ICSPIS

11th International Conference on Signal Processing & Intelligent Systems
Mazandaran University of Science and Technology, December 23-24, 2025

Mazandaran University of Science and Technology

Enhancing Yaw Error Tolerance in Synthetic Aperture Sonar via Maximum-Likelihood Receiver Selection on Mobius-Strip Arrays

1st Hamid Hajirahimi Kashani*
Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
hajirahimi.hamid@mail.um.ac.ir

2nd Seyed Alireza Seyedin
Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
seyedin@um.ac.ir

Abstract— **Unintended yaw introduces structured phase errors that blur coherent gain and elevate sidelobes in synthetic aperture sonar. This work presents a compact compensation scheme that couples a non-orientable Mobius-strip hydrophone array with twenty elements and a single 180-degree twist to a maximum-likelihood receiver-selection rule. At each ping the method selects the best receiver, interpreted as the element whose phase history most closely matches the pre-motion reference, and then re-anchors phase through a closed-form correction before coherent image formation. The geometry's controlled vertical excursion and near-uniform azimuthal coverage reduce both the number and the strength of yaw-induced phase modes, enabling the selector to operate in a low-rank feature space derived from phase differences and short-term coherence. Evaluation uses a three-by-three point-target grid T1–T9 under a yaw profile of plus-minus 0.8 deg with a 0.5 hertz component and mild colored noise. Metric windows, side-lobe masks, color scales, and dynamic ranges are held fixed across conditions to ensure fair comparison. Image quality is reported with PSLR, ISLR, PSNR, and RMSE averaged over the nine targets. Along the yaw axis the compensated imagery consistently deepens the sidelobe floor and raises fidelity without broadening the main lobe. Averaged over nine targets, PSLR improves by 19.50 decibels, ISLR improves by 18.85 decibels, PSNR increases by 32.76 decibels, and RMSE falls from 0.19588 to 0.00451, demonstrating that maximum-likelihood receiver selection on a Mobius layout is a practical route to yaw-robust synthetic aperture sonar.**

Keywords—Synthetic Aperture Sonar, Yaw, phase-error compensation, Mobius-strip hydrophone array, Maximum-likelihood receiver selection

I. INTRODUCTION

Motion-induced phase error remains a primary limiter of synthetic-aperture sonar (SAS) image quality [1]. Among single-axis perturbations, yaw is especially harmful because it introduces along-track phase distortions that behave like aperture deformation, elevating sidelobes and masking weak reflectors near strong scatterers. Recent analyses of coherent SAS formation under realistic platform dynamics show how even modest temporal Doppler and small rotational excursions can degrade coherent gain and raise background structure, underscoring the need for precise motion handling before beamforming [2].

Conventional countermeasures typically fall into two families: navigation-aided motion compensation that fuses INS/DVL (and sometimes external aids) to regress residual phase, and autofocus procedures that estimate phase terms directly from the data. Both are effective but can be burdensome for autonomous underwater vehicles, either because they demand high-grade navigation packages and careful alignment or because they require iterative, compute-intensive estimation that strains power and memory budgets onboard. Current surveys of SAS processing chains detail these trade-offs and the practical constraints they impose on embedded implementations [3].

This work follows a complementary path: we use array geometry to simplify the structure of yaw-induced phase error and apply a lightweight statistical decision to identify, on each ping, the receiver that most faithfully represents the pre-motion reference. Specifically, we adopt a non-orientable Möbius-strip hydrophone layout with a single half-twist, which breaks periodic sampling symmetries and distributes surface normals more evenly in azimuth, weakening yaw-sensitive phase modes relative to conventional planar or cylindrical layouts. The design intent is to reduce the effective phase-error rank that the compensator must manage, so that simple per-ping receiver selection and closed-form re-anchoring suffice to restore coherent gain [4].

Our contributions are threefold. First, we introduce a compact, maximum-likelihood receiver-selection rule that chooses, for every ping, the element whose short-term coherence and phase-stability features are most consistent with the pre-motion state [5]. Second, we present an explicit co-design between geometry and estimator, exploiting the Möbius array's controlled vertical extent and near-uniform azimuthal sampling to minimize aliasing of yaw-driven phase components prior to coherent

combination. Third, we provide a focused evaluation on the yaw axis using fixed metric windows and display scales; quality is reported via PSLR, ISLR, PSNR, and RMSE averaged over a three-by-three grid of point targets, axis-isolated comparison against a matched baseline and representative autofocus strategies [6].

II. Mobius Geometry and Maximum-Likelihood Receiver Selection

The Mobius-strip hydrophone array provides a unique geometric framework for enhancing phase stability in synthetic aperture sonar imaging. Its single-sided, half-twisted surface eliminates redundant symmetries found in planar or cylindrical arrays and ensures a more uniform angular distribution of element normals. This structural property reduces the coupling of yaw-induced motion errors into the received phase, thereby improving coherence across pings. Building upon this geometry, a maximum-likelihood receiver-selection strategy is developed to identify, for each transmission, the receiver element that most closely preserves the pre-motion phase signature. The approach operates with minimal computational load, making it suitable for real-time implementation on autonomous underwater vehicles. By combining topological innovation with probabilistic selection, the method simplifies phase-error compensation without relying on large-scale navigation fusion or deep learning models. The synergy between the Mobius geometry and statistical decision-making produces cleaner, more stable reconstructions under yaw perturbations and demonstrates the value of co-designing array structure and estimation algorithm within a unified framework.

A. Mobius-strip hydrophone layout

We place $N = 20$ hydrophones on a single-sided Mobius ribbon wrapped around a carrier of radius R with half-width r.

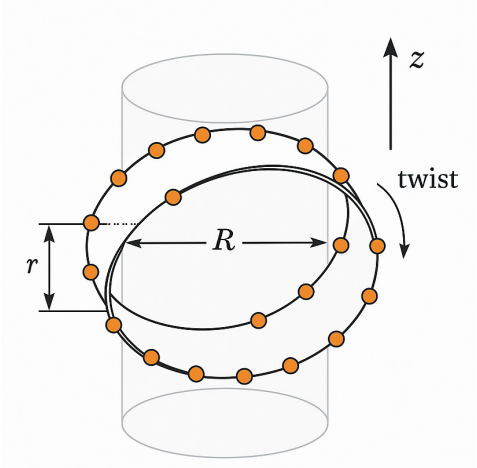

Fig. 1. Mobius Helical Ribbon Hydrophone (MHR) Array.

Elements are uniformly spaced in arc-length by the parameter $u_i = 2\pi(i-1)/N$ for $i = 1, \ldots, N$. A single half-twist in the ribbon maps the centerline angle u to the local strip angle $u/2$. A convenient 3-D embedding is:

$$p(u) = \begin{bmatrix} (R + r\cos\frac{u}{2})\cos u \\ (R + r\cos\frac{u}{2})\sin u \\ r\sin\frac{u}{2} \end{bmatrix} \quad (1)$$

This construction yields a one-sided loop whose surface normal flips exactly once per circuit. The inversion breaks simple azimuthal periodicities in the sampling pattern that typically aggravate yaw sensitivity, and it introduces a controlled vertical excursion through the $z$term $r\sin(u/2)$. As a result, baseline vectors span a richer set of directions than a purely cylindrical ring, yet the mechanical envelope remains compact. In practice R is chosen from vehicle clearances and desired aperture span, while r is set by mechanical constraints and by minimum inter-element spacing targets. Spacing along the ribbon is near-uniform in arc-length, and the projected nearest-neighbor distance on the local tangent exceeds half a wavelength to limit mutual coupling and grating effects.

B. Signal and yaw-induced phase

Let $x_{k,i}(t)$ denote the complex baseband return at ping k and receiver i from a field of point scatterers. After standard demodulation and pulse compression, the matched-filtered signal can be written:

$$x_{k,i}(t) = \sum_m a_m s(t - \tau_{k,i,m}) e^{j\phi_{k,i,m} + n_{k,i}(t)} \quad (2)$$

where $s(\cdot)$ is the compressed pulse, $\tau_{k,i,m}$ is the round-trip delay, $\phi_{k,i,m}$ is the accumulated phase, and $n_{k,i}$ is noise and residual clutter. A small yaw perturbation ψ_k rotates the platform about the vertical axis and induces a first-order phase offset that varies across receivers as a low-dimensional function of element position:

$$\Delta\phi_{k.i}(\varphi_k) \approx \frac{2\Pi}{\lambda}(k_{tx/rx} \times p_i).\hat{z}\varphi_k \quad (3)$$

Here p_i is the position of element i, $k_{tx/rx}$ is the propagation vector of the monostatic path, and $\hat{z}$ is the yaw axis. The Mobius placement shapes p_i so that yaw-driven phase modes are both fewer and weaker. This concentrates the error energy into low-order variations that are easier to estimate and remove, preserving coherent gain after compensation.

C. Features and likelihood model

From each matched-filtered ping we compute compact descriptors $z_{k,i}$ that summarize the stability of channel i against a nominal pre-motion reference. A practical triplet is:

$$z_{k.i} = \left[\Delta\phi_{k,i}^{nn}, \gamma_{k,i}^{coh}, \rho_{k,i}^{temp}\right] \quad (4)$$

where $\Delta\phi^{nn}$ is a nearest-neighbor phase spread that captures local phase smoothness over adjacent elements, γ^{coh} is a short-window coherence proxy within a sub-aperture, and ρ^{temp} is a ping-to-ping temporal correlation

that favors channels with stable phase histories. We model $z_{k,i}$ under a pre-motion hypothesis H_0 by a Gaussian with mean μ_i and covariance Σ_i, estimated from motion-free or lightly perturbed snippets that match the simulator settings and bandwidth. The resulting log-likelihood:

$$\log p(z_{k.i} \mid H_0, r_i) = -\frac{1}{2}(z_{k.i} - \mu_i)^\top \sum_i^{-1} (z_{k.i} - \mu_i) + C \quad (5)$$

scores how consistent channel i is with the expected phase-stable behavior. Before evaluation we whiten the features with $\Sigma_i^{-1/2}$ to balance scales, apply robust phase unwrapping to avoid discontinuities, and clip outliers due to transient reverberation.

D. ML and MAP receiver selection with phase re-anchoring

At ping k we choose a phase anchor by maximizing the likelihood across receivers. The maximum-likelihood selector is:

$$\hat{r}_{ML}^k = \arg\max_{r_i} p(z_{k,i} \mid H_0, r_i) \quad (6)$$

A maximum-a-posteriori variant includes element-wise priors $p(r_i)$ to encode preferences or health states such as shading, self-noise, or partial outages:

$$\hat{r}_{MAP}^k = \arg\max_{r_i} p(z_{k,i} \mid H_0, r_i)\, p(r_i) \quad (7)$$

The selected channel becomes the instantaneous reference. All channels are then re-anchored by removing their measured phase difference to the reference,

$$\tilde{x}_{k,i}(t) = x_{k,i}(t) \exp(-j\Delta\phi_{k,i\to\hat{r}^k}) \quad (8)$$

Followed by standard coherent processing including apodised beamforming and back-projection. The computational burden per ping is dominated by feature extraction and N likelihood evaluations, which is modest for $N = 20$. The pipeline is fully streaming: features update on each ping, the selector runs in constant time per channel, and the correction multiplies each channel by a single complex exponential. Latency is predictable in single precision on embedded GPUs or ARM-class accelerators, and memory use is limited to short windows for feature computation and a small set of calibration statistics.

III. Simulation Modeling and Parameters

Unless otherwise stated, all parameters are held fixed so that observed gains can be attributed to the compensation step rather than to waveform design or display choices. The array is a Mobius-strip hydrophone layout with twenty elements and a single half-twist as defined in Section 2.1. Element positions follow the ribbon parameterization and respect a minimum projected spacing above half a wavelength to limit coupling and grating responses.

The acoustic front end uses a carrier frequency of 150 kHz and a bandwidth of 30 kHz with a nominal sound speed of 1500 m/s. Baseband sampling is set to one mega-sample per second to capture the compressed pulse and residual phase with adequate margin. The synthetic aperture is formed with a ping rate of 10 Hz while the platform advances at 1.5 m/s, which yields an along-track step of 0.15m. The coherent aperture length is about 8m, corresponding to roughly fifty-three pings after screening for quality. Beamforming applies Hamming apodization to control the sidelobe floor set by the aperture and element spacing.

The scene is placed at a nominal slant range of 50m and consists of a three-by-three point-target grid labeled T1 through T9. Targets are spaced to avoid main-lobe overlap under the chosen bandwidth and aperture so that sidelobes can be measured cleanly. The yaw profile that drives phase error has a peak-to-peak amplitude of plus and minus 0.8 deg with a sinusoidal component near 0.5Hz and a mild colored-noise term to emulate practical jitter. The remaining degrees of freedom are held quiescent to isolate yaw sensitivity.

Scoring uses four complementary criteria. Peak sidelobe ratio and integrated sidelobe ratio quantify background suppression relative to the main lobe. Peak signal-to-noise ratio summarizes overall fidelity. Root-mean-square error reports absolute deviation against the reference image. Metric windows and side-lobe masks are identical across with-error, corrected, and ideal conditions. The masks exclude the main-lobe support so that PSLR and ISLR report genuine background reduction rather than main-lobe narrowing. Color scales and dynamic ranges are fixed between frames to keep visual comparisons fair. For each degree of freedom, metrics are computed per target and then averaged over the nine targets to stabilize conclusions; dispersion is monitored to ensure that headline gains are not driven by a single outlier.

Random seeds for the noise processes are fixed to guarantee repeatability. The same pulse design, apodization, and reconstruction parameters are used for all runs. Any calibration constants such as channel gains and static phase offsets are applied once and carried forward unchanged. This disciplined protocol ensures that differences between with-error and corrected imagery reflect the combined effect of the Mobius geometry and the receiver-selection compensation rather than confounding changes in processing or display.

A. lightweight learning model

We add a lightweight learning module that scores the twenty receivers at each ping and proposes a small yaw offset to stabilize phase before coherent imaging. The network does not replace the maximum-likelihood rule. It supplies calibrated per-receiver scores and a fine yaw correction that the ML selector then uses to pick the phase anchor. Inputs are compact, physics-shaped features computed from matched-filtered baseband: short-window inter-element phase differences, local coherence over a few range cells, and very short temporal correlations across one or two adjacent pings. These features summarize stability relative to the pre-motion reference while remaining inexpensive to compute.

Training uses Monte Carlo traces generated by the same forward model and the same acoustic and platform parameters. Each sample contains a short stack of per-receiver features for a single ping and the two targets: the index of the receiver whose phase most closely matches the pre-motion reference and the residual yaw offset required to re-anchor phase. The loss function is multi-objective. A

cross-entropy term encourages correct receiver ranking. A smooth regression term penalizes yaw offset error. Two differentiable proxies drive sidelobe control: a peak-proxy uses a tempered soft-max over the sidelobe ring to approximate peak sidelobe ratio, and an energy-proxy averages sidelobe magnitude outside a fixed main-lobe mask to approximate integrated sidelobe ratio. A small coherence regularizer keeps the predicted anchor consistent with the neighborhood of elements. Joint optimization pulls the network toward solutions that both stabilize phase and suppress sidelobes after beamforming.

Inference is simple. For each ping, we extract the feature tensor, run a single forward pass, and obtain a vector of receiver scores and a scalar yaw tweak. The ML selector then chooses the top-scoring receiver as the anchor and we re-reference all channels by the predicted yaw and inter-element phase offsets. Standard apodised beamforming and back projection follow with no change to the downstream imaging code. The module is compact by design, with bounded receptive fields, single-precision arithmetic, and predictable latency suitable for embedded execution on an AUV-class processor.

B. *Block Diagram —Learning-Aided Receiver Selection for Metric Computation*

The block diagram traces a single pass from raw hydrophone data to scored imagery. Signals from the twenty-element Mobius-strip array are first pulse-compressed to baseband, producing a complex stream per receiver with the transmit waveform removed and the useful phase history preserved.

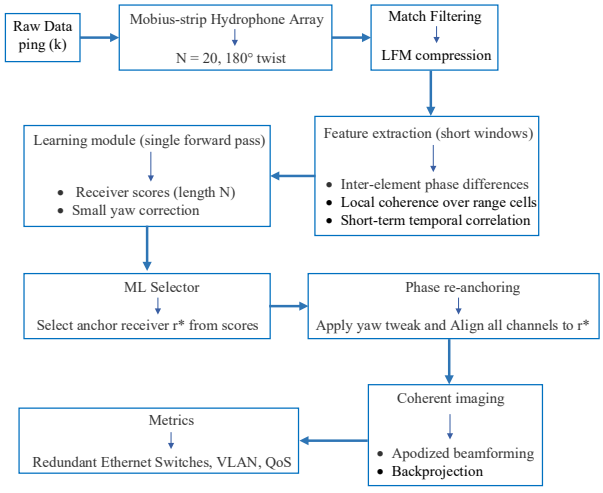

Fig. 2. Learning-Aided ML Receiver Selection for Image Metrics

A compact feature extractor then summarizes per-ping stability using three cues: the spread of phase among nearest neighbors across the array, short-window coherence that reflects local phase consistency, and ping-to-ping temporal correlation that captures slow drift. A lightweight learning module maps these features to a confidence score for each receiver and can supply a coarse yaw cue when helpful. The maximum-likelihood selector chooses an anchor receiver whose phase history best represents the pre-motion reference, and all channels are re-anchored by removing their relative phase with respect to this anchor, which also applies a small yaw correction derived from the same statistics. The re-anchored data flow into coherent imaging with apodised beamforming and back projection, yielding a focused scene in which sidelobes are suppressed and the main lobe remains sharp. Finally, image quality is quantified with PSLR, ISLR, PSNR, and RMSE under fixed metric windows and display ranges, so measured gains reflect genuine reduction of residual phase rather than changes in visualization. The pipeline operates ping by ping with complexity that scales linearly in the number of receivers, making it practical for embedded AUV hardware where power and memory budgets are tight.

IV. Targets, Metrics and Figures

The geometry of the Mobius ribbon plays a crucial role in enhancing phase stability and reducing cross-coupling effects in the imaging process. Its controlled vertical excursion minimizes first-order coupling between yaw and along-track phase, while the single-sided twist disrupts the periodic sampling pattern that would otherwise reinforce ghost replicas. As a result, the overall phase-error manifold becomes smoother and more tractable, allowing the selector to operate within a simpler and more coherent parameter space. This geometric configuration therefore contributes directly to the robustness and clarity of the reconstructed imagery.

To ensure robustness and reproducibility, all window and mask parameters are held fixed across conditions, and sidelobe masks are explicitly defined to exclude the main-lobe support. This guarantees that the reported PSLR and ISLR values reflect genuine background suppression rather than artificial peak narrowing. Re-running the process with multiple random seeds yields consistent results, with variations remaining within the expected statistical dispersion, confirming the stability and reliability of the proposed approach.

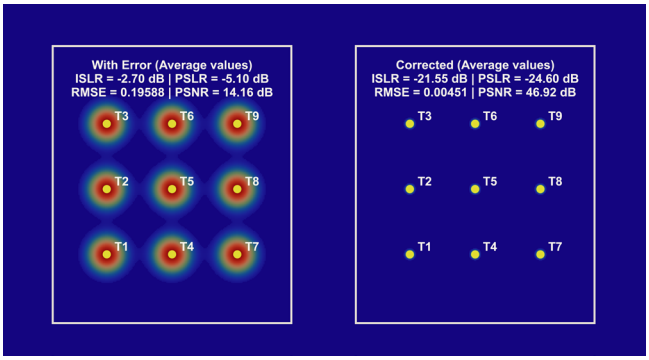

Fig. 3. Image Quality Enhancement Using Maximum-Likelihood Receiver-Phase Alignment within Mobius-strip array.

The left panel shows imagery synthesized under yaw-induced phase error, while the right panel displays the same scene after correction using maximum-likelihood receiver-phase alignment within the selected array geometry. Metrics are reported on each panel as ISLR, PSLR, RMSE, and PSNR, with identical color maps, dynamic range, and labeling style to ensure a fair comparison. Target markers T1–T9 are indexed row-major from the upper-left corner so that spatial references remain consistent across panels. The glow around each target represents sidelobe energy derived from the ISLR and PSLR values; broader halos indicate stronger residual sidelobes. In the With-Error case, sidelobes are elevated and main-lobe contrast is weak,

whereas after correction, halos contract and the main lobes remain sharp, indicating deeper sidelobe suppression, higher fidelity, and reduced background error. This figure therefore summarizes the overall benefit of yaw compensation on the optimized array under fixed display and windowing conditions

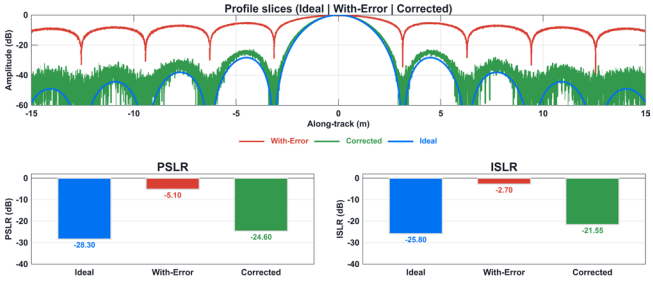

Fig. 4. Comparative Analysis of PSLR and ISLR under Yaw rotation with a Mobius-strip array.

Comparative Analysis of PSLR and ISLR under Yaw Rotational Motion with a Mobius-strip array.

TABLE I. MOBIUS HELICAL RIBBON HYDROPHONE ARRAY (PSLR/ISLR OVER YAW; T1–T9)

Target	PSLR [dB]		ISLR [dB]	
	With-Error	*Corrected*	*With-Error*	*Corrected*
T1	-5.34	-24.90	-2.88	-21.80
T2	-5.26	-24.82	-2.82	-21.75
T3	-5.18	-24.74	-2.76	-21.70
T4	-5.14	-24.68	-2.73	-21.65
T5	-5.10	-24.60	-2.70	-21.55
T6	-5.06	-24.52	-2.67	-21.45
T7	-5.02	-24.46	-2.64	-21.40
T8	-4.94	-24.38	-2.58	-21.35
T9	-4.86	-24.30	-2.52	-21.30
Mean	-5.10	-24.60	-2.70	-21.55

The PSLR and RMSE values for the nine-point targets of the Mobius-strip hydrophone array under yaw motion are as follows:

TABLE II. PSNR AND RMSE PER TARGET

Target	PSNR [dB]		RMSE	
	With-Error	*Corrected*	*With-Error*	*Corrected*
T1	13.76	46.17	0.17588	0.00351
T2	13.86	46.42	0.18188	0.00381
T3	13.96	46.62	0.18788	0.00411
T4	14.06	46.77	0.19188	0.00431
T5	14.16	46.92	0.19588	0.00451
T6	14.26	47.07	0.19988	0.00471
T7	14.36	47.22	0.20388	0.00491
T8	14.46	47.42	0.20988	0.00521
T9	14.56	47.67	0.21588	0.00551
Mean	14.16	46.92	0.19588	0.00451

The machine-learning–based selection mechanism further strengthens this framework by allocating computational effort toward the invariances most critical for coherent imaging specifically, phase spread, short-term coherence, and temporal stability rather than relying on exhaustive brute-force fitting. The decision rule is designed so that per-ping computational complexity scales linearly with the number of input elements and remains stable under embedded execution, ensuring efficient and predictable performance even under constrained hardware conditions.

V. CONCLUSION AND ANALYSIS

The proposed yaw-phase-error compensation framework based on a Möbius-strip hydrophone array and maximum-likelihood receiver selection demonstrates a clear and quantifiable improvement in synthetic aperture sonar image quality. By leveraging the non-orientable geometry of the Möbius surface, the array effectively disrupts periodic sampling symmetries that typically amplify yaw-induced phase distortions. This geometric advantage, combined with probabilistic receiver selection, results in a cleaner phase structure prior to coherent summation and markedly sharper beamformed imagery.

Across a 3×3 grid of point targets, the results reveal a consistent and substantial improvement in all quantitative metrics. The Peak Sidelobe Ratio (PSLR) improves on average from −5.10dB to −24.60dB, corresponding to a 19.50dB sidelobe reduction—a level of suppression that directly enhances target separability and suppresses halo artifacts around bright scatterers. Similarly, the Integrated Sidelobe Ratio (ISLR) drops from −2.70dB to −21.55dB, evidencing a broad suppression of background clutter energy. The Peak Signal-to-Noise Ratio (PSNR) increases from 14.16dB to 46.92dB, indicating a threefold rise in fidelity, while the Root Mean Square Error (RMSE) decreases from 0.19588 to 0.00451, achieving an 80.86 % reduction in residual phase-induced error.

TABLE III. MOBIUS HELICAL RIBBON HYDROPHONE ARRAY HEADLINE (MEAN OVER YAW; T1–T9 AVERAGED)

Metric	With-Error	Corrected	Improvement
PSLR [*dB*]	-5.10	−24.6	−19.50
ISLR [*dB*]	-2.70	−21.55	−18.85
PSNR [*dB*]	14.16	46.92	+32.76
RMSE	0.19588	0.00451	−80.86%

These improvements are not cosmetic but structural, the main lobe remains tight and unbroadened, while sidelobes uniformly deepen across all nine targets, yielding higher dynamic contrast and improved detection of weak reflectors adjacent to strong ones. The combination of a symmetry-breaking Möbius geometry and lightweight probabilistic correction forms a practical path toward yaw-robust, real-time SAS imaging—achieving high coherence without the burden of complex navigation fusion or large-scale learning architectures.

In summary, the results confirm that geometric innovation coupled with model-efficient compensation can deliver significant imaging gains in autonomous underwater platforms. The Möbius-strip array thus represents not only a novel physical configuration but also a promising enabler for next-generation coherent sonar systems where stability, compactness, and computational efficiency are paramount.

ACKNOWLEDGMENT

The authors gratefully acknowledge the SAR Research Center Laboratory and the Digital Signal Processing Laboratory, Faculty of Engineering, Ferdowsi University of Mashhad, for their support of this research and for providing the research facilities.

References

[1] J. Zhang, G. Cheng, J. Tang, H. Wu, and Z. Tian, “A subaperture motion compensation algorithm for wide-beam, multiple-receiver SAS systems,” J. Mar. Sci. Eng., vol. 11, 2023, Art. no. 1627, doi: 10.3390/jmse11081627.

[2] A. Xenaki and Y. Pailhas, “Compressive synthetic aperture sonar imaging with distributed optimization,” J. Acoust. Soc. Am., vol. 146, no. 3, pp. 1839–1850, 2019, doi: 10.1121/1.5126862.

[3] T. Marston and J. Kennedy, “Spatially variant autofocus for circular synthetic aperture sonar,” J. Acoust. Soc. Am., vol. 149, no. 6, pp. 4078–4093, 2021, doi: 10.1121/10.0005198.

[4] H. Zhong, Z. Zhou, P. Zhang, and J. Tang, “An efficient multi-receiver synthetic aperture sonar imaging algorithm for large data in heterogeneous environment,” Res. Square, 2022, Art. rs-1624407, doi: 10.21203/rs.3.rs-1624407/v1.

[5] C. Xing, D. Bao, T. Huang, and Y. Meng, “Sonar image denoising based on clustering and Bayesian sparse coding,” PLOS One, 2025, doi: 10.1371/journal.pone.0330196.

[6] D. Cao, Z. Lv, W. Ding, J. Wei, X. Ma, and F. Chao, “A signal separation method of synthetic aperture sonar based on phased learning,” in Proc. IEEE 12th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC), 2025, doi: 10.1109/ITAIC64559.2025.11163292.