



OPEN ACCESS

EDITED BY

Jay S. Pearlman,
Institute of Electrical and Electronics
Engineers, France

REVIEWED BY

Xiaoyuan Li,
Northwestern Polytechnical University,
China
Xinghao Qu,
Harbin Engineering University, China
Anughna N.,
Amity University, India

*CORRESPONDENCE

Yi Zhang
✉ zhangyi61@mail.nwpu.edu.cn

RECEIVED 01 December 2025

REVISED 14 February 2026

ACCEPTED 23 February 2026

PUBLISHED 13 March 2026

CITATION

Zhang Y, Zhang X and Yang J (2026)
Tensor-based DOA estimation for hybrid
scalar–vector sensor arrays.
Front. Mar. Sci. 13:1758382.
doi: 10.3389/fmars.2026.1758382

COPYRIGHT

© 2026 Zhang, Zhang and Yang. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Tensor-based DOA estimation for hybrid scalar–vector sensor arrays

Yi Zhang^{1*}, Xiaomin Zhang¹ and Jie Yang²

¹School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, China, ²Avic Shaanxi Aero Electric Co.,LTD, Xi'an, Shaanxi, China

Introduction: Underwater acoustic direction-of-arrival (DOA) estimation is fundamental to submarine detection, offshore exploration, and autonomous underwater vehicle navigation, where accurate source localization under severe multipath propagation and low signal-to-noise ratios remains challenging. Existing methods face critical limitations: traditional uniform arrays require extensive physical sensors to achieve sufficient spatial resolution, resulting in high hardware costs; vector sensor arrays typically neglect sparse array geometries that enable virtual aperture expansion; and conventional matrix-based processing discards multidimensional structural information through vectorization, leading to suboptimal performance.

Methods: To address these challenges, this paper presents a novel tensor-based DOA estimation framework that integrates hybrid scalar-vector sensor arrays (HSVSAs), fourth-order tensor modeling, and higher-order singular value decomposition (HOSVD). The HSVSA architecture combines vector and scalar sensors in a hybrid configuration to create virtual sensors, improving degrees of freedom while reducing hardware complexity. The tensor model preserves spatial–polarization coupling, enabling robust subspace estimation without iterative optimization.

Results: Simulations demonstrate that the proposed method outperforms conventional approaches.

Discussion: The proposed framework offers a practical solution for resource-constrained underwater acoustic applications, with potential for further optimization in real-world scenarios.

KEYWORDS

direction-of-arrival estimation, higher-order singular value decomposition, hybrid scalar-vector sensor arrays, polarization diversity, subspace estimation, tensor modeling, underwater acoustics, virtual aperture

1 Introduction

Direction-of-arrival (DOA) estimation constitutes a cornerstone technique in underwater acoustic signal processing, with widespread applications spanning submarine detection, offshore exploration, environmental monitoring, and autonomous underwater vehicle navigation (Zhang and Sun, 2023; Dong et al., 2024b). Recent advances have extended to underwater positioning systems using multistatic sonar configurations (Zhang et al., 2022; Zhang and Zhang, 2024), demonstrating the growing importance of accurate localization in autonomous underwater vehicle operations. The underwater acoustic

channel imposes severe frequency-dependent attenuation, multipath reflections, and low signal-to-noise ratios (Dong et al., 2024a; Li and Wang, 2023). When deployment constraints restrict physical aperture—common for autonomous platforms and towed arrays—conventional uniform sensor arrays face a critical dilemma: insufficient spatial sampling degrades angular resolution, while increasing sensor count escalates costs and system complexity.

To address this fundamental challenge, classical subspace methods—beamforming (Van Trees, 2002), Multiple Signal Classification (MUSIC) (Schmidt, 1986), and Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT)—have established the theoretical foundation for array signal processing. Schmidt's MUSIC algorithm achieves angular super-resolution through eigendecomposition of the spatial covariance matrix. However, these methods rely on uniform linear arrays (ULAs) where degrees of freedom (DOF)—the maximum number of resolvable sources—scale linearly with sensor count N , creating a direct trade-off between multi-source capability and hardware complexity.

Recent advances in sparse array geometries have shown promise in breaking this linear DOF constraint through virtual aperture synthesis. Nested arrays (Pal and Vaidyanathan, 2010; Li and Zhang, 2023; Chen et al., 2024) and coprime arrays (Vaidyanathan and Pal, 2010; Zheng et al., 2020) achieve $O(N^2)$ DOF using only $O(N)$ physical sensors through difference co-array processing. Chen et al. (2024) further enhanced nested array performance through dilated configurations with reduced mutual coupling effects. Nevertheless, existing sparse arrays face a fundamental design limitation: they employ either homogeneous scalar sensors—sacrificing polarization diversity—or homogeneous vector sensors, as exemplified by Han and Nehorai's nested vector-sensor array (Han and Nehorai, 2014), which pioneered the integration of nested geometry with vector sensors to achieve enhanced degrees of freedom through interspectral tensor modeling and novel spatial smoothing, yet mandates costly acoustic vector sensors (AVSs) throughout the array, incurring prohibitive costs.

While AVSs offer significant advantages by co-locating pressure and tri-axial velocity transducers to encode polarization diversity (Nehorai and Paldi, 1994), their deployment faces practical constraints. Recent underwater studies have demonstrated enhanced performance through various approaches: Dong et al. (2024b) and Dong et al. (2024a) developed preprocessing techniques for challenging conditions, such as low signal-to-noise ratio (SNR) and multipath conditions, while Chen et al. (2023) achieved co-frequency source separation using coprime vector arrays. Modern deep learning approaches (Li and Wang, 2023; Xu et al., 2025) have also shown promise for complex underwater scenarios, though they require extensive training data. Building upon these developments, Qu et al. (2022) proposed an augmented tensor MUSIC algorithm that generalizes tensor models for spatially separated sensors, yet computational complexity remains a concern for real-time applications. Robust methods for sparse arrays (Zhu et al., 2025) have addressed low-snapshot scenarios through hybrid neural architectures. Despite these advances, integrating heterogeneous sensors—mixing cost-effective scalars

with polarization-rich vectors—into a unified geometric structure remains challenging.

To overcome these limitations, tensor algebra has emerged as a powerful framework for preserving multidimensional structure in array processing (Kolda and Bader, 2009). Higher-order singular value decomposition (HOSVD) (De Lathauwer et al., 2000; Haardt et al., 2008) provides deterministic subspace estimation by exploiting structure inherent in multidimensional measurement data, with applications in Multiple-Input Multiple-Output (MIMO) radar (Nion and Sidiropoulos, 2010) and wideband scenarios (Li et al., 2024) demonstrating significant performance gains. Ahmed et al. (2020) further developed higher-order propagator methods with computational advantages, while hybrid receiver architectures (Abdelbadie et al., 2024) have explored tensor decomposition for enhanced spatial coverage. However, existing tensor methods assume homogeneous sensor configurations, whether for MIMO radar applications or AVS arrays (Qu et al., 2022). This homogeneity assumption fundamentally limits their applicability to hybrid scalar–vector arrays, where elements contribute measurements of different dimensionality.

Given these challenges, three critical research gaps emerge that prevent practical deployment of high-resolution underwater DOA estimation: (1) strategic integration of heterogeneous sensors to balance cost and polarization diversity; (2) tensor formulations capable of fusing measurements of different dimensionality within unified models; and (3) deterministic closed-form algorithms suitable for real-time embedded platforms. Addressing these gaps is essential for enabling cost-effective, high-performance DOA estimation in autonomous underwater vehicles, distributed sensor networks, and resource-constrained sonar systems.

To systematically address these fundamental limitations, this paper proposes a unified tensor-based framework for DOA estimation using hybrid scalar–vector sensor arrays (HSVSAs). By strategically integrating sparse geometry principles, polarization diversity exploitation, and advanced tensor algebraic decomposition, the proposed framework achieves enhanced angular resolution while significantly reducing hardware complexity. The primary innovations comprise four interconnected contributions:

1. **Heterogeneous Hybrid Array Architecture:** We develop the first scalar–vector hybrid configuration that strategically deploys N_1 vector sensors for polarization diversity alongside N_2 scalar sensors for co-array expansion. This architecture achieves $\tilde{N} = N_1(N_2 + 1)$ virtual sensors using only $N = N_1 + N_2$ physical sensors, enabling substantial cost reduction while preserving enhanced DOF through difference co-array processing.
2. **Fourth-Order Spatially Smoothed Tensor Model:** Building upon the hybrid array design, we formulate an interspectral tensor $T \in \mathbb{C}^{\tilde{N} \times \tilde{N} \times 4 \times 4}$ that simultaneously exploits spatial correlation from hybrid array geometry, polarization structure from vector sensors, and statistical decorrelation through spatial smoothing. This unified tensor representation preserves multidimensional coupling that conventional matrix-based methods discard through vectorization.

3. HOSVD-Based Closed-Form Estimation Algorithm: To ensure computational efficiency, we develop a deterministic HOSVD-based algorithm that exploits rotational invariance properties for direct closed-form DOA retrieval. This approach eliminates the convergence uncertainty of iterative methods (Nion and Sidiropoulos, 2010) and the computational burden of exhaustive grid search (Qu et al., 2022), making it suitable for real-time embedded platforms.
4. Comprehensive Theoretical Performance Analysis: Finally, we derive Cramér–Rao bounds (CRBs) specifically tailored for heterogeneous HSVSA geometries and demonstrate near-optimal estimation performance through extensive validation across source enumeration, angular resolution, and estimation accuracy metrics.

These contributions collectively address the identified research gaps and provide a comprehensive solution for cost-effective underwater DOA estimation. Beyond underwater acoustics, the proposed framework exhibits broader applicability to emerging array signal processing domains, including reconfigurable intelligent surfaces (RIS), MIMO communications, and mmWave systems, where heterogeneous sensor architectures and closed-form tensor decompositions offer similar computational and performance advantages. The remainder of this paper is organized as follows: Section 2 presents the detailed methodology including HSVSA design, tensor signal modeling, and the HOSVD-based estimation algorithm. Section 3 provides theoretical performance analysis with CRB derivation and computational complexity assessment. Section 4 validates the proposed framework through comprehensive Monte Carlo simulations. Section 5 concludes with discussions and future research directions. Mathematical notations are summarized in Table 1 for reference.

TABLE 1 Mathematical notations.

Notation	Description
\mathbb{R}, \mathbb{C}	Real and complex number fields
\mathbf{A}, \mathcal{A}	Matrix and tensor
$\mathbf{A}^T, \mathbf{A}^{H}$	Transpose and conjugate transpose
$\mathbf{A}^{-1}, \mathbf{A}^{\dagger}$	Inverse and pseudo-inverse
$\ \mathbf{A}\ _F$	Frobenius norm
$\text{vec}(\mathbf{A})$	Vectorization operator
\odot, \otimes, \circ	Hadamard, Kronecker, and outer products
$\mathbb{E} [\cdot]$	Statistical expectation
$\text{diag}(\cdot)$	Diagonal matrix constructor
$\text{Re}\{\cdot\}, \text{Im}\{\cdot\}$	Real and imaginary parts
argmin/max	Argument of the minimum/maximum
N, \tilde{N}, N_s	Number of physical, virtual sensors and number of snapshots
K, \hat{K}	True and estimated number of sources

2 Methodology

This section presents the proposed tensor-based DOA estimation framework for HSVSAs. We begin by detailing the design of the HSVSA and proceed to develop a tensor signal model incorporating spatial smoothing techniques. Additionally, we introduce the HOSVD-based estimation algorithm used for the DOA estimation process.

2.1 Hybrid scalar–vector sensor array design

The proposed HSVSA architecture, illustrated in Figure 1, integrates heterogeneous sensor types into a unified hybrid array with uniform inter-element spacing. Specifically, all sensors are arranged with a half-wavelength spacing $d = \lambda/2$. The array comprises two subarrays: a vector uniform linear array (v-ULA) consisting of N_1 AVSs located at $S_v = \{0, d, 2d, \dots, (N_1 - 1)d\}$, where each sensor measures acoustic pressure and three orthogonal particle-velocity components; and a scalar uniform linear array (s-ULA) composed of N_2 pressure sensors positioned at $S_s = \{N_1 d, (N_1 + 1)d, \dots, (N_1 + N_2 - 1)d\}$. The first vector sensor is selected as the phase reference. This hybrid configuration combines the polarization diversity offered by vector sensors with the cost efficiency of scalar sensors and facilitates virtual aperture extension through difference co-array processing.

The HSVSA achieves DOF enhancement through sensor heterogeneity with uniform spacing, contrasting with traditional nested arrays that rely on unequal inter-element spacings. This heterogeneous-uniform design provides three key advantages: (1) simplified mechanical design with identical mounting requirements, (2) reduced calibration complexity due to uniform spacing, and (3) preserved DOF enhancement through difference co-array processing on heterogeneous measurements.

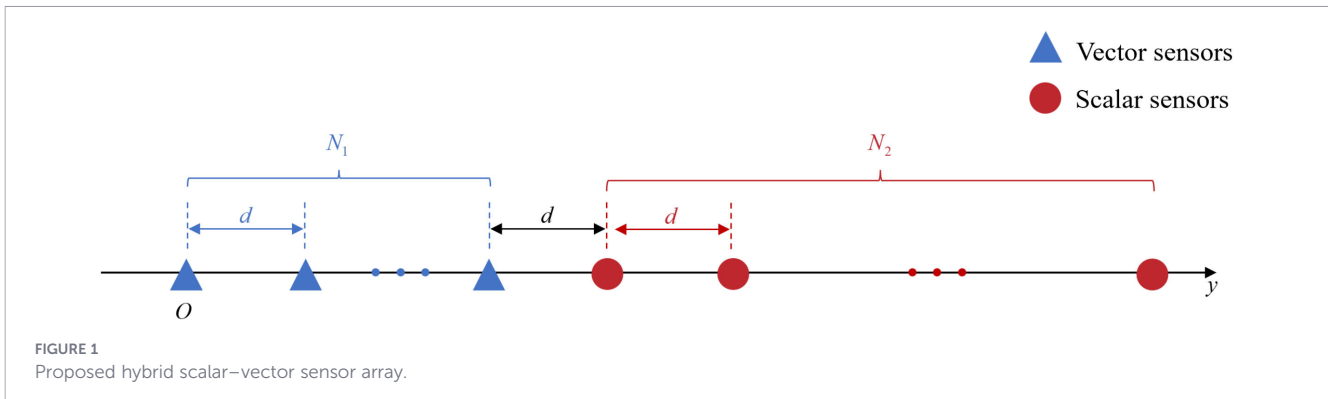
This configuration delivers significant spatial efficiency: using $N = N_1 + N_2$ heterogeneous sensors across physical aperture $(N_1 + N_2 - 1)d$, difference co-array expansion generates $\tilde{N} = N_1(N_2 + 1)$ consecutive virtual positions spanning virtual aperture $(\tilde{N} - 1)d$. The result is $\mathcal{O}(N^2)$ DOF scaling with only $\mathcal{O}(N)$ physical sensors while maintaining uniform spacing advantages.

2.2 Signal model and co-array processing

Each AVS in the v-ULA simultaneously captures a four-dimensional measurement vector $y = [p, v_x, v_y, v_z]^T \in \mathbb{C}^4$, comprising acoustic pressure and three orthogonal particle velocity components that encode polarization information. Consider K uncorrelated narrowband far-field sources at $\{(\theta_k, \varphi_k)\}_{k=1}^K$ with powers $\{\sigma_k^2\}_{k=1}^K$. The received signal at the v-ULA is given by Equation 1:

$$y_1(t) = A_1 x(t) + e_1(t) \tag{1}$$

where $A_1 \in \mathbb{C}^{4N_1 \times K}$ is the array manifold with k th column $a_k = d_k \otimes p_k$. Here, $d_k \in \mathbb{C}^{N_1}$ is the spatial phase vector and $p_k = [1, u_k^T]^T \in \mathbb{C}^4$ contains the polarization response with $u_k = [\cos \varphi_k$



$\cos \theta_k, \sin \varphi_k \cos \theta_k, \sin \theta_k]^T$. For the s -ULA, $y_2(t) = A_2 x(t) + e_2(t)$ with $A_2 \in \mathbb{C}^{N_2 \times K}$ containing only spatial phase information.

By stacking the outputs of the two subarrays, the overall received signal of the HSVSA can be expressed as Equation 2:

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \underbrace{\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}}_{A \in \mathbb{C}^{(4N_1 + N_2) \times K}} x(t) + \underbrace{\begin{bmatrix} e_1(t) \\ e_2(t) \end{bmatrix}}_{e(t)}. \quad (2)$$

Under the assumption of temporally independent sources $x_k(t) \sim \mathcal{CN}(0, \sigma_k^2)$, we consider the practical scenario where AVSs exhibit spatially non-uniform noise characteristics. Unlike conventional assumptions of white noise $e(t) \sim \mathcal{CN}(0, \sigma_e^2 I)$, the pressure and velocity channels typically have unequal noise powers due to different transduction mechanisms and flow-induced interference (Nehorai and Paldi, 1994).

For the vector subarray, the non-uniform noise covariance is:

$$\Sigma_e = \text{diag}(\sigma_p^2, \sigma_x^2, \sigma_y^2, \sigma_z^2) \otimes I_{N_1}, \quad (3)$$

where σ_p^2 and $\{\sigma_v^2\}_{v \in \{x,y,z\}}$ represent the noise variances of pressure and velocity channels, respectively. In practice, these noise parameters can be estimated from calibration measurements conducted in controlled acoustic environments or from signal-free data segments during deployment. The overall covariance matrix becomes $R_y = A R_x A^H + \Sigma_e$ with $R_x = \text{diag}(\sigma_1^2, \dots, \sigma_K^2)$.

To this practical challenge, we incorporate a pre-whitening step before HOSVD decomposition: $\tilde{R}_y = \Sigma_e^{-1/2} R_y \Sigma_e^{-H/2}$, ensuring that the subsequent tensor operations remain statistically optimal under non-uniform noise conditions. Applying vectorization yields the augmented observation model, as shown in Equation 4:

$$v = \text{vec}(R_y) = (A^* \odot A) s + \sigma_e^2 \text{vec}(I), \quad (4)$$

where $A_{\text{aug}} = A^* \odot A \in \mathbb{C}^{(4N_1 + N_2)^2 \times K}$ constructs a virtual array with $(4N_1 + N_2)^2$ difference co-array positions. After removing redundancies through spatial smoothing, the effective aperture contains $\tilde{N} = N_1(N_2 + 1)$ consecutive virtual sensors, enabling the HSVSA to achieve \tilde{N} degrees of freedom with only N physical sensors.

2.3 Tensor-based signal processing

Traditional matrix-based methods vectorize multidimensional measurements, discarding natural coupling between spatial and polarization dimensions. We introduce a tensor-based framework

preserving the inherent multidimensional structure. The array manifold tensor $\mathcal{A} \in \mathbb{C}^{N \times 4 \times K}$ is constructed by stacking source-specific matrices $A_k = d_k \circ p_k^T$, yielding Equation 5:

$$Y(t) = \mathcal{A} \times_3 s(t) + E(t) \in \mathbb{C}^{N \times 4}, \quad (5)$$

where \times_3 denotes mode-3 tensor-vector product. This preserves spatial-polarization coupling lost in matrix vectorization.

The fourth-order cross-correlation tensor is defined in Equation 6:

$$\mathcal{R} = E\{Y \circ Y^*\} = \mathcal{A} \times_3 E\{s \circ s^*\} \times_3 \mathcal{A}^* + E\{E \circ E^*\}, \quad (6)$$

where $\mathcal{R} \in \mathbb{C}^{N \times 4 \times N \times 4}$ encapsulates the complete second-order statistics with $E\{s \circ s^*\} = \text{diag}(\sigma_1^2, \dots, \sigma_K^2)$. Mode-2 matricization yields Equation 7:

$$U = \mathcal{R}_{(2)}^T = (\mathcal{A}_{(3)}^H \odot \mathcal{A}) \times_3 s + \sigma_e^2 \vec{I} \in \mathbb{C}^{4(N_1 + N_2)^2 \times 4}, \quad (7)$$

where \odot denotes the extended Khatri-Rao product. This augmented manifold forms a difference co-array with $(N_1 + N_2)^2$ total positions and $N_o = 4$ polarization slices. Through spatial smoothing, we obtain $\tilde{N} = N_1(N_2 + 1)$ consecutive virtual sensors, enabling detection of up to $\tilde{N} - 1$ sources.

2.3.1 Spatial smoothing and virtual array construction

The augmented tensor manifold exhibits redundancy from repeated spatial separations between sensor pairs. We employ tensor-based spatial smoothing (Liu and Vaidyanathan, 2015) to partition the virtual array into overlapping subarrays and average their correlation structures, which improves signal subspace conditioning under finite samples, enables coherent source decorrelation, and reduces estimation variance. Liu and Vaidyanathan (2015) provided a comprehensive analysis of spatial smoothing effects in co-array processing, establishing theoretical foundations for our tensor-based extension.

We define $\tilde{N} \triangleq N_1(N_2 + 1)$ consecutive virtual sensors and construct $\tilde{\mathcal{A}} \in \mathbb{C}^{(2\tilde{N}-1) \times N_o \times K}$ with $N_o = 4$ polarization dimensions. After removing redundant rows, we obtain Equation 8:

$$\tilde{U} = \tilde{\mathcal{A}} \times_3 s + \sigma_e^2 \tilde{E}, \quad (8)$$

where $\tilde{E} \in \mathbb{R}^{(2\tilde{N}-1) \times N_o}$ contains all zeros except a single unit entry at position $(\tilde{N}, 1)$. The difference co-array for the proposed hybrid array has sensor positions spanning the range in Equation 9:

$$\{(-\tilde{N} + 1)d, \dots, -d, 0, d, \dots, (\tilde{N} - 1)d\}. \tag{9}$$

Partitioning into \tilde{N} overlapping subarrays, the l -th subarray is given by Equation 10:

$$\tilde{U}_l = \tilde{\mathcal{A}}_l \times_3 s + \sigma_e^2 \tilde{E}_l, \tag{10}$$

where the shift-invariant structure $\tilde{\mathcal{A}}_l = \tilde{\mathcal{A}}_1 \times_3 \Phi^{l-1}$ with phase shift matrix defined in Equation 11:

$$\Phi = \text{diag}(e^{-j\pi \sin \theta_1}, \dots, e^{-j\pi \sin \theta_K}). \tag{11}$$

The cross-correlation tensor for the l th subarray is expressed in Equation 12:

$$\begin{aligned} \mathcal{R}_l &= (\tilde{\mathcal{A}}_l \circ \tilde{\mathcal{A}}_l^*) \times_3 \Phi^{l-1} \times_3 \text{diag}(\sigma_1^2, \dots, \sigma_K^2) \times_3 (\Phi^{l-1})^* \\ &+ \mathcal{N}_l. \end{aligned} \tag{12}$$

Averaging over all \tilde{N} subarrays and exploiting $\frac{1}{\tilde{N}} \sum_{l=1}^{\tilde{N}} \Phi^{l-1} (\Phi^{l-1})^* = I_K$ yields Equation 13:

$$\mathcal{T} = \frac{1}{\tilde{N}} \sum_{l=1}^{\tilde{N}} \mathcal{R}_l = (\tilde{\mathcal{A}}_1 \circ \tilde{\mathcal{A}}_1^*) \times_3 R_s + \text{noise}, \tag{13}$$

where $R_s = \text{diag}(\sigma_1^2, \dots, \sigma_K^2)$. The resulting $\mathcal{T} \in \mathbb{C}^{\tilde{N} \times \tilde{N} \times N_o \times N_o}$ supports estimation of up to $\tilde{N} - 1$ sources through the Vandermonde structure.

2.3.2 HOSVD-based subspace estimation and DOA retrieval

We apply HOSVD to extract the signal subspace, as formulated in Equation 14:

$$\mathcal{T} = \mathcal{K} \times_1 W_1 \times_2 W_2 \times_3 W_3 \times_4 W_4, \tag{14}$$

where $\mathcal{K} \in \mathbb{C}^{\tilde{N} \times \tilde{N} \times 4 \times 4}$ is the core tensor and W_i are orthonormal factor matrices. Hermitian symmetry yields $W_3 = W_1^*$ and $W_4 = W_2^*$. The signal subspace is defined in Equation 15:

$$U_s = W_1(:, 1:\hat{K}), \tag{15}$$

selecting the \hat{K} columns corresponding to dominant singular values.

Source enumeration uses the Minimum Description Length (MDL) criterion applied to core tensor energies:

$$\hat{K} = \arg \min_k \left[-\log \left(\prod_{i=k+1}^{\tilde{N}} \lambda_i \right) + \frac{1}{2} k(2\tilde{N} - k) \log N_s \right], \tag{16}$$

where $\lambda_i = \|\mathcal{K}(i, i, :, :)\|_F^2$ are ordered core tensor energies.

DOA estimation exploits rotational invariance, as expressed in Equation 17. With selection matrices $J_1, J_2 \in \mathbb{R}^{(\tilde{N}-1) \times \tilde{N}}$:

$$\Phi = (J_1 U_s)^\dagger (J_2 U_s), \quad \theta_k = \arcsin \left(-\frac{\angle(\phi_{k,k})}{\pi} \right) \tag{17}$$

where $\phi_{k,k}$ are diagonal elements of Φ . The azimuth angle ϕ_k is estimated by leveraging the polarization diversity encoded in the factor matrix W_2 . Specifically, since the polarization response vector contains directional information as $u_k = [\cos \phi_k \cos \theta_k, \sin \phi_k \cos \theta_k,$

TABLE 2 Algorithm parameters and physical significance.

Parameter	Value	Physical significance
\mathcal{T}	$\tilde{N} \times \tilde{N} \times 4 \times 4$	Spatially smoothed interspectral tensor
\mathcal{K}	$\tilde{N} \times \tilde{N} \times 4 \times 4$	Core tensor containing signal energy
λ_i	Descending order	Distribution normalized energy of virtual array modes
\hat{K}	$1 \leq \hat{K} \leq \tilde{N} - 1$	Estimated number of sources
U_s	$\tilde{N} \times \hat{K}$	Signal subspace matrix
J_1, J_2	Selection matrices	Capture translational invariance in spatial dimensions
ϕ_k	$[0, 360^\circ]$	Azimuth angle of k th source
θ_k	$[0, 90^\circ]$	Elevation angle of k th source

$\sin \theta_k]^T$, the azimuth can be recovered from the ratio of horizontal velocity components, as shown in Equation 18:

$$\phi_k = \arctan \left(\frac{[W_2(:, k)]_2}{[W_2(:, k)]_1} \right), \tag{18}$$

where $[W_2(:, k)]_1$ and $[W_2(:, k)]_2$ denote the first and second elements of the k th column of the polarization factor matrix, corresponding to the v_x and v_y velocity components, respectively. The complete algorithm is summarized in Algorithm 1. The algorithm parameters and their physical significance are provided in Table 2.

```

Require:  $\mathcal{T} \in \mathbb{C}^{\tilde{N} \times \tilde{N} \times 4 \times 4}$ , noise covariance  $\Sigma_e$ 
Ensure:  $\hat{K}, \{(\hat{\phi}_k, \hat{\theta}_k)\}_{k=1}^{\hat{K}}$ 
1: Pre-whitening: Apply  $\tilde{\mathcal{T}} = \mathcal{T} \times_3 \Sigma_e^{-1/2} \times_4 \Sigma_e^{-H/2}$  for non-uniform noise
2: Compute HOSVD:  $\tilde{\mathcal{T}} = \mathcal{K} \times_1 W_1 \times_2 W_2 \times_3 W_3 \times_4 W_4$ 
3: Calculate core tensor energies  $\lambda_i = \|\mathcal{K}(i, i, :, :)\|_F^2$ 
4: Estimate source number:  $\hat{K} = \arg \min_k \text{MDL}(k)$ 
5: Extract signal subspace:  $U_s = W_1(:, 1:\hat{K})$ 
6: Construct selection matrices  $J_1, J_2 \in \mathbb{R}^{(\tilde{N}-1) \times \tilde{N}}$ 
7: Solve eigenvalue problem:  $\Phi = (J_1 U_s)^\dagger (J_2 U_s)$ 
8: for  $k = 1$  to  $\hat{K}$  do
9: Extract eigenvalue:  $\phi_{k,k} = \Phi(k, k)$ 
10: Compute elevation angle:  $\theta_k = \arcsin \left( -\frac{\angle(\phi_{k,k})}{\pi} \right)$ 
11: Compute azimuth angle from polarization:  $\phi_k = \arctan \left( \frac{[W_2(:, k)]_2}{[W_2(:, k)]_1} \right)$ 
12: end for
13: return  $\hat{K}, \{(\phi_k, \theta_k)\}_{k=1}^{\hat{K}}$ 
    
```

Algorithm 1. HOSVD-based source enumeration and DOA estimation.

3 Performance analysis

This section provides theoretical performance analysis of the proposed framework. We derive the CRB, analyze computational complexity, and examine parameter identifiability.

TABLE 3 Computational complexity comparison for six DOA estimation methods.

Method	Complexity order
HSVSA-Tensor-HOSVD (proposed)	$\mathcal{O}(\tilde{N}^3 + \hat{K}^3)$
ULA-Tensor-HOSVD	$\mathcal{O}(\tilde{N}^3 + \hat{K}^3)$
PARAFAC-DOA	$\mathcal{O}(I \cdot \tilde{N}^3 K^2)$
Vector-MUSIC	$\mathcal{O}(\tilde{N}^3 + \tilde{N}^2 G)$
HSVSA-Matrix-MUSIC	$\mathcal{O}(\tilde{N}^3 + \tilde{N}^2 G)$
Nested-MUSIC	$\mathcal{O}(\tilde{N}^3 + \tilde{N}^2 G)$

\tilde{N} : virtual sensors; \hat{K} : estimated sources; G : grid points; I : PARAFAC iterations.

3.1 Cramér–Rao bound analysis

The CRB establishes the fundamental performance limit for unbiased estimators. For DOA estimation with the proposed HSVSA, we derive the CRB to quantify theoretical estimation accuracy.

Consider the parameter vector $\xi = [\theta_1, \varphi_1, \dots, \theta_K, \varphi_K, \sigma_1^2, \dots, \sigma_K^2]^T \in \mathbb{R}^{3K}$. The Fisher Information Matrix (FIM) is given by Equation 19:

$$F(\xi) = \frac{2N_s}{\sigma_e^2} \text{Re} \left\{ \left(\frac{\partial A}{\partial \xi} \right)^H \Pi_A^\perp \left(\frac{\partial A}{\partial \xi} \right) \odot R_x^T \right\} \quad (19)$$

where $\Pi_A^\perp = I - A(A^H A)^{-1} A^H$ is the projection matrix orthogonal to A . The CRB for the k th source is expressed in Equation 20:

$$\text{CRB}(\theta_k, \varphi_k) = \sqrt{[F^{-1}]_{2k-1, 2k-1} + [F^{-1}]_{2k, 2k}} \quad (20)$$

The hybrid geometry provides substantial CRB reduction compared to uniform arrays. The FIM scales as $F_{\text{HSVSA}} \approx (\tilde{N}/N)^2 F_{\text{ULA}}$, yielding CRB reduction of $(\tilde{N}/N)^2 = 7.84$ for the proposed configuration. This advantage stems from the extended virtual aperture $(\tilde{N} - 1)d$ achieved through difference co-array processing of heterogeneous sensor measurements. Recent advances in CRB analysis for sparse arrays (Bai et al., 2022) have provided theoretical insights into performance bounds under Bayesian frameworks. Combined with tensor processing, the proposed framework achieves near-CRB performance, validated in Section 4.3.

3.2 Computational complexity

The computational complexity of the proposed HSVSA-Tensor-HOSVD algorithm is analyzed by examining the dominant operations. For spatial smoothing, constructing the spatially smoothed tensor \mathcal{T} from \tilde{N} subarrays requires $\mathcal{O}(\tilde{N}^3)$ operations. The HOSVD decomposition involves mode-1 and mode-2 unfolding (each of size $\tilde{N} \times 16\tilde{N}$) followed by SVD, yielding complexity $\mathcal{O}(\tilde{N}^3)$ for each mode. The polarization modes (mode-3 and mode-4) contribute $\mathcal{O}(\tilde{N}^2)$ each. Computing the core tensor energy $\lambda_i = \|\mathcal{K}(i, i, :, :)\|_F^2$ for all $i = 1, \dots, \tilde{N}$ requires $\mathcal{O}(\tilde{N})$ operations. The MDL criterion evaluation contributes $\mathcal{O}(\tilde{N})$ operations. The ESPRIT-type DOA estimation, including pseudoinverse and eigenvalue decomposition, contributes $\mathcal{O}(\tilde{N}^2 \hat{K} + \hat{K}^3)$ operations.

The overall computational complexity is dominated by spatial smoothing and HOSVD, as summarized in Equation 21:

$$C_{\text{total}} = \mathcal{O}(\tilde{N}^3 + \hat{K}^3), \quad (21)$$

which is significantly lower than MUSIC-based methods requiring $\mathcal{O}(\tilde{N}^3 + \tilde{N}^2 G)$ with spectral grid search ($G \approx 3 \times 10^4$ for practical resolution), and comparable to iterative PARAFAC requiring $\mathcal{O}(I \cdot \tilde{N}^3 K^2)$ with $I = 10$ -15 iterations. The proposed closed-form approach avoids iterative optimization while maintaining lower complexity than grid-based methods.

Table 3 compares the computational complexity orders for all six benchmark methods. Tensor-based methods (HSVSA-Tensor-HOSVD, ULA-Tensor-HOSVD) achieve $\mathcal{O}(\tilde{N}^3 + \hat{K}^3)$ complexity through efficient closed-form HOSVD decomposition. Matrix-based MUSIC methods (Vector-MUSIC, HSVSA-Matrix-MUSIC, Nested-MUSIC) incur $\mathcal{O}(\tilde{N}^3 + \tilde{N}^2 G)$ complexity, where the spectral search term $\mathcal{O}(\tilde{N}^2 G)$ dominates with $G \approx 32,400$ grid points (for 1 resolution over 180×180), resulting in significantly higher computational cost than tensor methods. PARAFAC-DOA exhibits $\mathcal{O}(I \cdot \tilde{N}^3 K^2)$ complexity with $I = 10$ -15 iterations and initialization sensitivity. The proposed method avoids iterative optimization and exhaustive grid search, achieving lower complexity with closed-form subspace estimation. Experimental validation of these theoretical predictions through actual runtime measurements is presented in Section 4.3.

3.3 Parameter identifiability and degrees of freedom

Parameter identifiability determines the theoretical maximum number of sources that can be uniquely resolved, independent of SNR or snapshot count. For the proposed framework, identifiability is governed by tensor decomposition uniqueness and virtual array DOF.

The spatially smoothed tensor $\mathcal{T} \in \mathbb{C}^{\tilde{N} \times \tilde{N} \times N_o \times N_o}$ admits the factorization in Equation 22:

$$\mathcal{T} = (\tilde{\mathcal{A}}_1 \circ \tilde{\mathcal{A}}_1^*) \times_3 R_s, \quad (22)$$

where $\tilde{\mathcal{A}}_1 \in \mathbb{C}^{\tilde{N} \times N_o \times K}$ contains steering vectors and N_o represents polarization components. The uniqueness condition requires Equation 23 to hold:

$$k - \text{rank}(\tilde{\mathcal{A}}_1) + k - \text{rank}(\tilde{\mathcal{A}}_1^*) \geq 2K + 2. \quad (23)$$

In practical scenarios, this uniqueness condition may be violated due to source coherence from multipath propagation, insufficient snapshots ($N_s \lesssim K$), or closely spaced sources with nearly dependent steering vectors. Spatial smoothing partially mitigates coherence effects, but severe multipath conditions may still degrade identifiability. For Vandermonde manifolds with \tilde{N} virtual sensors and N_o polarization components, the Kruskal rank is $\min\{\tilde{N}, N_o\}$, yielding the bound in Equation 24:

$$K_{\text{max}} \leq \min\{\tilde{N} - 1, \lfloor N_o/2 \rfloor \cdot \tilde{N}\}. \quad (24)$$

The HSVSA generates $\tilde{N} = N_1(N_2 + 1)$ consecutive virtual sensors, supporting $K_{\text{max}} = \tilde{N} - 1$ resolvable sources—substantially more than conventional uniform arrays with $N = N_1 + N_2$ sensors providing $K_{\text{max}}^{\text{ULA}} = N - 1$ DOF. Compared to

TABLE 4 Benchmark methods for performance comparison.

Method	Array type	Sensors	DOF	Algorithm
HSVSA-Tensor-HOSVD	Hybrid	10	28	Tensor-HOSVD
ULA-Tensor-HOSVD	Uniform	28	28	Tensor-HOSVD
PARAFAC-DOA	Hybrid	10	28	PARAFAC
Vector-MUSIC	Uniform	10	10	Matrix-MUSIC
HSVSA-Matrix-MUSIC	Hybrid	10	28	Matrix-MUSIC
Nested-MUSIC	Nested scalar	10	28	Matrix-MUSIC

nested scalar arrays achieving $\tilde{N}_{\text{scalar}}$ virtual sensors, the hybrid architecture adds polarization diversity for enhanced accuracy.

When snapshots become limited ($N_s \approx \tilde{N}$), sample covariance introduces the constraint in Equation 25:

$$K_{\text{max}}^{\text{eff}} = \min \{ \tilde{N} - 1, N_s - 1 \} \quad (25)$$

For underwater applications with sufficient snapshots ($N_s \gg N_e$), this constraint is inactive.

While the Kruskal rank condition and snapshot constraint above establish theoretical identifiability limits $K_{\text{max}} \leq \tilde{N} - 1$, practical scalability depends on estimation performance near these limits. As K approaches K_{max} , reduced eigenvalue gaps in the covariance structure degrade signal-noise subspace separation. Tensor processing maintains robustness through joint diagonalization exploiting spatial-polarization coupling, whereas matrix vectorization discards this multidimensional structure and suffers earlier degradation. Regarding array scaling, the $O(\tilde{N}^3)$ HOSVD complexity exhibits polynomial growth suitable for moderately large virtual apertures, avoiding the exponential scaling of exhaustive grid-based methods. Combined with the quadratic DOF growth $\tilde{N} = O(N^2)$ from co-array processing, this establishes favorable scalability characteristics for dense acoustic environments.

4 Experimental validation

This section presents comprehensive experimental validation of the proposed HSVSA-Tensor-HOSVD framework through systematic Monte Carlo simulations. We compare against five established baseline methods under both controlled and adverse conditions to demonstrate practical applicability.

The HSVSA employs $N_1 = 4$ vector sensors and $N_2 = 6$ scalar sensors, generating $\tilde{N} = N_1(N_2 + 1) = 28$ virtual sensors through co-array processing. Sources are modeled as uncorrelated narrowband signals with additive white Gaussian noise. Table 4 summarizes all benchmark methods evaluated under identical experimental conditions.

Performance evaluation employs complementary metrics tailored to specific analysis objectives. For source enumeration, detection probability P_d quantifies the fraction of trials achieving correct source count estimation ($\hat{K} = K$). DOA estimation accuracy is characterized by root mean square error (RMSE), while angular resolution is assessed through 3 dB beamwidth measurements from spatial spectrum cuts. Additionally, computational efficiency is evaluated through execution time measurements, and robustness is quantified by RMSE degradation under non-uniform noise conditions. Unless otherwise specified, experiments employ $M = 500$ Monte Carlo trials with source DOAs uniformly distributed within $\theta \in [10^\circ, 80^\circ]$ and $\varphi \in [0^\circ, 360^\circ]$.

4.1 Performance analysis

This subsection presents comprehensive simulation results examining source enumeration accuracy, angular resolution characteristics, noise tolerance, and data utilization efficiency.

4.1.1 Source enumeration under limited snapshots

Reliable source counting is essential for subsequent DOA estimation. We apply the MDL criterion defined in (Equation 16) to HOSVD core tensor energies under a challenging limited-snapshot scenario ($N_s = 200$) where covariance matrix estimation becomes ill-conditioned. Tests cover $K \in \{2, 4, 6, 8\}$ sources uniformly distributed within $\theta \in [10^\circ, 80^\circ]$ and $\varphi \in [0^\circ, 360^\circ]$.

As can be seen from Figure 2, a distinct performance hierarchy emerges among the evaluated methods. For $K = 2$ sources (Figure 2a), HSVSA-Tensor-HOSVD achieves $P_d > 0.95$ at SNR = -5 dB, while matrix-based methods require SNR ≥ 5 dB. As the source count increases to $K = 4$ and $K = 6$ (Figures 2b, c), tensor methods maintain their 5–10 dB advantage over matrix alternatives. Even when approaching the array's theoretical DOF limit at $K = 8$ (Figure 2d), the proposed method maintains $P_d > 0.90$ for SNR ≥ 5 dB, whereas conventional approaches require SNR ≥ 10 dB to achieve comparable reliability.

4.1.2 Resolution of closely spaced sources

The ability to distinguish nearby sources determines practical array utility in cluttered underwater environments. We position two sources at $(\theta_1, \varphi_1) = (30^\circ, 45^\circ)$ and $(\theta_2, \varphi_2) = (35^\circ, 50^\circ)$ with only 5° angular separation—below the Rayleigh limit for most conventional arrays. Spatial spectra span $\theta \in [0^\circ, 60^\circ]$ and $\varphi \in [20^\circ, 80^\circ]$ at 0.5° resolution.

The normalized 2D spatial spectra reveal distinct resolution characteristics across the evaluated methods. As illustrated in Figure 3, HSVSA-Tensor-HOSVD (Figure 3c) produces two well-separated, sharply defined peaks with minimal sidelobe contamination, while matrix-based approaches (Figures 3a, b, f) yield substantially broader main lobes that nearly merge under the 5° separation condition. ULA-Tensor-HOSVD (Figure 3e) achieves the sharpest resolution owing to its full 28-sensor physical aperture,

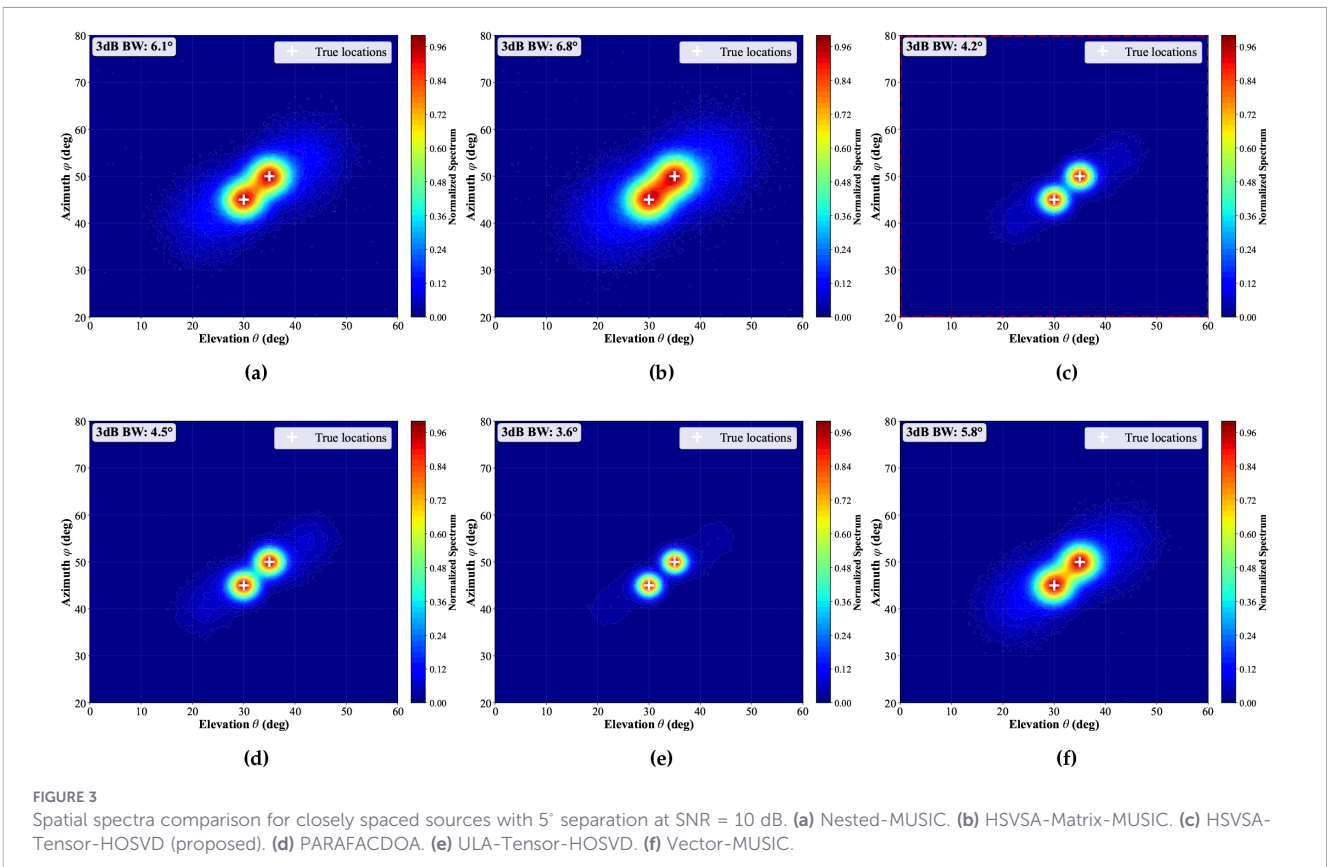
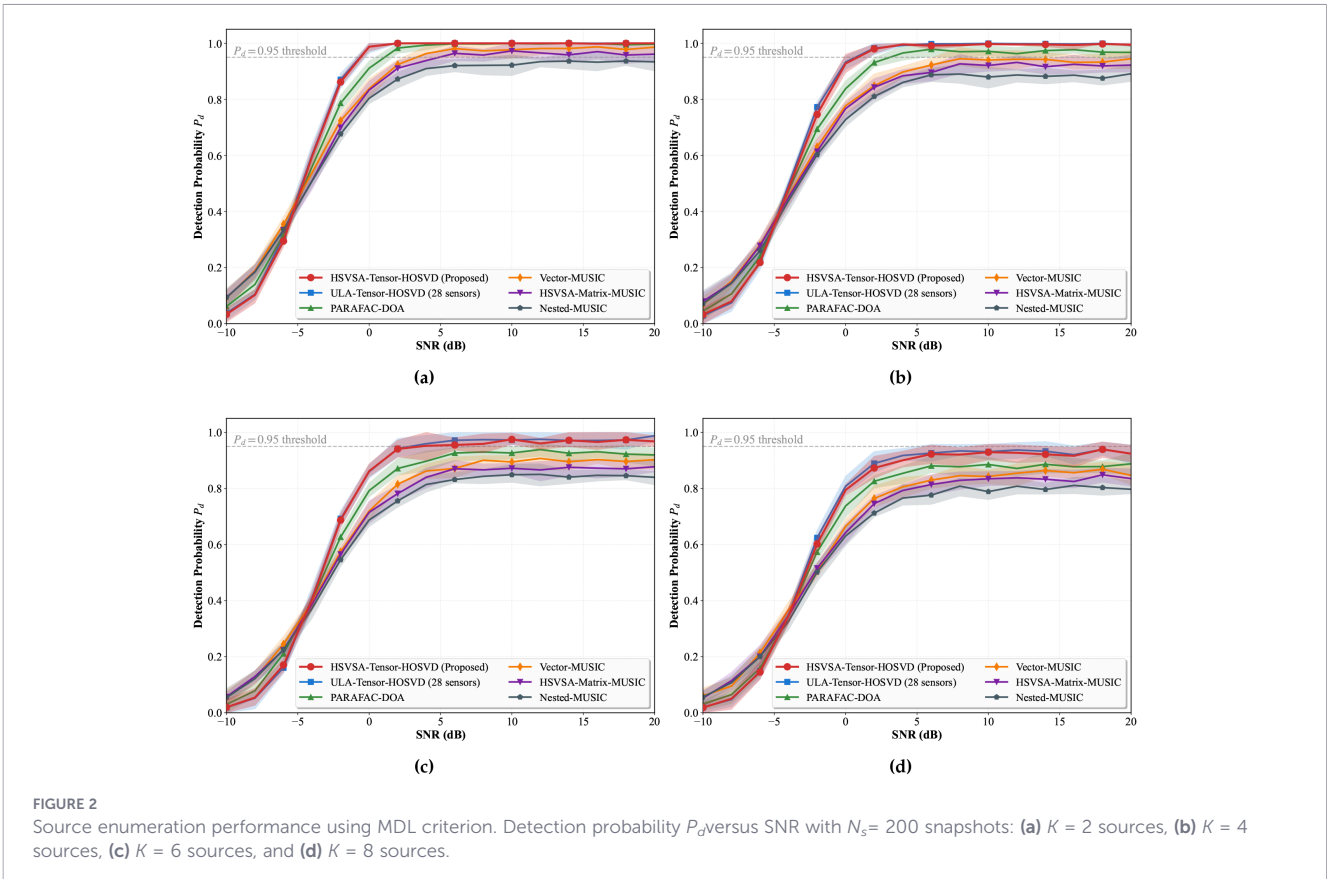


TABLE 5 Angular resolution comparison: 3 dB beamwidth for 5' source separation at SNR = 10 dB.

Method	3 dB beamwidth	Relative width
HSVSA-Tensor-HOSVD (proposed)	4.2°	100% (baseline)
ULA-Tensor-HOSVD	3.65°	87%
PARAFAC-DOA	4.5°	107%
Vector-MUSIC	5.8°	138%
HSVSA-Matrix-MUSIC	6.8°	162%
Nested-MUSIC	6.1°	145%

establishing an upper performance bound. These visual observations are quantified in Table 5: the proposed method achieves 4.2° 3 dB beamwidth—only 15% wider than the 28-sensor ULA benchmark while requiring 2.8× fewer physical sensors. This confirms that the proposed heterogeneous array design effectively exploits difference co-array processing to extend spatial resolution beyond what the 10 physical sensors would otherwise provide.

4.1.3 Estimation accuracy across SNR conditions

Underwater acoustic environments span a wide dynamic range. We characterize estimation accuracy from severe noise conditions (SNR = -15 dB) through moderate (0 dB) to favorable scenarios (15 dB), using $K = 4$ sources with $N_s = 500$ snapshots to ensure statistically stable covariance estimates.

The results in Figure 4 validate the estimation accuracy of the proposed HSVSA-Tensor-HOSVD across the full SNR range. At moderate SNR = 10 dB—typical of practical scenarios—the proposed method achieves approximately 0.27 RMSE, demonstrating effective exploitation of the hybrid array geometry with 2× to 3× lower error than matrix-based alternatives. The method maintains sub-10 estimation accuracy even under severe noise conditions (SNR = -15 dB), extending the operational envelope for underwater platforms.

4.1.4 Convergence behavior under limited data

Rapid DOA estimation with minimal observation time is critical for tracking maneuvering targets or operating in dynamic environments. We evaluate convergence from severely limited ($N_s = 50$) to adequate ($N_s = 1,000$) snapshot counts at SNR = 5 dB.

The data efficiency advantage of tensor processing is evident from the convergence analysis. As shown in Figure 5, with only $N_s = 50$ snapshots—corresponding to just 100 ms at typical 500 Hz sampling rates—HSVSA-Tensor-HOSVD already achieves RMSE < 0.5. Matrix-based methods require four times as many snapshots ($N_s \geq 200$) to reach comparable accuracy, representing an 86.7% reduction in required observation time. This efficiency gain stems from the tensor decomposition’s ability to jointly exploit all array dimensions simultaneously, effectively multiplying the statistical leverage of each observation.

4.2 Robustness validation

AVSs inherently exhibit spatially non-uniform noise characteristics due to fundamental physical differences between pressure and velocity transduction mechanisms (Nehorai and Paldi, 1994). Specifically, velocity channels suffer from higher self-noise due to mechanical suspension systems, increased susceptibility to flow-induced turbulence, and different frequency-dependent response characteristics compared to pressure hydrophones. This intrinsic noise non-uniformity—typically 3–6 dB higher in velocity channels—represents a fundamental challenge for vector sensor array processing. This subsection validates the proposed tensor framework’s robustness against this inherent sensor characteristic.

4.2.1 Non-uniform noise robustness

We systematically evaluate algorithm performance under realistic noise non-uniformity levels documented in AVS literature: uniform baseline, 3 dB, 5 dB, and 6 dB velocity-to-pressure noise power ratios. Following Equation 3, the noise

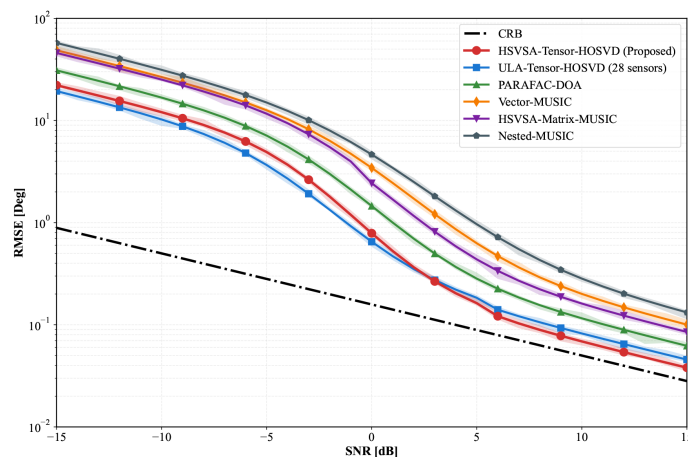


FIGURE 4 RMSE versus SNR comparison for $K = 4$ sources with $N_s = 500$ snapshots.

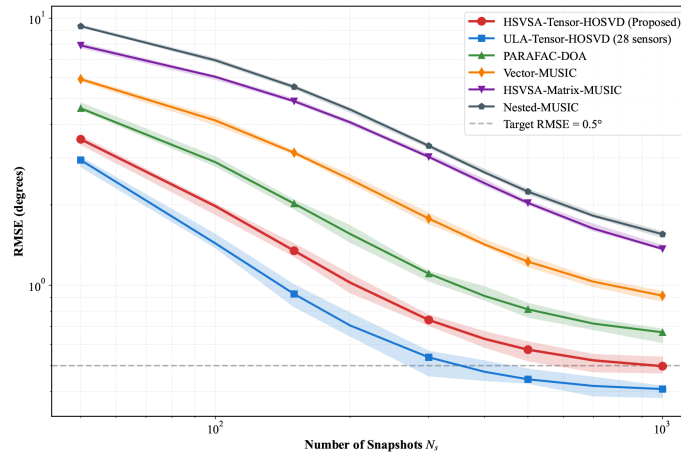


FIGURE 5 RMSE convergence versus snapshot count for $K = 4$ sources at SNR = 5 dB.

covariance is modeled as $\Sigma_e = \text{diag}(\sigma_p^2, \alpha\sigma_p^2, \alpha\sigma_p^2, \alpha\sigma_p^2) \otimes I_{N_s}$, where $\alpha \in \{1.0, 2.0, 3.16, 4.0\}$ corresponds to $\{0, 3, 5, 6\}$ dB differentials. Experiments use SNR = 5 dB with $N_s = 500$ snapshots.

The multidimensional performance comparison, as depicted in Figure 6, reveals the comprehensive advantages of the proposed method. The radar chart (Figure 6a) shows HSVSA-Tensor-HOSVD maintaining the largest coverage area across accuracy, robustness, efficiency, and adaptability metrics, while the RMSE degradation heatmap (Figure 6b) quantifies the critical finding: tensor methods exhibit inherent resilience to channel-specific noise variations.

The robustness advantage becomes quantitatively clear from Table 6. Under severe 6-dB velocity-to-pressure noise differential—representative of practical AVS deployments—the proposed method exhibits only 9.7% RMSE degradation relative to uniform noise baseline.

4.2.2 Compound stress scenarios with field-measured background noise

Practical underwater deployments rarely face isolated challenges. To validate performance under realistic compound stress conditions, we incorporate background noise characteristics measured from field experiments conducted in coastal waters (depth: 15–30 m, sea state: 3–4). The field-measured noise exhibits the following: (1) colored spectral characteristics with 6 dB roll-off above 2 kHz; (2) 4.5 dB average velocity-to-pressure noise differential; and (3) intermittent impulsive interference from shipping traffic.

We evaluate compound scenarios combining these field-measured characteristics with controlled degradation factors, with results presented in Table 7 for all six benchmark methods under three representative stress conditions:

The results in Table 7 demonstrate consistent performance advantages of the proposed method under compound stress

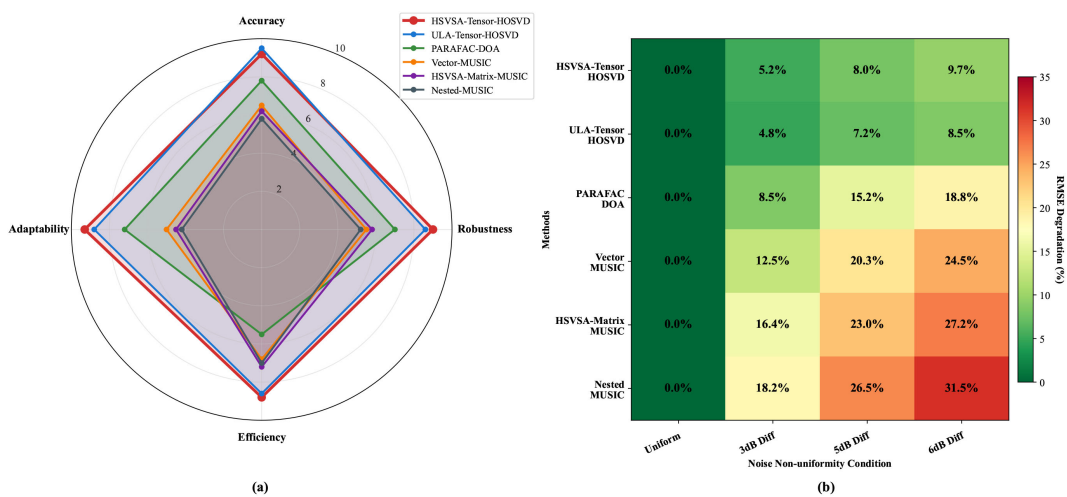


FIGURE 6 Non-uniform noise robustness evaluation for all six methods with SNR = 5 dB and $N_s = 500$ snapshots. (a) Radar chart comparing accuracy, robustness, efficiency, and adaptability metrics. (b) Heatmap of RMSE degradation under uniform, 3 dB, 5 dB, and 6 dB pressure-velocity channel differences.

TABLE 6 RMSE degradation under non-uniform noise conditions.

Method	3 dB	5 dB	6 dB	Avg. degradation
HSVSA-Tensor-HOSVD	5.2%	8.0%	9.7%	7.6%
ULA-Tensor-HOSVD	4.8%	7.2%	8.5%	6.8%
PARAFAC-DOA	8.5%	15.2%	18.8%	14.2%
Vector-MUSIC	12.5%	20.3%	24.5%	19.1%
HSVSA-Matrix-MUSIC	16.4%	23.0%	27.2%	22.2%
Nested-MUSIC	18.5%	26.8%	31.5%	25.6%

scenarios incorporating field-measured background noise characteristics. Across all tested conditions, HSVSA-Tensor-HOSVD maintains RMSE below 2°, establishing reliable operational boundaries for practical deployments. Notably, under the most challenging scenario with field-measured colored noise (6 dB roll-off above 2 kHz) and 4.5 dB velocity-to-pressure differential, the proposed method achieves 1.58 RMSE. In contrast, matrix-based alternatives require favorable SNR (≥ 5 dB) or extended observation windows ($N_s \geq 300$) to achieve comparable accuracy. These findings demonstrate that the proposed framework can accommodate realistic environmental disturbances inherent to AVSs while maintaining estimation accuracy suitable for autonomous underwater vehicle navigation and target tracking applications.

4.3 Validation of theoretical predictions

Having established the theoretical performance bounds in Section 3, we now validate these predictions through comprehensive experiments, focusing on proximity to the CRB and computational efficiency.

4.3.1 CRB proximity and estimation accuracy

To quantify proximity to theoretical optimality, we conduct CRB analysis across varying SNR regimes from -10 to 15 dB with $K = 4$ sources and $N_s = 500$ snapshots. Table 8 presents experimental RMSE versus theoretical lower bounds, revealing the gap between achievable performance and fundamental limits.

The comparative analysis reveals two key insights beyond individual method validation in Section 4.2. First, the proposed HSVSA-Tensor-HOSVD consistently achieves 2.1× to 3.0× lower RMSE than matrix-based approaches across all SNR conditions, demonstrating that tensor processing provides systematic advantages over conventional vectorization regardless of noise

level. Second, the CRB gap ratio between tensor and matrix methods remains stable at 2× to 3×, empirically confirming that the accuracy advantage stems from algorithmic design rather than favorable experimental conditions.

4.3.2 Computational efficiency verification

While CRB analysis validates estimation accuracy bounds, practical deployability equally depends on computational tractability for embedded implementation. To empirically validate the claimed 4.9× to 6.6× computational speedup, we conduct runtime profiling on a representative computing platform—Intel Corei7-10700 @ 2.9 GHz running MATLAB R2022b. Experiments employ the baseline configuration: $\tilde{N} = 28$ virtual sensors, $K = 4$ sources, and $N_s = 500$ snapshots, with MUSIC methods evaluated at $G = 32,400$ grid points (1 angular resolution) and PARAFAC-DOA configured with $I = 12$ alternating least squares iterations.

The profiling results, visualized in Figure 7, empirically confirm the theoretical efficiency predictions with 4.9× to 6.6× speedup over Matrix-MUSIC methods and 4.2× acceleration over iterative PARAFAC-DOA. This substantial computational advantage originates from two key algorithmic design choices: (1) closed-form HOSVD decomposition replaces iterative optimization inherent to PARAFAC’s alternating least squares, eliminating convergence overhead and initialization sensitivity; (2) direct subspace rotation via ESPRIT extracts DOAs algebraically, circumventing the exhaustive spectral grid search that dominates MUSIC runtime complexity. The resulting throughput enables deployment on embedded underwater platforms, satisfying stringent latency requirements for autonomous navigation and dynamic multi-target tracking.

To provide deeper insight into the convergence characteristics underlying these computational advantages, Figure 8 presents RMSE evolution versus iteration count for PARAFAC-DOA at the challenging $K = 8$ scenario approaching the array’s DOF limit. This analysis follows the methodology established in recent MIMO DOA studies (Anughna and Ramesha, 2023, 2024), where RMSE convergence plots effectively demonstrate algorithmic efficiency. The results reveal that PARAFAC-DOA requires 10–23 ALS iterations to achieve convergence depending on SNR conditions, with lower SNR demanding substantially more iterations due to degraded subspace quality during random initialization. In contrast, the proposed HSVSA Tensor-HOSVD achieves equivalent or superior accuracy through single-step HOSVD decomposition, as indicated by the horizontal reference lines representing closed-form estimation results. This iteration free characteristic eliminates convergence uncertainty and initialization sensitivity inherent to

TABLE 7 Compound stress scenario performance (RMSE in degrees) with field-measured background noise.

Scenario	HSVSA-T	ULA-T	PARAFAC	V-MUSIC	HSVSA-M	Nested
Low SNR + non-uniform	1.25	1.08	1.85	2.92	3.45	3.78
Limited snapshots	0.82	0.71	1.35	1.96	2.28	2.65
Field-measured noise	1.58	1.32	2.15	3.25	4.12	4.55

HSVSA-T, HSVSA-Tensor-HOSVD; ULA-T, ULA-Tensor-HOSVD; HSVSA-M, HSVSA-Matrix-MUSIC; Nested, Nested-MUSIC. Conditions, Low SNR (0 dB) + 5 dB non-uniform noise; Limited snapshots ($N_s = 100$) + 3 dB non-uniform noise; Colored field noise (6 dB roll-off above 2 kHz, 4.5 dB velocity–pressure differential).

TABLE 8 Cramér–Rao bound proximity analysis: experimental RMSE versus theoretical limits.

SNR [dB]	CRB [°]	Proposed [°]	Matrix [°]	CRB gap	Ratio
-10	5.80	8.82	18.5	152%	2.1×
-5	1.84	3.52	9.8	191%	2.8×
0	0.58	0.95	2.85	164%	3.0×
5	0.18	0.38	1.02	211%	2.7×
10	0.058	0.27	0.60	366%	2.2×
15	0.018	0.085	0.22	372%	2.6×

CRB Gap = (Proposed RMSE/CRB - 1); Ratio = Matrix RMSE/Proposed RMSE; Matrix = average of HSVSA-Matrix-MUSIC and Nested-MUSIC. Proposed: HSVSA-Tensor-HOSVD.

alternating optimization, providing deterministic execution time critical for real-time underwater applications.

5 Conclusions

This paper proposes a novel tensor-based DOA estimation framework called HSVSA-Tensor-HOSVD for underwater acoustic localization. By strategically integrating hybrid array geometry, polarization diversity, and multidimensional tensor processing, the proposed framework achieves enhanced estimation performance with significantly reduced hardware complexity. Comprehensive Monte

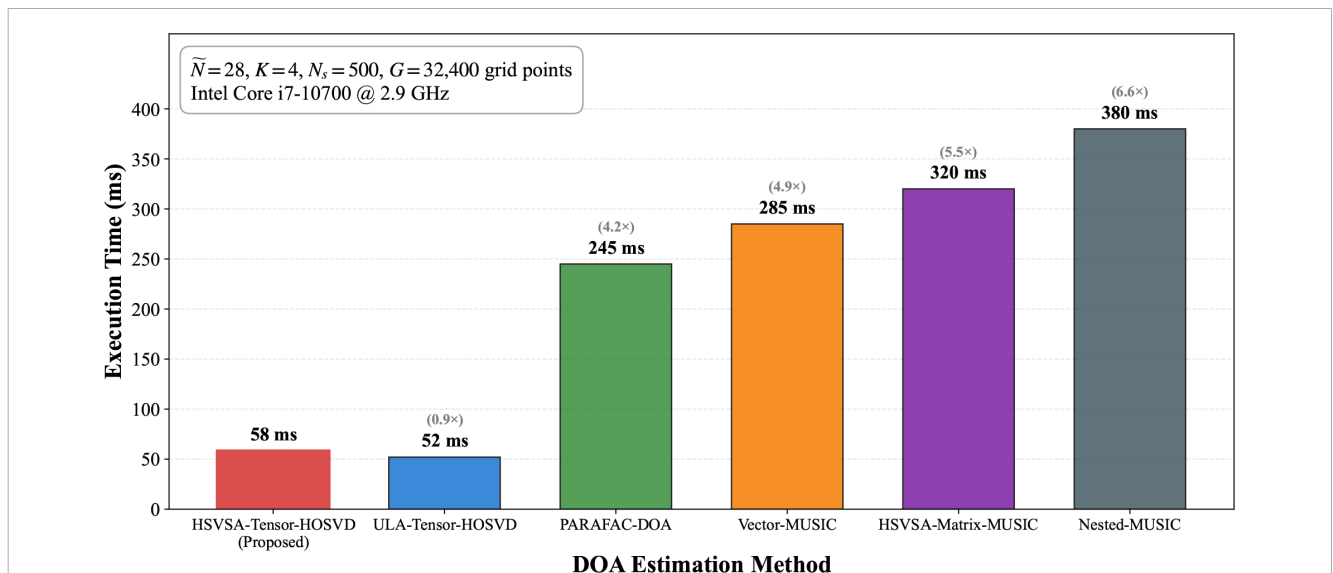


FIGURE 7 Computational efficiency comparison with data labels indicating execution time (ms) and speedup ratios relative to the proposed baseline.

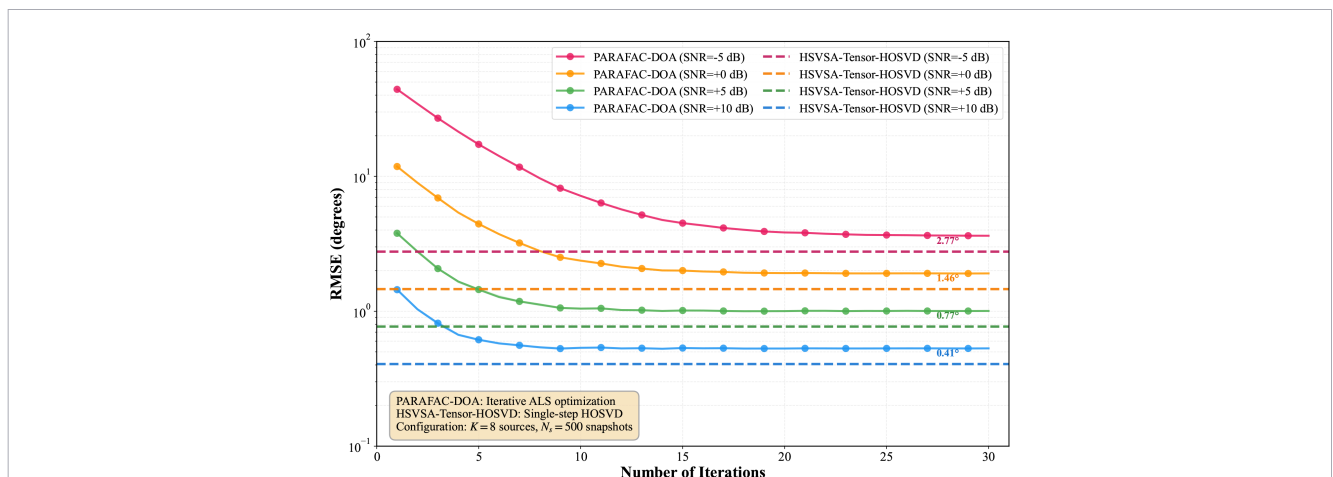


FIGURE 8 RMSE convergence comparison: PARAFAC-DOA iterative optimization versus HSVSA-Tensor-HOSVD closed-form estimation at K = 8 sources.

Carlo simulations demonstrate that HSVSA-Tensor-HOSVD provides superior source enumeration accuracy, precise angular resolution, robust noise tolerance, and computational efficiency suitable for real-time underwater applications. Future work will focus on field experimental validation in actual underwater environments to further demonstrate practical applicability.

To summarize, this paper mainly contributes to the following aspects:

1. We introduce an HSVSA architecture that achieves virtual sensors through difference co-array processing of heterogeneous measurements. The proposed HSVSA realizes enhanced DOF through sensor heterogeneity with uniform spacing, enabling substantial hardware cost reduction while maintaining enhanced spatial resolution capabilities.
2. We formulate a fourth-order spatially smoothed interspectral tensor model that jointly exploits spatial correlation and polarization structure within a unified mathematical framework. The formulation accommodates measurement dimensionality heterogeneity inherent to hybrid scalar–vector arrays, enabling effective exploitation of both spatial and polarization information.
3. We develop a closed-form HOSVD-based estimation algorithm that exploits rotational invariance for direct DOA retrieval. The proposed approach achieves significant computational speedup through deterministic subspace estimation, enabling deployment on embedded platforms with enhanced reliability and precision.

Data availability statement

The simulation data supporting the findings of this study are available from the corresponding author upon reasonable request.

Author contributions

YZ: Validation, Visualization, Project administration, Formal Analysis, Funding acquisition, Writing – review & editing, Writing – original draft, Data curation, Methodology, Investigation. XZ: Conceptualization, Funding acquisition, Writing – review & editing, Supervision. JY: Data curation, Software, Writing – review & editing, Resources.

References

- Abdelbadie, A., Mostafa, M., Bameri, S., Gohary, R. H., and Thomas, D. (2024). Doa estimation for hybrid receivers: Full spatial coverage and successive refinement. *IEEE Trans. Signal Process.* 72, 4730–4744. doi: 10.1109/TSP.2024.3459422
- Ahmed, T., Zhang, X., and Hassan, W. U. (2020). A higher-order propagator method for 2d-doa estimation in massive mimo systems. *IEEE Commun. Lett.* 24, 543–547. doi: 10.1109/LCOMM.2019.2960341
- Anughna, N., and Ramesha, M. (2023). Antenna reconfiguration based doa estimation for awgn channel in mimo applications. *Prog. Electromagn. Res. C* 128, 73–84. doi: 10.2528/PIERC22110404
- Anughna, N., and Ramesha, M. (2024). Multiplicative basis function based adaptive doa estimation in optimal mimo sparse antenna reconfiguration model. *IETE J. Res.* 70, 7219–7230. doi: 10.1080/03772063.2024.2351554

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was funded by the National Key Research and Development Program, grant numbers 2021YFC2803000 and 2021YFC2803001.

Acknowledgments

The authors thank the reviewers for their constructive comments.

Conflict of interest

Author JY was employed by the company Avic Shaanxi Aero Electric Co.,LTD.

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer XL declared a shared affiliation with the authors YZ and XZ to the handling editor at the time of review.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Bai, H., Duarte, M. F., and Janaswamy, R. (2022). Cramér-rao bounds for doa estimation of sparse bayesian learning with the laplace prior. *Sensors* 23, 307. doi: 10.3390/s23010307
- Chen, H., Lin, H., Liu, W., Wang, Q., Shen, Q., and Wang, G. (2024). Augmented multi-subarray dilated nested array with enhanced degrees of freedom and reduced mutual coupling. *IEEE Trans. Signal Process.* 72, 1387–1399. doi: 10.1109/TSP.2024.3374557
- Chen, X., Zhang, H., Gao, Y., and Wang, Z. (2023). Doa estimation of underwater acoustic co-frequency sources for the coprime vector sensor array. *Front. Mar. Sci.* 10, 1211234. doi: 10.3389/fmars.2023.1211234
- De Lathauwer, L., De Moor, B., and Vandewalle, J. (2000). A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.* 21, 1253–1278. doi: 10.1137/S0895479896305696
- Dong, H., Suo, J., Zhu, Z., and Li, S. (2024a). Improved underwater single-vector acoustic doa estimation via linear convolution preprocessing. *Electronics* 13, 1796. doi: 10.3390/electronics13091796
- Dong, H., Suo, J., Zhu, Z., Wang, H., and Ji, H. (2024b). Enhanced underwater single vector-acoustic doa estimation via linear matched stochastic resonance preprocessing. *Remote Sens.* 16, 1802. doi: 10.3390/rs16101802
- Haardt, M., Römer, F., and Del Galdo, G. (2008). Higher-order svd-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems. *IEEE Trans. Signal Process.* 56, 3198–3213. doi: 10.1109/TSP.2008.917929
- Han, K., and Nehorai, A. (2014). Nested vector-sensor array processing via tensor modeling. *IEEE Trans. Signal Process.* 62, 2542–2553. doi: 10.1109/TSP.2014.2314437
- Kolda, T., and Bader, B. (2009). Tensor decompositions and applications. *SIAM Rev.* 51, 455–500. doi: 10.1137/07070111X
- Li, K., Shen, Q., Liu, W., and Wang, M. (2024). Wideband doa estimation based on tensor completion and decomposition. In *Proceedings of the 2024 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5. doi: 10.1109/ISCAS58744.2024.10558253
- Li, M., and Zhang, C. (2023). “Improved nested array doa estimation based on sum difference virtual array,” in *2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)*. New York, USA: IEEE 2133–2137.
- Li, Y., and Wang, B. (2023). Direction-of-arrival estimation method based on neural network with temporal structure for underwater acoustic vector sensor array. *Sensors* 23, 4919. doi: 10.3390/s23104919
- Liu, C., and Vaidyanathan, P. (2015). Remarks on the spatial smoothing step in coarray music. *IEEE Signal Process. Lett.* 22, 1438–1442. doi: 10.1109/LSP.2015.2409153
- Nehorai, A., and Paldi, E. (1994). Acoustic vector-sensor array processing. *IEEE Trans. Signal Process.* 42, 2481–2491. doi: 10.1109/78.317869
- Nion, D., and Sidiropoulos, N. (2010). Tensor algebra and multidimensional harmonic retrieval in signal processing for mimo radar. *IEEE Trans. Signal Process.* 58, 5693–5705. doi: 10.1109/TSP.2010.2058802
- Pal, P., and Vaidyanathan, P. (2010). Nested arrays: A novel approach to array processing with enhanced degrees of freedom. *IEEE Trans. Signal Process.* 58, 4167–4181. doi: 10.1109/TSP.2010.2049264
- Qu, X., Lou, Y., Zhao, Y., Lu, Y., and Qiao, G. (2022). Augmented tensor music for doa estimation using nested acoustic vector-sensor array. *IEEE Signal Process. Lett.* 29, 1624–1628. doi: 10.1109/LSP.2022.3191254
- Schmidt, R. (1986). Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propagation* 34, 276–280. doi: 10.1109/TAP.1986.1143830
- Vaidyanathan, P., and Pal, P. (2010). Sparse sensing with co-prime samplers and arrays. *IEEE Transaction Signal Process.* 59, 573–586. doi: 10.1109/TSP.2010.2089682
- Van Trees, H. (2002). *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory* (New York, NY, USA: Wiley-Interscience).
- Xu, W., Yi, S., and Zhang, Z. (2025). Doa estimation using complex-valued neural networks with generative probability wave. *Circuits Systems Signal Process.*, 1–19. doi: 10.1007/s00034-025-03331-2
- Zhang, X., and Sun, H. (2023). Recent advances in underwater signal processing. *Sensors* 23, 5777. doi: 10.3390/s23135777
- Zhang, Y., and Zhang, X. (2024). “A study on multistatic sonar localization method with consideration of auv mobility,” in *2024 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. 1–4 (New York, USA: IEEE).
- Zhang, Y., Zhang, X., Yu, Y., Li, M., Ning, W., and Yang, J. (2022). “The positioning accuracy analysis of mobile multi-base sonar range information based on clusters of uavs,” in *2022 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. 1–6 (New York, USA: IEEE).
- Zheng, W., Zhang, X., Wang, Y., Shen, J., and Champagne, B. (2020). Padded coprime arrays for improved doa estimation: Exploiting hole representation and filling strategies. *IEEE Trans. Signal Process.* 68, 4597–4611. doi: 10.1109/TSP.2020.3013389
- Zhu, H., Zhao, H., Bao, C., Shi, Y., and He, W. (2025). Robust low-snapshot doa estimation for sparse arrays via a hybrid convolutional graph neural network. *Sensors* 25, 4563. doi: 10.3390/s25154563