

Supplementary material

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

Mingde Ye, Xiaohua Ding

September 2026

This supplement provides additional mathematical details, experimental settings, numerical comparisons and reference-solution checks. Exploratory and independent comparison results are reported separately. Complete configurations and run records accompany the code.

S1 Additional mathematical details

The main text gives the detailed proofs of the sharp mixed-energy equivalence and the componentwise-weighting results. We retain its notation and ellipticity assumptions. In particular,

$$\mathcal{J}_{\text{nat}} = \|r_c\|_{-1,K}^2 + \|\eta\|_{K^{-1}}^2, \quad \mathcal{J}_{\text{stab}} = \lambda_{\text{th}} \|r_c\|_{L^2}^2 + \|\eta\|_{K^{-1}}^2,$$

where $\lambda_{\text{th}} = C_P^2/\alpha$ and $\varphi = (1 + \sqrt{5})/2$. The following details supplement the weight comparisons and the probe construction.

S1.1 Pointwise covariance

For an invertible residual transformation $A(x)$ and a positive-definite scale $M(x)$,

$$(Ar)^\top (AMA^\top)^{-1} (Ar) = r^\top M^{-1} r, \quad (\text{S1})$$

since $(AMA^\top)^{-1} = A^{-\top} M^{-1} A^{-1}$. The integral objectives therefore agree wherever finite. If A and M are independent of the network parameters, all existing parameter derivatives agree as well.

S1.2 Trace-matched controls

The main text proves the one-sided energy bounds and the constant-flux approximation result for componentwise weights. Here a direct matrix comparison makes the contrast dependence of the isotropic weight explicit. Let $H = K^{-1}$ have eigenvalues $0 < h_1 \leq h_2$, write $\kappa = h_2/h_1$, and set $Q_{\text{iso}} = \text{tr}(H)I/2$. In the principal axes of H ,

$$\frac{\kappa + 1}{2\kappa} H \preceq Q_{\text{iso}} \preceq \frac{\kappa + 1}{2} H. \quad (\text{S2})$$

The lower factor is at least $1/2$, whereas the upper factor grows with κ . Rotating the axes by 45° makes both diagonal entries of H equal, so $Q_{\text{diag}} = Q_{\text{iso}}$ and the same dependence applies to that example.

S1.3 Two-bank operator-probe bounds

Draw independent standard Gaussian directions $a_i, b_i \in \mathbb{R}^m$ and reuse each draw across all locations. The probes $(u, q) = (0, a_i)$ and $(b_i^\top(x - x_0), 0)$, for fixed $x_0 \in \mathbb{R}^m$, have responses a_i and Kb_i . They need

not satisfy the target boundary condition. Their known zero population means give

$$P_N = N^{-1} \sum_i a_i a_i^\top, \quad S_N = N^{-1} \sum_i b_i b_i^\top, \quad Q_N = K S_N K,$$

without sample centering. For Gaussian draws, $N \geq m$ gives $P_N, S_N \succ 0$ almost surely; the following deterministic statement assumes the realized matrices are positive definite. Using the classical matrix geometric mean[1, 2], set

$$X \# Y = X^{1/2} (X^{-1/2} Y X^{-1/2})^{1/2} X^{1/2}, \quad M_N^\# = P_N \# Q_N. \quad (\text{S3})$$

Let $p_\pm, s_\pm$ be the extremal eigenvalues of P_N, S_N , and $m_N = \sqrt{p_- s_-}$, $h_N = \sqrt{p_+ s_+}$. Since $s_- K^2 \preceq Q_N \preceq s_+ K^2$, joint monotonicity, scalar homogeneity and $I \# K^2 = K$ imply

$$m_N K \preceq M_N^\# \preceq h_N K. \quad (\text{S4})$$

The same banks are used everywhere, so the envelope is simultaneous in space. If $P_N = S_N = I$, it gives $M_N^\# = K$. For congruence, $Z = X \# Y$ is the unique SPD solution of $Z X^{-1} Z = Y$: conjugation by $X^{-1/2}$ reduces this to the unique positive square root. Thus $A Z A^\top$ solves the transported equation and

$$(A X A^\top) \# (A Y A^\top) = A (X \# Y) A^\top.$$

Both response moments must be transported; resetting the first moment to I after changing coordinates defines a different metric.

Inverting (S4) and adding the common conservation term gives

$$\begin{aligned} \min\{1, h_N^{-1}\} \mathcal{J}_{\text{stab}} &\leq J_{N,\text{probe}} \leq \max\{1, m_N^{-1}\} \mathcal{J}_{\text{stab}}, \\ \mathcal{E}_{\text{mix}}^2 &\leq \varphi^2 \max\{1, h_N\} J_{N,\text{probe}}, \end{aligned} \quad (\text{S5})$$

where $J_{N,\text{probe}} = \lambda_{\text{th}} \|r_c\|_{L^2}^2 + \int \eta^\top (M_N^\#)^{-1} \eta$. The constants involve only the probe banks; the construction still uses the known action of K and the lower bound α for conservation scaling. The law of large numbers gives $P_N, S_N \rightarrow I$ and $M_N^\# \rightarrow K$ as the bank size grows. The executed method freezes one realization with $N = 32$ per bank. Neither the envelope nor the sampling limit implies superiority to analytic weighting at a finite training budget.

S2 Detailed experimental settings

S2.1 Problems, architecture and losses

The three manufactured problems are defined in the main text. Here C, D and O denote the constant rotated, variable diagonal and oscillatory rotated cases, respectively. Their ellipticity lower bounds are 0.3, 0.1 and 0.05, and $\lambda_{\text{th}} = 1/(2\pi^2\alpha)$.

The network has widths 2–32–32–32–32–3, tanh activations and 3363 parameters. Inputs are $(2x - 1, 2y - 1)$. The raw outputs are multiplied by $\text{diag}(16, 20, 20)$; the potential channel is then multiplied by $x(1 - x)y(1 - y)$. Weights are Xavier-normal and biases zero. Automatic differentiation evaluates potential gradients and flux divergence. All updates are full-batch in float64; the probe development and confirmation networks are initialized in float32 before casting to float64. Within a problem–target cell, compared methods share initial tensors and fixed training points. There is no resampling, manual parameter replacement or model selection by validation error.

Write $\widehat{\mathbb{E}}$ for the arithmetic point average. Table S1 defines $\lambda \widehat{\mathbb{E}}[r_c^2] + \widehat{\mathbb{E}}[\eta^\top W \eta]$. All constitutive quadratics sum components, except for the displayed 1/2 in the original Euclidean loss. The probe uses two independent Gaussian banks of 32 directions, 64 instances in total, fixed across space. The reference statistics are computed before training and remain fixed; neither ridge regularization nor a pseudoinverse is used. For N, 32 independently initialized untrained networks with the target architecture and output scales supply constitutive residuals; their sample-centered covariance uses divisor 32. Target PDE residuals are never centered. This construction differs from the separate response banks used by P.

Table S1: Weights in the probe studies. N appears only in the 24-run development study; original Euclidean appears in confirmation, not in that development matrix.

Method	λ	W
Probe (P)	λ_{th}	$(M_{32}^{\#})^{-1}$
Matched Euclidean (M)	λ_{th}	I
Analytic (A)	λ_{th}	K^{-1}
Trace (T)	λ_{th}	$\text{tr}(K^{-1})I/2$
Diagonal (D)	λ_{th}	$\text{diag}((K^{-1})_{11}, (K^{-1})_{22})$
Centered neural reference (N)	λ_{th}	inverse centered constitutive covariance
Original Euclidean	1	$I/2$

S2.2 Optimizer settings and training budgets

For the manufactured studies, Adam uses learning rate 10^{-3} and coefficients (0.9, 0.999). L-BFGS uses learning rate 1, strong-Wolfe line search, history size 100, at most 20 internal iterations and 25 evaluations per outer call, and gradient/change tolerances $10^{-12}/10^{-15}$. Each evaluation computes the objective and its gradient; an outer iteration may require several such evaluations. Budgets are checked after complete accepted outer steps; reported overshoots remain part of the executed work.

Longer development study. The 12 branches combine C/D/O, original Euclidean/analytic weights and F/I residual representations, with target seed 2111 and 1024 fixed uniform interior points drawn by a CPU generator with seed 100000 + 2111. Each branch takes 4000 Adam updates, then stops at a joint diagnostic, 60000 cumulative L-BFGS evaluations, or one cumulative active hour. The 33×33 closed validation grid participates in stopping and reported errors; it is not an independent test set. Absorb the output scales into the last layer to obtain canonical parameters θ_{can} . Joint stopping requires

$$\frac{\|\nabla_{\theta_{\text{can}}} \hat{J}\|_2}{1 + \|\theta_{\text{can}}\|_2} \leq 10^{-5}, \quad \frac{\max_j V_j - \min_j V_j}{\min_j V_j} \leq 10^{-4},$$

and conservation and constitutive validation-component max/min ratios at most 1.01, over five finite observations in the same optimizer phase.

F and I denote the two constitutive-energy implementations defined in the main text. Their analytical objectives are identical, while their original Euclidean objectives differ.

Factorial and trace development. Seeds 2111, 2222, 2333 and 2444 use only F, 4000 Adam updates and 6000 L-BFGS objective and gradient evaluations. The factorial design toggles conservation weight 1 versus λ_{th} and constitutive weight $I/2$ versus K^{-1} . Its 12 full-tensor fits are reused by the trace study, which adds 20 control fits. No longer-study endpoint is substituted into these pairs. The coordinate-diagonal control equals the analytic weight in D, so that duplicate is not run.

Probe development. C/O and targets 62111/62222 are paired one-to-one with reference seeds 73111/73222. Six methods A/M/T/D/P/N give 24 fits, each with 512 fixed training points, 2000 Adam updates and 6000–6021 L-BFGS evaluations. Endpoints are evaluated on the 33×33 grid used in the preceding experiments. This design does not separate target and reference variance. Short software test runs are excluded from performance comparisons.

S2.3 Paired comparison and statistical unit

The independent comparison with a prespecified design uses C/O and six fresh targets, with two reference realizations nested within each target (Table S2). Each realization comprises two independent banks and is reused across the two PDEs. Different targets use disjoint random streams. Each target–PDE cell has five control fits and two probe fits: 60 controls plus 24 probe fits. A control is reused in two paired ratios; the independent inferential unit is the target cluster, giving $n = 6$. Training uses 512 float64 uniform interior points from Torch seed 100000 + target, 2000 Adam updates and the final accepted L-BFGS state

Table S2: Prespecified target and nested reference seeds.

Target	Reference 1	Reference 2
64101	75101	76101
64202	75202	76202
64303	75303	76303
64404	75404	76404
64505	75505	76505
64606	75606	76606

after at least 6000 evaluations. The time limit is 1800 seconds, checked after complete accepted steps, with an additional ceiling of 6200 L-BFGS evaluations. Runs are not retried or resumed. The joint criterion requires all planned fits to finish with finite results before reaching the execution limits. Validation supplies diagnostics only. Complete execution limits and run records accompany the code.

The evaluation set consists of 16384 float64 i.i.d. uniform interior points drawn with NumPy PCG64 seed 83120260912 before the experimental design was finalized. These points were used for error evaluation only after all 84 fits were complete. Squared-norm sums are accumulated over all evaluation points before forming relative errors. The primary endpoint is

$$E_{\text{mix}} = \left(\frac{\|u - u_*\|_{K,G}^2 + \|q - q_*\|_{K^{-1},G}^2}{\|u_*\|_{K,G}^2 + \|q_*\|_{K^{-1},G}^2} \right)^{1/2}, \quad (\text{S6})$$

where G replaces integrals by point averages. This relative discrete error differs from the absolute continuous $\mathcal{E}_{\text{mix}}$. Component errors use the corresponding relative L^2 or energy norms. Residual RMS values are $R_c = (\hat{\mathbb{E}}_G[r_c^2])^{1/2}$ and $R_\eta = (\hat{\mathbb{E}}_G[|\eta|^2]/2)^{1/2}$.

For target j , PDE c , reference realization k and comparator b , define

$$\begin{aligned} d_{j,b} &= \frac{1}{4} \sum_{c=1}^2 \sum_{k=1}^2 \log(E_{jck}^P / E_{jc}^b), & G_b &= \exp(\bar{d}_b), \\ U_b &= \exp(\bar{d}_b + t_{0.95,5} s_b / \sqrt{6}), & G_{c,b} &= \exp\left(\frac{1}{12} \sum_{j,k} \log(E_{jck}^P / E_{jc}^b)\right), \end{aligned} \quad (\text{S7})$$

where $\bar{d}_b$ and s_b are the mean and sample standard deviation over six targets and $t_{0.95,5}$ is the 95th Student- t quantile. The prespecified intersection–union rule requires all 84 valid endpoints and, for both M and original Euclidean, $G_b < 1$, $U_b < 1$, and both $G_{c,b} < 1$. The procedure assumes independent, approximately Gaussian cluster log effects and conditions on the two fixed PDEs and common quadrature. Six clusters give limited information about that distribution. Analytic, trace and diagonal outcomes are descriptive; no equivalence margin or favorable-comparator selection is introduced after evaluation. The two predefined 8192-point halves are sensitivity checks, not alternative primary quadratures.

Software versions, configurations and complete run records accompany the reproducibility package.

S2.4 Application sampling and evaluation

The two application illustrations use target seed 92101 and probe seed 93101. Each case contains matched Euclidean, analytical and operator-probe runs, for six fits in total. A Torch generator with seed $192101 = 100000 + 92101$ draws 8192 fixed float64 uniform interior points. Initial model tensors, points, coefficients, sources and probe banks are generated once and shared across methods within each case. Xavier-normal initialization is performed in float32 before conversion to float64, with zero biases and unit output scales. Probe seed 93101 generates two independent Gaussian banks of 32 two-dimensional vectors.

Training uses 2000 full-batch Adam updates at learning rate 10^{-3} with coefficients (0.9, 0.999), followed by L-BFGS with learning rate 1, strong-Wolfe line search, history size 100, one internal iteration and

a nominal maximum of 25 evaluations per outer call. Gradient and change tolerances are zero. Recorded stages target 2000, 6000, 12000, 24000 and 48000 cumulative L-BFGS objective/gradient evaluations. All six runs complete the final common 48000-evaluation stage, with 48000–48002 actual objective and gradient evaluations, including overshoots from complete accepted steps. The figures and tables use the models at the end of this common training stage, without selection by reference error.

Field errors are evaluated against the $n = 512$ reference on a fixed 3×3 composite Gauss rule over 32×32 cells (9216 points). These are descriptive deterministic quadrature points used for diagnostics, not an independent test set. At the fixed endpoints, refinement to 64×64 cells (36864 points) changes the field errors and evaluation objectives by at most 0.03181%. With all norms evaluated by the 9216-point weighted quadrature, the application tables report

$$\begin{aligned} E_u &= \frac{\|u_\theta - u_h\|_{L^2}}{\|u_h\|_{L^2}}, & E_{q,K^{-1}} &= \frac{\|q_\theta - q_h\|_{K^{-1}}}{\|q_h\|_{K^{-1}}}, \\ E_{\text{mix}}^2 &= \frac{\|u_\theta - u_h\|_K^2 + \|q_\theta - q_h\|_{K^{-1}}^2}{2\|q_h\|_{K^{-1}}^2}, & q_h &= -K\nabla u_h. \end{aligned}$$

The flux-energy normalization is distinct from the unweighted flux L^2 error denoted by E_q in the manufactured studies. The figures compare the reference with all three selected methods using common scales. No statistical superiority or optimizer-convergence conclusion is drawn from these six illustrations.

S3 Additional numerical results

S3.1 Longer training and representation sensitivity

The longer study retains all 12 endpoints. Eight C/D branches satisfy the joint stopping rule; the four O branches reach the one-hour limit. Table S3 compares analytical and original Euclidean weighting at two common evaluation stages and at the final endpoints. The relative advantage is smaller at the final endpoints, which have unequal realized work and are not fixed-budget ablation pairs.

Table S3: Development comparison of analytic and original Euclidean losses. Entries are geometric means over the F and I paths of analytic-to-Euclidean mixed-energy error ratios. The common milestones are approximate L-BFGS evaluation counts after 4000 Adam updates. Final endpoints follow the stopping protocol.

Problem	Near 6000 evaluations	Near 12000 evaluations	Final endpoint
Constant rotated	0.2347	0.2668	0.3755
Variable diagonal	0.3323	0.3349	0.4295
Oscillatory rotated	0.4976	0.4659	0.8265

The improvements do not cover every component. In C/F, analytic weighting raises potential L^2 error from 6.654×10^{-5} to 7.296×10^{-5} (about 9.6%), and potential-energy error by 14.5%. In O/I, flux-energy error is about 29.9% higher. From the milestone near 30000 evaluations to the endpoint, analytic mixed errors increase by 13.3% in F and 10.7% in I, while both Euclidean branches improve. The four O endpoints do not meet the joint convergence diagnostic. A longer trajectory or smaller training loss therefore does not imply monotone improvement of validation physical error.

Table S4 reports the endpoint sensitivity $S_E = |\log_{10}(E_F/E_I)|$ between the two analytical implementations. Only three of nine values are at most 0.01, equivalent to a larger/smaller error ratio of at most 1.0233. These differences concern independent floating-point training trajectories; the same-state loss and gradient checks are reported in the main text.

Table S4: Analytic endpoint sensitivities; E_u, E_q are relative L^2 errors. F/I share a seed but are separate floating-point optimization trajectories.

Problem	S_{E_u}	S_{E_q}	$S_{E_{\text{mix}}}$
C	0.03802	0.00141	0.00060
D	0.04618	0.00895	0.01499
O	0.02916	0.01234	0.01419

S3.2 Scalar-tensor and trace-matched ablations

The factorial definitions and four-seed summaries are given in the main text. With $\ell_{ab} = \log_{10} E_{ab}$, the median interaction $\ell_{11} - \ell_{10} - \ell_{01} + \ell_{00}$ is -0.0814 , -0.0615 and $+0.0255$ for C, D and O, respectively. The joint effect therefore differs from the sum of two independent gains. For O, the median full/tensor-only error ratio is 1.019, consistent with its small adverse conservation-weight effect.

With conservation fixed at λ_{th} , the trace study fixes the mean eigenvalue of the constitutive weight. All twelve analytic/isotropic pairs favor the analytic tensor, but one O pair favors the coordinate-diagonal control (Table S5). Trace matching does not equate actual residual losses or parameter gradients, and this experiment does not separately identify principal directions, eigenvalue allocation and their interaction with optimization. It uses existing validation data and four seeds; conservation weights are not retuned for each control.

Table S5: Analytic/control mixed-error ratios at equal trace, summarized over four seeds. No duplicate diagonal control is run for D.

Problem	Control	Median	Range	Analytic lower
C	Isotropic	0.634	0.614–0.721	4/4
D	Isotropic	0.808	0.716–0.969	4/4
O	Isotropic	0.956	0.856–0.958	4/4
C	Diagonal	0.688	0.471–0.857	4/4
O	Diagonal	0.935	0.743–1.106	3/4

S3.3 Operator-probe endpoint comparisons

Table S6 retains the mixed errors of all 24 development fits. P/M ratios are 0.659646, 1.089033, 0.894832 and 0.475176 in row order, so the all-four-pairs criterion fails at C/62222. The four-pair geometric mean 0.743 is descriptive and does not replace that criterion. P beats N in all four cells, T in two and D in three; P/A ranges from 0.870069 to 1.178929. Residual components can move oppositely to solution

Table S6: All 24 development mixed-energy errors, in units of 10^{-3} . Method codes are defined in Table S1. Each cell uses its final accepted budget endpoint.

Case	Target	A	M	T	D	P	N
C	62111	0.4962	0.7655	0.6283	0.6120	0.5049	1.560
C	62222	0.7463	0.7464	0.9273	0.6155	0.8128	1.412
O	62111	14.18	18.68	14.72	17.28	16.71	82.28
O	62222	14.90	27.29	12.60	13.54	12.97	70.27

error. For O/62111, P lowers mixed error relative to M, while R_c rises from 0.2333 to 0.3925 and R_η from 0.04799 to 0.07304. These exploratory results motivated the subsequent mean-effect comparison; they are not pooled with the six-cluster confirmation.

The separate tensor-geometry diagnostic uses $N = 8, 16, 32, 64, 128, 100$ independent bank pairs per N , contrasts 1, 10, 100, and 17 orientations $\theta_\ell = \ell\pi/16$. The 500 bank pairs are reused across contrasts,

giving 1500 dependent records. No violation of the realized envelope is recorded; the maximum congruence defect is 9.55×10^{-14} . At contrast 100 and $N = 32$, the bank median of the orientation maximum of $\|K^{-1/2}M_N^\#K^{-1/2} - I\|_2$ is about 0.228. These quantify tensor sampling and algebraic implementation, not learned-solution error or optimizer ordering.

S3.4 Held-out quadrature and comparator diagnostics

All 84 planned fits reached valid endpoints, with no exclusions or failures. The maximum L-BFGS evaluation count was 6022. Error measures were checked against the saved fields.

Table S7 retains every comparator and both prescribed half-quadratures. Both prespecified Euclidean comparisons pass the full rule, including case-specific ratios (0.639, 0.617) for M and (0.391, 0.485) for original Euclidean. The trace comparison is descriptive despite its upper limit below one. The probe-to-analytic geometric-mean ratio is 0.931, indicating a slightly lower observed mixed-energy error with probe weighting (approximately 6.9% lower). The diagonal ratio is 0.917. Diagonal and analytic upper limits exceed one; their two-sided 95% intervals are [0.757, 1.110] and [0.777, 1.115], so neither statistical superiority nor equivalence is established against them.

Table S7: Confirmation and quadrature sensitivity. G, U are defined by (S7). Only the first two rows are prespecified primary comparisons; the full 16384-point quadrature determines the decision.

Comparator	Full G	Full U	First half G	Second half G
Matched Euclidean	0.628	0.806	0.632	0.624
Original Euclidean	0.435	0.501	0.440	0.431
Trace isotropic	0.760	0.893	0.760	0.759
Diagonal	0.917	1.065	0.916	0.917
Analytic	0.931	1.072	0.929	0.933

The statistical comparison concerns the two fixed manufactured problems, the stated network and six independent seed groups. It evaluates accuracy at the common training budget.

S3.5 Constant rotated benchmark

Figures S1 and S2 use target seed 64101 and probe bank 75101, as in the main-text oscillatory comparison. The plotting grid contains 129×129 nodes, including the boundary. These fields show the recorded final models.

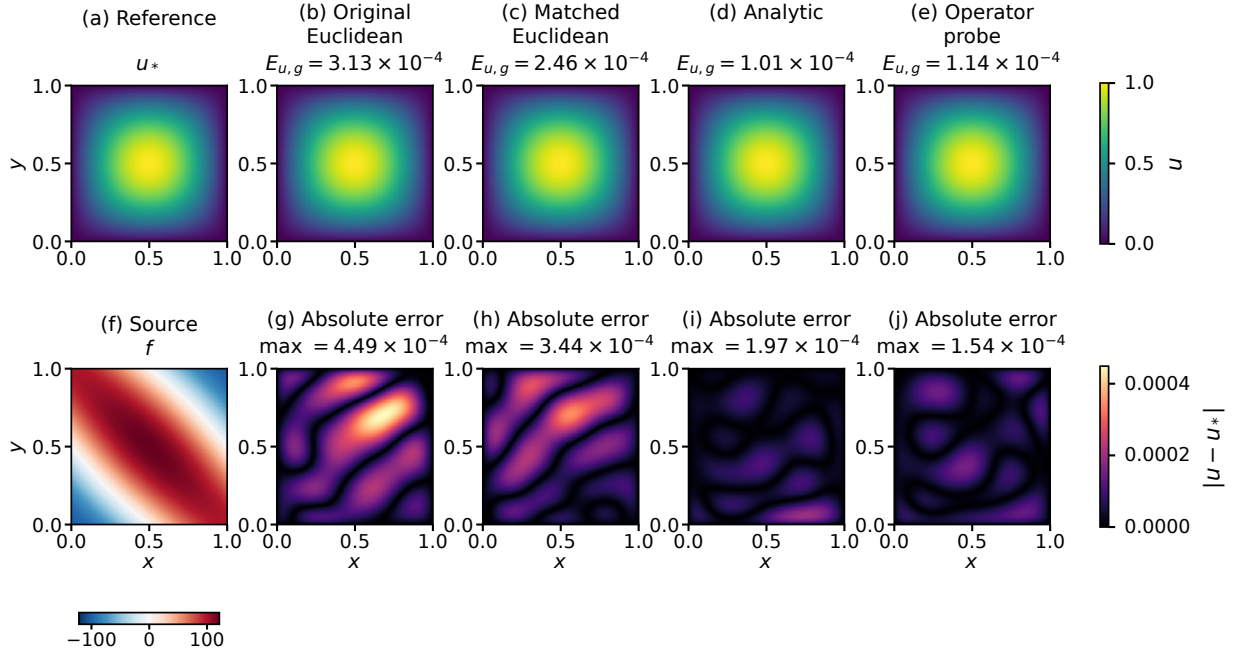


Figure S1: Constant rotated problem: exact potential, four final-model predictions and absolute potential errors. The potential and error panels respectively share linear color scales. Panel errors are evaluated on the plotting grid. The source field is also shown.

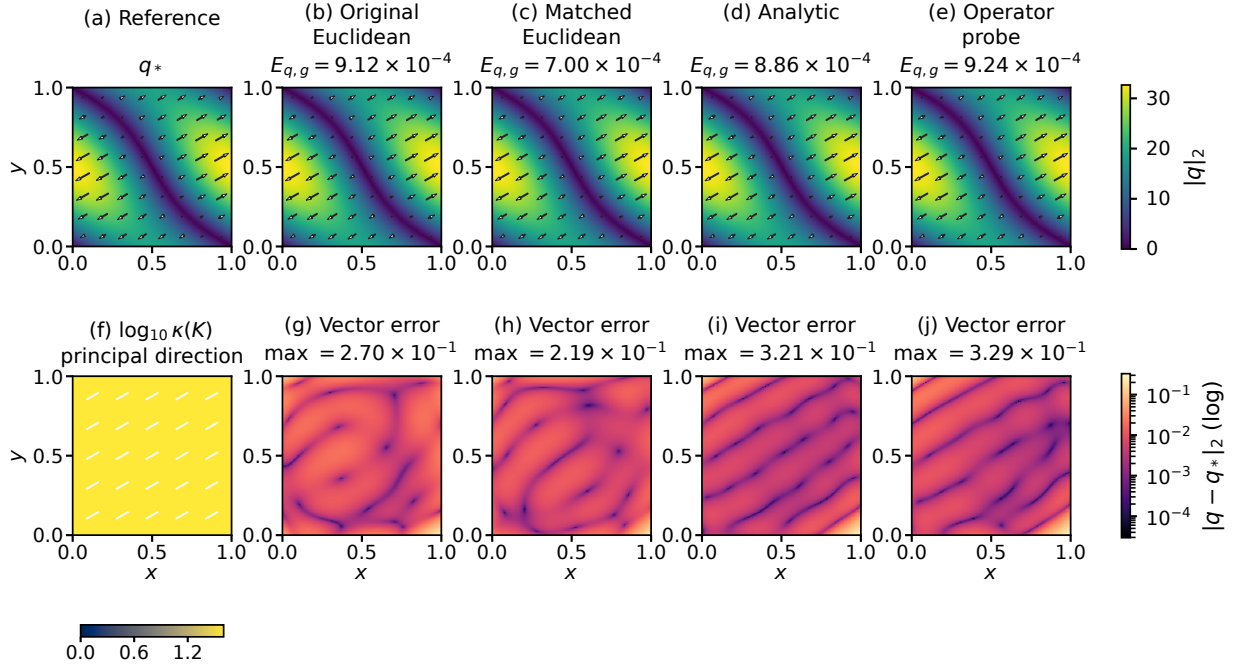


Figure S2: Constant rotated problem: physical flux magnitude and direction, and vector-error magnitude. Flux colors and arrows use common scales; vector errors use a shared logarithmic scale. The coefficient panel gives $\log_{10} \kappa(K)$ and its principal axis. The probe plotting-grid vector L^2 error is 9.245×10^{-4} , compared with 9.117×10^{-4} for original Euclidean, whereas their relative mixed-energy errors are 7.828×10^{-4} and 1.692×10^{-3} .

S4 Verification of numerical reference solutions

The Darcy and thermal references use uniform Q_2 meshes with $n = 16, 32, 64, 128, 256, 512$ elements per direction and 4×4 Gauss integration. The finest system has 1050625 potential degrees of freedom. A constant diagonal tensor reproduces the linear head solution with maximum potential error 2.89×10^{-15} and maximum flux-component error 3.77×10^{-14} . A constant-source quadratic solution checks the assembled load and gives corresponding errors 1.39×10^{-15} and 9.33×10^{-15} .

Table S8 reports the last two refinement differences, evaluated on the recorded analysis quadrature. Here $\Delta_u = \|u_h - u_{2h}\|_{L^2}$ and $\Delta_q = \|q_h - q_{2h}\|_{K^{-1}}$, with $q_h = -K\nabla u_h$. These are observed mesh differences, not rigorous bounds on the remaining reference error.

Table S8: Successive differences for the independent numerical references.

Problem	Mesh transition	Δ_u	Δ_q
Darcy	128 $\rightarrow$ 256	2.166×10^{-7}	2.265×10^{-4}
Darcy	256 $\rightarrow$ 512	2.701×10^{-8}	5.664×10^{-5}
Thermal	128 $\rightarrow$ 256	6.277×10^{-8}	5.152×10^{-5}
Thermal	256 $\rightarrow$ 512	7.846×10^{-9}	1.288×10^{-5}

The $n = 512$ linear solves have relative algebraic residuals 6.70×10^{-15} for Darcy and 5.21×10^{-11} for heat conduction. The raw constitutive reference flux can have inter-element normal jumps and does not satisfy the insulating data pointwise. It is therefore not treated as a global $H(\text{div})$ reconstruction or used to assert a guaranteed neural error bound. Refinement assesses the resolution of the displayed potential and flux reference fields.

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] 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.