Robust Variational Physics-Informed Neural Networks

Sergio Rojas[1], Paweł Maczuga[2], Judit Muñoz-Matute[3,4], David Pardo[3,5], Maciej Paszynski[2]

[1]Instituto de Matemáticas, Pontificia Universidad Católica de Valparaíso, Chile
[2]Institute of Computer Science, AGH University of Science and Technology, Poland
[3]Basque Center for Applied Mathematics, Bilbao (BCAM), Spain
[4]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, USA
[5]University of the Basque Country (UPV/EHU), Leioa, Spain

9th European Congress on Computational Methods
in Applied Sciences and Engineering

3-7 June 2024, Lisboa, Portugal

Most cited PINNs-related Karniadakis' works according to Google Scholar (01/05/2024)

- (8873) Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). *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.
- (3199) Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). *Physics-informed machine learning*. Nature Reviews Physics, 3(6), 422-440.
- (1305) Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2017). *Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations*. arXiv preprint arXiv:1711.10561.

⋮

- (422) Kharazmi, E., Zhang, Z., & Karniadakis, G. E. (2021). *hp-VPINNs: Variational physics-informed neural networks with domain decomposition*. Computer Methods in Applied Mechanics and Engineering, 374, 113547.

⋮

- (272) Kharazmi, E., Zhang, Z., & Karniadakis, G. E. (2019). *Variational physics-informed neural networks for solving partial differential equations*. arXiv preprint arXiv:1912.00873.

Outline

1. VPINNs: Variational Physics-Informed Neural Networks
2. RVPINNs: Robust Variational Physics-Informed Neural Networks
   - Construction
   - Error estimates
   - Numerical results
3. Conclusions & further directions

VPINNs: Original approach[1]

Given:

- A (possibly non-linear) well-posed variational formulation:

$$\text{Find } u \in U, \text{ such that } r(u,v) := a(u,v) - l(v) = 0, \forall\, v \in V.$$

- A discrete test space $V_M := \text{span}\{\varphi_m\}_{m=1}^{M} \subseteq V$.

Idea: Obtain a DNN approximation of u by solving the minimization problem:

$$\text{Find } u_{\theta^*}, \text{ s.t. } \theta^* = \arg\min_{\theta} \mathcal{L}_r\left(u_\theta\right) := \sum_{m=1}^{M} r(u_\theta, \varphi_m)^2 + C(u_\theta),$$

where u_θ denotes a DNN output with trainable parameters θ, and $C(\cdot)$ a quadratic functional to impose BC's for the DNN output.

Problem: This approach is generally not robust and strongly depends upon the selection of the basis.

[1] As presented in: Kharazmi, E., et al. (2019). *Variational physics-informed neural networks for solving partial differential equations*. arXiv:1912.00873

A wrong discrete test space example

- As an illustrative example, authors[1] consider the following variational problem:

  Find $u \in H^1(-1,1)$, such that $u(-1) = u(1) = 0$, and

$$a(u,v) := \int_{-1}^{1} u' v' dx = l(v) := \int_{-1}^{1} f\, v\, dx, \quad \forall\, v \in H_0^1(-1,1).$$

- Setting $V_M = \text{span}\,\{\varphi_m := \sin(m\pi x)\}_{m=1}^{M} \subseteq H_0^1(-1,1)$, they propose the following loss function:

$$\mathcal{L}_r\,(u_\theta) = \sum_{m=1}^{M} r(u_\theta, \varphi_m)^2 + u_\theta(-1)^2 + u_\theta(1)^2.$$

[1] Kharazmi, E., et al. (2019). *Variational physics-informed neural networks for solving partial differential equations*. arXiv:1912.00873

Wrong test example: Poisson's equation with delta source

Variational problem:

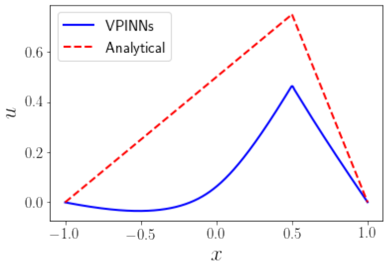
Find $u \in H_0^1(-1,1)$, such that:

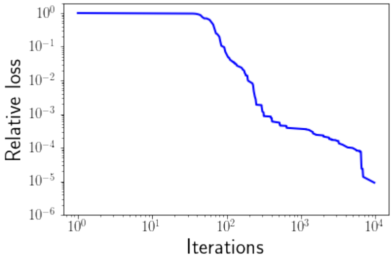
$$\int_{-1}^{1} u'v'\,dx = v(0.5),\ \forall v \in H_0^1(-1,1)$$

DNN Setting:

- Activation function: tanh
- Neurons: 25
- Hidden layers: 5
- N° test functions = 50
- Integration: Trapezoidal rule
- Integration nodes: 501
- Iterations: 10,000
- Optimizer: Adam
- Learning rate: 0.01

Best approximation (loss sense)





How can we define adequate loss functions?

We propose a general framework (RVPINNs) to define robust loss functions based on the computation of a discrete residual representative with respect to an adequate discrete norm.

RVPINNs assumptions

In the following, we assume a **linear** problem of the form:

$$\text{Find } u \in U, \text{ such that } r(u,v) := a(u,v) - l(v) = 0, \forall\, v \in V,$$

satisfying:

- **Boundedness:** There exists a constant $\mu > 0$, such that

$$a(w\,,v) \leq \mu \|w\|_U \|v\|_V, \quad \forall\, w \in U, v \in V,$$

- **Inf-sup stability:** There exists a constant $\alpha > 0$, such that

$$\sup_{0 \neq v \in V} \frac{a(w\,,v)}{\|v\|_V} \geq \alpha \|w\|_U, \quad \forall\, w \in U.$$

- **Adjoint injectivity**: For all $v \in V$,

$$a(w\,,v) = 0, \forall\, w \in U \implies v = 0.$$

Under previous assumptions, the problem is well-posed, and it holds:

$$\sup_{0 \neq w \in U} \frac{a(w\,,v)}{\|w\|_U} \geq \alpha \|v\|_V, \quad \forall\, v \in V.$$

VPINNs: An alternative definition

Given $V_M = \text{span}\{\varphi_m\}_{m=1}^{M}$ and the original loss functional from VPINNs

$$\mathcal{L}_r(u_\theta) = \sum_{m=1}^{M} r(u_\theta, \varphi_m)^2 + C(u_\theta),$$

if we define the following function in V_M

$$\widetilde{\varphi} = \sum_{m=1}^{M} r(u_\theta, \varphi_m)\, \varphi_m,$$

as a consequence of the linearity of the weak residual, it holds

$$\mathcal{L}_r(u_\theta) = \sum_{m=1}^{M} r(u_\theta, \varphi_m)^2 + C(u_\theta) = r(u_\theta, \widetilde{\varphi}) + C(u_\theta).$$

Our goal is to define the loss functional in terms of a particular discrete test function $\widetilde{\varphi} \in V_M$ such that VPINNs becomes robust.

RVPINNs: Robust Variational Physics-Informed Neural Networks

Main idea: For a given trainable parameter θ, we compute the **Riesz representation of the residual** $\phi := \phi(\theta) \in V_M$ as olution of the following Galerkin problem:

$$(\phi, v_M)_V = r(u_\theta, v_M), \quad \forall\, v_M \in V_M, \tag{1}$$

and define the loss function as:

$$\mathcal{L}_r^{\phi}(u_\theta) := r(u_\theta, \phi) + C(u_\theta), \tag{2}$$

Key observations:

- Defining: $\phi := \sum_{m=1}^{M} \eta_m(\theta)\varphi_m$, problem (1) leads to the resolution of:

  $$G\eta(\theta) = R(\theta), \text{ with } G_{nm} = (\varphi_m, \varphi_n)_V, \text{ and } R_n(\theta) = r(u_\theta, \varphi_n).$$

  Thus,

  $$\mathcal{L}_r^{\phi}(u_\theta) = R(\theta)^T G^{-1} R(\theta) + C(u_\theta).$$

- To minimize the loss functional (2) is equivalent, up to the constraint $C(u_\theta)$, to minimize the quantity $\|\phi\|_V^2$. Indeed, evaluating (1) with $v_M = \phi$, gives

  $$\|\phi\|_V^2 = (\phi, \phi)_V = r(u_\theta, \phi).$$

- $\|\phi\|_V^2$ is an aposteriori error estimator for $\|u - u_\theta\|_U^2$.

Orthonormal discrete basis and relation with other VPINNs

- When a test space V_M is the span of an orthonormal set $\{\varphi_m\}_{m=1}^{M}$ with respect to the $\|\cdot\|_V$-norm, the Gram matrix G coincides with the identity matrix; therefore, the corresponding residual representative has the form:

$$\phi = \sum_{m=1}^{M} r(u_\theta, \varphi_m)\varphi_m,$$

  and the loss function is explicitly written as (coinciding with the classical approach):

$$\mathcal{L}_r^{\phi}(u_\theta) = \sum_{m=1}^{M} r(u_\theta, \varphi_m)^2 + C(u_\theta).$$

- The Deep Fourier Residual method[2] is a particular case of RVPINNs.

[2] Taylor, J. M., Pardo, D., & Muga, I. (2023). *A Deep Fourier Residual method for solving PDEs using Neural Networks. Computer Methods in Applied Mechanics and Engineering, 405, 115850.*

A posteriori error estimates for linear problems

Main complexities for proving the robustness of $\|\phi\|_V$:

- The solution of the Petrov-Galerkin problem **may not have a solution** or, if there exists, **it may be non-unique** since the space of all possible realizations for the NN structure defines a manifold instead of a finite-dimensional space[3].
- Standard FEM arguments based on a discrete inf-sup condition cannot be applied in this context.

Approach:

- We introduce an equivalence class that allows us to neglect the part of the error that is a-orthogonal to V_M. For that equivalence class, we prove that the residual representative is a reliable and efficient a posteriori estimator for the error.
- For the full error, we demonstrate its equivalence to the residual error estimator up to an oscillation term and under the assumption of the existence of a local Fortin operator.

[3] See, e.g., Section 6.3 in Berrone, S., Canuto, C., & Pintore, M. (2022). *Variational physics informed neural networks: the role of quadratures and test functions. Journal of Scientific Computing, 92(3), 100.*

A posteriori error estimates for RVPINNs in an equivalence class sense

- Let us define the following Null space of the operator $A : U \mapsto V_M'$:

$$U_M^0 := \{w \in U \,:\, \langle A(w)\,, v_M\rangle := a(w\,,\, v_M) = 0, \forall\, v_M \in V_M\}\,,$$

  and the following norm for the quotient space U/U_M^0:

$$\|[w]\|_{U/U_M^0} := \inf_{w_0 \in U_M^0} \|w + w_0\|_U.$$

- We extend the definition of the bilinear form $a(\cdot\,,\cdot)$ to the product space $U/U_M^0 \times V_M$ as:

$$a([w]\,, v_M) := a(w\,, v_M), \quad \text{with } w \in [w] \text{ being any arbitrary representative of } [w].$$

A posteriori error estimates for RVPINNs in an equivalence class sense

Proposition

The following boundedness and semi-discrete inf-sup conditions are satisfied

$$a([w], v_M) \leq \mu \|[w]\|_{U/U_M^0} \|v_M\|_V, \quad \forall [w] \in U/U_M^0,\ v_M \in V_M,$$

$$\sup_{0 \neq v_M \in V_M} \frac{a([w], v_M)}{\|v_M\|_V} \geq \alpha \|[w]\|_{U/U_M^0}, \quad \forall [w] \in U/U_M^0.$$

Theorem (Lower and upper bounds in terms of the residual representative)

Let $u \in U$ *be the solution of the continuous problem;* $u_\theta \in U_{NN}$ *denote a DNN structure with trainable parameters* $\theta \in \mathbb{R}^S$*;* $V_M \subseteq V$ *denote a finite-dimensional space, equiped with norm* $\|\cdot\|_V$*; and* $\phi \in V_M$ *be the Riesz representative of the weak residual. It holds:*

$$\frac{1}{\mu}\|\phi\|_V \leq \|[u - u_\theta]\|_{U/U_M^0} \leq \frac{1}{\alpha}\|\phi\|_V.$$

Energy norm error estimates

Corollary (Lower bound for the true error)

Under the same hypothesis of the previous Theorem, it holds:

$$\frac{1}{\mu}\|\phi\|_V \leq \|u - u_\theta\|_U.$$

Proposition (Upper bound for the true error)

Under the same hypothesis as before, if there exists $R > 0$ such that, for all $\theta \in B(\theta^*, R)$, there is a local Fortin operator $\Pi_\theta : V \mapsto V_M$ with a θ-independent constant $C_\Pi > 0$, it holds:

$$\|u - u_\theta\|_U \leq \frac{1}{\alpha}\,\mathrm{osc}(u) + \frac{1}{C_\Pi \alpha}\|\phi\|_V, \quad \forall\, \theta \in B(\theta^*, R),$$

with

$$\mathrm{osc}(u) := \sup_{0 \neq v \in V} \frac{a(u, v - \Pi_\theta v)}{\|v\|_V}.$$

where $B(\theta^*, R)$ denotes an open ball of center θ^* and radius R, with respect to a given norm of $\mathbb{R}^S$.

Model problem: 1D advection-diffusion problem

We consider the model problem:

$$
\begin{aligned}
-\varepsilon u'' + \beta u' &= f, \qquad \text{in } \Omega = (-1,1),\\
u &= 0, \qquad \text{on } \partial\Omega.
\end{aligned}
$$

We then set $U = V = H_0^1(\Omega)$ and consider the following continuous variational formulation:

$$\text{Find } u \in U \;:\; r(u,v) := l(v) - a(u,v),\, \forall\, v \in V,$$

with

$$a(u\,,v) := (\varepsilon u' - \beta u\,, v')\,, \quad \text{and} \quad l(v) = \langle f\,, v\rangle,$$

where $(\cdot\,,\cdot)$ denotes the L^2-inner product, and $\langle\cdot\,,\cdot\rangle$ denotes the duality map between V' and V.

We finally equip the Hilbert spaces U, V with the norms

$$\|w\|_U^2 := \|w\|_V^2 := \varepsilon\,(w', w').$$

Discrete setting

We consider two discrete test spaces:

- $V_M = \text{span}\{\varphi_m\}_{m=1}^{M}$ where φ_m being the standard globally continuous and piece-wise linear functions defined over a uniform mesh partition of Ω. The loss function in this case is

$$\mathcal{L}_r^{\phi}(u_\theta) = R(\theta)^T G^{-1} R(\theta) + C(u_\theta).$$

where $G_{nm} = (\varphi_m, \varphi_n)_V,$ and $R_n(\theta) = r(u_\theta, \varphi_n).$

- $V_M = \text{span}\{\varphi_m = \dfrac{2 s_m}{\sqrt{\varepsilon \pi} m}\}_{m=1}^{M}$ where $s_m = \sin\left(m\pi \dfrac{x+1}{2}\right)$, so the function φ_m are orthonormal with respect to $(\cdot,\cdot)_V$. The loss function in this case is

$$\mathcal{L}_r^{\phi}(u_\theta) = \frac{4}{\varepsilon \pi^2} \sum_{m=1}^{M} \frac{1}{m^2} r\left(u_\theta\, , s_m\right)^2 + C(u_\theta).$$

Example: A smooth diffusion problem ($\varepsilon = 1$, $\beta = 0$, and $u(x) = x\sin(\pi(x+1))$)

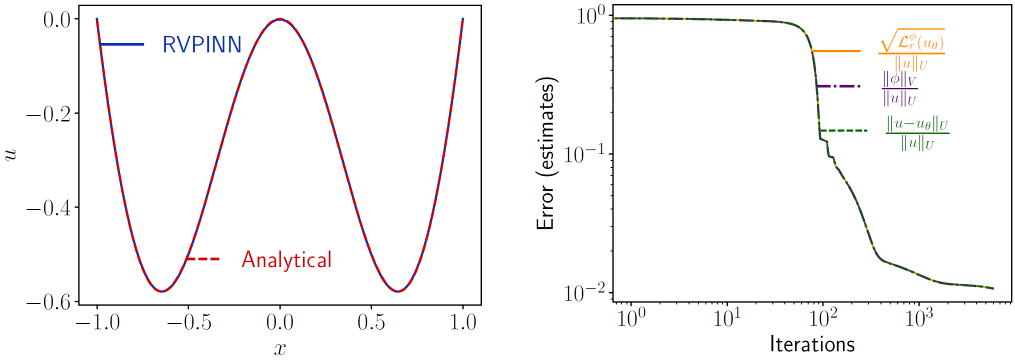


Figure: RVPINNs approximation with strong BCs imposition and 50 spectral test functions.

Example: A smooth diffusion problem ($\varepsilon = 1$, $\beta = 0$, and $u(x) = x\sin(\pi(x+1))$)

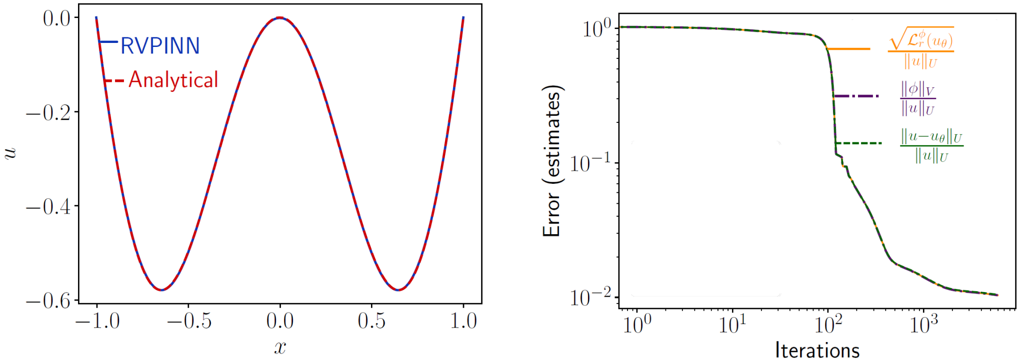


Figure: RVPINNs approximation with strong BCs imposition and 100 FE test functions.

Example: Delta source problem ($\varepsilon = 1$, $\beta = 0$, and $l(v) = \langle \delta_{1/2}, v \rangle$)

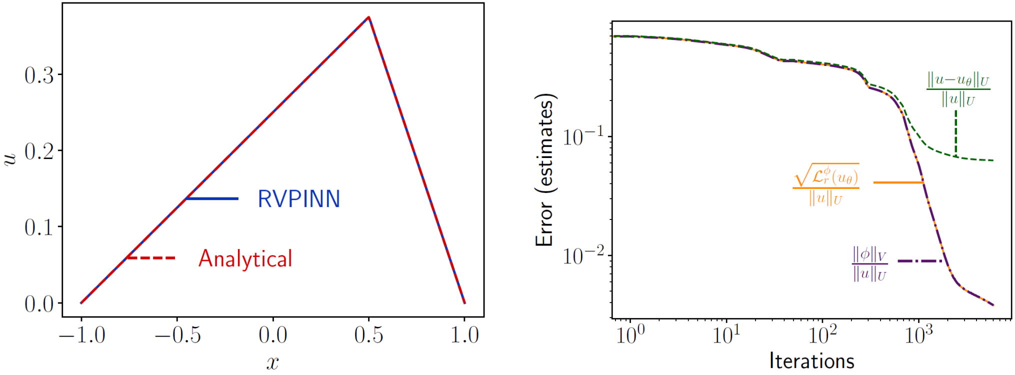
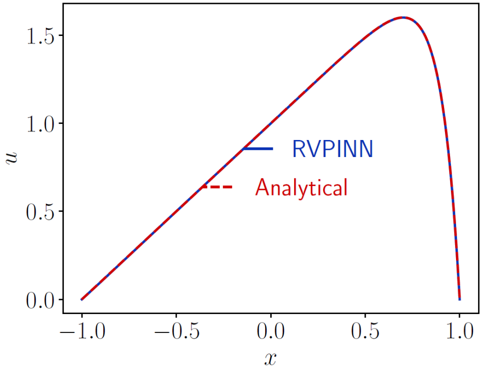


Figure: RVPINNs approximation with strong BCs imposition and 100 FE test functions.

Example: Advection-dominated-difussion problem

$$\left(\beta = 1,\ \epsilon = 0.1,\ f = 1, \text{ and } u(x) = \frac{2(1 - e^{\frac{x-1}{\varepsilon}})}{1 - e^{-\frac{2}{\varepsilon}}} + x - 1\right)$$



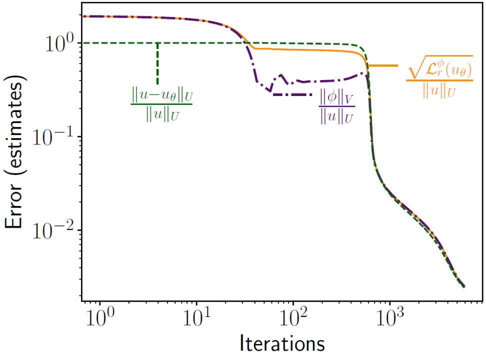


Figure: RVPINNs approximation with $\varepsilon = 0.1$, constrained BCs imposition, and 50 spectral test functions.

Example: Advection-dominated-difussion problem

$$\left(\beta = 1,\ \epsilon = 0.005,\ f = 1, \text{ and } u(x) = \frac{2(1 - e^{\frac{x-1}{\varepsilon}})}{1 - e^{-\frac{2}{\varepsilon}}} + x - 1\right)$$

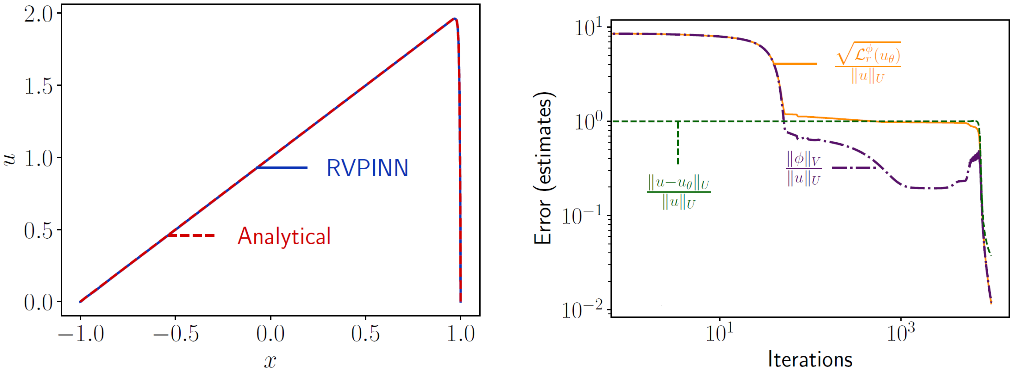


Figure: RVPINNs approximation with $\varepsilon = 0.005$, constrained BCs imposition, and 200 spectral test functions.

Example: 2D pure diffusion $\Delta u = f$ with exact solution $(u(x,y) = \sin(\pi x)\sin(\pi y))$

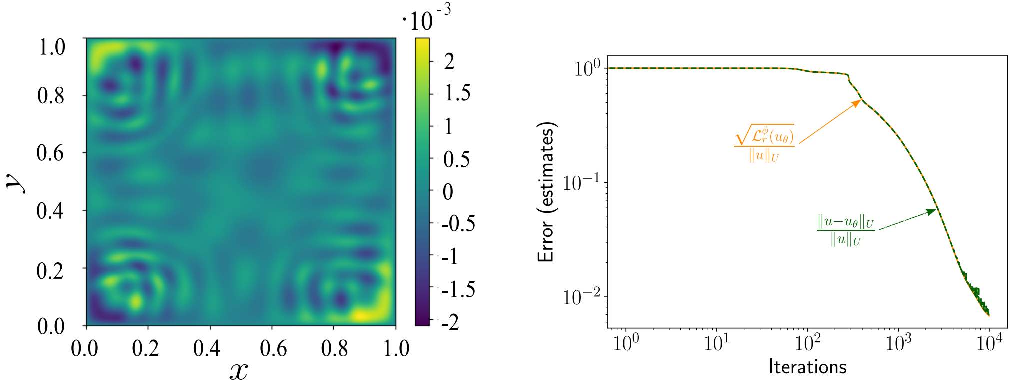


Figure: Error of RVPINNs approximation. 30×30 spectral test functions, strong imposition of zero Dirichlet BC.

Conclusions[4]

- We presented a general framework to define robust VPINNs losses.
- The strategy is based on defining the loss function in terms of a discrete Riesz representative for the residual.
- We derive a posteriori error estimates in terms of the DNN structure and the discrete test.

Further directions

- Other combinations of DNN structures, VFs, and discrete tests
- Extension to other PDEs
- Efficient integration techniques
- Adaptive strategies (in the test) to speed up computational time
- Application to data interpolation
- Hyperbolic conservation laws

---

[4] S. Rojas, P. Mazuga, J. Muñoz-Matute, D. Pardo, and M. Paszynski. (2024). *Robust Variational Physics-Informed Neural Networks*. Computer Methods in Applied Mechanics and Engineering, 425, 116904.