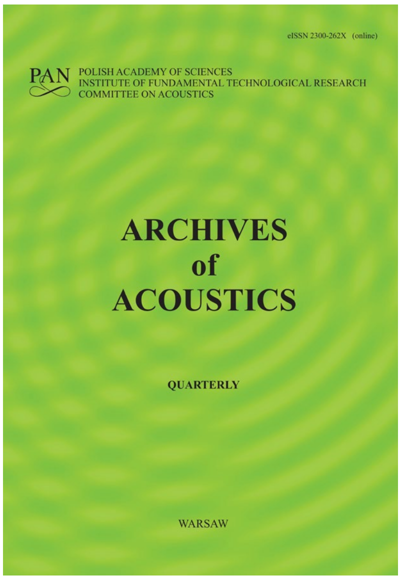
JOURNAL PRE-PROOF



Title: Optimized Hydrophone Array Empowering Deep Learning–Based Phase Error Correction in Synthetic Aperture Sonar

Author(s): Hamid Hajirahimi Kashani, Seyed Alireza Seyedin

DOI: https://doi.org/10.24423/archacoust.2026.4372

Journal: *Archives of Acoustics*

ISSN: 0137-5075, e-ISSN: 2300-262X

Publication status: In press

Received: 2025-11-07

Revised: 2026-04-13

Accepted: 2026-04-21

Published pre-proof: 2026-04-27

Please cite this article as:

Hajirahimi Kashani H., Sayedin S.A. (2026), Optimized Hydrophone Array Empowering Deep Learning–Based Phase Error Correction in Synthetic Aperture Sonar, *Archives of Acoustics*, https://doi.org/10.24423/archacoust.2026.4372

Archives of Acoustics

Optimized Hydrophone Array Empowering Deep Learning–Based Phase Error Correction in Synthetic Aperture Sonar

Hamid Hajirahimi Kashani [1,*], Seyed Alireza Seyedin [2]

[1] https://orcid.org/0009-0008-6255-7249
[2] https://orcid.org/0000-0003-4682-8770
Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
*Corresponding Author e-mail: Hajirahimi.hamid@mail.um.ac.ir

Abstract

This study addresses motion robust synthetic aperture sonar by coupling deep learning-based phase error estimation with hydrophone array optimization under 6 degrees of freedom. The objective is to deliver low sidelobe and high-fidelity imagery when surge, sway, heave, roll, pitch, and yaw inject residual phase errors that conventional autofocus and inertial aiding do not fully remove. The method is physics informed and trains an optimized array to predict residual phase from profile stacks and small image patches using a composite loss that blends image consistency, smooth proxies for peak and integrated sidelobe ratios, and phase smoothness. Array geometry is expressed in closed form and is co tuned in the loop with the corrector so that motion sensitivity is reduced before phase correction. Four hydrophone arrays are examined, namely a planar dual spiral, two helices on a cylinder, a hemispherical Fibonacci layout, and an optimized oblate hemisphere with equatorial densification and a mild tilt. Evaluation spans 24 scenarios across 6 motion axes over a 16 targets grid with PSLR, ISLR, PSNR, and RMSE as metrics. Three dimensional and nearly isotropic layouts consistently outperform planar and cylindrical baselines after correction. The optimized array ranks first overall, typically reaching side-lobe floors near −26 to −27 dB with PSNR near 49 to 50 dB and RMSE near 0.0033 and shows aggregated gains over the planar baseline of about −8 dB in sidelobe metrics and about +12 dB in PSNR, while retaining smaller but measurable advantages over the hemi-spherical Fibonacci design. The contribution unifies physics informed deep correction with interpretable array design that advances motion robust SAS image quality and offers actionable guidance for future hydrophone arrays.

Keywords: Synthetic Aperture Sonar; Deep Learning; Phase-Error Estimation; Motion Compensation; Hydrophone Array; Six Degrees of Freedom.

Acronyms

PSLR – Peak Sidelobe Level Ratio,

ISLR – Integrated Sidelobe Level Ratio,

PSNR – Peak Signal-to-Noise Ratio,

RMSE – Root Mean Square Error,

PDS – Planar Dual-Spiral Array,

CTH – Cylindrical Two-Helix Array,

HF-A – Hemispherical Fibonacci Array,

OPT – Optimized Oblate Hemispherical Array.

1. Introduction

Synthetic aperture sonar (SAS) achieves fine cross-range resolution by coherently integrating echoes acquired along the platform trajectory. The resulting image quality, however, depends critically on phase consistency across pings. Under six-degree-of-freedom (6-DoF) motion, residual translational and rotational errors perturb the two-way propagation path, reduce coherent gain, broaden the mainlobe, and elevate sidelobes. These effects are especially harmful in high-resolution SAS, where even small phase errors can noticeably degrade PSLR, ISLR, and overall point-spread-function (PSF) compactness (Hayes and Gough, 2009; Zhang X. et al., 2022).

The problem becomes more demanding in multi-receiver SAS. Because the receivers occupy different phase centers, the residual phase error is receiver-dependent and spatially variant. Existing approaches, including subaperture motion compensation and multi-receiver image formation, substantially reduce the motion-error budget, yet a non-negligible residual phase term often remains in practice (Zhang J. et al., 2023; Zhang X. et al., 2023). This residual is particularly detrimental when the scene is reconstructed with coherent back-projection, since uncompensated phase errors directly contaminate the complex summation.

Array geometry also plays a fundamental role in motion robustness. The spatial distribution of hydrophones determines angular coverage, beampattern regularity, nominal PSF behavior, and the first-order sensitivity of the received phase to platform perturbations. Recent work on three-dimensional sonar arrays and non-uniform sampling has shown that properly distributed arrays can reduce sidelobes and stabilize imaging performance over wider viewing conditions

(Baron et al., 2021; Reijniers et al., 2020; Zeng et al., 2022; Wu et al., 2018). In particular, Fibonacci-based point sets provide nearly uniform spherical distributions and offer a principled way to generate low-clustering sampling patterns on circular and spherical manifolds (González, 2010; Keinert et al., 2015; Marques et al., 2021).

In parallel, physics-informed deep learning has emerged as a promising strategy for phase-error estimation. Rather than replacing the forward model, such approaches use physics-based structure to regularize the learning problem and improve generalization under incomplete or noisy motion information (Karniadakis et al., 2021; Lee et al., 2021; Feng et al., 2023). However, most existing studies optimize either the reconstruction algorithm or the array geometry in isolation. The coupled design of a motion-robust hydrophone array together with a deep residual phase corrector remains insufficiently explored for SAS.

This paper addresses that gap by jointly studying hydrophone array geometry and physics-informed deep phase-error correction in multi-receiver SAS under 6-DoF motion. Four array geometries are considered: a planar dual-spiral array, a cylindrical two-helix array, a hemispherical Fibonacci array, and a proposed optimized oblate hemispherical array with equatorial densification and mild tilt. The proposed framework learns to estimate residual phase from beam-profile, patch, and signal-snippet features, while the geometry is selected to reduce motion sensitivity before correction.

The main contributions of this work are as follows:

(1) We formulate residual phase-error correction in multi-receiver SAS as a physics-informed learning problem under full 6-DoF motion;

(2) We compare four interpretable hydrophone array geometries within a unified back-projection and correction framework;

(3) We propose an optimized oblate hemispherical design that reshapes hemispherical Fibonacci sampling through vertical compression, equatorial densification, and controlled tilt;

(4) We evaluate the coupled geometry-correction framework over 24 scenarios and 16-target scenes using PSLR, ISLR, PSNR, and RMSE; and

(5) We show that near-isotropic 3-D layouts consistently outperform planar and cylindrical baselines after correction, with the proposed design achieving the best overall robustness.

The remainder of the paper is organized as follows. Section 2 presents the signal model, first-order motion sensitivity, and the array geometries considered in this study. Section 3 describes the physics-informed deep phase-error estimator and the in-the-loop array optimization framework. Section 4 details the proposed optimized array. Section 5 summarizes the simulation setup, motion models, mechanical constraints, and evaluation metrics. Section 6 reports the experimental results and comparative discussion. Section 7 concludes the paper.

2. Signal Model, First-Order Motion Sensitivity, and Baseline Array Geometries

This section presents the signal model adopted for synthetic aperture sonar (SAS) image formation and describes the first-order sensitivity of the received phase to translational and rotational platform motion. It then introduces the baseline array geometries considered in this study, including planar, cylindrical, and hemispherical layouts. Together, these elements establish the physical and geometrical framework used to analyze motion robustness and to compare the corrected imaging performance of the candidate hydrophone arrays. This formulation also provides the basis for understanding why three-dimensional and near-isotropic spatial sampling can improve robustness against residual phase errors.

2.1. Back-Projection Image Formation and First-Order 6-DoF Sensitivity

The back-projection (BP) image is reconstructed by coherently summing delayed and phase-corrected returns from all pings and receiver elements. For a scene point x, each contribution is sampled at the corresponding two-way delay and weighted before summation. BP is adopted here because it remains valid under arbitrary platform motion and near-field geometry, which makes it appropriate for a fair comparison among all arrays and motion axes. Throughout this study, the same BP operator, apodization, and time-gating are used so that the observed differences can be attributed to array geometry and phase correction rather than to changes in image formation.

$$I(x) = \sum_{m} \sum_{i=1}^{N} w_{i,m} \times s_{i,m} \times t \times e^{-j\frac{4\pi}{\lambda}R_{i,m}(x)} \tag{1}$$

This nonparametric imager tolerates near-field effects and arbitrary motion, at the cost of heavier computation. It is also the natural place to apply phase correction because phase enters multiplicatively inside the coherent sum. For consistency, we use the same BP across all arrays and DoFs. Apodization and time-gating are kept fixed to isolate geometry and phase-correction effects. Rotational sensitivity (first-order Jacobian). For a small rotation by angle α, the range sensitivity is:

$$\frac{\partial R}{\partial \alpha} \approx p \times \hat{u} = \begin{bmatrix} p_y \hat{u}_z - p_z \hat{u}_y \\ p_z \hat{u}_x - p_x \hat{u}_z \\ p_x \hat{u}_y - p_y \hat{u}_x \end{bmatrix} \tag{2}$$

Here, $p = [p_x, p_y, p_z]^{\top}$is the sensor position with respect to the phase center and $\hat{u} = [\hat{u}_x, \hat{u}_y, \hat{u}_z]^{\top}$ is the unit look vector. The structure shows why height matters: it creates a lever arm against roll/yaw components, reducing equivalent phase perturbations at broadside. Arrays restricted to $z = 0$ (planar) lack this protection and are more vulnerable to yaw/heave interactions. For receiver i in ping m, the first-order one-way range error to scene point x under small rigid-body offsets is:

$$\hat{u}_{i,m}(x) = \frac{x - p_{i,m}}{\left\| x - p_{i,m} \right\|} \tag{3}$$

$$\Delta R_{i,m}(x) \approx \hat{u}_{i,m}(x)^{\cdot}\ \Delta r_m + \left(p_{i,m} \times \hat{u}_{i,m}(x) \right)^{\cdot}\ \alpha_m \tag{4}$$

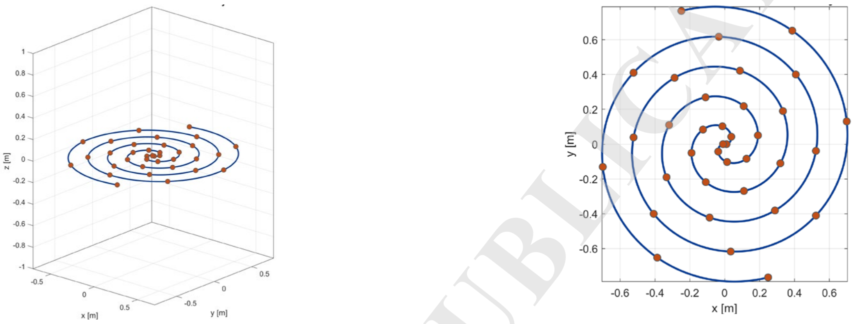
Here, $p = [p_x, p_y, p_z]^{\top}$is the sensor position with respect to the phase center and $\hat{u} = [\hat{u}_x, \hat{u}_y, \hat{u}_z]^{\top}$ is the unit look vector. The structure shows why height matters: it creates a lever arm against roll/yaw components, reducing equivalent phase perturbations at broadside. Arrays restricted to $z = 0$ (planar) lack this protection and are more vulnerable to yaw/heave interactions. Equations (2)–(4) show that the sensitivity of the range history to rigid-body motion depends directly on the receiver coordinates relative to the phase center. Genuine three-dimensional layouts provide vertical leverage against roll, pitch, and yaw components, whereas planar layouts lack this degree of freedom and are therefore more vulnerable to coupled translational and rotational perturbations. This observation motivates the comparison of the candidate array geometries considered below.

2.2. Baseline Array Geometries

All four candidate arrays contain 36 hydrophones. Their coordinates are expressed in a right-handed Cartesian frame (x,y,z), with dimensions in meters and angles in radians. The designs deliberately span planar, cylindrical, hemispherical, and optimized hemispherical layouts so that the roles of vertical extent, angular isotropy, and non-uniform sampling can be isolated. Fibonacci or golden-angle sampling is used where appropriate because it yields low-clustering point distributions and nearly uniform angular coverage on circular or spherical manifolds.

2.2.1. PDS (Planar Dual-Spiral) Array

A strictly planar layout on the z=0 plane built from two Archimedean spiral arms that interleave to fill the disk without clustering. Each arm carries 18 sensors. The sensor angles along each arm follow a Fibonacci progression so that successive points land at approximately irrational angular increments, preventing azimuthal aliasing. The second arm is rotated by π to interleave with the first and equalize tangential spacing. Radial growth is linear with angle, so near the origin the points are closer (higher sampling density) and gradually spread out, which mimics a well-apodised aperture in the plane.



(**a**) (**b**)

Fig. 1. Geometry of a Planar Dual-Spiral Sensor Array; (a): Side View, (b) Top View.

$$r(\theta) = a + b \times \theta \tag{5}$$

$$\theta_k^{(h)} = \theta_{\max} \left\langle \frac{k}{\tau = (1+\sqrt{5})/2} \right\rangle + (h-1)\pi, ak = 0,1,\ldots,17, h \in \{1,2\} \tag{6}$$

$$p_k^{(h)} = \begin{bmatrix} r(\theta_k^{(h)}) \cos\theta_k^{(h)} \\ r(\theta_k^{(h)}) \sin\theta_k^{(h)} \\ 0 \end{bmatrix} \tag{7}$$

That $p_k^{(h)} = [x\ y\ z]^{\mathsf{T}}$ is the sensor position. To better illustrate the geometry, the PDS array specifications are summarized in Table 1 that a is the initial spiral offset that prevents overlap at the origin, b is the spiral pitch governing radial growth per radian, θ_{max} is the maximum angular extent traced by each arm, h is the arm index (1 or 2), k is the within-arm sample index.

Table 1. PDS Array Specification.

Parameter	Value	Notes
Sensors (N)	36 (18 per arm)	Two planar arms, Phase shift π
a	0.01 $[m]$	Spiral offset
b	0.055 $[m/rad]$	Spiral pitch
θ_{max}	4.6π $[rad]$	Angular span
Fibonacci use	$\langle k/\tau \rangle$ in $\theta_k^{(h)}$	Blue-noise 2-D sampling

2.2.2. CTH (Two Helices on a Cylinder) Array

We consider a genuinely three-dimensional arrangement consisting of two counter-phased helices wrapped around a right circular cylinder. Each helix carries 18 sensors. A stratified Fibonacci parameter drives both the azimuthal progression and the vertical coordinate, ensuring that as the sensors wind around the cylinder they also distribute uniformly along height. The second helix is phase-shifted by π and slightly offset in the stratified index to decorrelate the two strands. This layout yields a tall, hollow aperture with balanced coverage around the vehicle and along the vertical axis useful for mitigating sensitivity to sway/roll while providing vertical leverage.

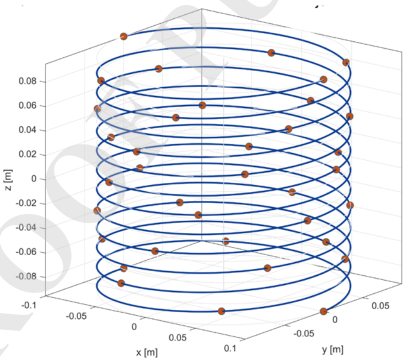


Fig. 2. Geometry of a Two Helices on a Cylinder Sensor Array.

$$s_k^{(h)} = \left\langle \frac{k + q_h}{\tau = (1+\sqrt{5})/2} \right\rangle \rightarrow q_1 = 0, q_2 = 1/2 \tag{8}$$

$$t_k^{(h)} = 2\pi \times T \times s_k^{(h)} + (h-1)\pi \tag{9}$$

$$p_k^{(h)} = \begin{bmatrix} R_c \cos t_k^{(h)} \\ R_c \sin t_k^{(h)} \\ z_k^{(h)} \end{bmatrix} \rightarrow k = 0, 1, \ldots, 17, h \in \{1, 2\} \tag{10}$$

That $p_k^{(h)}$ denotes the position vector of the k-th sensor element on the h-th helix. R_c denotes the cylinder radius, T represents the total number of turns, and $z_k^{(h)}$specifies the axial coordinate of each element along the cylinder. This configuration forms two intertwined helices uniformly distributed around the cylindrical surface, enabling improved angular coverage and denser spatial sampling compared to planar spiral designs.

Table 2. CTH Array Specification.

Parameter	Value	Notes
Sensors (N)	36 (18 per helix)	Two helices, π phase separation
R_c	0.1 $[m]$	Cylinder radius
H	0.19 $[m]$	Cylinder height
T	6	Helix turns
Fibonacci use	$s_k = \langle k + q_h/\tau \rangle$	Azimuth–height stratification

Consider a cylindrical array populated by $\mathrm{N} = 36$ sensors; let $\mathrm{d_{nn}}$ denote the approximate nearest-neighbor spacing between sensors on the lateral surface.

$$d_{nn} \approx \sqrt{\frac{2\pi R_c H}{N}} \approx 5.8cm \tag{11}$$

2.2.3. HF-A (Hemispherical Fibonacci Array)

The hemispherical Fibonacci array (HF-A) distributes N = 36 hydrophones over the upper hemisphere using spherical Fibonacci sampling. Let $\emptyset = (1 + \sqrt{5})/2$ denote the golden ratio. The element coordinates are generated by a golden-angle progression in azimuth together with an equal-area vertical coordinate on the hemisphere:

$$z_i = \frac{i - 0.5}{N} \tag{12}$$

$$\psi_i = \frac{2\Pi i}{\varphi^2} \tag{13}$$

$$p_i = \sqrt{1 - z_i^2} \tag{14}$$

$$p_i = R \begin{bmatrix} p_i \cos(\psi_i) \\ p_i \sin(\psi_i) \\ z_i \end{bmatrix} \tag{15}$$

Where R is the radius of the hemisphere, z_i controls the vertical position, ψ_i is the azimuth angle, ρ_i ensures the points lie on the hemisphere surface. This arrangement provides nearly uniform angular coverage and can be considered as a reference 3D isotropic array configuration.

3. Physics-Informed Deep Phase-Error Estimation

This study casts residual phase correction as a supervised (synthetic) and physics-informed learning problem (Karniadakis et al., 2021). The network receives motion-corrupted evidence (profiles/patches and signal snippets) and predicts a ping/sensor–indexed residual phase, which is then applied before coherent imaging. To stabilize learning and reduce the degrees of freedom, the residual phase is parameterized by a low-order basis with an optional space-variant term.

3.1. Network Inputs, Outputs, Loss, and Optimization

Residual phase correction is formulated here as a supervised synthetic-learning problem with physics-informed regularization. The network receives motion-corrupted evidence in the form of beam-profile stacks, local image patches, and short baseband signal snippets, and predicts ping- and receiver-indexed residual phase terms that are removed prior to coherent imaging. To stabilize learning, the residual phase is parameterized by a low-order physically interpretable basis, with an optional space-variant component for non-rigid residuals.

1) Inputs (features)

For each ping m and element i, we extract:

- A 1-D beam/profile stack (radial magnitude slices through the PSF);
- Small image patches around bright targets;
- Short baseband snippets $x_{i,m}(t)$ near the expected two-way delay.

2) Output (prediction)

The network returns coefficients a_m (rigid-like), b_m (non-rigid), and optional space-variant coefficients c_m over a small angular/range basis. The predicted residual phase is:

$$\hat{\varphi}_{i,m}(t) = \frac{4\pi}{\lambda}\left[\hat{u}^{\boldsymbol{\cdot}}_{i,m}\Delta r_m + (p_{i,m} \times \hat{u}_{i,m})^{\boldsymbol{\cdot}}\,\alpha_m\right] \tag{16}$$

Where λ is wavelength, $\hat{u}_{i,m}$ is unit look vector and $p_{i,m}$ is element position. In the application stage (complex demodulation), the estimated residual phase $\hat{\varphi}_{i,m}$ is removed by multiplying each complex baseband datum by $e^{-j\hat{\varphi}_{i,m}}$ for space-variant terms. This unit-modulus operation preserves amplitude/SNR while re-phasing all pings/sensors prior to coherent imaging. After demodulation, the corrected data sum coherently in back-projection/FFT beamforming, narrowing the main lobe and suppressing sidelobes (improving PSLR/ISLR).

$$\tilde{x}_{i,m}(t,\vartheta,\rho)=x_{i,m}(t)\times e^{-j\hat{\varphi}_{i.m}(t,\vartheta,\rho)} \tag{17}$$

$$\hat{I}=B\left(\left\{\tilde{x}_{i,m}\right\}\right) \tag{18}$$

That $x_{i,m}(t)$ is input baseband, $\tilde{x}_{i,m}(t)$ is corrected baseband, B is imaging operator (back-projection in our study) and $\hat{I}$ is corrected image.

3) Training data (synthetic supervision)

For each scenario we generate paired (I_{ref}, I_{err}) by simulating ideal phase, adding 6-DoF residuals, and forming images. The network sees $F_{i,m}$ from the erroneous case and learns to predict $\hat{\varphi}$ that recovers I_{ref}.

The total loss combines a physics informed data consistency term with decision driven image quality terms such as PSLR, ISLR and task accuracy, ensuring the network respects wave physics while optimizing final performance.

$$\begin{aligned} L = {} & \lambda_{\min 2}^{\;2} + \left\|\hat{I}-I_{ref}\right\| \lambda_{ps} L_{PSLR} + \lambda_{IS} L_{ISLR} \\ & + \lambda_{sm}\left\|\nabla\hat{\varphi}\right\|_2^2 + \lambda_{tv}\sum_{m}\left\|\hat{\varphi}_{0,m}-\hat{\varphi}_{0,m-1}\right\|_1 \end{aligned} \tag{19}$$

Where $\hat{I}$ denotes corrected image, I_{ref} is ideal image, L_{PSLR}, L_{ISLR} are differentiable proxies, $\nabla\hat{\varphi}$ is temporal gradient (smoothness), TV denotes total variation and its term enforces ping-to-ping continuity in the residual-phase estimates, PSLR and ISLR denote the peak sidelobe level ratio and the integrated sidelobe level ratio, which quantify sidelobe behavior in the reconstructed image; img labels the image-fidelity term and sm labels a smoothing regularizer. The total loss is a weighted sum with nonnegative weights $\lambda_{\mathrm{im}}, \lambda_{\mathrm{im}}, \lambda_{\mathrm{PS}}, \lambda_{\mathrm{IS}}, \lambda_{\mathrm{sm}}, \lambda_{\mathrm{tv}} \geq 0$ that tune each component: λ_{im} balances image fidelity, λ_{PS} , λ_{IS} penalize excessive sidelobes, λ_{sm} promotes smooth solutions, and λ_{tv} encourages continuity across pings; together these weights ensure the network respects wave physics while optimizing task performance.

4) PSLR proxy (smooth max on sidelobe set S)

The peak sidelobe level is approximated in a differentiable way using a smooth maximum over the sidelobe set S excluding the main lobe. This yields a stable PSLR surrogate that penalizes the strongest sidelobe while remaining suitable for gradient based optimization. The value can also be normalized by the main lobe peak to form the sidelobe ratio.

$$L_{PSLR}=20\log_{10}\left(\frac{soft\max_{\tau}\left\{\left|h(x)\right|: x\in S\right\}}{\max_{x\in M}\left|h(x)\right|}\right) \tag{20}$$

$$soft\ \max_{\tau}\{a_k\} = \tau \log \sum_k e^{a_k/\tau} \tag{21}$$

Where $h(x)$ denotes PSF magnitude around a bright point, M, S are main/side regions, $w_M,\ w_S$ related to smooth windows and τ is temperature.

5) Optimization

We train the network using the Adam optimizer with learning rate η, θ and early stopping based on validation PSLR/ISLR. The rigid head stabilizes training; the non-rigid head captures small, consistent deviations beyond first-order physics.

$$g_t = \nabla_{\theta} L_{batch}(\theta_t) \tag{22}$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{23}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{24}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{25}$$

$$\theta_{t+1} = \theta_t - \eta_{\theta} \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \tag{26}$$

where L_{batch} denotes the mini-batch version of the total training loss, g_t is the gradient with respect to the network parameters θ_t, and $\beta_1,\ \beta_2$ and ε are the standard Adam hyperparameters.

3.2. In-the-Loop Geometry Optimization

The proposed framework co-designs array geometry and the deep phase corrector so that the selected geometry maximizes corrected-image quality under explicit mechanical constraints. Let $\boldsymbol{\gamma}$ denote the geometry parameters (e.g., $\boldsymbol{k_z}, \boldsymbol{\beta}, \boldsymbol{k}, \boldsymbol{\theta_t}, \boldsymbol{R}$ for OPT), and $\boldsymbol{\theta}$ denote the network parameters. The ideal bilevel objective selects the geometry that yields the highest expected corrected-image score after inner-loop training of the phase estimator.

1) Bilevel objective (Ideal form)

This part in its ideal form separates the optimization into two levels where the inner problem enforces data fidelity and physical constraints while the outer problem optimizes high level performance metrics. The inner problem determines the optimal network parameters for a fixed geometry, whereas the outer problem selects the geometry that maximizes the expected corrected-image quality under the prescribed constraints.

$$\max_{\gamma \in C} E_w \ \square\ \Pi\left[Score(I(\gamma, \theta^*(\gamma); w)) \right] \tag{27}$$

$$\theta^*(\gamma) = \arg\min_{\theta} E_w \ \square \ \Pi\left[L(\gamma, \theta; w)\right] \tag{28}$$

Where, C: feasible set (spacing, envelope, mass, tilt); ω: scenario (array, DoF, motion draw); Π: scenario prior; I: corrected image; L: training loss; $Score$: ranking metric; $\theta^*(\gamma)$: inner optimum.

This expression states that the optimal array design $\boldsymbol{\gamma}$ is chosen within the feasible set $\boldsymbol{C_{max}}$ so as to maximize the expected score of the reconstructed image. The expectation is taken over environmental or scenario variations ω drawn from the distribution Π. Here $\boldsymbol{I}(\boldsymbol{\gamma}, \boldsymbol{\theta}^*(\boldsymbol{\gamma}); \boldsymbol{w})$ denotes the image formed using design $\boldsymbol{\gamma}$ with its corresponding optimal reconstruction parameters $\boldsymbol{\theta}^*(\boldsymbol{\gamma})$ and the function $\boldsymbol{Score}$ evaluates image quality through chosen performance metrics such as PSLR, ISLR, or resolution.

2) Practical alternating optimization

Exact bilevel optimization is costly, so it alternates:

- Inner step: train θ for K epochs at fixed γ via $min_\theta EL$
- Outer step: update γ to increase score using gradients or sample efficient search, followed by a projection onto C: $\Pi_C \to \gamma$.

3) Composite score

The composite score aggregates multiple quality metrics into a single performance index that reflects overall image fidelity. A higher value indicates better sidelobe suppression, sharper resolution, and improved robustness of the array design.

$$Score = \alpha(-PSLR) + \beta(-ISLR) + \gamma PSNR - \delta(10^3 \times RMSE); \alpha, \beta, \gamma, \delta \geq 0 \tag{29}$$

All metrics are computed on Corrected images. we use α=1.0, β=0.7, γ=0.15 and δ=0.2.

4) Gradient estimation

- Differentiable route: the geometry map $p_i^0 = G(i, \gamma)$ is analytic and our proxy losses are differentiable; use back-prop through β or its surrogate to obtain ∂Score/∂γ;
- Derivative-free route: low-dimensional γ (e.g., $[k_z, \beta, k, \theta_t]$, use CMA-ES/Bayesian optimization on batches of scenarios (Hansen, 2016); enforce constraints by projection.

Constraint projection: With minimum spacing d_{min} and envelope E, project by:

$$\Pi_C(\gamma) = \arg\min_{\tilde{\gamma} \in C} \left\|\tilde{\gamma} - \gamma\right\|_2^2 \tag{30}$$

$$C=\left\{\gamma : d_{\min}(G(.;\gamma)\geq d_0, G(.;\gamma)\in\varepsilon, \left|\theta_t\right|\leq\theta_{\max}\right\} \tag{31}$$

Where d_{min} is minimum inter-element spacing induced by γ , d_0 is required spacing, E is size envelope and θ_{max} is tilt limit. This joint design guarantees that the geometry that wins is the one that still wins after DL correction, not merely in an uncorrected or differently corrected regime.

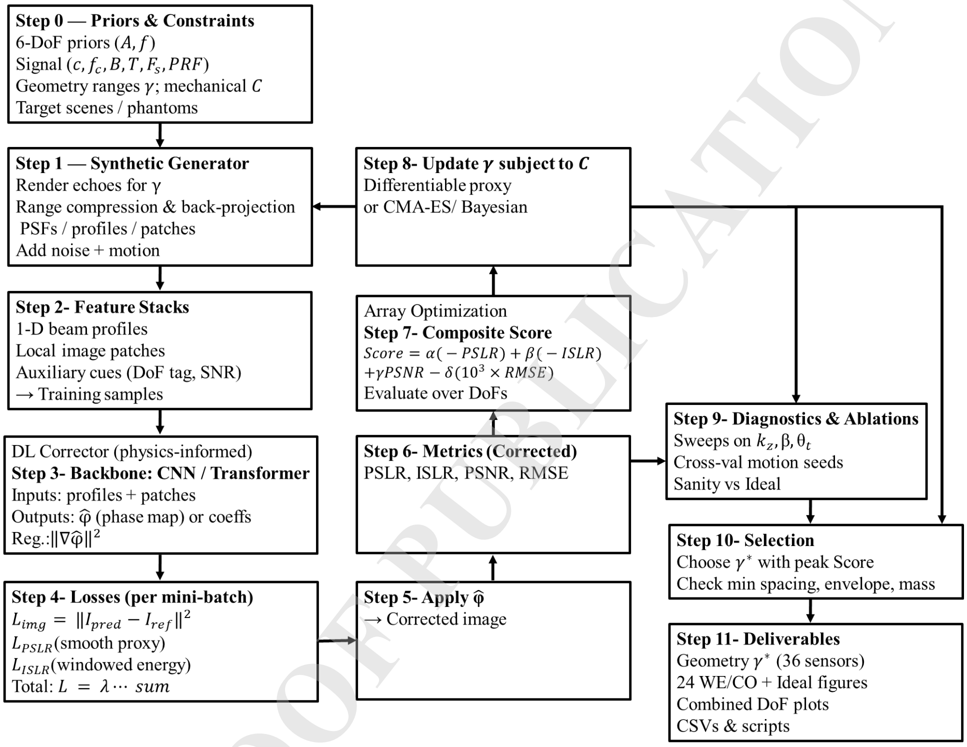


Fig. 3. Deep Learning Driven hydrophone Array optimization: System Flow.

3.3. System-Level Design Relations

Design formulas describe how system parameters affect imaging and they guide experiment setup, ensure adequate sampling, and predict SNR and resolution trends under coherent processing while clarifying where deep learning and geometry provide the greatest benefit, including relations for range resolution, unambiguous range, ping spacing, and time bandwidth product.

$$\delta_r(\gamma) = \frac{c}{2B}, \qquad R_{max}(\gamma) = \frac{c}{2PRF}, \qquad \Delta x = \frac{v}{PRF}, \qquad TBT = B \times T \tag{32}$$

Where $\boldsymbol{\delta_r}$ denotes range resolution; $\boldsymbol{R_{max}}$: unambiguous range; Δx: along-track ping spacing; $\boldsymbol{TBP}$: time–bandwidth product; $\boldsymbol{c}$: sound speed; $\boldsymbol{B}$: bandwidth; T: pulse duration; $\boldsymbol{PRF}$: ping repetition frequency and $\boldsymbol{v}$: platform speed.

Azimuth resolution characterizes the cross-range sharpness of the image and in synthetic aperture sonar it improves as the effective aperture length increases (Hayes, Gough, 2009). At broadside the resolution scales inversely with the synthetic aperture size so longer coherent integration yields narrower beams and finer target separation.

$$\Delta\theta_{az} \approx \frac{\lambda}{2L} \Rightarrow \delta_{az} \approx \frac{\lambda R}{2L} \qquad (33)$$

Where Δ_{az}: beamwidth; δ_{az}: cross-range resolution; λ: wavelength; $L = N_{pings}\Delta x$: synthetic aperture length and R: slant range.

Along track sampling must satisfy a Nyquist like condition to avoid azimuth aliasing in synthetic aperture formation. The ping spacing is therefore chosen small enough relative to the wavelength and geometry so that coherent processing produces an unambiguous and well-focused image.

$$\Delta x \le \frac{\lambda R}{2D_{eff}} \qquad (34)$$

Where $\boldsymbol{D_{eff}}$: effective real-aperture width in azimuth (depends on array's projected extent and apodization). This heuristic mirrors SAS sampling: longer real apertures (larger $\boldsymbol{D_{eff}}$) or shorter ranges demand denser pinging (smaller $\boldsymbol{\Delta x}$).

1) Coherent integration gain (ideal matched filtering)

Design the output SNR increases linearly with the number of coherently combined pings, or equivalently grows by ten times the logarithm of the ping count in decibels. In practice residual phase errors reduce this gain, which motivates the use of deep learning correction and robust array geometries.

$$SNR_{out}^{(dB)} \approx SNR_{single}^{(dB)} + 10\log_{10} N_{ping} \qquad (35)$$

Where $\boldsymbol{N_{ping}}$ denotes coherent pings in the aperture; $\boldsymbol{SNR_{single}}$: per-ping matched-filter SNR. Residual phase error reduces coherent gain, motivating DL correction and robust geometry. Phase-error variance and image blur (rule-of-thumb): It provides a rule of thumb link between motion induced phase noise and image blur. As the variance of phase errors grows beyond a fraction of a radian squared the main lobe widens and sidelobes rise, leading to degraded resolution and contrast in the reconstructed image.

$$E\left[e^{j(\varphi-\hat{\varphi})}\right] \approx e^{-\frac{1}{2}\sigma_{\Delta\varphi}^{2}} \Rightarrow Coherentloss(dB) \approx 10\log_{10}(e^{-\frac{1}{2}\sigma_{\Delta\varphi}^{2}}) \qquad (36)$$

Where $\sigma^2_{\Delta\varphi}$ denotes residual phase variance after correction (smaller is better); This captures how even zero-mean phase jitter broadens the main-lobe and raises sidelobes.

2) Element spacing for grating-lobe avoidance (instantaneous array)

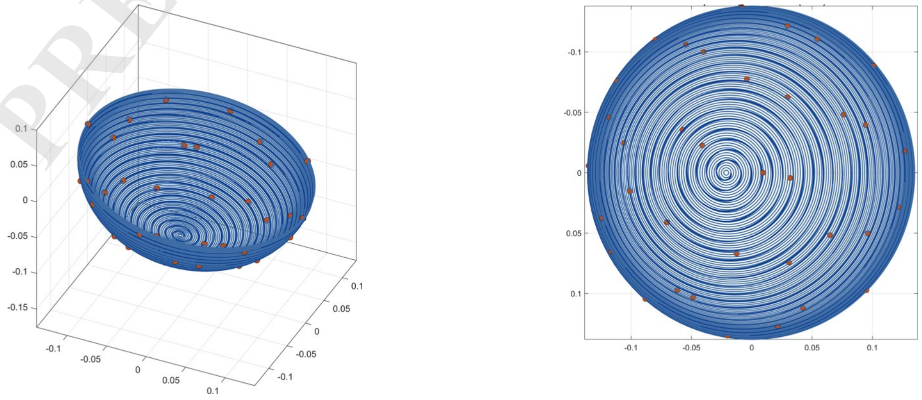
Design It requires that adjacent sensors be separated by less than half the acoustic wavelength. This constraint suppresses spatial aliasing and ensures that the array produces a single main lobe without spurious repetitions in the beampattern.

$$d_{\min} \le \frac{\lambda}{2}(broadside, periodic\, 1-D) \tag{37}$$

Where $\boldsymbol{d_{min}}$ denotes minimum inter-element spacing. For aperiodic/Fibonacci layouts, this becomes a soft limit-irregularity suppresses coherent grating lobes, but we still enforce $\boldsymbol{d_{min}}$ for coupling/mechanics. The system-level relations determine feasible $(\boldsymbol{B}, \boldsymbol{T}, \boldsymbol{PRF}, \boldsymbol{F_s}, \boldsymbol{v})$ and aperture length $\boldsymbol{L}$. Our DL correction improves the coherent gain factor by shrinking $\sigma^2_{\Delta\varphi}$, while our geometry reduces sensitivity Jacobians. Together they produce the observed improvements in PSLR/ISLR/PSNR/RMSE across 24 scenarios.

4. Proposed Optimized Oblate Hemispherical Array

Design This optimized variant starts from the HF-A directions and then reshapes them to better balance vertical sensitivity and equatorial isotropy. First, the vertical coordinate is compressed (oblate), reducing lever arms in **z** that amplify heave-induced phase. Next, points are gently densified near the equator to increase lateral angular sampling where motion blur is most critical. Finally, the cloud is scaled and subjected to a small tilt about the y axis so that the effective aperture is slanted toward the scene, improving along-track conditioning without sacrificing azimuthal uniformity.



(a) (b)

Fig. 4. Geometry of a Optimized Sensor Array: (a) Perspective; (b) Bottom View.

$$(x, y, z) = (\sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, -\cos\theta_i), i = 1, 2, ..., 36 \qquad (38)$$

$$z_1 = k_z z, \quad w = (1 - |z_1|)^{\beta}, \quad f = 1 + kw, \quad x_1 = f \times x\,, \qquad \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \times R_y(\theta_t) \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \qquad (39)$$

$$R_y(\theta_t) = \begin{bmatrix} \cos\theta_t & 0 & \sin\theta_t \\ 0 & 1 & 0 \\ -\sin\theta_t & 0 & \cos\theta_t \end{bmatrix} \qquad (40)$$

Where $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ are the seed direction cosines from HF-A (on the unit sphere), $\boldsymbol{z_1}$ is the ablated vertical coordinate, $\boldsymbol{k_z} \in (\mathbf{0}, \mathbf{1}]$ is the oblate factor (smaller $\boldsymbol{k_z}$, stronger vertical compression), $\boldsymbol{w} \in [\mathbf{0}, \mathbf{1}]$ is an equatorial weight that increases as $|\boldsymbol{z_1}|$ decreases, $\boldsymbol{\beta} > \mathbf{0}$ is the densification exponent shaping how sharply the weight grows near the equator, $\boldsymbol{f} = \mathbf{1} + \boldsymbol{kw}$ is a horizontal gain applied to (x,y), $\boldsymbol{k} \geq \mathbf{0}$ is the densification gain, $(\boldsymbol{x_1}, \boldsymbol{y_1})$ are the horizontally magnified coordinates, $\boldsymbol{R}$ is the physical radius scale, $\boldsymbol{\theta_t}$ is the small tilt angle about the $\boldsymbol{y}$ axis, $\boldsymbol{R_y(\theta_t)}$ is the right-hand rotation matrix and $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ are the final sensor coordinates.

Table 3. OPT array specification.

Parameter	Value	Notes
Sensors (N)	36	HF-A seeds: oblate/densify/tilt
R	0.12 $[m]$	Physical scale
k_z	0.85	Vertical compression
β	0.30	Densification exponent
k	0.15	Densification gain
θ_t	$12^0 \approx 0.209$ $[rad]$	Tilt about y
A_{hemi}	0.0816 $[m^2]$	hemisphere area

The optimized array places $\boldsymbol{N} = \mathbf{36}$ sensors on an oblate, tilted hemisphere of horizontal radius $\boldsymbol{a} = \boldsymbol{R}$ and polar (vertical) radius $\boldsymbol{c} = \boldsymbol{k_z R}$. Tilt $\boldsymbol{\theta_t}$ is a rigid rotation and does not change surface area, hence it does not affect the average spacing. Let $\boldsymbol{d_{nn}} = \mathbf{4.76}\ \boldsymbol{cm}$ denote the approximate nearest-neighbor spacing.

$$d_{nn} \approx \sqrt{\frac{\pi R^2 (1 + \frac{k_z^2}{\sqrt{1-k_z^2}} ar\tanh\sqrt{1-k_z^2}}{N}} \qquad (41)$$

5. Simulation Setup and Evaluation Metrics

This study casts residual phase correction as a supervised (synthetic) and physics-informed learning problem. The network receives motion-corrupted evidence (profiles/patches and signal snippets) and predicts a ping/sensor–indexed residual phase, we evaluate 24 scenarios (4 arrays

× 6 DoFs) in with Error (WE) and Corrected (CO) settings relative to an Ideal reference image. The simulation stack mirrors deployable constraints and allows parameter sweeps. All parameters, layouts, and figures are regenerated by our MATLAB scripts.

5.1. Scenarios

Each array is tested under surge, sway, heave, roll, pitch, and yaw separately to isolate sensitivities. For each DoF, we synthesize motion traces with specified amplitudes and frequencies and add optional colored noise. Metrics are averaged over 16 targets to stabilize conclusions. This factorial design produces 24 with Error (WE) and 24 Corrected (CO) images plus Ideal references. Signal and platform parameters jointly determine the achievable resolution and coverage of the synthetic aperture sonar system. Together they set the fundamental performance limits of coherent imaging.

Table 4. Signal and platform specification.

Parameter	Value	Notes
c	1500 [m/s]	Speed of sound in water
f_c	300 [kHz]	Carrier frequency
λ	5 [mm]	Acoustic wavelength
B	30 [kHz]	Signal bandwidth
δ_r	2.5 [cm]	Range resolution
T	8 [ms]	Pulse duration
TBF	240	Time–bandwidth product
F_s	256 [kb/s]	Sampling rate
PRF	12 [Hz]	Pulse repetition frequency
R_{max}	62.5 [m]	Unambiguous range
v	1.5 [m/s]	Platform speed
Δx	0.125 [m]	Along-track ping spacing
N_{ping}	200	Number of coherent pings
L	25 [m]	Synthetic aperture length

These values remain within practical AUV limits while providing manageable apertures, the PRF balances unambiguous range with platform speed, the TBP of about 240 ensures robust compression, and the settings are intentionally moderate so that the conclusions can be transferred to other systems.

5.2. Motion models (training and eval priors)

Per-DoF sinusoids use representative amplitudes and frequencies ($A_t = 3\ cm\ and\ A_r = 3°$, $f = 0.5\ Hz$). We optionally add colored noise to emulate sea-state variability. Training draws match evaluation priors (domain alignment), but we also test mild distributional shifts to verify generalization. The result is a DL corrector that is robust and fast at inference.

5.3. Mechanical constraints

The minimum inter-element spacing is constrained to d_{min} ≥12mm, corresponding to 2.4λ at 300 kHz. The allowable wet-end envelope is limited to 0.26 m in diameter and 0.22 m in length. For the hemispherical designs (HF-A and OPT), the radius is constrained to R≤0.12 m; for CTH, $R_c = 0.10m$ and H = 0.19m; and for PDS, the footprint does not exceed 0.24m×0.24m. The wet-end mass is limited to 3 kg, and the allowable tilt range is $\theta_t \leq 15°$, while the proposed OPT array uses $\theta_t = 12°$. These constraints reflect realistic nose-cone and harness limits.

5.4. Targets and Metrics

The evaluation covers six motion axes, namely surge, sway, heave, roll, pitch, and yaw. Each axis is treated separately to isolate sensitivity. The evaluation focuses on sidelobe suppression and reconstruction fidelity across these axes. For every array and motion axis, both the with-error (WE) and corrected (CO) images are compared with the ideal reference. Metrics include PSLR and ISLR for sidelobe characterization, and PSNR and RMSE for quantitative fidelity assessment. Four metrics are used:

(1) Peak sidelobe level ratio (PSLR) is defined as:

$$PSLR = 20\log_{10}\left(\frac{\max_{x\in S}|I(x)|}{\max_{x\in M}|I(x)|}\right) \tag{42}$$

where M and S denote the mainlobe and sidelobe regions, respectively.

(2) Integrated sidelobe level ratio (ISLR) is defined as:

$$ISLR = 10\log_{10}\left(\frac{\sum_{x\in S}|I(x)|^2}{\sum_{x\in M}|I(x)|^2}\right) \tag{43}$$

(3) Peak signal-to-noise ratio (PSNR) is defined with respect to the ideal reference image as:

$$PSNR = 10\log_{10}\left(\frac{(\max(I_{ref}))^2}{MSE}\right) \tag{44}$$

where $\max(I_{ref})$ denotes the maximum possible pixel intensity value of the reference image I_{max}.

(4) Root mean square error (RMSE) is defined as:

$$RSME = \sqrt{\frac{1}{N}\sum (I(x) - I_{ref}(x))^2} \tag{45}$$

The same windowing, mainlobe masks, sidelobe masks, and pixel support are used across all arrays and all scenarios to ensure a fair comparison. Consistent windows and pixel masks are applied across arrays to maintain fairness.

The optimized oblate hemisphere (OPT) consistently outperforms all alternatives, achieving the lowest sidelobe levels (PSLR ≈ −26.5 dB) and the highest PSNR (> 49 dB). The hemispherical Fibonacci array (HF-A) follows closely, typically 1–2 dB behind, confirming its stable three-dimensional isotropy. The cylindrical twin-helix (CTH) delivers moderate improvements over the planar dual-spiral (PDS), benefiting from its limited vertical aperture that mitigates motion-induced blur. The PDS configuration remains the weakest due to its purely planar geometry, exhibiting elevated sidelobes and higher RMSE.

Overall, the results demonstrate that equatorial densification and oblate shaping in the OPT configuration yield the most robust performance, validating the consistency between the simulation figures and their corresponding analytical interpretation.

5.5. Reporting Protocol

The main comparative discussion is presented in Section 6, where aggregate ranking and per-axis behavior are analyzed systematically. Appendix A contains supplementary montage figures and detailed per-axis tables for visual completeness.

6. Results and Discussion

This section compares the four array geometries after deep-learning-based residual phase correction. Across the 24 evaluated scenarios, a stable ranking emerges: the proposed optimized oblate hemispherical array (OPT) performs best overall, followed by the hemispherical Fibonacci array (HF-A), the cylindrical two-helix array (CTH), and the planar dual-spiral array (PDS). This ordering is consistent across sidelobe-oriented metrics (PSLR and ISLR) and fidelity-oriented metrics (PSNR and RMSE), indicating that genuine three-dimensional and near-isotropic sampling reduces motion sensitivity before correction and enables more effective phase compensation after correction.

The goal is to identify the superior geometry overall and to characterize the pattern of degradation by motion axis after deep learning-based phase correction. All statements refer to the corrected state and therefore reflect the joint effect of array geometry and the learned estimator applied before coherent imaging. The prose is continuous and formal and includes a small number of precise numerical instances only where they sharpen the comparison.

The optimized oblate hemisphere named OPT is the most robust geometry in aggregate. A modest vertical flattening combined with equatorial densification and a small controlled tilt distributes sensing directions in 3D and lowers first order sensitivity of phase to translations and rotations. After correction OPT repeatedly produces deep sidelobes with high image fidelity. Typical corrected peak sidelobe ratios fall near −26 to −27 dB while peak signal to noise ratios approach 50 dB and root mean squared error remains close to 3.3×10^{-3} . The hemispherical Fibonacci array named HF-A is the second most robust design. Its near isotropic 3D coverage achieves very similar sidelobe floors with slightly lower fidelity than OPT. The cylindrical two helices array named CTH ranks in the middle. Its vertical extent mitigates height coupled errors but its azimuthal sampling is less uniform than the hemispherical families. The planar dual spiral named PDS is consistently the weakest because it lacks height and therefore couples more strongly to yaw and heave which leaves larger residuals after correction.

Across all motion axes, three-dimensional layouts with near-isotropic angular coverage deliver the best corrected imagery. The optimized oblate hemisphere is consistently first, the hemispherical Fibonacci is a strong second, the cylindrical two helices is intermediate, and the planar dual spiral trails. The pattern holds because height and angular uniformity reduce phase-error sensitivity before learning-based correction is applied.

In surge, the optimized array preserves coherent gain most effectively and achieves very high fidelity, for example PSNR 49.56 dB with RMSE 0.00333, while the hemispherical Fibonacci follows closely and the cylindrical array remains clearly ahead of the planar layout. Sway shows the same ordering as it stresses azimuthal regularity, where equatorial densification in the optimized design yields deeper sidelobes and cleaner backgrounds than planar sampling. Heave produces the sharpest separation between three-dimensional and planar layouts. The planar array falls to PSLR −14.61 dB and PSNR 33.02 dB, whereas the optimized array reaches PSLR −23.11 dB and PSNR 45.56 dB, illustrating the value of real height. Roll pushes the top two toward a near-saturation regime. The optimized array attains PSLR −27.14 dB, ISLR −25.02 dB, PSNR 49.55 dB, and RMSE 0.00332, with the hemispherical Fibonacci very close and the cylindrical array still noticeably better than the planar layout. Pitch mirrors roll with slightly wider gaps in favor of the optimized design. Yaw emphasizes angular uniformity most strongly, where the optimized array secures about a two-decibel PSLR advantage over the hemispherical Fibonacci, for example −24.31 dB versus −22.31 dB, while the cylindrical array improves over planar but does not close the gap.

Viewed in aggregate, the optimized array's corrected gains over the planar baseline are large and consistent. Typical averages show about −8 dB in PSLR, about −8 dB in ISLR, roughly +12 dB in PSNR, and an absolute RMSE drop near 0.012, with smaller but steady advantages over the cylindrical geometry and modest yet measurable gains over the hemispherical Fibonacci. These outcomes confirm that the best image quality is obtained when geometry first lowers first-order sensitivities and deep learning then removes the remaining residual phase.

Taken across all motions a stable hierarchy emerges. Arrays that offer true height and near isotropic angular coverage yield the lowest corrected sidelobes and the highest corrected fidelity. The optimized oblate hemisphere and the hemispherical Fibonacci form the top tier. Their separation is modest in motions where sidelobes approach the floor and somewhat larger in heave and yaw where OPT gains from reduced height sensitivity and smoother azimuthal density. The cylindrical array forms a reliable middle tier and consistently outperforms the planar layout because height alone secures a substantial part of the robustness. The planar layout is the most sensitive in yaw heave and sway because residual phase aligns more strongly with the main lobe in the absence of vertical diversity. The learned corrector improves every case but delivers the largest gains when the geometry already lowers first order sensitivity. This explains why the optimized array simultaneously produces the deepest sidelobes and the highest fidelity. When performance nears the sidelobe floor improvements become incremental and are bounded by apodization and by small unmodeled dynamics. Below a short numerical summary quantifies typical corrected improvements of OPT relative to other geometries. Values are averaged over the 6 degrees of freedom.

Table 5. Averaged Corrected PSLR/ISLR/ PSNR/ RMSE Advantage of OPT.

Array	PSLR [dB]	ISLR [dB]	PSNR [dB]	RMSE
PDS	−8.5	−8.08	12.18	−0.0124
CTH	−6	−5.96	8.14	−0.0063
HF-A	−1.55	−1.58	2.28	−0.0013

For PSLR and ISLR, more negative values are better; therefore, a negative Δ indicates that OPT achieves deeper sidelobe suppression than the baseline. For PSNR, larger values are better; therefore, a positive Δ indicates higher reconstruction fidelity. For RMSE, smaller values are better; therefore, a negative Δ indicates lower error.

This study fixes the hydrophone count at 36 and uses one acoustic setup near 300 kHz with 30 kHz bandwidth and nominal sound speed 1500 m/s, motion priors are largely single-axis with mild colored noise. Image formation assumes a homogeneous water column with limited multipath and gentle bathymetry, and navigation/calibration imperfections such as residual jitter, clock drift, mutual coupling, and gain/phase bias are only partially modeled. The deep network is trained on simulator-matched distributions, so a domain gap may appear under out-of-distribution sea states, clutter, or hardware latencies. The sidelobe floor is bounded by fixed apodization, minimum spacing, and mechanical envelope; thus near-floor results reflect practical limits as much as algorithmic ones. The composite score uses fixed weights for PSLR, ISLR, PSNR, and RMSE, alternative weightings could slightly reshuffle close rankings without altering the overall hierarchy.

Recommended next steps include compact pool and sea trials to validate generalization under real navigation errors, reverberation, and range-dependent sound speed; a differentiable end-to-end pipeline to jointly optimize geometry, apodization, motion priors, and the DL corrector under explicit mechanical constraints, training with domain adaptation, uncertainty estimation, and robustness to calibration drift and mutual coupling; extension to multi-receiver SAS and to mixed-axis motions.

6.1. Overall Ranking and Aggregate Comparison

Table 5 summarizes the averaged corrected advantage of OPT relative to the other geometries. Relative to PDS, OPT improves PSLR and ISLR by approximately 8 dB, increases PSNR by about 12 dB, and reduces RMSE by roughly 0.012. Relative to CTH, the gains remain substantial, whereas the margin over HF-A is smaller but still consistent. These results show that hemispherical sampling already provides strong robustness, and that the additional vertical compression, equatorial densification, and mild tilt introduced in OPT produce a measurable further benefit.

From a physical standpoint, the ranking follows the degree of three-dimensional leverage and angular regularity offered by each array. PDS is most vulnerable because it lacks height altogether. CTH benefits from vertical extent but remains less isotropic than the hemispherical families. HF-A provides strong near-isotropic three-dimensional coverage and therefore serves as a robust baseline. OPT further improves that baseline by reshaping the sampling density and reducing sensitivity in the most critical directions.

6.2. Axis-Specific Behavior

1) Surge

Under surge motion, OPT achieves the best corrected result, reaching PSLR = -26.50 dB, ISLR = -24.08 dB, PSNR = 49.50 dB, and RMSE = 0.00332. HF-A follows closely, whereas CTH remains clearly ahead of PDS but behind the hemispherical designs. This behavior is consistent with the expectation that three-dimensional and nearly isotropic layouts better preserve coherent gain under translational perturbation.

2) Sway

Sway motion again preserves the same ordering, with OPT providing the lowest sidelobes and the highest PSNR. Relative to PDS, the improvement is particularly clear, which indicates that lateral motion strongly rewards uniform azimuthal coverage. The advantage of OPT over HF-A remains moderate but consistent, suggesting that equatorial densification further improves stability under lateral perturbations.

3) Heave

Heave produces the sharpest separation between truly three-dimensional arrays and the planar baseline. PDS drops to PSLR = -14.61 dB and PSNR = 33.02 dB, whereas OPT reaches PSLR = -23.11 dB and PSNR = 45.56 dB. HF-A also performs strongly, but OPT retains a noticeable advantage, which supports the design rationale of combining vertical compression with equatorial densification.

4) Roll

Under roll motion, both OPT and HF-A approach a saturation regime in sidelobe performance. OPT attains PSLR = -27.14 dB, ISLR = -25.02 dB, PSNR = 49.55 dB, and RMSE = 0.00332, while HF-A remains very close. This suggests that once sufficient three-dimensional leverage and angular isotropy are present, further improvements become incremental and are partly limited by the sidelobe floor.

5) Pitch

Pitch behavior parallels roll but preserves a clearer ranking among the four arrays. OPT reaches the best overall corrected performance, followed by HF-A, CTH, and PDS. The result indicates that balanced spatial support in both the horizontal and vertical directions remain important when the platform experiences rotational motion about the lateral axis.

6) Yaw

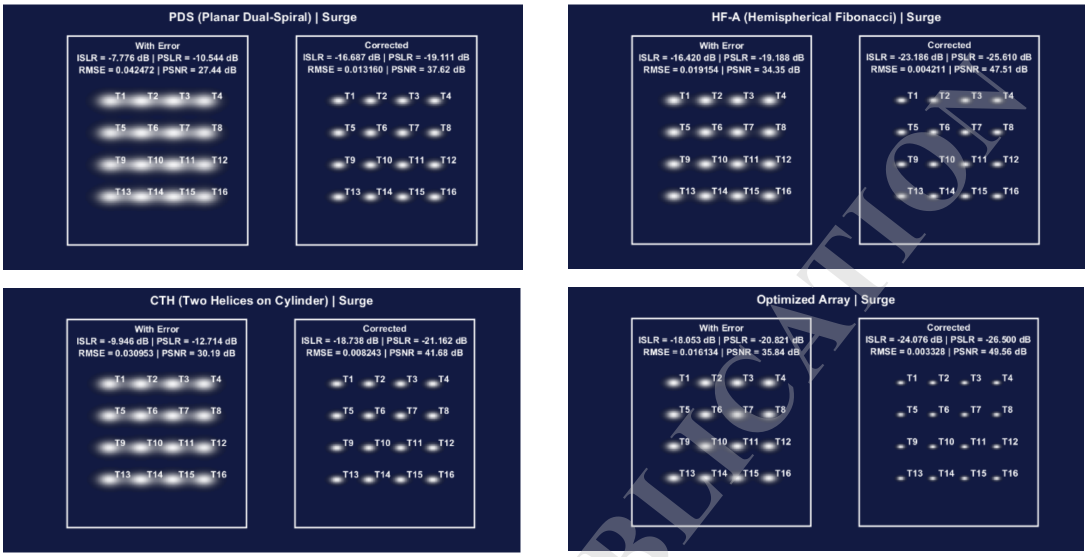
Yaw emphasizes angular regularity more strongly than the other axes. OPT achieves PSLR = -24.31 dB and PSNR = 46.75 dB, compared with PSLR = -22.31 dB and PSNR = 44.11 dB for HF-A. CTH improves over PDS but does not close the gap to the hemispherical layouts. This confirms that azimuthal isotropy is especially valuable when the platform rotates about the vertical axis.

7. Conclusion

This paper investigated the joint effect of hydrophone array geometry and physics-informed deep residual phase correction in multi-receiver synthetic aperture sonar under six-degree-of-freedom motion. Across 24 scenarios, genuine three-dimensional layouts consistently outperformed the planar baseline after correction. The proposed optimized oblate hemispherical array achieved the best overall ranking, combining deep sidelobe suppression with high image fidelity. The hemispherical Fibonacci array ranked second and confirmed the value of near-isotropic three-dimensional coverage, whereas the cylindrical two-helix design provided an intermediate but still meaningful improvement over the planar case. These results show that geometry and phase correction should be designed jointly rather than sequentially. A favorable three-dimensional geometry reduces first-order motion sensitivity before correction, and the physics-informed network then removes the remaining residual phase. The strongest benefits appear in heave and yaw, where vertical extent and angular uniformity are most valuable. Future work should validate the framework with pool and sea-trial data, incorporate calibration uncertainty and mutual coupling more explicitly, and extend the optimization to mixed-axis motion priors and end-to-end differentiable design.

Appendix A. Supplementary Per-Axis Figures and Tables

This appendix provides supplementary montage figures and detailed per-axis corrected metrics supporting the discussion in Section 6. Metrics emphasize sidelobe suppression (PSLR/ISLR) and fidelity (PSNR/RMSE). All integrals/sums use consistent windows and pixel masks across arrays and the optimized oblate hemisphere (OPT) achieves the strongest sidelobe suppression and the highest fidelity, indicating near-saturation of the corrected PSF. HF-A is a close second, trailing OPT by 1–1.5 dB in PSLR/ISLR and 2 dB in PSNR, with only a modest RMSE gap. CTH sits mid-pack, improving 2 dB in PSLR and ISLR over PDS and gaining 4 dB in PSNR, reflecting the benefit of vertical extent for along track blur. PDS remains the weakest due to its purely planar aperture, which couples more strongly to surge-induced phase drift. The 7–8 dB advantage of OPT over PDS in both PSLR and ISLR underscores the combined effect of equatorial densification and mild tilt in stabilizing the main lobe. As PSLR approaches −26.5 dB, marginal returns diminish, explaining the smaller OPT–HF-A gap.



(a) (b)

Fig. 5. Surge-Direction Image Quality Comparison across Arrays: (a) PDS and CTH array; (b) HF-A array and Optimized array.

Overall, surge behavior corroborates the Jacobian analysis, 3-D layouts mitigate translation coupling more effectively than planar designs.

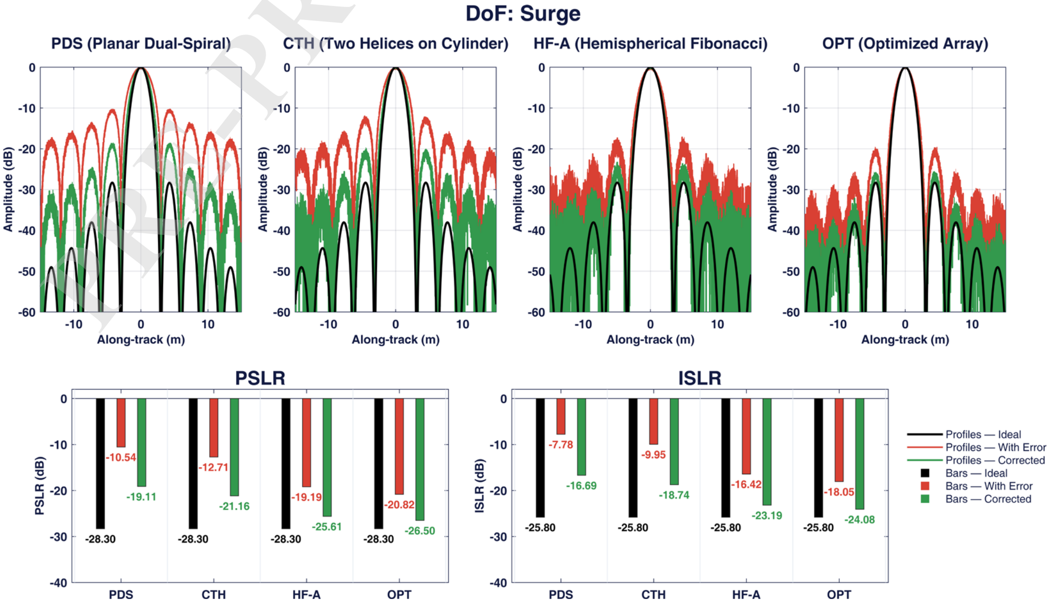
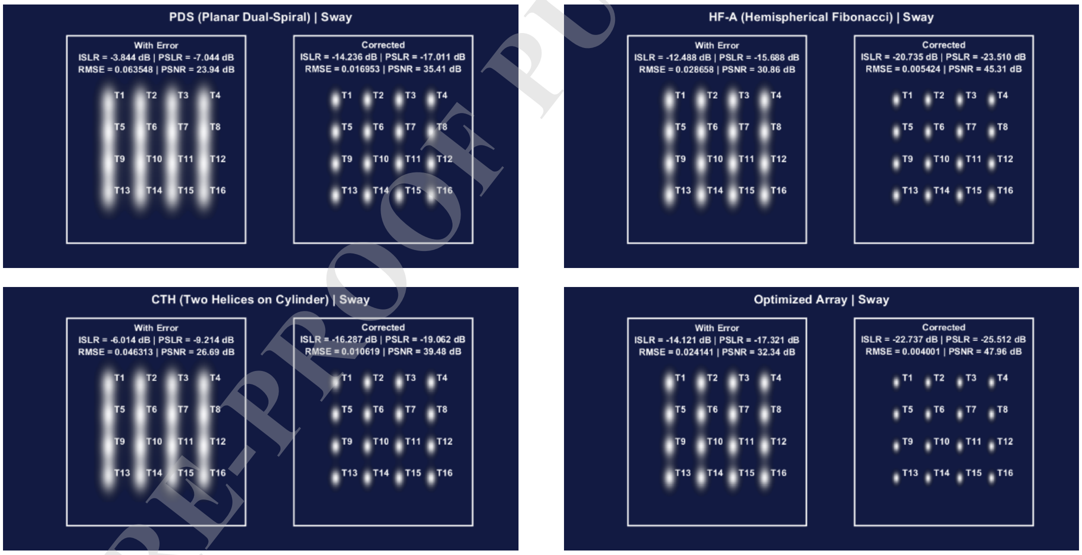


Fig. 6. Surge Motion: Comparative PSLR/ISLR of Arrays.

Table 6. Surge — Corrected Metrics per Array.

Array	PSLR [dB]	ISLR [dB]	PSNR [dB]	RMSE
PDS	−19.11	−16.69	37.6	0.01316
CTH	−21.16	−18.74	41.6	0.00824
HF-A	−25.61	−23.19	47.5	0.00421
OPT	−26.5	−24.08	49.5	0.00332

Under sway, OPT again leads with the most negative sidelobe metrics and top PSNR, while HF-A trails by 1.8–2 dB in PSLR/ISLR and 2.6 dB in PSNR. CTH provides a consistent, moderate improvement relative to PDS (2 dB better PSLR/ISLR and 4 dB PSNR), highlighting that cylindrical height partially offsets lateral translation errors. PDS exhibits elevated sidelobes and lower PSNR, consistent with its lack of z-diversity and less uniform azimuthal sampling. The 8.5 dB PSLR/ISLR margin of OPT versus PDS is the largest among the four arrays here, mirroring the value of angular isotropy against lateral motion. RMSE trends mirror PSNR, with OPT attaining the lowest error (4.0×10^{-3}). These results emphasize that lateral translation stresses azimuthal sampling; hemispherical families (HF-A/OPT) are inherently better conditioned.



(a) (b)

Fig. 7. Sway-Direction Image Quality Comparison across Arrays: (a) PDS and CTH array; (b) HF-A array and Optimized array.

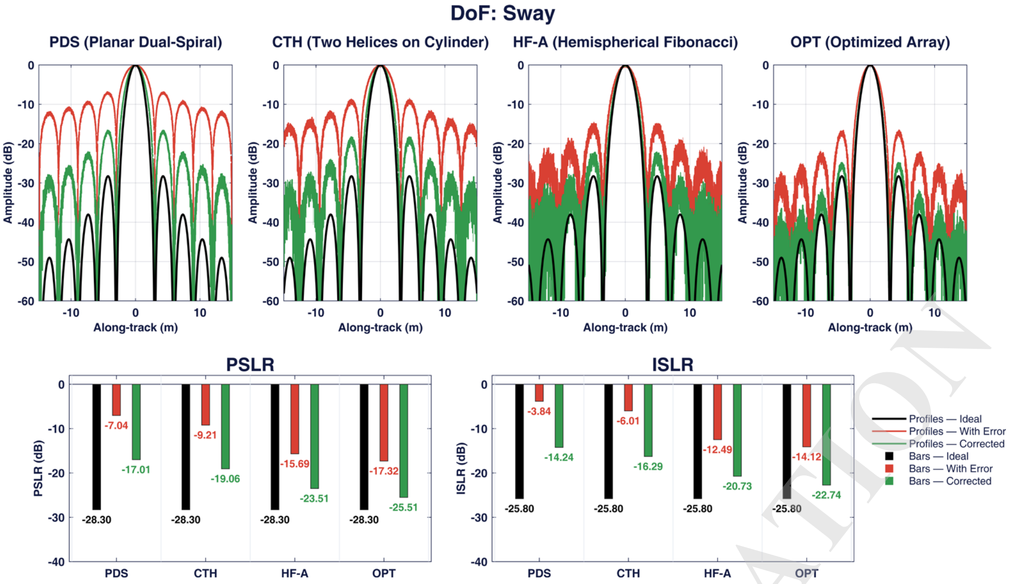
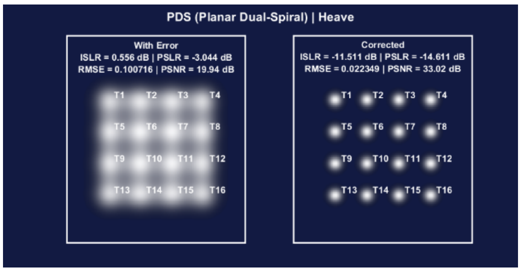


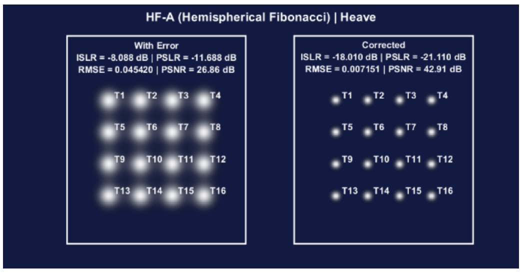
Fig. 8. Sway Motion: Comparative PSLR/ISLR of Arrays.

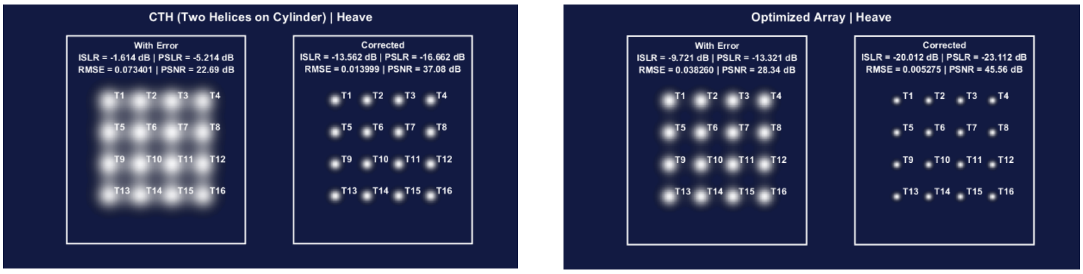
Table 7. Sway — Corrected Metrics per Array.

Array	PSLR [dB]	ISLR [dB]	PSNR [dB]	RMSE
PDS	−17.01	−14.24	35.4	0.01695
CTH	−19.06	−16.29	39.4	0.01061
HF-A	−23.51	−20.73	45.3	0.00542
OPT	−25.51	−22.74	47.9	0.00400

Heave presents the sharpest separation between 3-D and planar designs: OPT and HF-A deliver the lowest sidelobes and highest PSNR, while PDS lags by 8.5 dB in PSLR/ISLR and 12.5 dB in PSNR versus OPT. CTH benefits from vertical extent but still trails HF-A due to less uniform azimuthal coverage. The RMSE hierarchy (OPT < HF-A < CTH < PDS) indicates robust phase recovery where vertical leverage is available. Heave-induced path-length errors couple strongly to height; thus, oblate shaping ($k_z < 1$) plus equatorial densification in OPT reduces sensitivity and tightens the corrected main lobe. The 2.4 dB PSLR gap between OPT and HF-A highlights the added value of tilt and density shaping beyond a pure hemispherical Fibonacci distribution. Overall, heave results most clearly validate the geometric sensitivity arguments and the benefit of DL correction applied on a favorable geometry.







(a) (b)

Fig. 9. Heave-Direction Image Quality Comparison across Arrays: (a) PDS and CTH array; (b) HF-A array and Optimized array.

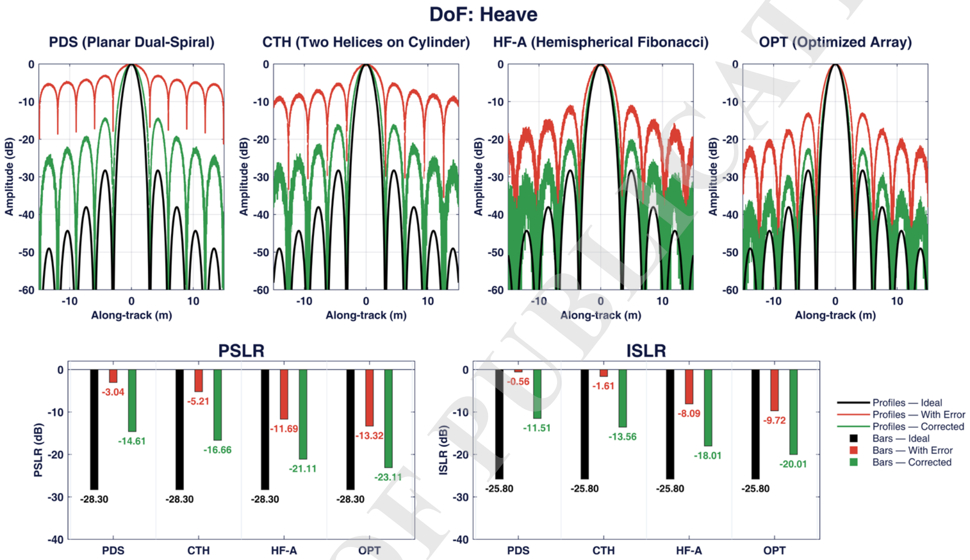
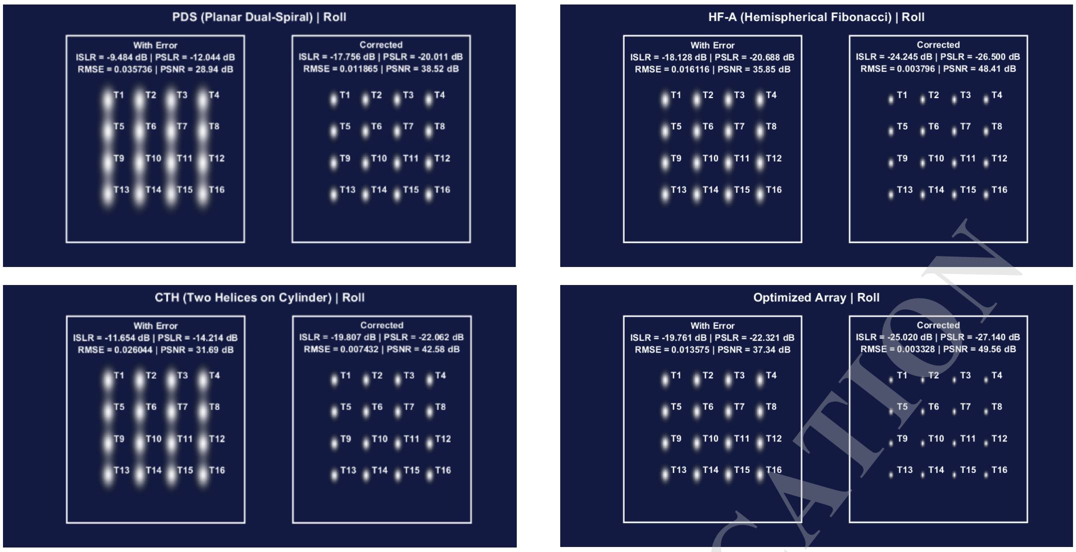


Fig. 10. Heave Motion: Comparative PSLR/ISLR of Arrays.

Table 8. Heave — Corrected Metrics per Array.

Array	PSLR [dB]	ISLR [dB]	PSNR [dB]	RMSE
PDS	− 14.61	−11.51	33.02	0.02234
CTH	−16.66	−13.56	37.08	0.01399
HF-A	−21.11	−18.01	42.91	0.00715
OPT	−23.11	−20.01	45.56	0.00527

For roll, both OPT and HF-A reach the PSLR floor and the same ISLR, indicating a saturation regime where sidelobes approach the apodization limit. OPT retains a small advantage in PSNR and RMSE, but differences are practically minor. CTH shows meaningful gains over PDS (2.05 dB in PSLR/ISLR and 4.06 dB in PSNR), reflecting the stabilizing role of height under rotational perturbations. PDS again trails due to lack of vertical extent, with higher residual sidelobes and larger RMSE after correction. The saturation seen in HF-A and OPT suggests that, past a geometry threshold, DL correction is limited by sidelobe floors rather than phase estimation error. These findings confirm that roll primarily rewards 3-D lever arms and isotropy; once those are present, improvements converge.



(a) (b)

Fig. 11. Roll-Direction Image Quality Comparison across Arrays: (a) PDS and CTH array; (b) HF-A array and Optimized array.

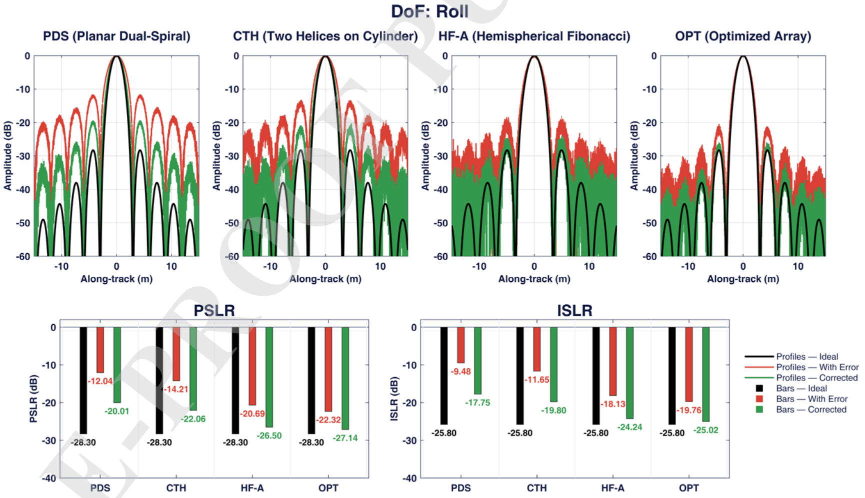


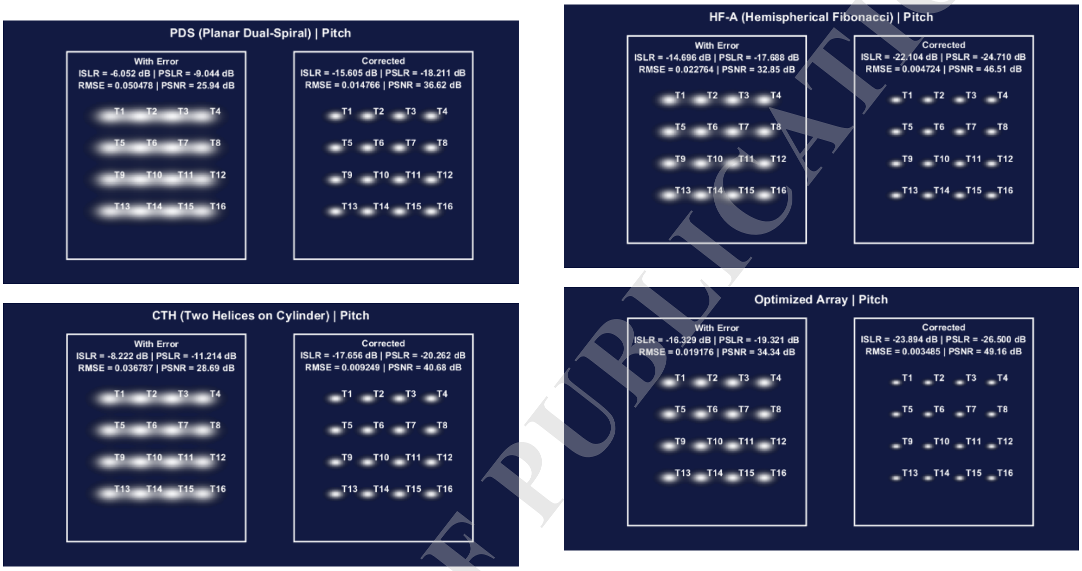
Fig. 12. Roll Motion: Comparative PSLR/ISLR of Arrays.

Table 9. Roll — Corrected Metrics per Array.

Array	PSLR [dB]	ISLR [dB]	PSNR [dB]	RMSE
PDS	−20.01	−17.75	38.51	0.01186
CTH	−22.06	−19.80	42.57	0.00743
HF-A	−26.50	−24.24	48.41	0.00379
OPT	−27.14	−25.02	49.55	0.00332

Pitch behavior parallels roll but with clearer ranking: OPT ≥ HF-A > CTH > PDS across all metrics. OPT reaches the PSLR cap (−26.5 dB) and pairs it with superior ISLR and PSNR

(49.16 dB), reflecting exceptionally clean corrected PSFs. HF-A trails narrowly (1.8 dB in PSLR/ISLR and 2.6 dB in PSNR), indicating that equatorial densification and tilt provide incremental benefit beyond a standard hemisphere. CTH’s intermediate performance shows vertical extent helps, but azimuthal non-uniformity leaves some residual phase sensitivity. PDS remains the weakest, consistent with pitch’s reliance on balanced x-z distribution. RMSE ordering matches PSLR/ISLR, reinforcing the consistency of the metrics. In practice, pitch robustness is a key rationale for adopting 3-D arrays with controlled vertical profiles.



(a)

(b)

Fig. 13. Pitch-Direction Image Quality Comparison across Arrays: (a) PDS and CTH array; (b) HF-A array and Optimized array.

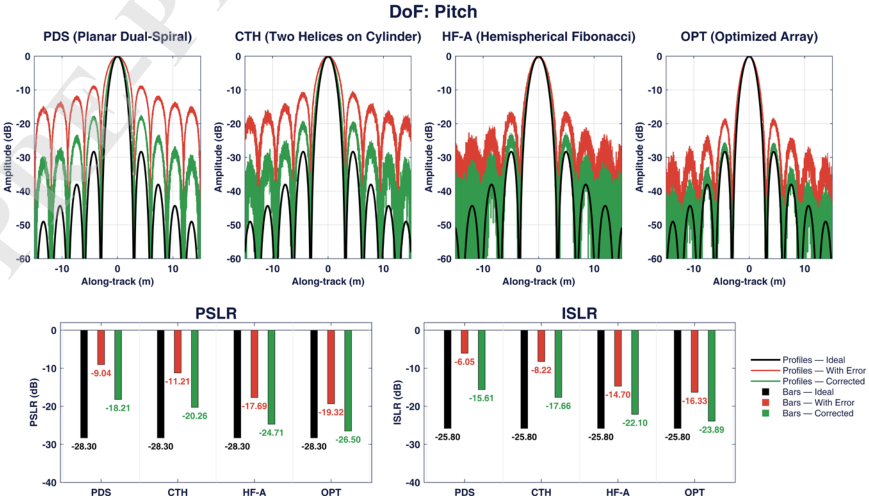
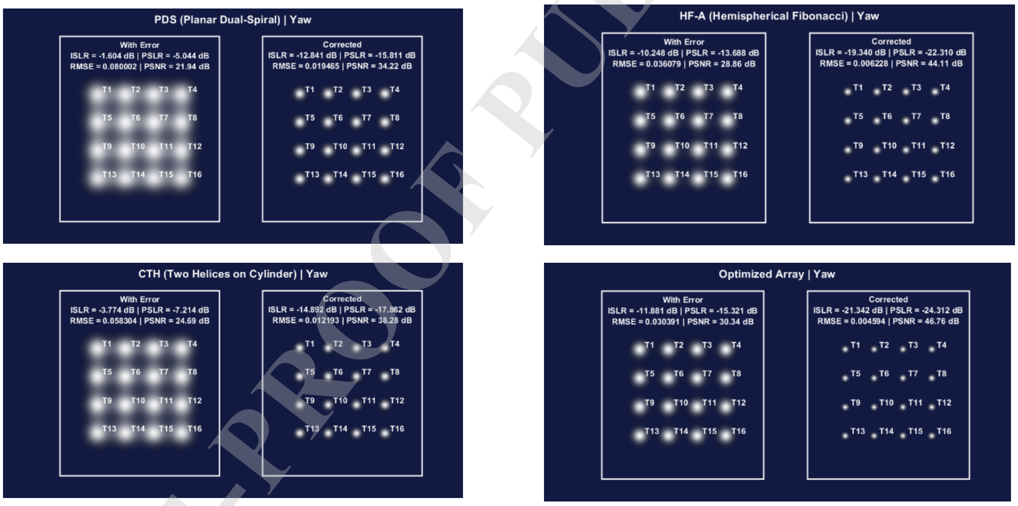


Fig. 14. Pitch Motion: Comparative PSLR/ISLR of Arrays.

Table 10. Pitch — Corrected Metrics per Array.

Array	PSLR [dB]	ISLR [dB]	PSNR [dB]	RMSE
PDS	−18.21	−15.60	36.61	0.01476
CTH	−20.26	−17.65	40.67	0.00924
HF-A	−24.71	−22.10	46.51	0.00472
OPT	−26.50	−23.89	49.15	0.00348

Yaw clearly rewards angular uniformity: OPT and HF-A beat CTH and PDS by large margins in PSLR/ISLR and PSNR. Compared with PDS, OPT gains 8.5 dB in PSLR/ISLR and 12.5 dB in PSNR, with HF-A close behind, showing that equatorial coverage dominates rotational stability about z. CTH's added height helps less, since yaw mainly probes azimuthal sampling regularity, while planar PDS shows the highest sidelobes and RMSE, evidencing strong phase aliasing under rotation. The 2 dB PSLR/ISLR and 2.6 dB PSNR gap between OPT and HF-A is consistent with densification and tilt further smoothing the azimuthal response, and yaw thus offers the clearest evidence that phyllotactic 3-D layouts with DL correction yield robust, low-sidelobe images under rotational motion.



(a) (b)

Fig. 15. Yaw-Direction Image Quality Comparison across Arrays: (a) PDS and CTH array; (b) HF-A array and Optimized array.

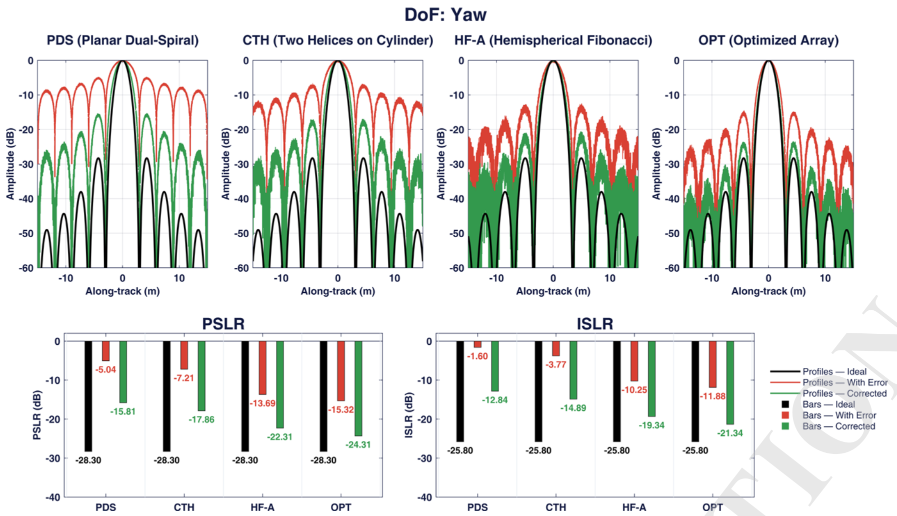


Fig. 16. Yaw Motion: Comparative PSLR/ISLR of Arrays.

Table 11. Yaw — Corrected Metrics per Array.

Array	PSLR [dB]	ISLR [dB]	PSNR [dB]	RMSE
PDS	−15.81	−12.84	34.21	0.01946
CTH	−17.86	−14.89	38.27	0.01219
HF-A	−22.31	−19.34	44.11	0.00622
OPT	−24.31	−21.34	46.75	0.00459

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' contributions

Hamid Hajirahimi Kashani (conceptualized the study, designed the hydrophone arrays, and wrote the original draft). Professor Seyed Alireza Seyedin (supervised the research, contributed to methodology and analysis, and reviewed and edited the manuscript). All authors reviewed and approved the final manuscript.

Acknowledgments

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. Baron V. et al. (2021), Improving performance of three-dimensional imaging sonars through optimally distributed arrays, IEEE Journal of Oceanic Engineering, 46(2): 555–563, https://doi.org/10.1109/JOE.2020.3004018.

2. Chen X., Wan M., Xing M., Sun G.-C. (2022), Azimuth-variant motion error compensation algorithm for airborne SAR imaging based on Doppler adjustment, IEEE Geoscience and Remote Sensing Letters, 19: 1–5, https://doi.org/10.1109/LGRS.2021.3137903.

3. Feng J., Gao K., Wu G., Xu Y., Jiang H. (2023), A deep learning-based InSAR framework for abnormal displacement deformation prediction of bridges, Remote Sensing, 15(5): 1220, https://doi.org/10.3390/rs15051220.

4. Hansen N. (2016), The CMA evolution strategy: A tutorial, arXiv, arXiv:1604.00772, https://arxiv.org/abs/1604.00772.

5. Hayes M.P., Gough P.T. (2009), Synthetic aperture sonar: A review of current status, IEEE Journal of Oceanic Engineering, 34(3): 207–224, https://doi.org/10.1109/JOE.2009.2020931.

6. Karniadakis G.E. et al. (2021), Physics-informed machine learning, Nature Reviews Physics, 3: 422–440, https://doi.org/10.1038/s42254-021-00314-5.

7. Kongsberg Maritime (n.d.), HISAS-1030 interferometric synthetic aperture sonar, Product brochure, available at: https://www.kongsberg.com, accessed 15 March 2026.

8. Lee H., Jung C.-S., Kim K.-W. (2021), Feature preserving autofocus algorithm for phase error correction of SAR images, Sensors, 21(21): 6986, https://doi.org/10.3390/s21216986.

9. Reijniers J., Kerstens R., Steckel J. (2020), An optimized spatial sampling strategy for wide-view planar array 3D sonar sensors, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 67(6): 1236–1241, https://doi.org/10.1109/TUFFC.2020.2964991.

10. Wu H., Tang J., Zhong H. (2018), A correction approach for the inclined array of hydrophones in synthetic aperture sonar, Sensors, 18(7): 2338,

https://doi.org/10.3390/s18072338.

11. Zeng S., Fan W., Du X. (2022), Three-dimensional imaging of circular-array synthetic aperture sonar for unmanned surface vehicle, Sensors, 22(10): 3797, https://doi.org/10.3390/s22103797.

12. Zhang J. et al. (2023), A subaperture motion compensation algorithm for wide-beam, multiple-receiver SAS systems, Journal of Marine Science and Engineering, 11(8): 1627, https://doi.org/10.3390/jmse11081627.

13. Zhang X. et al. (2022), Wide-bandwidth signal-based multireceiver SAS imagery using extended chirp scaling algorithm, IET Radar, Sonar & Navigation, 16(4): 531–541, https://doi.org/10.1049/rsn2.12200.

14. Zhang X. et al. (2023), Multireceiver SAS imagery with generalized PCA, IEEE Geoscience and Remote Sensing Letters, 20: 1502205, https://doi.org/10.1109/LGRS.2023.3286180.

15. Zhong, H.; Zhou, Z.; Zhang, P.; Tang, J. An Efficient Multi-Receiver Synthetic Aperture Sonar Imaging Algorithm for Large Data in Heterogeneous Environment. Res. Sq. 2022, Preprint, rs-1624407. https://doi.org/10.21203/rs.3.rs-1624407/v1.

16. González Á. (2010), Measurement of Areas on a Sphere Using Fibonacci and Latitude–Longitude Lattices, Mathematical Geosciences, 42(1): 49–64, https://doi.org/10.1007/s11004-009-9257-x.

17. Keinert B., Innmann M., Sänger M., Stamminger M. (2015), Spherical Fibonacci Mapping, ACM Transactions on Graphics, 34(6): 193, https://doi.org/10.1145/2816795.2818131.

18. Marques R., Bouville C., Bouatouch K., Blat J. (2021), Extensible Spherical Fibonacci Grids, IEEE Transactions on Visualization and Computer Graphics, 27(4): 2341–2354, https://doi.org/10.1109/TVCG.2019.2952131.