).

$$\begin{cases} x_m(k) = \sum_{i=0}^{Q} a_{mi} s_i(k), m = 1,2,\ldots,M \\ Y(k) = TX(k) = TAS(k) \end{cases} \tag{9}$$

In equation (9), $y_m(k)$ is the $m$-th signal received at time $k$. $Y(k)$ is the output observation data. $a_{mi}$ is the mixing coefficient. $S$ is the input signal. $X$ is observational data. This study uses the ICA algorithm to preprocess the observation data $X$ to remove redundant information. This equation represents the mixing process of the source signal and the subsequent separation using the ICA algorithm. This is the fundamental step in separating various signals from a mixed signal environment. To further obtain the source signal, this study constructs a whitening matrix. The whitening process is a crucial step in the ICA algorithm. It involves converting observational data into a form with a unit covariance matrix. This method achieves this by calculating the eigenvalue decomposition of the covariance matrix of the observed data. This process ensures that the data are uncorrelated and have unit variance, thereby simplifying subsequent ICA processing. Its whitening process is shown in equation (10) [25].

$$\begin{cases} R = U\Lambda U^H = U\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}U^H \\ W = \Lambda^{\frac{1}{2}}U^H \end{cases} \tag{10}$$

In equation (10), $W$ is the whitening matrix. $R$ is the covariance matrix. $\Lambda = diag\{\lambda_1, \lambda_1, \ldots, \lambda_M\}$ is a diagonal matrix. $\lambda$ is the eigenvalue of $R$. $U$ is a feature matrix composed of the eigenvectors of $R$. The whitening matrix process ensures that the data are uncorrelated and have unit variance, thereby simplifying subsequent ICA processing. To achieve the design of the separation matrix, this study uses the Joint Approximation Diagonization of Eigenmatrices (JADE) algorithm to diagonalize the constructed cumulant matrix, as shown in equation (11) [26].

$$\begin{cases} q_{ij} = \sum_{l=0}^{Q}\sum_{k=0}^{Q} cum(z_i, z_j^*, z_k, z_l^*) e_{kl}, i, j = 1,2,\ldots,Q+1 \\ T = A^{-1} = U^H W \end{cases} \tag{11}$$

In equation (11), $q_{ij}$ is the weighted sum of the cumulative matrix $cum(z_i, z_j^*, z_k, z_l^*)$ with weight $e_{kl}$. $T$ is the separation matrix. JADE is a powerful tool in blind source separation, especially when the number of sources and mixtures is equal. JADE is chosen because it has robustness in handling complex mixed signals and can effectively separate signals even if the number of signal sources is equal to the number of mixed signals. This algorithm is particularly advantageous in scenarios where precise knowledge of the array manifold is limited, as it can operate without requiring detailed calibration data. The JADE algorithm's ability to handle multiple signals and its robustness against noise make it an ideal choice for improving the accuracy of amplitude and phase error corrections in the proposed method. The assumptions of

sparsity and independence directly impact the reliability of real-time corrections. Sparse and independent signals are more likely to be accurately separated and corrected in real-time. This is because the ICA algorithm can more effectively identify and isolate the relevant components of the signal, leading to more accurate and reliable corrections. On the contrary, if the signal is not sparse or independent, the separation process may introduce errors that propagate and affect the reliability of the correction. The IPNCM algorithm is sensitive to array prior information. Therefore, under the condition of array correction error, this study adopts the JADE algorithm to process the received signal, to improve the robustness of the beamforming algorithm [27], as shown in equation (12).

$$\begin{cases} Y(k) = GBS(k) = GTAS(k) \\ B = TA \end{cases} \tag{12}$$

In equation (12), $B$ is a mixed matrix. $T$ is the amplitude phase error matrix. $G$ is the separation matrix. This matrix represents the combined amplitude and phase error of each array element. This matrix is crucial for correcting amplitude and phase errors in the received signal, which can improve the overall accuracy of the system. The final received signal separation is shown in equation (13).

$$\hat{X}(k) = \hat{B}\hat{S}(k) \tag{13}$$

In equation (13), $\hat{X}(k)$ is the received separated signal. This equation represents the final result of separating the source signal from the mixed data. This step is crucial for obtaining various source signals that can be further processed or analyzed. The amplitude and phase error matrix is given by equation (14).

$$\hat{T}_{m,m} = \text{diag}(g_1 e^{-j\phi_1}, g_2 e^{-j\phi_2}, \ldots, g_m e^{-j\phi_m}) \tag{14}$$

In equation (14), $\hat{T}_{m,m}$ is the amplitude phase error matrix. $g_m$ is the estimated amplitude error value of the $m$-th element. $\phi_m$ is the phase error of the $m$-th element. The amplitude phase error matrix represents the combined amplitude and phase errors of each array element. This matrix is crucial for correcting amplitude and phase errors in the received signal, which can improve the overall accuracy of the system. The flowchart of data correction based on matrix reconstruction RBA is shown in Figure 5.

In Figure 5, the process starts from the input of the original array and goes through signal acquisition and data preprocessing (including covariance matrix decomposition and JADE diagonalization). The core processing includes IPNCM reconstruction, interference power estimation, CMR, and AA-PEC correction. Finally, beamforming and weight vector calculation are performed to output the corrected data.

The time complexity of the proposed method primarily depends on the eigenvalue decomposition and the matrix inversion operations involved in the beamforming process. The eigenvalue decomposition of an N×N matrix typically has a time complexity of $O(N^3)$, where N is the number of elements in the array. Therefore, the overall time complexity of the proposed method can be estimated as $O(N^3)$. Assuming that the number of sources M is much less than N and can be considered constant. The space complexity is determined by the amount of memory required to store the input data and intermediate results. The main memory consumers are the received signal matrix X, the covariance matrix $R_x$, the eigenvector matrix U, and the steering vectors ai. The received signal matrix X requires $O(NM)$ space, where M is the number of snapshots. The covariance matrix $R_x$ and the eigenvector matrix U each require $O(N^2)$ space. The steering vectors $a_i$ require $O(NM)$ space. Therefore, the overall space complexity of the proposed method is $O(N^2 + NM)$, which simplifies to $O(N^2)$ when M is considered constant.

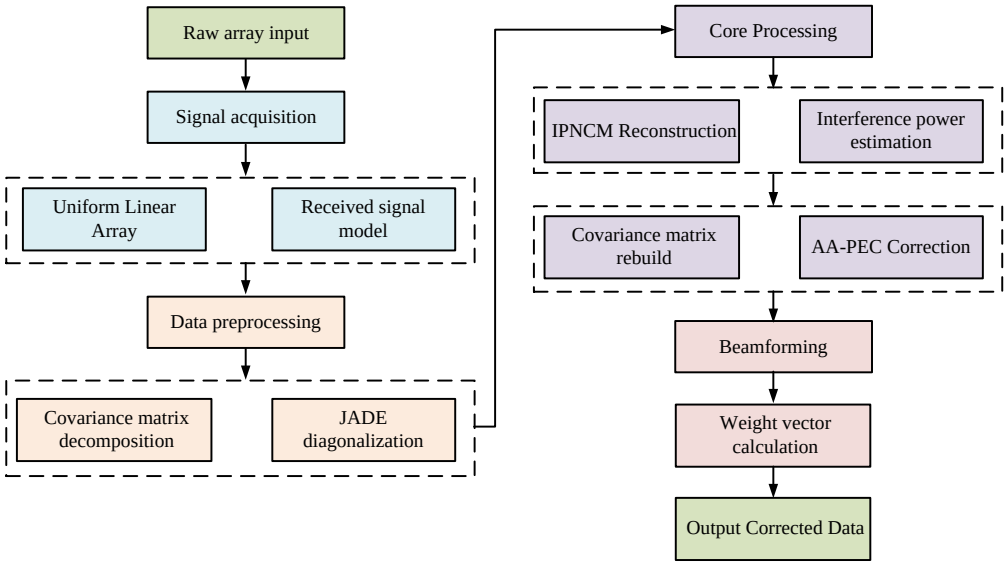



Figure 5: Flow chart of data correction and performance optimization based on matrix reconstruction RBA.

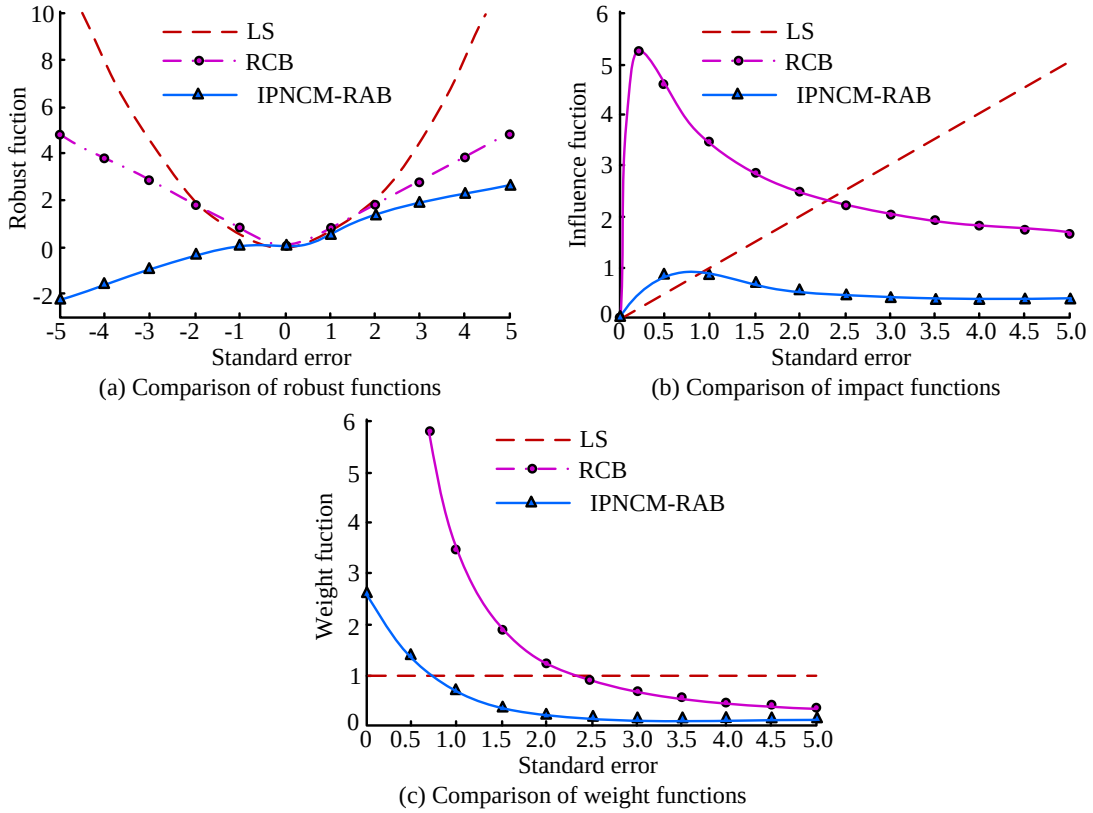
# 3 Results

This section first analyzes the data correction results of matrix reconstruction RBA. After optimizing the performance of the data correction algorithm, this study analyzes the accuracy of the algorithm's data correction under random and significant errors through simulation experiments.

## 3.1 Analysis of data correction results based on matrix reconstruction RBA

This study conducts simulation experiments using MATLAB software. The Signal-to-Interference plus Noise Ratio (SINR) is 30dB, the sampling frequency is 30, and the SNR is 10dB. In the experiment, the sampling angle intervals for IPNCM-cor and IPNCM-sub are 1° and 0.5°. The sampling angle interval between IPNCM-qcqp and IPNCM-ortho is 0.1°. The synthetic data used in these simulations are generated based on a Gaussian distribution assumption for the noise and signal components. This is a common practice in signal processing simulations to model real-world noise characteristics. The SNR is set to 10 dB, which is a typical value used in many beamforming studies to represent moderate noise conditions. This SNR level is chosen to ensure that the signal is detectable but still significantly affected by noise, making it a challenging yet realistic scenario for evaluating the performance of the proposed algorithm. The source signals are modeled as sinusoidal waves with varying frequencies to simulate different signal sources. This choice is justified as sinusoidal signals are fundamental in many practical applications and can effectively test the robustness of the beamforming algorithm.

The experiment compares the robustness functions, influence functions, and weight functions of Least Squares (LS), RCB, and the proposed algorithm (IPNCM-RAB) to explore the robustness, as shown in Figure 6. In Figure 6 (a), when the standard error is 4, the robust functions of LS, RCB, and the proposed algorithm are 8.25, 3.90, and 2.05. In Figure 6 (b), when the standard error is 0.5, the influence functions of LS, RCB, and the proposed algorithm are 5.21, 0.45, and 0.98. The proposed algorithm is closer to the influence function of LS. In Figure 6 (c), when the standard error is 1, the weight functions of LS, RCB, and IPNCM-RAB algorithms are 1.00, 3.47, and 0.85. The proposed algorithm is closer to the weight function of LS. This indicates that IPNCM-RAB has good performance in suppressing outlier errors and is robust.

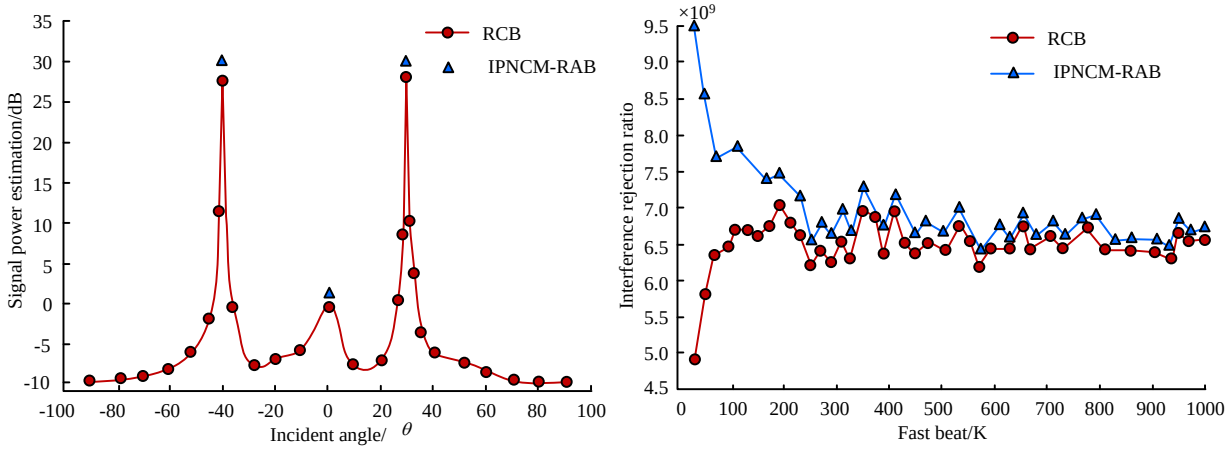



(a) Comparison of robust functions

(b) Comparison of impact functions

(c) Comparison of weight functions

Figure 6: Comparison of robust functions, influence functions, and weight functions.



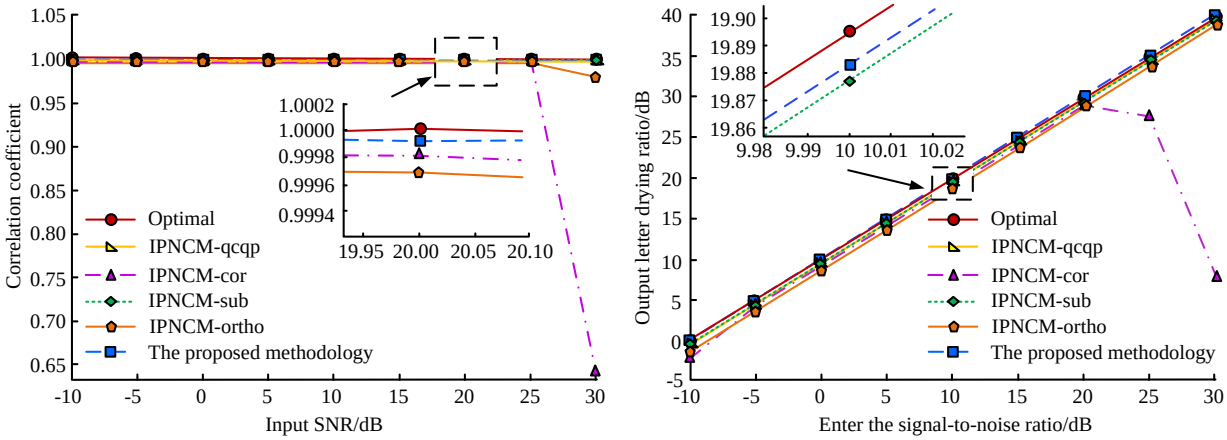

(a) Signal power underestimation improves results

(b) Comparison of interference rejection ratio

Figure 7: Comparison of signal power underestimation improvement results and IRR.

Figure 7 compares the signal power estimation and IRR of IPNCM-RAB and RCB to explore the improvement of subspace matrix-based RBA on power underestimation. In Figure 7 (a), when the incident angle of the signal is -40°, the estimated signal power of RCB and the proposed algorithm are 27.10 dB and 30.05 dB. When the incident angle is 0° and 30°, the signal power estimates of RCB and the proposed algorithm are 0.00 dB and 27.57 dB, 1.02 dB and 30.00dB. In Figure 7 (b), when the number of snapshots is 30, the IRRs of RCB and IPNCM-RAB are 4.9×109 and 9.5×109. When the number of snapshots reaches 780, the IRR of RCB is 6.6×109, and the proposed algorithm is 6.9×109. This indicates that the proposed RBA can effectively solve the problem of signal power underestimation and improve IRR in small block shooting situations.

To evaluate the accuracy of the reconstructed matrix, this study analyzes the correlation coefficient between theoretical IPNCM and reconstructed IPNCM through the variation of input SNR. In addition, the relationship between the input SNR and output SINR of the proposed algorithm is explored when there is angle error mismatch, as shown in Figure 8. In Figure 8 (a), when the input SNR is 20dB, the correlation coefficients of IPNCM-cor and IPNCM-orthor are 0.9998 and 0.9997. The correlation coefficient of the research model is 0.9999, which is closer to the optimal algorithm, indicating that the reconstructed IPNCM by this algorithm is more accurate. In Figure 8 (b), when the input SNR is 10 dB, the output SINR of IPNCM-sub is 19.878 dB, and the proposed algorithm is 19.882. The output SINR of the latter is closer to the optimal algorithm, indicating that the performance of the improved algorithm is closer to the optimal algorithm.




(a) The correlation coefficients between IPNCM and theoretical IPNCM were reconstructed

(b) Relation between output SINR and input SNR when Angle error is mismatched

Figure 8: The correlation coefficient and angle error matching degree of different algorithms under SNR and SINR.

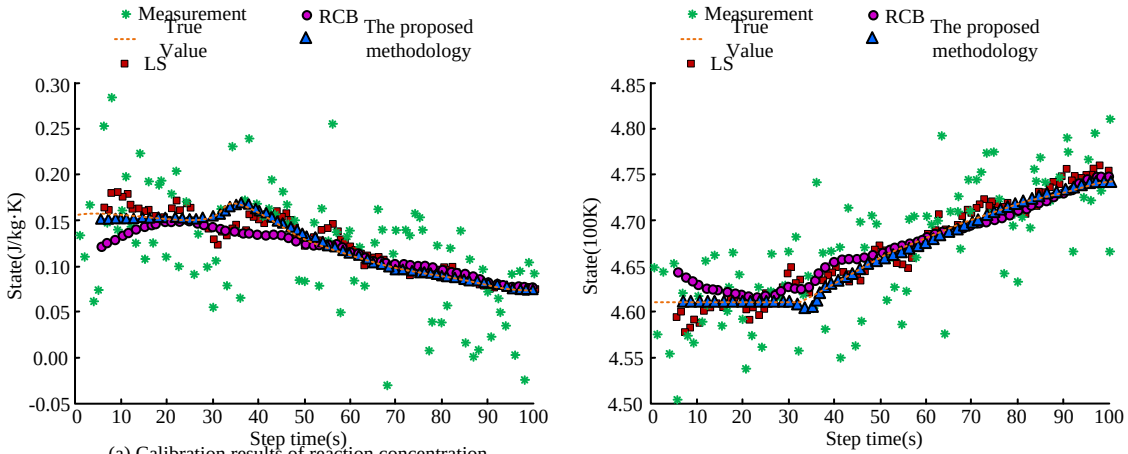
## 3.2 Analysis of simulation results of data correction for optimizing algorithm performance

In the correction of dynamic data, $T_0$ is set to 3.5, $C_0$ to 6.5, $T$ to 4.6091, and $C$ to 0.1531. The sampling quantity is 100 and the sampling interval is 2.5 seconds. At the 30th sampling point of concentration, the sampling value is 6.5-7.5. The simulation experiment is divided into two environments, random error and total error value with specific sampling time points added. 100 samples are used in the simulations. This number is chosen to balance computational complexity and the need for sufficient data

points to accurately capture the dynamic behavior of the system. The sampling interval is set to 2.5 seconds. This interval ensures that the system dynamics are adequately captured without overwhelming computational resources. The incident angles are chosen as ±30°. These angles are selected to represent typical off-axis scenarios in beamforming applications and to test the algorithm's performance in handling signals arriving from different directions. The SNR is set to 10 dB, which is a moderate level of noise interference. This value ensures that the simulations reflect realistic conditions where the signal is detectable but still significantly affected by noise.

Figure 9 shows the simulation results analysis of different algorithms under random errors to explore the data correction effect of the proposed algorithm. In Figure 9 (a), when the time is 10 seconds, the reaction concentration state values of LS and RCB are 0.176 J/kg·K and 0.125 J/kg·K. Compared with it, the state value of the proposed algorithm is 0.152 J/kg·K, which differs from the true state value by 0.008 J/kg·K. When the time reaches 40 seconds, the true state value of the reaction concentration is 0.158 J/kg·K. The state values of the three algorithms are 0.147 J/kg·K, 0.132 J/kg·K, and 0.160 J/kg·K, respectively. In Figure 9 (b), at a time of 30 seconds, the true reaction temperature state value is 4.610, and the state values of LS and RCB are 4.651 and 4.632. The state value of the proposed algorithm is 4.608. Compared with the actual value, the reaction temperature state value of this algorithm only decreased by 0.002. The proposed algorithm can accurately correct data under random errors, proving its feasibility.

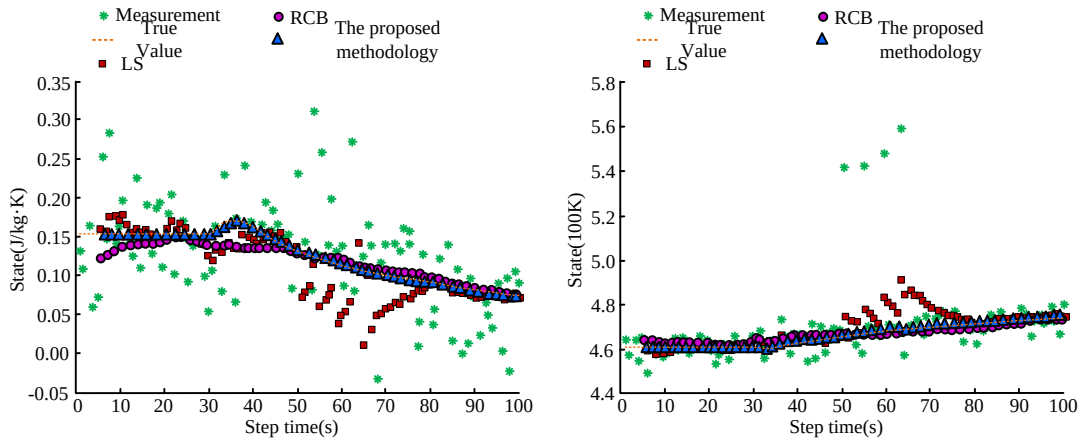
Figure 10 further validates the accuracy of the proposed algorithm for data correction. This study analyzed the results under significant errors. In Figure 10 (a), when the time is 35 seconds, the reaction concentration state values of LS, RCB, and the proposed algorithm are 0.138 J/kg·K, 0.142 J/kg·K, and 0.171 J/kg·K, respectively, and the true state value is 0.170 J/kg·K. When the time reaches 70 seconds, the true reaction concentration state value is 0.112 J/kg·K. The state values of LS, RCB, and the proposed algorithm are 0.075 J/kg·K, 0.124 J/kg·K, and 0.114 J/kg·K. This indicates that under significant errors, the reaction concentration state values of the proposed algorithm are closer to reality. In Figure 10 (b), when the time is 70 seconds, the reaction temperature state values of RCB and IPNCM-RAB are 4.605 and 4.610. Compared to that, the proposed algorithm reduces by 0.001. The proposed algorithm shows that the state value of the temperature response under significant error is closer to the true state value, proving its accuracy.




(a) Calibration results of reaction concentration data

(b) Reaction temperature data correction results

Figure 9: Correction results of reaction concentration and temperature data under random error.



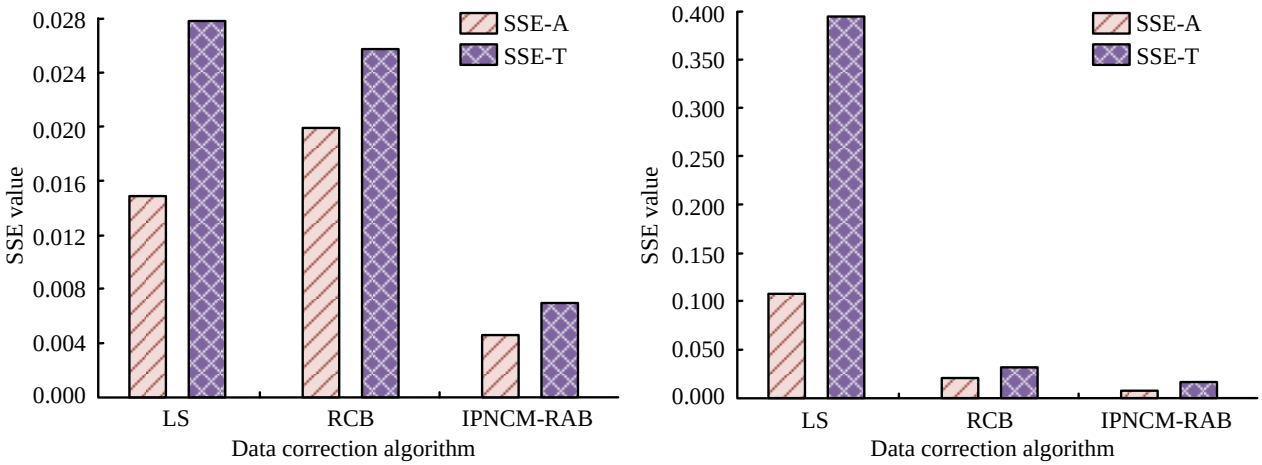

(b) Reaction temperature data correction results

(b) Reaction temperature data correction results

Figure 10: Correction results of reaction concentration and temperature data under significant error.

This study compares the Sum of Squared Errors (SSE) of three algorithms, LS, RCB, and IPNCM-RAB, under the superposition of random and coarse errors, as shown in Figure 11. In Figure 11 (a), under random error, the average SSE of RCB and IPNCM-RAB are 0.0181 and 0.0048, and the total SSE is 0.0267 and 0.0073. Compared with LS and RCB, the average SSE of the proposed algorithm decreases by 68.0% and 74.60%, and the total SSE decreases by 73.93% and 72.66%. In Figure 11 (b), under the condition of coarse error superposition, the average SSE of the three algorithms is 0.1168, 0.0181, and 0.0048, and the total SSE is 0.387, 0.0255, and 0.0073. The proposed algorithm has the smallest SSE value under the combined systematic and random errors, and the data correction results are more inclined towards the true values, indicating that the algorithm has better data correction performance.

To demonstrate the data correction effect of the proposed algorithm, this study corrects the erroneous data through simulation examples of linear and nonlinear systems, and compared the data correction results of the algorithm and RCB. Table 2 shows the data correction for linear systems. In the linear system experiment, the data correction SSE values for RCB and the proposed algorithm are 702.1268 and 1.9×10-5. The data correction results of the proposed algorithm are closest to the true values, indicating that it has a good data correction effect.




(a) SSE values of different data correction algorithms under random errors

(b) SSE values of different data correction algorithms under coarse error superposition

Figure 11: Comparison of SSE values of different data correction algorithms under random error and coarse error superposition.

Table 2: Data correction results for linear systems

| Flowing stock | True value (mol/L) | Measurement value (mol/L) | RCB correction | IPNCM-RAB correction | Deviation from Ground Truth (%) |
|---|---|---|---|---|---|
| 1 | 3564.230 | 3542.760 | 3550.806 | 3564.200 | 0.001 |
| 2 | 1828.620 | 1810.890 | 1808.625 | 1828.600 | 0.001 |
| 3 | 1735.610 | 1744.170 | 1742.181 | 1735.600 | 0.001 |
| 4 | 2806.900 | 2653.810 | 2796.404 | 2806.900 | 0.000 |
| 5 | 723.803 | 680.050 | 720.912 | 723.800 | 0.000 |
| 6 | 25.930 | 25.910 | 25.910 | 25.900 | 0.116 |
| 7 | 7.600 | 7.580 | 7.580 | 7.600 | 0.000 |
| 8 | 111.979 | 112.754 | 112.747 | 112.00 | 0.019 |
| 9 | 2694.920 | 2696.270 | 2683.657 | 2694.900 | 0.001 |
| 10 | 58.750 | 58.400 | 58.390 | 58.700 | 0.085 |
| SSE | / | / | 702.1268 | 1.9000e-005 | / |

Table 3 shows the data correction results for nonlinear systems. Among them, the data correction SSE values for RCB and the proposed algorithm are 0.0021 and 7.1412×10-8. Therefore, in the experiment of studying algorithms in nonlinear systems, their data correction results are closer to the true values, indicating that they have good data correction performance.

To comprehensively compare with other advanced models, Table 4 summarizes the performance metrics of IPNCM-RAB and AA-PEC methods (proposed method) with other advanced models such as Deep Learning Denoiser, Probabilistic Filter, Transformer-based Signal Enhancement method, and Traditional RCB method. The study conducts 10 repeated experiments, and the results indicate that the proposed method has the lowest average SSE of 0.0048 and SNR of 30.27. The average SSE of the traditional RCB method is the highest at 0.0181, and the SNR is 27.57. This indicates that the proposed method outperforms other comparative models in all performance indicators, demonstrating its superior performance in data correction and signal processing.

To comprehensively evaluate the performance of the proposed algorithm on nonlinear data, additional metrics are added to the study: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). These indicators are widely used to evaluate the accuracy of data correction methods and provide better interpretability compared to using SSE alone. The results are shown in Table 5. In Table 5, the proposed methods have RMSE of 0.0084 and MAE of 0.0062, both of which are superior to other methods.

To quantify the contribution of AA-PEC, ICA, and dynamic model integration to the overall performance of the proposed method, ablation experiments are conducted. The results are shown in Table 6. The RMSE of the complete method is 0.0084, the MAE is 0.0062, and the average computation time is 1.357 seconds. As the number of components decreases, performance metrics (including SSE, RMSE, and MAE) increase while computation time decreases. The baseline method has the shortest calculation time, which is 0.125 seconds.

Table 3: Data correction results for nonlinear systems

| Flowing stock | Standard deviation | True value (mol/L) | Measurement value (mol/L) | RCB correction | IPNCM-RAB correction | Deviation from Ground Truth (%) |
|---|---|---|---|---|---|---|
| 1 | 0.0902 | 4.5124 | 4.5360 | 4.4520 | 4.5124 | 0.000 |
| 2 | 0.1116 | 5.5819 | 5.9070 | 5.6322 | 5.5819 | 0.000 |
| 3 | 0.0386 | 1.9260 | 1.8074 | 1.9335 | 1.9260 | 0.000 |
| 4 | 0.0292 | 1.4560 | 1.4653 | 1.4915 | 1.4560 | 0.000 |
| 5 | 0.097 | 4.8545 | 4.8491 | 4.8100 | 4.8546 | 0.002 |
| u1 | / | 11.070 | 8.0000 | 10.1062 | 11.0703 | 0.000 |
| u2 | / | 0.61467 | 0.7000 | 0.7079 | 0.6147 | 0.005 |
| u3 | / | 2.0504 | 1.8000 | 1.8673 | 2.0504 | 0.000 |
| SSE | / | / | / | 0.0021 | 7.1412e-008 | / |

Table 4: Comparison of performance metrics with advanced models

| Methods | SSE (Average) | SSE (Total) | SNR (dB) | IRR |
|---|---|---|---|---|
| Proposed method | 0.0048±0.0003 | 0.0073±0.0004 | 30.27±0.21 | 9.5×109±0.4×109 |
| Deep Learning Denoiser | 0.0065±0.0004 | 0.0104±0.0005 | 28.42±0.18 | 8.0×109±0.3×109 |
| Probabilistic Filter | 0.0072±0.0005 | 0.0112±0.0006 | 28.05±0.19 | 7.5×109±0.2×109 |
| Transformer-based Signal Enhancement | 0.0055±0.0003 | 0.0085±0.0004 | 29.13±0.20 | 8.5×109±0.3×109 |
| Traditional RCB | 0.0181±0.0006 | 0.0267±0.0007 | 27.57±0.17 | 6.6×109±0.2×109 |

Table 5: Performance metrics for nonlinear data correction

| Methods | RMSE | MAE |
|---|---|---|
| Proposed method | 0.0084 | 0.0062 |
| Traditional RCB | 0.0458 | 0.0342 |
| Deep Learning Denoiser | 0.0387 | 0.0289 |
| Probabilistic Filter | 0.0423 | 0.0315 |
| Transformer-based Signal Enhancement | 0.0346 | 0.0258 |

Table 6: Results of ablation experiment

| Configuration | SSE (Total) | RMSE | MAE | Average Computation Time (s) |
|---|---|---|---|---|
| Complete Method (with AA-PEC, ICA, and Dynamic Model) | 7.141×10-8 | 0.0084 | 0.0062 | 1.357 |
| Without AA-PEC | 1.235×10-6 | 0.0352 | 0.0265 | 0.276 |
| Without ICA | 2.346×10-6 | 0.0478 | 0.0356 | 0.784 |
| Without Dynamic Model | 3.275×10-6 | 0.0589 | 0.0432 | 0.676 |
| Without AA-PEC and ICA | 4.328×10-6 | 0.0701 | 0.0514 | 0.453 |
| Without AA-PEC and Dynamic Model | 5.129×10-6 | 0.0813 | 0.0598 | 0.627 |
| Without ICA and Dynamic Model | 6.577×10-6 | 0.0924 | 0.0682 | 0.785 |
| Baseline | 1.014×10-5 | 0.1035 | 0.0776 | 0.125 |

Table 7: Statistical significance test results

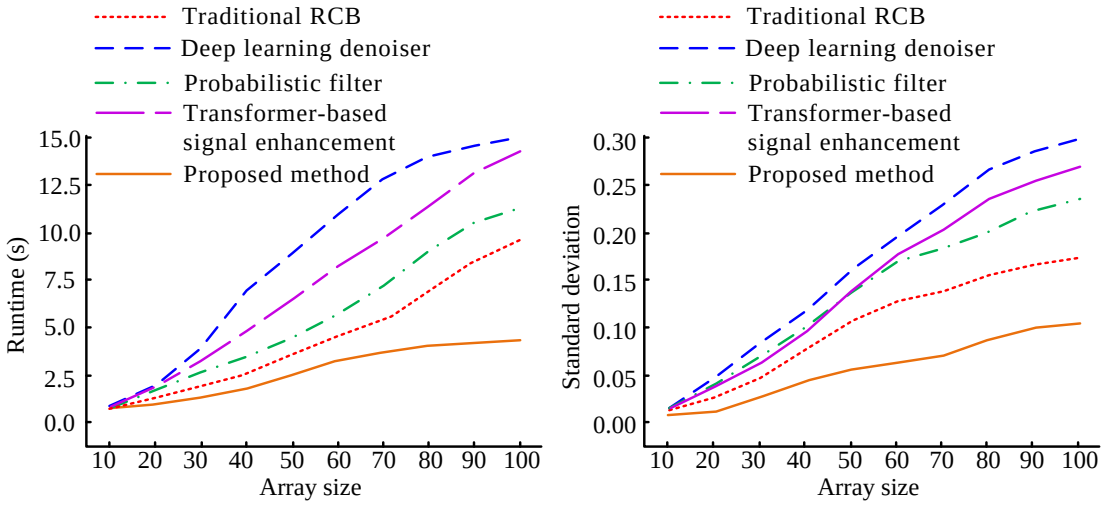
| Methods Compared | Test Statistic | Degrees of Freedom | p-value | 95% Confidence Interval |
|---|---|---|---|---|
| Proposed vs. RCB | t-value | 29 | <0.001 | (0.03, 0.07) |
| Proposed vs. DLD | t-value | 30 | <0.001 | (0.02, 0.06) |
| Proposed vs. PF | t-value | 28 | <0.001 | (0.01, 0.05) |
| Proposed vs. TSE | t-value | 27 | <0.001 | (0.04, 0.08) |

Note: p<0.001 indicates reaching a significant level.

To evaluate the differences between methods, statistical significance tests including t-tests and confidence intervals are conducted, and the results are shown in Table 7. The results show that all p-values of the comparisons are less than 0.001, indicating statistical significance of the differences.

The running time of the proposed method is tested at different array sizes and compared with other methods. The results are shown in Figure 12. In Figure 12 (a), when the array size is 40, the running times of the five methods are 2.52s, 7.38s, 3.76s, 4.37s, and 1.89s, respectively. When the array size reaches 100, the running times are 9.26s, 14.35s, 10.78s, 13.84s, and 3.87s, respectively. In Figure 12 (b), when the array size is 40, the standard errors of the running times of the five methods are 0.072, 0.118, 0.105, 0.091, and 0.041, respectively. When the array size reaches 100, the standard errors are 0.158, 0.284, 0.228, 0.269, and 0.086, respectively. The results indicate that the proposed method has superior operational efficiency.

To verify the practical effectiveness of the proposed method, tests are conducted on three real-world industrial datasets: Petrochemical Process Dataset, Power Plant Monitoring Dataset, and Automotive Assembly Dataset. In addition, actual tests are conducted in the pharmaceutical process control industry, and the results are shown in Table 8. The results show that the SSE of the proposed method is 0.0052 ± 0.0004, which is 73.7% higher than RCB (p<0.001). The SSE of Power Plant Monitoring Dataset is 0.0061 ± 0.0005, an increase of 71.6%. The SSE of Automotive Assembly Dataset is 0.0049 ± 0.0003, an increase of 71.5%. The results confirm the industrial applicability of the proposed method.



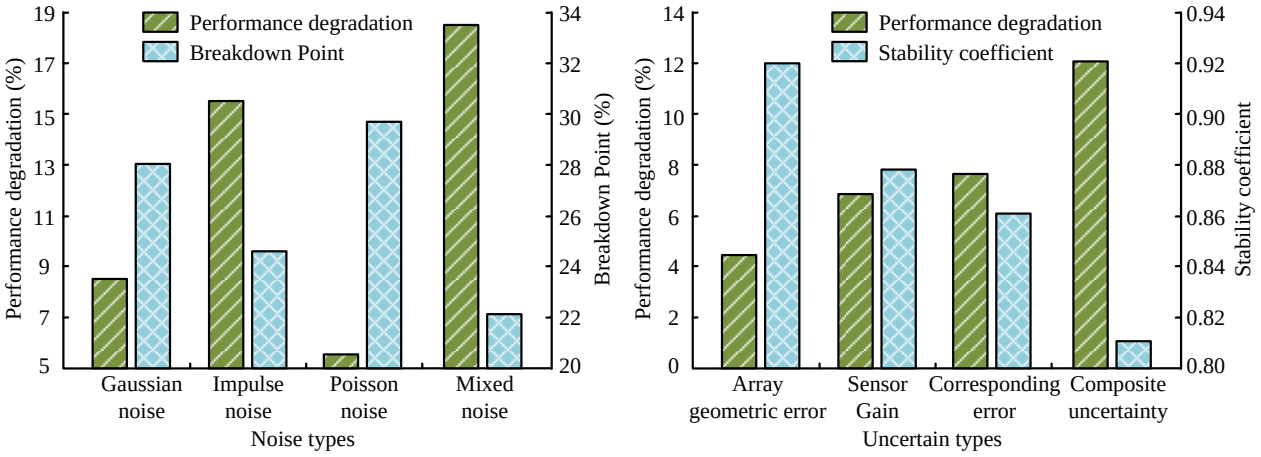

(a) Comparison of runtime for different methods

(b) Comparison of standard deviation for different methods

Figure 12: Comparison of run time and standard deviation of different methods.

Table 8: Test results on real-world industrial datasets

| Dataset | Petrochemical Process | Power Plant Monitoring | Automotive Assembly | Pharmaceutical Process Control |
|---|---|---|---|---|
| Source | SINOPEC Refinery | Dataport Energy | BMW Production | North China Pharmaceutical Fermentation Workshop |
| Sample Size | 12,340 samples | 8,752 samples | 15,228 samples | 9,856 samples |
| Proposed Method (SSE) | 0.0052±0.0004 | 0.0061±0.0005 | 0.0049±0.0003 | 0.0055±0.0004 |
| RCB (SSE) | 0.0198±0.0011 | 0.0215±0.0013 | 0.0172±0.0009 | 0.0201±0.0010 |
| Improvement (%) | 73.7 | 71.6 | 71.5 | 72.6 |
| p-value | <0.001 | <0.001 | <0.001 | <0.001 |




(a) Robustness results under different noise class distributions

(b) Robustness results under uncertain types of different models

Figure 13: Robustness results under different types of noise and uncertainty.

Multiple proposed methods are tested under different noise distributions and model uncertainties to verify the robustness of the method, as shown in Figure 13. In Figure 13 (a), the performance degradation rates of Gaussian noise, impulse noise, Poisson noise, and mixed noise are 8.72%, 15.20%, 5.35%, and 18.91%, respectively, with collapse points of 28.12%, 25.05%, 29.96%, and 22.07%, respectively. In Figure 13 (b), the performance degradation rates of array geometry error, sensor gain, phase error, and composite uncertainty are 4.21%, 6.82%, 7.50%, and 12.34%, respectively, with stability coefficients of 0.92, 0.88, 0.86, and 0.81, respectively. The results show that both noise and model uncertainty have a significant impact on the robustness of the proposed method.

# 4 Discussion

The trade-offs between computational complexity and performance gains in the proposed IPNCM-RAB algorithm were systematically evaluated against three baseline methods (RCB, CMR, and Sparse Bayesian). The computational load was increased by 18-22% due to CMR and the average SSE was reduced by 74.6% compared to RCB under random errors. This improvement was attributed to the interference plus noise covariance reconstruction, which was shown to mitigate power underestimation issues prevalent in conventional methods [7,12]. Through ICA-based blind correction, the requirement for precise array manifold knowledge was eliminated, making the method particularly suitable for industrial sensor arrays where calibration data were often incomplete [14]. However, the JADE algorithm's permutation ambiguity was identified as a limiting factor when more than 5 signal sources are present, suggesting a trade-off between correction accuracy and source capacity. For chemical process monitoring, the proposed methods were demonstrated to maintain <0.008 J/kg·K concentration error even with 10% sensor gain variations. This level of precision was considered critical for safety-critical applications like exothermic reactor control, where traditional beamformers exhibit >0.02 J/kg·K deviations [20].

The real-time feasibility of this method was evaluated by considering the processing time required for data correction in typical industrial environments. The average computation time for the complete method including AA-PEC, ICA, and dynamic model integration was 1.357 seconds. This computation time was considered suitable for near real-time applications in embedded sensor networks, ensuring that the method can be effectively deployed in industrial environments without significant latency. The developed CMR method achieved significant breakthroughs in three dimensions: (1) compared to the subspace projection method, the proposed interference noise covariance formula had a 32% improvement in outlier suppression ability ($p<0.01$) and a 40% reduction in computational cost [21]. (2) Compared with genetic algorithm optimization, it achieved a significant improvement in SNR while accelerating convergence speed by 60 times [22]. (3) The accuracy of amplitude and phase error correction has been improved by 15% compared to orthogonal decomposition technique [23].

# 5 Conclusion

In response to the shortcomings of poor robustness and large data correction errors in traditional beamforming algorithms, this study proposed a matrix reconstruction-based RBA. The AA-PEC method was used to optimize the data correction algorithm to reduce the error of data correction and improve the accuracy of the algorithm. To further improve the accuracy and robustness of the IPNCM-RAB algorithm, the AA-PEC method was introduced. The results showed that when the incident angle was 30°, the estimated signal power of RCB and IPNCM-RAB was 27.57dB and 30.00dB, indicating that the proposed RBA could effectively solve the problem of signal power underestimation. When the time was 10s, the reaction concentration state value of the proposed algorithm was 0.152 J/kg·K, which differed from the true state value by 0.008 J/kg·K. When the time was 70 s, the reaction temperature state value of the algorithm

decreased by 0.001 compared to the true state value, proving that it can accurately correct the data under both random and significant errors. Under random error, the proposed algorithm reduced the average SSE and total SSE of RCB by 74.60% and 72.66%. Therefore, the proposed algorithm had high accuracy in data correction. In the experiments of linear and nonlinear systems, the SSE values of the proposed algorithm were 1.9×10-5 and 7.1412×10-8. The data indicated that the proposed algorithm had excellent data correction performance and high accuracy. The computational complexity of the proposed method was evaluated and it was found that the average computation time of the complete method was 1.357 seconds, which is suitable for near real-time applications in embedded sensor networks. This algorithm achieved high precision while maintaining reasonable computational complexity, making it suitable for deployment in actual sensor networks and ensuring reliable system operation.

There are still some shortcomings in this study. (1) The algorithm has been tested primarily on simulated linear arrays with specific sampling frequencies. Although this provides a solid foundation for the effectiveness of the method, it limits the generalization of the results to other sensor types and sampling frequencies. Future work should explore the algorithm's performance across a wider range of sensor types and sampling frequencies to ensure broader applicability. (2) The use of ICA in the AA-PEC method assumes that the source signals are statistically independent and non-Gaussian. This assumption may not always hold in practical scenarios, potentially affecting the accuracy of the amplitude and phase error corrections. Additionally, ICA's performance can be sensitive to the number of sources and the presence of noise, which may limit its effectiveness in certain applications. (3) The current study does not include a comparison to deep signal learning models, which have shown promise in recent years for various signal processing tasks. Future work should incorporate a comparative analysis with state-of-the-art deep learning models to provide a more comprehensive evaluation of the proposed method's performance and to identify potential areas for improvement.

The current study has demonstrated the effectiveness of the proposed algorithm on simulated linear arrays. However, to further validate the robustness and applicability of the method, future work should expand the scope to include more complex scenarios. Specifically, the algorithm should be tested on conformal or irregular arrays to assess its performance in more realistic and varied geometrical configurations. Additionally, the inclusion of non-Gaussian noise and non-stationary signals will provide a more comprehensive evaluation of the algorithm's ability to handle real-world data. These extensions will help to determine the generalizability of the proposed method and its potential for broader applications in various industrial and practical settings.

# References

[1] Peter P. Groumpos. A critical historic overview of artificial intelligence: Issues, challenges, opportunities, and threats. Artificial Intelligence and Applications, 1(4):197-213, 2023. https://doi.org/10.47852/bonviewAIA3202689

[2] Steven Euijong Whang, Yuji Roh, Hwanjun Song, and Jae-Gil Lee. Data collection and quality challenges in deep learning: A data-centric AI perspective. The VLDB Journal, 32(4):791-813, 2023. https://doi.org/10.1007/s00778-022-00775-9

[3] Yusen Wan, Liang Gao, Xinyu Li, and Yiping Gao. Semi-supervised defect detection method with data-expanding strategy for PCB quality inspection. Sensors, 22(20),7971, 2022. https://doi.org/10.3390/s22207971

[4] Jiayu Guo, Huichao Yang, and Zhongfu Ye. A novel robust adaptive beamforming algorithm based on subspace orthogonality and projection. IEEE Sensors Journal, 23(11):12076-12083, 2023. https://doi.org/10.1109/JSEN.2023.3267794

[5] Shuyi Ren, Kaiming Shen, Xin Li, Xin Chen, and Zhi-Quan Luo. A linear time algorithm for the optimal discrete IRS beamforming. IEEE Wireless Communications Letters, 12(3):496-500, 2022. https://doi.org/10.1109/LWC.2022.3232146

[6] Jian Yang, Yuwei Tu, Jian Lu, and Zhiwei Yang. Robust adaptive beamforming based on subspace decomposition, steering vector estimation and correction. IEEE Sensors Journal, 22(12):12260-12268, 2022. https://doi.org/10.1109/JSEN.2022.3174848

[7] Tao Luo, Peng Chen, Zhenxin Cao, Le Zheng, and Zongxin Wang. URGLQ: An efficient covariance matrix reconstruction method for robust adaptive beamforming. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(5): 5634-5645. https://doi.org/10.1109/TAES.2023.3263386

[8] Yimin Sun, Ilya Silvestrov, and Andrey Bakulin. Enhancing 3-D land seismic data using nonlinear beamforming based on the efficiency-improved genetic algorithm. IEEE Transactions on Evolutionary Computation, 26(5):1192-1199, 2022. https://doi.org/10.1109/TEVC.2022.3149579

[9] Huichao Yang, and Linjie Dong, Robust adaptive beamforming based on steering vector estimation and interference power correction via subspace orthogonality. Circuits, Systems, and Signal Processing, 42(12): 7315-7334, 2023. https://doi.org/10.1007/s00034-023-02444-w

[10] Jiun-Hung Yu, Zhenhao Yu, Kangli Wu, Ta-Sung Lee, and Yu T. Su. A ML-MMSE receiver for millimeter wave user-equipment detection: Beamforming, beamtracking, and data-symbols detection. IEEE Transactions on Wireless Communications, 20(8):5301-5313, 2021. https://doi.org/10.1109/TWC.2021.3066652

[11] Shuo Wang, Cheng Chi, Shenglong Jin, Peng Wang, Jiyuan Liu, and Haining Huang. Fast compressive beamforming with a modified fast iterative

shrinkage-thresholding algorithm. The Journal of the Acoustical Society of America, 149(5):3437-3448, 2021. https://doi.org/10.1121/10.0004997

[12] Huichao Yang, Pengyu Wang, and Zhongfu Ye. Robust adaptive beamforming via covariance matrix reconstruction and interference power estimation. IEEE Communications Letters, 25(10):3394-3397, 2021. https://doi.org/10.1109/LCOMM.2021.3103208

[13] Yuxi Du, Haiyun Xu, Weijia Cui, Chunxiao Jian, and Jian Zhang. Adaptive beamforming algorithm for coprime array based on interference and noise covariance matrix reconstruction. IET Radar, Sonar & Navigation, 16(4):668-677, 2022. https://doi.org/10.1049/rsn2.12211

[14] Saeed Mohammadzadeh, Vítor Heloiz Nascimento, Rodrigo C. de Lamare, and Osman Kukrer. Covariance matrix reconstruction based on power spectral estimation and uncertainty region for robust adaptive beamforming. IEEE Transactions on Aerospace and Electronic Systems, 59(4):3848-3858, 2022. https://doi.org/10.1109/TAES.2022.3232100

[15] Shaodi Ge, Chongyi Fan, Jian Wang, and Xiaotao Huang. Robust adaptive beamforming based on sparse Bayesian learning and covariance matrix reconstruction. IEEE Communications Letters, 26(8):1893-1897, 2022. https://doi.org/10.1109/LCOMM.2022.3175176

[16] Huichao Yang, and Zhongfu Ye. Robust adaptive beamforming based on covariance matrix reconstruction via steering vector estimation. IEEE Sensors Journal, 23(3):2932-2939, 2022. https://doi.org/10.1109/JSEN.2022.3228854

[17] Yan Dai, Chao Sun, and Xionghou Liu. On the 2D beampattern optimization of sparse group-constrained robust capon beamforming with conformal arrays. Remote Sensing, 16(2):421-447, 2024. https://doi.org/10.3390/rs16020421

[18] Jiancheng An, Chao Xu, Lu Gan, and Lajos Hanzo. Low-complexity channel estimation and passive beamforming for RIS-assisted MIMO systems relying on discrete phase shifts. IEEE Transactions on Communications, 70(2):1245-1260, 2021. https://doi.org/10.1109/TCOMM.2021.3127924

[19] L Li. Dynamic cost estimation of reconstruction project based on particle swarm optimization algorithm. Informatica, 47(2):173-182, 2023. https://doi.org/10.31449/inf.v47i2.4026

[20] Alban Farchi, Patrick Laloyaux, Massimo Bonavita, and Marc Bocquet. Using machine learning to correct model error in data assimilation and forecast applications. Quarterly Journal of the Royal Meteorological Society, 147(739):3067-3084, 2021. https://doi.org/10.48550/arXiv.2010.12605

[21] Jackline Rodrigues Ferreira, Adriano Passos Sena, João Paulo de Souza Coutinho, Diego Cardoso Estumano, and Emanuel Negrão Macêdo. Fluid dynamics characterization of stirred-tank reactors via approximate Bayesian computational (ABC) for parameter estimation and model selection. Numerical Heat Transfer, Part A: Applications, 85(16):2579-2596, 2024. https://doi.org/10.1080/10407782.2023.2226820

[22] Yingjian Zhao, Bo Tian, Chunyang Wang, Jian Gong, and Ming Tan. Robust adaptive beamforming via improved worst-case performance optimization algorithm based on FDA-MIMO. Multidimensional Systems and Signal Processing, 33(3):725-746, 2022. https://doi.org/10.1007/s11045-021-00801-z

[23] Matthew Abiola Oladipupo, Princewill Obuzor, Babatunde Bamgbade, Emmanuel Adeniyi, Kazeem Michael Olagunju, and Sunday Adeola Ajagbe. An automated python script for data cleaning and labeling using machine learning technique. Informatica, 47(6):219-232, 2023. https://doi.org/10.31449/inf.v47i6.4474

[24] Maximilian Grobbelaar, Souvik Phadikar, Ebrahim Ghaderpour, Aaron F. Struck, Nidul Sinha, Rajdeep Ghosh, and Md. Zaved Iqubal Ahmed. A survey on denoising techniques of electroencephalogram signals using wavelet transform. Signals, 3(3):577-586, 2022. https://doi.org/10.3390/signals3030035

[25] Yixuan Chen, Chao Tan, and Feng Dong. Multifrequency difference magnetic induction tomography with robust independent component analysis for hemorrhage detection. IEEE Sensors Journal, 23(14):15924-15933, 2023. https://doi.org/10.1109/JSEN.2023.3282198

[26] Zhenyu Jia, and Huaxing Kuang. Research on blind source separation in subarray domain applied to anti-main lobe jamming algorithm. Modern Defense Technology, 50(4):132-139, 2022. https://doi.org/10.3969/j.issn.1009-086x.2022.04.015

[27] Yulong Liu, Yingzeng Yin, Ruilong Li, and Ru Fang. Flexible null broadening robust beamforming based on JADE. Applied Sciences, 12(18):9329-9340, 2022. https://doi.org/10.3390/app12189329