The 9th European Congress on Computational Methods in Applied Sciences and Engineering
ECCOMAS Congress 2024
3–7 June 2024, Lisboa, Portugal

A PDE-INFORMED AND PHYSICS-INFORMED NEURAL NETWORK FOR LEARNING WEAK SOLUTIONS IN TWO-DIMENSIONAL CONSERVATION LAWS

EDUARDO ABREU[1], JOAO B. FLORINDO[1] AND WANDERSON LAMBERT[2]

[1] Universidade Estadual de Campinas - UNICAMP/IMECC, SP, Brazil, {eabreu,florindo}@unicamp.br

[2] Federal University of Alfenas, MG, Brazil, wanderson.lambert@unifal-mg.edu.br

Key words: Physics-Informed Neural Networks, Weak Solutions, Entropic Flux

Summary. Here we propose a hybrid PINN model to learn weak solutions of conservation laws. The entropy flux, which satisfies admissibility condition and, as a consequence, grants uniqueness to the weak solution, is learned by a fully-connected neural network. The entropy rate principle is used to compose the PINN residual in our model. The proposed approach is also extended to problems in the two-dimensional domain. Considering that ensuring convexity of a deep neural network output is in general not possible, here we combine the learned function with Kruzkhov entropy. By setting appropriate weights to the Kruzkhov and the learned residuals, we ensure a valid weak solution under the admissibility criteria at the same time that the learning process of the entropy provides more flexibility for the model optimizer to search for the most accurate solution, i.e., *weak PDE-informed*. The model was validated in the inviscid Burgers' equation with rarefaction waves, moving shock, and Riemann initial conditions in 2D, as well as with the linear advection equation with periodic initial data. Our results suggest that the proposed model is capable of providing qualitatively good solutions with fewer training epochs.

1 INTRODUCTION

Deep neural networks (DNN) have been recently used to solve partial differential equations (PDE) in the most diverse problems [Raissi et al., 2019, Abreu and Florindo, 2021, Carniello et al., 2022]. Nevertheless, in most of those studies, they are applied to relatively "well-behaved" configurations, for example, with continuous initial data and smooth solutions.

Furthermore, several "ad-hoc" mechanisms have been employed, with strategies that can be applied only to a limited number of scenarios. In this regard, we observe that little theoretical foundation has been developed and more in-depth discussion is necessary.

Hyperbolic transport models [Galvis et al., 2020, Alibaud et al., 2020] are a challenging example that has not been sufficiently explored by learning-based methods. Inspired by [De Ryck et al., 2024], here we explore a hybrid weak PINN formulation in two dimensions. We also introduce the minimization of an entropy rate as in [Dafermos, 2009] to the PINN neural network.

We test the model in challenging initial conditions, including moving shocks, periodic function, and rarefaction fan. Our results are qualitatively compared with the exact solution when available or with reference numerical schemes. Promising performance is achieved if we consider that the DNN model used here is reasonably simple. We also observe evidences that the weak entropic formulation of the functional as in [De Ryck et al., 2024] is beneficial for the solution, but further improvements are still necessary.

2 RELATED WORKS

Since the seminal work of Raissi et al. [Raissi et al., 2019], physics-informed neural networks have been widely used in the most diverse applications involving PDEs. This includes, for example, heat transfer [Cai et al., 2021], power systems [Misyris et al., 2020], fluid mechanics [Mao et al., 2020], and several others.

When it comes to hyperbolic conservation laws, on the other hand, the number of studies is significantly reduced and adjustments over the original model are necessary. In [Patel et al., 2022], the authors propose a reformulation of PINN adopting space-time control volume, obtaining in this way a hybrid model that combines the learning-based approach with classical strategies of numerical analysis. In [Rodriguez-Torrado et al., 2022], the focus is on the network architecture and attention mechanisms are introduced over the PINN model, which in this case also is designed using recurrent neural networks instead of the fully-connected architecture in [Raissi et al., 2019].

Another possibility in these more challenging scenarios is to adapt the problem formulation. This is explored in [De Ryck et al., 2024], where the deep neural networks is settled up to solve conservation laws in their weak formulations. Considering that this approach is more natural to the context of PDE solvers, this is also our choice here to develop PDE-informed PINN solutions for conservation laws.

3 PHYSICS-INFORMED NEURAL NETWORKS

The basic version of PINN model is proposed in [Raissi et al., 2019] and here we extend it for two dimensional problems. The partial differential equation to be solved is appropriately formulated as

$$u_t + \mathcal{D}(u) = 0, \qquad (x, y) \in \Omega \subset \mathbb{R}^2, t \in [0, T], \tag{1}$$

where $\mathcal{D}(\cdot)$ is a differential operator and $u(x, y, t)$ is its proper solution. The PINN model corresponds to a feed-forward neural network $\mathcal{N}^m : \mathbb{R}^{d_0} \to \mathbb{R}^{d_m}$ that is constructed to represent the solution of the PDE with m layers and d_k neurons in the k^{th} layer. The input layer should be d_0-dimensional and the output should have d_m neurons. For any $1 \leq k \leq m$, the output of the k^{th} layer is obtained by

$$C_k(z_{k-1}) = W_k z_{k-1} + b_k, \quad \text{for} \quad W_k \in \mathbb{R}^{d_k \times d_{k-1}}, z_{k-1} \in \mathbb{R}^{d_{k-1}}, b_k \in \mathbb{R}^{d_k}, \tag{2}$$

where z_{k-1} represents the outputs from the previous layer, whereas W_k and b_k are the learnable weights and biases, respectively. Denoting the collection of all trainable parameters by $\Theta = \{W_k, b_k\}$, the neural network output is given by

$$u_\Theta(z_0) = C_m \circ \sigma \circ C_{m-1} ... \circ \sigma \circ C_2 \circ \sigma \circ C_1(z_0), \tag{3}$$

where $\circ$ represents function composition, σ is the non-linear activation function, and z_0 in the neural network input.

The cost function of the neural network is defined as f, corresponding to the left hand side of each PDE:

$$f := u_t + \mathcal{D}(u). \tag{4}$$

Two losses are defined over f, u, the initial condition u_0, and the boundary condition g:

$$\begin{aligned}
\mathcal{L}_f(u) &= \frac{1}{N_f}\sum_{i=1}^{N_f} |f(x_f^i, y_f^i, t_f^i, u^i)|^2, \\
\mathcal{L}_{u_0}(u, u_0) &= \frac{1}{N_u}\sum_{i=1}^{N_u} |u(x_u^i, y_u^i, 0) - u_0^i|^2, \\
\mathcal{L}_{bc}(u, g) &= \frac{1}{N_b}\sum_{(x_b^i, y_b^i)\in\partial D} |u(x_b^i, y_b^i, t_b^i) - g(x_b^i, y_b^i, t_b^i)|^2,
\end{aligned} \tag{5}$$

where $\{x_f^i, y_f^i, t_f^i\}_{i=1}^{N_f}$ represents collocation points over f, $\{x_u^i, y_u^i, t_u^i, u_0^i\}_{i=1}^{N_u}$ corresponds to the initial values at specific points, and $\{x_b^i, y_b^i, t_b^i\}_{i=1}^{N_b}$ are values sampled over the domain boundary, where the value of the solution is constrained to be $g(x_b^i, y_b^i, t_b^i)$.

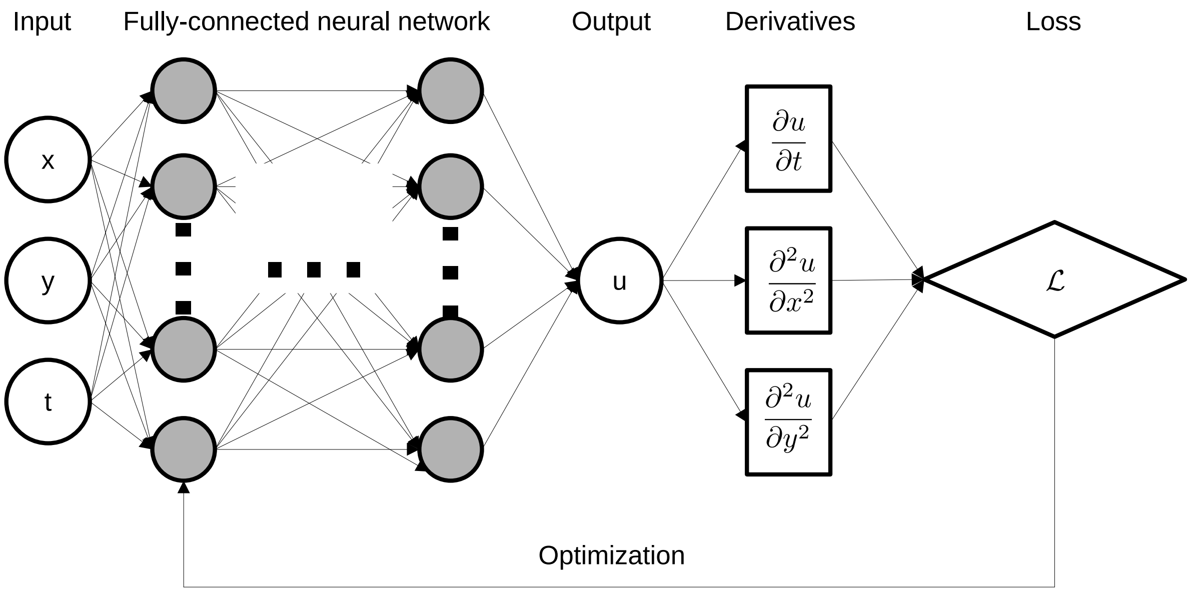


Figure 1: Physics-informed deep learning approach illustration.

The solution $u(x, y, t)$ provided by PINN is obtained by jointly minimizing both losses at the same time:

$$u(x, y, t) \approx \arg\min[\mathcal{L}_f(u) + \mathcal{L}_{u_0}(u, u_0) + \mathcal{L}_{bc}(u, g)]. \tag{6}$$

A diagramatic illustration of the PINN architecture is depicted on Figure 1.

4 PROPOSED METHOD

Let $\Omega \subset \mathbb{R}^2$ be the spatial domain and the conservation law

$$\begin{aligned} u_t + H(u)_x + H(u)_y = 0 \text{ in } \Omega \times [0,T] \\ u = u_0 \text{ in } \Omega \times \{0\}. \end{aligned} \tag{7}$$

Function $u \in L^\infty(\Omega \times [0,T])$ is a weak solution of (7) with $u_0 \in L^\infty(\Omega)$ if

$$\int_{[0,T]} \int_\Omega (u\phi_t + H(u)(\phi_x + \phi_y))dxdydt + \int_\Omega u_0(x,y)\phi(x,y,0)dxdy = 0 \tag{8}$$

for all test function $\phi \in \mathcal{C}_c^1(\mathbb{R}^2 \times \mathbb{R}_+)$ with compact support in $\Omega \times [0,T]$.

Nevertheless, weak solutions are not unique. In this scenario, we need extra admissibility constraints. This is the case, for example, of Kruzhkov entropy:

$$\partial_t |u - c| + (\partial_x + \partial_y) Q[u;c] \leq 0, \tag{9}$$

where $Q[u;c] = \text{sign}(u-c)(H(u) - H(c))$ for all $c \in \mathbb{R}$.

Equations 8 and 9 can be adapted to the PINN context by the establishment of a new residual term adapted from [De Ryck et al., 2024], which in two dimensions is given by:

$$\mathcal{L}_0(u,\phi,c) = -\int_{[0,T]} \int_\Omega (|u-c|\partial_t\phi + Q[u;c](\partial_x\phi + \partial_y\phi))dxdydt. \tag{10}$$

To further explore the role of entropies in weak formulations, we also take into consideration the rate of change of the total entropy in [Dafermos, 2009]. In the PINN residual, this contributes with the following extra term:

$$\mathcal{L}_1(u,\phi,c) = \mathcal{L}_0(u,\phi,c) + \frac{d}{dt}\int_\Omega \eta(u)dxdy, \tag{11}$$

where $\eta(u)$ is a general entropy function as defined in [Dafermos, 2009]. These entropy functions η are assumed to be $\mathcal{C}^2$ and convex. However, following [Panov, 1994], we can consider a specific case: the Kruzhkov entropy, which is the same entropy used in Eq. (9).

Finally, to stabilize the numerical solution around discontinuities, we add a diffusion term $\epsilon\Delta u$ (in weak form) for $\epsilon \in \mathbb{R}$ a small constant:

$$\mathcal{L}_2(u,\phi,c) = \mathcal{L}_1(u,\phi,c) - \epsilon \int_{[0,T]} \int_\Omega \nabla u \cdot \nabla\phi dxdydt. \tag{12}$$

In the numerical implementation, functional (12) is discretized by Monte-Carlo:

$$\begin{aligned} \mathcal{R}_{int} = & \sum_{i=1}^{N_{int}} \phi(x_i,y_i,t_i)\partial_t|u(x_i,y_i,t_i) - c| - \\ & Q[u;c](\partial_x\phi(x_i,y_i,t_i) + \partial_y\phi(x_i,y_i,t_i)) \\ & + \sum_{\substack{i=1 \\ t_i}}^{N_{int}} \partial_t \sum_{\substack{i=1 \\ x_i,y_i}}^{N_{int}} |u(x_i,y_i,t_i) - c| \\ & - \epsilon \sum_{i=1}^{N_{int}} [\partial_x u(x_i,y_i,t_i)\partial_x\phi(x_i,y_i,t_i) + \partial_y u(x_i,y_i,t_i)\partial_y\phi(x_i,y_i,t_i)]. \end{aligned}$$

Notice that here the time derivative taken over $|u-c|$ as in [De Ryck et al., 2024] accelerates the learning process. Following [De Ryck et al., 2024], we also apply L_2 normalization by the test function:

$$\mathcal{L}(u_\theta, \phi, c) = \frac{ReLU(\mathcal{R}_{int})}{\sum_{i=1}^{N_{int}} [\partial_x \phi(x_i, y_i, t_i)^2 + \partial_y \phi(x_i, y_i, t_i)^2]}, \tag{13}$$

where $ReLU(\cdot) = \max(0, \cdot)$ is the rectifier linear unit activation function.

Finally, initial and boundary conditions are expressed as usual in the PINN framework, by adding the respective terms:

$$\mathcal{L}_{u_0}(u, u_0) = \frac{1}{N_u} \sum_{i=1}^{N_u} |u(x^i, y^i, 0) - u_0^i|^2, \qquad \mathcal{L}_{bc}(u, g) = \frac{1}{N_u} \sum_{(x_i, y_i) \in \partial D} |u(x^i, y^i, t^i) - g(x^i, y^i, t^i)|^2. \tag{14}$$

PINN total loss function is finally provided by:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \max_{\phi} (\mathcal{L}(u_\theta, \phi, c) + \lambda(\|\mathcal{L}_{u_0}(u_\theta, u_0)\|^2 + \|\mathcal{L}_{bc}(u_\theta, g)\|^2)). \tag{15}$$

5 SOLVED PROBLEMS

To verify the feasibility of the proposed model to solve challenging PDEs in the real world, we solve two equations under different conditions: two-dimensional inviscid nonlinear Burgers and linear advection.

5.1 Burgers

Inviscid 2D nonlinear Burgers equation is given by scalar equation

$$u_t + \left(\frac{u^2}{2}\right)_x + \left(\frac{u^2}{2}\right)_y = 0. \tag{16}$$

Here we solve it with discontinuous initial data (a Riemann type problem) for the following conditions:

- *rarefaction fan initial condition*:

$$u(x, y, 0) = \begin{cases} -1, & x \le 0, \\ 1, & x > 0, \end{cases} \\ x \in [-5.5, 5.5], \quad y \in [-1.5, 1.5], \quad t \in [0, 2.5], \tag{17}$$

- *moving shock condition*

$$u(x, y, 0) = \begin{cases} 1, & x \le 0, \\ 0, & x > 0, \end{cases} \\ x \in [-1, 1], \quad y \in [1, 1], \quad t \in [0, 0.5], \tag{18}$$

- *and the Riemann condition*

$$u(x, y, 0) = \begin{cases} 2, & x < 0.25, y < 0.25, \\ 3, & x > 0.25, y > 0.25, \\ 1, & otherwise, \end{cases} \\ x \in [0, 1], \quad y \in [0, 1], \quad t \in [0, 1/12]. \tag{19}$$

5.2 Linear Advection

We also solve the 2D linear advection equation

$$u_t + u_x + u_y = 0, \quad x \in [0,1], \quad x \in [0,1], \quad t \in [0,1], \tag{20}$$

with smooth periodic initial condition:

$$u(x, y, 0) = \sin^2(\pi x) \sin^2(\pi y). \tag{21}$$

6 RESULTS

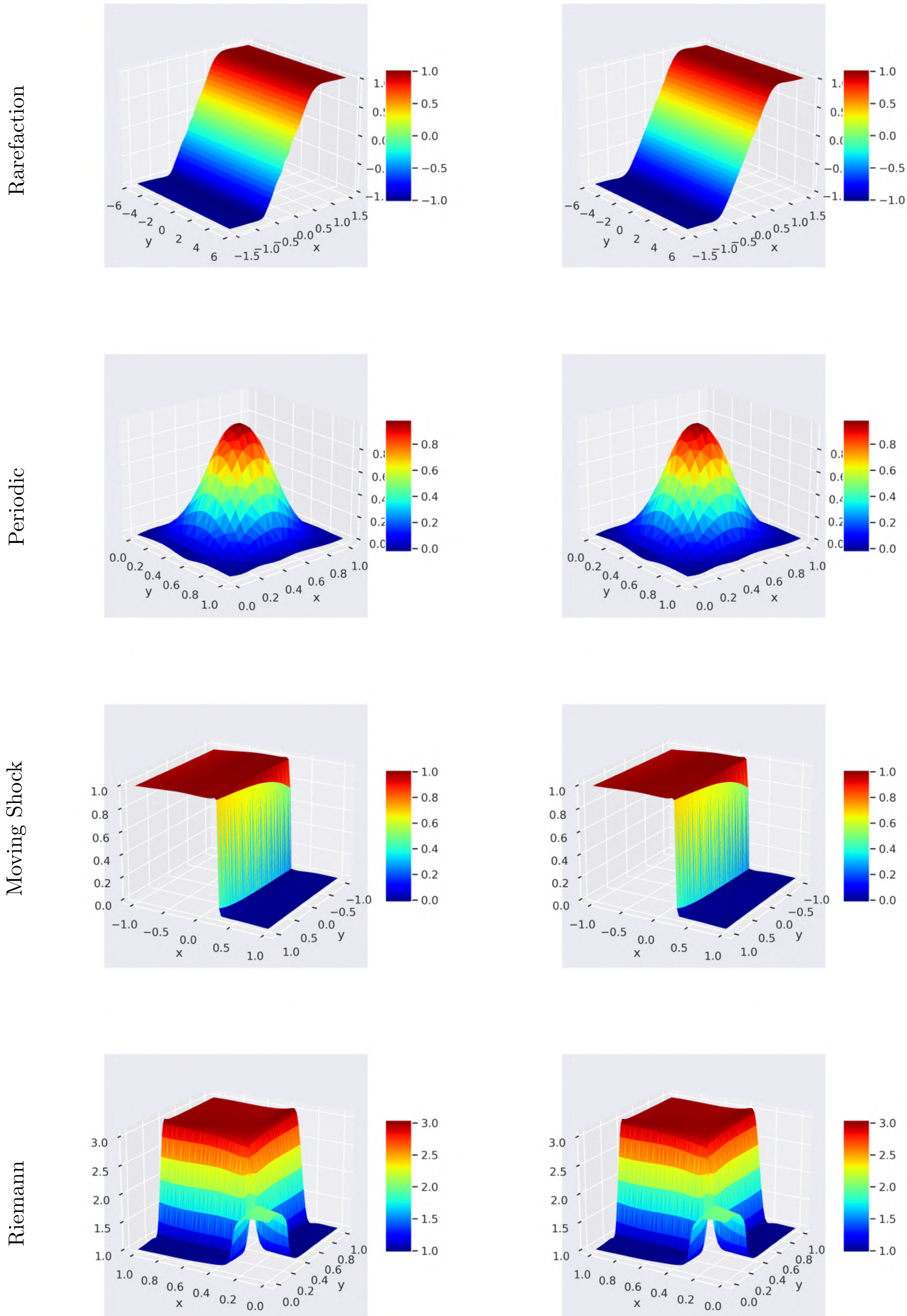
Figure 2 shows the resulting solution for the problems described in Section 5 with a two-dimensional version of the PINN neural network. That approach is similar to what was employed in [Carniello et al., 2022]. For the network model, the following settings are used: $N_{int} = 4 \times 10^6$, 4 layers, 20 neurons, 15000 epochs, $\epsilon = 10^{-3}$.

Despite some expected instability on the discontinuity regions, we observe that the neural network provided acceptable solutions. However, that comes at the price of a high computational cost, which is a consequence of the large number of sample points (N_{int}) and epochs. In this scenario, a natural question arising is whether such relatively precise solutions are worth the cost when well-established numerical solutions achieve comparable performance at a fraction of the computational time [Abreu and Pérez, 2019, Abreu et al., 2021, Galvis et al., 2020].

Figure 3, on the other hand, illustrates the solution for the proposed hybrid model, either with and without the diffusion term. In this case we adopt $N_{int} = 1.6 \times 10^4$, 6 layers/20 neurons (solution), 4 layers/10 neurons (test function), 1000 epochs, and $\epsilon = 10^{-3}$. Here it is important to highlight that the number of collocation points is $\approx 250\times$ smaller than in the basic 2D PINN of Figure 2, whereas the number of epochs is reduced by 1/15.

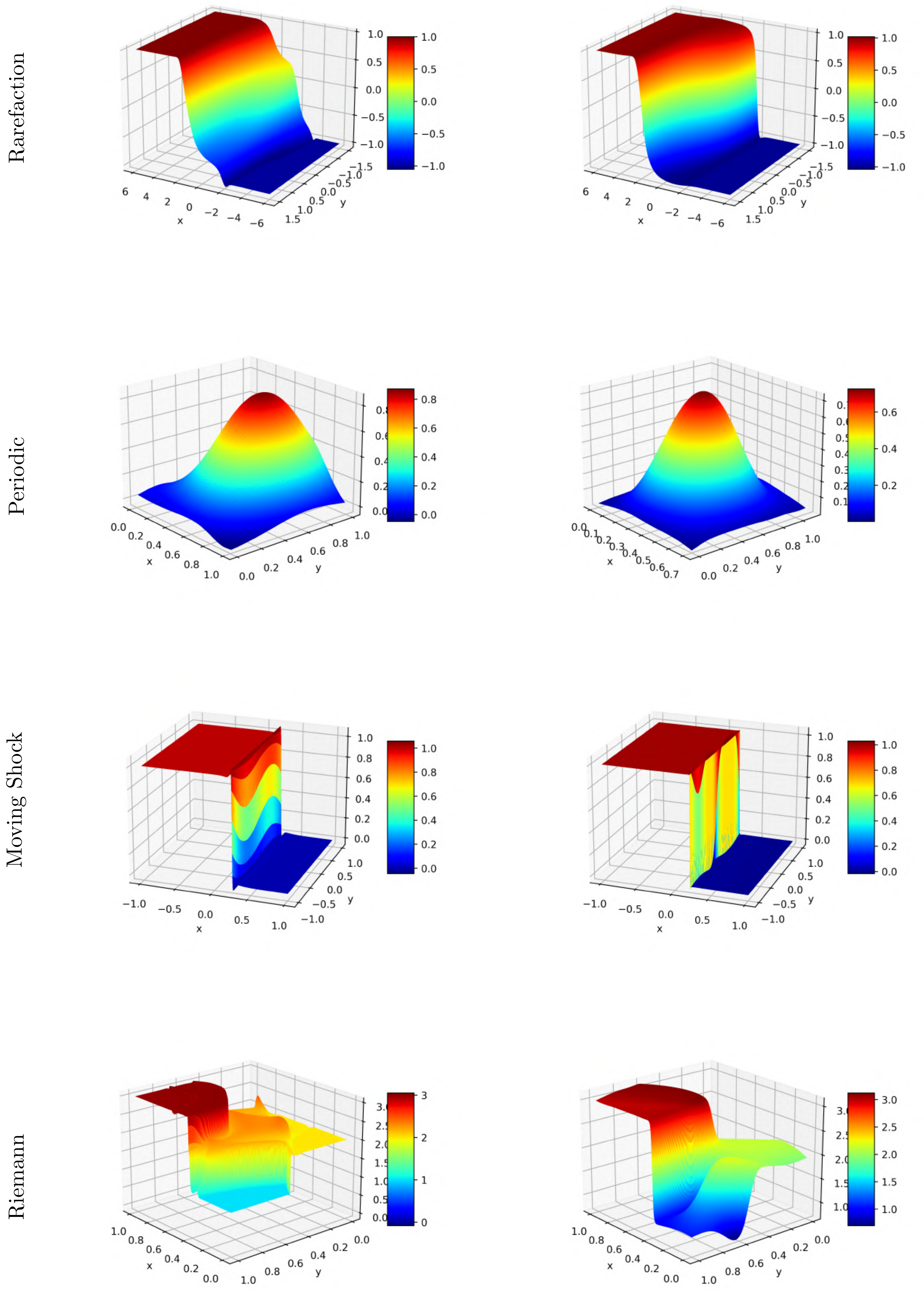
The role of the diffusion term (on the right in Figure 3) was paramount. This term here works as a regularizer of the network and, in this regard, it plays in some sense a similar role to the ensemble modeling in [De Ryck et al., 2024]. Nonetheless, the simple addition of an extra term to the PDE functional adds much less computational burden to the model. In comparison with the basic weak model in [De Ryck et al., 2024] we also notice that the extra entropy term (11) helped to smooth out artifacts on the discontinuity regions. This is particularly evident on the rarefaction and Riemann problems. Distortions like those on the periodic problem was also mitigated. This is an interesting observation as it opens the possibility of further studies on the role of the entropy in the numerical solution. Whilst in [De Ryck et al., 2024] that entropy is fixed, the approach in [Dafermos, 2009] raises the possibility of learning that entropy by obeying some specific rules, which are encompassed in the term 11. That might potentially improve the obtained solution in qualitative terms.

In summary, the qualitative analysis over the obtained numerical solutions present some promising directions for the development of PINN models focused on the weak formulation of the PDE. More precisely, the outcomes opens an avenue for an approach that here we name as "PDE-informed", where intrinsic properties of the PDE, such as its irregularity and entropic nature, are further explored by the deep learning model to provide more accurate and robust numerical solutions. We leverage in this way the power of PINN architectures in the most challenging scenarios found in the real-world applications of PDEs.



Basic PINN   Basic PINN + Diffusion term

Figure 2: Two-dimensional version of classical PINN (PINN2D).



WeakPINN　　　Weak Hybrid + Diffusion term (Proposed)

Figure 3: Proposed Weak Hybrid model.

7 CONCLUSIONS AND NEXT STEPS

7.1 Conclusions

In this work, we proposed and investigated a hybrid model to solve PDEs in weak formulation, following what we call a PDE-informed approach. PDE-Informed solutions are compared with the pure weak formulation of PINN in [De Ryck et al., 2024].

The general structure of the solution was correctly identified by the proposed DNN in different benchmark problems, which included nonlinear inviscid Burgers equation and linear advection with discontinuous initial data. The outcomes are promising considering the complexity of the PDE solved here and the relative straightforwardness of the DNN model. Our results also showed evidences that more general entropy formulation and hybrid approaches constitute a fruitful avenue and deserve more in-depth investigation.

7.2 Next Steps More PDE-Informed

We plan to introduce a more general formulation of entropy as in [Dafermos, 2009]. Going further, we intend to learn the entropy function by a neural network coupled with the PINN model.

The entropy is any $\mathcal{C}^2$ and convex function η, as in [Panov, 1994]. The entropy η and entropy flux q (pair entropy-entropy flux) for Eq. (7) should satisfy

$$\begin{cases} Dq(u) = D\eta(u)DH(u) \\ \partial_t\eta(u) + \partial_x q(u) + \partial_y q(u) \leq 0, \end{cases} \tag{22}$$

in the sense of distribution in $\Omega \times [0, T]$. Second condition represents the dissipation of entropy. If necessary we add diffusive terms to enforce uniqueness.

Moreover, admissible solutions should also minimize the rate of change of the total entropy:

$$\mathcal{H} = \frac{d}{dt}\int_{\Omega} \eta(u(x, y, t))dxdy. \tag{23}$$

We expect in this way to develop a fully-learnable model to accomplish with the entropy formulation of the PDE. This is also expected to provide more accurate and robust numerical solutions as the learned entropy may be more favorable to neural network training than a fixed form as the Kruzhkov entropy employed here and in [De Ryck et al., 2024].

REFERENCES

[Abreu and Florindo, 2021] Abreu, E. and Florindo, J. B. (2021). A study on a feedforward neural network to solve partial differential equations in hyperbolic-transport problems. In *Computational Science–ICCS 2021: 21st International Conference, Krakow, Poland, June 16–18, 2021, Proceedings, Part II 21*, pages 398–411. Springer.

[Abreu et al., 2021] Abreu, E., Matos, V., Pérez, J., and Rodríguez-Bermúdez, P. (2021). A class of Lagrangian-Eulerian shock-capturing schemes for first-order hyperbolic problems with forcing terms. *Journal of Scientific Computing, https://doi.org/10.1007/s10915-020-01392-w*, 86(14).

[Abreu and Pérez, 2019] Abreu, E. and Pérez, J. (2019). A fast, robust, and simple Lagrangian-Eulerian solver for balance laws and applications. *Computers & Mathematics with Applications*, 77(9):2310 – 2336.

[Alibaud et al., 2020] Alibaud, N., Andreianov, B., and Ouédraogo, A. (2020). Nonlocal dissipation measure and $\mathcal{L}^1$ kinetic theory for fractional conservation laws. *Communications in Partial Differential Equations*, 45(9):1213–1251.

[Cai et al., 2021] Cai, S., Wang, Z., Wang, S., Perdikaris, P., and Karniadakis, G. E. (2021). Physics-informed neural networks for heat transfer problems. *Journal of Heat Transfer*, 143(6):060801.

[Carniello et al., 2022] Carniello, R. A. F., Florindo, J. B., and Abreu, E. (2022). A pinn computational study for a scalar 2d inviscid burgers model with riemann data. *8th European Congress on Computational Methods in Applied Sciences and Engineering.*

[Dafermos, 2009] Dafermos, C. M. (2009). A variational approach to the riemann problem for hyperbolic conservation laws. *Discrete Contin. Dyn. Syst*, 23(1-2):185–195.

[De Ryck et al., 2024] De Ryck, T., Mishra, S., and Molinaro, R. (2024). wpinns: Weak physics informed neural networks for approximating entropy solutions of hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 62(2):811–841.

[Galvis et al., 2020] Galvis, J., Abreu, E., Díaz, C., and Pérez, J. (2020). On the conservation properties in multiple scale coupling and simulation for darcy flow with hyperbolic-transport in complex flows. *Multiscale Modeling and Simulation*, 18(4):1375–1408.

[Mao et al., 2020] Mao, Z., Jagtap, A. D., and Karniadakis, G. E. (2020). Physics-informed neural networks for high-speed flows. *Computer Methods in Applied Mechanics and Engineering*, 360:112789.

[Misyris et al., 2020] Misyris, G. S., Venzke, A., and Chatzivasileiadis, S. (2020). Physics-informed neural networks for power systems. In *2020 IEEE power & energy society general meeting (PESGM)*, pages 1–5. IEEE.

[Panov, 1994] Panov, E. Y. (1994). Uniqueness of the solution of the Cauchy problem for a first order quasilinear equation with one admissible striclty convex entropy . *Math Notes*, 55():517–525.

[Patel et al., 2022] Patel, R. G., Manickam, I., Trask, N. A., Wood, M. A., Lee, M., Tomas, I., and Cyr, E. C. (2022). Thermodynamically consistent physics-informed neural networks for hyperbolic systems. *Journal of Computational Physics*, 449:110754.

[Raissi et al., 2019] Raissi, M., Perdikaris, P., and Karniadakis, G. (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.

[Rodriguez-Torrado et al., 2022] Rodriguez-Torrado, R., Ruiz, P., Cueto-Felgueroso, L., Green, M. C., Friesen, T., Matringe, S., and Togelius, J. (2022). Physics-informed attention-based neural network for hyperbolic partial differential equations: application to the buckley–leverett problem. *Scientific reports*, 12(1):7557.