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. 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. The governing nonlinear equations are typically simulated using numerical methods such as the essentially non-oscillatory method, weighted essentially non-oscillatory method, and the discontinuous Galerkin method, 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 and physics-based methods for more efficient and accurate simulations for porous media flows.

Physics-informed neural network (PINN) has emerged as a noteworthy method in the context of physics-based methods and has gained considerable attention, particularly for its efficacy in...
solving complex problems within diverse engineering domains such as fluids, materials, and structures. 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. Various modifications of vanilla PINN have been proposed to capture shocks efficiently. Infusing thermodynamics with PINN was proposed for inverse problems. The inverse problem was also simulated 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. A similar study was carried out by subdividing the computational domain into smaller subdomains and employing a different neural network for each subdomain to address Burgers and Euler equations. Recently, 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. Instead, we follow the work presented in and propose to weight the governing PDE in the loss function with a physics-based term. In 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.

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 presents the PINN-WE framework, our tool for solving the underlying PDEs. Section 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.

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 \( D \) governing the PDE, \( D[u(x, t)] = 0 \), along with an abstract operator \( B \) governing the initial and boundary conditions, \( B[u(x, t)] = 0 \). 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 \( D \) along with \( B \) simultaneously. The operator \( D \) acts as a physics-informed component, where no data of the solutions is known or available. The operator \( B \) 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,

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

The physics and data parts could be weighted through several strategies, as discussed in the review paper. 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 \( \lambda \) defined as,

\[
\lambda = \frac{1}{\epsilon_2(\nabla \cdot u) - |\nabla \cdot u| + 1}
\]

where, \( \epsilon_2 \) is a hyperparameter, and \( \nabla \cdot u \) represents the divergence of \( u \). Subsequently, the physics based term becomes \( \frac{1}{\epsilon_2(\nabla \cdot u) - |\nabla \cdot u| + 1} ||D[u(x, t)]||^2 \). In addition, a hyperparameter \( \epsilon_1 \) 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

<table>
<thead>
<tr>
<th>Method</th>
<th>Convex $M = 0.5$</th>
<th>Non-Convex $M = 1$</th>
<th>Non-Convex $M = 2$</th>
<th>Non-Convex $M = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PINN-WE</td>
<td>1.05</td>
<td>2.96</td>
<td>4.70</td>
<td>0.17</td>
</tr>
<tr>
<td>Learnable global AV</td>
<td>3.05</td>
<td>6.72</td>
<td>7.63</td>
<td>6.88</td>
</tr>
<tr>
<td>Parametric AV map</td>
<td>3.33</td>
<td>6.93</td>
<td>6.42</td>
<td>5.94</td>
</tr>
<tr>
<td>Residual-based AV map</td>
<td>3.34</td>
<td>8.70</td>
<td>10.04</td>
<td>8.33</td>
</tr>
<tr>
<td>Nonadaptive global AV</td>
<td>4.50</td>
<td>6.55</td>
<td>7.23</td>
<td>7.05</td>
</tr>
</tbody>
</table>

PINN-WE is defined as follows,

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

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 $L_{\text{PINN}}$. Traditionally, in PINN literature, the hyperparameter choice of $\lambda$ and $\epsilon_1$ 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 ($\lambda$) 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,

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

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,

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

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 $\text{tanh}$ activation function with L-BFGS-B 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,

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

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

<table>
<thead>
<tr>
<th>Convex</th>
<th>Non-Convex</th>
<th>Convex</th>
<th>Non-Convex</th>
<th>Convex</th>
<th>Non-Convex</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M = 0.5$</td>
<td>$M = 10$</td>
<td>$M = 0.5$</td>
<td>$M = 1$</td>
<td>$M = 2$</td>
<td>$M = 10$</td>
</tr>
<tr>
<td>$\epsilon_1$</td>
<td>0.1</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>$\epsilon_2$</td>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

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:

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

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

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

Additionally, $u(\frac{x}{t})$ is defined for $\frac{x}{t} \leq f'_w(u^*)$ as $u(\frac{x}{t}) = (f'_w)^{-1}(\frac{x}{t})$. Due to its self-similarity, the analytical solution 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. Finally, Table 2 presents the choice of hyperparameters $\epsilon_1$ and $\epsilon_2$ 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


