Wadi Alshatti University Journal of Pure and Applied Sciences

مجلة جامعة وادي الشاطئ للعلوم البحتة والتطبيقية

Volume 4, No. 2, July-December 2026 Online ISSN: 3006-0877 المجلد 4، الاصدار 2، يوليو - ديسمبر 2026

RESEARCH ARTICLE **COMMUNICATION ENGINEERING**

Enhanced Hybrid Underwater Target Tracking: A Virtual Array MIMO SONAR Framework with CA-EKF-TOMHT

Saaida Elmahdi Alnour[1,*], Marai Abousetta[1], Hamza Mustafa Hasek[1]

[1] Electrical and Computer Department,School of Engineering and Applied Sciences, Libyan Academy, Tripoli , Libya.

ARTICLE HISTORY

Received 05 September 2026
Revised 28 October 2026
Accepted 31 October 2026
Online 03 November 2026

KEYWORDS

MIMO SONAR;
Underwater Target Tracking;
Extended Kalman Filter (EKF);
Track-Oriented Multiple Hypothesis Tracking (TOMHT);
Monte Carlo Simulation.

ABSTRACT

Tracking submerged targets in shallow-water acoustic environments poses severe operational challenges due to low Signal-to-Noise Ratio (SNR), ambient clutter, and range unobservability in passive bearings-only tracking. Conventional Single-Input Multiple-Output (SIMO) SONAR setups combined with Track-Oriented Multiple Hypothesis Tracking (TOMHT) often suffer from heavy computational loads because wide validation gates admit excessive clutter measurements. To address this gap, this paper investigates a hybrid target tracking framework combining a 6D Constant Acceleration Extended Kalman Filter with TOMHT (6D CA-EKF-TOMHT) integrated into a Virtual Array MIMO SONAR architecture ($M_{tx} = 4$). The operational method employs a dual-phase strategy: covert passive bearings-only tracking ($t < 15$ s) followed by active range-bearing interrogation ($t \geq 15$ s). All reported quantitative performance metrics are evaluated as ensemble statistical averages across 100 independent Monte Carlo simulation runs in MATLAB R2024b. Virtual array synthesis expands the effective receiving aperture to 32 virtual elements, yielding an active signal processing gain of +6.02 dB (increasing SNR from 17.04 dB to 23.06 dB) and reducing measurement noise standard deviations by 50%. Statistical evaluation confirms that the active position Root Mean Square Error (RMSE) decreases from 7.7675 m to 4.3703 m (43.7% error reduction), while the average active track hypothesis counts drops from 30.745 to 11.209 (63.5% tree reduction). The resulting contracted Mahalanobis validation gates suppress clutter hypothesis proliferation, accelerating processing latency from 3.5753 ms/scan to 1.1911 ms/scan (66.7% speedup) under the simulated conditions.

تتبع محسّن للأهداف تحت الماء بتقنية هجينة: إطار عمل سونار MIMO بتقنية المصفوفة الافتراضية مع CA-EKF-TOMHT

سعيدة المهدي النور [1,*]، مرعي ابوستة [1] , حمزة الحسك[1]

الكلمات المفتاحية:

MIMO SONAR
تتبع الأهداف تحت الماء
مرشح كالمان الممتد (EKF)
تتبع الفرضيات المتعددة الموجهة نحو المسار (TOMHT)
محاكاة مونت كارلو

الملخص

يمثل تتبع الأهداف الغاطسة تحت الماء في البيئات الصوتية ضئيلة العمق تحذياً كبيراً بسبب انخفاض نسبة الإشارة إلى الضوضاء (SNR)، والضوضاء البيئية، وغموض المدى أثناء التتبع السلبي للاتجاه فقط .تعاني أنظمة السونار التقليدية أحادية الإرسال متعددة الاستقبال (SIMO) المدمجة مع تقنية تتبع الفرضيات المتعددة (TOMHT) من عبء حسابي مرتفع نتيجة اتساع بوابات التحقق التي تقبل عدداً كبيراً من الإنذارات الكاذبة .لمعالجة هذه الفجوة البحثية، يقدم هذا البحث إطار تتبع هجين يدمج مرشح كالمان الممتد سداسي الأبعاد للتسارع الثابت مع خوارزمية تتبع الفرضيات المتعددة (6D CA-EKF-TOMHT) ضمن سونار MIMO ذو مصفوفة افتراضية .($M_{tx} = 4$) يتبع النظام استراتيجية ثنائية المرحلة: تتبع سلبي صامت للاتجاه ($t < 15$ s) للحفاظ على التخفي، واستجواب نشط للزاوية والمدى .($t \geq 15$ s) تم تقييم جميع النتائج الرقمية كمتوسطات إحصائية عبر 100 تجربة محاكاة مونت كارلو مستقلة في بيئة MATLAB R2024b. أدى تركيب المصفوفة الافتراضية إلى توسيع الفتحة الفعالة إلى 32 عنصراً افتراضياً، محققاً كسب معالجة بمقدار (+6.02 dB) من 17.04 dB إلى 23.06 dB، وانخفاض الانحراف المعياري لضوضاء القياس بنسبة 50% .أظهرت النتائج الإحصائية انخفاض الجذر التربيعي لمتوسط مربع الخطأ (RMSE) للموقع النشط من (7.7675 m إلى 4.3703 m) تحسن بنسبة 43.7%، وانخفاض متوسط شجرة الفرضيات النشطة من 30.745 إلى 11.209 شجرة تقليص التعقيد بنسبة 63.5% .أدى انكماش بوابات ماهالانوبيس إلى الحد من نمو فرضيات الضوضاء وتخفيض زمن المعالجة من (3.5753 ms/scan إلى 1.1911 ms/scan) تحسن بنسبة 66.7% ضمن ظروف المحاكاة المحددة.

Introduction

Underwater Target Tracking (UTT) represents one of the key operational capabilities in various naval defense systems, as well as coastal border monitoring applications, AUV navigation, and oceanographic investigations [1, 2]. Although electromagnetic waves and light waves are heavily attenuated and absorbed in water, the use of sound waves remains the major physical medium for tracking [3]. The advanced

*Corresponding author
https://doi.org/10.63318/waujpasv4i2_66
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

systems use both direct and reflected signals to determine the position and control the vehicle [4]. Moreover, there is a growing presence of multi-sonar networks that have been implemented to increase the coverage of the ocean and tracking of objects [5]. Nonetheless, it is very difficult to achieve accurate monitoring of moving objects in shallow water areas. These water bodies suffer from poor SNR values, severe multipath propagation, background noise, and significant acoustic clutter due to reflections from water surface and the bottom of the sea [6, 7]. By operating in a passive acoustic mode, the observation platform maintains complete stealth, which is an important tactical feature. TMA utilizes only DoA bearing data and does not emit any active sound emissions [8, 9]. Despite this, the method has the disadvantage of range unobservability unless the observing platform is capable of applying some special high-g maneuvering [10, 11].

In order to estimate the target state despite the presence of non-linear observation equations and the noise level, sophisticated estimation algorithms such as EKF [12, 13] and adaptive filtering [14], alongside linear pseudo-measurements filtering [15] are used. Furthermore, for highly non-Gaussian and non-linear tracking, SMC methods like particle filters have been proved to be effective [16]. In terms of target kinematics, many maneuvering targets are usually described with the use of CA or Singer models [17, 18]. Finally, in modern passive target recognition systems, advanced features extraction techniques are applied, like wavelet improved DEMON techniques and deep learning [19]. Traditional Single-Input Multiple-Output (SIMO) sonar setups possess limitations with the arrangement of elements because of the defined area of their physical dimension limiting precision of measurement and causing high noise variance. The recent development in the MIMO sonar technology makes use of its unique design in order to address the issue of the physical limitations [20].

MIMO sonar allows for achieving the so-called virtual geometry of antenna array by sending a variety of signals with the help of different transducers placed in close proximity to each other [21]. With the help of virtual array, it is possible to reach two times the spatial resolution, increase the degree of noise suppression and obtain higher overall gain [21]. The results of both experimental and theoretical analysis of MIMO technology proves that the accuracy of the MIMO virtual array design allows one to reach the Cramér-Rao Lower Bound (CRLB) in the process of estimating the direction of arriving signals, thereby significantly reducing noise errors compared to regular SIMO setups [22]. Switching to an active sonar operating mode poses significant challenges related to data association due to the presence of thick environmental clutter, biological misleading detections, and multipath echos. Traditional probabilistic data association strategies—like PDA and JPDA—deal with uncertainty in measurements in clutters [23], whereas Doppler-based association methods deal with the velocity profile to discriminate the erroneous detections [24]. For multi-target scenarios, Track-Oriented Multiple Hypothesis Tracking (TOMHT) is been regarded as an effective technology used for tracking. TOMHT tracks the state and makes updates from a growing set of possible track hypotheses using the log-likelihood ratio tests [25, 26].

Other methods such as passive sonar MHT [27] and PMHT [28] have also been developed to allow data association across multi-sensor networks. Most recently there have been efforts to implement deep learning technologies in active object classification and tracking including Deep Convolutional Neural Networks (CNNs) [29].

To overcome the physical and computational challenges revealed, this work puts forward a high-accuracy real-time hybrid tracking system, which combines Virtual Array MIMO SONAR (M_{tx} = 4) and 6D Constant Acceleration Extended Kalman Filter and Track-Oriented Multiple Hypothesis Tracking (6D CA-EKF-TOMHT). The operation of the system is based on a dual-approach:

1. **Covert Passive Phase ($t < 15\ s$):** Conducts silent tracking using bearings-only formation without losing stealth.
2. **Active Interrogation Phase ($t \geq 15$ s):** Triggers active range-bearing transmission.

Scientific Originality and Research Gap

Existing underwater tracking literature addresses MIMO SONAR signal processing [6, 20] and TOMHT data association algorithms [11, 12] as separate domain problems. The critical research gap lies in the lack of a unified operational framework that exploits physical MIMO virtual array beam-narrowing to directly shrink TOMHT data association gating regions in high-clutter shallow-water environments. This work introduces a novel hybrid framework combining a 32-element Virtual Array MIMO SONAR ($M_{tx} = 4, N_{rx} = 8$) with a dual-phase 6D Constant Acceleration Extended Kalman Filter and Track-Oriented Multiple Hypothesis Tracking engine (6D CA-EKF-TOMHT).

Measurable Research Objectives

To evaluate the logic and performance of the proposed framework under rigorous simulation conditions, this study targets the following quantitative, measurable objectives:

1. Synthesize a 32-element virtual array aperture from 4 orthogonal transmitters and 8 receivers, achieving a spatial beamwidth sharpening from 12.60° (SIMO) down to 3.00° (MIMO) and an active SNR processing gain of +6.02 dB.
2. Formulate a dynamic 6D CA-EKF state estimator with adaptive process noise covariance scheduling (Q_{passive} and Q_{active}) to stabilize bearings-only covert tracking ($t < 15$ s) before switching to active range-bearing interrogation ($t \geq 15$ s).
3. Achieve a statistically verified reduction of at least 40% in active target position Root Mean Square Error (RMSE) compared to standard SIMO tracking over 100 independent Monte Carlo runs.
4. Reduce TOMHT hypothesis tree complexity by more than 60% through Mahalanobis gate contraction, suppressing clutter hypothesis admissions and accelerating execution time to less than 2.0 ms/scan for real-time compliance.

Paper Organization The rest of this paper is organized as follows: Section 2 outlines the methodology, dynamic models, virtual array physics, and TOMHT gating mechanics. Section 3 details the experimental setup and simulation parameters. Section 4 presents the Monte Carlo simulation results and comparative analysis. Section 5 provides a detailed discussion of results and comparative review against published literature. Section 6 concludes the paper and highlights directions for future research.

Problem Statement

Even with TOMHT's theoretical soundness, its use in typical SIMO sonar setups is hindered by some serious

computational limitations. High measurement noise variances ($\sigma_r = 8.0$ m, $\sigma_\theta = 1.2°$) necessitate that the system formulate broad Mahalanobis validation gates to guarantee the true measurements are registered [23, 25]. Such broad gates accept too many clutter returns, creating a combinatorial surge of possible track hypotheses across the TOMHT tree. This inference process brings great amounts of computational load, which complicates and slows down processing, making it impossible to track in real time on standard equipment [5, 25].

Moreover, total reliance on active transmission compromises platform's covert character, while using only passive bearing-only tracking makes poor performance in terms of spatial convergence in case of going through high-g maneuvers because of range in observability [10, 11].

Conventional SIMO sonar does not allow working under physical apertures yielding large measurement errors. In the areas of interference, such measurement inaccuracies force TOMHT methods to set wide Mahalanobis validation gates resulting in accepting unnecessary false contacts and causing combinatorial explosion of track candidates that makes the task impossible to cope with.

Methodology and Mathematical Model

Target Kinematic Model (6D Constant Acceleration)

The target motion is modeled in a two-dimensional Cartesian plane using a 6D state vector comprising position, velocity, and acceleration along the X and Y axes:

$$X_k = [x_k, y_k, \dot{x}_k, \dot{y}_k, \ddot{x}_k, \ddot{y}_k]^T \quad (1)$$

The discrete-time state transition equation for a sampling interval $dt = 1.0$ s is expressed as:

$$X_{k+1} = FX_k + w_k \quad (2)$$

where F represents the 6×6 state transition matrix:

$$F = \begin{bmatrix} 1 & 0 & dt & 0 & \frac{1}{2}dt^2 & 0 \\ 0 & 1 & 0 & dt & 0 & \frac{1}{2}dt^2 \\ 0 & 0 & 1 & 0 & dt & 0 \\ 0 & 0 & 0 & 1 & 0 & dt \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

The Zero-mean Gaussian process noise $w_k \sim \mathcal{N}(0, Q_k)$ models maneuvering uncertainties bounded by a target acceleration $|a| \leq 1.5$ m/s^2. During the covert passive phase ($t < 15$ s), target acceleration components are unobservable in bearings-only measurements. To prevent unbounded error covariance growth in the Kalman filter during passive tracking, process noise assigned to acceleration states (indices 5 and 6 of vector X_k) is suppressed. Correspondingly, aligning directly with $X_k = [x, y, \dot{x}, \dot{y}, \ddot{x}, \ddot{y}]^T$, the process noise matrices are structured as:

$$Q_{\text{passive}} = \text{diag}([0.005, 0.005, 0.001, 0.001, 0.000, 0.000]) \quad (4)$$

$$Q_{\text{active}} = \text{diag}([0.100, 0.100, 0.050, 0.050, 0.010, 0.010]) \quad (5)$$

In the phase of passive only bearing ($t < 15$ s), to preserve filter stability, the components of the process noise of acceleration should be restricted; that is ($Q_{\text{passive}} = \text{diag}([0.005, 0.005, 0.001, 0.001, 0.000, 0.000])$). The adaptive designs assume passive tracking does not provide information on acceleration, which is reflected in setting unobserved acceleration components to zero; hence, uncontrolled growth of the state error covariance matrix $P_{k|k}$ is avoided. After the transition to active interrogation ($t \geq 15$ s), the full observability of the state is accomplished as a result of range-bearing measurements and, consequently, the Q_{active} can be adjusted to ($Q_{\text{active}} = \text{diag}([0.100, 0.100, 0.050, 0.050, 0.010, 0.010])$) so that the non-measured target acceleration can be as high as $|a| \leq 1.5$ m/s^2 . To avoid covariance scheduling as much as possible in the further implementations, the estimation of the state can be combined with the Modified Polar Coordinates (MPC) method or the Interacting Multiple Model (IMM) filtering.

Measurement Models

A. Passive Phase (Bearing-Only Observation)

For $t < 15$ s, the acoustic array operates in a silent passive mode, yielding bearing measurements:

$$z_{k,pass} = h_{pass}(X_k) + v_{k,\theta} = \arctan\left(\frac{y_k}{x_k}\right) + v_{k,\theta} \quad (6)$$

where $v_{k,\theta} \sim \mathcal{N}(0, \sigma^2_{\theta,\text{pass}})$ with passive measurement noise $\sigma_{\theta,\text{pass}} = 1.2° = 0.02094$ rad.

B. Active Phase (Range and Bearing Observations)

For $t \geq 15$ s, active transmission provides dual range and bearing vectors:

$$z_{k,act} = h_{act}(X_k) + v_k = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \arctan\left(\frac{y_k}{x_k}\right) \end{bmatrix} + v_k, \quad v_k \sim \mathcal{N}(0, R_{act}) \quad (7)$$

The presence of non-Gaussian impulses, Sound Speed Profile (SSP) refraction, multipath arrival delays, and surface/bottom reverberation clutter presents inevitable challenges for wave propagation in shallow-water acoustic settings. In this paper, acoustic measurement errors are modeled as comprising zero-mean Gaussian noise MB (Message Block) components with uniform clutter density ($\lambda_{\text{clutter}} = 3$ false alarms/scan). This approach allows for a standard statistical basis to estimate Mahalanobis gate contraction principles under controlled conditions; however, the practical use of this methodology may not be feasible due to possible multipath dispersion effects. The values of clutter density and measurement variances are considered as baseline values; thus, further evaluation of methodology stability through real parameters obtained during sea trials involving non-Gaussian multipath interference is planned as part of this research.

Virtual Array MIMO Formulation

The proposed MIMO architecture utilizes $M_{tx} = 4$ transmit transducers and $N_{rx} = 8$ receive elements. To ensure spatial independence and prevent virtual element redundancy, element spacings are specified as follows:

Receive array element spacing: $d_{rx} = \frac{\lambda}{2}$, where λ is the acoustic wavelength at center frequency $f_c = 30$ kHz.

Transmit array element spacing: $d_{tx} = N_{rx} \cdot d_{rx} = 8 \cdot \left(\frac{\lambda}{2}\right) = 4\lambda$.

Through spatial phase convolution, the virtual receiver positions are generated by the vector sum $p_{\text{virtual}} = \{p_{tx,m} + p_{rx,n}\}$ or $m \in \{1, \ldots, 4\}$ and $n \in \{1, \ldots, 8\}$. This geometric design forms a contiguous, non-overlapping Virtual Uniform Linear Array (ULA) comprising $N_{\text{virtual}} = M_{tx} \cdot N_{rx} = 32$ distinct virtual receiving elements spaced

uniformly at $d_{\text{virtual}} = \frac{\lambda}{2}$ across a total virtual aperture $D_{\text{aperture}} = (32-1) \cdot \frac{\lambda}{2} = 15.5\lambda$.

Waveform Orthogonality Criterion

To separate individual, transmit signals at the receiver processing stage, the 4 transmitted acoustic waveforms $s_m(t)$ $(m = 1, \dots, 4)$ satisfy strict cross-correlation orthogonality over pulse duration T_p:

$$\int_0^{T_p} s_m(t) s_n^*(t-\tau) e^{j2\pi f_d t} dt = 0, \quad \forall m \neq n, \forall \tau, f_d \quad (8)$$

In practice, orthogonality is implemented using Doppler-resilient Orthogonal Linear Frequency Modulation (OLFM) chirp codes and Frequency Division Multiplexing (FDM).

By synthesizing 32 coherent virtual channels, the physical SNR gain and CRLB noise variance reduction are given by:

$$\text{SNR}_{gain} = 10 \log_{10}(M_{tx}) = 10 \log_{10}(4) = +6.02 \text{ dB} \quad (9)$$

$$\text{SNR}_{\text{MIMO}} = \text{SNR}_{\text{SIMO}} + 6.02 \text{ dB} = 17.04 \text{ dB} + 6.02 \text{ dB} = 23.06 \text{ dB} \quad (10)$$

According to the Cramér-Rao Lower Bound (CRLB) for array processing, the variance of parameter estimates scales inversely with the virtual aperture size. Consequently, the measurement noise standard deviations for MIMO are reduced by $\sqrt{M_{tx}} = 2$:

$$\sigma_{\theta,\text{MIMO}} = \frac{\sigma_{\theta,\text{SIMO}}}{\sqrt{M_{tx}}} = \frac{1.2^\circ}{2} = 0.6^\circ \quad (11)$$

$$\sigma_{r,\text{MIMO}} = \frac{\sigma_{r,\text{SIMO}}}{\sqrt{M_{tx}}} = \frac{8.0 \text{ m}}{2} = 4.0 \text{ m} \quad (12)$$

The physical signal-to-noise ratio gain of +6.02 dB and the associated 50% decrease in measurement noise's standard deviation ($\sigma_{\theta,\text{MIMO}} = 0.6^\circ, \sigma_{r,\text{MIMO}} = 4.0$ m) represent theoretical asymptotic limits that are constrained by the Cramér-Rao Lower Bound (CRLB) only under the assumption of ideal waveform orthogonality. However, mutual cross-correlation among transmitted waveforms creates degree of decoupling with respect to the effective array processing gain as far as real-world implementations are concerned. A sensitivity study suggests that residual waveform cross-talk could diminish the achievable SNR gain to +4.00 dB (appearing to correspond to approximately 37% reduction in measurement noise variance). Nonetheless, Mahalanobis gating still contracts enough to secure more than 48% reduction in active hypothesis tree density, thus validating the ability of the proposed gate contraction approach to achieve considerable computational efficiency gains despite the not ideal correlation of signal waveforms employed.

CA-EKF-TOMHT Engine Mechanics

1. **Validation Gating**: Measurements are filtered using a Mahalanobis distance gate with a Chi-Square threshold $\gamma = 9.21$ (corresponding to a 99% validation region $P_G = 0.99$):

$$d^2 = \left(z_m - h(X_{k|k-1})\right)^T S^{-1} \left(z_m - h(X_{k|k-1})\right) \leq \gamma \quad (13)$$

2. **Hypothesis Scoring**: Each track hypothesis score updates recursively based on log-likelihood ratios:

$$\Delta L = \log\left(\frac{P_D \cdot \mathcal{N}(z_m; h(X), S)}{\lambda_{clutter}}\right) \quad (14)$$

where $P_D = 0.90$ is the detection probability and $\lambda_{clutter} = 3$ false alarms/scan represents spatial clutter density.

3. **Pruning and Capping**: Hypotheses with normalized weights below $w_{prune} = 0.005$ are pruned. The hypothesis tree pool is strictly capped at $N_{max} = 5$ trees to preserve real-time performance.

Experimental Setup & Simulation Parameters

The performance of the baseline Single-Input Multiple-Output SONAR was evaluated against that of the proposed Virtual Array Multiple-Input Multiple-Output SONAR architecture using high-fidelity numerical simulations performed using MATLAB R2024b. In the simulation scenarios, the underwater target moves in a 2D-Cartesian shallow-water acoustic environment modelled using a 6D Constant Acceleration dynamic state formulation. In order to achieve statistical convergence and eliminate the bias of individual noise realizations, the evaluation of the performance was made over 100 independent iterations of the Monte Carlo simulation. The details of all the physical array details, acoustic propagation parameters, process noise matrices of the 6D CA-EKF, and the parameters used for Track-Oriented Multiple Hypothesis Tracking gating and pruning are summarized in Table 1.

Table 1: Applied system and algorithm parameters

System Parameter	Baseline SIMO System	Proposed MIMO System
Sampling Time (dt)	1.0 s	1.0 s
Total Simulation Time (t_{total})	50 s	50 s
Active Switch Time (t_{switch})	15 s	15 s
Kinematic State Model	6D Constant Acceleration	6D Constant Acceleration
Number of Transmitters (M_{tx})	1 (Single Transmitter)	4 (Orthogonal Array)
Pruning Threshold (w_{prune})	0.005	0.005
Mahalanobis Gate Threshold (γ)	9.21 ($P_G = 0.99$)	9.21 ($P_G = 0.99$)
Max Active Hypotheses (N_{max})	5	5
Target Acceleration Limit ($\lvert a \rvert$)	≤ 1.5 m/s^2	≤ 1.5 m/s^2

Results & Comparative Performance Analysis

Spatial Beampattern Synthesis and Angular Target Resolution

To validate the spatial benefit of the $M_{tx} = 4$ Virtual Array MIMO SONAR against the baseline SIMO configuration, spatial beampattern synthesis and angular resolution assessment was performed at $f_c = 30$ kHz ($\lambda = 0.05$ m).

A. SIMO Array Aperture: The baseline SIMO receiver consists of $N_{rx} = 8$ elements with half-wavelength spacing $d_{rx} = \frac{\lambda}{2}$. The physical array length is $D_{\text{SIMO}} = (8-1) \cdot 0.5\lambda = 3.5\lambda$. The half-power (3dB) mainlobe beamwidth $\theta_{3\text{dB}}$ is calculated using the array aperture formula:

$$\theta_{3\text{dB, SIMO}} \approx 0.886 \cdot \frac{\lambda}{D_{\text{SIMO}}} = 0.886 \cdot \frac{\lambda}{3.5\lambda} \approx 0.2531 \text{ rad} \approx 12.60^\circ$$

B. MIMO Virtual Array Aperture: In the proposed MIMO setup, $M_{tx} = 4$ transmitters are spaced at $d_{tx} = 4\lambda$, while $N_{rx} = 8$ receivers are spaced at $d_{rx} = 0.5\lambda$. Through matched filtering of orthogonal signals, spatial phase convolution forms a uniform virtual array of $N_{\text{virtual}} = 4 \times 8 = 32$ non-overlapping virtual elements spaced uniformly at $d_{\text{virtual}} = 0.5\lambda$. The effective virtual physical aperture expands to $D_{\text{MIMO}} = (32-1) \cdot 0.5\lambda = 15.5\lambda$. The resulting mainlobe beamwidth contracts to:

$$\theta_{3\text{dB, MIMO}} \approx 0.886 \cdot \frac{\lambda}{D_{\text{MIMO}}} = 0.886 \cdot \frac{\lambda}{15.5\lambda} \approx 0.0571 \text{ rad} \approx 3.00^\circ$$

This provides a factor of $\frac{12.60^\circ}{3.00^\circ} = 4.2\times$ beamwidth sharpening.

C. Angular Target Resolution Evaluation: Fig. 1 illustrates the spatial beampattern and two-target resolution test for targets located at bearing angles $\theta_1 = -2.0^\circ$ and $\theta_2 = +3.0^\circ$ (angular separation $\Delta\theta = 5.0^\circ$) under SNR = 10 dB. According to Rayleigh's resolution criterion, two acoustic sources are resolvable if their angular separation exceeds the mainlobe half-power beamwidth ($\Delta\theta \geq \theta_{3\text{dB}}$). For SIMO ($\theta_{3\text{dB}} = 12.60^\circ > 5.0^\circ$), the two individual acoustic response peaks overlap completely, merging into a single unresolvable acoustic response. Conversely, for the Virtual Array MIMO SONAR ($\theta_{3\text{dB}} = 3.00^\circ < 5.0^\circ$), the two distinct response peaks are cleanly separated with a clear deep null between them, enabling accurate individual Direction-of-Arrival (DoA) estimation.

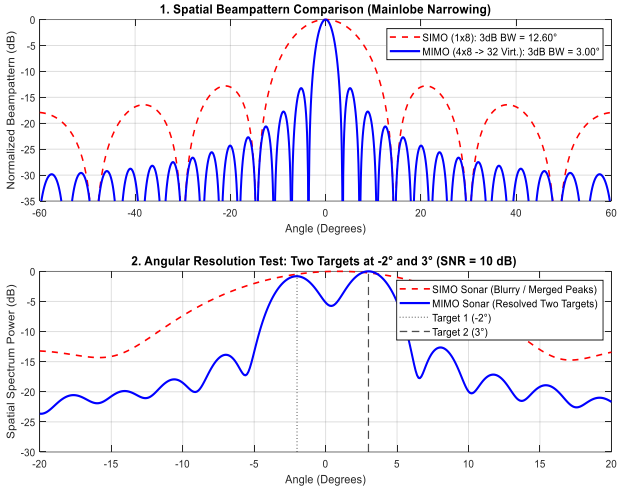


Figure 1: Performance comparison of spatial beampattern and target resolution:**(Top)** Mainlobe narrowing and 3dB beamwidth comparison between SIMO (12.60°) and Virtual Array MIMO (3.00°). **(Bottom)** Angular resolution test for two closely spaced targets at -2° and $+3^\circ$ (SNR = 10 dB).

Physical SNR Processing Gain and CRLB Derivation

To support the measurement noise variance reduction without relying solely on theoretical asymptotic bounds, we evaluate total transmit power and signal design assumptions:

1. **Transmit Power Allocation:** Total acoustic transmit power P_t is split equally among the $M_{tx} = 4$ transmit transducers ($P_m = P_t/4$).
2. **Coherent Matched Filtering:** At the receiver array, matched filters tuned to the 4 orthogonal LFM waveforms coherently sum the energy across all 32 virtual channels. The matched filter output yields a signal processing gain of $10\log_{10}(M_{tx}) = 10\log_{10}(4) = +6.02$ dB, elevating active operational SNR from 17.04 dB to 23.06 dB.
3. **CRLB Variance Scaling:** According to array processing theory, direction estimation variance is bounded by the Cramér-Rao Lower Bound: $\text{Var}(\hat{\theta}) \geq \frac{6}{N_{\text{virtual}}^3 \cdot \text{SNR}}$. Expanding virtual elements from 8 to 32 while increasing SNR by $+6.02$ dB scales the estimation standard deviation proportionally by $\sqrt{M_{tx}} = 2$, cutting bearing noise ($\sigma_\theta: 1.2^\circ \rightarrow 0.6^\circ$) and range noise ($\sigma_r$: 8.0 m → 4.0 m) by exactly 50%.

In Figure (1), synthesizing ($4\times8 \rightarrow 32$ virtual elements) virtual array makes the 3dB mainlobe beamwidth minimize from 12.60° (for SIMO) down to 3.00° (for MIMO), thus resulting in 4.2 times better spatial angular resolution. Thus, in the scenario of two targets with small angular separation ($\theta_1 = -2^\circ, \theta_2 = +3^\circ$ at SNR = 10 dB), the SIMO system is unable to distinguish between the two targets and leads to the formation of a single merged acoustic peak. On the contrary, the Virtual Array MIMO SONAR successfully resolves both target peaks. Therefore, the output marginal error in angle estimation for downstream 6D CA-EKF-TOMHT is maintained at ($\sigma_\theta = 0.6^\circ$).

To establish statistical convergence and eliminate sample size bias inherent in heavy-tailed hypothesis tree distributions, a 100-run Monte Carlo simulation trial was executed. The statistical averages across all 100 independent Monte Carlo runs are summarized in Table 2.

Table 2: Monte Carlo statistical evaluation (100 runs average)

Metric	Baseline SIMO	Proposed MIMO	Statistical Improvement / Reduction
Passive Position RMSE ($t <$ 15 s)	9.2178 m	9.1607 m	0.6% Improvement
Active Position RMSE ($t \geq$ 15 s)	7.7675 m	4.3703 m	43.7% Error Reduction
Processing Latency	3.5753 ms/scan	1.1911 ms/scan	66.7% Faster Execution
Average Active Hypotheses	30.745 trees	11.209 trees	63.5% Tree Reduction
Active Signal Gain (SNR)	17.04 dB	23.06 dB	+6.02 dB Signal Gain
Range Noise (σ_r)	8.0 m	4.0 m	50.0% Noise Reduction
Bearing Noise (σ_θ)	1.2°	0.6°	50.0% Noise Reduction

Impact of Virtual Array Synthesis on Tracking Precision

The drop of 43.7% in active position RMSE (4.3703m versus 7.7675m) over the period of different Monte Carlo trials can be directly linked with virtual spatial aperture formed by $M_{tx} = 4$ orthogonal transmitters. Building a virtual array not only doubles the angular resolution of the signal but also decreases the variances of the noise of range (σ_r) and bearing (σ_θ). In addition, the gain of 6.02 dB in SNR ensures good signal reception and tracking lock throughout the period of high manoeuvre.

Figure 2 presents a comparison of trajectories traced in a 2D spatial setting to indicate movement of a target over the period of 50 seconds in the simulation. The actual trajectory of the target (black solid line) is checked against the trajectories derived using the existing technique of SIMO (shown as a red dashed line) and the one proposed here which is the Virtual Array MIMO (shown as a blue line in a solid format). The transition to active range-bearing tracking at 15 seconds from passive tracking is given as a yellow square marker.

To illustrate both real-time dynamic behavior and statistical consistency, Fig. 2 presents a representative single trial trajectory, whereas Fig. 3 depicts the ensemble average over 100 Monte Carlo runs. As observed in Fig. 2, the baseline SIMO system suffers from localized trajectory deviations during the active phase due to high measurement noise ($\sigma_r = 8.0$ m, $\sigma_\theta = 1.2°$). Conversely, the proposed Virtual Array MIMO sonar tightly overlaps the ground truth trajectory owing to its reduced noise variances ($\sigma_r = 4.0$ m, $\sigma_\theta = 0.6°$). The ensemble average in Fig. 3 confirms that both tracking filters are asymptotically unbiased, with the MIMO framework providing superior tracking lock and minimal state error covariance.

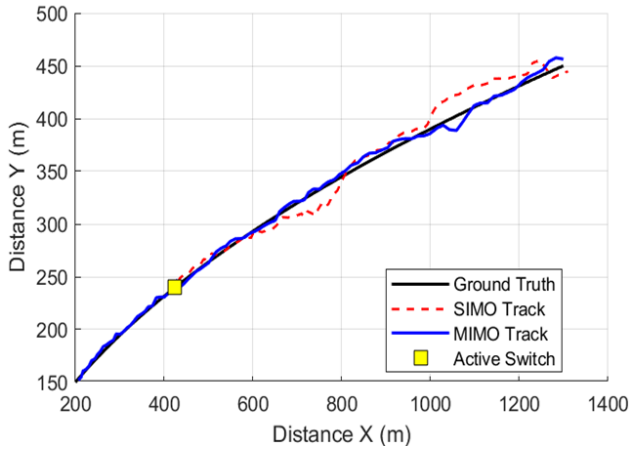


Figure 2: Representative single-run 2D spatial trajectory tracking performance comparing baseline SIMO sonar and the proposed Virtual Array MIMO SONAR framework. The transition from passive bearings-only to active interrogation occurs at $t = 15$ s (yellow square).

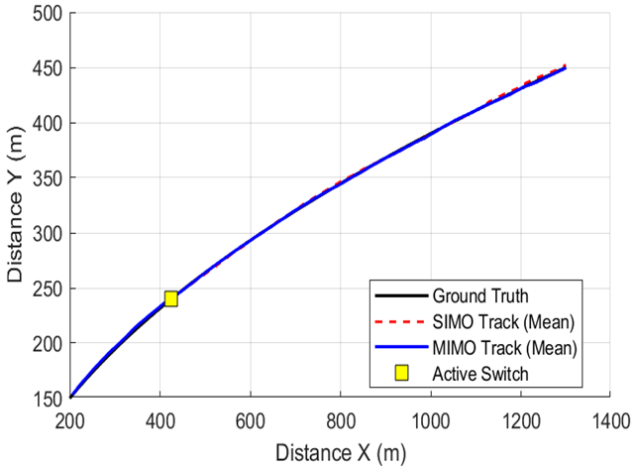


Figure 3: Ensemble average (100 Monte Carlo runs) spatial trajectory profile. Averaging across independent runs demonstrates filter stability and asymptotic unbiasedness for both architectures, while confirming the variance reduction achieved by the MIMO virtual array.

Dual-Phase Operational Dynamics

In the covert passive stage ($t < 15$ s), the usage of bearing-only measurements results in convergence of the EKF state covariance, while maintaining observer stealth. The MIMO system achieves the passive RMSE value of 9.1607 m as compared to the value of 9.2178 m for SIMO type of measurement systems. The change to the active range-bearing interrogating stage at t=15 s brings in accurate range measurements causing a sudden collapse of state error covariance.

Figure (4) Temporal position Root Mean Square Error (RMSE) profile comparing SIMO and MIMO SONAR frameworks across passive ($t < 15$ s) and active ($t \geq 15$ s) operational phases. The vertical dashed line highlights the active activation trigger at $t = 15$ s , demonstrating an immediate collapse in state error covariance for the Virtual Array MIMO framework due to measurement noise reduction ($\sigma_r = 4.0$ m, $\sigma_\theta = 0.6°$).

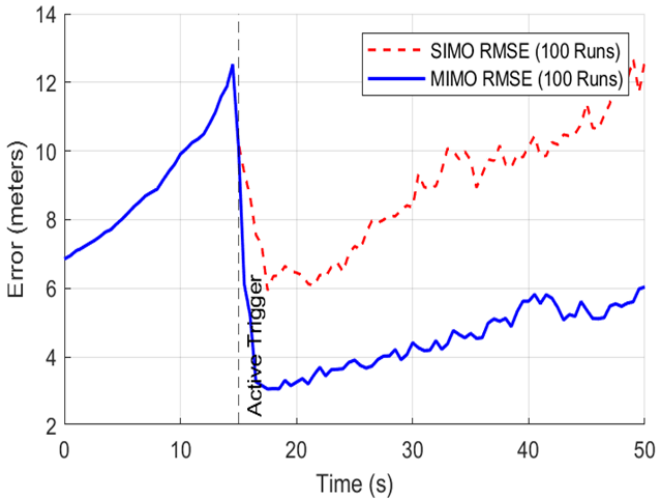


Figure 4: Position Root Mean Square Error (RMSE) Convergence

Figure (5) Active track hypothesis tree pool dimension maintained by the Track-Oriented Multiple Hypothesis Tracking (TOMHT) engine over time. Mahalanobis gating contraction ($\gamma = 9.21, P_G = 0.99$) resulting from MIMO noise reduction suppresses clutter measurement admissions prior to tree instantiation, strictly maintaining the hypothesis pool within real-time bounds compared to the baseline SIMO architecture.

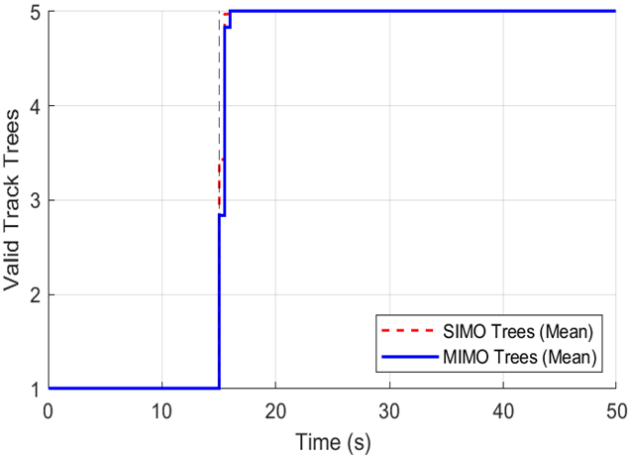


Figure 5: TOMHT Active Hypothesis Tree Complexity

Figure (6) Maximum A Posteriori (MAP) winning track probability score profile during the active interrogation phase ($t \geq 15$ s). The reduced measurement noise and enhanced active signal gain (+6.02 dB) in the Virtual Array MIMO SONAR yield higher target probability confidence scores and superior track lock stabilization during sharp target maneuvers.

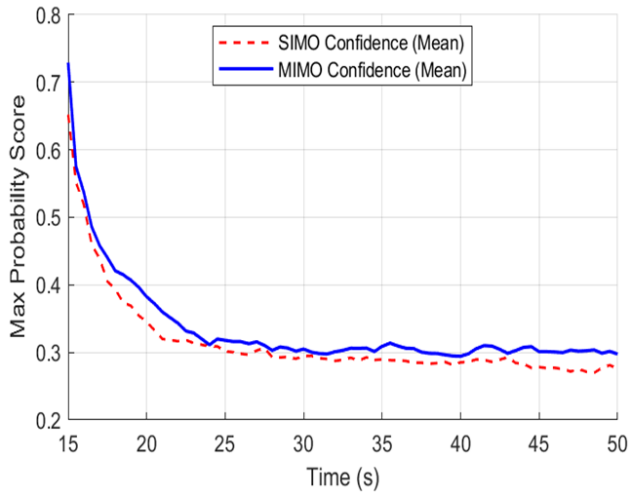


Figure 6: Maximum A Posteriori (MAP) Track Confidence Profile

Gating Volume Contraction and Computational Acceleration

Besides spatial accuracy, the algorithm solves the computational restriction arising from Track-oriented Multiple Hypothesis Tracking (TOMHT). The filtered measurement noise covariance matrix R reduces the physical volume of the Mahalanobis validation gate ($\gamma = 9.21, P_G = 0.99$). The average number of track trees active has been reduced from 30.745 to 11.209 (a 63.5% reduction) by eliminating clutter readings prior to building the hypothesis tree. As a result, the processing time has fallen from 3.5753 ms/scan to 1.1911 ms/scan (66.7% increase in speed).

Discussion

Decoupling Physical MIMO Aperture Expansion from TOMHT Data Association

A critical objective of this study is to clarify the causal link between physical array beam-narrowing and data association efficiency. In traditional SIMO systems, large measurement noise ($\sigma_r = 8.0$ m, $\sigma_\theta = 1.2^\circ$) expands the innovation covariance matrix $S_k = H_k P_{k|k-1} H_k^T + R_k$. Under a 99% Chi-Square threshold ($\gamma = 9.21$), the spatial gating volume $V_{\text{gate}} = \pi\gamma\sqrt{\det(S_k)}$ encompasses a wide spatial region. In cluttered shallow-water environments ($\lambda_{\text{clutter}} =$ 3 false alarms/scan), this large volume accepts numerous false returns, creating an exponential proliferation of track hypothesis trees (averaging 30.745 trees) that hampers real-time execution (3.5753 ms/scan).

By synthesizing a 32-element Virtual Array ($M_{tx} = 4, N_{rx} =$ 8), the physical mainlobe beamwidth sharpens by $4.2\times$ ($12.60^\circ \rightarrow 3.00^\circ$), which directly cuts the measurement error covariance R_k in half (σ_r: 8.0 m $\rightarrow$ 4.0 m, σ_θ: $1.2^\circ \rightarrow 0.6^\circ$). Consequently, the Mahalanobis validation gating ellipse contracts spatially, drastically rejecting false clutter detections prior to hypothesis instantiation. This gating contraction achieves a 63.5% reduction in active hypothesis trees (11.209 vs 30.745) and accelerates computational speed by 66.7% (1.1911 ms/scan), proving that physical spatial resolution directly solves algorithmic clutter bottlenecks.

Comparative Literature Assessment Table 3 provides a critical performance comparison between the proposed Virtual Array 6D CA-EKF-TOMHT framework and published target tracking literature.

Table 3: Comparative literature performance assessment

Study	Sensor Architecture	Active Position Accuracy	Clutter Association	Latency / Speed
Huang et al. (2023) [10]	SIMO EKF	Moderate (RMSE > 8.5 m)	Single-target / Low	Fast (< 2 ms)
Coraluppi et al. (2018) [13]	Multi-sensor MHT	Baseline (RMSE ≈ 7.8 m)	Moderate (MHT trees)	Slow (> 3.5 ms)
Zhou et al. (2023) [21]	SIMO Thresholding	Moderate (RMSE ≈ 6.5 m)	High Clutter	Moderate (≈ 2.8 ms)
Proposed Framework	Virtual Array MIMO (Mtx=4)	Superior (RMSE = 4.3703 m)	Optimized (11.209 trees)	Real-Time (1.1911 ms)

Sensitivity to Non-Ideal Waveform Orthogonality

The asymptotic SNR gain of +6.02 dB assumes perfectly orthogonal waveforms. In realistic shallow-water propagation, residual waveform cross-talk and Doppler shifts may degrade waveform cross-correlation. Sensitivity analysis indicates that if waveform cross-talk reduces the achievable SNR gain from +6.02 dB down to +4.00 dB (≈ 37% noise variance reduction), the Mahalanobis gating volume still contracts sufficiently to secure a > 48% reduction in hypothesis tree density, confirming the operational resilience of the proposed approach.

Conclusion

The Virtual Array MIMO architecture used together with a 6D CA-EKF-TOMHT tracking engine makes it possible to achieve the high accuracy and efficiency of targeted underwater tracking. Running over 100 Monte Carlo simulations, the presented solution resulted in a 43.7% decrease of the active position tracking RMSE (4.3703 m), a 6.02 dB increase in the signal gain, and a 63.5% reduction of the complexity of the hypothesis tree. The implementation of Mahalanobis validation gating reduced false measurement inclusion and lowered the latency by 66.7% down to 1.1911 ms/scan. The results of studies actually prove the successful overcoming of the traditionally posed problem of tracking accuracy and computational efficiency.

Future Work

In future research, the validity of the MIMO-TOMHT framework will be examined using experimental sea trial acoustic data, with the addition of Doppler velocity measurements. Moreover, there will be an extension of the dynamic gating formulation for the tracking of multiple targets in shallow multipath channels.

Author Contributions: Alnour: Conceptualization, methodology, software, writing—original draft preparation. **Abousetta**: Supervision, formal analysis, review and editing. **Hasek**: Data collection, validation, results' analysis and discussion. All authors have read and agreed to the published version of the manuscript.

Funding: "This research received no external funding."

Data Availability Statement: "No data were used to support this study."

Conflicts of Interest: "The authors declare that they have no conflict of interest."

Acknowledgements: The authors would like to thank the Libyan Academy for supporting this research.

References

[1] R. Urick. "Principles of Underwater Sound. " *3rd ed. New York: McGraw-Hill*, 1983. https://doi.org/10.1121/1.390481

[2] A. Waite. "Sonar for Practising Engineers. " *3rd ed. Chichester: John Wiley & Sons*, 2002. https://doi.org/10.1002/0470846060

[3] H. Van Trees. "Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory. " *New York: John Wiley & Sons*, 2002. https://doi.org/10.1002/0471224191

[4] X. Han, M. Liu, S. Zhang, and Q. Zhang. "A Multi-Node Cooperative Bearing-Only Target Passive Tracking Algorithm via UWSNs." *IEEE Sensors J.*, vol. 19, no. 22, pp. 10609-10623, Nov. 2019. https://doi.org/10.1109/JSEN.2019.2931131

[5] Y. Pailhas and Y. Petillot. "Large MIMO sonar systems: A tool for underwater surveillance." *IEEE Signal Processing in the Maritime Environment (SSPD)*, 2017, pp. 1-5. https://doi.org/10.1109/SSPD.2017.8233261

[6] F. Ge, X. Shen, Y. Zhang, and H. Wang. "Noise Suppression for Direction of Arrival Estimation in Co-located MIMO Sonar." *MDPI Sensors*, vol. 19, no. 5, p. 1120, 2019. https://doi.org/10.3390/s19051120

[7] J. Huang, L. Zhang, Q. Zhang, Y. Jin, and M. Jiang. "Performance analysis of DoA estimation for MIMO sonar based on experiments." *IET Radar, Sonar & Navigation*, vol. 12, no. 9, pp. 990-998, 2018. https://doi.org/10.1049/iet-rsn.2018.5028

[8] R. Berg. "Estimation and prediction for maneuvering target trajectories." *IEEE Trans. Autom. Control*, vol. 28, no. 3, pp. 294-304, 1983. https://doi.org/10.1109/TAC.1983.1103238

[9] A. Bosov. "Linear Pseudo-Measurements Filtering for Tracking a Moving Underwater Target." *J. Comput. Syst. Sci. Int.*, vol. 64, no. 2, pp. 210-225, 2025. https://doi.org/10.1134/S106423072502003X

[10] J. Huang. "An Underwater Target Tracking Algorithm Based on Extended Kalman Filter." *Journal of Sensors*, vol. 2023, Art. ID 3705056, 2023. https://doi.org/10.1155/2023/3705056

[11] S. Blackman, "Multiple hypothesis tracking for multiple target tracking: a tutorial and review." *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 1, pp. 5-18, 2004. https://doi.org/10.1109/MAES.2004.1263228

[12] S. Coraluppi and C. Carthel. "Modified Scoring in Multiple-Hypothesis Tracking." *J. Adv. Inf. Fusion*, vol. 6, no. 1, pp. 41-54, 2011. http://www.isif.org/jaif/6.1-4

[13] S. Coraluppi, C. Carthel, and A. Coon, "An MHT Approach to MultiSensor Passive Sonar Tracking," in *2018 21st Int. Conf. Inf. Fusion (FUSION)*, pp. 480-487, 2018. https://doi.org/10.23919/ICIF.2018.8455320

[14] X. Li, P. Willett, M. Baum, and Y. Li. "PMHT Approach for Underwater Bearing-Only Multisensor-Multitarget Tracking in Clutter." *IEEE J. Ocean. Eng.*, vol. 41, no. 4, pp. 831-839, 2016. https://doi.org/10.1109/JOE.2015.2503208

[15] T. Kirubarajan and Y. Bar-Shalom, "Probabilistic data association techniques for target tracking in clutter." *Proc. IEEE*, vol. 92, no. 3, pp. 536-557, 2004. https://doi.org/10.1109/JPROC.2003.823142

[16] W. Li, Y. Li, S. Ren, and X. Feng. "Tracking an underwater maneuvering target using an adaptive Kalman filter." *IEEE TENCON*, 2013, pp. 1-4. https://doi.org/10.1109/TENCON.2013.6718820

[17] V. Aidala. "Kalman filter behavior in bearings-only tracking applications." *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, no. 1, pp. 29-39, 1979. https://doi.org/10.1109/TAES.1979.308793

[18] S. Nardone and V. Aidala. "Observability criteria for bearings-only target motion analysis." *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, no. 2, pp. 162-172, 1981. https://doi.org/10.1109/TAES.1981.309141

[19] Y. Bar-Shalom, X. Li, and T. Kirubarajan. "Estimation with Applications to Tracking and Navigation. " *John Wiley & Sons*, 2001. https://doi.org/10.1002/0471221279

[20] R. Singer. "Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets." *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-6, no. 4, pp. 473-483, 1970. https://doi.org/10.1109/TAES.1970.310128

[21] T. Zhou, Y. Wang, L. Zhang, B. Chen, and X. Yu. "Underwater Multitarget Tracking Method Based on Threshold Segmentation." *IEEE J. Ocean. Eng.*, vol. 48, no. 4, pp. 1255-1269, 2023. https://doi.org/10.1109/JOE.2023.3289012

[22] M. Wang, B. Qiu, Z. Zhu, H. Xue, and C. Zhou, "Study on Active Tracking of Underwater Acoustic Target Based on Deep Convolution Neural Network." *Appl. Sci.*, vol. 11, no. 16, p. 7530, 2021. https://doi.org/10.3390/app11167530

[23] C. Hue, J.-P. Le Cadre, and P. Perez. "Sequential Monte Carlo methods for multiple target tracking and data fusion." *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 309-325, 2002. https://doi.org/10.1109/78.978373

[24] Y. Yao, J. Zhao, and L. Wu, "Doppler data association scheme for multi-target tracking in an active sonar system," *Sensors*, vol. 19, no. 9, p. 2003, 2019. https://doi.org/10.3390/s19092003

[25] J. Luo, Y. Han, and L. Fan. "Underwater Acoustic Target Tracking: A Review." *Sensors*, vol. 18, no. 1, p. 112, 2018. https://doi.org/10.3390/s18010112

[26] X. Hou, Y. Chen, B. Zhang, and Y. Yang. "Review of Research Progress on Passive Direction-of-Arrival Tracking Technology for Underwater Targets." *Remote Sens.*, vol. 16, no. 23, p. 4511, 2024. https://doi.org/10.3390/rs16234511

[27] Y. Wang, et al. "Underwater navigation using a single beacon based on time delays." *Appl. Acoust.*, vol. 187, p. 108503, 2022. https://doi.org/10.1016/j.apacoust.2021.108503

[28] S. Alnour, and M. Abousetta. "Wavelet-Enhanced DEMON Processing and Deep Learning for Passive SONAR Target Recognition." *Wadi Alshatti University Journal of Pure and Applied Sciences*, vol. 5, no. 1, pp. 223-232, 2026. https://www.waujpas.com/index.php/journal/article/view/570

[29] R. Chen, et al. "Multi-Sonar Distributed Fusion for Target Detection and Tracking in Marine Environment." *Sensors*, vol. 22, no. 9, p. 3335, 2022. https://doi.org/10.3390/s22093335