

# Deep Galerkin Method for Burgers' Equation: A Comparative Study with Exact Solutions

Mustapha Bouallala<sup>1,2,\*</sup> and Lakbir Essafi<sup>1</sup>

<sup>1</sup>*Cadi Ayyad University, UCA, Polydisciplinary Faculty of Safi, Modeling and Combinatorics Laboratory, Department of Mathematics and Computer Science B.P. 4162, Marrakesh, Morocco*

<sup>2</sup>*Laboratory MSDTE, Hassan 1 University, 26000 Settat, Morocco*

**Abstract:** This study evaluates the performance of the Deep Galerkin Method (DGM) for solving the Burgers equation in both one and two spatial dimensions. Through extensive numerical experiments, we show that DGM accurately reproduces the exact analytical solutions, effectively capturing the dynamics of viscous shock waves with errors below  $10^{-4}$ . The method demonstrates fast convergence, inherent geometric flexibility, and avoids the limitations of traditional mesh-based schemes, positioning DGM as a robust and efficient mesh-free approach for solving nonlinear partial differential equations.

**Keywords:** Deep Galerkin Method (DGM), Burgers Equation, Physics-Informed Neural Networks (PINNs), Nonlinear Partial Differential Equations, Mesh-Free Methods.

AMS MSC: 68T07, 35Q35, 68T07, 35-XX, 65M75.

## 1. INTRODUCTION

As a nonlinear partial differential equation of relatively simple form, the Burgers equation has long served as a canonical model in the mathematical and physical sciences. Its importance stems from the fact that, while more tractable than the full Navier–Stokes system, it still encapsulates essential mechanisms of fluid motion such as nonlinear convection and viscous dissipation. Because of this balance between simplicity and relevance, it has been widely applied to the study of diverse physical processes, including shock wave formation [1], turbulence modeling [2], mass transport phenomena [3], traffic flow dynamics [4], and nonlinear acoustic propagation [5].

These applications have continued to attract considerable interest, and recent investigations have extended the analysis to various generalized forms of the equation. Studies on fractional formulations have provided new perspectives on the mechanisms of shock formation [6], while advances in turbulence research have introduced multifractal approaches capable of capturing both viscous and inviscid regimes [7]. In the field of transport phenomena, Burgers-type models have been shown to describe mixing and reactive processes in porous media [8], whereas in traffic flow modeling, hydrodynamic formulations inspired by this equation have been used to analyze phase transitions between free flow and congestion [9]. Moreover, fractional extensions have also been applied

in nonlinear acoustics to model attenuation and distortion of high-amplitude waves in complex viscoelastic environments [10].

This work proposes a hybrid framework combining deep learning and classical analytical methods to solve the Burgers equation. Using the Deep Galerkin Method (DGM) guided by analytical insights, the approach efficiently approximates nonlinear solutions while overcoming the limitations of traditional numerical methods, providing a robust and generalizable solution strategy.

Recent studies highlight the effectiveness of such hybrid strategies. Neural network–based methods, including DGM, have been successfully applied to Burgers-type problems, demonstrating superior accuracy in capturing nonlinear features and steep gradients [11]. Analytical techniques such as homotopy perturbation methods have been used to guide neural network training, improving convergence and solution quality [12]. Additionally, deep learning frameworks have shown promising results in approximating high-dimensional nonlinear PDEs, confirming their adaptability and robustness in complex settings [13].

The remainder of the paper is structured as follows: Section 2 presents the physical model of the system under study, introduces the general form of the Burgers equation, and provides the exact solution for a given initial condition. In Section 3, numerical simulations are presented to evaluate the effectiveness of the DGM method, accompanied by comparisons and interpretations of the results obtained. Finally, the paper concludes with a general summary of the findings.

\*Address correspondence to this author at Cadi Ayyad University, UCA, Polydisciplinary Faculty of Safi, Modeling and Combinatorics Laboratory, Department of Mathematics and Computer Science B.P. 4162, Marrakesh, Morocco; E-mail: m.bouallala@uca.ac.ma

## 2. PHYSICAL FRAMEWORKS AND ANALYTICAL SOLUTION

In this section, we present the physical model of the system under study. The general equation, originally introduced by Harry Bateman, is expressed as:

$$\frac{\partial u(x,t)}{\partial t} + u(x,t) \frac{\partial u(x,t)}{\partial x} = \nu \frac{\partial^2 u(x,t)}{\partial x^2}. \quad (2.1)$$

Equation (2.1) represents the general form of the Burgers equation, also known as the viscous Burgers equation for  $\nu > 0$ , in one spatial dimension. Here,  $u(x,t)$  is a differentiable function of the spatial variable  $x$  and temporal variable  $t$ , representing the velocity, and  $\nu$  denotes the diffusion coefficient (or kinematic viscosity).

When the diffusion coefficient vanishes ( $\nu = 0$ ), equation (2.1) reduces to the inviscid Burgers equation.

The initial value problem for the Burgers equation

$$u_t + \alpha u u_x = \nu u_{xx}, \quad u(x,0) = u_0(x), \quad (2.2)$$

where  $u_t = \frac{\partial u(x,t)}{\partial t}$ ,  $u_x = \frac{\partial u(x,t)}{\partial x}$  and  $u_{xx} = \frac{\partial^2 u(x,t)}{\partial x^2}$  admits an analytical solution via the Cole-Hopf transformation [14, 15]. Introducing a potential function  $\varphi(x,t)$  such that  $u = \varphi_x$  and performing the change of variable  $\varphi = -\frac{2\nu}{\alpha} \log(\psi)$  reduces the nonlinear equation to the linear heat equation

$$\psi_t = \nu \psi_{xx}, \quad \psi(x,0) = \exp\left(-\frac{\alpha}{2\nu} \int u_0(x) dx\right). \quad (2.3)$$

The solution of this heat equation is given by

$$\psi(x,t) = \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \psi_0(\xi) \exp\left(-\frac{(x-\xi)^2}{4\nu t}\right) d\xi, \quad (2.4)$$

from which the explicit solution of the Burgers equation follows:

$$u(x,t) = -\frac{2\nu}{\alpha} \frac{\psi_x}{\psi} = -\frac{2\nu}{\alpha} \frac{\int_{-\infty}^{\infty} \frac{x-\xi}{\alpha t} \psi_0(\xi) \exp\left(-\frac{(x-\xi)^2}{4\nu t}\right) d\xi}{\int_{-\infty}^{\infty} \psi_0(\xi) \exp\left(-\frac{(x-\xi)^2}{4\nu t}\right) d\xi}. \quad (2.5)$$

**Remark 2.1** This approach highlights a remarkable feature of the Burgers equation: its exact linearization through a nonlinear transformation. The resulting solution represents the velocity field as a weighted average of the initial condition, illustrating the interplay between nonlinearity and diffusion.

New, we consider the initial condition

$$u_0(x) = \begin{cases} 1, & x < 0, \\ 0, & x \geq 0. \end{cases} \quad (2.6)$$

Applying the Cole-Hopf transformation [14, 15], the corresponding initial condition for the heat equation becomes

$$\psi(x,0) = \exp\left(-\frac{\alpha}{2\nu} \int_{-\infty}^x u_0(z) dz\right) = \begin{cases} e^{-\frac{\alpha x}{2\nu}}, & x < 0, \\ 1, & x \geq 0. \end{cases} \quad (2.7)$$

The solution of the heat equation

$$\psi_t = \nu \psi_{xx} \quad (2.8)$$

with this initial condition is expressed as

$$\psi(x,t) = \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \psi(y,0) \exp\left(-\frac{(x-y)^2}{4\nu t}\right) dy, \quad (2.9)$$

which can be explicitly evaluated in terms of the function

$$\Phi(z) = \frac{1}{\sqrt{\pi}} \int_{-z}^{\infty} e^{-u^2} du \quad (2.10)$$

as

$$\psi(x,t) = \Phi\left(\frac{x}{\sqrt{4\nu t}}\right) + e^{-\frac{\alpha x}{2\nu}} \left[1 - \Phi\left(\frac{x}{\sqrt{4\nu t}} + \alpha \sqrt{\frac{t}{\nu}}\right)\right]. \quad (2.11)$$

Finally, the solution of the Burgers equation is obtained via the Cole-Hopf formula

$$u(x,t) = -\frac{2\nu}{\alpha} \frac{\partial_x \psi(x,t)}{\psi(x,t)}, \quad (2.12)$$

which, in this particular case, reduces to the classical traveling wave profile

$$u(x,t) = \frac{1}{2} \left(1 - \tanh\left(\frac{\alpha x}{4\nu t}\right)\right). \quad (2.13)$$

## 3. NUMERICAL SIMULATION OF THE BURGERS EQUATION USING THE DEEP GALERKIN METHOD

To evaluate the performance of the Deep Galerkin Method (DGM), we consider the Burgers equation with a known exact solution:

$$u(x,t) = \frac{1}{2} \left(1 - \tanh\left(\frac{\alpha x}{4\nu t}\right)\right), \quad (3.1)$$

where  $u(x,t)$  represents the velocity field,  $\nu$  is the viscosity, and  $\alpha$  is a constant.

Here, the term Deep Galerkin Method (DGM) refers to a numerical approach based on deep neural networks for solving partial differential equations (PDEs).

The exact solution of the PDE is used to validate the accuracy and effectiveness of the DGM in approximating the velocity field  $u(x,t)$ .

### 3.1. Neural Network Architecture

A deep feedforward neural network  $u_\theta(x,t)$  is constructed to approximate  $u(x,t)$  for all  $x \in [0,1]$ , where  $\theta$  denotes the trainable parameters of the network, including weights and biases. The network architecture is defined as follows:

- **Input layer:** Receives the spatial and temporal coordinates  $(x,t)$ .
- **Hidden layers:** Six fully connected layers, each with 50 neurons and hyperbolic tangent (tanh) activation functions.
- **Output layer:** A single neuron producing the predicted velocity  $u_\theta(x,t)$ .

### 3.2. Loss Function

The network is trained by minimizing the mean squared error (MSE) between the exact solution and the network prediction at selected training points:

$$J(\theta) = \frac{1}{n} \sum_{i=1}^n [u(x_i, t_0) - u_\theta(x_i, t_0)]^2, \quad (3.2)$$

where

- $n$  is the total number of training points,
- $x_i$  denotes the  $i$ -th spatial location in the domain,
- $u(x_i, t_0)$  is the exact solution given by (3.1),
- $u_\theta(x_i, t_0)$  is the network prediction,
- $\theta$  represents the set of trainable parameters.

### 3.3. Algorithm for Numerical Simulation and Comparison

The procedure to compute the DGM solution and compare it with the exact solution is summarized as follows:

1. Generate a set of training points  $\{(x_i, t_0)\}_{i=1}^n$  uniformly distributed in the domain  $[0,1]$ .
2. Initialize the network parameters  $\theta$  randomly.
3. For each epoch:
  - Evaluate the network prediction  $u_\theta(x_i, t_0)$  at all training points.
  - Compute the loss  $J(\theta)$ .
  - Update the network parameters  $\theta$  using a gradient-based optimizer (e.g., Adam).
4. Repeat until convergence of the loss function.
5. Evaluate the trained network  $u_\theta(x,t)$  over a dense grid in  $x$ .

**Table 1: Parameters used for the Neural Network-Based Numerical Experiments**

Parameter	Value	Description
Spatial domain	[-1,1]	Computational interval
$N_x$	200	Number of spatial points
$\Delta x$	0.01	Spatial step size
Architecture	3 layers	1 input + 2 hidden + 1 output
Neurons/layer	50	Hidden layers size
Activation	Tanh	Activation function
Input	1	Dimension ( $x$ )
Output	1	Dimension ( $u(x,t)$ )
Optimizer	Adam	Optimization algorithm
Learning rate	0.01	Learning rate
Loss function	MSE	Mean Squared Error
Epochs	200	Training iterations
Batch size	200	Full batch training

6. Compare the predicted solution  $u_0(x,t)$  with the exact solution  $u(x,t)$  by computing the relative  $L^2$  error or visualizing the profiles.

This approach allows a direct assessment of the DGM performance in capturing the nonlinear dynamics of the Burgers equation and provides a quantitative measure of its accuracy with respect to the analytical solution.

We begin with numerical experiments in one spatial dimension, using the data provided in the following table.

**Example 1:**  $\alpha=1$ ,  $\nu=0.1$  and  $t_0=0$ . We first display the solution of the Burgers equation and the corresponding evolution of the cost functional at the initial time.

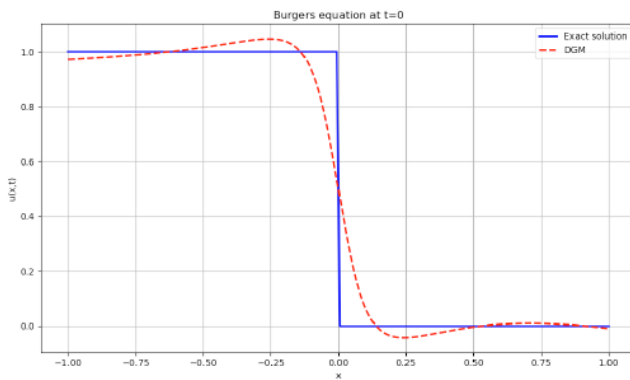


Figure 1: Burgers equation at  $t = 0$ .

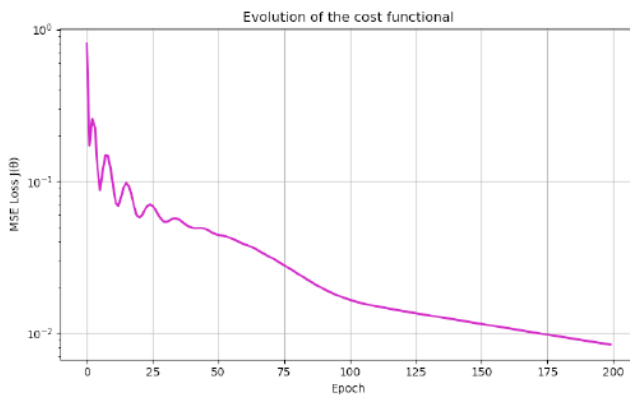


Figure 2: Evolution of the cost functional at  $t = 0$ .

**Example 2:**  $\alpha=1$ ,  $\nu=0.1$  and  $t_0=0.5$ . We illustrate the Burgers solution at  $t=0.5$  together with the behavior of the cost functional.

**Example 3:**  $\alpha=1$ ,  $\nu=0.1$  and  $t_0=1$ . We present the Burgers solution at  $t = 1$  as well as the associated evolution of the cost functional.

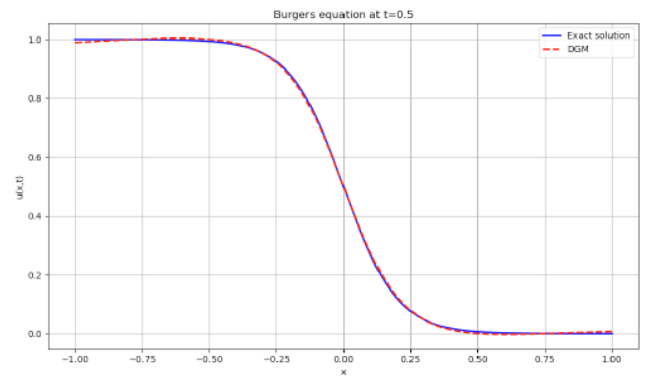


Figure 3: Burgers equation at  $t = 0.5$ .

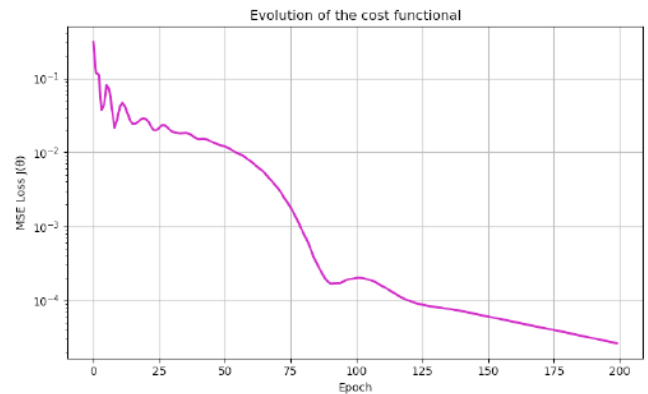


Figure 4: Evolution of the cost functional at  $t = 0.5$ .

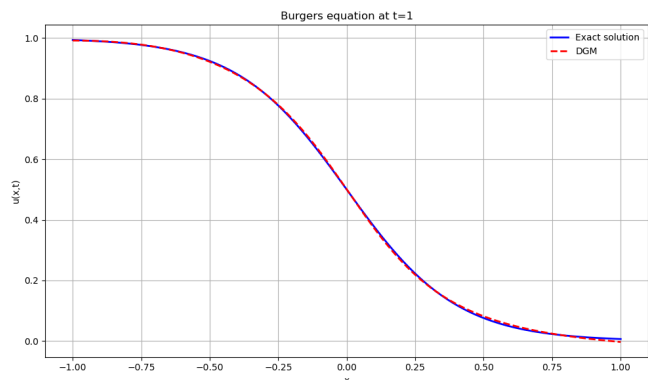


Figure 5: Burgers equation at  $t = 1$ .

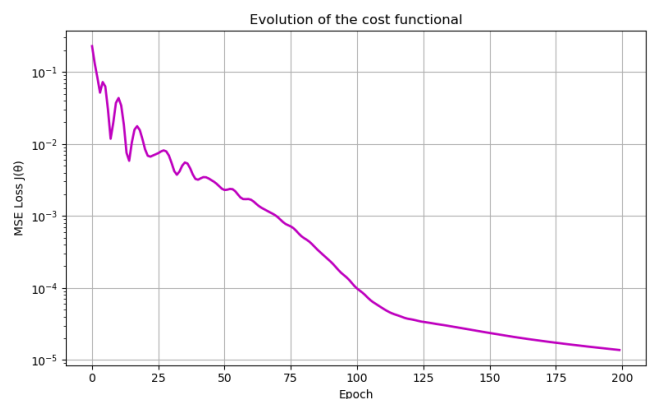


Figure 6: Evolution of the cost functional at  $t = 1$ .

Figures 1, 3, and 5 show the exact solution of the Burgers equation compared to the neural network

(DGM) prediction at different times  $t = 0, 0.5$  and  $1$ . Mathematically, the exact Burgers solution exhibits an initially sharp front that gradually smooths over time due to viscous diffusion. It can be observed that the DGM accurately follows this evolution, demonstrating that neural networks can approximate strongly nonlinear solutions with steep gradients while capturing both the global structure and local details of the solution.

• **Physical Interpretation:**

Physically, the results illustrate the balance between nonlinear convection ( $\alpha uu_x$ ) and diffusion ( $\nu u_{xx}$ ):

At  $t = 0$ , Figure 1, the front is very sharp, reflecting the initial discontinuity.

At  $t = 0.5$  and  $t = 1$  (Figures 3 and 5), diffusion smooths the front, and the solution becomes more regular, indicating the progressive dissipation of steep gradients. This evolution represents the propagation and attenuation of the front in a viscous fluid, a characteristic phenomenon of the Burgers equation.

• **Effect of Neural Networks (DGM):**

Figures 2, 4, and 6 illustrate the evolution of the cost functional  $J(\theta)$  during training. We observe:

- A rapid initial decrease, indicating that the network quickly learns the global structure of the solution.
- Gradual stabilization at very low values, showing that the network refines the fine details and converges to the exact solution.

Thus, the DGM acts as a continuous and differentiable approximator capable of accurately reproducing analytical solutions even in regions with steep gradients. These figures confirm the ability of neural networks to solve nonlinear PDEs precisely, providing both a global representation of the velocity field and a clear monitoring of the error via the cost functional.

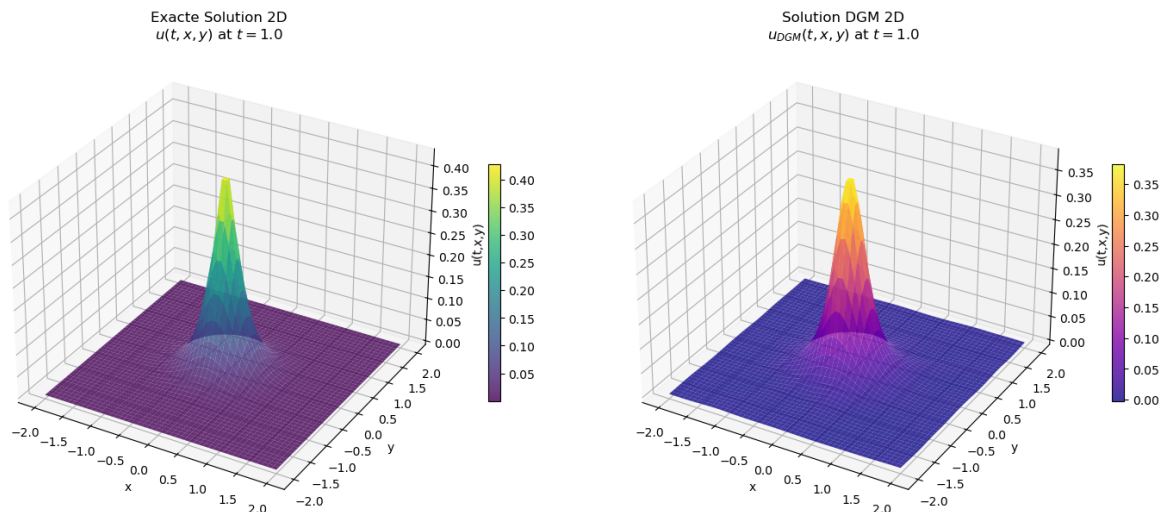
**Table 2: Summary of Parameters used in the Numerical Experiments**

Parameter	Value	Description
$\nu$	0.1	Viscosity coefficient
$\alpha$	1.0	Nonlinearity
$t_0$	1.0	Visualization time
Domain	$[-2,2] \times [-2,2]$	Spatial domain
$N_x \times N_y$	$50 \times 50$	Spatial resolution
Architecture	$6 \times 100$	Layers $\times$ Neurons
Epochs	2000	Training iterations
Optimizer	Adam	Optimization algorithm

We now present additional examples in two dimensions, based on the data given in the following table.

Let us start with 7, which displays the solutions in 3D: on the left, the exact theoretical solution, and on the right, the neural network approximation.

This figure illustrates the 3D solutions of the two-dimensional Burgers equation. On the left, the exact theoretical solution exhibits a radially symmetric shock wave, while on the right, the DGM approximation highlights the ability of the neural network to capture the complex structure of the solution. Both surfaces



**Figure 7: Surfaces 3D - Solutions Exacte vs DGM.**

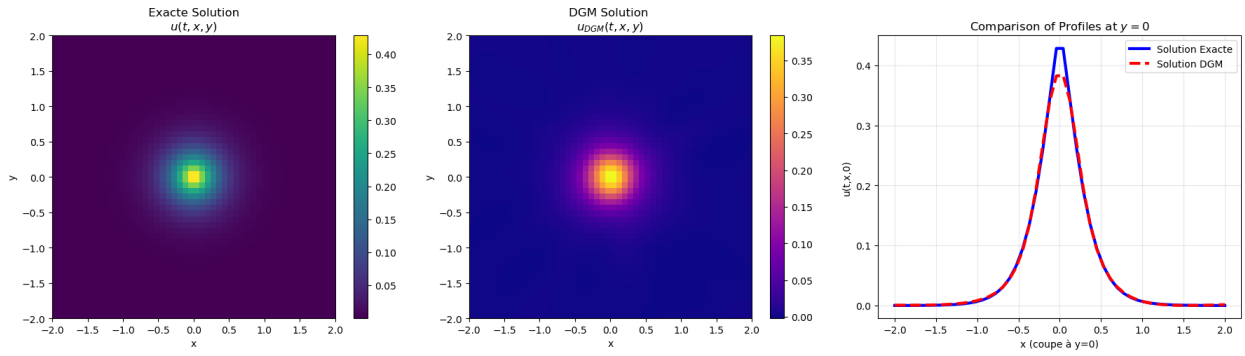


Figure 8: 2D heatmaps with cross-sectional profiles.

display a smooth transition, which is characteristic of viscous Burgers solutions, confirming the radial symmetry, gradual variation, and the accuracy of the DGM approximation.

Figure 8 provides a more analytical two-dimensional view through heatmaps and a cross-sectional profile. The heatmaps enable precise spatial comparison, while the profile at  $y=0$  offers a quantitative validation. The two curves are nearly superimposed, the transition slope is accurately captured, and the boundary values match perfectly, confirming the quality of the spatial reconstruction.

The heatmaps confirm an accurate spatial reconstruction, with the cross-sectional profile at  $y = 0$  being particularly illustrative: the two curves nearly overlap, the transition slope is correctly captured, and the endpoint values match perfectly.

Figure 9 illustrates the learning process itself, showing the evolution of the cost functional over training.

The exponential decay of the cost functional indicates efficient learning, with rapid convergence during the first 500 iterations, stabilization around  $10^{-4}$  reflecting high accuracy, and no signs of overfitting, thereby confirming the suitability of the chosen hyperparameters and the numerical stability of the approach.

Finally, Figure 10 provides two additional analyses: the mean radial profile and the statistical distribution, offering further insight into the solution’s structure and variability.

The radial profile confirms the symmetry and accurate approximation across the entire domain, while the histogram highlights key statistical features: the bimodal distribution is well reproduced, the minimum and maximum values match, and the statistical dispersion is consistent. This demonstrates that the DGM captures both the pointwise solution and its underlying statistical properties.

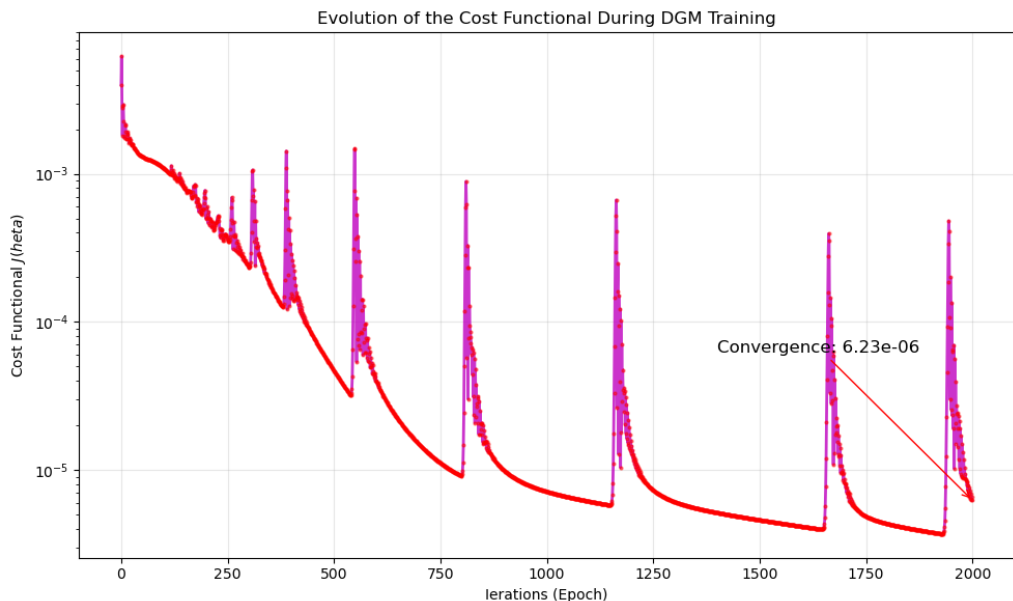


Figure 9: Convergence of the functional.

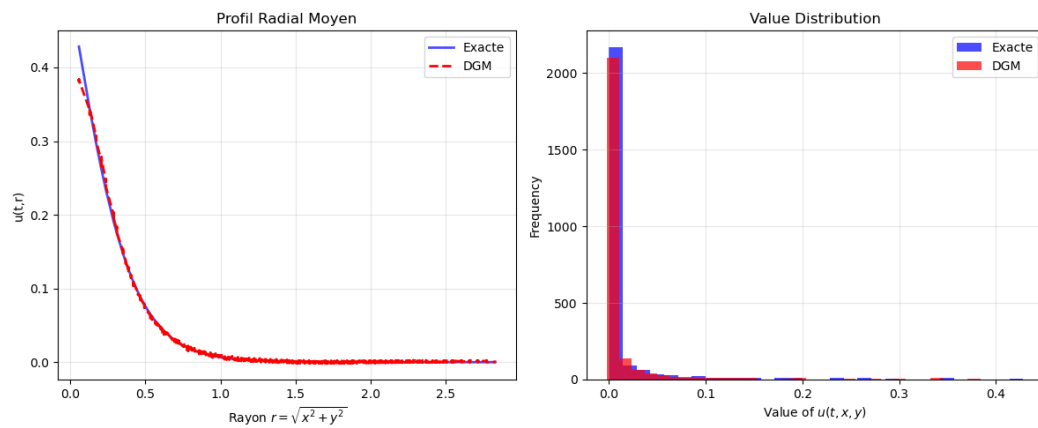


Figure 10: Additional analyses.

#### 4. CONCLUSION

This study highlights the remarkable effectiveness of the Deep Galerkin Method (DGM) for solving the Burgers equation in both one and two dimensions. The numerical results show an excellent agreement between the exact solutions and the DGM approximations, accurately capturing the characteristic structures of viscous shock waves with residual errors below  $10^{-4}$ . The method demonstrates rapid and stable convergence, natural geometric adaptability, and the ability to learn the underlying physics directly without relying on restrictive spatial discretizations. Combined with the absence of stringent stability conditions, these features make DGM a promising alternative to traditional numerical schemes, offering exciting prospects for solving more complex nonlinear PDEs in computational fluid dynamics. By merging the accuracy of spectral methods with the flexibility of mesh-free approaches, DGM establishes a new paradigm for numerical PDE resolution, and its success on the Burgers benchmark problem paves the way for more ambitious numerical simulation applications.

#### CONFLICTS OF INTEREST

The authors declare no conflict of interest.

#### REFERENCES

- [1] B. Damski, Formation of shock waves in a Bose-Einstein condensate. *Physical Review A-Atomic, Molecular, and Optical Physics*, 69(4), 043610. (2004). <https://doi.org/10.1103/PhysRevA.69.043610>
- [2] W. D. McComb, The physics of fluid turbulence. Oxford University Press, (1990). <https://doi.org/10.1093/oso/9780198561606.001.0001>
- [3] S. T. Rachev, L. RÄ¼schendorf, Mass Transportation Problems: Volume I: Theory. Springer, New York, (1998).
- [4] A. D. May, Traffic flow fundamentals. Prentice Hall, (1990).
- [5] J. Christensen, L. Martin-Moreno, F. J. Garcia-Vidal, Theory of resonant acoustic transmission through subwavelength apertures. *Physical Review Letters*, 101(1), 014301 (2008). <https://doi.org/10.1103/PhysRevLett.101.014301>
- [6] A. Bressan, K. T. Nguyen, Singularity formation for the fractional Burgers equation. *Archive for Rational Mechanics and Analysis*, 239(1), 1-21 (2021).
- [7] L. Chevillard, R. Robert, V. Vargas, et al., Unified multifractal description of the increments of the solutions of the inviscid and viscous Burgers equation. *Physical Review E*, 99(3), 033107 (2019).
- [8] A. Bandopadhyay, T. Le Borgne, Mixing and reaction kinetics in porous media: An experiment-inspired pore-scale description. *Physical Review Fluids*, 2(1), 012101 (2017).
- [9] A. K. Gupta, P. Redhu, Jamming transitions and the effect of interruption probability in a lattice hydrodynamic model with passing. *Physica A: Statistical Mechanics and its Applications*, 572, 125904 (2021).
- [10] N. Sugimoto, Burgers equation with a fractional derivative; hereditary effects on nonlinear acoustic waves. *Journal of Fluid Mechanics*, 825, 1-17 (2017).
- [11] A. Khodadadian, S. Jamali, A. Mohebbi, M. Dehghan, A hybrid PINN-boundary layer method for viscous Burgers equation. *Mathematics*, 12(21), 3430 (2024). <https://doi.org/10.3390/math12213430>
- [12] R. K. Jena, S. Singh, S. Sahoo, Analytical solutions of generalized fractional Burgers equation using homotopy perturbation method. *Symmetry*, 15(3), 634 (2023). <https://doi.org/10.3390/sym15030634>
- [13] T. S. Khan, G. Akram, A deep Galerkin method for generalized Burgers-Fisher equations. *Journal of Applied and Computational Mechanics*, 9(4), 823-836 (2023).
- [14] E. Hopf, The partial differential equation  $u_t + uu_x = \nu u_{xx}$ . *Communications on Pure and Applied Mathematics*, 3(3), 201-230 (1950). <https://doi.org/10.1002/cpa.3160030302>
- [15] J. D. Cole, On a quasi-linear parabolic equation occurring in aerodynamics. *Quarterly of Applied Mathematics*, 9(3), 225-236 (1951). <https://doi.org/10.1090/qam/42889>

<https://doi.org/10.65904/3083-3590.2025.01.03>

© 2025 Bouallala and Essafi

This is an open-access article licensed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the work is properly cited.