ENRICHED HOLOMORPHIC NEURAL NETWORKS FOR 2D AND 3D LAPLACE PROBLEMS

Anonymous authors
Paper under double-blind review

ABSTRACT

Boundary-value problems for the Laplace equation develop singular behaviour near corners, edges and junctions that smooth approximators resolve poorly. We combine holomorphic neural networks with singular functions derived from the local geometry and boundary conditions. The representation is harmonic by construction, so no interior residual is needed: the network fits the smooth remainder from boundary data alone, and the enrichment terms capture the singular behaviour. Singular exponents are trainable, so the construction also applies when the local expansion is not known in advance. We extend the approach to three dimensions using independent networks evaluated along complex null directions. The enriched model is the most accurate tested method on every problem. Its margin is several orders of magnitude where the enrichment captures the singularity exactly, and a factor of about 1.7–2.1 over the next-best method on a CdZnTe radiation detector with many singular sites, in both two and three dimensions.

1 INTRODUCTION

Solving Partial Differential Equations (PDEs) is difficult on domains with non-trivial geometries or boundary conditions. Physics-Informed neural networks (PINNs) (Raissi et al., 2019) offer an alternative by incorporating the governing equations directly into the training objective. The solution is represented by a neural network, so it is continuous and its derivatives come from automatic differentiation. The method is therefore mesh-free and differentiable end-to-end, which motivates its use for irregular geometries and inverse problems (Luo et al., 2025).

Unfortunately, the resulting objective suffers from competing loss terms, ill-conditioning and imbalanced gradients, which slow convergence or degrade accuracy (Wang et al., 2021; 2022; Rahaman et al., 2019). Much work has therefore improved the optimization and representation of PINNs. Adaptive loss weighting balances the terms of the objective (Kendall et al., 2018; Heydari et al., 2019; Wang et al., 2022; McClenny & Braga-Neto, 2023), and adaptive collocation changes where the residual is evaluated (Wu et al., 2023). Other approaches change the representation, through periodic activations or Fourier features (Sitzmann et al., 2020; Liu et al., 2025), or by replacing the multilayer perceptron with a Kolmogorov–Arnold network (Liu et al., 2024). Domain-decomposition approaches such as conservative PINNs (cPINNs) (Jagtap et al., 2020) and extended PINNs (XPINNs) (Jagtap & Karniadakis, 2020) partition the domain into subdomains coupled through interface conditions.

We address the representation side of these difficulties. We focus on the Laplace equation, $\Delta u = 0$, and build on the Physics-Informed Holomorphic Neural Networks (PIHNNs) formulations introduced by Calafà et al. (2024) for two-dimensional problems and further developed by Ballini et al. (2026) in three dimensions. These representations have also been shown to benefit from enrichment with known singularities for specific problems, such as crack tips in plane elasticity (Calafà et al., 2025). Our approach addresses the enrichment for a set of local boundary configurations rather than for one particular problem: we construct a dictionary mapping each local boundary configuration to its enrichment family, together with a rule for orienting the branch cuts so that the representation remains single-valued inside the domain. We develop the construction in two and three dimensions.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

2 Related work

Representations that satisfy the governing equation in the interior reduce a boundary value problem to determining model parameters from boundary information. Classical examples include the boundary element method, the method of fundamental solutions (Fairweather & Karageorghis, 1998), harmonic kernel collocation (Schaback, 2008), and rational approximations such as the lightning Laplace solver (Gopal & Trefethen, 2019). These methods construct the solution from fixed analytical families, such as harmonic polynomials or fundamental solutions centred outside the domain (Axler et al., 2001). Neural representations instead learn part of the approximating family, at the cost of a nonconvex fit. Ghosh et al. (2023) represent harmonic functions as the real part of a complex-valued neural network, while Calafà et al. (2024) develop PIHNNs for two-dimensional problems. In 3D, where the correspondence with complex analysis is lost, Ballini et al. (2026) write the solution as an integral over projection directions, approximate it by a weighted sum over a finite set of directions, and share one network across them.

Exact satisfaction of the governing equation, however, does not guarantee an efficient representation of solutions with reduced regularity. At corners and boundary-condition junctions, local singularities arise whose exponents depend on the geometry and boundary conditions (Grisvard, 1985). Enriching the approximation space with known singular functions is a classical remedy in numerical PDE methods (Fix et al., 1973; Li & Lu, 2000). Among neural solvers, the singularity-enriched PINN (SEPINN) of Hu et al. (2024) augments a real-valued PINN with singular functions for polygonal corners, mixed boundary conditions and edges in 3D, while Calafà et al. (2025) enrich PIHNNs with the square-root term from Williams' expansion at crack tips. In these approaches, the singularity and its exponent are prescribed for the problem at hand.

3 Method

3.1 Problem formulation

We consider the Laplace boundary-value problem on a bounded domain $\Omega \subset \mathbb{R}^d$, with $d \in \{2, 3\}$. The boundary is partitioned into disjoint Dirichlet and Neumann portions, $\partial\Omega = \Gamma_D \cup \Gamma_N$, and the solution $u : \Omega \to \mathbb{R}$ satisfies

$$\begin{aligned} \Delta u(\boldsymbol{x}) &= 0, && \boldsymbol{x} \in \Omega, \\ u(\boldsymbol{x}) &= g_D(\boldsymbol{x}), && \boldsymbol{x} \in \Gamma_D, \\ \nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x}) &= g_N(\boldsymbol{x}), && \boldsymbol{x} \in \Gamma_N, \end{aligned} \tag{1}$$

where g_D and g_N denote the prescribed Dirichlet and Neumann data, respectively, and $\boldsymbol{n}$ is the outward unit normal. We assume that Γ_D has positive measure, so that the solution is unique.

We represent the boundary as a collection of segments in two dimensions or faces in three dimensions. We refer to points at which adjacent boundary pieces meet, or at which the boundary condition or its prescribed value changes, as *junctions*. These local boundary configurations determine the singular terms used to enrich the neural representation.

3.2 Holomorphic functions

A function $f : U \to \mathbb{C}$ on an open set $U \subset \mathbb{C}$ is holomorphic if it is complex differentiable at every point of U. Writing $z = x + iy$ and $f(z) = p(x, y) + iq(x, y)$, and using the subscript notation for partial derivatives throughout so that $p_x = \frac{\partial p}{\partial x}$, for continuously differentiable p and q, complex differentiability is equivalent to the Cauchy–Riemann equations $p_x = q_y$ and $p_y = -q_x$ (Axler et al., 2001; Evans, 2010). Differentiating these once more gives $\Delta p = p_{xx} + p_{yy} = q_{yx} - q_{xy} = 0$, and the same for q, so both parts of a holomorphic function are harmonic. Conversely, on a simply connected domain every harmonic function is the real part of a holomorphic one. A function that is holomorphic on all of $\mathbb{C}$ is called *entire*.

Two properties of holomorphic functions are central to this work. First, they are closed under composition and linear combination: a network assembled from holomorphic layers is therefore holomorphic, and adding further holomorphic terms to its output keeps it so. Second, the real part of any holomorphic function is harmonic so a holomorphic network represents a harmonic function

for every value of its weights. Together these mean the network is harmonic by construction, with no interior residual required. The Cauchy–Riemann equations also give the complex derivative as $f'(z) = p_x - ip_y$, from which the normal derivative follows directly.

3.3 Enrichment functions

Near corners and junctions, solutions to the Laplace equation develop singular behaviour that smooth approximators resolve poorly: the solution may be bounded, but its gradient can blow up. Rather than relying on the network to learn this from boundary samples, we augment the representation with enrichment functions drawn from local singular expansions for the given geometry and boundary conditions (Grisvard, 1985; Hu et al., 2024; Calafà et al., 2025).

Consider a planar wedge with interior angle ω and polar coordinates (r, θ) centred at its vertex. For a homogeneous Dirichlet condition at $\theta = 0$ and a homogeneous Neumann condition at $\theta = \omega$, the leading singular mode is $r^\lambda \sin(\lambda\theta)$ with $\lambda = \pi/(2\omega)$. The solution stays bounded as $r \to 0$, but its gradient grows as $r^{\lambda-1}$ and diverges whenever $\lambda < 1$, i.e. when $\omega > \pi/2$. This is illustrated in Figure 1 for $\omega = 2\pi/3$, $\lambda = 3/4$, where the field looks smooth but the gradient concentrates at the vertex.

In the local complex coordinate $w = re^{i\theta}$, the singular modes are the real and imaginary parts of $w^\lambda = \exp(\lambda \operatorname{Log} w)$. We place the branch cut of Log along a fixed exterior direction of the junction (Appendix A). The complex coefficient of each term then absorbs the phase of the wall from which θ is measured and the order of the two conditions. At each junction there is not one singular exponent but a countably infinite sequence $\lambda_1 < \lambda_2 < \cdots$, each indexing a valid local mode. For a D→N (Dirichlet to Neumann) junction at angle ω, for instance, $\lambda_k = (2k-1)\pi/(2\omega)$. The leading term ($k = 1$) is the most singular and dominates near the vertex, and higher terms are progressively smoother. The number of terms per junction is a hyperparameter we call the *ladder depth*. We use ladder depth one except on Motz's problem and the cracked beam, where the first four exponents are included (Appendix F). Table 1 lists the enrichment terms.

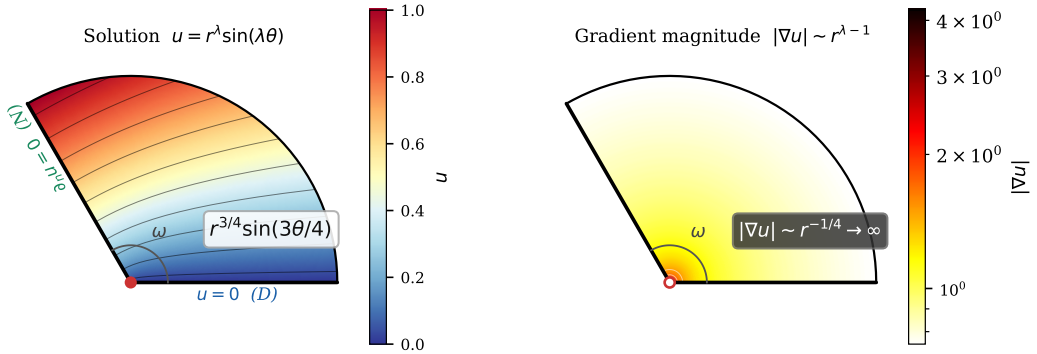

Figure 1: Singular behaviour at a D→N junction with $\omega = 2\pi/3$, $\lambda = 3/4$. Left: the leading mode $r^\lambda \sin(\lambda\theta)$ is bounded at the vertex. Right: the gradient magnitude $|\nabla u| \sim r^{\lambda-1} = r^{-1/4}$ diverges there. The enrichment term w^λ represents this mode directly.

Trainable exponents. The exponents in Table 1 and Figure 5 can be fixed at their theoretical values or trained. Since w^λ is holomorphic for every $\lambda > 0$, varying the exponent does not affect harmonicity, but only the exponents of the table satisfy the local boundary conditions, so a trained λ is a fitted parameter rather than the singular exponent itself. Each term carries one real parameter τ, with $\lambda = \varepsilon + \operatorname{softplus}(\tau)$ initialised at the predicted value, except the conical degree, which starts at $1/2$. This allows the construction to be applied when the exponent is not known in advance, whereas prior enrichment schemes fix λ (Hu et al., 2024; Calafà et al., 2025).

Table 1: Enrichment dictionary. Each junction matches one row and contributes the term E_λ with a trainable coefficient c_q. Local coordinates are centred at the junction z_c, with $w = z - z_c = re^{i\theta}$; ω is the interior angle and $k = 1, 2, \ldots$ indexes the ladder. Every term is holomorphic, so the representation remains exactly harmonic. Figure 5 is a visualization of this table.

Junction	Local solution	$\lvert\nabla u\rvert$	Term E_λ
Dirichlet to Neumann	$r^\lambda \sin(\lambda\theta)$, $\lambda = (2k-1)\pi/(2\omega)$	$r^{\lambda-1}$	w^λ
Both Dirichlet, value jump Δg	$(\Delta g/\omega)\,\theta$	r^{-1}	$\arg w$
Both Neumann, flux jump, $\omega \in \pi\mathbb{N}$	$\Re[w \operatorname{Log} w]$	$\lvert\ln r\rvert$	$w \operatorname{Log} w$
Both Neumann, equal flux, $\omega > \pi$	$r^\lambda \cos(\lambda\theta)$, $\lambda = k\pi/\omega$	$r^{\lambda-1}$	w^λ
Both Dirichlet, equal values, $\omega > \pi$	$r^\lambda \sin(\lambda\theta)$, $\lambda = k\pi/\omega$	$r^{\lambda-1}$	w^λ
Otherwise	C^∞ away from vertex	bounded	none

Extension to 3D. For a straight edge with unit tangent t and transverse orthonormal frame e_1, e_2, we set

$$w = (\boldsymbol{x} - \boldsymbol{x}_e) \cdot (e_1 + ie_2), \qquad E_\lambda(\boldsymbol{x}) = w^\lambda. \tag{2}$$

The direction $e_1 + ie_2$ satisfies $a \cdot a = 0$, so E_λ is harmonic wherever its branch is holomorphic.

3.4 Enriched Holomorphic neural networks

PIHNNs exploit the compositional property of holomorphic functions to construct a holomorphic neural representation. Let $f_\eta : \mathbb{C} \to \mathbb{C}$ denote such a network, with complex-valued parameters η. Since holomorphic functions are closed under composition, f_η is holomorphic by construction, and its real part is therefore harmonic (Ghosh et al., 2023; Calafà et al., 2024).

The PIHNN represents the remaining part of the solution, and Q enrichment terms E_{λ_q}, one per singular site and ladder index, represent the behaviour near the sites. The output of the model is

$$u_{\eta,\boldsymbol{c}}(x, y) = \Re\left[f_\eta(z) + \sum_{q=1}^{Q} c_q E_{\lambda_q}(z)\right], \tag{3}$$

where c_q determine the amplitude and phase of each term E_{λ_q}.

Since the representation is harmonic by construction, no interior residual is needed, and the parameters are fitted to the boundary conditions by minimizing

$$\mathcal{L}(\eta, \boldsymbol{c}) = \frac{1}{|X_D|} \sum_{\boldsymbol{x} \in X_D} \left\|u_{\eta,\boldsymbol{c}}(\boldsymbol{x}) - g_D(\boldsymbol{x})\right\|^2 + \frac{1}{|X_N|} \sum_{\boldsymbol{x} \in X_N} \left\|\partial_n u_{\eta,\boldsymbol{c}}(\boldsymbol{x}) - g_N(\boldsymbol{x})\right\|^2, \tag{4}$$

where X_D and X_N denote the corresponding boundary sample sets, and a term is omitted when the corresponding boundary condition is absent.

Extension to 3D In 3D there is no single complex coordinate, so we use complex projections of the form $a \cdot \boldsymbol{x}$, where $a \in \mathbb{C}^3$. For a holomorphic function f and a fixed direction a, the chain rule gives

$$\Delta_{\boldsymbol{x}} f(a \cdot \boldsymbol{x}) = (a \cdot a) f''(a \cdot \boldsymbol{x}), \qquad a \cdot a = \sum_{k=1}^{3} a_k^2. \tag{5}$$

Thus, choosing a null direction satisfying $a \cdot a = 0$ makes the projected function harmonic.

We use K such projections and sum their outputs:

$$u_{\boldsymbol{\eta},\boldsymbol{c}}(\boldsymbol{x}) = \Re\left[\sum_{k=1}^{K} f_{\eta_k}(\zeta_k) + \sum_{q=1}^{Q} c_q E_{\lambda_q}(\boldsymbol{x})\right], \tag{6}$$

where $\zeta_k = a_k \cdot \hat{\boldsymbol{x}} + b_k$, with a complex offset b_k and $\hat{\boldsymbol{x}} = (\boldsymbol{x} - \boldsymbol{x}_0)/L_0$ (Section 4). Each branch is harmonic, and so is their sum. The normal derivative follows directly from the chain rule. For an

outward unit normal $\boldsymbol{n} \in \mathbb{R}^3$,

$$\partial_n u_{\boldsymbol{\eta},\boldsymbol{c}}(\boldsymbol{x}) = \Re\left[\frac{1}{L_0}\sum_{k=1}^{K}(\boldsymbol{n}\cdot a_k) f'_{\eta_k}(\zeta_k) + \sum_{q=1}^{Q} c_q \boldsymbol{n}\cdot\nabla_{\boldsymbol{x}} E_{\lambda_q}(\boldsymbol{x})\right]. \tag{7}$$

The model is fitted with the same boundary loss as in 2D.

Since each network branch is entire and the enrichment terms are harmonic away from their branch cuts, the full representation is harmonic in Ω (Appendix A). For pure Dirichlet problems, $\Gamma_D = \partial\Omega$, this bounds the interior error by the boundary misfit. The error $e = u_{\boldsymbol{\eta},\boldsymbol{c}} - u^\star$ is harmonic, and the junctions J, points in 2D and edges in 3D, form a polar set. If e is bounded and continuous up to $\partial\Omega \setminus J$, the extended maximum principle (Ransford, 1995, Thm. 3.6.9; Armitage & Gardiner, 2001 in 3D) gives

$$\sup_{\Omega} |u_{\boldsymbol{\eta},\boldsymbol{c}} - u^\star| \;\le\; \sup_{\partial\Omega\setminus J} |u_{\boldsymbol{\eta},\boldsymbol{c}} - g_D|\,. \tag{8}$$

The error is bounded when every enrichment term is, which rules out a $\ln|w|$ component in the value-jump terms (Appendix A). We evaluate the right-hand side on a finite boundary sample, which can only underestimate the supremum, so the reported value is an error estimate rather than a certified bound.

4 Model Setup

In both dimensions the model has the same three parts: a holomorphic core, the enrichment terms that Table 1 assigns to the junctions, and a real part that turns their sum into a potential. The core is a small complex-valued network in which each layer applies a complex affine map followed by the complex exponential,

$$h_0 = \zeta, \qquad h_{\ell+1} = \exp(W_\ell h_\ell + b_\ell), \qquad f_\eta(\zeta) = W_{\text{out}} h_m + b_{\text{out}}, \tag{9}$$

where ζ is a complex scalar, m is the number of hidden layers, the exponential acts on each component, and η collects the complex weights and biases. The exponential is entire, so f_η is entire for any value of η.

2D model In two dimensions the input is the normalized coordinate $\hat{z} = (z - z_0)/L_0$, with z_0 the centre of the domain and L_0 its characteristic radius. The core has $m = 2$ hidden layers of width 16 with complex exponential activations. The enrichment terms are evaluated in the physical coordinate, since their sites, cut directions and interior angles come from the geometry of the problem, and one term is placed per junction and ladder index. All 2D problems have unit length scale, so w needs no separate reference length.

3D model The 3D model first projects the point onto several complex coordinates. Figure 2 shows how a point $\boldsymbol{x}$ becomes a potential $u(\boldsymbol{x})$.

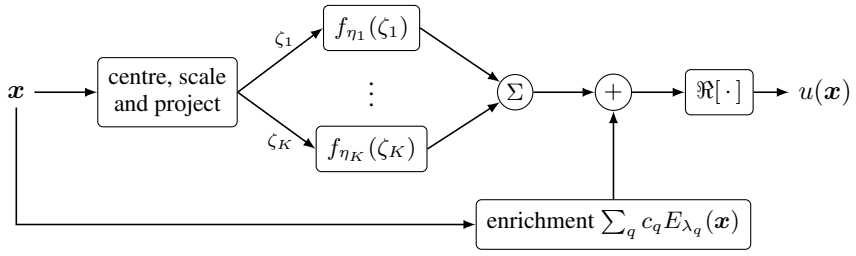

Figure 2: Forward pass of the 3D model. The point is projected onto K complex coordinates $\zeta_k = a_k \cdot \hat{\boldsymbol{x}} + b_k$ along null directions, each coordinate is passed through its own holomorphic network, and the outputs are summed. The enrichment terms are evaluated at the physical point and added before the real part is taken.

The point is centred and scaled, $\hat{\boldsymbol{x}} = (\boldsymbol{x} - \boldsymbol{x}_0)/L_0$, and projected onto K complex coordinates $\zeta_k = a_k \cdot \hat{\boldsymbol{x}} + b_k$, with a complex offset b_k. By Equation 5, a holomorphic function of ζ_k is harmonic

in $\boldsymbol{x}$ when $a_k \cdot a_k = 0$, so we write each direction as

$$a(\xi, \rho) = \rho \left(\frac{1 - \xi^2}{2},\ \frac{i(1 + \xi^2)}{2},\ \xi \right), \qquad \xi, \rho \in \mathbb{C}, \tag{10}$$

which is null for every ξ and ρ. The directions can therefore be trained freely. The same length scale is used for all coordinates, since a separate scale per coordinate would break the null condition. Each ζ_k is passed through its own network of the form Equation 9, with two hidden layers of width 8, and the $K = 16$ outputs are added. This sum is a discrete counterpart of the integral over directions used by Ballini et al. (2026), with separate parameters for each direction instead of one network shared across them. The number and size of the branches are the same for all problems (Appendix B).

The enrichment terms are evaluated at the physical point, since their location is part of the geometry. A straight edge uses the term w^λ of Equation 2, with the exponent that Table 1 gives for its dihedral angle and the conditions on its two faces. A circular conical vertex uses the separated mode

$$S_\nu(R, \vartheta) = R^\nu P_\nu(\cos \vartheta), \tag{11}$$

where R is the distance from the vertex, ϑ the angle from the cone axis, and ν the first positive root of $P_\nu(\cos \vartheta_{\text{wall}}) = 0$. A trained ν enters both factors, and P_ν is evaluated by a 400-term hypergeometric series whose tail is below 10^{-27} on the domain. Site positions, edge frames and branch cuts are fixed by the geometry and are not trained, since moving them could place a singular line inside the domain. Appendix H writes out the model for the L-shaped prism.

4.1 Parameter fitting

The model is fitted to boundary samples only. Both loss terms of Equation 4 are made dimensionless by a characteristic potential U_0 and length L_0: the Dirichlet term is divided by U_0^2 and the Neumann term multiplied by L_0^2/U_0^2; both scales are one on the sampled suite. Enrichment coefficients start at zero and exponents at the values of Table 1 (the conical degree at $1/2$); exponents are fixed on Motz's problem, the cracked beam and the sampled suite. All parameters are optimized jointly with Adam followed by L-BFGS.

Every sample carries gradient on λ through $\partial_\lambda w^\lambda = w^\lambda \operatorname{Log} w$, but far from a site the change is smooth and the rest of the model can absorb it, so we expect uniform boundary samples to carry little information about the exponent. Since the coefficients start at zero, an exponent also receives no gradient until its coefficient is nonzero. We therefore add, except on the sampled suite, samples whose distance to each junction or edge is drawn log-uniformly between 10^{-4} and 10^{-1} of the local length. Step counts and sample sizes are given in Appendices F and J.

4.2 Evaluation

Where a reference solution $u^\star$ is available, we report the relative L^2 error

$$\epsilon_u(X) = \left(\frac{\sum_{\boldsymbol{x} \in X} |u(\boldsymbol{x}) - u^\star(\boldsymbol{x})|^2}{\sum_{\boldsymbol{x} \in X} |u^\star(\boldsymbol{x})|^2} \right)^{1/2}, \tag{12}$$

estimated on a set X of interior points. For the CZT detector, where the reference is numerical, we report the root-mean-square error instead. Where no closed-form solution exists, $u^\star$ is a singular-series or numerical reference, described with each problem. Unless stated otherwise, results are medians over three random seeds.

5 Experiments in two dimensions

5.1 Problems

We use six named benchmarks, each with a single singular site (Appendix C). Four have a closed-form solution: two circular sectors, an L-shaped domain and a slit disk, whose solution is the leading wedge mode of the opening angle (Grisvard, 1985). The other two, Motz's problem (Motz, 1947) and the cracked beam (Lu et al., 2004), are scored against their truncated singular expansion, whose leading coefficient D_0 is known to 17 significant digits (Lu et al., 2004).

As an application we use a CdZnTe (CZT) drift strip radiation detector (Kuvvetli et al., 2010) geometry. CZT detectors are semiconductor gamma-ray detectors that operate at room temperature, and the electric potential inside the crystal determines how the charge released by a photon drifts to the electrodes. The electrodes on the crystal faces are held at fixed potentials and separated by insulating gaps, so every electrode edge is a junction between a Dirichlet and a Neumann boundary. A single detector therefore contains many singular sites of the same type. In 2D we use a cross-section of the detector, scored by the root-mean-square error against a numerical reference (Appendix D). To test many sites on simpler geometries, we also use a suite of 24 randomly sampled problems with 4 to 12 sites each (Appendix E).

5.2 Results

Baselines We compare three representations. The MultiLayer Perceptron (MLP) PINN is a real-valued network trained on the Laplace residual at interior points and on the boundary misfit, with the weighting tuned per problem. The plain holomorphic model is Equation 3 with $Q = 0$, and the enriched model adds the terms of Table 1. The last two are the same network trained with the same budget and restarts, so they differ only in the enrichment.

Table 2: Errors in 2D. All entries report ϵ_u, median over three seeds at the best budget of Appendix F, chosen by the same rule for every method. *Plain* and *Enriched* are the same network without and with the terms of Table 1.

Problem	MLP PINN	Plain	Enriched
Closed-form solution			
Sector $3\pi/2$	9.5×10^{-2}	7.7×10^{-4}	$\mathbf{3.3 \times 10^{-12}}$
Sector $7\pi/4$	1.8×10^{-1}	1.9×10^{-2}	$\mathbf{1.4 \times 10^{-16}}$
L-shape	7.4×10^{-2}	7.5×10^{-4}	$\mathbf{2.9 \times 10^{-16}}$
Slit disk	1.8×10^{-1}	3.2×10^{-1}	$\mathbf{2.8 \times 10^{-12}}$
Series reference			
Motz	1.7×10^{-1}	5.4×10^{-4}	$\mathbf{2.5 \times 10^{-8}}$
Cracked beam	8.6×10^{-2}	6.4×10^{-4}	$\mathbf{8.6 \times 10^{-8}}$
Application			
CZT detector	7.8×10^{-1}	1.7×10^{-2}	$\mathbf{1.0 \times 10^{-2}}$

The plain holomorphic model improves on the MLP PINN by a factor of 9 to 315 on five of the six benchmarks (Table 2). The two configurations also differ in precision, architecture and boundary sampling (Appendix F), so this compares solver configurations rather than isolating exact harmonicity. The exception is the slit disk, where the plain model is worse than the MLP PINN: its solution vanishes on both slit faces but its transverse Cartesian derivative has opposite one-sided limits. An entire core extends smoothly across the slit and cannot reproduce this discontinuity, but the branch enrichment represents it directly.

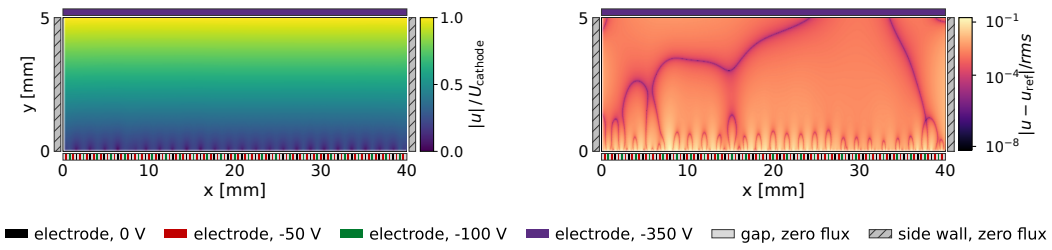

Figure 3: Enriched model on the 2D CZT detector. Left: normalized potential. Right: error against the reference

The enrichment gives a further 4 to 14 orders of magnitude, and the enriched model reaches roundoff on the $7\pi/4$ sector and the L-shape. On these closed-form benchmarks the solution is itself a dictionary term, so they test recovery of the supplied mode. On Motz's problem and the cracked beam it

Table 3: Relative error ϵ_u in 3D, median over three converged seeds at each method's best schedule. The detector is scored against the finite-element reference. ‡Extension of WAN to mixed boundary data.

Problem	PINN	Deep Ritz	WAN	Ballini et al.	Enriched
Closed-form solution					
L-prism	8.7×10^{-4}	2.7×10^{-3}	5.6×10^{-3}	1.3×10^{-3}	$\mathbf{6.9 \times 10^{-9}}$
u_A	8.9×10^{-4}	4.5×10^{-3}	1.2×10^{-2}	1.9×10^{-3}	$\mathbf{1.8 \times 10^{-4}}$
u_B	1.1×10^{-3}	5.9×10^{-3}	1.2×10^{-2}	1.6×10^{-3}	$\mathbf{1.6 \times 10^{-4}}$
Smooth cube	1.5×10^{-5}	5.9×10^{-4}	2.2×10^{-3}	2.0×10^{-4}	$\mathbf{4.6 \times 10^{-6}}$
Series reference					
Conical notch	1.2×10^{-3}	1.0×10^{-2}	7.4×10^{-2}	4.1×10^{-2}	$\mathbf{6.0 \times 10^{-9}}$
Extruded Motz	1.1×10^{-2}	4.1×10^{-3}	‡3.4×10^{-1}	5.8×10^{-3}	$\mathbf{3.4 \times 10^{-6}}$
Application					
CZT detector	7.5×10^{-2}	2.7×10^{-2}	‡7.5×10^{-2}	2.4×10^{-2}	$\mathbf{1.3 \times 10^{-2}}$

recovers D_0 to relative errors of 8.2×10^{-7} and 2.0×10^{-6}, with the network supplying the correction that reduces the error of the same four singular terms fitted alone, 5.5×10^{-3} and 2.2×10^{-2} (Appendix C), to the reported values. With the same network and budget, tripling the ADAM steps and boundary samples improves the plain model by less than an order of magnitude, while the enriched model improves by several (Appendix G). On the sampled suite, the enrichment improves all 72 paired runs and reduces the median ϵ_u by a factor of 3.2×10^6 on the four closed-form instances.

On the detector (Figure 3), the plain model reduces the error of the MLP PINN by a factor of 46, and the enrichment reduces it by a further factor of 1.7. The enrichment gain is much smaller than on the benchmarks: the terms of many junctions must be fitted together with the smooth field between the electrodes. The enriched model is nevertheless 78 times more accurate than the MLP PINN on this detector cross-section.

On the pure Dirichlet benchmarks, the sampled right-hand side of Equation 8 is close to the true error: wherever the error lies above roundoff, the largest interior error of the enriched model reaches 94% to 100% of it, making it a useful estimate when no reference solution is available.

6 Experiments in three dimensions

6.1 Problems

We group the problems as in 2D (Appendix I). Four have a closed-form solution. The L-shaped prism has the cross-section of a re-entrant wedge and the solution $g(x, y) = r^{2/3} \sin(2\theta/3)$, which does not depend on the axial coordinate. Two further fields on the same prism vary along the edge,

$$u_A(x, y, z) = (1 + 0.5z)g(x, y) + 0.1xyz, \qquad u_B(\boldsymbol{x}) = u_A(\boldsymbol{x}) + \frac{0.1}{\|\boldsymbol{x} - (-2, -2, 2)\|}, \quad (13)$$

and a smooth harmonic field on the unit cube has no singular site at all, so it tests the core alone. Two problems are scored against a series reference: the conical notch, whose solution is the vertex mode of Equation 11, and the extruded Motz problem, which adds zero-flux end faces to the planar problem and uses a 60-term singular series. As an application we use the full 3D CZT detector, posed in millimetres and volts, whose 238 electrode edges are all Dirichlet to Neumann junctions. It is scored against a graded finite-element reference whose own error is about 10^{-4}.

6.2 Results

Baselines We compare against four published mesh-free methods that impose the equation in different ways. The PINN (Raissi et al., 2019) minimizes the strong-form residual, the Deep Ritz method (E & Yu, 2018) minimizes the Dirichlet energy, and the weak adversarial network (WAN) of Zang et al. (2020) minimizes the weak form against an adversarially trained test function. The holomorphic network of Ballini et al. (2026) is exactly harmonic but has no singular terms. Each

baseline follows its published protocol, with the authors' code where it runs on our problems, and is also run with its steps and boundary samples multiplied by 1.5, 2 and 3. We report the best implementation and budget of each method (Appendix J).

None of the four baselines reaches an error below 8×10^{-4} on any problem with a singularity (Table 3), and the exactly harmonic network of Ballini et al. (2026) is not consistently more accurate than the PINN. As in 2D, exact harmonicity is not sufficient without the singular terms. With them, the enriched model outperforms the best baseline by 5 or more orders of magnitude on the L-prism and the conical notch, whose solution is the supplied mode, and by three orders of magnitude on the extruded Motz problem.

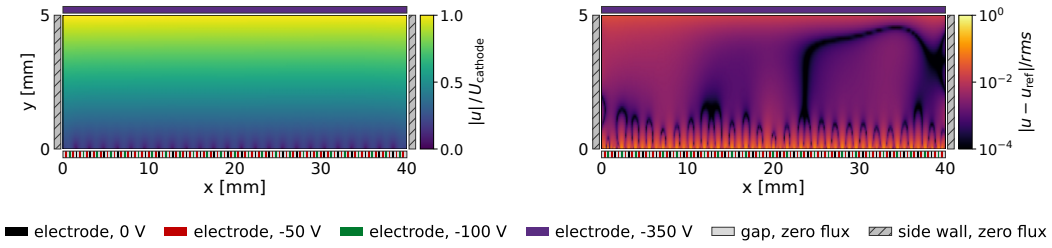

Figure 4: Enriched model on the 3D CZT detector, plane $z = 21$ mm. Left: normalized potential. Right: error against the reference, scaled by the root mean square of the reference solution.

The gain is smaller when the solution varies along the edge. The edge terms carry a constant amplitude, so the network must carry the axial variation of u_A and u_B, and the margin falls to a factor of 5 to 7. On the smooth cube there is nothing to enrich, and the enriched model is three times more accurate than the PINN.

On the detector, the enriched model reaches 1.3×10^{-2}, against 2.7×10^{-2} for Deep Ritz and 7.5×10^{-2} for the PINN. The network of Ballini et al. (2026) reaches 2.4×10^{-2}. The smaller margin may reflect the 238 edge terms fitted together with the field of the whole crystal, the unenriched strip ends, or conditioning. It is still the lowest potential error of the tested methods on this geometry.

7 Conclusion

We have shown that building the analytical structure of the Laplace equation directly into the hypothesis space is more effective than enforcing it through a residual penalty. By leveraging holomorphic network whose real part is harmonic by construction and using a set of singular enrichment terms we remove the need to use the interior points residual during training and handle difficult local behaviours that smooth neural methods cannot handle correctly. Furthermore, we extend the formulation to three dimensions through null complex projections. Together they reduce errors by orders of magnitude over both residual-based and exactly harmonic baselines across a range of 2D and 3D problems, with the largest gains where the dictionary captures the solution's singular behaviour exactly.

The advantage shrinks when many junctions share the domain, when edge amplitudes vary axially, or when the domain geometry falls outside the tested configurations. Several extensions follow naturally from the construction. A dictionary for general polyhedral vertices would extend the 3D results beyond the edge and conical cases. Edge terms with an affine axial amplitude would close the remaining gap on problems where the singular strength varies along the edge. More broadly, since the Laplace equation is linear, a single model with N output heads sharing the same holomorphic core and enrichment structure could represent N solutions on the same geometry simultaneously, as the junction structure and branch cuts are identical for all of them and only the smooth remainder and enrichment coefficients differ. For the CZT detector, where 100 weighting potentials share the same electrode geometry, this would reduce training cost by an order of magnitude while exploiting the shared singular structure across all outputs.

AI USE STATEMENT

We used generative AI tools (large language model assistants) to write and debug experiment code, to plan, submit and monitor cluster jobs, and to collect and check results. We also used these tools as a writing aid, for tasks such as suggesting edits, checking consistency, and flagging potential errors in the manuscript. All text was written and substantially revised by the authors, who reviewed all content, verified the reported numbers against stored run records, and take full responsibility for the final work.

REPRODUCIBILITY STATEMENT

Section 4, Appendices F and J and Table 9 give the architectures, budgets, sampling and selection rules; Appendices C to E and I define every problem and its reference solution; and Table 10 gives the per-seed results behind Table 3. We will release the code on the camera-ready version.

REFERENCES

David H. Armitage and Stephen J. Gardiner. *Classical Potential Theory*. Springer Monographs in Mathematics. Springer, London, 2001. doi: 10.1007/978-1-4471-0233-5.

Sheldon Axler, Paul Bourdon, and Wade Ramey. *Harmonic Function Theory*, volume 137 of *Graduate Texts in Mathematics*. Springer, 2 edition, 2001. ISBN 9780387952185.

Enrico Ballini, Allan Peter Engsig-Karup, and Tito Andriollo. A holomorphic neural network framework for 3D boundary value problems governed by harmonic potentials. *arXiv preprint arXiv:2605.31231*, 2026.

Matteo Calafà, Emil Hovad, Allan P. Engsig-Karup, and Tito Andriollo. Physics-informed holomorphic neural networks (PIHNNs): Solving 2D linear elasticity problems. *Computer Methods in Applied Mechanics and Engineering*, 432:117406, 2024. doi: 10.1016/j.cma.2024.117406.

Matteo Calafà, Henrik Myhre Jensen, and Tito Andriollo. Solving plane crack problems via enriched holomorphic neural networks. *Engineering Fracture Mechanics*, 322:111133, 2025. doi: 10.1016/j.engfracmech.2025.111133.

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.

Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2 edition, 2010.

Graeme Fairweather and Andreas Karageorghis. The method of fundamental solutions for elliptic boundary value problems. *Advances in Computational Mathematics*, 9:69–95, 1998. doi: 10.1023/A:1018981221740.

G. J. Fix, S. Gulati, and G. I. Wakoff. On the use of singular functions with finite element approximations. *Journal of Computational Physics*, 13(2):209–228, 1973. doi: 10.1016/0021-9991(73)90023-5.

Atiyo Ghosh, Antonio Andrea Gentile, Mario Dagrada, Chul Lee, Seong-Hyok Sean Kim, Hyukgeun Cha, Yunjun Choi, Dongho Kim, Jeong-Il Kye, and Vincent Emanuel Elfving. Harmonic neural networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 11340–11359. PMLR, 2023. URL https://proceedings.mlr.press/v202/ghosh23b.html.

Abinand Gopal and Lloyd N. Trefethen. Solving Laplace problems with corner singularities via rational functions. *SIAM Journal on Numerical Analysis*, 57(5):2074–2094, 2019. doi: 10.1137/19M125947X.

Pierre Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, 1985. Wedge exponents in chapters 4–5.

A. Ali Heydari, Craig A. Thompson, and Asif Mehmood. SoftAdapt: Techniques for adaptive loss weighting of neural networks with multi-part loss functions, 2019. URL https://arxiv.org/abs/1912.12355.

Tianhao Hu, Bangti Jin, and Zhi Zhou. Solving Poisson problems in polygonal domains with singularity enriched physics informed neural networks. *SIAM Journal on Scientific Computing*, 46(4): C369–C398, 2024. doi: 10.1137/23M1601195.

Ameya D Jagtap and George Em Karniadakis. Extended physics-informed neural networks (xpinns): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. *Communications in Computational Physics*, 28(5):2002–2041, 2020.

Ameya D Jagtap, Ehsan Kharazmi, and George Em Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365:113028, 2020.

Alex Kendall, Yarin Gal, and Roberto Cipolla. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics, 2018. URL https://arxiv.org/abs/1705.07115.

I. Kuvvetli, C. Budtz-Jørgensen, E. Caroli, and N. Auricchio. Czt drift strip detectors for high energy astrophysics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 624(2):486–491, 2010. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2010.03.172. URL https://www.sciencedirect.com/science/article/pii/S0168900210008247. New Developments in Radiation Detectors.

Z. C. Li and T. T. Lu. Singularities and treatments of elliptic boundary value problems. *Mathematical and Computer Modelling*, 31(8):97–145, 2000. doi: 10.1016/S0895-7177(00)00062-5.

Dong C. Liu and Jorge Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1–3):503–528, 1989.

Yarong Liu, Hong Gu, Xiangjun Yu, and Pan Qin. Diminishing spectral bias in physics-informed neural networks using spatially-adaptive Fourier feature encoding. *Neural Networks*, 182:106886, 2025. doi: 10.1016/j.neunet.2024.106886.

Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljacić, Thomas Y. Hou, and Max Tegmark. KAN: Kolmogorov–Arnold networks, 2024. URL https://arxiv.org/abs/2404.19756.

T. T. Lu, H. Y. Hu, and Z. C. Li. Highly accurate solutions of Motz's and the cracked beam problems. *Engineering Analysis with Boundary Elements*, 28(11):1387–1403, 2004. doi: 10.1016/j.enganabound.2004.03.005.

Kuang Luo, Jingshang Zhao, Yingping Wang, Jiayao Li, Junjie Wen, Jiong Liang, Henry Soekmadji, and Shaolin Liao. Physics-informed neural networks for PDE problems: A comprehensive review. *Artificial Intelligence Review*, 58, 2025. doi: 10.1007/s10462-025-11322-7.

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.

H. Motz. The treatment of singularities of partial differential equations by relaxation methods. *Quarterly of Applied Mathematics*, 4(4):371–377, 1947. URL http://www.jstor.org/stable/43633577.

Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, and Aaron Courville. On the spectral bias of neural networks. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 5301–5310, 2019.

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.

Thomas Ransford. *Potential Theory in the Complex Plane*, volume 28 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1995. doi: 10.1017/CBO9780511623776.

Robert Schaback. Solving the Laplace equation by meshless collocation using harmonic kernels. *Advances in Computational Mathematics*, 31:457–470, 2008. doi: 10.1007/s10444-008-9078-3.

Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *Advances in Neural Information Processing Systems*, 33, 2020.

Sifan Wang, Yujun Teng, and Paris Perdikaris. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5):A3055–A3081, 2021. doi: 10.1137/20M1318043.

Sifan Wang, Xinling Yu, and Paris Perdikaris. When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, 449:110768, 2022. doi: 10.1016/j.jcp.2021.110768.

Chenxi Wu, Min Zhu, Qinyang Tan, Yadhu Kartha, and Lu Lu. A comprehensive study of non-adaptive and residual-based adaptive sampling for physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, 403:115671, 2023. doi: 10.1016/j.cma.2022.115671.

Yaohua Zang, Gang Bao, Xiaojing Ye, and Haomin Zhou. Weak adversarial networks for high-dimensional partial differential equations. *Journal of Computational Physics*, 411:109409, 2020. doi: 10.1016/j.jcp.2020.109409.

A HARMONICITY AND BRANCH CUTS

This section states when the model of Section 3 is harmonic and how the enrichment terms are placed so that the conditions hold.

Lemma 1 (Harmonicity)**.** *Assume that each network branch is entire and that every enrichment term is twice continuously differentiable and harmonic in* Ω*. In 3D, also assume* $a_k \cdot a_k = 0$ *for every projection direction. Then the representations in Equations* 3 *and* 6 *are harmonic in* Ω*.*

Proof. In 2D, the real part of each holomorphic term is harmonic. In 3D, each branch has zero Laplacian by Equation 5, and the centring and scaling of the input only multiply it by L_0^{-2}. The enrichment terms are harmonic by assumption, and taking real parts and summing preserves harmonicity by linearity of the Laplacian. □

The network branches satisfy the lemma automatically; the enrichment terms satisfy it only if their branch cuts stay outside the domain. We therefore orient each cut along a direction pointing out of the domain: the outward normal on a segment or face, and the exterior bisector at a corner (Figure 5). On the convex domains (the rectangles, the square and disk of the suite, the detector) such a cut meets the closure of the domain only at its site. On the sectors and the L-shape it runs through the removed wedge, on the slit disk it lies on the slit, and on the L-prism it lies in the removed quadrant. The cone mode has no cut and is singular only on the ray $\vartheta = \pi$, outside the domain. Site positions and orientations are fixed during training.

The cone mode is evaluated by the series $P_\nu(\cos\vartheta) = {}_2F_1(-\nu, \nu+1; 1; (1-\cos\vartheta)/2)$ truncated at 400 terms. Its argument is at most 0.854 on the domain, so the truncation error is below 10^{-27}.

The value-jump term of Table 1 is $\arg w = \Re(-i \operatorname{Log} w)$, which carries the jump in the data and stays bounded at the junction. A general complex coefficient on $\operatorname{Log} w$ would add a radial part, since $\Re(c \operatorname{Log} w) = \Re(c) \ln|w| - \Im(c) \arg w$, and $\ln|w|$ is unbounded at the junction. This is the component excluded in Section 3: with it, the error would be unbounded and Equation 8 would not apply.

B CHOICE OF ARCHITECTURE

The 2D network size and optimizer settings come from the ablation of Table 7. The 3D size, 16 branches of two hidden layers of width eight, was fixed after a preliminary single-seed screen of 135 configurations on four development problems, varying depth, width, activation, initialization, the number of branches and the treatment of the directions. We keep one configuration per dimension for all problems.

C NAMED BENCHMARKS IN TWO DIMENSIONS

Table 4 defines the six benchmarks. Each has a single singular site; every other junction has an integer exponent, so Table 1 emits no term there.

Table 4: The named benchmarks. λ is the leading exponent at the singular site. The reference coefficients are from Lu et al. (2004).

Problem	Domain	λ	Reference
Sector $3\pi/2$	$r < 1, 0 < \theta < 3\pi/2$	$2/3$	exact
Sector $7\pi/4$	$r < 1, 0 < \theta < 7\pi/4$	$4/7$	exact
L-shape	$(-1,1)^2 \setminus [0,1] \times [-1,0]$	$2/3$	exact
Slit disk	$r < 1$ cut along $0 \le x < 1, y = 0$	$1/2$	exact
Motz	$(-1,1) \times (0,1)$	$1/2$	$D_0 = 401.162453745234416$
Cracked beam	$(-1,1) \times (0,1)$	$1/2$	$D_0 = 540.56512271362749$

The first four carry Dirichlet data from the exact solution on the whole boundary. On the sectors, the data are zero on both radii and $r^\lambda \sin(\lambda\theta)$, $\lambda = \pi/\omega$, on the arc. On the L-shape, they are zero

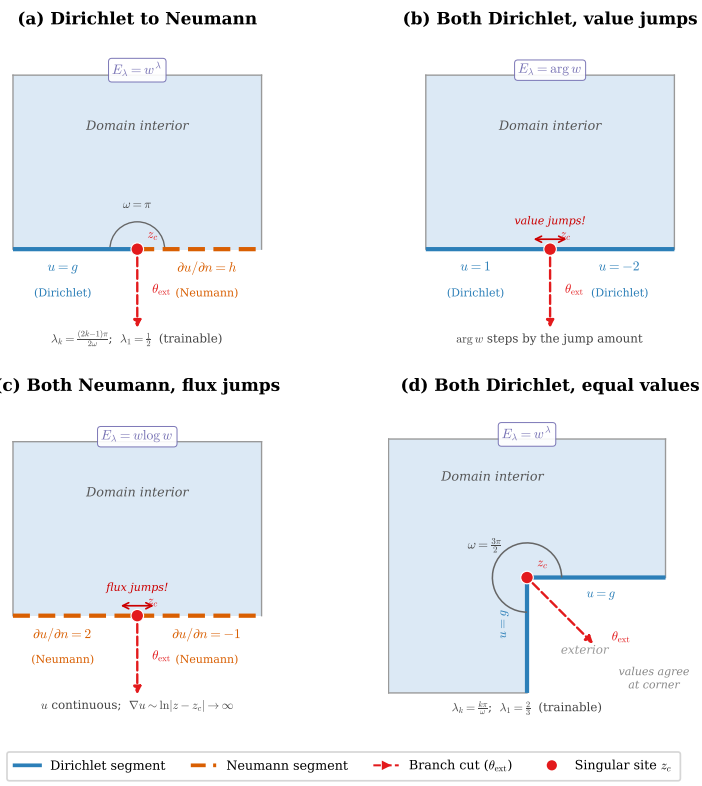

Figure 5: Enrichment terms and branch cuts for the rows of Table 1. **(a)** A Dirichlet to Neumann junction gives w^λ. **(b)** A Dirichlet value jump gives $\arg w$, which carries the jump. **(c)** A Neumann flux jump gives $w \operatorname{Log} w$. **(d)** A re-entrant corner with equal Dirichlet values gives w^λ; the equal-flux Neumann row has the same term. Each cut leaves the site along the exterior direction.

on the two edges meeting at the re-entrant corner and $r^{2/3}\sin(2\theta/3)$ on the outer edges. On the slit disk, they are zero on both faces of the slit and $\sin(\theta/2)$ on the rim, so that $u = r^{1/2}\sin(\theta/2)$.

Motz's problem and the cracked beam have mixed data. Both have $u = 0$ on the bottom edge for $x < 0$ and zero flux there for $x > 0$. Motz's problem has zero flux on the top and left edges and $u = 500$ on the right edge; the cracked beam has zero flux on the left and right edges and $u = 500$ on the top edge. Both are scored against the series $\sum_{i=0}^{59} D_i r^{i+1/2}\cos((i+\frac{1}{2})\theta)$, which satisfies the bottom-edge conditions exactly, with coefficients fitted by least squares on the other three edges. It reproduces the published D_0 to relative errors of 7×10^{-16} and 1.1×10^{-15}. Since the model's exponents are fixed on these problems, its leading coefficient compares directly with D_0. Table 5 shows how far the singular terms go without the network.

Table 5: Least-squares fit of the first n singular terms without the network: ϵ_u on 20,000 random interior points and the relative error of D_0. The model uses $n = 4$.

	Motz		Cracked beam	
n	ϵ_u	D_0 error	ϵ_u	D_0 error
1	1.9×10^{-1}	7.9×10^{-2}	3.0×10^{-1}	1.5×10^{-1}
4	5.5×10^{-3}	4.4×10^{-3}	2.2×10^{-2}	1.1×10^{-2}
8	1.9×10^{-4}	3.5×10^{-5}	4.4×10^{-4}	8.6×10^{-6}
12	1.6×10^{-5}	5.1×10^{-7}	2.6×10^{-5}	2.8×10^{-7}

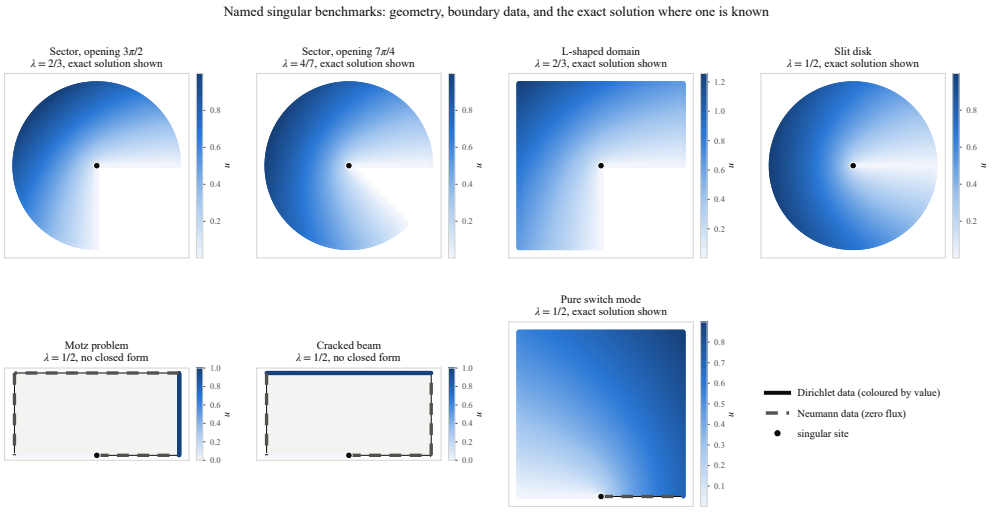

Figure 6: The six benchmarks of Table 4, with the exact solution drawn where one exists.

D CZT DETECTOR CROSS-SECTION

The 2D detector is a cross-section, 40 mm wide and 5 mm thick, of the crystal of Appendix I. One face carries 24 anode strips at 0 V, 0.2 mm wide at a pitch of 1.6 mm, with 25 groups of three drift strips at -50, -100 and -50 V between them. The opposite face is a single cathode at -350 V, and potentials are divided by -350 V. The gaps and side walls carry zero flux, so each of the 198 strip edges is a Dirichlet to Neumann junction and receives one term w^{λ}, with λ trained from $1/2$. The reference is a finite-difference solution, and Table 6 gives the errors on its grid.

Table 6: Potential errors on the CZT cross-section, in units of 350 V, median over three seeds.

Model	MSE	RMSE
MLP PINN	2.40×10^{-1}	4.90×10^{-1}
Plain	1.18×10^{-4}	1.09×10^{-2}
Enriched	$\mathbf{4.19 \times 10^{-5}}$	$\mathbf{6.47 \times 10^{-3}}$

E SAMPLED SUITE

The suite tests many sites on simple geometries. It has 24 instances, twelve on the unit square and twelve on the unit disk, in three groups of eight: Dirichlet data with value jumps, mixed Dirichlet and Neumann data, and configurations with many segments (Figure 7). The instances have 4 to 12 sites, with a median of 6.5, and are drawn with seed 20260713; the tuning suite of Table 7 is drawn separately with seed 20260804. The four Dirichlet-jump instances on the disk have a closed-form solution and are scored by ϵ_u, the eight Dirichlet-jump instances also by the sampled right-hand side of Equation 8, and all instances by their boundary loss. Exponents are fixed at their dictionary values and no near-site samples are used.

F TRAINING IN TWO DIMENSIONS

The model is implemented in JAX in double precision, since the reported errors go well below the 10^{-7} resolution of single precision. Weights are drawn from the scaled complex He initialization of Calafà et al. (2024). Training runs 4×10^4 full-batch Adam steps at learning rate 10^{-2} with cosine decay, followed by at most 6×10^3 L-BFGS iterations with a history of 100 (Liu & Nocedal, 1989), and keeps the best of three restarts. Each segment receives 1,024 uniform samples, and

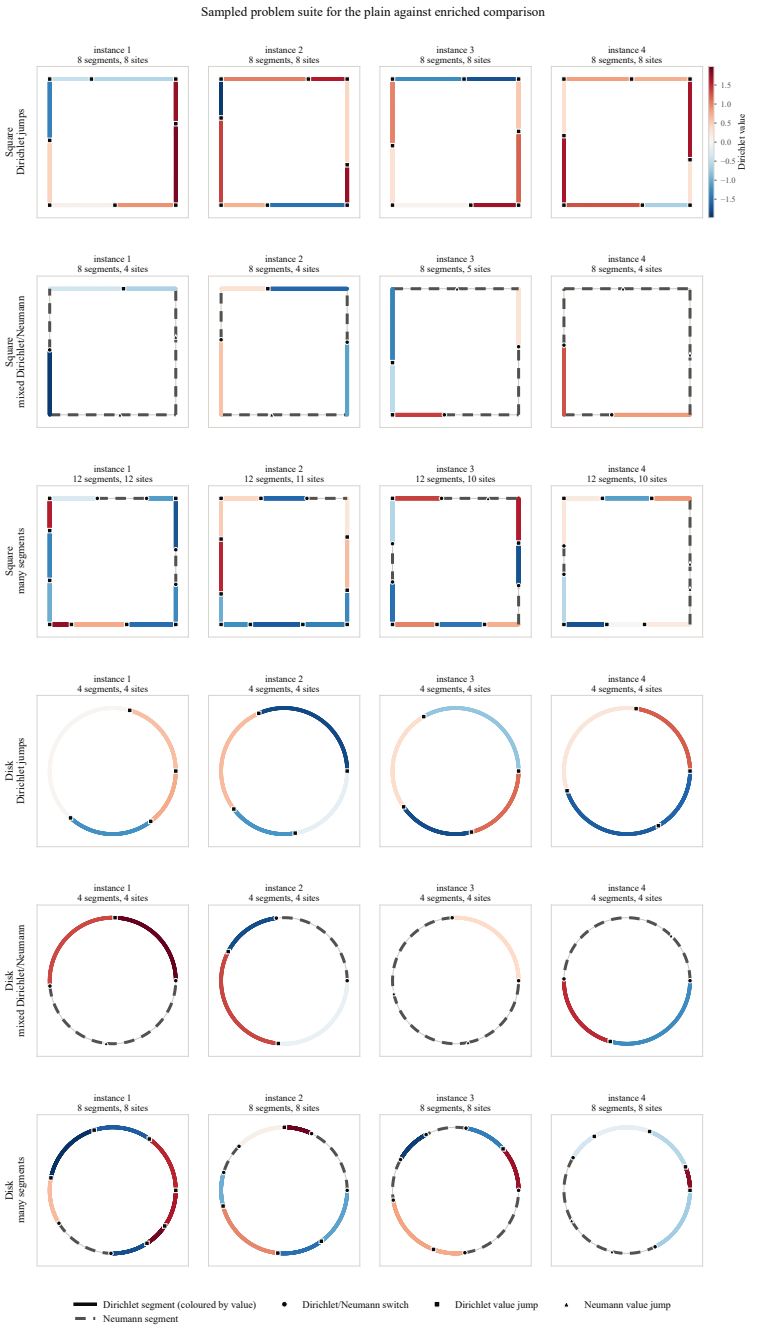

Figure 7: The 24 instances of the sampled suite. Dirichlet segments are coloured by their prescribed value, and Neumann segments are dashed. Markers show the sites the dictionary emits, with the shape giving the row of Table 1.

each endpoint 128 further samples at log-uniform distances between 10^{-4} and 10^{-1} of the segment length. The sectors, the L-shape, the slit disk and the detector train their exponents; Motz's problem and the cracked beam fix them at $1/2$, $3/2$, $5/2$ and $7/2$.

The budget ladder multiplies the Adam steps and the samples per segment by 1, 1.5, 2 and 3, and each rung is run with seeds 0, 1 and 2. Table 2 reports the best rung of each method. All other settings were chosen on the tuning suite (Table 7), which is disjoint from the reported problems.

Baseline. The MLP PINN is a real-valued tanh network with one hidden layer of width 128, trained in single precision on 5,000 interior points and $400b$ boundary points at budget b. Its loss is the Laplace residual plus the boundary misfit, with boundary weight 10 on the sectors and 100 elsewhere, minimized by Adam at learning rate 10^{-3} followed by 6,000 L-BFGS iterations, on the same budget ladder and seeds.

Table 7: Ablation on the tuning suite (six instances, three seeds, 18 runs per row), changing one factor at a time from a starting configuration of width 32, depth 3, learning rate 0.003, 20 000 Adam steps, 6 000 L-BFGS steps, 512 points per segment, ladder depth 1 and fixed exponents. *Ratio* is the median boundary loss relative to the starting configuration; below one is an improvement. *Converged* counts runs with boundary loss below 10^{-6}. The configuration of Section 4 combines the five best changes.

Configuration	Boundary loss	Ratio	Converged	Time (s)
starting configuration	6.4×10^{-5}	1.00	3/18	402
depth 2	3.3×10^{-6}	0.03	7/18	280
learning rate 0.01	8.1×10^{-6}	0.14	7/18	381
Adam steps 40 000	1.4×10^{-5}	0.41	5/18	644
width 16	1.4×10^{-5}	0.42	7/18	275
points per segment 1 024	8.5×10^{-6}	0.43	6/18	645
depth 4	6.0×10^{-5}	0.84	1/18	512
L-BFGS steps 12 000	3.8×10^{-5}	0.85	3/18	505
initialization empirical	7.6×10^{-5}	0.95	3/18	381
trained exponents	5.1×10^{-5}	0.98	3/18	457
ladder depth 2	6.9×10^{-5}	1.00	3/18	415
ladder depth 4	8.0×10^{-5}	1.00	3/18	487
L-BFGS steps 3 000	7.2×10^{-5}	1.13	3/18	325
learning rate 0.001	6.5×10^{-5}	1.38	4/18	385
Adam steps 10 000	1.4×10^{-4}	2.33	1/18	254
width 64	1.4×10^{-4}	2.36	3/18	1 079
Adam steps 1 000	3.8×10^{-4}	3.54	4/18	138
Adam steps 3 000	5.3×10^{-4}	7.07	2/18	165
points per segment 256	1.6×10^{-3}	11.39	3/18	238
plain model	4.0×10^{-3}	60.97	0/18	369

G PLAIN AGAINST ENRICHED

The plain and enriched models share the network, budget, restarts and seeds, and differ only in the singular terms. On the named benchmarks (Figure 8), tripling the budget improves the plain model by less than an order of magnitude and the enriched model by several. On the sampled suite (Figure 9), all 72 paired runs improve with enrichment. The median paired ratio is 1.5×10^3 for the boundary loss, 1.1×10^7 for the sampled right-hand side of Equation 8 on the eight Dirichlet-jump instances, and 3.2×10^6 for ϵ_u on the four closed-form instances. The gain is largest on value jumps and smallest on the many-segment instances.

To compare the error with the sampled bound of Equation 8, each Dirichlet segment is sampled at 2×10^4 points, excluding points within 10^{-3} of a site. On the two sectors, the largest error of the enriched model on a 400×400 interior grid reaches 94% to 100% of this bound; on the L-shape and the $7\pi/4$ sector both are at roundoff. For the plain model the ratio is 83% to 90% on all four.

Ladder depth. Table 8 adds terms on the $7\pi/4$ sector, whose solution is the leading mode alone. The leading exponent stays within 2×10^{-8} of its true value at every depth, while the surplus exponents cluster and carry amplitudes that cancel in pairs. The error stays near 10^{-10}, so extra terms neither help nor harm here. These runs use budget one, which is why the depth-one error is above that of Table 2.

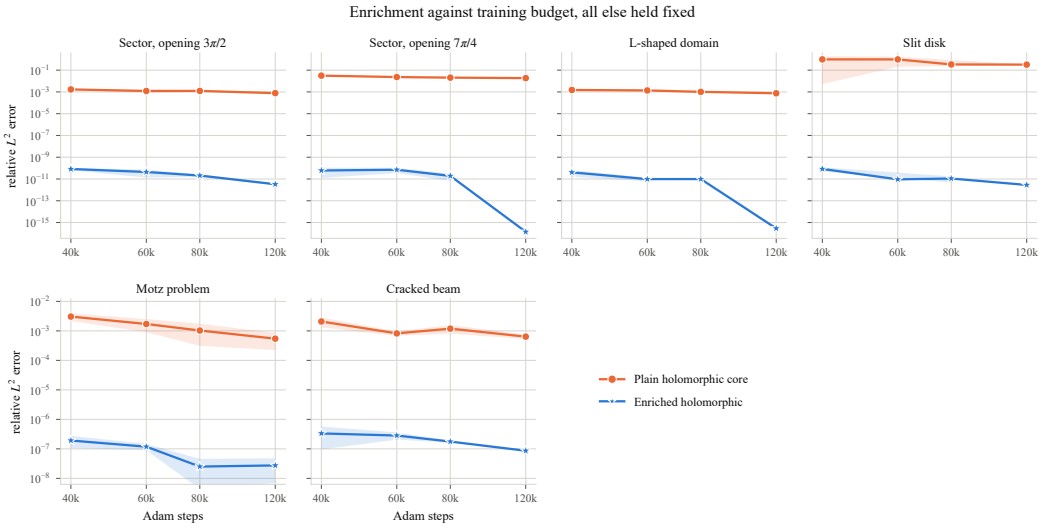

Figure 8: The plain and enriched models across the budget ladder on the named benchmarks, median over three seeds with the spread shaded.

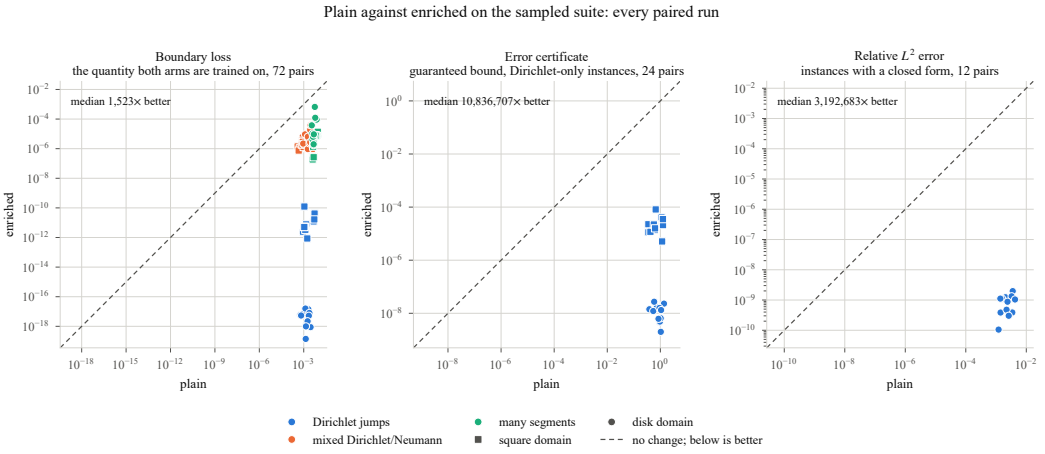

Figure 9: Plain against enriched over the sampled suite, one point per instance and seed. Points below the diagonal improved with enrichment; annotations give the median paired ratio.

H THE 3D MODEL ON THE L-SHAPED PRISM

On the L-shaped prism the only singular site is the re-entrant edge, with $\omega = 3\pi/2$ and zero Dirichlet data on both walls. Table 1 assigns it the term $w^{2/3}$, and the model is

$$u(\boldsymbol{x}) = \Re\left[\sum_{k=1}^{16} f_{\eta_k}(\zeta_k) + c\, w^{2/3}\right].$$

The exact solution $r^{2/3}\sin(2\theta/3)$ equals $\Re[-i\, w^{2/3}]$, so $c = -i$ represents it exactly and the branches need only represent zero. On u_A and u_B the singular amplitude varies along the edge, and the branches must carry that variation.

Table 8: Ladder depth on the $7\pi/4$ sector, $\lambda^\star = 4/7$. λ_1 is the trained exponent closest to $\lambda^\star$, the gap is the smallest separation between trained exponents, and the spurious amplitude is the largest $|c_q|$ on a non-leading term.

Terms	$\lvert\lambda_1 - \lambda^\star\rvert$	Smallest gap	Spurious $\lvert c_q\rvert$	ϵ_u
1	4.6×10^{-11}	–	0	5.7×10^{-11}
2	2.3×10^{-10}	6.3×10^{-1}	1.9×10^{-9}	3.4×10^{-11}
3	7.4×10^{-10}	9.5×10^{-8}	3.7×10^{-2}	2.2×10^{-11}
4	8.8×10^{-10}	1.0×10^{-7}	1.8×10^{-1}	3.0×10^{-11}
6	2.0×10^{-8}	2.5×10^{-6}	1.0×10^{-1}	1.3×10^{-10}
8	1.9×10^{-9}	4.7×10^{-7}	2.2×10^{-1}	1.1×10^{-10}

I THREE-DIMENSIONAL PROBLEMS

Closed-form solution. The L-prism is the L-shape of Table 4 extruded over $0 < z < 1$, with Dirichlet data from the exact solution on every face; u_A and u_B use the same geometry. The smooth cube is $[0, 1]^3$ with Dirichlet data from $1/\|\boldsymbol{x} - (1.6, 1.6, 1.6)\|$, whose pole lies outside the cube.

Series reference. The conical notch is the unit ball with a cone of half-angle $\pi/4$ removed around the positive z axis, so that, measured from the negative axis, the wall lies at $\vartheta = 3\pi/4$. The data are zero on the wall and exact on the sphere, and the solution is Equation 11 with $\nu = 0.4630985617801066$; the trained degree, started at $1/2$, reaches 0.463099. The extruded Motz problem has the planar Motz data on its cross-section and zero flux on its end faces. Its 60-term series reference changes by 5.5×10^{-13} when reduced to 40 terms and matches the published D_0 to 7×10^{-16}.

CZT detector. The detector is a $40 \times 5 \times 40$ mm crystal, posed in millimetres and volts. One face carries the anode and drift strips of Appendix D, running the full length of the crystal, and the opposite face 20 cathode strips at -350 V, 1.8 mm wide at a pitch of 2 mm. The gaps and the four sides carry zero flux, so each of the 119 electrodes has two Dirichlet to Neumann edges on a flat face, 238 in total, each with one edge term and λ trained from $1/2$; the strip ends are not enriched. Zero flux idealizes an insulating surface, and space charge is neglected, so this is a numerical study of the operating potential rather than a validation of the device. The reference is a graded finite-element solution whose own error is about 10^{-4}.

J TRAINING AND EVALUATION IN THREE DIMENSIONS

Table 9 lists the settings of the enriched model. The core has 3,104 real parameters; with the directions and enrichment terms the model has between 3,200 and 4,390, depending on the number of sites. Table 10 gives the per-seed errors behind Table 3.

Table 9: Settings of the enriched model in 3D.

Setting	Value
Branches / hidden layers / width	$16/2/8$
Adam steps	120,000, learning rate 0.01, cosine decay
Maximum L-BFGS iterations	4,000
Uniform boundary samples	1,536 per face (1,024 on the detector)
Near-site samples	128 per site, 10^{-4} to 10^{-1} of the local length
Restarts	3 per seed, selected on a fresh boundary sample
Seeds	$0, 1, 2$
Precision	double
Evaluation	10,000 interior points per problem

Baselines. Each baseline runs under its own published protocol: the architecture, optimizer, sampling and step counts of its paper, with the authors' code where it runs on our problems. It is also

Table 10: Per-seed relative error ϵ_u of the enriched model in Table 3.

Problem	Seed 0	Seed 1	Seed 2
L-prism	1.3×10^{-8}	1.7×10^{-9}	6.9×10^{-9}
u_A	1.8×10^{-4}	1.9×10^{-4}	1.6×10^{-4}
u_B	1.8×10^{-4}	1.6×10^{-4}	1.6×10^{-4}
Smooth cube	4.4×10^{-6}	4.9×10^{-6}	4.6×10^{-6}
Conical notch	4.5×10^{-9}	6.0×10^{-9}	6.1×10^{-9}
Extruded Motz	3.4×10^{-6}	2.1×10^{-6}	4.6×10^{-6}

run with its steps and boundary samples multiplied by 1.5, 2 and 3, and Table 3 reports its lowest three-seed median. Deep Ritz uses three residual blocks of width ten with cubic ReLU activations and a boundary penalty of 10^3. WAN is formulated for Dirichlet data; on the extruded Motz problem and the detector, the zero-flux condition enters through the weak form, and we mark these entries as an extension. Baselines that require unit scales are fitted on the detector with lengths divided by 5 mm and potentials by 350 V, and mapped back before scoring.