

Mixed physics-informed neural networks for anisotropic diffusion: Energy error and directional coupling

Mingde Ye

ye_mingde@163.com

Harbin Institute of Technology

Xiaohua Ding

Harbin Institute of Technology

Research Article

Keywords: physics-informed neural networks, anisotropic diffusion, mixed formulation, energy error, directional coupling, residual minimization

Posted Date: October 12th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11128449/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Mixed physics-informed neural networks for anisotropic diffusion: Energy error and directional coupling

Mingde Ye, Xiaohua Ding*

College of Science, Harbin Institute of Technology at Weihai, China

E-mail addresses: Ye_mingde@163.com (M. Ye); mathdxh@hit.edu.cn (X. Ding)

*Corresponding author

September 2026

Abstract

Mixed physics-informed neural networks (PINNs) offer a natural way to approximate potential and flux together in anisotropic diffusion. Balancing individual residual components can be useful, but leaves open how their directional coupling should enter the loss. We examine how this coupling affects the relation between residual and mixed energy error, and use this connection to study analytical and operator-probe weights. A weighted orthogonal decomposition yields sharp equivalence constants for the ideal residual functional, independent of coefficient contrast in the diffusion-weighted norms. A simple but revealing constant-flux example shows how every fixed positive componentwise weight can select the same suboptimal approximation. The exact error ratio is determined by off-diagonal coupling. This result holds for continuous and discrete losses with positive quadrature weights, even with exact conservation and minimization. The classical weight K^{-1} recovers the best approximation in this trial family. An operator-probe alternative combines the flux and potential-gradient response moments through their matrix geometric mean. Its spectral bounds yield a continuous upper bound on the energy error. Numerical comparisons show a particularly clear benefit from retaining coupling for constant rotated coefficients. In an independent 84-run study on two manufactured problems, probe weighting reduces geometric-mean mixed-energy errors by approximately 37% and 56% relative to the conservation-matched and original Euclidean baselines, respectively. The observed geometric-mean error is also slightly lower than with analytical weighting. Darcy-flow and heat-conduction examples illustrate the resulting potential, flux and error fields.

Keywords: physics-informed neural networks; anisotropic diffusion; mixed formulation; energy error; directional coupling; residual minimization

1 Introduction

Anisotropic diffusion describes transport with a preferred direction, as in flow through porous media and heat conduction in composite materials. In these applications, the flux is often just as important as the potential. Yet an accurate hydraulic head or temperature does not by itself ensure an accurate flow rate or heat flux. Physics-informed neural networks (PINNs) approximate solutions of partial differential equations (PDEs) by minimizing equation residuals[16, 25]. A mixed formulation predicts the potential and flux together, so their errors can be assessed in a common energy norm. Choosing a residual norm that reflects this error is therefore an important part of the method.

This choice becomes particularly important when the principal diffusion directions are oblique to the coordinate axes. Separate component weights adjust the penalties along those axes, whereas the energy norm also contains cross terms that describe how the components interact. Omitting these terms may seem a modest simplification, but its consequence can go beyond a change in scale. Even when minimized exactly, the resulting loss can select a different member of the same approximation class. Our central question is how directional coupling affects both the control of energy error and the approximation selected by residual minimization.

This question fits naturally into the development of energy-based neural solvers. The Deep Ritz method minimizes an energy functional, while neural least-squares methods minimize first-order system residuals[2, 8, 11]. Variational physics-informed neural operators such as VINO extend energy-based training to learning solution operators[12]. Here we consider individual diffusion problems with known coefficients, for which the relation between a mixed residual norm and the potential–flux error can be examined directly.

The analytical foundation is well established. Classical first-order system least-squares (FOSLS) and mixed a posteriori analysis relate residual norms to errors in appropriate function spaces[4, 5, 7, 26, 30]. For diffusion, the constitutive energy uses the full tensor K^{-1} , while a Poincaré estimate supplies a computable weight for the conservation residual. The analytical loss used here is one half of the squared functional error majorant of Repin et al. [26] with unit balancing parameter.

These ideas have also been carried into neural solvers through coefficient-weighted least squares and functional majorants[8, 14, 15]. ASTRAL, for example, jointly approximates the primal variable and flux within this majorant family[15]. Robust Deep FOSLS studies energy-based Poincaré weighting for scalar-coefficient transmission problems, using an error norm that also includes flux divergence[10]. This body of work provides a sound starting point for examining what the full tensor contributes beyond separate component weights, and how a related tensor can be constructed from operator responses.

A complementary line of work addresses how the training objective is balanced and minimized. Adaptive loss weights and multi-objective updates balance training terms[3, 18, 20, 21, 29], and operator preconditioning uses properties of the differential operator to improve optimization[9, 19]. Deep Fourier Residual and robust variational PINNs approximate dual residual norms through structured transforms or discrete Riesz representations[27, 28]. For our purposes, a basic distinction remains: adjusting component magnitudes does not supply the cross terms of a quadratic loss. The underlying geometry of inner-product changes and diagonal scaling is classical[17, 24]. We examine its effect in an admissible mixed diffusion example where directional coupling can be studied separately from conservation, spatial sampling and iterative optimization errors.

Our analysis begins with a sharp comparison between the ideal residual functional and the mixed energy error. A weighted orthogonal decomposition gives the optimal constants, independently of coefficient contrast in the diffusion-weighted norms. Isotropic and coordinate-diagonal weights still provide an upper error bound, although their reverse estimates can deteriorate as anisotropy increases.

A reliable error bound is valuable, but it does not by itself tell us which approximation the loss will select. A simple constant-flux trial family makes this distinction especially clear. Every fixed positive diagonal weight selects the same approximation; when the off-diagonal coupling is nonzero, its energy error exceeds the best available by an explicit factor determined by that coupling. The factor applies to continuous and sampled losses alike. Since the loss is already minimized exactly, the discrepancy in this example comes from the chosen quadratic form rather than incomplete optimization.

The two implementations follow naturally from this analysis. The analytical weight uses K^{-1} directly. For the operator-probe alternative, we probe the flux and potential-gradient terms separately and combine their response moments through the matrix geometric mean. Spectral bounds relate the resulting tensor to K , allowing the continuous probe-weighted loss to bound the energy error. In the constant-flux setting, these bounds also compare the selected approximation with the best one in the prescribed class. Both constructions have a useful consistency property: rewriting the constitutive equations leaves the objective unchanged when the residuals and weights are transformed together.

We assess the practical consequences using matched networks and training settings. By fixing the conservation weight and the trace of the constitutive weight matrix, we compare how penalties are distributed across directions. The benefit of retaining the full tensor is clearest for constant rotated coefficients; when the principal axes vary in space, the gains are more modest and less consistent across seeds. A separate 84-run comparison on two manufactured problems supports reduced mixed-energy errors for probe weighting relative to two Euclidean baselines. In this study, probe weighting also gives a slightly lower geometric-mean mixed-energy error than analytical weighting. Finally, potential, flux and error fields for manufactured solutions, Darcy flow and heat conduction show where the remaining errors occur, complementing the aggregate measures with a direct view of the computed solutions.

Section 2 develops the weighting methods and their error estimates. Section 3 presents the numerical comparisons and applications, and Section 4 concludes the paper. The Supplementary Material contains additional proofs, experimental settings and numerical results.

2 Method

2.1 Mixed formulation of anisotropic diffusion

Let $\Omega \subset \mathbb{R}^m$ be a bounded Lipschitz domain, $f \in L^2(\Omega)$, and $K(x)$ a known measurable symmetric tensor satisfying

$$\alpha|z|^2 \leq z^\top K(x)z \leq \beta|z|^2, \quad 0 < \alpha \leq \beta < \infty, \quad (1)$$

for every $z \in \mathbb{R}^m$ and almost every $x \in \Omega$. We consider

$$-\operatorname{div}(K\nabla u_*) = f \quad \text{in } \Omega, \quad u_* = 0 \quad \text{on } \partial\Omega, \quad q_* = -K\nabla u_*. \quad (2)$$

The potential belongs to $H_0^1(\Omega)$ and the flux to $H(\operatorname{div}; \Omega)$, the space of square-integrable vector fields with square-integrable distributional divergence. The weak equation is

$$\int_{\Omega} \nabla u_*^\top K \nabla v \, dx = \int_{\Omega} f v \, dx, \quad v \in H_0^1(\Omega). \quad (3)$$

Uniform ellipticity and the Poincaré inequality give a unique weak solution[6, 13].

For an admissible approximation $(u, q) \in H_0^1(\Omega) \times H(\operatorname{div}; \Omega)$, define the errors $e = u - u_*$, $p = q - q_*$ and residuals

$$r_c = \operatorname{div} q - f, \quad \eta = q + K\nabla u, \quad \xi = K^{-1}q + \nabla u = K^{-1}\eta. \quad (4)$$

The energy and dual norms are

$$\begin{aligned} \|v\|_K^2 &= \int_{\Omega} \nabla v^\top K \nabla v \, dx, & \|w\|_{K^{-1}}^2 &= \int_{\Omega} w^\top K^{-1} w \, dx, \\ \|g\|_{-1,K} &= \sup_{0 \neq v \in H_0^1(\Omega)} \frac{|\langle g, v \rangle|}{\|v\|_K}, & \mathcal{E}_{\text{mix}}^2 &= \|e\|_K^2 + \|p\|_{K^{-1}}^2. \end{aligned} \quad (5)$$

Here $\langle g, v \rangle$ denotes the H^{-1} - H_0^1 dual pairing. In particular, $\langle \operatorname{div} w, v \rangle = -\int_{\Omega} w \cdot \nabla v$ for $w \in L^2(\Omega)^m$. The mixed error $\mathcal{E}_{\text{mix}}$ measures both potential gradients and physical flux in the diffusion energy scales.

2.2 From componentwise to tensor weighting

A componentwise loss has the form $\sum_j w_j r_j^2$. For anisotropic diffusion, the constitutive energy generally contains an additional term. In two dimensions, writing $H = K^{-1}$ gives

$$\eta^\top H \eta = H_{11}\eta_1^2 + 2H_{12}\eta_1\eta_2 + H_{22}\eta_2^2. \quad (6)$$

When $H_{12} \neq 0$, no fixed weights on η_1^2 and η_2^2 reproduce this energy for every residual. The cross term accounts for the orientation of the principal diffusion axes. Section 2.5 quantifies its effect on the approximation selected by loss minimization.

Full tensor weights also allow equivalent equations to use the same loss. Under an invertible change of residual coordinates $r' = Ar$, transport a positive-definite weight W as

$$W' = A^{-\top} W A^{-1}, \quad (Ar)^\top W' (Ar) = r^\top W r. \quad (7)$$

For $\xi = K^{-1}\eta$, the corresponding weights are K^{-1} and K , so both formulations give the same constitutive energy. Even a diagonal weight can become a full matrix under a change of coordinates. We next connect this tensor structure to the potential and flux errors and describe its analytical and probe-based implementations.

2.3 Analytical residual weighting

We choose the residual norm to control the mixed error in (5). Classical mixed diffusion estimates[5, 26] suggest measuring conservation in the energy-dual norm and the constitutive residual with the full coefficient weight:

$$\mathcal{J}_{\text{nat}}(u, q) = \|r_c\|_{-1, K}^2 + \|\eta\|_{K^{-1}}^2. \quad (8)$$

Evaluating the dual norm requires an auxiliary variational solve. For pointwise training, we replace it by a Poincaré upper bound. Let C_P satisfy $\|v\|_{L^2} \leq C_P \|\nabla v\|_{L^2}$ on $H_0^1(\Omega)$. Then

$$\|r_c\|_{-1, K} \leq \frac{C_P}{\sqrt{\alpha}} \|r_c\|_{L^2}, \quad \mathcal{J}_{\text{stab}} = \lambda_{\text{th}} \|r_c\|_{L^2}^2 + \|\eta\|_{K^{-1}}^2, \quad \lambda_{\text{th}} = \frac{C_P^2}{\alpha}. \quad (9)$$

The analytical method uses K^{-1} directly and retains all directional coupling in the constitutive residual. This loss is an existing construction: in the notation of Repin et al. [26], the squared error majorant at unit balancing parameter is $M_1^2 = 2\mathcal{J}_{\text{stab}}$. Neural implementations of coefficient-weighted least squares and functional majorants include[8, 15]. Here the analytical weight also provides the reference for the probe construction and for comparisons with isotropic and componentwise weights.

2.4 Residual weighting from operator probes

The analytical construction identifies the desired coefficient scale. We now recover a related scale from the responses of the constitutive operator. Its two terms act on the flux and potential gradient in different ways, so we probe them separately. Draw independent standard Gaussian directions $a_i, b_i \in \mathbb{R}^m$, $i = 1, \dots, N$, in nondimensional physical coordinates and reuse them at every spatial point. For any fixed $x_0 \in \mathbb{R}^m$, the fields $(u, q) = (0, a_i)$ and $(b_i^\top(x - x_0), 0)$ give the responses a_i and $K(x)b_i$. These fields serve only to evaluate the local operator; the trained potential satisfies its prescribed boundary condition.

Using the known zero population means, form the uncentered second moments

$$P_N = \frac{1}{N} \sum_i a_i a_i^\top, \quad S_N = \frac{1}{N} \sum_i b_i b_i^\top, \quad Q_N(x) = K(x) S_N K(x). \quad (10)$$

For positive-definite P_N, S_N , define

$$M_N^\# = P_N \# Q_N, \quad X \# Y = X^{1/2} (X^{-1/2} Y X^{-1/2})^{1/2} X^{1/2}. \quad (11)$$

The population second moments of these responses are I and K^2 . Their arithmetic mean would mix two different scales, whereas their matrix geometric mean is $I \# K^2 = K$. This identity motivates (11); the geometric-mean operation itself is classical[1, 23].

The probe loss keeps the analytic conservation weight,

$$\mathcal{J}_{N, \text{probe}} = \lambda_{\text{th}} \|r_c\|_{L^2}^2 + \int_{\Omega} \eta^\top (M_N^\#)^{-1} \eta \, dx. \quad (12)$$

We use $N = 32$ directions in each set, giving 64 probe responses per realization, and keep the resulting tensor fixed during training. The construction uses the known constitutive operator and the ellipticity lower bound α ; it does not require the exact solution or the current network parameters. As the number of probes increases, the tensor approaches K . The bounds below quantify how a finite set of probes changes the analytical scale and its error estimate.

2.5 Energy error and directional coupling

We first relate the ideal residual norm to the mixed energy error. We then examine what changes when directional coupling is removed: both the loss-to-error comparison and the approximation selected by minimization. Finally, the probe bounds connect the construction to these error estimates, and the transformation rule ensures consistency between equivalent constitutive formulations.

Theorem 2.1 (Sharp mixed-energy equivalence). *Under (1)–(2), every admissible pair satisfies*

$$\varphi^{-1} \mathcal{J}_{\text{nat}}^{1/2} \leq \mathcal{E}_{\text{mix}} \leq \varphi \mathcal{J}_{\text{nat}}^{1/2}, \quad \varphi = \frac{1 + \sqrt{5}}{2}. \quad (13)$$

Both constants are independent of α, β in the stated weighted norms and are optimal uniformly over this problem class.

Proof. The error and residual definitions give $r_c = \operatorname{div} p$ and $\eta = p + K \nabla e$. We first separate p into a weighted gradient and a divergence-free part. For every $v \in H_0^1(\Omega)$, weighted Cauchy–Schwarz gives

$$\left| \int_{\Omega} p \cdot \nabla v \, dx \right| \leq \|p\|_{K^{-1}} \|v\|_K.$$

The integral therefore defines a bounded linear functional of v in the energy norm. Uniform ellipticity and the Poincaré inequality make the energy bilinear form coercive on $H_0^1(\Omega)$. The Lax–Milgram theorem therefore gives a unique $z \in H_0^1(\Omega)$ satisfying

$$\int_{\Omega} \nabla z^\top K \nabla v \, dx = \int_{\Omega} p \cdot \nabla v \, dx \quad \text{for every } v \in H_0^1(\Omega).$$

Set $p_0 = p - K \nabla z$. Subtracting the two sides shows that

$$\int_{\Omega} p_0 \cdot \nabla v \, dx = 0 \quad (v \in H_0^1(\Omega)),$$

so $\operatorname{div} p_0 = 0$ in the distributional sense. Moreover,

$$\langle r_c, v \rangle = - \int_{\Omega} p \cdot \nabla v \, dx = - \int_{\Omega} \nabla z^\top K \nabla v \, dx.$$

Cauchy–Schwarz gives $|\langle r_c, v \rangle| \leq \|z\|_K \|v\|_K$. If $z \neq 0$, the choice $v = z$ attains equality in the supremum defining the dual norm; if $z = 0$, both sides vanish. Hence $\|r_c\|_{-1,K} = \|z\|_K$.

We now express the error and loss in the same orthogonal components. Define

$$X = K^{1/2} \nabla e, \quad Y = K^{1/2} \nabla z, \quad W = K^{-1/2} p_0.$$

All unsubscripted norms and inner products below are those of $L^2(\Omega)^m$. Taking $v = e$ and $v = z$ in the identity for p_0 gives

$$(W, X)_{L^2} = \int_{\Omega} p_0 \cdot \nabla e \, dx = 0, \quad (W, Y)_{L^2} = \int_{\Omega} p_0 \cdot \nabla z \, dx = 0.$$

Since $K^{-1/2} p = Y + W$ and $K^{-1/2} \eta = X + Y + W$, these orthogonality relations yield

$$\mathcal{E}_{\text{mix}}^2 = \|X\|^2 + \|Y\|^2 + \|W\|^2, \quad \mathcal{J}_{\text{nat}} = \|Y\|^2 + \|X + Y\|^2 + \|W\|^2. \quad (14)$$

Expanding the two terms involving X and Y in $\mathcal{J}_{\text{nat}}$ gives

$$\|Y\|^2 + \|X + Y\|^2 = \|X\|^2 + 2(X, Y)_{L^2} + 2\|Y\|^2.$$

For each spatial point and vector component, this quadratic form has matrix

$$B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \det(B - \lambda I) = \lambda^2 - 3\lambda + 1.$$

Its eigenvalues are $(3 - \sqrt{5})/2 = \varphi^{-2}$ and $(3 + \sqrt{5})/2 = \varphi^2$. Applying the corresponding bounds to each component and integrating gives

$$\varphi^{-2}(\|X\|^2 + \|Y\|^2) \leq \|Y\|^2 + \|X + Y\|^2 \leq \varphi^2(\|X\|^2 + \|Y\|^2).$$

Because $\varphi^{-2} \leq 1 \leq \varphi^2$, adding the term $\|W\|^2$ and using (14) yields

$$\varphi^{-2} \mathcal{E}_{\text{mix}}^2 \leq \mathcal{J}_{\text{nat}} \leq \varphi^2 \mathcal{E}_{\text{mix}}^2.$$

Rearranging and taking square roots proves (13). The constants come entirely from B , which contains no coefficient parameters.

To prove optimality, take $K = I$ and $f = 0$, so that $u_* = 0$ and $q_* = 0$. Choose any nonzero $e \in C_c^\infty(\Omega)$ and set $u = e$, $q = t \nabla e$, where $t \in \mathbb{R}$. This pair is admissible because u has zero boundary trace and q has square-integrable divergence $t \Delta e$. Its projection is $z = te$, so $p_0 = 0$, $Y = tX$ and

$$\mathcal{E}_{\text{mix}}^2 = (1 + t^2) \|\nabla e\|_{L^2}^2, \quad \mathcal{J}_{\text{nat}} = [t^2 + (1 + t)^2] \|\nabla e\|_{L^2}^2.$$

Consequently,

$$\frac{\mathcal{J}_{\text{nat}}}{\mathcal{E}_{\text{mix}}^2} = \frac{1 + 2t + 2t^2}{1 + t^2} = \begin{cases} \varphi^2, & t = \varphi, \\ \varphi^{-2}, & t = -\varphi^{-1}. \end{cases}$$

The first choice gives equality in the lower bound for $\mathcal{E}_{\text{mix}}$ in (13), and the second gives equality in the upper bound. Neither constant can therefore be improved uniformly over the stated problem class. $\square$

The absence of the coefficient contrast from the constants follows from the choice of energy spaces. In these spaces, the diffusion operator satisfies

$$\|-\operatorname{div}(K \nabla v)\|_{-1,K} = \|v\|_K, \quad v \in H_0^1(\Omega). \quad (15)$$

The residual and the error are therefore measured using the same coefficient field. The orthogonal decomposition reduces their comparison to the fixed two-by-two matrix in the proof, rather than an estimate involving the largest and smallest eigenvalues of K .

The coefficient independence concerns the energy norms themselves; converting to unweighted norms reintroduces coefficient dependence. The upper constant also follows by applying a two-by-two spectral bound to the mixed estimate of Repin et al. [26]. The decomposition above makes both sharp constants and their equality cases explicit, and also gives $\|p\|_{K^{-1}}^2 \leq \mathcal{J}_{\text{nat}}$.

Consequently,

$$\|e\|_K \leq \sqrt{2} \mathcal{J}_{\text{stab}}^{1/2}, \quad \|p\|_{K^{-1}} \leq \mathcal{J}_{\text{stab}}^{1/2}, \quad \mathcal{E}_{\text{mix}} \leq \varphi \mathcal{J}_{\text{stab}}^{1/2}. \quad (16)$$

On $(0, 1)^2$, $C_P = 1/(\pi\sqrt{2})$ and $\lambda_{\text{th}} = 1/(2\pi^2\alpha)$. This replacement retains a continuous upper error bound. Its frequency weighting differs from that of the dual norm: on $(0, 1)$, $K = 1$, $f = 0$ and $(u_n, q_n) = (0, \cos(n\pi x))$ give $\mathcal{E}_{\text{mix}}^2 = 1/2$, $\mathcal{J}_{\text{nat}} = 1$, and $\mathcal{J}_{\text{stab}} = (n^2 + 1)/2$. Thus a frequency-independent reverse bound for $\mathcal{J}_{\text{stab}}$ is unavailable over all admissible pairs. Finite-point approximation introduces a further quadrature error. For scalar diffusion, Robust Deep FOSLS[10] instead compares the weighted L^2 loss with a graph error containing an additional term $C \|\operatorname{div} p\|_{L^2}^2$, where $C > 0$. That target norm differs from $\mathcal{E}_{\text{mix}}$ in (13).

What is lost by removing directional coupling? In two dimensions, write $H = K^{-1}$ and compare the full weight with

$$Q_{\text{iso}} = \operatorname{tr}(H)I/2, \quad Q_{\text{diag}} = \operatorname{diag}(H_{11}, H_{22}). \quad (17)$$

All three weights have the same trace. The isotropic weight treats all directions equally. The diagonal weight retains separate penalties along the coordinate axes but omits the cross term associated with H_{12} . If the principal axes of K coincide with the coordinate axes, $Q_{\text{diag}} = H$; otherwise the two weights generally differ.

Proposition 2.2 (Energy bounds for isotropic and componentwise weights). *Under the assumptions of Theorem 2.1, in two dimensions let Q be either weight in (17) and define*

$$J_Q = \|r_c\|_{-1,K}^2 + \int_{\Omega} \eta^\top Q \eta \, dx.$$

Then

$$\mathcal{E}_{\text{mix}}^2 \leq (2 + \sqrt{2})J_Q, \quad \|p\|_{K^{-1}}^2 \leq 2J_Q. \quad (18)$$

Both bounds remain valid when the dual conservation norm is replaced by its L^2 upper bound. A reverse bound with a constant independent of the coefficient contrast does not hold uniformly for either weight.

Proof. We first compare each weight with H pointwise. Since the eigenvalues of H are positive, its largest eigenvalue is at most their sum. Thus

$$H \preceq \text{tr}(H)I = 2Q_{\text{iso}}.$$

For the diagonal weight, positive definiteness gives $H_{12}^2 < H_{11}H_{22}$. Hence, for any $a, b \in \mathbb{R}$,

$$2|H_{12}ab| \leq 2\sqrt{H_{11}H_{22}}|ab| \leq H_{11}a^2 + H_{22}b^2.$$

Writing $\zeta = (a, b)^\top$, we obtain

$$\zeta^\top H \zeta = H_{11}a^2 + 2H_{12}ab + H_{22}b^2 \leq 2(H_{11}a^2 + H_{22}b^2) = 2\zeta^\top Q_{\text{diag}} \zeta.$$

Both weights therefore satisfy $Q \succeq H/2$ almost everywhere.

Use the fields X, Y, W from the proof of Theorem 2.1. The weight comparison and orthogonality give

$$\begin{aligned} J_Q &\geq \|r_c\|_{-1,K}^2 + \frac{1}{2} \|\eta\|_{K^{-1}}^2 \\ &= \|Y\|^2 + \frac{1}{2} \|X + Y\|^2 + \frac{1}{2} \|W\|^2. \end{aligned}$$

The terms involving X and Y expand as

$$\|Y\|^2 + \frac{1}{2} \|X + Y\|^2 = \frac{1}{2} \|X\|^2 + (X, Y)_{L^2} + \frac{3}{2} \|Y\|^2.$$

The associated matrix and its eigenvalues are

$$B_{1/2} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 3/2 \end{pmatrix}, \quad \lambda_{\pm} = 1 \pm \frac{1}{\sqrt{2}}.$$

Let $c = 1 - 1/\sqrt{2}$. Applying the lower spectral bound and using $0 < c < 1/2$, we find

$$J_Q \geq c(\|X\|^2 + \|Y\|^2) + \frac{1}{2} \|W\|^2 \geq c\mathcal{E}_{\text{mix}}^2.$$

Since $c^{-1} = 2 + \sqrt{2}$, this proves the first bound in (18). For the flux, the same estimate before the spectral comparison gives directly

$$2J_Q \geq 2\|Y\|^2 + \|X + Y\|^2 + \|W\|^2 \geq \|Y\|^2 + \|W\|^2 = \|p\|_{K^{-1}}^2.$$

Finally, (9) implies

$$\lambda_{\text{th}} \|r_c\|_{L^2}^2 + \int_{\Omega} \eta^\top Q \eta \, dx \geq J_Q.$$

Replacing J_Q by this larger quantity therefore preserves both upper error bounds.

To examine the reverse estimate, we construct admissible pairs for which $J_Q/\mathcal{E}_{\text{mix}}^2$ becomes arbitrarily large. On $\Omega = (0, 1)^2$, take the constant tensors

$$H = \text{diag}(1, \kappa), \quad K = H^{-1} = \text{diag}(1, \kappa^{-1}), \quad \kappa \geq 1,$$

and choose $f = 0$, $u = 0$, $q = (1, 0)^\top$. For each finite κ , the coefficient is uniformly elliptic and the exact solution and flux are zero. The constant field q belongs to $H(\operatorname{div}; \Omega)$; no flux boundary condition is imposed in this fully Dirichlet problem. Since $r_c = 0$ and $\eta = q$, the unit area of Ω gives

$$\mathcal{E}_{\text{mix}}^2 = \int_{\Omega} q^\top H q \, dx = 1, \quad J_{Q_{\text{iso}}} = \int_{\Omega} q^\top \frac{1 + \kappa}{2} I q \, dx = \frac{1 + \kappa}{2}.$$

Thus

$$\frac{J_{Q_{\text{iso}}}}{\mathcal{E}_{\text{mix}}^2} = \frac{1 + \kappa}{2}. \quad (19)$$

For the diagonal weight, keep the same domain and rotate the coefficient and flux by 45° . With

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \tilde{H} = R H R^\top = \frac{1}{2} \begin{pmatrix} 1 + \kappa & 1 - \kappa \\ 1 - \kappa & 1 + \kappa \end{pmatrix},$$

set $\tilde{K} = \tilde{H}^{-1}$ and $\tilde{q} = R q = (1, 1)^\top / \sqrt{2}$, retaining $f = u = 0$. The flux is still constant, so its conservation residual is zero. Moreover,

$$\tilde{q}^\top \tilde{H} \tilde{q} = q^\top H q = 1, \quad \tilde{Q}_{\text{diag}} = \frac{1 + \kappa}{2} I.$$

The mixed energy error is again one, while $J_{\tilde{Q}_{\text{diag}}} = (1 + \kappa)/2$. Both examples have coefficient contrast κ , and their loss-to-error ratios diverge as $\kappa \rightarrow \infty$. Hence no inequality $J_Q \leq C \mathcal{E}_{\text{mix}}^2$ can hold for either weight with C independent of that contrast. The same examples also apply after the L^2 replacement, because their conservation residuals vanish. $\square$

These weights retain an upper error bound, but can penalize some directions too strongly relative to the energy error. Comparing weights with the same trace separates this directional distribution from the overall scale. The next example shows how the missing coupling can affect the minimizer itself.

How coupling changes the selected approximation. An upper error bound alone does not determine which approximation a loss will select. The following example isolates this distinction. It is a mixed diffusion instance of the classical geometry of changing an inner product[17, 24].

Proposition 2.3 (Approximation selected by componentwise weighting). *Let $\Omega = (0, 1)^2$, let K be constant and symmetric positive definite, and prescribe $f = 0$ and homogeneous Dirichlet data. Write*

$$H = K^{-1} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad \rho = \frac{b}{\sqrt{ac}}, \quad a, c > 0, \quad ac > b^2.$$

For fixed $s > 0$, consider the admissible trial family

$$u_t = 0, \quad q_t = s(1, t)^\top, \quad t \in \mathbb{R}. \quad (20)$$

Let $D(x) = \operatorname{diag}(d_1(x), d_2(x))$ be independent of t , with positive integrable entries. The loss with constitutive weight D selects $t_D = 0$. The same minimizer is obtained with any finite nonempty set of sampling points, positive quadrature weights and positive finite values of d_1, d_2 at those points. The best mixed-energy approximation has $t_H = -b/c$, and

$$\frac{\mathcal{E}_{\text{mix}}(t_D)}{\min_t \mathcal{E}_{\text{mix}}(t)} = \frac{1}{\sqrt{1 - \rho^2}}. \quad (21)$$

If $H = R(\gamma) \operatorname{diag}(1, \kappa) R(\gamma)^\top$, with $R(\gamma)$ a planar rotation and $\kappa \geq 1$, then

$$\frac{1}{1 - \rho^2} = 1 + \frac{(\kappa - 1)^2}{4\kappa} \sin^2(2\gamma). \quad (22)$$

Proof. The exact potential and flux are zero. Every q_t is constant and belongs to $H(\text{div}; \Omega)$, so $r_c = 0$ and $\eta = q_t$. No flux boundary condition is imposed in this fully Dirichlet problem. The conservation term therefore vanishes, and the continuous loss is

$$J_D(t) = s^2(A + Bt^2), \quad A = \int_{\Omega} d_1 dx > 0, \quad B = \int_{\Omega} d_2 dx > 0.$$

Its unique minimizer is $t_D = 0$. For the discrete loss, the coefficients become $A = \sum_i w_i d_1(x_i)$ and $B = \sum_i w_i d_2(x_i)$. Both remain positive, so the minimizer is unchanged. This argument does not require exact quadrature of the variable weights.

Since Ω has unit area, the mixed-energy error satisfies

$$\begin{aligned} \mathcal{E}_{\text{mix}}(t)^2 &= s^2(a + 2bt + ct^2) \\ &= s^2 \left[c \left(t + \frac{b}{c} \right)^2 + a - \frac{b^2}{c} \right]. \end{aligned}$$

Thus $t_H = -b/c$, with minimum squared error $s^2(ac - b^2)/c$. At $t_D = 0$ the squared error is s^2a . Dividing and taking square roots proves (21). Finally, rotation gives $\det H = \kappa$ and $b^2 = (\kappa - 1)^2 \sin^2(2\gamma)/4$. Substitution into $ac = \det H + b^2$ proves (22). $\square$

At $\gamma = \pi/4$, the error ratio is $(\kappa + 1)/(2\sqrt{\kappa})$, regardless of the positive diagonal weights. It grows with the coefficient contrast, whereas aligned principal axes give a ratio of one. The cause is the missing term $2bt$: changing the coefficients of 1 and t^2 cannot recover it. This discrepancy is present at the exact minimum of a strictly convex scalar loss, for any sampling locations and arbitrarily small s .

The example concerns the restricted family (20), which does not contain the exact zero flux. It does not prescribe the outcome over a larger neural network class. Its conclusion applies to fixed weights in the stated coordinates; rotating the residual components or allowing weights to depend on t changes the problem.

The analytical weight retains the cross term and gives $J_H(t) = \mathcal{E}_{\text{mix}}(t)^2$, hence selects t_H . This equality also holds for the discrete loss when the quadrature weights sum to one, since the integrand is constant. It is a simple instance of the energy orthogonality for an exactly conserved flux[26]. The next result shows how closely the probe construction retains this coefficient scale.

Proposition 2.4 (Finite-probe scale bounds). *Let $P_N, S_N \succ 0$, and set*

$$m_N = [\lambda_{\min}(P_N)\lambda_{\min}(S_N)]^{1/2}, \quad h_N = [\lambda_{\max}(P_N)\lambda_{\max}(S_N)]^{1/2}.$$

Then, almost everywhere,

$$m_N K \preceq M_N^{\#} \preceq h_N K. \quad (23)$$

The constants depend only on the shared probe banks. If $P_N = S_N = I$, then $M_N^{\#} = K$. Transporting both response moments by $A(\cdot)A^{\top}$ transports $M_N^{\#}$ by the same congruence.

Indeed, if $p_-I \preceq P_N \preceq p_+I$ and $s_-I \preceq S_N \preceq s_+I$, then $s_-K^2 \preceq Q_N \preceq s_+K^2$. Monotonicity and $I \# K^2 = K$ yield (23). Congruence follows from the characterization of $X \# Y$ as the unique positive-definite solution of $ZX^{-1}Z = Y$. The full proof is in Supplementary Section S1.

The spectral bounds quantify how finite probe banks change the analytical loss. Inverting (23) gives

$$\min\{1, h_N^{-1}\} \mathcal{J}_{\text{stab}} \leq \mathcal{J}_{N, \text{probe}} \leq \max\{1, m_N^{-1}\} \mathcal{J}_{\text{stab}}. \quad (24)$$

The constants depend on the sampled directions, not on the coefficient contrast. Combining the left inequality with (16) gives

$$\mathcal{E}_{\text{mix}}^2 \leq \varphi^2 \max\{1, h_N\} \mathcal{J}_{N, \text{probe}}. \quad (25)$$

For the constant coefficient and zero-source problem in Proposition 2.3, the same bounds also control the selected approximation. Let $\mathcal{F}$ be any nonempty prescribed set of constant flux candidates with $u = 0$, and suppose q_N minimizes $J_N(q) = q^\top (M_N^\#)^{-1} q$ on $\mathcal{F}$. Then

$$h_N^{-1} \mathcal{E}_{\text{mix}}(q)^2 \leq J_N(q) \leq m_N^{-1} \mathcal{E}_{\text{mix}}(q)^2,$$

so, for every $q \in \mathcal{F}$,

$$\mathcal{E}_{\text{mix}}(q_N)^2 \leq h_N J_N(q_N) \leq h_N J_N(q) \leq (h_N/m_N) \mathcal{E}_{\text{mix}}(q)^2.$$

Taking the infimum gives

$$\mathcal{E}_{\text{mix}}(q_N) \leq \sqrt{h_N/m_N} \inf_{q \in \mathcal{F}} \mathcal{E}_{\text{mix}}(q). \quad (26)$$

The factor depends only on the probe banks. With constant K and shared banks, $M_N^\#$ and the integrand are spatially constant, so normalized positive quadrature gives the same result. This establishes a direct link between the probe construction and approximation quality in the conserved constant-flux setting.

Finally, the transformation rule from Section 2.2 applies to both constructions. To state it in terms of a scale tensor, let $r_\theta(x) \in \mathbb{R}^d$ be a residual block evaluated at a common point and with a common integration measure. Write the weight as M^{-1} , where $M(x)$ is measurable, symmetric positive definite and fixed during training. The objective is

$$\mathcal{Q}[r, M](\theta) = \frac{1}{d} \int_{\Omega} r_\theta^\top M^{-1} r_\theta \, dx. \quad (27)$$

For any measurable invertible matrix field $A(x)$ independent of θ , set $r' = Ar$ and $M' = AMA^\top$. Then

$$(Ar)^\top (AMA^\top)^{-1} (Ar) = r^\top M^{-1} r. \quad (28)$$

Thus the objectives and all parameter derivatives that exist are identical. In exact arithmetic, identical initial states, batches, random inputs, schedules and deterministic update rules also produce identical finite training trajectories. Floating-point implementations can differ through their arithmetic and differentiation paths.

For the diffusion residuals in (4), choose $A = K^{-1}$ and $M = K$. The common constitutive energy is

$$\eta^\top K^{-1} \eta = \xi^\top K \xi. \quad (29)$$

This is the coefficient weighting used in mixed least-squares analysis. A metric selected in one residual coordinate system is transported by (28) to any other. Recomputing an isotropic ridge term or a diagonal approximation separately in each system generally changes the metric. Supplementary Section S1 gives the transport proof and its assumptions.

2.6 Neural approximation and boundary conditions

The analysis specifies how the residuals are weighted. We now describe the network, boundary conditions and sampled loss used to compute the approximations. The network predicts the flux independently of the potential. Setting the flux equal to $-K \nabla u_\theta$ would instead eliminate the constitutive residual and define a different approximation method.

We use a fully connected 2–32–32–32–32–3 tanh network with 3363 trainable parameters. It takes $(2x - 1, 2y - 1)$ as inputs and supplies one potential channel and two independent flux channels. On the fully Dirichlet manufactured problems, multiplying the potential channel by $x(1 - x)y(1 - y)$ imposes the homogeneous data exactly. Output scales, initialization and arithmetic precision for each study are specified in Section 3.1.

For the Darcy and heat-conduction applications, split the boundary into the two vertical edges Γ_D and the two horizontal edges Γ_N . Prescribe the hydraulic head on Γ_D and the outward normal flux on Γ_N . If these data are imposed exactly by both the exact solution and the neural approximation, their errors belong to

$$V = \{v \in H^1(\Omega) : v|_{\Gamma_D} = 0\}, \quad W_N = \{w \in H(\text{div}; \Omega) : w \cdot n|_{\Gamma_N} = 0\}. \quad (30)$$

The energy-dual norm is then defined using test functions in V . For $p \in W_N$ and $v \in V$, the boundary contribution in integration by parts vanishes: $v = 0$ on Γ_D and $p \cdot n = 0$ on Γ_N . The weighted projection in the proof of Theorem 2.1 therefore applies with $H_0^1(\Omega)$ replaced by V , giving the same sharp equivalence and reliability constants.

On the unit square, the one-dimensional Poincaré inequality in the x direction gives $\|v\|_{L^2} \leq \pi^{-1} \|\nabla v\|_{L^2}$ on V . This constant is sharp, as seen from $v(x, y) = \sin(\pi x)$. Accordingly, the conservation weight for this boundary partition is

$$C_{P,D} = \pi^{-1}, \quad \lambda_{\text{th},D} = \frac{1}{\pi^2 \alpha}. \quad (31)$$

The change from the full Dirichlet case follows from the error space. Applying the full-boundary constant without this adjustment would not give the stated bound.

For unit head at $x = 0$, zero head at $x = 1$, and impermeable horizontal edges, an admissible network representation is

$$u_\theta = 1 - x + x(1 - x)\tilde{u}_\theta, \quad q_{\theta,x} = \tilde{q}_{\theta,x}, \quad q_{\theta,y} = y(1 - y)\tilde{q}_{\theta,y}. \quad (32)$$

These constraints impose the prescribed head and normal flux exactly. The residual definitions and constitutive weights remain those in Sections 2.3 and 2.4.

For the heat-conduction problem with zero temperature on both vertical edges, omit the term $1 - x$ in (32).

2.7 Training objective and optimization

Let $\{x_i\}_{i=1}^n$ be fixed uniform interior points and let W_i denote the selected constitutive weight. The sampled objective is

$$\hat{J}(\theta) = \frac{1}{n} \sum_{i=1}^n \left[\lambda (\text{div } q_\theta(x_i) - f(x_i))^2 + \eta_\theta(x_i)^\top W_i \eta_\theta(x_i) \right]. \quad (33)$$

The unit-square area is one; on a general domain, the corresponding uniform Monte Carlo quadrature includes its volume. The coefficient values and weight matrices are computed before optimization and remain fixed during training. Differentiation is through the network outputs and the first spatial derivatives needed for the mixed residual. No reference solution is used to construct W_i or to evaluate the training loss.

For the analytic choice, the quadratic form can be evaluated by solving $K(x_i)w_i = \eta_\theta(x_i)$ and forming $\eta_\theta(x_i)^\top w_i$. The probe choice uses the same operation with $M_N^\#(x_i)$. Equivalently, for fixed training points, the weight matrices can be prepared once and contracted with the residual at each update; the fixed-point probe and application comparisons use this implementation. The scale matrix remains fixed during training, so its construction does not add trainable parameters or a feedback rule based on the current residual. In two dimensions these are small local matrix operations; their purpose is to change the residual geometry seen by the optimizer.

The numerical procedure is as follows.

1. Specify K , f , the essential boundary conditions and the ellipticity lower bound α . Choose a boundary-admissible potential-flux network.
2. Fix the training points and initial network parameters. For a probe run, draw the two direction banks, form P_N , S_N and the matrices $M_N^\#(x_i)$, and retain the banks for reproducibility.
3. Select the conservation weight λ and constitutive weight W_i . Keep the architecture, points, initialization and optimizer budget common across the paired comparisons.
4. Evaluate the mixed residuals by automatic differentiation and minimize (33) using the specified Adam and L-BFGS stages. Retain the final endpoint and its realized work.

5. Evaluate potential, flux and mixed-energy errors on points excluded from optimization. Inspect the solution and error fields using common scales across methods.

The sampled loss approximates a continuous integral. For general neural fields, applying the continuous reliability bound to this loss requires a quadrature estimate, as in residual-stability analyses of PINN generalization[22]. Proposition 2.3 avoids this issue only for its specific constant-flux family: positive quadrature leaves the diagonal minimizer unchanged, and integrates the analytical and probe energies exactly when normalized. For the trained networks, we evaluate solution errors independently of the training objective.

Automatic differentiation evaluates ∇u_θ and $\text{div } q_\theta$. Write $\widehat{\mathbb{E}}$ for the arithmetic mean over training points. The six compared losses are specified in Table 1; all use the physical flux residual η . The constitutive quadratic form is summed over components, except for the explicit factor $1/2$ in the original Euclidean baseline.

Table 1: Losses in the operator-probe comparison. Each objective is $\lambda \widehat{\mathbb{E}}[r_c^2] + \widehat{\mathbb{E}}[\eta^\top W \eta]$. The two Euclidean baselines are the prespecified primary comparators.

Method	λ	W
Operator probe	λ_{th}	$(M_{32}^\#)^{-1}$
Matched Euclidean	λ_{th}	I
Original Euclidean	1	$I/2$
Trace-matched isotropic	λ_{th}	$\text{tr}(K^{-1})I/2$
Coordinate diagonal	λ_{th}	$\text{diag}((K^{-1})_{11}, (K^{-1})_{22})$
Analytic	λ_{th}	K^{-1}

Two constitutive-energy implementations. The longer study evaluates the analytic energy through flux and inverse-diffusivity paths, denoted F and I. Both first form $\eta = q + K\nabla u$. Path F solves $Kw = \eta$ and evaluates $\eta^\top w$; path I uses the same solve to form $\xi = K^{-1}\eta$ and evaluates $\xi^\top K\xi$. The comparison therefore tests two energy-evaluation paths after shared residual formation. Its Euclidean development baselines use $\widehat{\mathbb{E}}[r_c^2] + \frac{1}{2}\widehat{\mathbb{E}}[|\eta|^2]$ in F and $\widehat{\mathbb{E}}[r_c^2] + \frac{1}{2}\widehat{\mathbb{E}}[|\xi|^2]$ in I; these are distinct coordinate-specific objectives. Same-state loss and gradient differences are measured separately from the final errors of independent training trajectories.

3 Results

3.1 Experimental setup

The experiments examine three questions: whether retaining directional coupling changes accuracy, how the probe construction compares with analytical and simpler weights, and whether equivalent constitutive-energy evaluations agree numerically. All methods use the network and boundary constructions in Section 2.6. Manufactured problems use output scales $D_1 = \text{diag}(16, 20, 20)$, Xavier-normal initialization and zero biases; the applications use unit output scales. Parameters and residuals are evaluated in float64. The independent 84-run comparison initializes the networks in float32 before conversion.

The independent comparison uses two manufactured problems and six random seeds for network initialization and point sampling. For each seed, the two PDEs share two independently drawn probe realizations. Each PDE is solved with five control methods and the two probe realizations, giving 60 control runs and 24 probe runs in total. Because the same control enters both probe-to-control ratios, uncertainty is calculated from the six independent seed groups, rather than from 84 independent observations. Within each seed-PDE group, all methods use the same initial parameters and 512 training points. Training consists of 2000 Adam updates at learning rate 10^{-3} , followed by L-BFGS with strong-Wolfe line search and at least

6000 objective and gradient evaluations. We compare the models after the final complete accepted step, without selecting them by validation error. Supplementary Section S2 gives the seeds and stopping rules.

The primary outcome on a separate point set G is the relative mixed-energy error

$$E_{\text{mix}} = \left(\frac{\|u_\theta - u_*\|_{K,G}^2 + \|q_\theta - q_*\|_{K^{-1},G}^2}{\|u_*\|_{K,G}^2 + \|q_*\|_{K^{-1},G}^2} \right)^{1/2}, \quad (34)$$

where the squared discrete norms replace the integrals in (5) by point averages. We also report relative potential and flux L^2 errors, E_u and E_q , and separate energy and residual diagnostics. Exact manufactured fields are used for evaluation. The 16384 independent uniform test points are used for error evaluation only after all 84 runs are complete.

For seed group j , PDE c and probe realization k , let E_{jck}^P and E_{jc}^b denote the probe and comparator errors. We first average the logarithms of their ratios within each seed group, then calculate the geometric-mean ratio G_b and its one-sided upper confidence limit U_b :

$$\begin{aligned} d_{j,b} &= \frac{1}{4} \sum_{c=1}^2 \sum_{k=1}^2 \log \frac{E_{jck}^P}{E_{jc}^b}, & G_b &= \exp(\bar{d}_b), \\ U_b &= \exp\left(\bar{d}_b + t_{0.95,5} \frac{s_b}{\sqrt{6}}\right), & \bar{d}_b &= \frac{1}{6} \sum_{j=1}^6 d_{j,b}, \end{aligned} \quad (35)$$

where s_b is the sample standard deviation of the six group means. The comparison rule was fixed before the experiments: all runs must complete successfully and, for each Euclidean baseline, both G_b and U_b must be below one, as must the ratio within each PDE. This is an intersection–union test. The t calculation assumes independent, approximately Gaussian group means and treats the two PDEs and evaluation points as fixed. The analytical, isotropic and diagonal comparisons are descriptive.

Separate exploratory experiments help interpret the weights. A four-seed factorial comparison changes the conservation and constitutive weights independently. A trace-matched comparison holds the conservation weight and the sum of the constitutive eigenvalues fixed, so that the remaining difference is how the weight is distributed across directions. These studies use 4000 Adam updates followed by approximately 6000 L-BFGS evaluations. Earlier probe comparisons and longer training histories are reported in Supplementary Sections S2 and S3; their results are not pooled with the independent six-seed comparison.

The Darcy and heat-conduction cases illustrate predicted solution and error fields for prescribed nondimensional boundary-value problems. Each case uses three methods: matched Euclidean weighting, analytical weighting and the 32-direction operator-probe weighting from Table 1. A single target seed, 92101, and probe seed, 93101, are fixed for these six runs. Within each case the methods share the initialization, 8192 fixed uniform interior points, architecture, boundary construction and training schedule. Application outputs use unit scales and float64 arithmetic.

Each application run uses 2000 Adam updates at learning rate 10^{-3} , followed by L-BFGS until at least 48000 objective and gradient evaluations have been completed. Finishing the accepted step gives 48000–48002 evaluations across the six runs. The figures and tables use these final models, without selection by reference error. The applications illustrate the fields for one seed; the statistical comparison uses the manufactured problems.

Independent tensor-product Q_2 finite-element solves on 512×512 elements provide the reference potentials, using 4×4 Gauss integration per element. Dirichlet values are imposed strongly; insulating conditions enter the weak formulation naturally. The raw constitutive flux $q_h = -K \nabla u_h$ is generally not globally in $H(\text{div}; \Omega)$, and its normal trace is not constrained to equal the prescribed flux pointwise. It serves as a mesh-refined field reference, not an exactly equilibrated flux. Tables 4 and 5 give relative potential L^2 , flux-energy ($E_{q,K^{-1}}$) and mixed-energy errors for the displayed endpoints. No reference values enter training. Error normalizations, evaluation quadrature and the observed reference-refinement differences are specified in Supplementary Sections S2 and S4.

3.2 Manufactured-solution benchmarks

The experiments use three smooth manufactured solutions with different coefficient structures. Let $R(\vartheta)$ denote the two-dimensional rotation matrix. The constant rotated problem is

$$K = R(\pi/6) \text{diag}(12, 0.3) R(\pi/6)^\top, \quad u_* = \sin(\pi x) \sin(\pi y), \quad (36)$$

with $\alpha = 0.3$. The variable diagonal problem is

$$K = \text{diag}(0.5 + 4x, 0.1 + 0.9y), \quad u_* = \sin(2\pi x) \sin(\pi y), \quad (37)$$

with $\alpha = 0.1$. The oscillatory rotated problem is

$$\begin{aligned} K &= R(\vartheta) \text{diag}(1 + 9x^2, 0.05 + 0.45y) R(\vartheta)^\top, & \vartheta &= \frac{\pi}{3} \sin(2\pi x) \sin(\pi y), \\ u_* &= \sin(\pi x) \sin(2\pi y), & \alpha &= 0.05. \end{aligned} \quad (38)$$

In each case, $q_* = -K \nabla u_*$ and $f = \text{div } q_*$. The confirmation uses the constant and oscillatory problems. The diagonal problem provides an additional development case in which the analytic and coordinate-diagonal constitutive weights coincide.

Figure 1 shows the coefficient contrast and principal diffusion direction. The constant problem isolates a fixed rotation, the diagonal problem varies the principal values without rotating the axes, and the oscillatory problem combines both changes. These differences matter for interpreting the diagonal control: it is identical to the analytic metric in the diagonal case, but discards cross-direction coupling in the rotated cases.

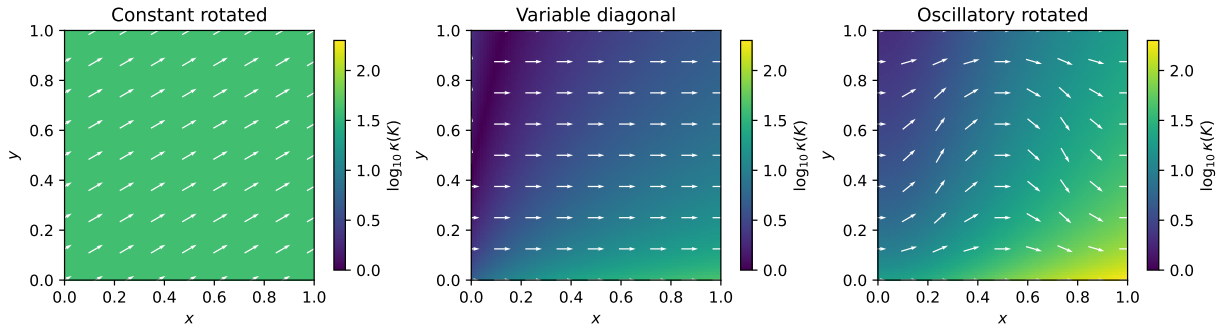


Figure 1: Coefficient fields of the manufactured benchmarks. Color gives $\log_{10} \kappa(K)$, where $\kappa(K)$ is the ratio of the largest to smallest eigenvalue. Arrows show a principal direction associated with the larger eigenvalue; they represent coefficient orientation, not physical flux. The common color scale permits comparison across problems.

Table 2 compares probe weighting with five alternatives, from Euclidean losses to the full analytical tensor. The geometric-mean mixed-energy error is approximately 37% lower than with the conservation-matched Euclidean loss and 56% lower than with the original Euclidean loss. Both comparisons satisfy the prespecified rule in Section 3.1: the PDE-specific ratios and one-sided upper confidence limits are below one. Each of the six independent seed groups also favors probe weighting over these two baselines (Figure 2).

The differences become smaller when the comparison weights contain more coefficient information. The probe-to-isotropic ratio is 0.760, corresponding to a 24% reduction. Probe weighting also gives a slightly lower geometric-mean mixed-energy error than analytical weighting: the ratio is 0.931, or a reduction of approximately 6.9%. The PDE-specific ratios are 0.972 and 0.891 for the constant and oscillatory problems. Against the coordinate-diagonal weight, the ratio is 0.917. The 95% confidence intervals for the analytical and diagonal comparisons both include one, so these smaller observed errors do not establish statistical superiority. The controlled comparisons below examine the effect of directional coupling more directly.

All 84 runs completed successfully. Repeating the error calculation on the two halves of the independent test set gives results close to those on the full set. Supplementary Section S3 reports these checks and additional comparisons.

Table 2: Probe-to-control mixed-energy error ratios on 16384 held-out points. G and U are defined in (35); CI_{95} is the two-sided 95% interval for G . The first two rows are the prespecified primary comparisons; the other rows are descriptive. The last two columns give PDE-specific geometric means.

Comparator	G	U	CI_{95}	Constant	Oscillatory
Matched Euclidean	0.628	0.806	[0.457, 0.864]	0.639	0.617
Original Euclidean	0.435	0.501	[0.364, 0.521]	0.391	0.485
Trace-matched isotropic	0.760	0.893	[0.618, 0.934]	0.735	0.785
Coordinate diagonal	0.917	1.065	[0.757, 1.110]	0.880	0.955
Analytic K^{-1}	0.931	1.072	[0.777, 1.115]	0.972	0.891

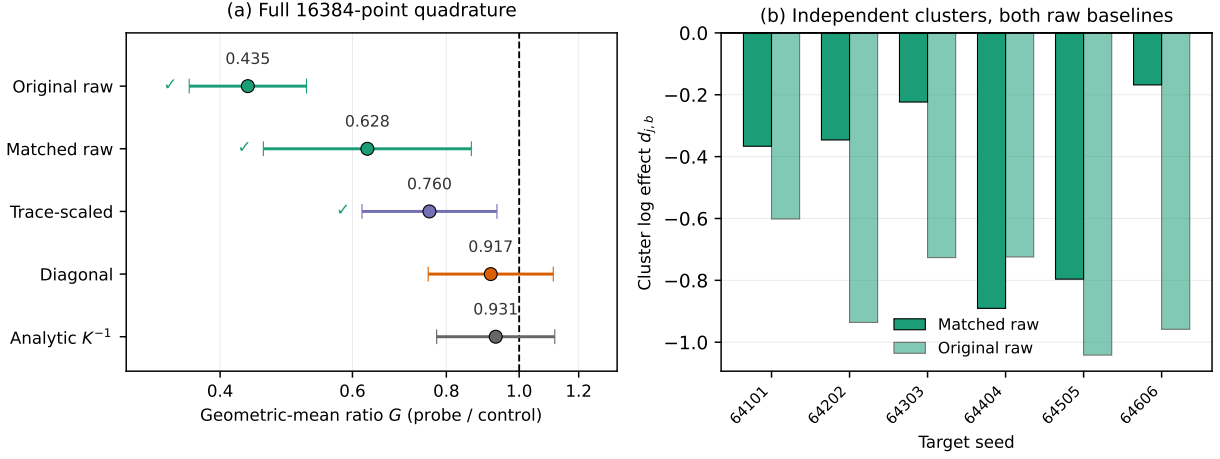


Figure 2: Held-out comparison at the fixed training budget. Left: geometric-mean probe-to-control ratios with two-sided 95% intervals. Right: the six independent target-cluster log effects for each primary comparator. The check marks in the left panel indicate ratios and one-sided upper limits below one; only the two Euclidean comparisons are prespecified primary outcomes.

Aggregate error measures do not show the spatial distribution of errors. Figures 3 and 4 compare predictions at the end of training for the oscillatory problem. We use target seed 64101 and its first probe bank, 75101, selected before viewing the fields. The figures compare the two Euclidean baselines, analytical weighting and probe weighting. The quantitative tables include all six methods.

The potential maps recover the two principal lobes of the reference solution, while their error maps reveal spatial differences that are difficult to identify from the predictions alone. On the boundary-inclusive plotting grid, the relative potential L^2 errors are 6.470×10^{-3} , 7.068×10^{-3} , 2.893×10^{-3} and 3.446×10^{-3} for the original Euclidean, matched Euclidean, analytic and probe methods, respectively. Thus the analytic metric gives the smallest potential error among these four training results. The probe gives the smallest vector-flux L^2 error, 8.581×10^{-3} , compared with 1.500×10^{-2} , 1.046×10^{-2} and 1.844×10^{-2} for the other three methods. The potential and flux rankings consequently differ.

The flux-error maps show errors near the boundary that are less apparent in the potential field. These local errors are not apparent from E_{mix} alone. The constant rotated example in Supplementary Section S3 shows another tradeoff: for the displayed target, the probe has a slightly larger vector-flux L^2 error than the original Euclidean method, despite a smaller mixed-energy error. These plots describe individual training runs; the six-group estimates in Table 2 remain the basis for the prespecified comparison.

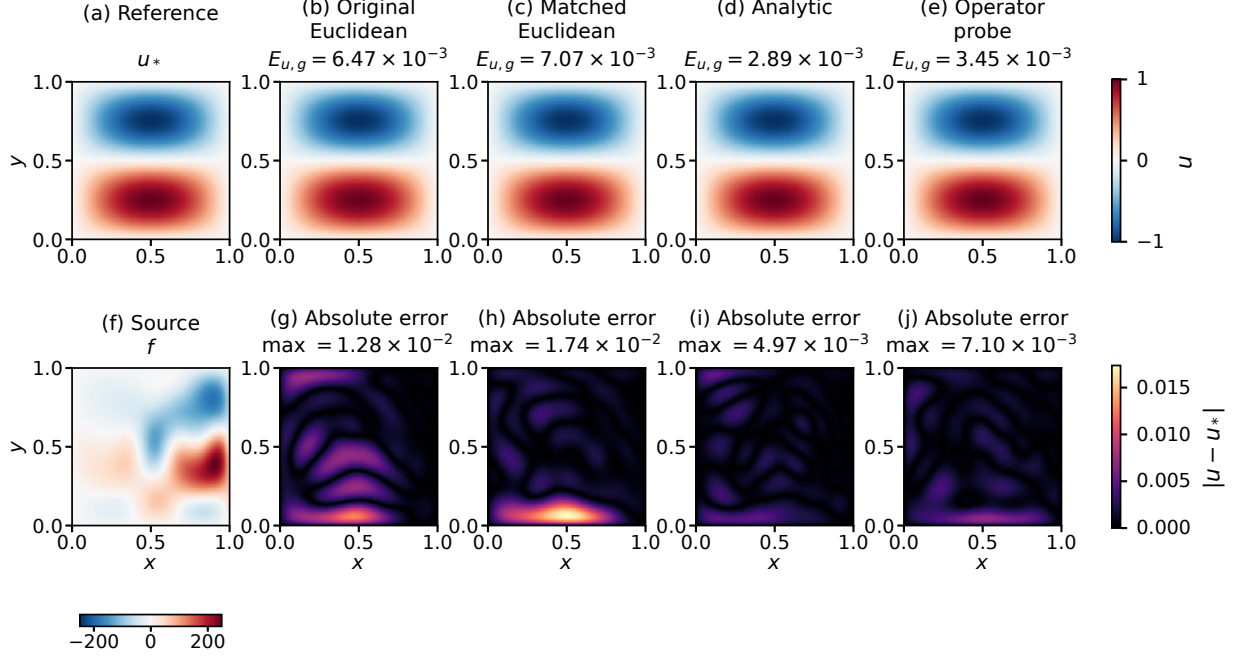


Figure 3: Potential predictions and absolute errors for the oscillatory rotated problem. The reference is the manufactured exact solution; predictions use the final models for target seed 64101 and probe bank 75101. Predictions share a linear color scale, and their absolute errors share a second scale. Panel values $E_{u,g}$ are relative L^2 errors on a 129×129 plotting grid including boundary nodes. The source field is shown at lower left. These descriptive grid values are distinct from the prespecified quadrature on independent uniform interior points.

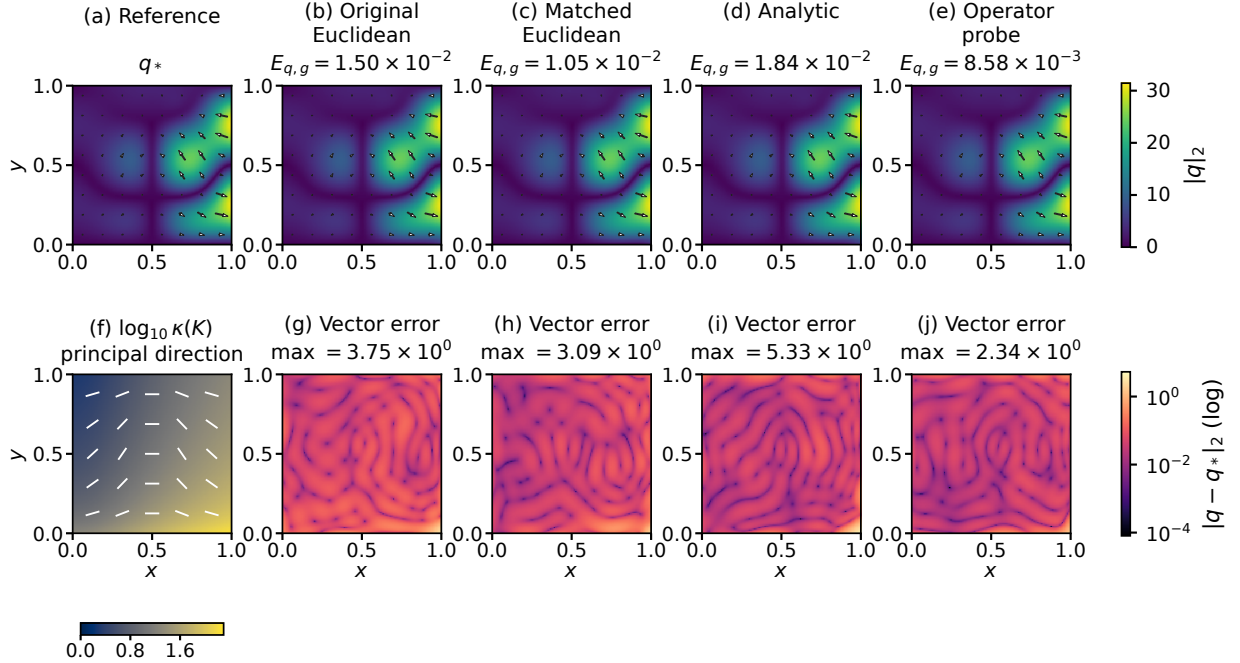


Figure 4: Physical fluxes and vector errors for the same endpoints as Figure 3. Colors in the prediction panels show $|q|_2$ and arrows show flux direction and magnitude, using common scales. The displayed flux is the network's independent flux output. Error panels show $|q - q_*|_2$ with a shared logarithmic color scale. The lower-left panel shows coefficient anisotropy and principal orientation. Panel values $E_{q,g}$ are relative vector L^2 errors on the plotting grid.

3.3 Effects of residual weighting and representation

The analytical loss changes two things at once: the conservation scale and the constitutive weight. To distinguish their effects, a four-seed factorial comparison uses

$$\begin{aligned} J_{00} &= \widehat{\mathbb{E}}[r_c^2] + \frac{1}{2}\widehat{\mathbb{E}}[|\eta|^2], & J_{10} &= \lambda_{\text{th}}\widehat{\mathbb{E}}[r_c^2] + \frac{1}{2}\widehat{\mathbb{E}}[|\eta|^2], \\ J_{01} &= \widehat{\mathbb{E}}[r_c^2] + \widehat{\mathbb{E}}[\eta^\top K^{-1}\eta], & J_{11} &= \lambda_{\text{th}}\widehat{\mathbb{E}}[r_c^2] + \widehat{\mathbb{E}}[\eta^\top K^{-1}\eta]. \end{aligned} \quad (39)$$

For each seed, let $\ell_{ab} = \log_{10} E_{\text{mix}}(J_{ab})$. The conservation and tensor effects are $\Delta_c = [(\ell_{10} - \ell_{00}) + (\ell_{11} - \ell_{01})]/2$ and $\Delta_t = [(\ell_{01} - \ell_{00}) + (\ell_{11} - \ell_{10})]/2$. Table 3 reports medians over seeds. Both changes help on the two easier problems. On the oscillatory problem, the tensor change remains beneficial, whereas conservation rescaling has a slightly adverse median effect. The full loss beats both single-change controls in only one of its four oscillatory runs. Across the three problems, this occurs in nine of twelve runs.

Table 3: Four-seed factorial ablation at the common development budget. Negative Δ_c or Δ_t indicates a reduction in log error. The tensor change includes both directional weighting and its specified normalization.

Problem	med Δ_c	med Δ_t	med(E_{11}/E_{00})	$E_{11} < \min(E_{10}, E_{01})$
Constant rotated	−0.2170	−0.2434	0.3454	4/4
Variable diagonal	−0.1884	−0.2353	0.4118	4/4
Oscillatory rotated	+0.0080	−0.3507	0.4566	1/4

Propositions 2.2 and 2.3 motivate a comparison that removes coupling while keeping the overall penalty scale fixed. We hold the conservation weight and the trace of the constitutive weight fixed, and compare K^{-1} with Q_{iso} and Q_{diag} . This requires 20 additional control runs and reuses the 12 analytical runs at the same training settings. The analytical-to-isotropic median error ratios are 0.634, 0.808 and 0.956 for the constant, diagonal and oscillatory problems; all twelve paired ratios are below one.

Against the componentwise weight, the median ratios are 0.688 for the constant rotated problem and 0.935 for the oscillatory problem. The full tensor gives smaller errors in all four runs of the former and three of four runs of the latter, whose ratios range from 0.743 to 1.106. For the variable diagonal problem the two weights coincide, so duplicate runs are unnecessary. Retaining cross-direction coupling is therefore beneficial in these comparisons, most consistently for the fixed rotation. Its effect is smaller and varies more across seeds when the principal axes also vary in space. Supplementary Section S3 gives the paired medians and ranges.

The representation identity concerns the loss evaluated at the same network parameters. We check it by computing the two analytical energy expressions and their parameter gradients from the same residuals. The maximum relative loss difference is 2.993×10^{-16} and the maximum normalized gradient difference is 4.055×10^{-14} , consistent with (29) to near machine precision.

Training the two implementations separately can still produce different final models because floating-point differences accumulate during optimization. For the analytic objective, the mixed-error sensitivities $|\log_{10}(E_F/E_I)|$ are 0.00060, 0.01499 and 0.01419 in the constant, diagonal and oscillatory problems. Across potential, flux and mixed errors, only three of the nine sensitivities are at most 0.01, a threshold corresponding to an error ratio of about 1.0233. Thus agreement of the loss and gradient at fixed parameters does not ensure identical predictions after separate training runs. The complete sensitivity table is in Supplementary Section S3.

An earlier operator-probe development study also gives a mixed result: the probe beats the matched Euclidean control in three of four target–PDE pairs, while the constant problem with target seed 62222 has an error ratio of 1.089. That study does not meet its four-of-four criterion. It motivates, but is not pooled with, the independent confirmation above; its full 24-run record is retained with the code.

3.4 Darcy flow in anisotropic porous media

The unit-square domain has hydraulic head $u = 1$ at $x = 0$ and $u = 0$ at $x = 1$, with impermeable horizontal edges and no interior source, $f = 0$. Let

$$\begin{aligned} K_D &= R(\vartheta_D) \text{diag}(k_{\parallel}, k_{\perp}) R(\vartheta_D)^{\top}, \\ k_{\parallel} &= 1 + 4 \exp \left[- \left(\frac{y - 0.5 - 0.15 \sin(2\pi x)}{0.18} \right)^2 \right], \\ k_{\perp} &= 0.05 + 0.15 \sin^2(\pi x) \sin^2(\pi y), \quad \vartheta_D = \frac{\pi}{4} \sin(2\pi x) \sin(\pi y). \end{aligned} \quad (40)$$

The higher principal permeability has a meandering spatial feature, while the principal axes rotate smoothly. The tensor has the uniform lower bound $\alpha_D = 0.05$. The head difference drives the flow, and the predicted flux provides both the local flow direction and the total throughflow.

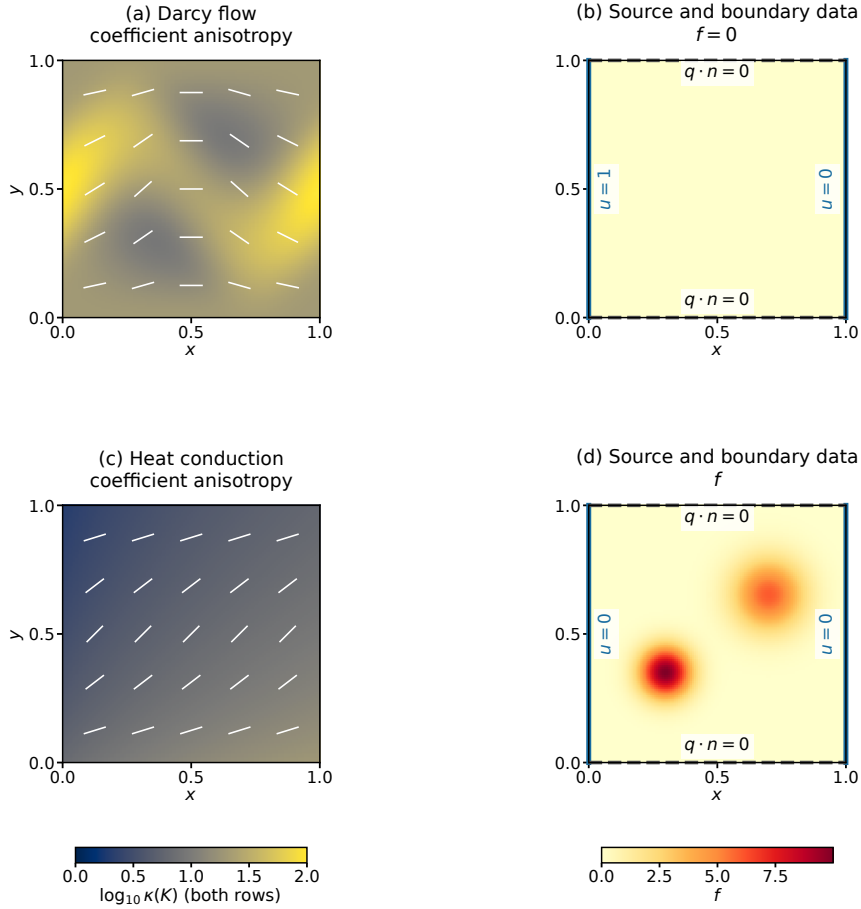


Figure 5: Material anisotropy, prescribed sources and boundary conditions for the two application benchmarks. Left panels show $\log_{10} \kappa(K)$ and the principal axis, with a common coefficient scale. Right panels show the source and the prescribed head/temperature and normal-flux data. All quantities are nondimensional.

Figure 5 summarizes the two physical settings. Their boundary values and sources are specified before solving; neither problem is constructed by choosing a target solution.

The Q_2 reference uses 512 elements per direction. From 256 to 512 elements, the potential L^2 and flux-energy differences are 2.701×10^{-8} and 5.664×10^{-5} . These observed differences assess reference resolution; Supplementary Section S4 gives the verification details.

Figures 6 and 7 compare the reference head and flux with the three neural approximations and their pointwise errors. Flux concentrates along the meandering high-permeability band. The plots use common

Table 4: Darcy flow: relative errors (percent) on 9216 Gauss points for the three displayed runs at target seed 92101 and the final common 48000-evaluation stage.

Method	E_u	$E_{q,K-1}$	E_{mix}
Matched Euclidean	0.0405	0.3114	0.2974
Analytical	0.0121	0.1586	0.1476
Operator probe	0.0124	0.1502	0.1442

field and error scales across methods. Table 4 gives the corresponding errors for these three runs.

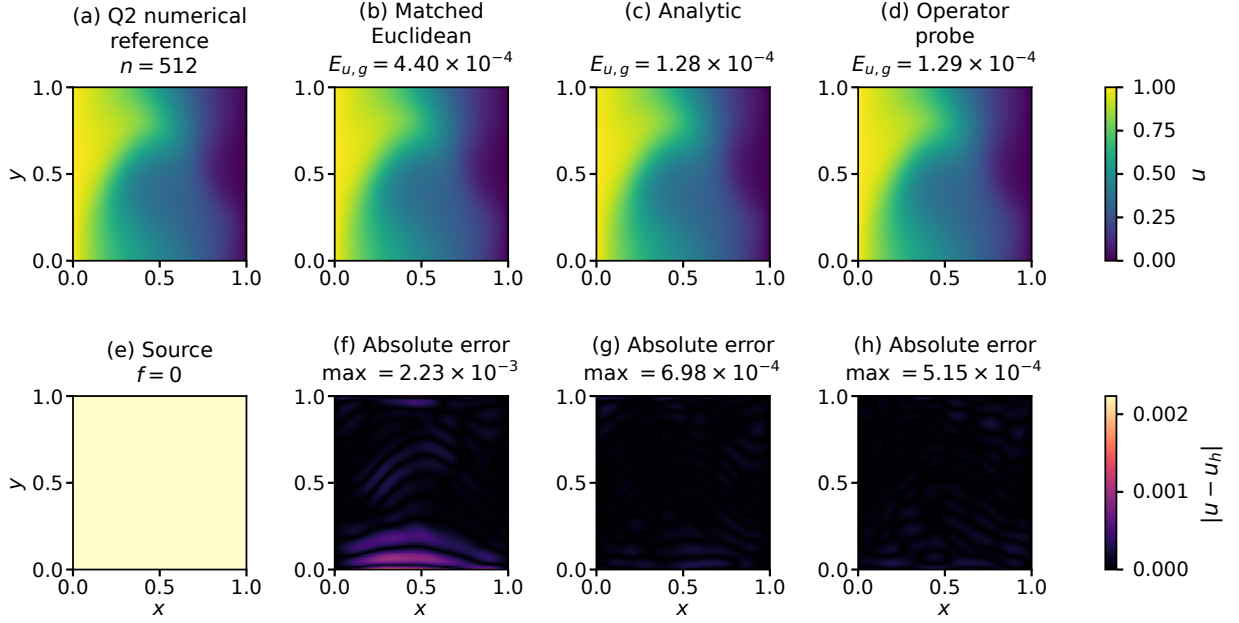


Figure 6: Darcy flow: reference potential, matched Euclidean, analytical and operator-probe predictions, and absolute errors. Field and error panels respectively share linear color scales. Relative L^2 annotations use the equally weighted 129×129 plotting grid, including its boundary; Table 4 uses 9216 Gauss points.

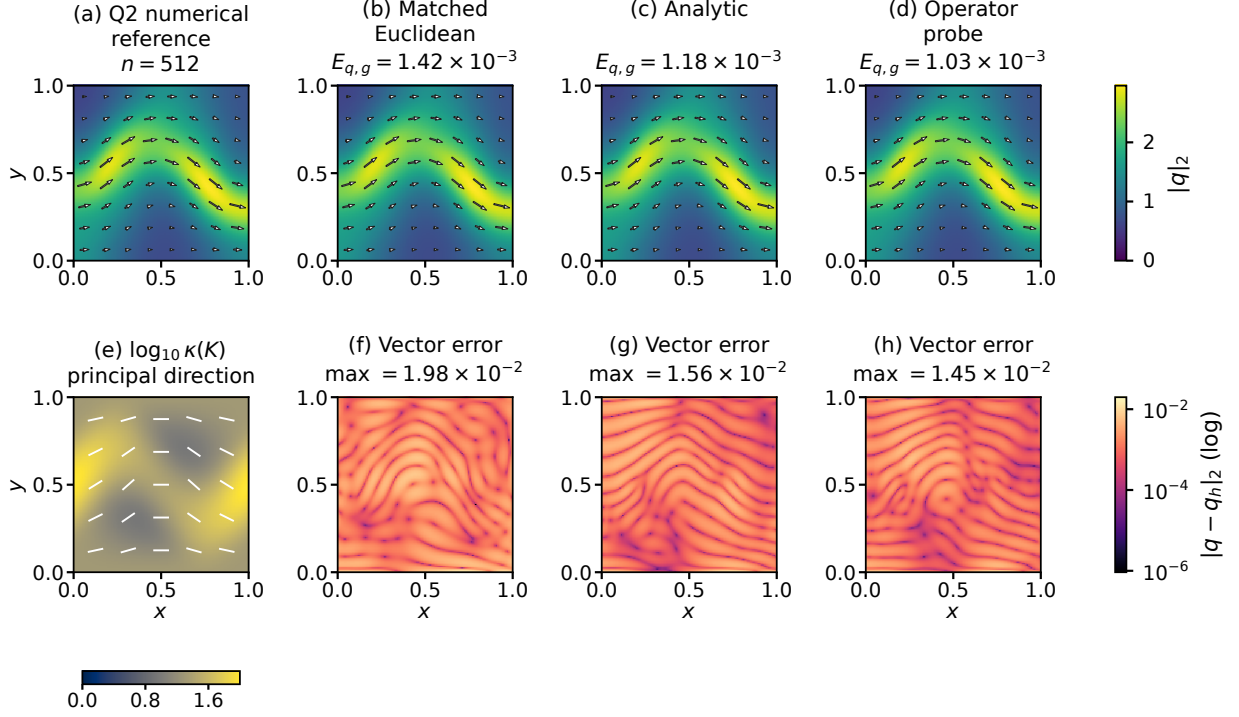


Figure 7: Darcy flow: reference and independent neural fluxes, with vector-error magnitude against $q_h = -K\nabla u_h$. Flux magnitudes share a linear scale, errors a logarithmic scale, and arrows a common length scale. Annotations are unweighted relative L^2 errors on the 129×129 plotting grid; the table reports flux-energy errors on 9216 Gauss points.

3.5 Anisotropic heat conduction with internal sources

The vertical edges are held at zero temperature and the horizontal edges are insulated. The conductivity tensor and heat source are

$$K_T = R(\vartheta_T) \text{diag}(1 + 2x, 0.15 + 0.35y) R(\vartheta_T)^\top, \quad \vartheta_T = \frac{\pi}{4} \sin(\pi y),$$

$$f_T = 10 \exp\left[-\frac{(x - 0.3)^2 + (y - 0.35)^2}{0.1^2}\right] + 6 \exp\left[-\frac{(x - 0.7)^2 + (y - 0.65)^2}{0.14^2}\right]. \quad (41)$$

Here $\alpha_T = 0.15$. The prescribed sources supply heat at two separated locations, and the cold edges remove it. In contrast to the Darcy case, the potential is generated by an interior source and both vertical boundary values are equal. The flux balance now requires the net outward heat flow to equal the integrated source.

The reference mesh and evaluation rules match the Darcy example. From 256 to 512 elements per direction, the potential L^2 and flux-energy differences are 7.846×10^{-9} and 1.288×10^{-5} . The prescribed source and boundary configuration is shown in Figure 5; Supplementary Section S4 gives the reference checks.

Table 5: Heat conduction: relative errors (percent) on 9216 Gauss points for the three displayed runs at target seed 92101 and the final common 48000-evaluation stage.

Method	E_u	$E_{q,K^{-1}}$	E_{mix}
Matched Euclidean	0.0389	0.4454	0.3645
Analytical	0.0221	0.3442	0.2750
Operator probe	0.0182	0.2552	0.2036

Figures 8 and 9 show the reference temperature and heat flux, three neural predictions and their pointwise

errors. The two sources create overlapping warm regions, with heat leaving through the cooled boundaries. Table 5 gives the corresponding single-run errors.

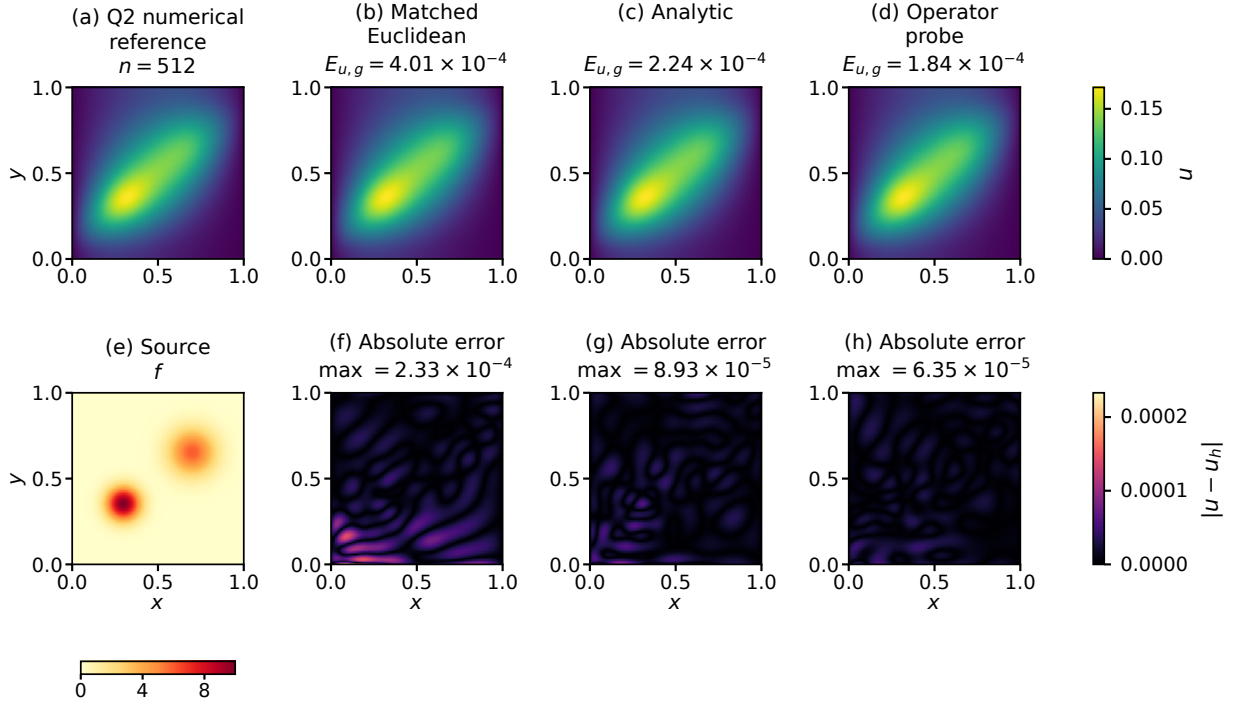


Figure 8: Heat conduction: reference potential, matched Euclidean, analytical and operator-probe predictions, and absolute errors. Field and error panels respectively share linear color scales. Relative L^2 annotations use the equally weighted 129×129 plotting grid, including its boundary; Table 5 uses 9216 Gauss points.

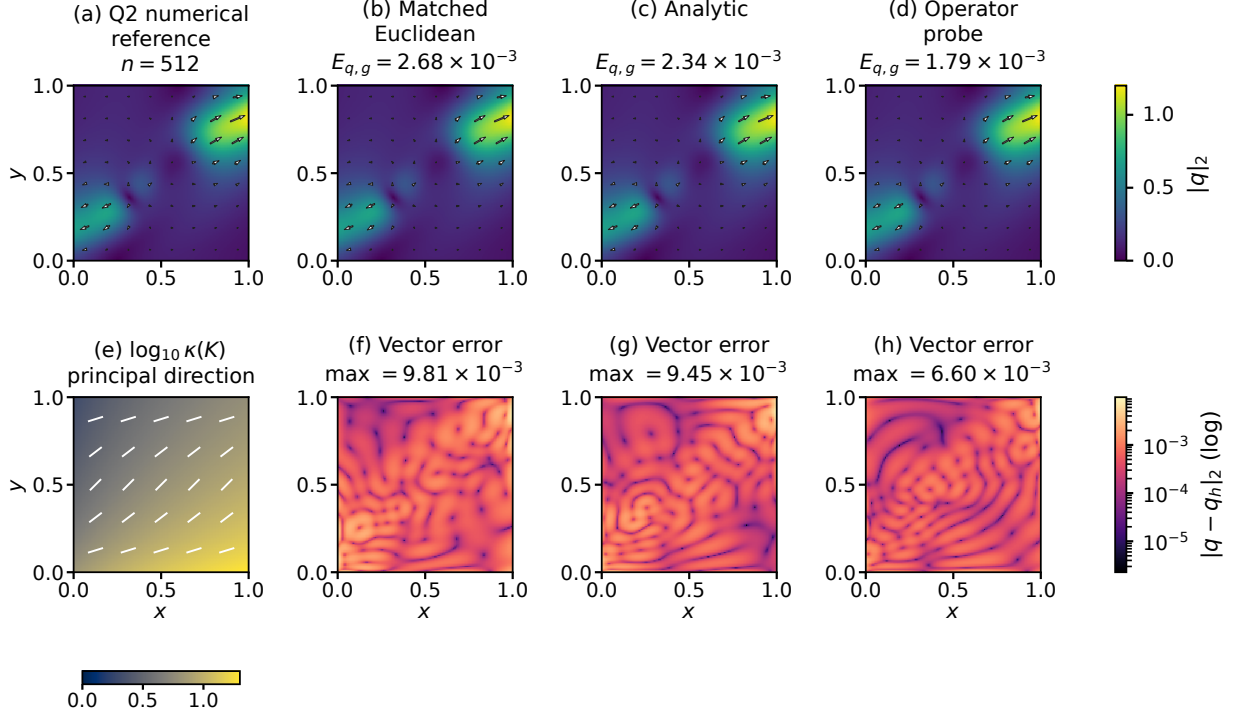


Figure 9: Heat conduction: reference and independent neural fluxes, with vector-error magnitude against $q_h = -K \nabla u_h$. Flux magnitudes share a linear scale, errors a logarithmic scale, and arrows a common length scale. Annotations are unweighted relative L^2 errors on the 129×129 plotting grid; the table reports flux-energy errors on 9216 Gauss points.

3.6 Discussion and limitations

The analysis distinguishes two roles of the residual norm: controlling the energy error and selecting an approximation from the available class. Theorem 2.1 addresses the first role. Proposition 2.3 makes the second explicit: missing cross terms can move the minimizer away from the best energy approximation even when conservation and optimization are exact. The trace-matched experiments examine this choice in neural approximations. Their gains are most consistent for constant rotation and smaller when the principal axes vary in space. The probe bounds explain how operator responses retain the analytical energy scale; they do not predict superiority to direct analytical weighting.

The constant-flux example isolates directional coupling in a restricted affine family; its exact error ratio is not a prediction for the full neural network class. For general neural fields, the contrast-independent two-sided estimate applies to the ideal continuous functional, while training uses an L^2 conservation term and finite-point quadrature. The frequency example in Section 2.5 rules out a general reverse bound after the L^2 replacement, and spatial quadrature requires its own error control[22]. The longer runs in Supplementary Section S3 include component errors that increase as the training loss decreases. In the Darcy example, mixed-energy errors are still decreasing between 24000 and 48000 L-BFGS evaluations, so the final models are not claimed to have reached optimization convergence.

The numerical results concern known, smooth, uniformly elliptic diffusion tensors. The independent statistical comparison uses two manufactured problems, six seed groups and one network family. It supports improvement over the two Euclidean baselines and records a slightly lower geometric-mean error than analytical weighting. The smaller observed errors relative to analytical and coordinate-diagonal weighting remain subject to the uncertainty quantified by their confidence intervals. The earlier comparisons use separate exploratory data. The Darcy and heat-conduction examples show one seed per method and case, so they describe spatial errors without estimating variation across seeds. Their mesh-refinement differences assess reference resolution rather than bound the remaining reference error. The finite-element solutions serve as references, and no equal-cost comparison with a finite-element solver is made. Discontinuous

coefficients and more complex geometries remain to be studied.

4 Conclusion

For mixed PINNs in anisotropic diffusion, directional coupling affects both the relation between residual and energy error and the approximation selected by minimizing the loss. The weighted orthogonal decomposition gives sharp equivalence constants for the ideal residual functional. The constant-flux example shows why an upper error bound alone is insufficient: every fixed positive diagonal weight selects the same approximation, and its error relative to the best member of the trial family grows with the off-diagonal coupling. This remains true for positive quadrature and exact optimization.

The analytical weight K^{-1} preserves the full constitutive energy. The operator-probe construction recovers a related tensor from the geometric mean of two response moments, with spectral bounds that connect it to the analytical weight. In the constant-flux setting, these bounds also control the selected approximation relative to the best available. Both implementations preserve the objective under consistent changes of residual coordinates.

The controlled experiments find the clearest benefit from retaining coupling for constant rotated coefficients. In the independent two-problem study, probe weighting reduces geometric-mean mixed-energy errors by approximately 37% and 56% relative to the conservation-matched and original Euclidean baselines. Its observed geometric-mean error is also slightly lower than with analytical weighting. Darcy flow and heat conduction illustrate the corresponding potential, flux and error fields. Together, the results support designing mixed residual norms around the energy error and the directional information needed to represent it.

Declarations

Data and code. Numerical data were generated from manufactured solutions and from prescribed Darcy-flow and heat-conduction boundary-value problems. No external or restricted datasets were used.

Data and code availability. The original manufactured-benchmark data and simulation code, together with the experimental settings and software environment, are available in the repository https://github.com/pixel-tt/MixedPINNs_Anisotropic_Diffusion and at the fixed release https://github.com/pixel-tt/MixedPINNs_Anisotropic_Diffusion/releases/tag/v1.0.0. Data, code and field-plotting scripts for the additional application experiments accompany this revision in a separate reproducibility package. The repository code is released under a MIT licence and its data under a CC BY 4.0 licence.

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

Author contributions. Mingde Ye: theoretical analysis, experimental design, data analysis, and manuscript writing (Methodology, Formal analysis, Writing – original draft). Xiaohua Ding: topic selection, theorem verification, and literature search (Conceptualization, Validation, Investigation).

Competing interests. The authors declare no competing interests.

References

- [1] T. Ando, Chi-Kwong Li, and Roy Mathias. Geometric means. *Linear Algebra and its Applications*, 385:305–334, 2004. doi: 10.1016/j.laa.2003.11.019.

- [2] Francisco M. Bersetche and Juan Pablo Borthagaray. A deep first-order system least squares method for solving elliptic PDEs. *Computers & Mathematics with Applications*, 129:136–150, 2023. doi: 10.1016/j.camwa.2022.11.014.
- [3] Rafael Bischof and Michael Kraus. Multi-objective loss balancing for physics-informed deep learning. *Computer Methods in Applied Mechanics and Engineering*, 439:117914, 2025. doi: 10.1016/j.cma.2025.117914.
- [4] Pavel B. Bochev and Max D. Gunzburger. *Least-Squares Finite Element Methods*, volume 166 of *Applied Mathematical Sciences*. Springer, New York, 2009. doi: 10.1007/b13382.
- [5] James H. Bramble, Raytcho D. Lazarov, and Joseph E. Pasciak. A least-squares approach based on a discrete minus one inner product for first order systems. *Mathematics of Computation*, 66(219): 935–955, 1997. doi: 10.1090/S0025-5718-97-00848-X.
- [6] Franco Brezzi and Michel Fortin. *Mixed and Hybrid Finite Element Methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer, New York, 1991. doi: 10.1007/978-1-4612-3172-1.
- [7] Zhiqiang Cai, Raytcho D. Lazarov, Thomas A. Manteuffel, and Stephen F. McCormick. First-order system least squares for second-order partial differential equations: Part I. *SIAM Journal on Numerical Analysis*, 31(6):1785–1799, 1994. doi: 10.1137/0731091.
- [8] Zhiqiang Cai, Jingshuang Chen, Min Liu, and Xinyu Liu. Deep least-squares methods: An unsupervised learning-based numerical method for solving elliptic PDEs. *Journal of Computational Physics*, 420: 109707, 2020. doi: 10.1016/j.jcp.2020.109707.
- [9] Tim De Ryck, Florent Bonnet, Siddhartha Mishra, and Emmanuel de Bézenac. An operator preconditioning perspective on training in physics-informed machine learning. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/f122a6778999d1d2b25cdc6927930f1a-Abstract-Conference.html.
- [10] Alejandro Duque-Salazar, Paulina Sepúlveda, Carlos Uriarte, Jamie M. Taylor, and David Pardo. Robust deep FOSLS for transmission problems. arXiv preprint arXiv:2604.17549, 2026. URL <https://arxiv.org/abs/2604.17549v1>. Version 1, April 2026.
- [11] Weinan E and Bing Yu. The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1):1–12, 2018. doi: 10.1007/s40304-018-0127-z.
- [12] Mohammad Sadegh Eshaghi, Cosmin Anitescu, Manish Thombre, Yizheng Wang, Xiaoying Zhuang, and Timon Rabczuk. Variational physics-informed neural operator (VINO) for solving partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 437:117785, 2025. doi: 10.1016/j.cma.2025.117785. URL <https://arxiv.org/abs/2411.06587>.
- [13] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2 edition, 2010. doi: 10.1090/gsm/019.
- [14] Vladimir Fanaskov, Alexander Rudikov, and Ivan Oseledets. Neural functional a posteriori error estimates. arXiv preprint arXiv:2402.05585, 2024. URL <https://arxiv.org/abs/2402.05585>.
- [15] Vladimir Fanaskov, Tianchi Yu, Alexander Rudikov, and Ivan Oseledets. Astral: Training physics-informed neural networks with error majorants. arXiv preprint arXiv:2406.02645, 2024. URL <https://arxiv.org/abs/2406.02645v2>. Version 2, March 2026; accepted to the ICLR 2026 AI&PDE workshop.
- [16] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 2021. doi: 10.1038/s42254-021-00314-5.

- [17] Minghua Lin and Gord Sinnamon. The generalized Wielandt inequality in inner product spaces. *Eurasian Mathematical Journal*, 3(1):72–85, 2012. URL <https://arxiv.org/abs/1201.6294>.
- [18] Qiang Liu, Mengyu Chu, and Nils Thuerey. ConFIG: Towards conflict-free training of physics informed neural networks. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/94e85561a342de88b559b72c9b29f638-Abstract-Conference.html.
- [19] Songming Liu, Chang Su, Jiachen Yao, Zhongkai Hao, Hang Su, Youjia Wu, and Jun Zhu. Pre-conditioning for physics-informed neural networks. *arXiv preprint arXiv:2402.00531*, 2024. doi: 10.48550/arXiv.2402.00531.
- [20] Suryanarayana Maddu, Dominik Sturm, Christian L. Müller, and Ivo F. Sbalzarini. Inverse-dirichlet weighting enables reliable training of physics informed neural networks. *Machine Learning: Science and Technology*, 3:015026, 2022. doi: 10.1088/2632-2153/ac3712.
- [21] Levi D. McClenny and Ulisses M. Braga-Neto. Self-adaptive physics-informed neural networks. *Journal of Computational Physics*, 474:111722, 2023. doi: 10.1016/j.jcp.2022.111722.
- [22] Siddhartha Mishra and Roberto Molinaro. Estimates on the generalization error of physics informed neural networks (PINNs) for approximating PDEs. *IMA Journal of Numerical Analysis*, 43(1):1–43, 2023. doi: 10.1093/imanum/drab093.
- [23] W. Pusz and S. L. Woronowicz. Functional calculus for sesquilinear forms and the purification map. *Reports on Mathematical Physics*, 8(2):159–170, 1975. doi: 10.1016/0034-4877(75)90061-0.
- [24] Zhaonan Qu, Wenzhi Gao, Oliver Hinder, Yinyu Ye, and Zhengyuan Zhou. Optimal diagonal preconditioning. *arXiv preprint arXiv:2209.00809v2*, 2022. URL <https://arxiv.org/abs/2209.00809v2>.
- [25] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.
- [26] Sergey Repin, Stefan Sauter, and Anton Smolianski. Two-sided a posteriori error estimates for mixed formulations of elliptic problems. *SIAM Journal on Numerical Analysis*, 45(3):928–945, 2007. doi: 10.1137/050641533. URL <https://www.math.uzh.ch/d.php/publication/357>.
- [27] Sergio Rojas, Paweł Maczuga, Judit Muñoz-Matute, David Pardo, and Maciej Paszyński. Robust variational physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, 425:116904, 2024. doi: 10.1016/j.cma.2024.116904.
- [28] Jamie M. Taylor, David Pardo, and Ignacio Muga. A Deep Fourier Residual method for solving PDEs using Neural Networks. *Computer Methods in Applied Mechanics and Engineering*, 405:115850, 2023. doi: 10.1016/j.cma.2022.115850.
- [29] Jiachen Yao, Chang Su, Zhongkai Hao, Songming Liu, Hang Su, and Jun Zhu. MultiAdam: Parameter-wise scale-invariant optimizer for multiscale training of physics-informed neural networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, 2023. URL <https://proceedings.mlr.press/v202/yao23c.html>.
- [30] Shun Zhang. A simple proof of coerciveness of first-order system least-squares methods for general second-order elliptic PDEs. *Computers & Mathematics with Applications*, 130:98–104, 2023. doi: 10.1016/j.camwa.2022.11.021.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ERIPINNEnglishSupplementpdflatex.zip](#)