The 8th European Congress on Computational Methods in Applied Sciences and Engineering
ECCOMAS Congress 2022
5-–9 June 2022, Oslo, Norway

A PINN COMPUTATIONAL STUDY FOR A SCALAR 2D INVISCID BURGERS MODEL WITH RIEMANN DATA

RAFAEL CARNIELLO[1], JOAO B. FLORINDO[1] AND EDUARDO ABREU[1]

[1] Institute of Mathematics, Statistics and Scientific Computing (IMECC)
University of Campinas
Rua Sérgio Buarque de Holanda, 651, Cidade Universitária "Zeferino Vaz" - Distr. Barão Geraldo, CEP 13083-859, Campinas, SP, Brasil
http://www.ime.unicamp.br
e-mail: r204844@dac.unicamp.br, {eabreu,florindo}@unicamp.br

Key words: Physics Informed Neural Networks (PINN), Scalar 2D inviscid Burgers model, Riemann data, Computing Methods for hyperbolic-transport models.

Abstract. In this work, we present an application of modern deep learning methodologies to the numerical solution of two dimensional hyperbolic partial differential equations in transport models. More specifically, we employ a supervised deep neural network that takes into account of initial-boundary value problems for a scalar, 2D inviscid Burgers model including the case with Riemann data, whose solutions develop discontinuity, containing both shock wave and rarefaction wave. We also apply the proposed PINN approach to the linear advection equation with periodic sinusoidal initial condition. Our results suggest that a relatively simple deep learning model was capable of achieving promising results in the linear advection and inviscid Burgers equation with rarefaction, providing numerical evidence of good approximation of weak-entropy solutions to the case of nonlinear 2D inviscid Burgers model. For the Riemann problems, the neural network performed better when rarefaction wave is predominant. The premises underlying these preliminary results as an integrated physics-informed deep learning approach are promising. However, there are hints of evidence suggesting specific fine tuning on the PINN methodology for solving hyperbolic-transport problems in the presence of shock formation in the solutions.

1 INTRODUCTION

The recent popularity of *physics-informed* techniques [12, 25, 31, 32] as a approximation method to numerically solve certain classes of models of partial differential equations (PDEs) (e.g., [7, 11, 16, 22, 29]) represents an astonishing advance in scientific computing [28, 20], especially considering the relatively simple neural networks employed for this purpose [4].

To the best of our knowledge, an important and distinct step in the numerical treatment of models involving PDEs was given in [25] (2019) for forward problems (directly solving a differential equation with boundary and/or initial condition) and inverse problems (approximate a parametric PDE based on an experimental data set), by combining PDE modeling and data-driven approach; see also [7, 11]. Since then, only a few more approaches have been developed with rigorous mathematical and numerical analysis; see [7, 26, 27] and the references cited therein.

For the specific purpose of the current work, we consider essential to highlight some recent advances from which we gain additional insights, namely, the proof of the consistency for linear second-order elliptic and parabolic PDEs [5]; extending this strategy, the generalized proof of consistency for linear equations [6], including the linear advection equation, which is a linear hyperbolic PDE; and the estimate of error bounds for direct and inverse problems [7].

The hyperbolic-transport PDE models treated in this work are related to the scalar 2D conservation law $u_t + H_x(u) + H_y(u) = 0$ and the corresponding augmented hyperbolic-parabolic equation $u_t^\epsilon + H_x(u^\epsilon) + H_y(u^\epsilon) = \epsilon(u_{xx}^\epsilon + u_{yy}^\epsilon)$, $\epsilon > 0$. To put in perspective the complexity of the current study, we mention the classical works [18, 19] regarding rigorous mathematical and numerical analysis results via classical and well-established techniques for scalar and non-linear 2D conservation laws as discussed in this work. In [19], the authors presented a study via viscosity approximations to multidimensional scalar conservation laws related to the model problem $u_t^\epsilon + H_x(u^\epsilon) + H_y(u^\epsilon) = \epsilon(u_{xx}^\epsilon + u_{yy}^\epsilon)$, $\epsilon > 0$, in the context of the spectral viscosity method. The authors proved that the spectral viscosity is large enough to enforce the correct amount of entropy dissipation satisfying the entropy condition, that is, converges to the unique entropy solution. In [18], the authors introduced a new framework for studying two-dimensional conservation laws $u_t + H_x(u) + H_y(u) = 0$ via compensated compactness arguments. Moreover, the authors demonstrated convergence results in the context of vanishing viscosity, kinetic Bhatnagar-Gross-Krook (BGK) schemes and finite volume approximations. Roughly speaking in this above mentioned context, such authors proved that if u^ϵ is a family of uniformly bounded approximate solutions of such equations with H^{-1} - compact entropy production and with uniform time regularity, then a subsequence of u^ϵ strongly converges to a weak solution u linked to the purely hyperbolic 2D conservation laws $u_t + H_x(u) + H_y(u) = 0$. Therefore, independently of the approach chosen, the literature of approximation methods for hyperbolic problems primarily concern in the fundamental issues of conservation and the ability of the scheme to compute the correct entropy solution to the underlying nonlinear conservation laws (see, e.g., [14, 21, 23, 24, 30]), when computing shock fronts, in transporting discontinuities at the correct speed, and in giving the correct shape of continuous waves. This is of utmost importance among computational practitioners, theoretical mathematicians and numerical analysts.

Despite the recent success of learning-based approaches to solve PDEs in relatively "well-behaved" configurations, we still have points in these methodologies and applications that deserve more profound discussion, both in theoretical and practical terms. Furthermore, with respect to learning-based schemes to solve PDEs in physical models, we have seen little discussion about such procedures on more challenging problems such as multidimensional equations, fractional conservation laws [8], compressible turbulence and Navier-Stokes equations [13], stochastic conservation laws [17] and simulation for *Darcy flow* with hyperbolic-transport in complex flows with discontinuous coefficients [15]. We also mention the very recent review paper [11], which discusses machine learning for fluid mechanics, but highlighting that such approach could augment existing efforts for the study, modeling and control of fluid mechanics, keeping in mind the importance of honoring engineering principles and the governing equations [2, 3], mathematical [8, 30] and physical foundations [13, 14] driven by unprecedented volumes of data from experiments and advanced simulations at multiple spatiotemporal scales.

Thus, the motivation of our work comes also from the references mentioned above, once we have observed that until moment, there are no PINN results based on rigorous mathematical and numerical analysis for scalar 2D conservation laws and related multidimensional problems in general. Here we investigate a simple feed-forward PINN architecture, based on the physics-informed model proposed in [25], and applied to complex problems involving PDEs in transport models. More specifically, we extend the model initially investigated in [4] for one dimensional hyperbolic transport problem, but now for two dimensions. This represents a class of harder problems, since wave interaction may occur. For that, we analyze the numerical solutions of four initial-value problems: one periodic sinusoidal solution for the two dimensional linear advection equation, and three problems on the two dimensional nonlinear inviscid Burgers equation involving shock and rarefaction waves for distinct initial conditions.

The neural network consists of 4 hidden layers with tanh activation and geared towards minimizing the approximation error both for the initial values and for values of the PDE functional, calculated by automatic differentiation. By optimizing hyper-parameters we managed to obtain a significant reduction of the error for the two fundamental models under consideration, namely, the linear advection equation with sinusoidal periodic initial condition and inviscid Burgers equation for representative configurations of Riemann data, leading for solutions with *rarefaction wave in the presence sonic line* and the interactions of elementary waves (*shock waves and centered rarefaction waves*), respectively. Even though for the oblique Riemann datum case the results were qualitatively unsatisfactory, it seems we reach the limits of the approach at some extent, suggesting the development of an *enhanced PDE informed method*, but also keeping in mind the mathematical complexity of weak-entropic solutions for nonlinear conservation laws as briefly highlighted in the motivation as above. It also strongly suggests more in-depth studies on deep learning models that account for the underlying equations. They seem to be a quite promising line to be explored for challenging problems arising in physics, engineering, and many other areas.

Three major contributions are presented in this study:

- We apply a straightforward fully connected neural network to complex problems in transport models achieving promising results;
- We present an extension of [4] to two dimensions, increasing the challenge of the problem in this way, but the encouraging performance of the model persists;
- We verify the effect of hyperparameters of the model in these problems, opening space for further discussion on optimized models especially designed for the challenging conditions studied here.

The paper is organized as follows. In Section 2, we introduce some basic aspects of hyperbolic partial differential equations in transport problems, along with a benchmark numerical scheme for qualitatively comparing the results yielded by the neural network. The mathematical setup and proposed methodology considered in this work are presented in Section 3, aiming for stability and entropy of the feedforward neural network approximations. In Section 4, we present the numerical results obtained by the approach, evaluating its accuracy. Finally, in the last Section 5, we present some concluding remarks.

2 HYPERBOLIC PROBLEMS IN TRANSPORT MODELS

Hyperbolic partial differential equations in transport models describe a wide range of wave-propagation and transport phenomena arising from scientific and industrial engineering areas. This is a fundamental research that is in active progress since it involves complex multiphysics and advanced simulations due to a lack of general mathematical theory for closed-analytical solutions. For instance, see the noteworthy book by C. M. Dafermos [14] devoted to the mathematical theory of hyperbolic conservation and balance laws and their generic relations to continuum physics with a large bibliography list as well as some very recent work references cited therein related to recent advances covering distinct aspects, theoretical [8], numerical [33] and applications [13].

In addition, just to name some very recent works, see some interesting results covering distinct aspects, such as theoretical [8] (uniqueness for scalar conservation laws with nonlocal and nonlinear diffusion terms) and [21] (non-uniqueness of dissipative Euler flows), and well-posedness [30] for multi-dimensional inviscid Burgers equation with unbounded initial data, numerical analysis [33] and numerical computing for stochastic conservation laws [17] and applications [13] (Euler equations for barotropic fluids) and the references cited therein for complimentary discussion to highlight difficulties on the unavailability of universal results for finding the global explicit solution for Cauchy problems to the relevant class of hyperbolic-transport problems involving scalar and systems of conservation laws. Many problems in engineering and physics are modeled by hyperbolic systems and scalar nonlinear equations [14]. As examples for these equations, just to name a few of relevant situations, we can mention the Euler equations of compressible gas dynamics, the Shallow water equations of hydrology, the Magnetohydrodynamics equations of plasma physics and the Buckley-Leverett scalar equation in petroleum engineering [15].

For the sake of simplicity, we consider the scalar 2D Cauchy problem

$$\frac{\partial u}{\partial t} + \frac{\partial H(u)}{\partial x} + \frac{\partial H(u)}{\partial y} = 0, \quad (x,y) \in \mathbb{R}^2, \quad t > 0, \quad u(x,y,0) = u_0(x,y), \tag{1}$$

where $H \in C^2(\Omega)$, $H : \Omega \to \mathbb{R}$, $u_0(x,y) \in L^\infty(\mathbb{R}^2)$ and $u = u(x,y,t) : \mathbb{R}^2 \times \mathbb{R}^+ \longrightarrow \Omega \subset \mathbb{R}$. By using arguments viscosity approximations [19] or compensated compactness [18], one can conclude that solutions of (1) are limits of solutions $u^\epsilon(x,y,t) \to u(x,y,t)$, where $u(x,y,t)$ is given by (1) and $u^\epsilon(x,y,t)$ is given by the augmented parabolic equation $(x,y) \in \mathbb{R}^2$,

$$\frac{\partial u^\epsilon}{\partial t} + \frac{\partial H(u^\epsilon)}{\partial x} + \frac{\partial H(u^\epsilon)}{\partial y} = \epsilon \left(\frac{\partial^2 u^\epsilon}{\partial x^2} + \frac{\partial^2 u^\epsilon}{\partial y^2} \right), \quad t > 0, \quad u^\epsilon(x,y,0) = u_0^\epsilon(x,y), \tag{2}$$

with $\epsilon > 0$ and the same initial data as in (1).

In addition, from the well-established literature in this area of conservation laws (see, e. g., [14, 17, 18, 19, 23, 24, 30, 33, 34] and the references cited therein), we remember some distinct and basic facts in the classical theory for nonlinear hyperbolic conservation laws of type (1): (a) The existence of possibly infinitely many weak solutions; (b) Despite of a given smooth initial data, the evolution in time of the corresponding solution might lead to formation of shock discontinuities which requires weak-entropy solutions; (c) The notion of entropy condition for uniqueness; The notion of a viscosity limit solution as is in Eq. (2) is closely related to the notion of an entropy solution to model problem (1).

For concreteness, in this work we will specifically consider two hyperbolic-transport models related to the prototype models in Eq. (1) and in Eq. (2), namely, the 2D linear advection equation, where the flow function is $H(u) = u$, and the 2D inviscid Burgers scalar equation, where the flow function is $H(u) = u^2/2$ in Eq. (1). A nonlinear phenomenon that arises with the inviscid Burgers equation, even for smooth initial data, is the formation of shock, which is a discontinuity that may appear after finite time. Thus, these two models, the linear advection equation and inviscid Burgers equation, are suitable and effective fundamental problems for testing new approximation algorithms to the above mentioned properties as presented and discussed in the current work. In this regard, a typical flux function $H(u)$ associated to fundamental prototype models (1) and (2) depends on the application under consideration, for instance, such as modeling flow in fluid mechanics [3].

From the above discussion of basic facts of conservation laws it is noteworthy that in practice the calibration of function $H(u)$ for the PINN approach can be difficult to achieve, due to unknown parameters and data assimilation. We intend to design a unified approach that combines both PDE modeling and fine tuning machine learning techniques aiming as a first step to an effective tool for advanced simulations related to hyperbolic problems in transport models such as in (1) and (2). For evaluating qualitatively the results presented in this work, we define a reference benchmark scheme [2]. That is semi-discrete Lagrangian-Eulerian scheme based on the so-called no-flow curves applied for one-dimensional (1D) as well as for two-dimensional (2D) scalar hyperbolic conservation laws. The scheme satisfies a maximum principle property and a Kruzhkov entropy condition, handling discontinuous solutions with low numerical dissipation quite well and shows a very good resolution of rarefaction waves with no spurious glitch effect in the vicinity of the sonic points.

3 PROPOSED APPROACH

The proposed approach followed the same structure used in [4]. In that work the capability of approximating the solutions for one dimensional hyperbolic transport models was competitive compared to the classical numerical schemes. For that reason we extend it for the two dimensional problems. So we start with a partial differential equation

$$u_t + \mathcal{D}(u) = 0, \qquad (x, y) \in \Omega, t \in [0, T], \tag{3}$$

where $\mathcal{D}(\cdot)$ is a differential operator and $u(x, y, t)$ is the desired solution. Let $\mathcal{N}^m : \mathbb{R}^{d_0} \to \mathbb{R}^{d_m}$ be a feed-forward neural network model built to represent the solution of the problem with m layers and d_k neurons in the k^{th} layer, where the input layer has d_0 and the output d_m neurons. For any $1 \leq k \leq m$, the k^{th} layer output is defined as

$$C_k(z_k) = W_k z_k + b_k, \quad \text{for} \quad W_k \in \mathbb{R}^{d_{k+1} \times d_k}, z_k \in \mathbb{R}^{d_k}, b_k \in \mathbb{R}^{d_{k+1}}, \tag{4}$$

where z_k, W_k and b_k are the input, weights and biases, respectively. Let $\Theta = \{W_k, b_k\} \in \mathcal{V}$ be the collection of all trainable parameters, with $\mathcal{V}$ being the parameter space, then, the neural network output can be written as

$$u_\Theta(z_0) = C_m \circ \sigma \circ C_{m-1} ... \circ \sigma \circ C_2 \circ \sigma \circ C_1(z_0). \tag{5}$$

Here, $\circ$ refers to the composition of functions and σ is a non-linear activation function.

For the optimization of the neural network we should define f as the left hand side of each PDE, i.e.,

$$f := u_t + \mathcal{D}(u), \tag{6}$$

Two quadratic loss functions are defined over f, u and the initial condition u_0:

$$\begin{aligned} \mathcal{L}_f(u) &= \frac{1}{N_f}\sum_{i=1}^{N_f} |f(x_f^i, y_f^i, t_f^i,)|^2, \\ \mathcal{L}_{u_0}(u, u_0) &= \frac{1}{N_u}\sum_{i=1}^{N_u} |u(x_u^i, y_u^i, t_u^i) - u_0^i|^2, \end{aligned} \tag{7}$$

where $\{x_f^i, y_f^i, t_f^i\}_{i=1}^{N_f}$ correspond to collocation points over f, whereas $\{x_u^i, y_u^i, t_u^i, u_0^i\}_{i=1}^{N_u}$ correspond to the initial values at pre-defined points.

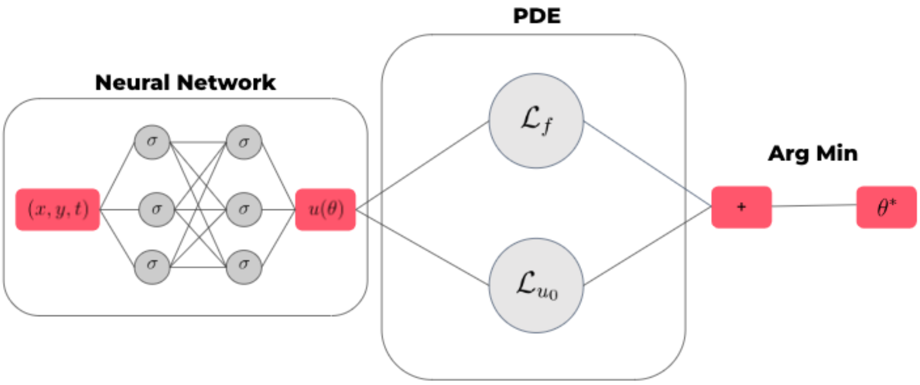

Figure 1: Physics-informed deep learning approach illustration.

Finally, the solution $u(x, y, t)$ is approximated by minimizing the sum of both objective functions at the same time, i.e.,

$$u(x, y, t) \approx \arg\min[\mathcal{L}_f(u) + \mathcal{L}_{u_0}(u, u_0)]. \tag{8}$$

We can view a schematic illustration representing the architecture employed on Figure 1.

The PINN architecture proposed here is composed of 4 hidden layers, each one with 20 neurons and a hyperbolic tangent used as activation function. The loss function is defined in the strong form, following the same approach originally used in [25]. In contrast, here we do not have an explicit boundary condition and the neural network is optimized only over the initial conditions, i.e., Cauchy problem. Here the optimization is performed by the Adam algorithm, instead of L-BFGS-B initially used by [25]. As for the optimizer hyperparameters we use the default values of learning rate 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$. These were empirically found to be suitable values in our experiments. For the distribution of collocation points we employed Latin hypercube sampling, following the previous procedure adopted in [25]. In the tests, the number of collocation points for the initial values was fixed at $N_u = 256$.

Regarding differences and similarities of our proposal with respect to the original PINN model in [25], we have the following ones:

- Differences:
  - We use Adam optimizer instead of the original BFGS algorithm;
  - Reduced number of layers and neurons. This reduced the computational burden whilst not impairing the precision;
  - We also employ $N_f = 4 \times 10^6$, which is significantly larger than the original values. This was motivated both by the extra dimension of the problem as well as by the identification that increasing N_f positively impacts the precision of the method;
  - We have $N_u = 256$ to match the granularity of SDLE method;
  - Our loss function does not include a boundary term considering the nature of our problems.
- Similarities:
  - A fully-connected network is also used here;
  - Distribution of N_u and N_f points is the same used in [25];
  - We also use the strong form of the loss function.

The "physics-informed" aspect of the method consists in taking into account the original equation governing the physical behavior by explicitly including the PDE functional and initial and/or boundary conditions in the objective function. The derivatives take advantage of automatic differentiation widely used in the optimization of classical neural networks.

We focus on the following problems: On the two dimensional linear advection equation

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0, \qquad (x, y, t) \in [0,1]^2 \times [0,1], \tag{9}$$

with periodic sinusoidal initial condition

$$u(x, y, 0) = \sin^2(\pi x)\sin^2(\pi y), \tag{10}$$

and exact solution

$$u(x, y, t) = \sin^2(\pi(x - t))\sin^2(\pi(y - t)). \tag{11}$$

On the inviscid nonlinear two dimensional inviscid Burgers equation

$$u_t + \left(\frac{u^2}{2}\right)_x + \left(\frac{u^2}{2}\right)_y = 0, \tag{12}$$

with Riemann datum, given by

$$u(x,y,0) = \begin{cases} 1, & x > 0, \\ -1, & x < 0, \end{cases} \tag{13}$$

wherein the solution is a rarefaction wave evolving in the domain $(x,y,t) \in [-6,6] \times [-1.5,1.5] \times [0,2.5]$, the jump is defined in such a way that the change in the sign of the wave speed occurs when $x = 0$; the Riemann problem initial condition

$$u(x,y,0) = \begin{cases} 2, & x < 0.25, y < 0.25, \\ 3, & x > 0.25, y > 0.25, \\ 1, & \text{otherwise}, \end{cases} \tag{14}$$

on the domain $(x,y,t) \in [0,1]^2 \times [0,1/12]$, and the oblique Riemann initial condition

$$u(x,y,0) = \begin{cases} -1.0, & x > 0.5, y > 0.5, \\ -0.2, & x < 0.5, y > 0.5, \\ +0.5, & x < 0.5, y < 0.5, \\ +0.8 & x > 0.5, y < 0.5, \end{cases} \tag{15}$$

on the domain $(x,y,t) \in [0,1]^2 \times [0,0.5]$. On each Riemann problem the predominance of rarefaction or shock waves is given by the largest jumps on the initial condition. That means that for the initial condition given by Equation (14) rarefaction is predominant. In contrast, for the oblique case, given by Equation (15), shock is the predominant phenomenon.

In our problem we have

$$\mathcal{D}(u) = u_x + u_y, \tag{16}$$

in the linear advection, and

$$\mathcal{D}(u) = \left(\frac{u^2}{2}\right)_x + \left(\frac{u^2}{2}\right)_y, \tag{17}$$

in the inviscid Burgers equation. Here we also have an important novelty which is the introduction of a derivative (w.r.t. x, y) in $\mathcal{D}(u)$, which was not present in [25].

The function f is mainly responsible for capturing the physical structure (i.e., selecting the qualitatively correct entropy solution) of the problem and inputting that structure as a primary element of the machine learning problem. Nevertheless, here to ensure the correct entropy solution, we add a small diffusion term to f $(0.0025(u_{xx} + u_{yy}))$ and $(0.00025(u_{xx} + u_{yy}))$ for better comprehension of shock formation on hyperbolic-parabolic equations, but in view of the modeling problems (1) and (2). It is crucial to mention at this point that the numerical approximation of entropy solutions (with respect to the neural network) to hyperbolic-transport problems also requires the notion of entropy-satisfying weak solution.

4 RESULTS AND DISCUSSION

In the following we present the solutions for the investigated hyperbolic problems obtained by the neural network model. We would like to point out that no comparison is made here with the more general approach in [3] as the main focus of the current work is a PINN computational study for a scalar 2D inviscid Burgers model with some specific Riemann data. We only use the method described in [3] to produce the correct weak-entropic pertinent solutions to the specific 2D scalar inviscid Burgers equation along with the initial data under investigation.

A thorough fine tuning process was explored for the models, achieving an optimal architecture with 4 hidden layers, 20 neurons in each one and $N_f = 4 \cdot 10^6$ collocation points for the functional. To illustrate the effect of the hyperparameters, we show the results for inviscid Burgers equation with a rarefaction wave in the solution as presented in the plots from Figure 3 through Figure 5, but the same happened for all cases. We prioritized maximizing the number of training points enforcing the *PDE functional*, inspired by recent results in [7]. The problem of determining the optimal relation between the error and the values of these parameters is not trivial.

For the two dimensional linear advection case (see Figure 2) 15000 epochs were sufficient to obtain promising results. In both models, inviscid and augmented one, the solutions are similar. We also highlight that the models did not show spurious effects such as excessive numerical diffusion as is the typical situation for the well-known Lax–Friedrichs scheme.

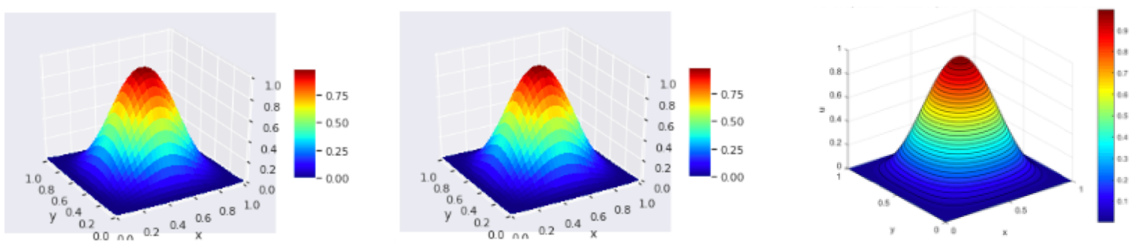

Figure 2: Neural network solution for two-dimensional linear advection equation with periodic initial condition. The architecture employed was composed of 4 hidden layers, 20 neurons each, $N_f = 4 \cdot 10^6$ and 15.000 epochs. Left: no viscosity. Middle: $2.5 \cdot 10^{-3}$ viscosity. Right: SDLE solution for two-dimensional linear advection equation with periodic initial condition.

In Figure 5 the neural network model yields qualitatively correct approximations as expected from the classical literature. Even though inviscid Burgers equation is nonlinear, for these cases, only 5000 epochs were necessary for a good result, in contrast with the linear advection problem. A very important aspect to analyze in problems with a rarefaction wave is the presence of a *"sonic line"* in the sense of a sonic point (glitch effect) as discussed, for instance, in the work [9]. That is a nonphysical effect generated by the change in the sign of the wave speed, and it is present in some numerical schemes, such as Godunov's scheme; for more details see [9] and Section 3.2.3 in [33]. We see that neural networks do not present such glitch effect.

On the Riemann problem in two dimensions, both shock and rarefaction waves interact during propagation, complexifying the physics to be captured by the PINN model. It is important to state that, in this case, the rarefaction wave is *predominant* over shock. For both approaches, inviscid ($\epsilon = 0$) and augmented ($\epsilon > 0$), we needed to increase the number of epochs from

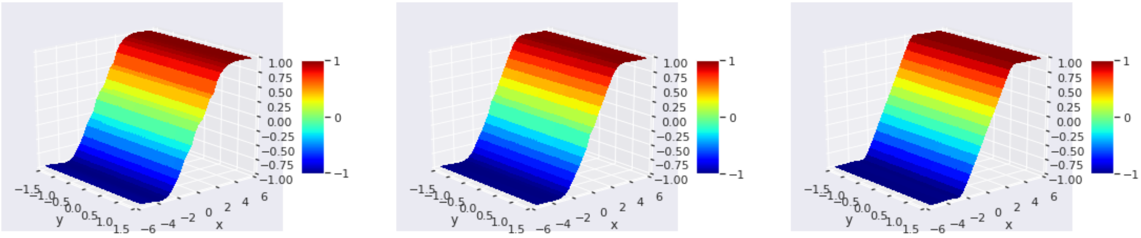

Figure 3: Neural network solution for two-dimensional inviscid Burgers equation with rarefaction wave, with $N_f = 10^4$ and 5.000 epochs. Left: 2 hidden layers, 20 neurons each. Middle: 2 hidden layers, 40 neurons each. Right: 4 hidden layers, 20 neurons each.

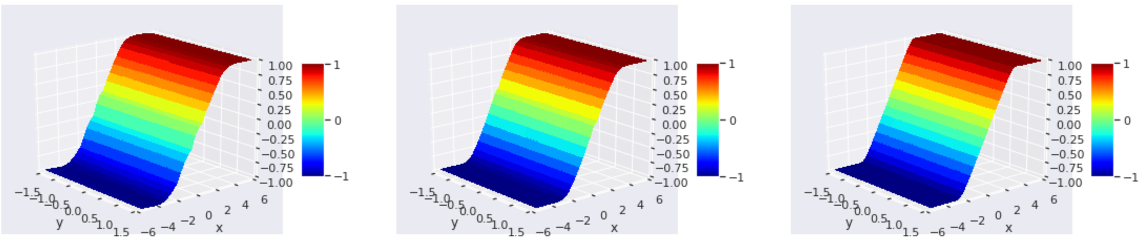

Figure 4: Neural network solution for two-dimensional inviscid Burgers equation with rarefaction wave, with $N_f = 10^4$, $2.5 \cdot 10^{-3}$ viscosity and 5.000 epochs. Left: 2 hidden layers, 20 neurons each. Middle: 2 hidden layers, 40 neurons each. Right: 4 hidden layers, 20 neurons each.

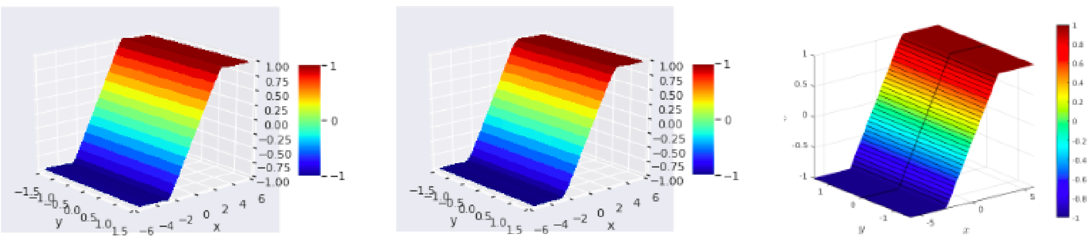

Figure 5: Neural network solution for two-dimensional inviscid Burgers equation with rarefaction wave. The architecture employed was composed of 4 hidden layers, 20 neurons each, $N_f = 4 \cdot 10^6$ and 5.000 epochs. Left: No viscosity. Middle: $2.5 \cdot 10^{-3}$ viscosity. Right: SDLE solution for two-dimensional inviscid Burgers equation with rarefaction fan.

15000 to 30000. Nevertheless, the neural network presents difficulties in approximating the part with shock in as shown in Figure 6. Introducing the diffusive term on the PDE, at Figure 7, reduced the height of the shock level closer to 2, which is the expected correct solution from the literature, but, as a consequence, it "smooths" the discontinuity, also as expected.

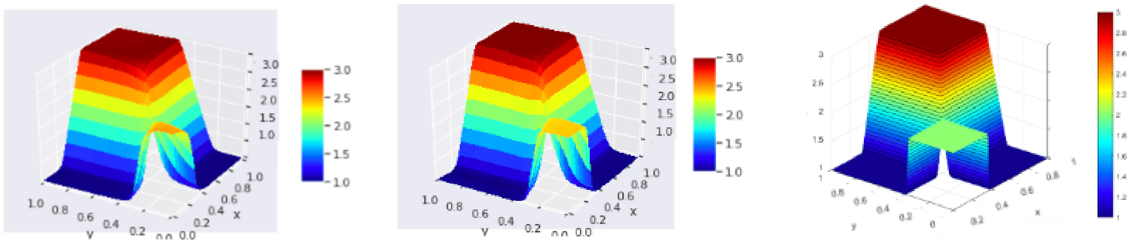

Figure 6: Neural network solution for two-dimensional inviscid Burgers equation with Riemann condition. The architecture employed was composed of 4 hidden layers, 20 neurons each, $N_f = 4 \cdot 10^6$. Left: 15.000 epochs. Middle: 30.000 epochs. Right: SDLE solution for two-dimensional inviscid Burgers equation with Riemann condition.

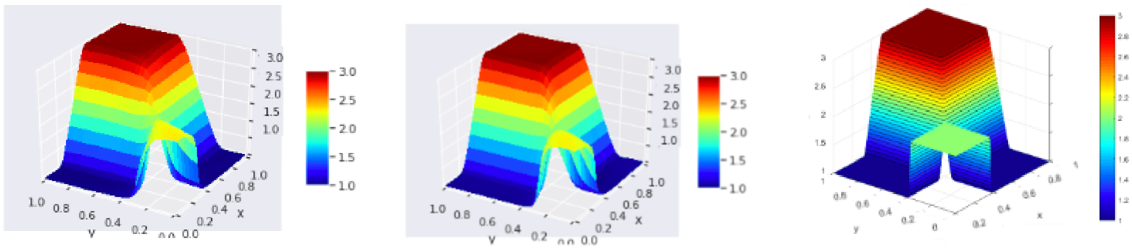

Figure 7: Neural network solution for two-dimensional inviscid Burgers equation equation with Riemann condition. The architecture employed was composed of 4 hidden layers, 20 neurons each, $N_f = 4 \cdot 10^6$, $2.5 \cdot 10^{-3}$ viscosity. Left: 15.000 epochs. Middle: 30.000 epochs. Right: SDLE solution for two-dimensional inviscid Burgers equation with Riemann condition.

Another case of the Riemann problems is the *oblique-type initial condition*, also containing shock and rarefaction waves, but shock being the *predominant nonlinear phenomenon*. This problem leverages the difficulties of the model already shown for the previous Riemann problems. We investigated for 15000 and 30000 epochs, and varied the ϵ between $2.5 \cdot 10^{-4}$ and $2.5 \cdot 10^{-3}$. In particular, the augmented results (i.e., $\epsilon > 0$) loosely resembles the desired solution, at Figure 9. As the viscosity decreases we clearly see the recover of the inviscid solution, that is not entropic for the proposed approach, see Figure 8.

In general, we observe from the results that the physics-informed model investigated here and especially tuned for such challenging conditions provided promising results, especially on the linear advection and inviscid Burgers equation with rarefaction wave. Further studies are necessary for the Riemann problems, mainly when shock is predominant.

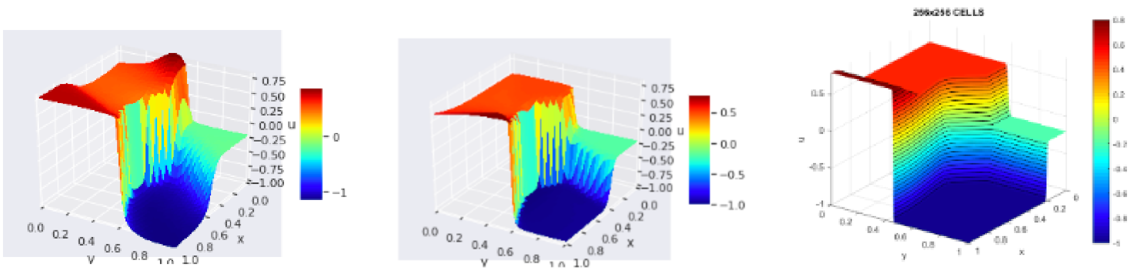

Figure 8: Neural network solution for two-dimensional inviscid Burgers equation with oblique Riemann condition. The architecture employed was composed of 4 hidden layers, 20 neurons each, $N_f = 4 \cdot 10^6$. Left: 15.000 epochs. Middle: 30.000 epochs. Right: SDLE solution for two-dimensional inviscid Burgers equation with oblique Riemann condition.

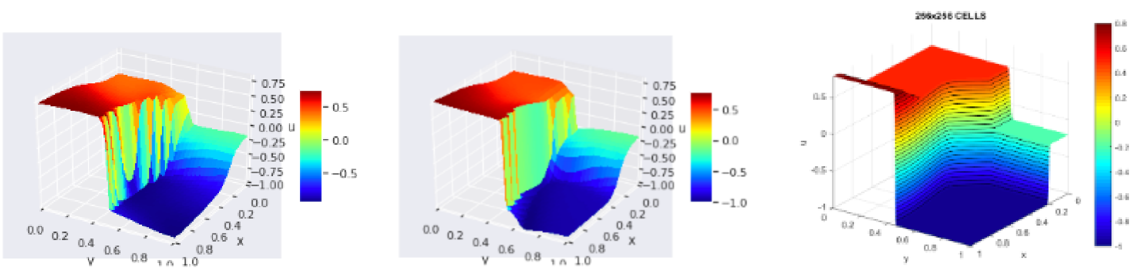

Figure 9: Neural network solution for two-dimensional inviscid Burgers equation with oblique Riemann condition. The architecture employed was composed of 4 hidden layers, 20 neurons each, $N_f = 4 \cdot 10^6$ and 15.000 epochs. Left: $2.5 \cdot 10^{-3}$ viscosity. Middle: $2.5 \cdot 10^{-4}$ viscosity. Right: SDLE solution for two-dimensional inviscid Burgers equation with oblique Riemann condition.

5 CONCLUSIONS

This work presented a PINN-based computational application of a feed-forward neural network to solve challenging hyperbolic problems in two dimensional transport models. More specifically, we solve the linear advection equation with a periodic initial condition, as well as the inviscid Burgers equation for several configurations within Riemann problems, which lead to the well-known solutions that comprise discontinuities. Our network was experimentally tuned according to each problem and interesting findings were observed, building upon the knowledge studied in [4]. For the linear advection and the inviscid Burgers equation with rarefaction wave, the results did not show spurious effects such as numerical diffusion and sonic point (*glitch line*), providing solutions pretty similar to those from the current literature. For the Riemann problems, the neural network performed better in the inviscid Burgers equation (with and without the vanishing viscosity limit) when rarefaction is predominant. In contrast, when shock is predominant, the results do not present the same effectiveness by the proposed approach.

The PINN-based approach incorporating the physical properties of the governing equations outperforms the solely data-driven type methods. We already glimpse the development of an enhanced PDE-informed method for improving the unsatisfactory results, and applying to improve simulation for hyperbolic problems in transport models. In summary, the obtained results share both practical and theoretical implications. In practical terms, the results confirm the potential of a relatively simple deep learning model in the solution of an intricate two dimensional numerical problem. In theoretical terms, this also opens an avenue for formal as well as rigorous studies on these networks as mathematically valid and effective numerical methods.

Acknowledgment

E. Abreu acknowledges the support from CNPq (Grant 306385/2019-8).

DECLARATIONS

Competing interests: The authors have no competing interests to declare that are relevant to the content of this article, that is, the authors have no financial or proprietary interests in any material discussed in this article.

Conflicts of interest: The authors have no financial or non-financial interests to disclose.

Availability of data and material: Not applicable.

Code availability: Available via request to the authors.

Consent to participate: Not applicable.

Consent for publication: Not applicable.

REFERENCES

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

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

[3] Abreu, E., François, J., Lambert, W. and Pérez, J. A semi-discrete Lagrangian–Eulerian scheme for hyperbolic-transport models. Journal of Scientific Computing, (2022). DOI `https://doi.org/10.1016/j.cam.2021.114011`.

[4] Abreu, E. and Florindo, JB. A study on a feedforward neural network to solve partial differential equations in hyperbolic-transport problems. International Conference on Computational Science, pp. 398–411, (2021).

[5] Shin, Y., Darbon, J. and Karniadakis, G.E. On the convergence of physics informed neural networks for linear second-order elliptic and parabolic type PDEs. Communications in Computational Physics, (2020). DOI 10.4208/cicp.OA-2020-0193.

[6] Shin, Y., Zhang, Z. and Karniadakis, G.E. Error estimates of residual minimization using neural networks for linear equations. Avaliable at arXiv preprint: 2010.08019, (2020).

[7] Mishra, S. and Molinaro, R. Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs. IMA Journal of Numerical Analysis, (2021).

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

[9] Tang, H. On the sonic point glitch. Journal of Computational Physics, **202**(2), pp. 507–532, (2005).

[10] Berg, J. and Nyström, K. Data-driven discovery of pdes in complex datasets. Journal of Computational Physics, **384**, pp. 239–252, (2019).

[11] Brunton, S.L., Noack, B.R. and Koumoutsakos, P. Machine learning for fluid mechanics. Annual Review of Fluid Mechanics, **52**(1), pp. 477–508, (2020).

[12] Chen, C., Seff, A., Kornhauser, A. and Xiao, J. Deep-Driving: Learning Affordance for Direct Perception in Autonomous Driving. IEEE International Conference on Computer Vision, pp. 2722–2730, (2015).

[13] Chen, G.Q.G. and Glimm, J. Kolmogorov-type theory of compressible turbulence and inviscid limit of the Navier-Stokes equations in $\mathcal{R}^3$. Physica D: Nonlinear Phenomena, **400**, (2019). DOI 132138.

[14] Dafermos, C.M. Hyperbolic conservation laws in continuous physics. Springer, (2016).

[15] Galvis, J., Abreu, E., Díaz, C. and Pérez, J. 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), pp. 1375–1408, (2020).

[16] He, K., Zhang, X., Ren, S. and Sun, J. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, (2016).

[17] Hoel, H., Karlsen, K.H., Risebro, N.H. and Storrosten, E.B. Numerical methods for conservation laws with rough flux. Stochastics and Partial Differential Equations-Analysis and Computations, **8**(1), pp. 186–261, (2020).

[18] Tadmor, E., Rascle, M. and Bagnerini, P. Compensated compactness for 2D conservation laws. Journal of hyperbolic differential equations, **2**(03), pp. 697–712, (2005).

[19] Chen, G. Q., Du, Q. and Tadmor, E. Spectral viscosity approximations to multidimensional scalar conservation laws. Mathematics of Computation, **61**(204), pp. 629–643, (1993).

[20] Kepuska, V. and Bohouta, G. Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). IEEE 8TH Annual Computing and Communication Workshop and Conference, pp. 99–103, (2018).

[21] Lellis, C.D. and Kwon, H. On non-uniqueness of Hölder continuous globally dissipative euler flows. appear in Analysis and PDEs Available at arXiv preprint: `https://arxiv.org/abs/2006.06482`, (2020).

[22] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B. and Sanchez, C.I. A survey on deep learning in medical image analysis. Medical Image Analysis, **42**, pp. 60–88, (2017).

[23] Oleinik, O.A. Discontinuous solutions of nonlinear differential equations. Uspekhi Matematicheskikh Nauk, Vol. 12, **26**(2), pp. 3–73, (1957).

[24] Quinn, B.K. Solutions with shocks: an example of an L_1–contraction semigroup. Communications on Pure and Applied Mathematics, **24**(2), pp. 125–132, (1971).

[25] Raissi, M., Perdikaris, P. and Karniadakis, G. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, **378**, pp. 686–707, (2019).

[26] Raissi, M., Yazdani, A., and Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. Science, 367(6481) (2020) 1026-1030.

[27] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., Yang, L. Physics-informed machine learning. Nature Reviews Physics, 3(6) (2021) 422-440.

[28] Roscher, R., Bohn, B., Duarte, M.F. and Garcke, J. Explainable machine learning for scientific insights and discoveries. IEEE Access, **8**, pp. 42200–42216, (2020). DOI https://doi.org/10.1109/ACCESS.2020.2976199.

[29] Rudy, S.H., Brunton, S.L., Proctor, J.L. and Kutz, J.N. Data-driven discovery of partial differential equations. Science Advances, **3**(4), (2017).

[30] Serre, D. and Silvestre, L. Multi-dimensional Burgers equation with unbounded initial data: Well-posedness and dispersive estimates. Archive for Rational Mechanics and Analysis, **234**, pp. 1391–1411, (2019).

[31] Tang, M., Liu, Y. and Durlofsky, L.J. A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. Journal of Computational Physics, **413**(109456), pp. 1–28, (2020).

[32] Young, T., Hazarika, D., Poria, S., Cambria, E. Recent Trends in Deep Learning Based Natural Language Processing. IEEE Computational Intelligence Magazine, **13**(3), pp. 55–75, (2018).

[33] E. Godlewski and P-A. Raviart, Numerical Approximation of Hyperbolic Systems of Conservation Laws (2nd ed. 2021) Springer.

[34] S. Bianchini and A. Bressan. Vanishing viscosity solutions of nonlinear hyperbolic systems. Ann. of Math. (2) 161(1) (2005) 223-342.