Gradient weighted physics-informed neural networks for capturing shocks in porous media flows

Somiya Kapoor
DST-Centre for
Interdisciplinary Mathematical Sciences
Banaras Hindu University
Varanasi, India
somiya.kaps@gmail.com

Abhishek Chandra
Department of Electrical Engineering
TU Eindhoven
Eindhoven, The Netherlands

Taniya Kapoor
Department of Civil Engineering and Geosciences
TU Delft
Delft, The Netherlands

Mitrofan Curti
Department of Electrical Engineering
TU Eindhoven
Eindhoven, The Netherlands

Abstract

Physics-informed neural networks (PINNs) seamlessly integrate physical laws into machine learning models, enabling accurate simulations while preserving the underlying physics. However, PINNs are still *suboptimal in approximating discontinuities* in the form of *shocks* compared to the traditional numerical shock-capturing methods. This paper proposes a framework to approximate shocks arising in dynamic *porous media flows* by weighting the governing nonlinear partial differential equation (PDE) with a physical gradient-based term in the loss function. The applicability of the proposed framework is investigated on the forward problem of immiscible two-phase fluid transport in porous media governed by a nonlinear first-order hyperbolic *Buckley–Leverett* PDE. Particularly, convex and non-convex flux functions are studied involving shocks and rarefaction. The results demonstrate that the proposed framework consistently learns accurate approximations containing shocks and rarefaction by weighting the underlying PDE with a physical gradient term and *outperforms state-of-the-art* artificial viscosity-based neural network methods to capture shocks on the standard L^2-norm metric.

1 Introduction

Porous media flows are ubiquitous in physical sciences and manifest in diverse industrial applications [1]. Understanding and simulating these flows is essential for optimizing design, predicting transport phenomena, and addressing environmental challenges. Nonlinear partial differential equations (PDEs) are often used to model and simulate these flows [2, 3]. The governing nonlinear equations are typically simulated using numerical methods such as the essentially non-oscillatory method [4], weighted essentially non-oscillatory method [5], and the discontinuous Galerkin method [6], among others. These methods ensure a *precise and oscillation-free* solution of shock waves and analogous discontinuities, a crucial aspect in dealing with hyperbolic equations for porous media flows. However, the surge in machine learning techniques has sparked significant interest in leveraging data-driven [7, 8] and *physics-based* [9, 10] methods for more efficient and accurate simulations for porous media flows.

Physics-informed neural network (PINN) [11] has emerged as a noteworthy method in the context of physics-based methods [12] and has gained considerable attention, particularly for its efficacy in

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

solving complex problems within diverse engineering domains such as fluids [13], materials [14], and structures[15]. However, when applied to porous media flows, a notable *challenge* arises due to the presence of *shock waves*. Vanilla PINN faces difficulties in accurately capturing shocks, presenting a significant obstacle in its application to these problems [9]. Various modifications of vanilla PINN have been proposed to capture shocks efficiently. Infusing thermodynamics with PINN [16] was proposed for inverse problems. The inverse problem was also simulated in [17] in the context of supersonic flows. In another study, Euler equations involving shock waves in high-gradient regions were simulated by appropriately clustering training data points and reducing error propagation across the entire domain [18]. A similar study was carried out in [19] by subdividing the computational domain into smaller subdomains and employing a different neural network for each subdomain to address Burgers and Euler equations. Recently, [10] proposed adaptive *artificial viscosity*-based strategies to solve porous media flows governed by Buckley-Leverett equations.

In contrast, in this work, we do not use artificial viscosity to change the underlying hyperbolic PDE into a parabolic PDE [20]. Instead, we follow the work presented in [21] and propose to *weight the governing PDE* in the loss function with a physics-based term. In [21], discontinuities in the form of shocks were approximated for the inviscid Burgers equation and Sod and Lax problems for the Euler equation. We use the method therein and propose to resolve the paradoxical challenges associated with shock capturing in porous media flows with PINNs. Primarily, the method *prioritizes* training in *smooth regions* while *diminishing* the training influence in *highly compressed regions* associated with shock waves. This approach allows for a precise and sharp shock capture. We implement this strategy by introducing a *positive weight* related to compression into the governing equations, thereby adjusting the high gradients at specific spatial locations. For the rest of this paper, we refer to this method as *PINN-WE*, as proposed in [21].

In this paper, we explore using the PINN-WE approach to address the forward problem of nonlinear two-phase transport within porous media. In particular, we simulate the *nonlinear Buckley-Leverett equation* with convex and non-convex flux functions, whose solutions contain shocks and rarefaction. The rest of the paper is structured as follows: Section 2 presents the PINN-WE framework, our tool for solving the underlying PDEs. Section 3 presents the application of PINN-WE on the Buckley-Leverett equation modeling the porous media flow. Finally, the conclusions drawn from the paper are collated in Section 4.

2 Method

The presented method, PINN-WE, is based on a *minor modification* of the PINN loss function. Hence, we first describe PINN and its corresponding loss function in brief. We define an abstract operator $\mathcal{D}$ governing the PDE, $\mathcal{D}[u(x,t)] = 0$, along with an abstract operator $\mathcal{B}$ governing the initial and boundary conditions, $\mathcal{B}[u(x,t)] = 0$. Here, $u \in \mathbb{R}$ denotes the *unknown* physical quantity of interest and $x \in \mathbb{R}$, and $t \in \mathbb{R}$ represent the space and time variables, respectively.

PINNs approximate the solution of the PDE by minimizing $\mathcal{D}$ along with $\mathcal{B}$ simultaneously. The operator $\mathcal{D}$ acts as a physics-informed component, where *no data* of the solutions is known or available. The operator $\mathcal{B}$ acts as a data part where initial and boundary conditions act as data, which are known a priori for a forward problem to be well-posed. The data part fits the provided data, and the physics part regularizes the neural network's approximation towards the *optimal solution* from the possible envelope of solutions. The loss function of PINNs is defined as,

$$L_{\text{PINN}} = ||\mathcal{D}[u(x,t)]||^2 + ||\mathcal{B}[u(x,t)]||^2 \tag{1}$$

The physics and data parts could be weighted through several strategies, as discussed in the review paper [22]. In practice, these methods have shown tremendous success. However, we employ PINN-WE to address the challenges of capturing shocks in porous media flow. For PINN-WE, the physics-based term is weighted with λ defined as,

$$\lambda = \frac{1}{\epsilon_2(|\nabla \cdot u| - \nabla \cdot u) + 1} \tag{2}$$

where, ϵ_2 is a hyperparameter, and $\nabla \cdot u$ represents the divergence of u. Subsequently, the physics based term becomes $\frac{1}{\epsilon_2(|\nabla \cdot u| - \nabla \cdot u)+1}||\mathcal{D}[u(x,t)]||^2$. In addition, a hyperparameter ϵ_1 is also weighted with the data term to make the loss function more generic and *expressive*. The loss function for

Table 1: L^2 relative percent error in simulating the considered cases of the Buckley-Leverett equation

Method	Convex	Non-Convex $M = 0.5$	Non-Convex $M = 1$	Non-Convex $M = 2$	Non-Convex $M = 10$
PINN-WE	**1.05**	**2.96**	**4.70**	**0.17**	**1.66**
Learnable global AV [10]	3.05	6.72	7.63	6.88	7.52
Parametric AV map [10]	3.33	6.93	6.42	5.94	7.66
Residual-based AV map [10]	3.34	8.70	10.04	8.33	7.19
Nonadaptive global AV [10]	4.50	6.55	7.23	7.05	8.75

PINN-WE is defined as follows,

$$L_{\text{PINN-WE}} = \frac{1}{\epsilon_2(|\nabla \cdot u| - \nabla \cdot u) + 1}||\mathcal{D}[u(x,t)]||^2 + \epsilon_1||\mathcal{B}[u(x,t)]||^2 \tag{3}$$

It is evident that equation 3 reduces to equation 1 for $\epsilon_2 = 0$ and $\epsilon_1 = 1$, making $L_{\text{PINN-WE}}$ as special case of weighted L_{PINN}. Traditionally, in PINN literature, the hyperparameter choice of λ and ϵ_1 is made as a *pre-defined constant* [23, 24] or are trained *adaptively* to modulate the influence of distinct components within the loss function. It is done to emphasize that the points located within the interior and those on boundaries exhibit *varying degrees of significance*. A common practice involves assigning greater significance to points located on boundaries.

For each distinct component of the loss function, traditional PINN-based approaches involve computing the average impact of each computational point on the overall loss. This approach is appropriate for scenarios characterized by smooth solutions but may not be effective for solutions with shocks and discontinuities. Consequently, the proposed framework alleviates the impact of points located within highly compressible regions by incorporating a local positive physics-based gradient weight (λ) into the governing equations to capture shocks efficiently.

3 Numerical Experiments

The *Buckley-Leverett* PDE is a mathematical representation for characterizing the displacement of immiscible and incompressible two-phase flow within porous media [25]. This one-dimensional nonlinear hyperbolic transport equation is expressed as,

$$\mathcal{D}[u(x,t)] := \frac{\partial u}{\partial t} + \frac{\partial f_w(u)}{\partial x} = 0 \tag{4}$$

where $f_w(u)$ is the flux function and $x, t \in [0, 1]$, and u represents water saturation for instance in a water-oil medium. Different flux functions result in diverse types of waves within the solution. Additionally, we employ uniform initial and boundary conditions for all the numerical experiments representing the injection of water at one end of a 1-D reservoir filled with oil,

$$\mathcal{B}[u(x,t)] := u(x,t) = \begin{cases} 0, & \forall x, \wedge t = 0, \\ 1, & x = 0, \wedge t > 0 \end{cases} \tag{5}$$

Five numerical experiments are performed, encompassing one convex and four non-convex flux functions. The neural network has two inputs x, t and one output u. 8 hidden layers with 20 neurons each and $tanh$ activation function with L-BFGS-B [26] optimizer is used to train the network. Initial trainable parameters are generated through Xavier initialization [27]. For all cases, 300 random points on initial time and boundary with 10000 collocation points are used. A comparison for all experiments is also presented with state-of-the-art artificial viscosity (AV) based methods [10]. The convex flux function for the first numerical experiment is $f_w(u) = u^2$. For the next four experiments, a non-convex flux function is taken depending on a parameter M as,

$$f_w(u) = \frac{u^2}{u^2 + \frac{(1-u)^2}{M}} \tag{6}$$

Equation 4- 6 represent a typical scenario for the Buckley-Leverett problem in porous media flow.

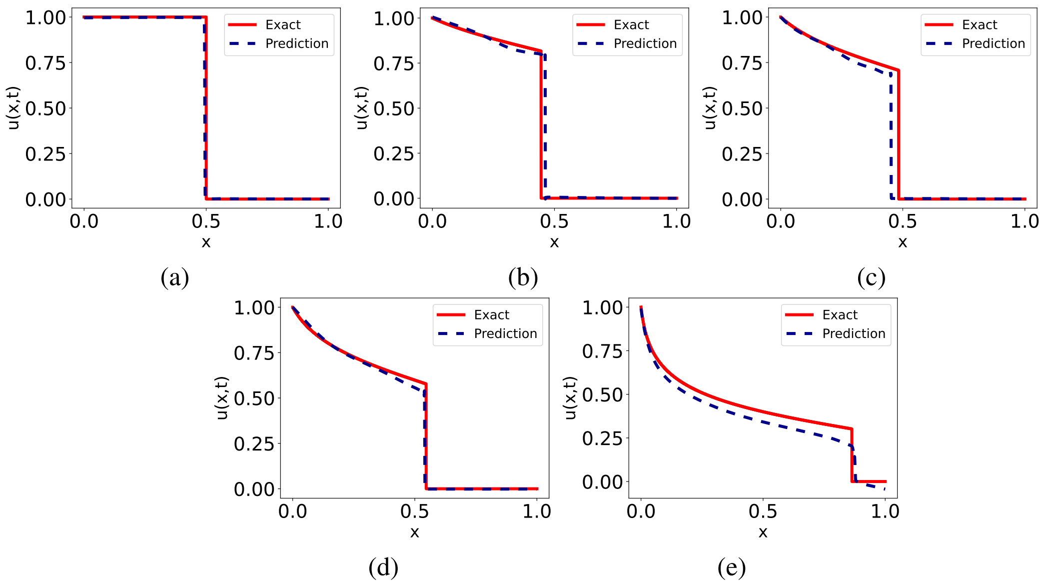

Figure 1: Comparison of the predicted solution with the exact solution for flux function (a) convex at $t = 0.5$. At $t = 0.4$ for non-convex (b) $M = 0.5$ (c) $M = 1$ (d) $M = 2$ (e) $M = 10$.

Table 2: Hyperparameters

	Convex	Non-Convex $M = 0.5$	Non-Convex $M = 1$	Non-Convex $M = 2$	Non-Convex $M = 10$
ϵ_1	0.1	10	1	1	0.1
ϵ_2	0.5	0.1	0.1	0.1	0.1

For the non-convex flow, we chose four distinct values of M as $0.5, 1, 2$, and 10. The analytical solution to this problem comprises both a shock and a rarefaction wave and is defined as follows:

$$u(x,t) = \begin{cases} 0, & \frac{x}{t} > f'_w(u^*) \\ u\left(\frac{x}{t}\right), & f'_w(u^*) \geq \frac{x}{t} \geq f'_w(u=1) \\ 1, & f'_w(u=1) \geq \frac{x}{t} \end{cases} \tag{7}$$

Where, u* represents the shock location, which is determined by the Rankine-Hugoniot condition

$$f'_w(u^*) = \frac{f_w(u^*) - f_w(u)|_{u=0}}{u^* - u|_{u=0}}$$

Additionally, $u(\frac{x}{t})$ is defined for $\frac{x}{t} \leq f'_w(u^*)$ as $u\left(\frac{x}{t}\right) = (f'_w)^{-1}(\frac{x}{t})$. Due to its self-similarity, the analytical solution 7 depends on just one governing parameter, the similarity variable $\frac{x}{t}$.

The predicted solution for all five experiments and the analytical solution are presented in Fig. 1. In addition, Table 1 presents a comparison for the L^2 relative percent error with state-of-the-art AV-based methods [10]. Finally, Table 2 presents the choice of hyperparameters ϵ_1 and ϵ_2 used to train the model.

4 Conclusions

We proposed a framework to address the challenge of accurately *approximating shocks* in dynamic porous media flows using PINNs. By introducing a *physical gradient-based* term in the loss function, the framework significantly improves the performance of PINNs in capturing discontinuities. We applied the framework to the forward problem of immiscible two-phase fluid transport in porous media, considering both convex and non-convex flux functions with shocks and rarefaction. The results consistently demonstrate the effectiveness of our approach in learning precise approximations for *several non-convex* cases governing *distinct* shock scenarios. The framework also *outperformed the state-of-the-art* artificial viscosity-based neural network methods for the Buckley-Leverett PDE in approximating the discontinuous solution. The results motivate the utilization of the proposed methodology in simulating intricate physical phenomena involving shocks, thereby broadening their potential utility across various scenarios.

Acknowledgements

The authors thank Emilio Jose Rocha Coutinho for providing analytical solutions that enriched the comparison of our proposed methodology.

References

[1] Adrian E Scheidegger. *The physics of flow through porous media*. University of Toronto press, 1957.

[2] Siddhartha Mishra and Jérôme Jaffré. On the upstream mobility scheme for two-phase flow in porous media. *Computational Geosciences*, 14:105–124, 2010.

[3] Kimoon Um, Eric J Hall, Markos A Katsoulakis, and Daniel M Tartakovsky. Causality and bayesian network pdes for multiscale representations of porous media. *Journal of Computational Physics*, 394:658–678, 2019.

[4] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy. Uniformly high order accurate essentially non-oscillatory schemes, iii. pages 218–290. Springer, 1987.

[5] G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted eno schemes. *Journal of Computational Physics*, 126(1):202–228, 1996.

[6] B. Cockburn and C.-W. Shu. The local discontinuous galerkin method for time-dependent convection-diffusion systems. *SIAM Journal on Numerical Analysis*, 35(6):2, 1998.

[7] Vinicius LS Silva, Pablo Salinas, Matthew D Jackson, and Christopher C Pain. Machine learning acceleration for nonlinear solvers applied to multiphase porous media flow. *Computer Methods in Applied Mechanics and Engineering*, 384:113989, 2021.

[8] Ying Da Wang, Traiwit Chung, Ryan T Armstrong, and Peyman Mostaghimi. Ml-lbm: Machine learning aided flow simulation in porous media. *arXiv preprint arXiv:2004.11675*, 2020.

[9] Olga Fuks and Hamdi A. Tchelepi. Limitations of physics informed machine learning for nonlinear two-phase transport in porous media. *Journal of Machine Learning Modeling and Computations*, 1(1), 2020.

[10] Emilio Jose Rocha Coutinho, Marcelo Dall'Aqua, Levi McClenny, Ming Zhong, Ulisses Braga-Neto, and Eduardo Gildin. Physics-informed neural networks with adaptive localized artificial viscosity. *Journal of Computational Physics*, page 112265, 2023.

[11] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[12] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[13] Maziar Raissi, Alireza Yazdani, and George Em Karniadakis. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, 367(6481):1026–1030, 2020.

[14] Enrui Zhang, Ming Dao, George Em Karniadakis, and Subra Suresh. Analyses of internal structures and defects in materials using physics-informed neural networks. *Science advances*, 8(7):eabk0644, 2022.

[15] Taniya Kapoor, Hongrui Wang, Alfredo Núñez, and Rolf Dollevoet. Physics-informed neural networks for solving forward and inverse problems in complex beam systems. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–15, 2023.

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

[17] A. D. Jagtap, Z. Mao, N. Adams, and G. E. Karniadakis. Physics-informed neural networks for inverse problems in supersonic flows. *arXiv preprint arXiv:2202.11821*.

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

[19] A. D. Jagtap, E. Kharazmi, and G. E. Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365:113028, 2020.

[20] Rajat Arora. A deep learning framework for solving hyperbolic partial differential equations: Part i. *arXiv preprint arXiv:2307.04121*, 2023.

[21] Li Liu, Shengping Liu, Hui Xie, Fansheng Xiong, Tengchao Yu, Mengjuan Xiao, Lufeng Liu, and Heng Yong. Discontinuity computing using physics-informed neural network. *arXiv preprint arXiv:2206.03864*, 2022.

[22] Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.

[23] Abhishek Chandra, Mitrofan Curti, Koen Tiels, Elena A Lomonova, and Daniel M Tartakovsky. Physics-informed neural networks for modelling anisotropic and bi-anisotropic electromagnetic constitutive laws through indirect data. In *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1451–1459. IEEE, 2022.

[24] Taniya Kapoor, Hongrui Wang, Alfredo Núñez, and Rolf Dollevoet. Physics-informed machine learning for moving load problems. *arXiv preprint arXiv:2304.00369*, 2023.

[25] Ruben Rodriguez-Torrado, Pablo Ruiz, Luis Cueto-Felgueroso, Michael Cerny Green, Tyler Friesen, Sebastien Matringe, and Julian Togelius. Physics-informed attention-based neural network for hyperbolic partial differential equations: application to the buckley–leverett problem. *Scientific reports*, 12(1):7557, 2022.

[26] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, New York, NY, USA, second edition, 2006.

[27] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. pages 249–256, 2010.