A study on a feedforward neural network to solve partial differential equations in hyperbolic-transport problems*

Eduardo Abreu[1[0000−0003−1979−3082]] and Joao B. Florindo[1[0000−0002−0071−0227]]

Institute of Mathematics, Statistics and Scientific Computing - 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
{eabreu,florindo}@unicamp.br

Abstract. In this work we present an application of modern deep learning methodologies to the numerical solution of partial differential equations in transport models. More specifically, we employ a supervised deep neural network that takes into account the equation and initial conditions of the model. We apply it to the Riemann problems over the inviscid nonlinear Burger's equation, whose solutions might develop discontinuity (shock wave) and rarefaction, as well as to the classical one-dimensional Buckley-Leverett two-phase problem. The Buckley-Leverett case is slightly more complex and interesting because it has a non-convex flux function with one inflection point. Our results suggest that a relatively simple deep learning model was capable of achieving promising results in such challenging tasks, providing numerical approximation of entropy solutions with very good precision and consistent to classical as well as to recently novel numerical methods in these particular scenarios.

Keywords: Neural networks · Partial differential equation · Transport models · Numerical approximation methods for PDEs · Approximation of entropy solutions.

1 Introduction

In this work, we are interested in the study of a unified approach which combines both data-driven models (regression method by machine learning) and physics-based models (PDE modeling).

Deep learning techniques have been applied to a variety of problems in science during the last years, with numerous examples in image recognition [11], natural language processing [23], self driving cars [7], virtual assistants [13], healthcare [15], and many others. More recently, we have seen a growing interest on applying those techniques to the most challenging problems in mathematics and the

* Supported by São Paulo Research Foundation (FAPESP), National Council for Scientific and Technological Development, Brazil (CNPq), and PETROBRAS - Brazil.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

solution of differential equations, especially partial differential equations (PDE), is a canonical example of such task [18].

Despite the success of recent 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. One of such points is that many of these models are based on complex structures of neural networks, sometimes comprising a large number of layers, recurrences, and other "ad-hoc" mechanisms that make them difficult to be trained and interpreted. 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 conservation laws, 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 and theoretical mathematicians and numerical analysts.

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, for instance, such as fractional conservation laws [3], compressible turbulence and Navier-Stokes equations [8], stochastic conservation laws [12] and simulation for darcy flow with hyperbolic-transport in complex flows with discontinuous coefficients [10, 2]. Burgers equation has been extensively studied in the literature (see, e.g., [9]). Burgers equations have been introduced to study different models of fluids. Thus even in the case of classical scalar one-dimensional Burgers equation, where the classical entropy condition (e.g., [16]) singles out a unique weak solution, which coincides with the one obtained by the vanishing viscosity method, there is no rigorous convergence proof for learning-based schemes. See [21] for a recent study of multi-dimensional Burgers equation with unbounded initial data concerning well-posedness and dispersive estimates. Roughly speaking, in solving the Riemann problem for systems of hyperbolic nonlinear equations, we might have nonlinear waves of several types, say, shock fronts, rarefactions, and contact discontinuities [9].

Related to the transport models treated in this work, the purelly hyperbolic equation $u_t + H_x(u) = 0$ and the corresponding augmented hyperbolic-parabolic equation $u_t^\epsilon + H_x(u^\epsilon) = \epsilon u_{xx}^\epsilon$, we mention the very recent review paper [6], 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, 10], mathematical [21, 3] and physical foundations [8, 9] driven by unprecedented volumes of data from experiments and advanced simulations at multiple spatiotemporal scales. We also mention the work [19], where the issue of *domain knowledge* is addressed as a prerequisite essential to gain explainability to enhance scientific consistency from machine learning and foundations of physics-based given in terms of mathematical equations and physical laws. However, we have seen much less discussion on more challenging PDE modeling problems, like those involving discontinuities and shock' solutions

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

numerical approximation of entropy solutions in hyperbolic-transport problems, in which the issue of conservative numerical approximation of entropy solutions is crucial and mandatory [1, 12].

This is the motivation for the study accomplished in this work, where we investigate a simple feed-forward architecture, based on the physics-informed model proposed in [18], applied to complex problems involving PDEs in transport models. More specifically, we analyze the numerical solutions of four initial-value problems: three problems on the inviscid nonlinear Burgers PDE (involving shock wave and smooth/rarefaction fan for distinct initial conditions) and on the one-dimensional Buckley-Leverett equation for two-phase configurations, which is a rather more complex and interesting because it has a non-convex flux function with one inflection point. The neural network consists of 9 stacked 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. The achieved results are promising.

Based upon a feedforward neural network approach and a simple algorithm construction, we managed to obtain a significant reduction of the error by simply controlling the input parameters of the simulations for the two fundamental models under consideration, namely, to the Burgers equation (cases rarefaction, shock and smooth) as well as to the Buckley-Leverett problem, respectively. Such results are pretty interesting if we consider the low complexity of the neural model and the challenge involved in these discontinuous cases. It also strongly suggests more in-depth studies on deep learning models that account for the underlying equation. They seem to be a quite promising line to be explored for challenging problems arising in physics, engineering, and many other areas.

What remains of this paper is organized as follows. In Section 2, we introduce the key aspects of hyperbolic problems in transport models we are considering in this work, along with a benchmark numerical scheme for comparison purposes with the traditional approach found in the specialized literature. We also offer a brief overview of the relevant approximation results for data-driven models and physics-based models in the context of PDE modeling linked to the feed-forward neural network approach. The proposed methodology considered in this work is presented in Section 3, considering stable computations and conservation properties of the feedforward neural network approximations. In Section 4, we present some numerical experiments to show the efficiency and accuracy of the computed solutions verifying the available theory. Finally, in the last Section 5, we present our 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 area. 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

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

instance, see the noteworthy book by C. M. Dafermos [9] 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, theorecical [3], numerical [5] and applications [8]. In addition, just to name some very recent works, see some interesting results covering distinct aspects, such as, theoretical [3] (uniqueness for scalar conservation laws with nonlocal and nonlinear diffusion terms) and [14] (non-uniqueness of dissipative Euler flows), and well-posedness [21] for multi-dimensional Burgers equation with unbounded initial data, numerical analysis [5] (*a posteriori* error estimates) and numerical computing for stochastic conservation laws [12] and applications [8] (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.

A basic fact of nonlinear hyperbolic transport problems is the possible loss of regularity in their solutions, namely, even solutions which are initially smooth (i.e., initial datum) may become discontinuous within finite time (blow up in finite time) and after this singular time, nonlinear interaction of shocks and rarefaction waves will play an important role in the dynamics. For the sake of simplicity, we consider the scalar 1D Cauchy problem

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

where $H \in C^2(\Omega)$, $H : \Omega \to \mathbb{R}$, $u_0(x) \in L^\infty(\mathbb{R})$ and $u = u(x,t) : \mathbb{R} \times \mathbb{R}^+ \longrightarrow \Omega \subset \mathbb{R}$. Many problems in engineering and physics are modeled by hyperbolic systems and scalar nonlinear equations [9]. 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 [10] as considered in this work. For this latter model, the flux function is smooth, namely, $H(u) = \frac{u^2}{u^2 + a(1-u)^2}$ in Eq. (1), $0 < a < 1$ ($H(u)$ is not convex with one inflection point, then an associated Riemann problem may be more complicated and the solution can involve both shock and rarefaction waves). Another interesting model is the inviscid Burgers' scalar equation used in many problems in fluid mechanics, where the flux function is $H(u) = u^2/2$ in Eq. (1). A nonlinear phenomenon that arises with the Burgers equation, even for smooth initial data, is the formation of shock, which is a discontinuity that may appear after the finite time. Together these two models [2], the Buckley-Leverett equation and Burgers' equation, are suitable and effective fundamental problems for testing new approximation algorithms to the above mentioned properties as is presented and discussed in the present work.

By using an argument in terms of traveling waves to capture the viscous profile at shocks, one can conclude that solutions of (1) satisfy Oleinik's entropy

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

condition [16], which are limits of solutions $u^\epsilon(x,t) \to u(x,t)$, where $u(x,t)$ is given by (1) and $u^\epsilon(x,t)$ is given by the augmented parabolic equation [17]

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

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

Thus, in many situations it is of importance to consider and study both hyperbolic-transport problems (1) and (2) and related conservation laws as treated in this work, and some others, of which are described in [3, 12]. 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 porous media [10] and problems in fluid mechanics [2]. Moreover, it is noteworthy that in practice the calibration of function $H(u)$ can be difficult to achieve due to unknown parameters and, thus, for instance, data assimilation can be an efficient method of calibrating reliable subsurface flow forecasts for the effective management of oil, gas and groundwater resources in geological models [22] and PDE models [4, 20]. We intend to design a unified approach which combines both PDE modeling and fine tuning machine learning techniques aiming as a fisrt step to an effective tool for advanced simulations related to hyperbolic problems in transport models such as in (1) and (2).

2.1 A benchmark numerical scheme for solving model (1)

First, we define a fixed x-uniform grid with a non-constant time step (x_j, t^n), where $(j,n) \in \mathbb{Z} \times \mathbb{N}^0$. In the space coordinates, the cells' middle points are $x_j = jh$, and the cells' endpoints are $x_{j\pm\frac{1}{2}} = jh \pm \frac{h}{2}$. The cells have a constant width $h = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}} = x_{j+1} - x_j = \Delta x$. Time step $\Delta t^n = t^{n+1} - t^n$ is non-constant subject to some Courant–Friedrichs–Lewy (CFL) stability condition. For simplicity of notation we simply use $k = \Delta t^n$. In order to numerically solve Eq. (1), instead of functions $u(\cdot,t) \in L^p(\mathbb{R})$ for $t \geq 0$, we will consider, for each time level t^n, the sequence $(U_j^n)_{j\in\mathbb{Z}}$ of the average values of $u(\cdot, t^n)$ over the x-uniform grid as follows

$$U_j^n = \frac{1}{h}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x,t^n)\,\mathrm{d}x, \tag{3}$$

for all time steps t^n, $n = 0, 1, 2, \cdots$ and in the cells $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, $j \in \mathbb{Z}$, where for t^0 we have the sequence $(U_j^0)_{j\in\mathbb{Z}}$ as an approximation of the pertinent Riemann data under study. Note that, in Equation (3), the quantity $u(x,t)$ is a solution of (1). The discrete counterpart of the space $L^p(\mathbb{R})$ is l_h^p, the space of sequences $U = (U_j)$, with $j \in \mathbb{Z}$, such that $\|U\|_{l_h^p} = \left(h\sum_{j\in\mathbb{Z}} |U_j|^p\right)^{\frac{1}{p}}$, $1 \leq p < \infty$ (for each time step t^n, as above). Following [1, 2], now suppose that the approximate solution U^h has been defined in some strip $\mathbb{R} \times [0, t_n)$, $n \geq 1$. Then we define U^h in $\mathbb{R} \times [t_n, t_{n+1})$ as setting $U^h(x,t)$ constant and equal to U_j^n, by using (3), in the

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

rectangle $(x_{j-1/2}, x_{j+1/2}] \times [t_n, t_{n+1})$ where we see that the Lagrangian-Eulerian numerical scheme applied to (1) reads the conservative method

$$U_j^{n+1} = U_j^n - \frac{k}{h}\Big[F(U_j^n, U_{j+1}^n) - F(U_{j-1}^n, U_j^n)\Big], \tag{4}$$

with the associated Lagrangian-Eulerian numerical flux [1, 2],

$$F(U_j^n, U_{j+1}^n) = \frac{1}{4}\left[\frac{h}{k}(U_j^n - U_{j+1}^n) + 2\Big(H(U_{j+1}^n) + H(U_j^n)\Big)\right]. \tag{5}$$

The classical Lax-Friedrichs numerical flux for (4), found elsewhere, is given by:

$$F(U_j^n, U_{j+1}^n) = \frac{1}{2}\left[\frac{h}{k}(U_j^n - U_{j+1}^n) + \Big(H(U_{j+1}^n) + H(U_j^n)\Big)\right]. \tag{6}$$

Here, both schemes (6) and (5) should follow the stability CFL condition

$$\max_j \left\{ |H'(U_j^n)|, \left|\frac{H(U_j^n)}{U_j^n}\right| \right\} \frac{k}{h} < \frac{1}{2}, \tag{7}$$

for all time steps n, where $k = \Delta t^n$ and $h = \Delta x$, $H'(U_j^n)$ is the partial derivative of H, namely $\dfrac{\partial H(u)}{\partial u}$ for all U_j^n in the mesh grid.

3 Proposed methodology

The neural network employed here is based on that described in [18]. There a "physics-informed" neural network is defined to solve nonlinear partial differential equations. That network takes into account the original equation by explicitly including the PDE functional and initial and/or boundary conditions in the objective function and taking advantage of automatic differentiation widely used in the optimization of classical neural networks. It follows a classical feed-forward architecture, with 9 hidden layers, each one with a hyperbolic tangent used as activation function. More details are provided in the following.

The general problem solved here has the form

$$u_t + \mathcal{N}(u) = 0, \qquad x \in \Omega, t \in [0, T], \tag{8}$$

where $\mathcal{N}(\cdot)$ is a non-linear operator and $u(x,t)$ is the desired solution. Unlike the methodology described in [18], here we do not have an explicit boundary condition and the neural network is optimized only over the initial conditions of each problem.

We focus on four problems: the inviscid nonlinear Burgers equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \qquad x \in [-10, 10], \qquad t \in [0, 8], \tag{9}$$

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

with shock initial condition

$$u(x, 0) = 1, x < 0 \text{ and } u(x, 0) = 0, x > 0, \tag{10}$$

discontinuous initial data (hereafter *rarefaction fan initial condition*)

$$u(x, 0) = -1, x < 0 \text{ and } u(x, 0) = 1, x > 0, \tag{11}$$

smooth initial condition

$$u(x, 0) = 0.5 + \sin(x), \tag{12}$$

and the two-phase Buckley-Leverett

$$\begin{aligned} &u_t + \left(\frac{u^2}{u^2 + a(1-u)^2}\right)_x = 0, \qquad x \in [-8, 8], \qquad t \in [0, 8], \\ &u(x, 0) = 1, x < 0 \text{ and } u(x, 0) = 0, x > 0. \end{aligned} \tag{13}$$

In this problem we take $a = 1$.

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

$$f := u_t + \mathcal{N}(u), \tag{14}$$

such that

$$\mathcal{N}(u) = \left(\frac{u^2}{2}\right)_x \tag{15}$$

in the inviscid Burgers and

$$\mathcal{N}(u) = \left(\frac{u^2}{u^2 + a(1-u)^2}\right)_x \tag{16}$$

in the Buckley-Leverett. Here we also have an important novelty which is the introduction of a derivative (w.r.t. x) in $\mathcal{N}(u)$, which was not present in [18].

The function f is 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 ($0.01u_{xx}$) for better stabilization, but in view on the modeling problems (1) and (2). It is crucial to mention at this point that numerical approximation of entropy solutions (with respect to the neural network) to hyperbolic-transport problems also require the notion of entropy-satisfying weak solution. The neural network computes the expected solution $u(x, t)$ and its output and the derivatives present in the evaluation of f are obtained by automatic differentiation.

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

$$\begin{aligned} \mathcal{L}_f(u) &= \frac{1}{N_f}\sum_{i=1}^{N_f} |f(x_f^i, t_f^i,)|^2, \\ \mathcal{L}_u(u) &= \frac{1}{N_u}\sum_{i=1}^{N_u} |u(x_u^i, t_u^i) - u^i|^2, \end{aligned} \tag{17}$$

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

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

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

$$u(x,t) \approx \arg\min_u [\mathcal{L}_f(u) + \mathcal{L}_u(u)]. \tag{18}$$

Inspired by the great results in [18], here the minimization is performed by the L-BFGS-B optimizer and the algorithm stops when a loss of 10^{-6} is reached. Figure 1 illustrates the evolution of the total loss function (in $\log_{10}$ scale to facilitate visualization) for the inviscid Burgers equation with shock initial condition.

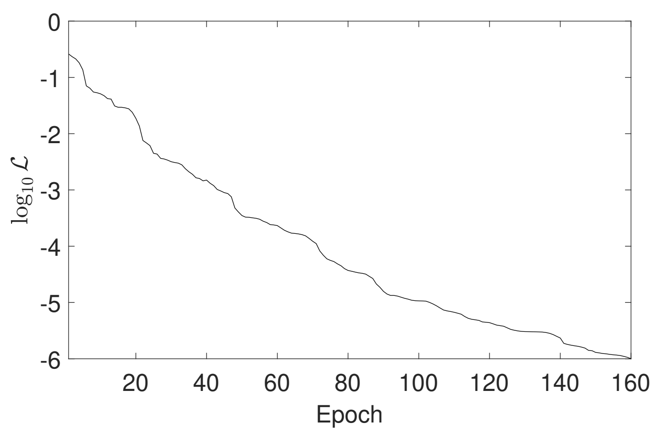

Fig. 1. Loss function evolution (in $\log_{10}$ scale) for the inviscid Burgers equation with shock initial condition.

4 Results and Discussion

In the following we present results for the solutions of the investigated problems obtained by the neural network model. We compare these solutions with two numerical schemes: Lagrangian-Eulerian and Lax-Friedrichs. These are very robust numerical methods with a solid mathematical basis. Here we use one scheme to validate the other. In fact, the solutions obtained by each scheme are very similar. For that reason, we opted for graphically showing curves only for the Lagrangian-Eulerian solution. However, we exhibit the errors of the proposed methodology both in comparison with Lagrangian-Eulerian (EEL) and Lax-Friedrichs (ELF).

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

Here such error corresponds to the average quadratic error, i.e.,

$$\begin{aligned} ELF(t) &= \frac{\sum_{i=1}^{N_u}(u_{NN}(x^i,t) - u_{LF}(x^i,t))^2}{N_u}, \\ EEL(t) &= \frac{\sum_{i=1}^{N_u}(u_{NN}(x^i,t) - u_{LE}(x^i,t))^2}{N_u}, \end{aligned} \tag{19}$$

where u_{NN}, u_{LF}, and u_{LE} correspond to the neural network, Lax-Friedrichs, and Lagrangian-Eulerian solutions, respectively. In our tests, we used $N_f = 10^4$ unless otherwise stated, and $N_u = 100$. These parameters of the neural network are empirically determined and are not explicitly related to the collocation points used in the numerical scheme. In fact, theoretical studies on optimal values for these parameters is an open research problem that we also intend to investigate in future works. For the numerical reference schemes we adopted CFL condition 0.4 for Lax-Friedrichs and 0.2 for Lagrangian-Eulerian. We also used $\Delta x = 0.01$.

For the rarefaction case, we observed that using $N_f = 10^4$ collocation points was sufficient to provide good results. In this scenario, we also verified the number of neurons, testing 40 and 60 neurons. Figure 2 shows the obtained solution compared with reference and the respective errors. Interestingly, the error decreases when time increases, which is a consequence of the solution behavior, which becomes smoother (smaller slope) for larger times, showing good accuracy and evidence that we are computing the correct solution in our numerical simulation.

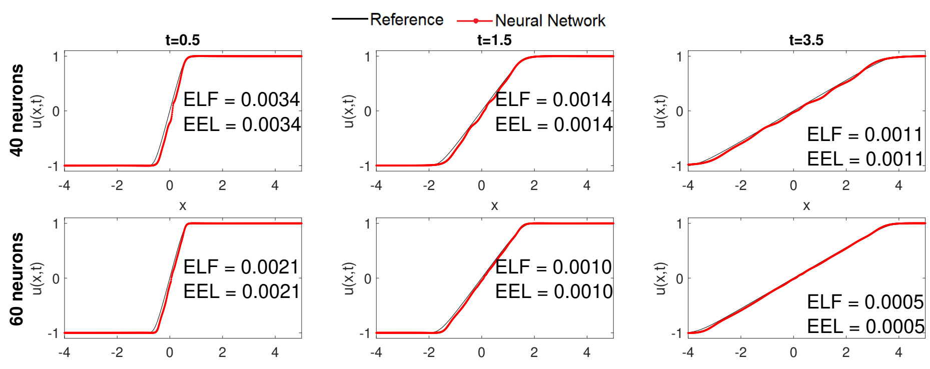

Fig. 2. Burgers: Rarefaction.

Figure 3 illustrates the performance of the neural network model for the inviscid Burgers equation with shock initial condition. Here we had to add a small viscous term ($0.01u_{xx}$) to obtain the entropy solution. Such underlying viscous mechanism did not bring significant reduction in error, but the general structure of the obtained solution is better, attenuating spurious fluctuations around the discontinuities. It was also interesting to see that the addition of

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

more neurons did not reduce the error for this initial condition. This is a typical example of overfitting caused by over-parameterization. An explanation for that is the relative simplicity of the initial condition, assuming only two possible values.

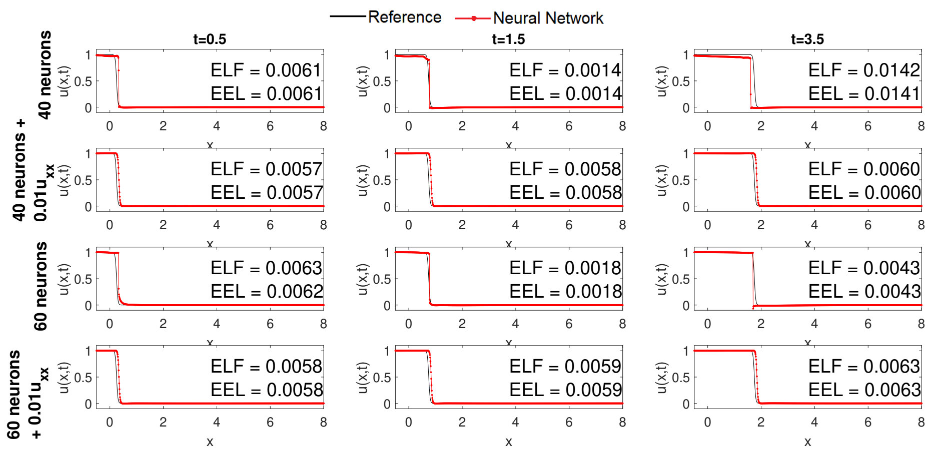

Fig. 3. Burgers: Shock.

Figure 4 depicts the solutions for the smooth initial condition in the inviscid Burgers equation. Here, unlike the previous case, increasing the number of neurons actually reduced the error. Indeed, it turned out better than expected considering that now both initial condition and solution are more complex. Nevertheless, we identified that tuning only the number of neurons was not enough to achieve satisfactory solutions in this situation. Therefore we also tuned the parameter N_f. In particular, we discovered that combining the same small viscous term used for the shock case with $N_f = 10^6$ provided excellent results, with quite promising precision in comparison with our reference solutions.

Another case characterized by solutions with more complex behavior is Buckley-Leverett with shock initial condition (Figure 5). In this example, similarly to what happened in the smooth case, again, the combination of $N_f = 10^6$ with the small viscous term was more effective than any increase in the number of neurons. While the introduction of the small viscous term attenuated fluctuations in the solution when using 40 neurons, at the same time when using $N_f = 10^4$, we observe that increasing the number of neurons causes an increase in the delay between the solution provided by the network and the reference.

Generally speaking, the neural networks studied here were capable of achieving promising results in challenging situations involving different types of discontinuities and nonlinearities. Moreover, our numerical findings might also suggest some good evidence on the robustness of a feedforward neural network as numerical approximation procedure for solving nonlinear hyperbolic-transport

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

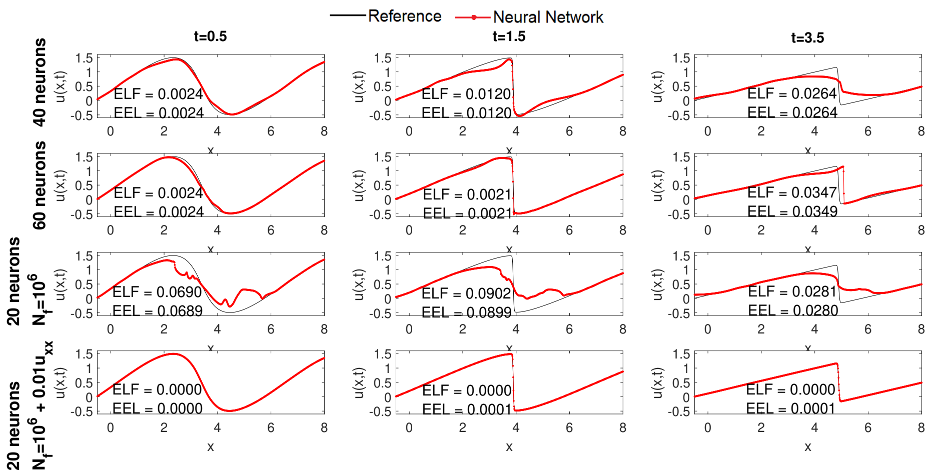

Fig. 4. Burgers: Smooth.

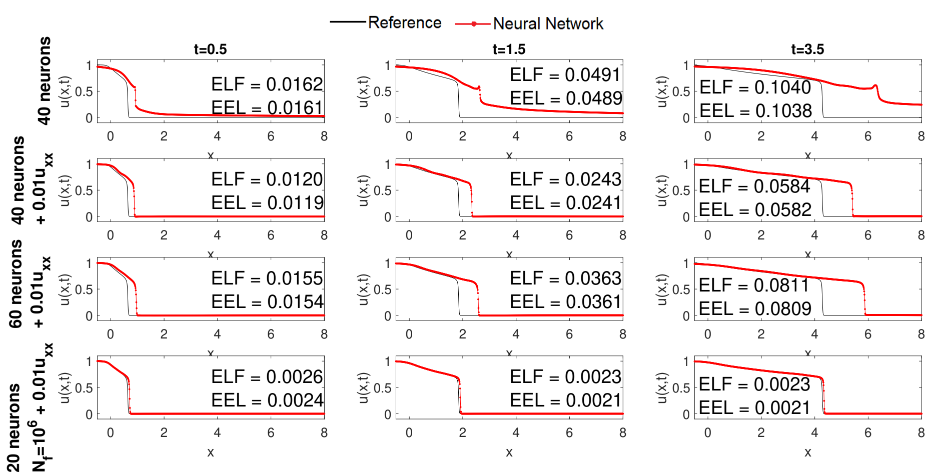

Fig. 5. Buckley-Leverett: Rarefaction + Shock.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

problems. Mathematical theoretical foundation of this model is to be pursued in a further work. In particular, the neural networks obtained results pretty close to those provided by entropic numerical schemes like Lagrangian-Eulerian and Lax-Friedrichs. Going beyond the analysis in terms of raw precision, these results give us evidences that our neural network model possess some type of entropic property, which from the viewpoint of a numerical method is a fundamental and desirable characteristic.

5 Conclusions

This work presented an application of a feed-forward neural network to solve challenging hyperbolic problems in transport models. More specifically, we solve the inviscid Burgers equation with shock, smooth and rarefaction initial conditions, as well as the Buckley-Leverett equation with classical Riemann datum, which lead to the well-known solution that comprises a rarefaction and a shock wave. Our network was tuned according to each problem and interesting findings were observed. At first, our neural network model was capable of providing solutions pretty similar to those obtained by two numerical schemes used as references: Lagrangian-Eulerian and Lax-Friedrichs. Besides, the general structure of the obtained solutions also behaved as expected, which is a remarkable achievement considering the intrinsic challenge of these problems. In fact, the investigated neural networks showed evidences of an entropic property to the scalar hyperbolic-tranport model studies, which is an important attribute of any numerical scheme when dealing with weak solution of scalar conservation laws.

Our approach is substantially distinct from the current trend of merely data-driven discovery type methods for recovery governing equations by using machine learning and artificial intelligence algorithms in a straightforward manner. We glimpse the use of novel methods, fine tuning machine learning algorithms and very fine mathematical and numerical analysis to improve comprehension of regression methods aiming to identify the potential and reliable prediction for advanced simulation for hyperbolic problems in transport models as well as the estimation of financial returns and economic benefits.

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 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.

Acknowledgements

J. B. Florindo gratefully acknowledges the financial support of São Paulo Research Foundation (FAPESP) (Grant #2016/16060-0) and from National Council for Scientific and Technological Development, Brazil (CNPq) (Grants #301480 /2016-8 and #423292/2018-8). E. Abreu gratefully acknowledges the financial support of São Paulo Research Foundation (FAPESP) (Grant #2019/20991-8),

DOI: 10.1007/978-3-030-77964-1_31

from National Council for Scientific and Technological Development - Brazil (CNPq) (Grant #2 306385/2019-8) and PETROBRAS - Brazil (Grant #2015/ 00398-0). E. Abreu and J. B. Florindo also gratefully acknowledge the financial support of Red Iberoamericana de Investigadores en Matemticas Aplicadas a Datos (MathData).

References

1. Abreu, E., Matos, V., Pérez, J., Rodrguez-Bermúdez, P.: 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) (2021)
2. Abreu, E., Prez, J.: A fast, robust, and simple Lagrangian-Eulerian solver for balance laws and applications. Computers & Mathematics with Applications **77**(9), 2310 – 2336 (2019)
3. Alibaud, N., Andreianov, B., Oudraogo, A.: Nonlocal dissipation measure and $\mathcal{L}^1$ kinetic theory for fractional conservation laws. Communications in Partial Differential Equations **45**(9), 1213–1251 (2020)
4. Berg, J., Nystrm, K.: Data-driven discovery of pdes in complex datasets. Journal of Computational Physics **384**, 239 – 252 (2019)
5. Bressan, A., Chiri, M.T., Shen, W.: A posteriori error estimates for numerical solutions to hyperbolic conservation laws (2020), Available at https://arxiv.org/abs/2010.00428
6. Brunton, S.L., Noack, B.R., Koumoutsakos, P.: Machine learning for fluid mechanics. Annual Review of Fluid Mechanics **52**(1), 477–508 (2020)
7. Chen, C., Seff, A., Kornhauser, A., Xiao, J.: DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving. In: 2015 IEEE International Conference on Computer Vision (ICCV). pp. 2722–2730. IEEE International Conference on Computer Vision (2015)
8. Chen, G.Q.G., 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**, 132138 (2019)
9. Dafermos, C.M.: Hyperbolic conservation laws in continuous physics. Springer (2016)
10. Galvis, J., Abreu, E., Díaz, C., 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), 1375–1408 (2020)
11. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778. IEEE Conference on Computer Vision and Pattern Recognition (2016)
12. Hoel, H., Karlsen, K.H., Risebro, N.H., Storrosten, E.B.: Numerical methods for conservation laws with rough flux. Stochastics and Partial Differential Equations-Analysis and Computations **8**(1), 186–261 (2020)
13. Kepuska, V., Bohouta, G.: Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In: 2018 IEEE 8TH Annual Computing and Communication Workshop and Conference (CCWC). pp. 99–103 (2018)
14. Lellis, C.D., Kwon, H.: On non-uniqueness of Hölder continuous globally dissipative euler flows (2020), Available at https://arxiv.org/abs/2006.06482

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31

15. 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., Sanchez, C.I.: A survey on deep learning in medical image analysis. Medical Image Analysis **42**, 60–88 (2017)
16. Oleinik, O.A.: Discontinuous solutions of nonlinear differential equations. Uspekhi Matematicheskikh Nauk, Vol. 12, 1957, pp. 3-73. (English translation Transactions of the American Mathematical Society **26**(2), 95–172 (1963)
17. Quinn, B.K.: Solutions with shocks: an example of an L_1–contraction semigroup. Communications on Pure and Applied Mathematics **24**(2), 125–132 (1971)
18. Raissi, M., Perdikaris, P., 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**, 686 – 707 (2019)
19. Roscher, R., Bohn, B., Duarte, M.F., Garcke, J.: Explainable machine learning for scientific insights and discoveries. IEEE Access **8**, 42200–42216 (2020). https://doi.org/10.1109/ACCESS.2020.2976199
20. Rudy, S.H., Brunton, S.L., Proctor, J.L., Kutz, J.N.: Data-driven discovery of partial differential equations. Science Advances **3**(4) (2017)
21. Serre, D., Silvestre, L.: Multi-dimensional Burgers equation with unbounded initial data: Well-posedness and dispersive estimates. Archive for Rational Mechanics and Analysis **234**, 1391–1411 (2019)
22. Tang, M., Liu, Y., Durlofsky, L.J.: A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. Journal of Computational Physics **413**(109456), 1–28 (2020)
23. Young, T., Hazarika, D., Poria, S., Cambria, E.: Recent Trends in Deep Learning Based Natural Language Processing. IEEE Computational Intelligence Magazine **13**(3), 55–75 (2018)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77964-1_31