



Available online at [www.qu.edu.iq/journalcm](http://www.qu.edu.iq/journalcm)  
**JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS**  
 ISSN:2521-3504(online) ISSN:2074-0204(print)



# Rigorous Error Quantification and Adaptive Sampling Diagnostics for Physics-Informed Neural Networks

**Bilal Ahmed Shihab\***

Holy Karbala Directorate General of Education, Karbala, Iraq

Email: [bilal\\_ahmed@karbala.edu.iq](mailto:bilal_ahmed@karbala.edu.iq)

## ARTICLE INFO

### Article history:

Received: 17 /06/2026

Revised: 07 /08/2026

Accepted: 09 /08/2026

Available online: 30 /09/2026

### Keywords:

Physics-informed neural networks

Residual-based adaptive sampling

Generalization error

Conditional stability

Quadrature error

Partial differential equations

Scientific machine learning

## ABSTRACT

Physics-informed neural networks (PINNs) approximate partial differential equations by embedding differential operators and auxiliary conditions in the training objective. Their reliability cannot be inferred from the final training loss alone because approximation, optimization, sampling, quadrature, and PDE stability errors interact. This study develops a stability-based framework that separates empirical residual losses from continuous residual norms, introduces explicit quadrature defects, and derives a sampling-aware residual-to-solution estimate under a stated weighted stability assumption. The framework is examined for Poisson, Heat, Wave, and Stokes equations. Archived single-run experiments compare fixed collocation sampling with residual-based adaptive sampling and analyze sensitivity to collocation-set size and network depth. Residual-based adaptive sampling reduces the independent-set error for Heat and Wave by 25.29% and 35.90%, respectively, but increases it for Poisson and Stokes, confirming that adaptive enrichment is problem dependent. The results also show non-monotone behavior with respect to training-point count and depth. The revised validation protocol distinguishes numerical test error from continuous-domain generalization error and identifies the replication, cost, and reporting requirements needed for defensible PINN assessment.

MSC 2020: 65M70; 65N35; 68T07

<https://doi.org/10.29304/jqcm.2026.18.33006>

## 1. Introduction

Physics-informed neural networks (PINNs) approximate the unknown solution of a differential equation with a neural network whose parameters are trained by minimizing residuals of the governing equation, boundary conditions, initial conditions, and available data. This mesh-free formulation supports forward and inverse problems and has become a major approach in scientific machine learning. The original PINN framework combined automatic differentiation with neural approximation for nonlinear partial differential equations (PDEs) [1]. Subsequent studies established approximation and generalization estimates for several PDE classes [2-5] and identified important limitations related to optimization, gradient imbalance, and spectral bias [6-8]. DeepXDE also provides a widely used computational framework for implementing PINNs and related differential-equation solvers [11].

A small empirical training loss does not by itself guarantee a small solution error. The neural hypothesis class may not represent the target solution accurately. The optimizer may stop far from the best element of that class. A residual evaluated at finitely many points may underestimate the continuous residual norm. In addition, a small residual may still produce a large solution error when the underlying PDE or inverse problem is weakly stable. These mechanisms must be separated before a PINN prediction can be treated as a reliable numerical approximation.

Adaptive sampling reduces the mismatch between empirical and continuous residuals by adding collocation points in regions where the current residual is large. Large benchmark studies show that sampling design can strongly

\*Corresponding author

Email address: [Billal\\_ahmed@karbala.edu.iq](mailto:Billal_ahmed@karbala.edu.iq)

Communicated by: For office use

affect PINN accuracy [9], while failure-informed methods interpret residual concentration as an a posteriori indicator [10]. However, adaptive sampling also changes the measure represented by the training set. Without importance weighting or a carefully controlled target measure, the empirical objective may become biased toward a small part of the domain.

This paper has four objectives. First, it presents a clear decomposition of approximation, optimization, quadrature, and stability effects. Second, it derives a sampling-aware residual-to-solution estimate under an explicit conditional-stability assumption. Third, it compares fixed collocation sampling with residual-based adaptive sampling for Poisson, Heat, Wave, and Stokes problems. Fourth, it defines a reproducible validation protocol that distinguishes numerical test error from continuous-domain generalization error. The contribution is methodological rather than a universal unconditional theorem. Its validity depends on the stated PDE stability inequality, the selected norms, the regularity assumptions, and controlled quadrature defects.

## 2. Mathematical Formulation

### 2.1 Abstract Pde Problem

Let  $\Omega \subset \mathbb{R}^d$  be a bounded domain with boundary  $\Gamma = \partial\Omega$ . For time-dependent problems, let  $Q = \Omega \times (0, T]$ . The exact solution  $u$  satisfies

$$\begin{aligned}\mathcal{N}[u] &= f \quad \text{in } Q, \\ \mathcal{B}[u] &= g \quad \text{on } \Gamma \times (0, T],\end{aligned}$$

and, when applicable,

$$\mathcal{I}[u] = u_0 \quad \text{on } \Omega \times \{0\}.$$

The neural approximation is denoted by  $u_\theta$ , where  $\theta \in \mathbb{R}^P$  contains all trainable parameters. Define

$$\begin{aligned}r_\Omega(\mathbf{z}; \theta) &= \mathcal{N}[u_\theta](\mathbf{z}) - f(\mathbf{z}), \\ r_\Gamma(\mathbf{z}; \theta) &= \mathcal{B}[u_\theta](\mathbf{z}) - g(\mathbf{z}), \\ r_0(\mathbf{x}; \theta) &= \mathcal{I}[u_\theta](\mathbf{x}) - u_0(\mathbf{x}).\end{aligned}$$

### 2.2 Continuous and Empirical Residual Objectives

The continuous weighted residual objective is

$$\mathcal{J}(\theta) = \sum_{s \in \{\Omega, \Gamma, 0\}} \lambda_s \|r_s(\cdot; \theta)\|_{L^2(\mu_s)}^2,$$

where  $\lambda_s > 0$  and each  $\mu_s$  is a normalized target measure. Its empirical counterpart is

$$\hat{\mathcal{J}}_N(\theta) = \sum_{s \in \{\Omega, \Gamma, 0\}} \lambda_s \hat{L}_s(\theta), \quad \hat{L}_s(\theta) = \frac{1}{N_s} \sum_{i=1}^{N_s} |r_s(\mathbf{z}_{s,i}; \theta)|^2.$$

The weights are part of the mathematical objective. They control the relative contribution and gradient scale of the PDE, boundary, and initial-condition terms. Poorly balanced weights can cause severe training imbalance [6,7].

### 2.3 Error Decomposition

Let  $\mathcal{H}_p$  denote the neural hypothesis class. The total numerical error can be interpreted through the schematic decomposition

$$E_{\text{total}} \lesssim E_{\text{app}} + E_{\text{opt}} + E_{\text{quad}} + E_{\text{stab}}.$$

The approximation error is

$$E_{\text{app}} = \inf_{v \in \mathcal{H}_p} \|u - v\|_X.$$

The optimization error measures the gap between the trained parameters and the empirical minimizer:

$$E_{\text{opt}} = \hat{\mathcal{J}}_N(\theta_{\text{train}}) - \inf_{\theta} \hat{\mathcal{J}}_N(\theta).$$

The quadrature error measures the difference between continuous residual norms and their empirical approximations. The stability term describes how residual perturbations propagate to the solution norm. PDE analysis determines this final term; neural-network training alone does not.

## 3. Sampling-Aware Residual-To-Solution Estimate

### 3.1 Weighted Stability Assumption

Assume that the exact solution  $u$  and an admissible approximation  $v$  satisfy

$$\|u - v\|_X \leq C_{\text{stab}} \left( \sum_{s \in \{\Omega, \Gamma, 0\}} \lambda_s^{1/2} \|r_s(v)\|_{Y_s} \right)^\alpha,$$

where  $C_{\text{stab}} > 0$ ,  $0 < \alpha \leq 1$ , and  $Y_s = L^2(\mu_s)$  in the present formulation. For well-posed forward problems,  $\alpha$  is often equal to one. Conditional inverse or unique-continuation problems may satisfy  $0 < \alpha < 1$ .

### 3.2 QUADRATURE ASSUMPTIONS

Suppose that the empirical squared residuals satisfy

$$\left| \|r_s\|_{L^2(\mu_s)}^2 - \hat{L}_s \right| \leq \delta_s, \quad s \in \{\Omega, \Gamma, 0\},$$

where  $\delta_s \geq 0$  is the quadrature defect associated with residual component  $s$ .

### 3.3 PROPOSITION 1

**Proposition 1.** Under the weighted stability and quadrature assumptions, the trained PINN satisfies

$$\|u - u_\theta\|_X \leq C_{\text{stab}} \left[ \sum_{s \in \{\Omega, \Gamma, 0\}} \lambda_s^{1/2} (\hat{L}_s + \delta_s)^{1/2} \right]^\alpha.$$

**Proof.** For every residual component,

$$\|r_s\|_{L^2(\mu_s)}^2 \leq \hat{L}_s + \delta_s,$$

and therefore

$$\|r_s\|_{L^2(\mu_s)} \leq (\hat{L}_s + \delta_s)^{1/2}.$$

Substitution into the weighted stability estimate with  $v = u_\theta$  gives the stated bound.  $\square$

The proposition explains why a squared empirical loss cannot be inserted directly into a residual-norm stability estimate without taking square roots. It also shows that reducing the training loss is insufficient when quadrature defects remain large.

### 3.4 IMPORTANCE-WEIGHTED ADAPTIVE SAMPLING

Let adaptive points be drawn from a density  $q_s$ , while the target residual norm is defined by a density  $p_s$ . The importance-weighted estimator is

$$\hat{L}_s^{\text{IW}} = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{p_s(\mathbf{z}_{s,i})}{q_s(\mathbf{z}_{s,i})} |r_s(\mathbf{z}_{s,i}; \theta)|^2, \quad \mathbf{z}_{s,i} \sim q_s.$$

When  $q_s > 0$  wherever  $p_s > 0$ , this estimator targets the original continuous residual norm rather than the adaptive sampling measure. A coverage-preserving mixture density is

$$q_s^{(k)}(\mathbf{z}) = (1 - \eta)p_s(\mathbf{z}) + \eta \frac{(|r_s(\mathbf{z}; \theta_k)| + \varepsilon)^\beta}{\int (|r_s(\mathbf{y}; \theta_k)| + \varepsilon)^\beta d\mathbf{y}},$$

with  $0 < \eta < 1$ ,  $\beta > 0$ , and  $\varepsilon > 0$ . The first component maintains global coverage, while the second concentrates points in difficult regions.

## 4. REPRESENTATIVE PDE SYSTEMS

### 4.1 POISSON EQUATION

For the elliptic problem

$$-\Delta u = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega,$$

a manufactured solution is differentiated analytically to define  $f$  and  $g$ . Validation should report both  $L^2(\Omega)$  and  $H^1(\Omega)$  errors because a small pointwise error may coexist with an inaccurate gradient field.

### 4.2 HEAT EQUATION

The parabolic problem is

$$\frac{\partial u}{\partial t} - \kappa \Delta u = f \quad \text{in } \Omega \times (0, T],$$

with prescribed boundary and initial data. The validation protocol should report errors at several time levels and the space-time relative norm

$$E_{L^2(Q)} = \left( \frac{\sum_{m=1}^{N_{\text{test}}} |u_\theta(\mathbf{z}_m) - u(\mathbf{z}_m)|^2}{\sum_{m=1}^{N_{\text{test}}} |u(\mathbf{z}_m)|^2} \right)^{1/2}.$$

### 4.3 WAVE EQUATION

The hyperbolic model is

$$\frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = f.$$

Oscillatory modes are often learned slowly by standard multilayer perceptrons. Fourier features, sinusoidal activations, multiscale architectures, or adaptive activation functions [12] therefore provide useful ablations. The energy error is

$$E_{\text{energy}}(t) = \frac{1}{2} \int_{\Omega} (|\partial_t(u_{\theta} - u)|^2 + c^2 |\nabla(u_{\theta} - u)|^2) d\mathbf{x}.$$

#### 4.4 STOKES EQUATIONS

For incompressible flow,

$$-\nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \quad \nabla \cdot \mathbf{u} = 0.$$

Pressure is determined only up to an additive constant, so the formulation must impose a gauge such as

$$\int_{\Omega} p d\mathbf{x} = 0.$$

Separate velocity, pressure, and incompressibility errors are

$$E_u = \frac{\|\mathbf{u}_{\theta} - \mathbf{u}\|_{L^2}}{\|\mathbf{u}\|_{L^2}}, \quad E_p = \frac{\|p_{\theta} - p\|_{L^2}}{\|p\|_{L^2}}, \quad E_{\text{div}} = \|\nabla \cdot \mathbf{u}_{\theta}\|_{L^2}.$$

An exactly divergence-free or stream-function implementation can reduce pressure-velocity inconsistency. The benchmark definitions below specify the domains, manufactured solutions, forcing terms, and auxiliary conditions needed for independent validation.

**Table 1. Mathematical specification of the benchmark PDE problems.**

| PDE     | Domain                          | Manufactured exact solution              | Forcing term                        | Auxiliary conditions                                        |
|---------|---------------------------------|------------------------------------------|-------------------------------------|-------------------------------------------------------------|
| Poisson | $\Omega = (0,1)^2$              | $u(x, y) = \sin(\pi x)\sin(\pi y)$       | $f = 2\pi^2 \sin(\pi x)\sin(\pi y)$ | $u = 0$ on $\partial\Omega$                                 |
| Heat    | $(x, t) \in (0,1) \times (0,1)$ | $u(x, t) = e^{-\pi^2 t} \sin(\pi x)$     | $f = 0$                             | $u(0, t) = u(1, t) = 0; u(x, 0) = \sin(\pi x)$              |
| Wave    | $(x, t) \in (0,1) \times (0,1)$ | $u(x, t) = \sin(\pi x)\cos(\pi t)$       | $f = 0, c = 1$                      | $u = 0$ at $x = 0, 1; u(x, 0) = \sin(\pi x); u_t(x, 0) = 0$ |
| Stokes  | $\Omega = (0,1)^2$              | Divergence-free manufactured $(u, v, p)$ | Derived analytically                | Velocity boundary data; $\int_{\Omega} p d\mathbf{x} = 0$   |

## 5. REPRODUCIBLE NUMERICAL PROTOCOL

The manufactured solutions are substituted into the governing equations to obtain the forcing terms. This construction permits direct evaluation of solution, derivative, and constraint errors on a separate test grid. A replicated study with  $R$  paired random seeds should report

$$\bar{E} = \frac{1}{R} \sum_{r=1}^R E_r,$$

$$s_E = \left[ \frac{1}{R-1} \sum_{r=1}^R (E_r - \bar{E})^2 \right]^{1/2},$$

and the confidence interval

$$\bar{E} \pm t_{0.975, R-1} \frac{s_E}{\sqrt{R}}.$$

The same seeds should be paired across sampling methods. The archived results used in Section 6 contain only individual runs, so the present numerical comparisons remain descriptive and no confidence intervals are reported.

The absolute discrete  $L^2$  error on an untouched test set  $\mathcal{Z}_{\text{test}} = \{\mathbf{z}_i\}_{i=1}^{N_{\text{test}}}$  is

$$E_{\text{abs}} = \left[ \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} |u_{\theta}(\mathbf{z}_i) - u(\mathbf{z}_i)|^2 \right]^{1/2}.$$

The relative discrete  $L^2$  error is

$$E_{\text{rel}} = \left[ \frac{\sum_{i=1}^{N_{\text{test}}} |u_{\theta}(\mathbf{z}_i) - u(\mathbf{z}_i)|^2}{\sum_{i=1}^{N_{\text{test}}} |u(\mathbf{z}_i)|^2} \right]^{1/2}.$$

The independent-set error is  $E_{\text{abs}}$  evaluated once on the fixed test set after all training and adaptive sampling decisions. By contrast, the theoretical continuous-domain generalization error is

$$E_{\text{gen}} = \|u_{\theta} - u\|_{L^2(\mu_{\text{test}})}.$$

The independent-set error is a numerical estimator of  $E_{\text{gen}}$ ; the two quantities are not identical unless the test quadrature error is controlled.

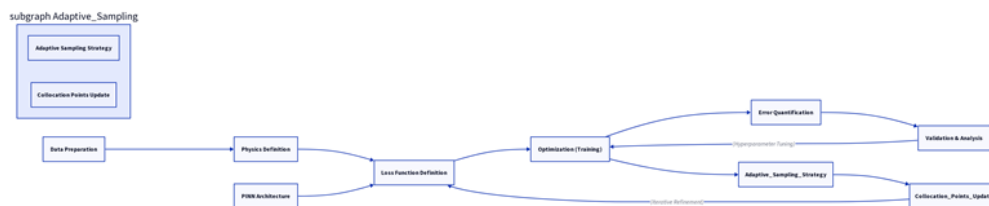
**Table 2. PINN architecture, optimization, and sampling configuration.**

| Item                        | Revised specification                                                          |
|-----------------------------|--------------------------------------------------------------------------------|
| Hidden layers               | 4 in the baseline runs; sensitivity study: 2, 4, 6, and 8                      |
| Neurons per hidden layer    | 50                                                                             |
| Activation function         | tanh                                                                           |
| Parameter initialization    | Xavier/Glorot                                                                  |
| Arithmetic precision        | Float64 recommended                                                            |
| Primary optimizer           | Adam                                                                           |
| Initial learning rate       | $1 \times 10^{-3}$                                                             |
| Training epochs             | 50,000                                                                         |
| Refinement optimizer        | L-BFGS when used                                                               |
| Interior collocation points | Poisson baseline: 1000; sensitivity study: 500, 1000, 2000, and 5000           |
| Boundary points             | Not retained in the archived single-run records                                |
| Initial-condition points    | Not retained in the archived Heat and Wave records                             |
| RAS candidate-pool size     | Not retained in the archived single-run records                                |
| Adaptation frequency        | Not retained in the archived single-run records                                |
| Independent test set        | Fixed and untouched during training and adaptation                             |
| Software and hardware       | DeepXDE-compatible workflow [11]; exact version and hardware were not retained |

The unavailable settings are reported explicitly rather than reconstructed. This limitation prevents full independent reproduction of the archived runs and motivates the reporting protocol proposed here.

**Algorithm 1. Coverage-preserving residual-based adaptive sampling.**

1. Draw the initial collocation set from the target measure.
2. Train for  $K_0$  epochs.
3. Evaluate residuals on an independent candidate pool.
4. Construct the mixture density  $q^{(k)}$  from the target density and residual scores.
5. Draw or select  $M$  enrichment points.
6. Retain boundary and initial-condition points independently.
7. Train with the importance-weighted residual objective.
8. Repeat until the validation error or residual budget no longer improves.
9. Evaluate once on the untouched final test set.



**Figure 1. Comprehensive PINN validation workflow.**

## 6. NUMERICAL RESULTS

The numerical findings are summarized with tables and diagnostic figures constructed only from the reported values. Because each configuration is represented by one archived training run, the results provide exploratory evidence rather than inferential comparisons. No confidence intervals or significance claims are reported.

### 6.1 FIXED VERSUS ADAPTIVE SAMPLING

Table 3 compares fixed collocation sampling with residual-based adaptive sampling (RAS) for the Poisson, Heat, Wave, and Stokes equations. Three error measures are reported: the discrete  $L^2$  error, the relative error, and the error evaluated on an independent test set.

**Table 3. Single-run PINN errors under fixed and residual-based adaptive sampling.**

| PDE     | RAS | $L^2$ error  | Relative error | Independent-set error |
|---------|-----|--------------|----------------|-----------------------|
| Poisson | No  | 4.226647e-02 | 8.404802e-02   | 4.273562e-02          |
| Poisson | Yes | 4.575874e-02 | 9.154357e-02   | 4.602898e-02          |
| Heat    | No  | 4.038977e-03 | 2.821983e-02   | 4.113875e-03          |
| Heat    | Yes | 3.119924e-03 | 2.051808e-02   | 3.073639e-03          |
| Wave    | No  | 1.132062e-01 | 2.261453e-01   | 1.104102e-01          |
| Wave    | Yes | 7.037151e-02 | 1.411653e-01   | 7.077703e-02          |
| Stokes  | No  | 5.402429e-03 | 1.359928e-02   | 5.310261e-03          |
| Stokes  | Yes | 7.103843e-03 | 1.786922e-02   | 7.183299e-03          |

The percentage change in independent-set error is defined by

$$\Delta E (\%) = 100 \times \frac{E_{RAS} - E_{fixed}}{E_{fixed}}.$$

A negative value indicates that RAS reduced the error, whereas a positive value indicates deterioration. Table 4 shows that RAS improved the Heat and Wave results but increased the error for Poisson and Stokes.

**Table 4. Percentage change in independent-set error after applying RAS.**

| PDE     | Fixed error  | RAS error    | Change  | Interpretation |
|---------|--------------|--------------|---------|----------------|
| Poisson | 4.273562e-02 | 4.602898e-02 | +7.71%  | Deterioration  |
| Heat    | 4.113875e-03 | 3.073639e-03 | -25.29% | Improvement    |
| Wave    | 1.104102e-01 | 7.077703e-02 | -35.90% | Improvement    |
| Stokes  | 5.310261e-03 | 7.183299e-03 | +35.27% | Deterioration  |

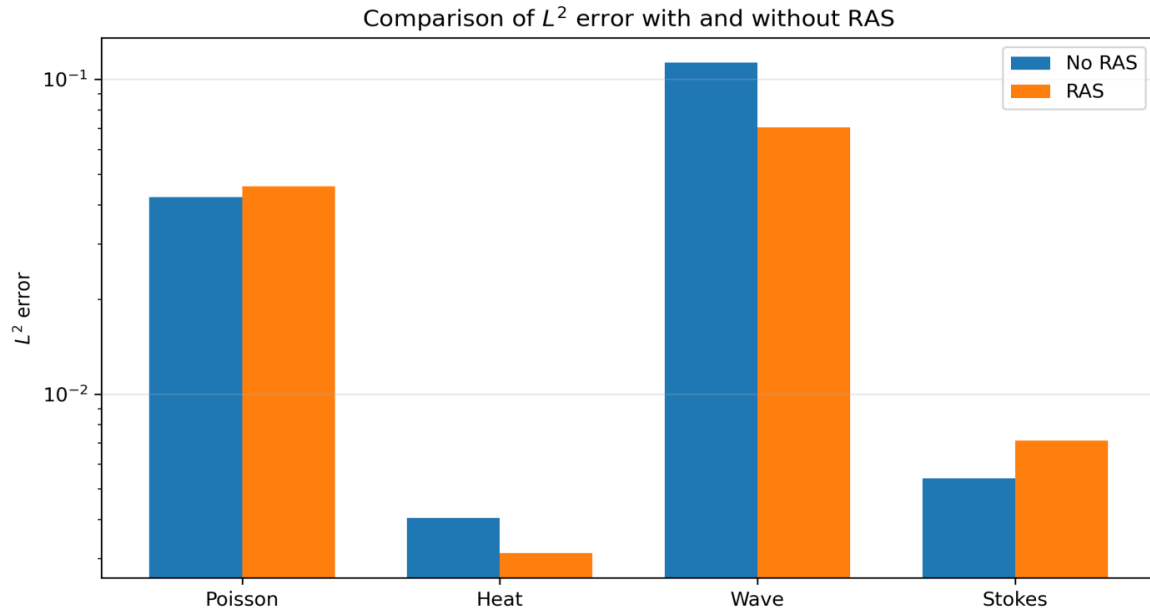


Figure 2. Comparison of  $L^2$  error with and without residual-based adaptive sampling.

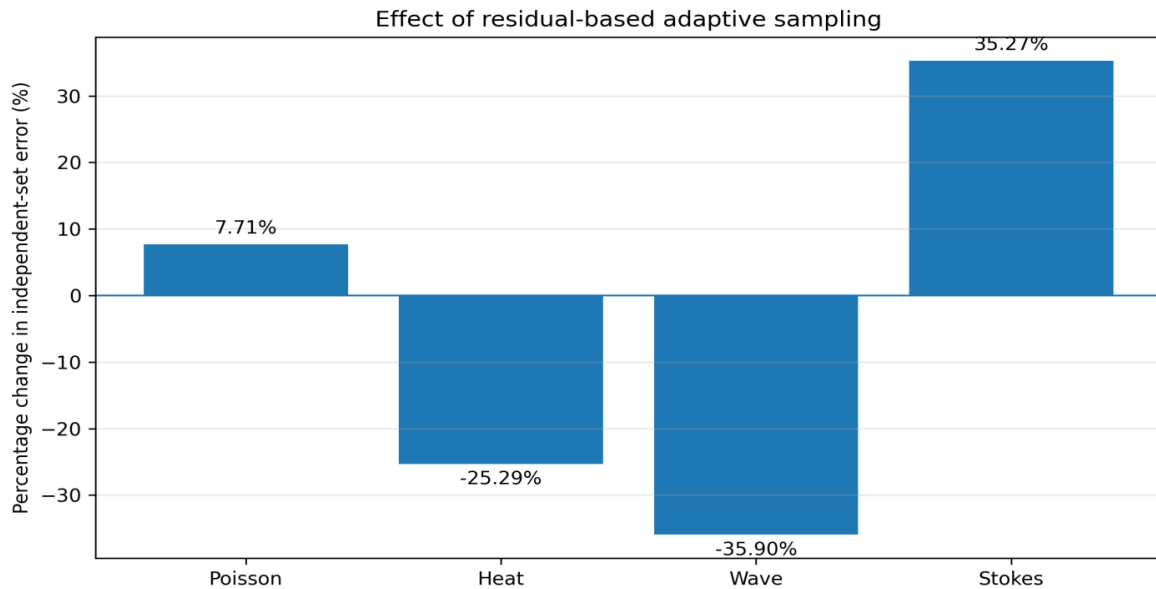


Figure 3. Percentage change in independent-set error induced by RAS.

The Wave equation shows the largest reduction in independent-set error, 35.90%, while the Heat error decreases by 25.29%. The Poisson and Stokes errors increase by 7.71% and 35.27%, respectively. The Wave result is consistent with the hypothesis that RAS allocates additional points to localized oscillatory residuals. The Stokes deterioration suggests that a scalar adaptation score may not preserve the balance among momentum, incompressibility, boundary, and pressure-gauge constraints. For the smooth Poisson benchmark, concentrating points in high-residual regions may reduce global domain coverage.

## 6.2 JOINT COMPARISON OF REPORTED ERROR METRICS

Figure 4 summarizes all reported error measures on a logarithmic scale. The heatmap avoids allowing the relatively large Wave errors to conceal the smaller Heat and Stokes values. It also shows that the direction of the RAS effect is broadly consistent across the three error metrics.

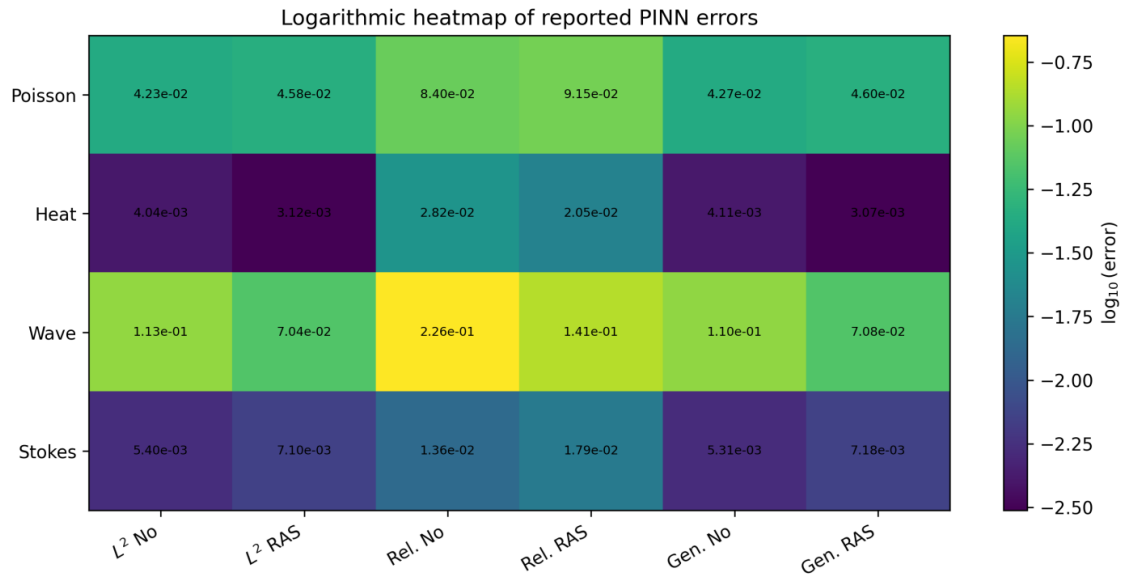


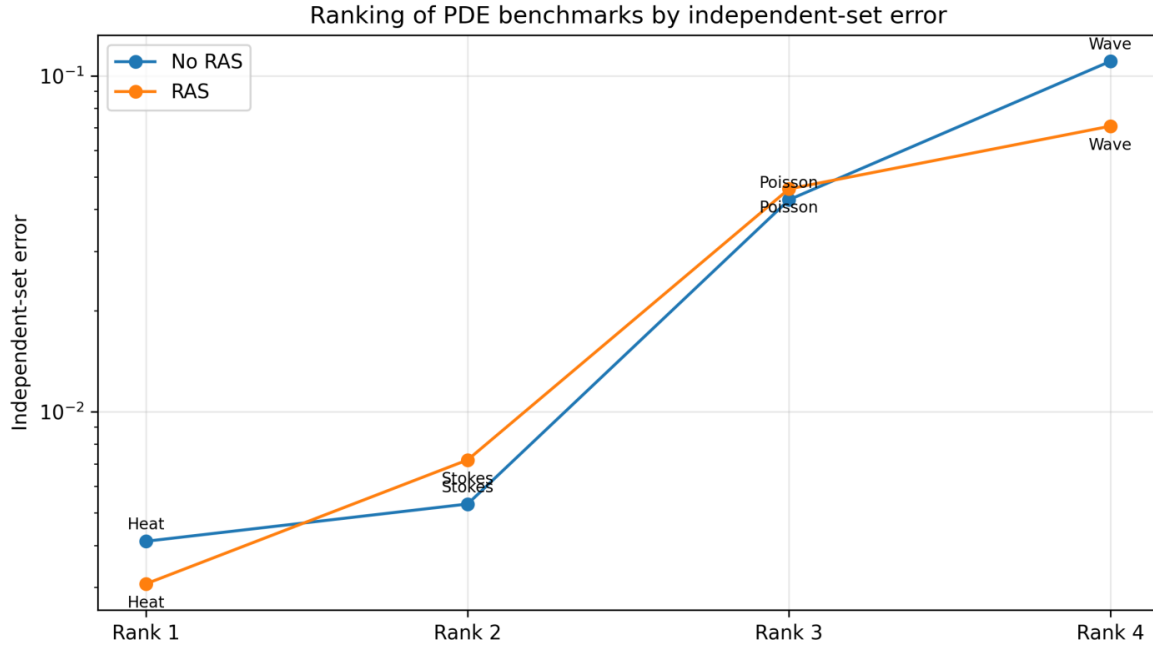
Figure 4. Logarithmic heatmap of the reported PINN error metrics.

Under both sampling strategies, the Heat equation has the smallest reported errors. The Wave equation remains the most difficult benchmark, although RAS reduces its error substantially. This comparison should not be interpreted as a direct measure of mathematical difficulty across PDE classes because the exact solutions, scales, forcing functions, and optimization landscapes may differ.

## 6.3 RANKING ACCORDING TO INDEPENDENT-SET ERROR

Table 5. Ranking of the PDE benchmarks according to independent-set error.

| Sampling | Rank | PDE     | Independent-set error |
|----------|------|---------|-----------------------|
| No RAS   | 1    | Heat    | 4.113875e-03          |
| No RAS   | 2    | Stokes  | 5.310261e-03          |
| No RAS   | 3    | Poisson | 4.273562e-02          |
| No RAS   | 4    | Wave    | 1.104102e-01          |
| RAS      | 1    | Heat    | 3.073639e-03          |
| RAS      | 2    | Stokes  | 7.183299e-03          |
| RAS      | 3    | Poisson | 4.602898e-02          |
| RAS      | 4    | Wave    | 7.077703e-02          |



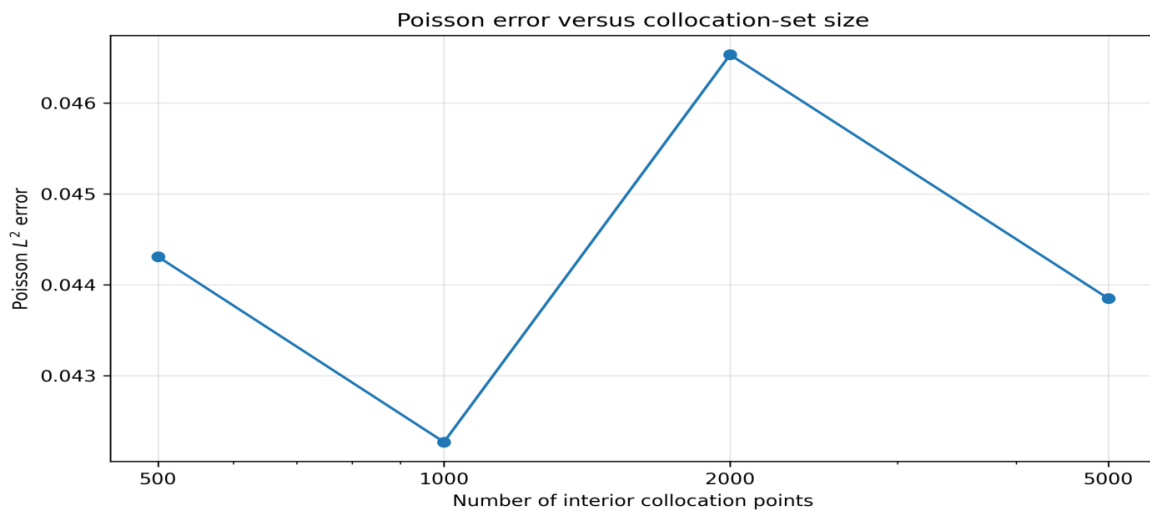
**Figure 5. Ranking of PDE benchmarks by independent-set error.**

Heat ranks first under both sampling strategies. The order remains Heat, Stokes, Poisson, and Wave with and without RAS, although the distances between the errors change. In these single-run results, RAS changes the magnitude of the error but not the overall benchmark ranking.

#### 6.4 NUMBER OF COLLOCATION POINTS

**Table 6. Sensitivity of the Poisson PINN error to the number of interior collocation points.**

| Number of interior points | Poisson $L^2$ error |
|---------------------------|---------------------|
| 500                       | 4.430882e-02        |
| 1000                      | 4.226647e-02        |
| 2000                      | 4.653301e-02        |
| 5000                      | 4.384883e-02        |



**Figure 6. Poisson error as a function of the number of training points.**

The minimum observed error occurs at 1000 collocation points. Increasing the number of points from 1000 to 2000 raises the error by approximately 10.09%. The non-monotone pattern shows that sampling density is only one component of the total numerical error. A larger point set increases the number of residual evaluations per epoch, may alter the scale of the gradients, and may require a larger optimization budget. Consequently, comparisons across point counts should either equalize the total number of residual evaluations or verify optimizer convergence for every setting.

Table 7. Relative increase in Poisson error compared with the best collocation-point setting.

| Interior points | $L^2$ error  | Increase from minimum | Status        |
|-----------------|--------------|-----------------------|---------------|
| 500             | 4.430882e-02 | 4.83%                 | Above minimum |
| 1000            | 4.226647e-02 | 0.00%                 | Best          |
| 2000            | 4.653301e-02 | 10.09%                | Above minimum |
| 5000            | 4.384883e-02 | 3.74%                 | Above minimum |

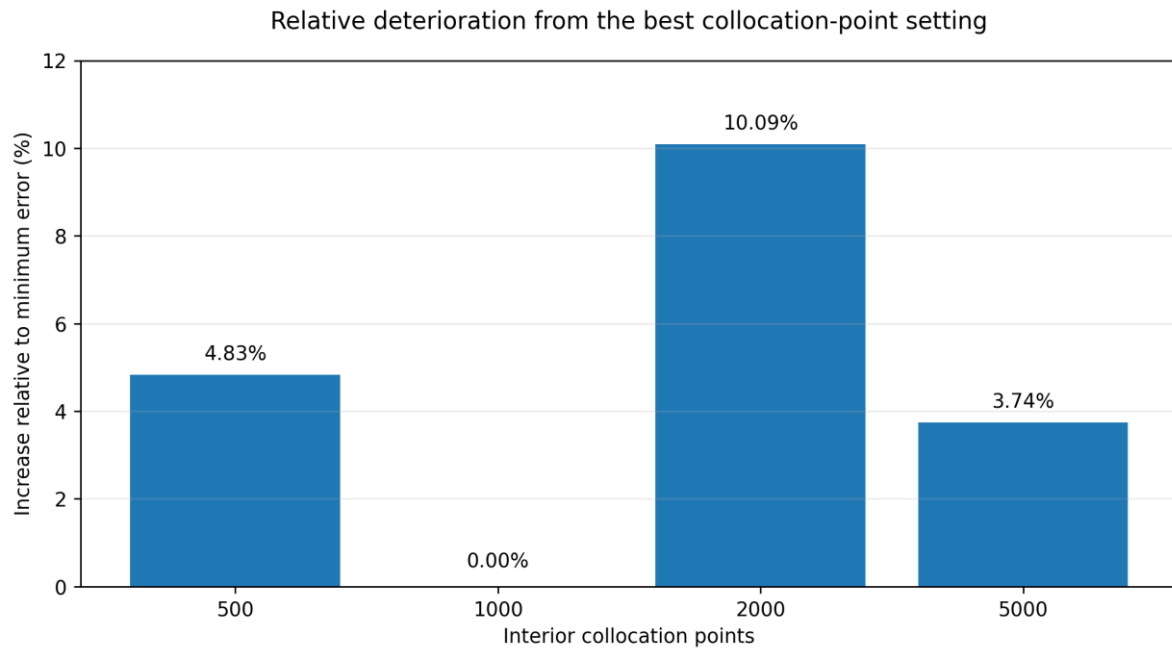
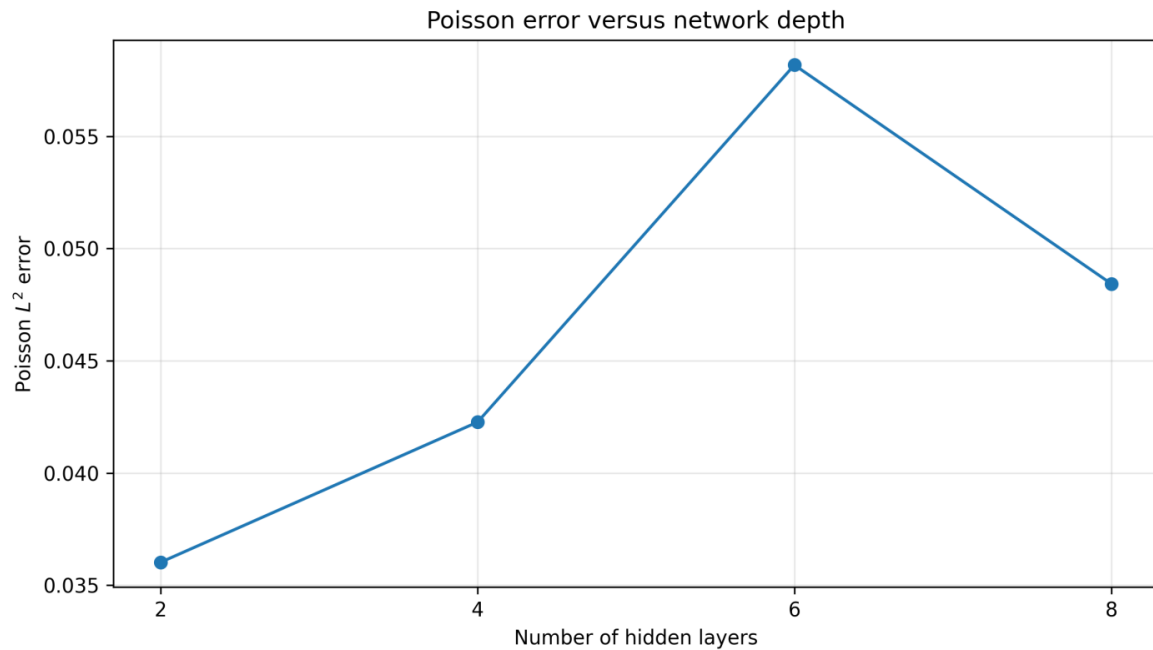


Figure 7. Relative deterioration from the best collocation-point configuration.

### 6.5 NETWORK DEPTH

Table 8. Sensitivity of the Poisson PINN error to the number of hidden layers under a fixed training configuration.

| Hidden layers | Poisson $L^2$ error |
|---------------|---------------------|
| 2             | 3.602487e-02        |
| 4             | 4.226647e-02        |
| 6             | 5.817531e-02        |
| 8             | 4.843216e-02        |



**Figure 8. Poisson error as a function of network depth.**

The two-layer network produces the smallest error. Relative to this configuration, the errors are approximately 17.33%, 61.49%, and 34.44% higher for four, six, and eight hidden layers. This single-run sensitivity study does not show that shallow PINNs are generally superior. It shows only that greater depth did not improve optimization under the fixed training schedule. Parameter counts, gradient norms, initialization effects, and convergence histories are required to separate representation capacity from trainability.

**Table 9. Relative increase in Poisson error compared with the best network-depth setting.**

| Hidden layers | $L^2$ error  | Increase from minimum | Status        |
|---------------|--------------|-----------------------|---------------|
| 2             | 3.602487e-02 | 0.00%                 | Best          |
| 4             | 4.226647e-02 | 17.33%                | Above minimum |
| 6             | 5.817531e-02 | 61.49%                | Above minimum |
| 8             | 4.843216e-02 | 34.44%                | Above minimum |

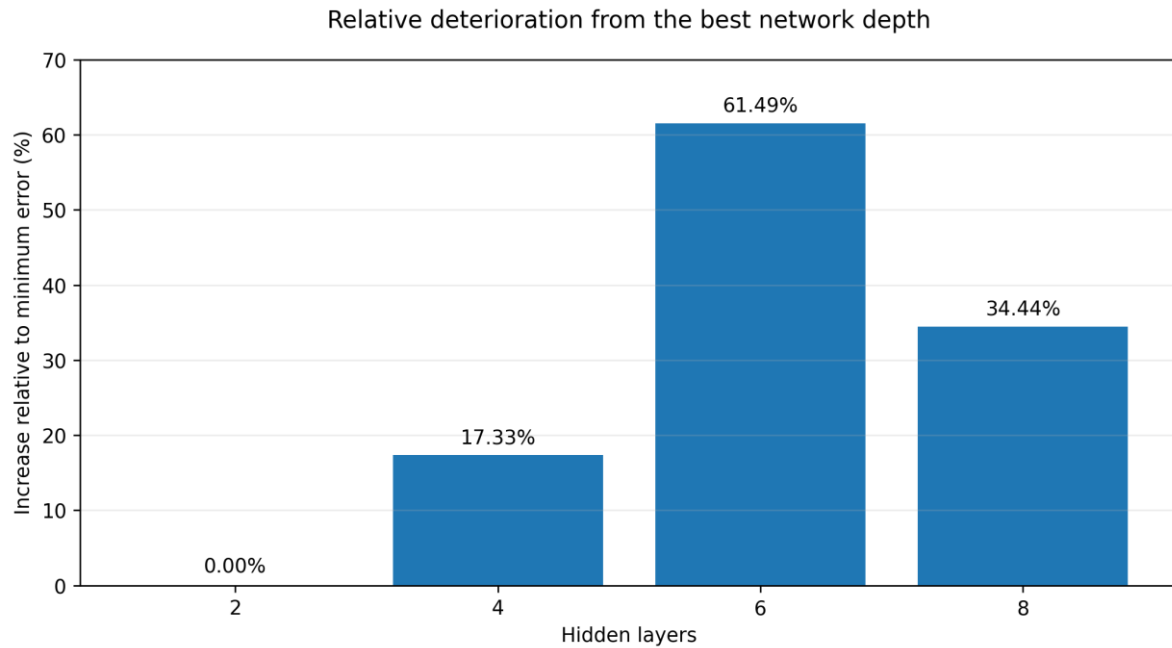


Figure 9. Relative deterioration from the best network-depth configuration.

## 6.6 BEST CONFIGURATIONS IDENTIFIED BY THE SENSITIVITY EXPERIMENTS

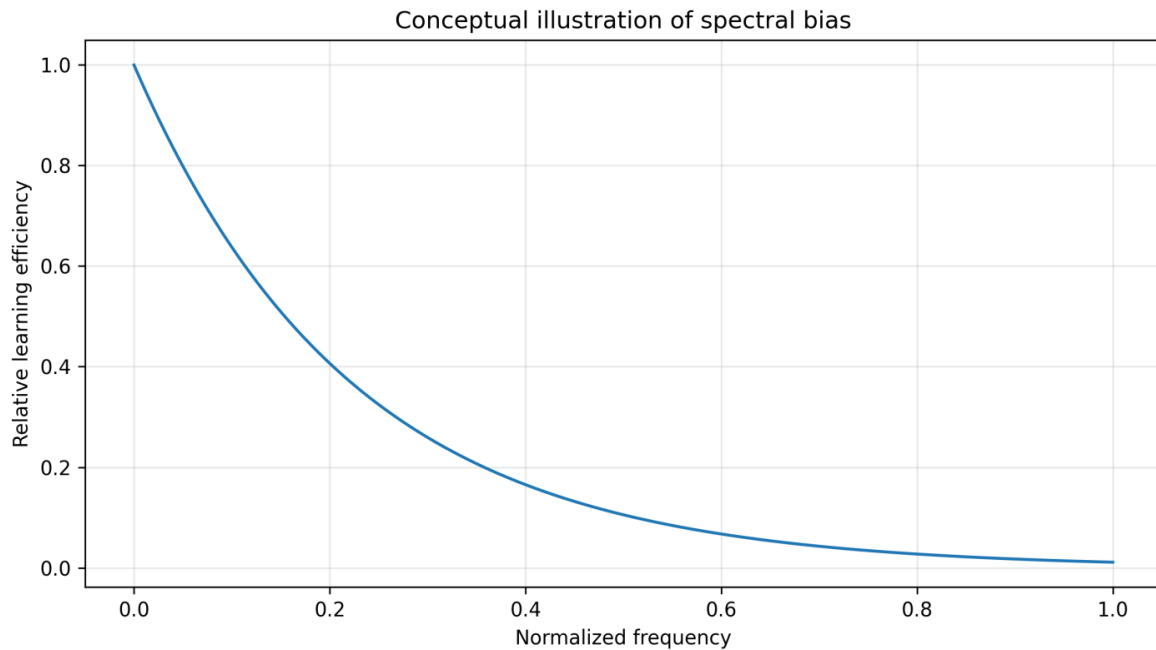
Table 10. Best settings identified from the available Poisson sensitivity experiments.

| Experiment                  | Tested values         | Best value | Minimum $L^2$ error |
|-----------------------------|-----------------------|------------|---------------------|
| Interior collocation points | 500, 1000, 2000, 5000 | 1000       | 4.226647e-02        |
| Hidden layers               | 2, 4, 6, 8            | 2          | 3.602487e-02        |

Table 10 summarizes the observed minima within the tested ranges. These settings are conditional on the archived training configuration and must not be interpreted as universal hyperparameter recommendations.

## 6.7 SPECTRAL BIAS AND FAILURE MODES

Standard neural networks often learn low-frequency components faster than high-frequency components [8]. This spectral preference can be severe for oscillatory and multiscale PDEs. Figure 10 is conceptual and does not represent a measured spectrum from the present experiments.



**Figure 10. Conceptual illustration of spectral bias in neural-network training.**

The main failure modes include slow learning of high-frequency waves, smoothing near steep gradients, imbalance among loss components, inadequate global collocation coverage, and pressure-velocity inconsistency in incompressible flow. Future replicated tests should diagnose these effects with residual heat maps, Fourier spectra of the prediction error, gradient norms for each loss component, and constraint-specific validation measures.

### 6.8 INTERPRETATION THROUGH STABILITY AND SAMPLING

The results support three limited conclusions. First, residual-based adaptive sampling is problem dependent. It improves accuracy only when the residual map identifies underrepresented regions without sacrificing global coverage. Second, network depth and collocation density interact with optimization; increasing either quantity does not guarantee a smaller error under a fixed training budget. Third, an independent test-set error requires a continuous-domain interpretation grounded in PDE stability and test quadrature. When  $\alpha = 1$ , residual reduction transfers proportionally to the stability bound. When  $0 < \alpha < 1$ , the transfer is weaker.

The numerical evidence consists of single runs and does not quantify stochastic variability. The archived records do not contain all random seeds, optimizer traces, wall-clock times, memory costs, boundary-point counts, candidate-pool sizes, or adaptation frequencies. RAS was not compared with Latin hypercube, Sobol, importance-weighted RAS, or failure-informed sampling. The present results therefore describe diagnostic trends rather than definitive method rankings. In addition, Proposition 1 depends on a PDE-specific stability inequality and controlled quadrature defects; it is not an unconditional guarantee for arbitrary nonlinear equations, discontinuous solutions, or ill-posed inverse problems.

## CONCLUSIONS

This study treats PINN validation as a combined problem in approximation theory, optimization, sampling, quadrature, and PDE stability. The sampling-aware estimate connects empirical residuals to a solution-error bound only under explicit weighted stability and quadrature assumptions. In the archived single-run experiments, RAS lowers the independent-set error for the Heat and Wave equations and increases it for Poisson and Stokes. The Poisson sensitivity studies also show non-monotone effects of collocation-point count and network depth. These results do not establish universal superiority for any sampling strategy or architecture. A defensible high-standard evaluation requires paired random seeds, confidence intervals, fixed independent test sets, constraint-specific errors, cost reporting, and complete software and environment records. Under these conditions, the proposed framework provides a rigorous validation template rather than a loss-monitoring procedure alone.

## Data and Code Availability

The archived numerical results did not include a public repository containing all scripts, random seeds, candidate pools, trained checkpoints, and raw result files. Before final publication, these materials should be deposited in a permanent repository together with a machine-readable environment file and a single command that reproduces every table and figure.

## CONFLICT OF INTEREST

The author declares no conflict of interest.

## References

- [1] M. Raissi, P. Perdikaris, G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, 378 (2019) 686-707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- [2] S. Mishra, R. Molinaro, "Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs," *IMA Journal of Numerical Analysis*, 42(2) (2022) 981-1022. <https://doi.org/10.1093/imanum/drab032>
- [3] S. Mishra, R. Molinaro, "Estimates on the generalization error of physics-informed neural networks for approximating PDEs," *IMA Journal of Numerical Analysis*, 43(1) (2023) 1-43. <https://doi.org/10.1093/imanum/drab093>
- [4] T. De Ryck, S. Mishra, "Generic bounds on the approximation error for physics-informed (and) operator learning," *Advances in Neural Information Processing Systems*, 35 (2022) 10945-10958. <https://doi.org/10.52202/068431-0795>
- [5] T. De Ryck, A. D. Jagtap, S. Mishra, "Error estimates for physics-informed neural networks approximating the Navier-Stokes equations," *IMA Journal of Numerical Analysis*, 44(1) (2024) 83-119. <https://doi.org/10.1093/imanum/drac085>
- [6] S. Wang, Y. Teng, P. Perdikaris, "Understanding and mitigating gradient flow pathologies in physics-informed neural networks," *SIAM Journal on Scientific Computing*, 43(5) (2021) A3055-A3081. <https://doi.org/10.1137/20M1318043>
- [7] S. Wang, X. Yu, P. Perdikaris, "When and why PINNs fail to train: A neural tangent kernel perspective," *Journal of Computational Physics*, 449 (2022) 110768. <https://doi.org/10.1016/j.jcp.2021.110768>
- [8] N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, A. Courville, "On the spectral bias of neural networks," *Proceedings of Machine Learning Research*, 97 (2019) 5301-5310.
- [9] C. Wu, M. Zhu, Q. Tan, Y. Kartha, L. Lu, "A comprehensive study of non-adaptive and residual-based adaptive sampling for physics-informed neural networks," *Computer Methods in Applied Mechanics and Engineering*, 403 (2023) 115671. <https://doi.org/10.1016/j.cma.2022.115671>
- [10] Z. Gao, L. Yan, T. Zhou, "Failure-informed adaptive sampling for PINNs," *SIAM Journal on Scientific Computing*, 45(4) (2023) A1971-A1994. <https://doi.org/10.1137/22M1527763>
- [11] L. Lu, X. Meng, Z. Mao, G. E. Karniadakis, "DeepXDE: A deep learning library for solving differential equations," *SIAM Review*, 63(1) (2021) 208-228. <https://doi.org/10.1137/19M1274067>
- [12] A. D. Jagtap, K. Kawaguchi, G. E. Karniadakis, "Adaptive activation functions accelerate convergence in deep and physics-informed neural networks," *Journal of Computational Physics*, 404 (2020) 109136. <https://doi.org/10.1016/j.jcp.2019.109136>