# Evaluating Uncertainty Quantification approaches for Neural PDEs in scientific application

**Anonymous Author(s)**
Affiliation
Address
email

## Abstract

The accessibility of geographically dispersed data, enabled by affordable sensors, field and numerical experiments, has led to an abundance of data that can be used to develop data-driven solutions for scientific problems. Neural Partial Differential Equations (PDEs), which employ deep learning (DL) techniques alongside domain expertise in PDEs for parameterization, have proven to be effective in capturing valuable correlations within spatiotemporal datasets. However, data noise and sparsity of measurements coupled with model overparameterization introduce aleatoric and epistemic uncertainties. Therefore, quantifying uncertainties propagated from model inputs to outputs remains a challenge and an essential goal for establishing the trustworthiness of these method's predictions. This work evaluates various Uncertainty Quantification (UQ) approaches, which are crucial for both Forward and Inverse Problems in scientific applications. Specifically, we investigate the effectiveness of Bayesian methods, such as Hamiltonian Monte Carlo (HMC) and Monte-Carlo Dropout (MCD), and a more conventional approach, Deep Ensembles (DE). To illustrate their performance, we take two canonical PDEs: Burger's equation and the Navier-Stokes equation. Our results indicate that these approaches accurately reconstruct flow systems and predict the associated parameters. However, it is noteworthy that results derived from Bayesian methods, in our observation, tend to display a higher degree of certainty in predictions than warranted, as compared to those obtained using the Deep Ensembles (DE) method. This elevated certainty in predictions implies that Bayesian techniques might underestimate the actual uncertainty present in the data, thereby appearing more confident in their predictions than the DE approach.

## 1 Introduction

The abundant geographically dispersed data facilitated by affordable sensors, numerical and field experiments, and satellite imagery provided a unique opportunity to tackle ongoing challenges in climate change, weather prediction, and resilient urban development. However, due to their relatively low sampling density, the spatiotemporal measurements are limited in providing a comprehensive view of complex flow systems.

While conventional deep learning-based approaches provide promising solutions, they often struggle to satisfy physical constraints effectively. Physics-Informed Neural Networks (PINNs), as introduced by Raissi et al. [2019], incorporate physical constraints within the neural network optimization process, leading to improved physical realizability of the solution. However, due to noisy and limited data, the accuracy of these models can degrade significantly. [e.g., He and Jiang, 2023]

Uncertainty quantification (UQ) in scientific machine learning provides promising avenues to tackle challenging problem arising due to various factors, such as the stochastic nature of scientific processes,

Submitted to 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Do not distribute.

model overparameterization, and data noise. This study aims to compare various UQ techniques, including Hamiltonian Monte Carlo (HMC), Monte Carlo Dropouts (MCD), and Deep Ensembles (DE) and the robustness of DL techniques. We apply these methods to forward and inverse problems in two canonical PDEs - Burger's and the Navier-Stokes equations, illustrating their performance in reconstructing flow systems and predicting unknown parameters from sparse and noisy measurements.

## 2 Forward Problems

Consider a parameterized and non-linear PDE that characterizes the behavior of a physical system, defined as

$$\mathcal{L}_{\mathbf{x}}[\mathbf{u};\lambda] = \mathbf{f}(\mathbf{x},t), \mathbf{x} \in \Omega, t \in [0,T], \tag{1}$$

where $\mathbf{u}(\mathbf{x},t)$ denotes the latent state (aka solution field), the $\mathcal{L}_{\mathbf{x}}[.;\lambda]$ is a general differential operator parameterized by $\lambda$, $\mathbf{f}(\mathbf{x},t)$ is the forcing term which refers to any external influences on the system, while $\Omega \subset \mathbb{R}^D$ is the bounded domain in a d-dimensional physical space.

Given this framework and noisy measurements of $\mathbf{u}(\mathbf{x},t), \mathbf{f}(\mathbf{x},t)$, the goal is to infer the latent state $\mathbf{u}(\mathbf{x},t)$ of the dynamical system. In forward problems, PINNs as well as their Bayesian variants B-PINNs are typically used as surrogates $\widetilde{\mathbf{u}}(\mathbf{x},t;\theta)$, to infer either point estimates or posterior distributions of this latent state. In the Bayesian framework, the parameters $\theta$ of the surrogates have a prior distribution $P(\theta)$ and its formulation is defined as:

$$\widetilde{\mathbf{f}}(\mathbf{x},t;\theta) := \mathcal{L}_{\mathbf{x}}[\widetilde{\mathbf{u}}(\mathbf{x},t;\theta);\lambda] \tag{2}$$

$P(\mathcal{D}|\theta)$ represents the likelihood while the Bayes' Theorem estimates the final posterior distribution.

$$p(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})} \cong P(\mathcal{D}|\theta)P(\theta) \tag{3}$$

To approximate the posterior distribution, we employ both Bayesian methods like HMC and MCD as well as deterministic DE approach. HMC is an efficient Markov Chain Monte Carlo (MCMC) sampling method that uses concepts from Hamiltonian Dynamics and utilizes momentum variables to guide the proposals in the Markov chain, which can lead to faster convergence and better exploration of the target distribution. Given the continuous nature of Hamiltonian dynamics, leapfrog integration is used as a numerical technique to discretize and update the momentum and position variables in a staggered manner over discrete time steps. In our Bayesian methodology, we posit an independent Gaussian distribution as the prior $P(\theta)$. For HMC, parameters for Burger's (Navier-Stokes) equation include a leapfrog step of 50 (50), an initial time step of 0.1 (0.01), 1000 (5000) burn-in steps, and a sampling size of 100 (100). With DE, we assemble an ensemble of PINNs equivalent in number to the HMC samples, set at 100 (200). For MCD, we induce variance by sporadically dropping neurons at a 1% (1%) dropout rate during each training iteration. To gauge prediction uncertainty, we execute 100 (200) inferences with HMC. For DE, we acquire 100 (200) predictions from each ensemble member, and for MCD, we undertake forward network propagation 100 (200) times, maintaining the established dropout rate.

### 2.1 1-D Burger's Equation

Burger's equation is a PDE that represents a combination of diffusion and convection processes. It has wide applications in various scientific domains, including traffic flow modeling, acoustics, and sound propagation, and material transport in porous media. here, we consider a one-dimensional Burger's equation with Dirichlet boundary condition and sinusoidal initial conditions as follows:

$$\begin{gathered}\frac{\partial u}{\partial t} + u \cdot \nabla u - \frac{0.01}{\pi}\nabla^2 u = 0, \;\; x \in [-1,1], \; t \in [0,1], \\ u(0,x) = -sin(\pi x), \\ u(t,-1) = u(t,1) = 0, \end{gathered} \tag{4}$$

where $x$ and $t$ is the spatial location and time, $u(x,t)$ is the velocity of the fluid, and $\nabla$ and $\nabla^2$ represents gradient and Laplacian operators.

Chebfun package has been employed to exact solution and for more details, please refer Rico-Martinez et al. [1994] and Raissi et al. [2019]. Here, we assume an unknown exact solution and instead use

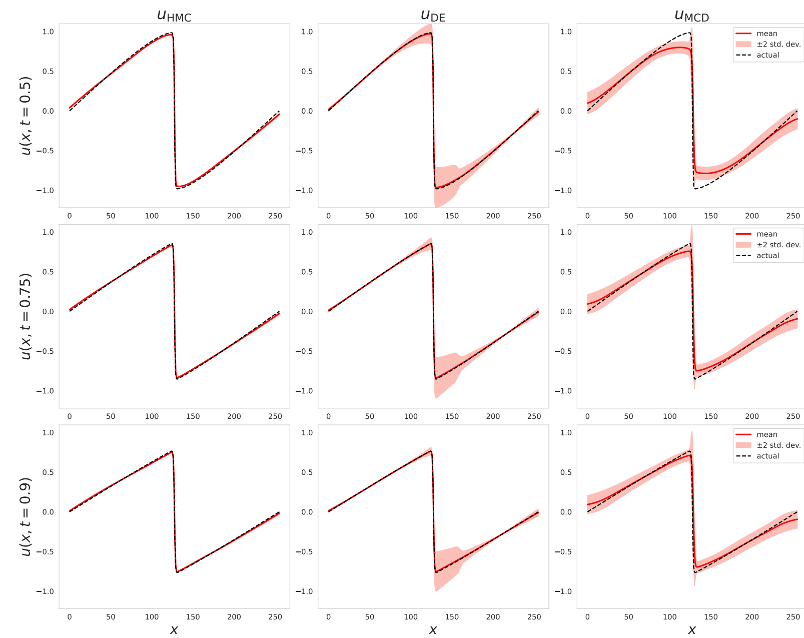



Figure 1: One-dimensional Burgers equation - forward problem: comparison of the predicted and exact solutions corresponding to the three temporal snapshots denoted by $t \in \{0.50s, 0.75s, 0.90s\}$ from different methods.

noisy sensors to collect 2000 spatiotemporal readings for $u$ and $f$. The sensor measurements have Gaussian noise with scales $\epsilon_f$ and and $\epsilon_u$ as $\mathcal{N}(0, 0.1^2)$. A multilayer perceptron (MLP) neural network consisting of eight hidden layers, each comprising 20 neurons with tanh non-linearity is employed to approximate the solution.

The predictive means ($\mu$) and the corresponding two standard deviations ($\pm 2\sigma$) using three different methods, HMC, MCD, and DE at three distinct time snapshots, particularly $t \in \{0.50s, 0.75s, 0.90s\}$ is illustrated in figure 1. From visual inspection, it is evident that both the HMC and DE approaches provide reasonably accurate posterior estimations of the variable $u$ at all three time-snapshots. Moreover, the error between these predictive means and the actual solution remains predominantly within $\pm 2\sigma$. In contrast, the MCD approach exhibits discrepancies from the actual solution across all temporal snapshots, although these discrepancies tend to diminish as time progresses. It is noteworthy that, for $t = 0.50s$, a significant portion of the error falls outside the two standard deviation confidence intervals. However, as time advances, the performance of the MCD approach noticeably improves. It is also important to highlight that all three approaches effectively capture the formation of shocks, a challenging task even for classical numerical methods.

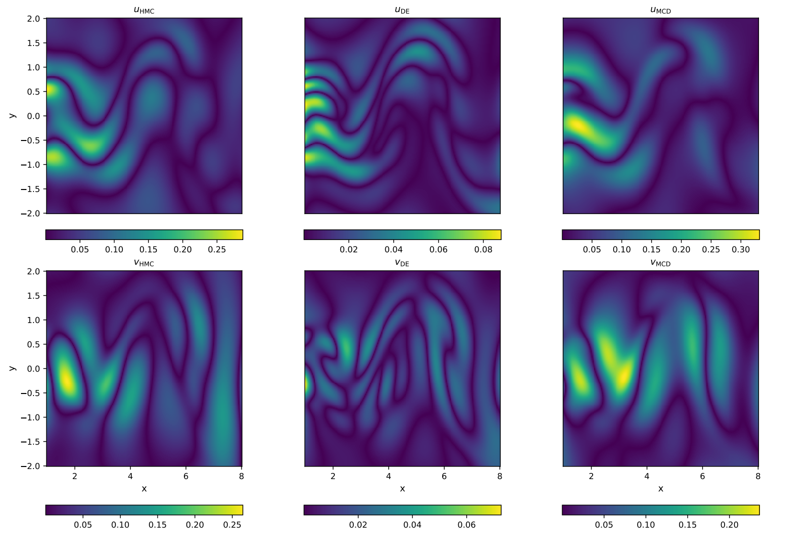
### 2.2 2-D Navier Stokes Equation

In our next example, we explore the practical scenario of incompressible fluid flow described by the Navier-Stokes equations. These equations are fundamental in science and engineering, with applications in various fields like climate prediction, aerodynamics, and blood circulation. An incompressible flow past a cylinder case is considered, and the associated governing equation are defined as follows:

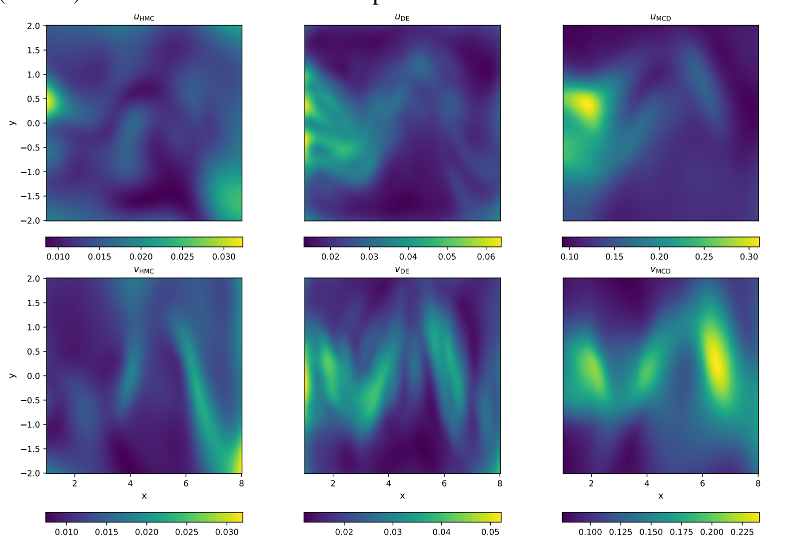
$$\frac{\partial \mathbf{u}}{\partial t} + \lambda_1 \mathbf{u} \cdot \nabla \mathbf{u} + \nabla \mathbf{u} - \lambda_2 \nabla^2 u = 0, \tag{5}$$

$$\nabla \cdot \mathbf{u} = 0,$$

where $\mathbf{u} = \{u, v\}$ and $p$ are the 2-D velocity and pressure fields, $t$ is time, and $\lambda = \{\lambda_1, \lambda_2\}$ are the parameters and for the forward problems $\lambda_1$ is set to 1 and $\lambda_2$ to $10^{-2}$. Given the multidimensional nature of this problem, it offers a challenging testbed for the Bayesian approach to quantify uncertainties in both the $\mathbf{u}$ and $p$ fields. It is important to emphasize that $p$ measurements are not included in




(a) Predictive errors for velocity component in $x$-direction $u$ (top) and $y$-direction $v$ (bottom) from different methods at a representative time instant.




(b) Two standard deviations $(\sigma)$ for velocity component in $x$-direction $u$ (top) and $y$-direction $v$ (bottom) from different methods at a representative time instant.

Figure 2: Navier-Stokes equation - forward problem.

the model training; instead, the neural network predicts them based on the governing equation. To generate the exact solutions, we leverage the data provided for the work by Raissi et al. [2019], and readers are advised to refer to the same for more details. Similar to section 2.1, noisy sensors capture 5000 spatiotemporal readings for both $u$ and $f$ with Gaussian noise $\mathcal{N}(0, 0.1^2)$. A 10-layer MLP network with 20 neurons per layer and a tanh non-linearity is used to approximate the latent variables ($u$, $v$, and $p$).

$\mathrm{L}_1$ norm-based error $\epsilon$ between the actual and predictive $\mu$ values and $\pm 2\sigma$ are presented in figure 2a. Notably, the DE approach exhibits the closest agreement with the actual solutions. In contrast, the $\epsilon_{\mathrm{HMC}}$ and $\epsilon_{\mathrm{MCD}}$ approaches are roughly three times higher than $\epsilon_{\mathrm{DE}}$ for $\mathbf{u}$ field. The $\epsilon$ consistently remains within the $\pm 2\sigma$ for all considered methodologies, as illustrated in figure 2b and underscores our confidence in the predictions generated using various approaches, as they remain well within the established confidence interval.

## 3 Inverse Problems

Inverse problems involve determining a system's underlying parameters $\lambda$ and physical properties from observable data. This study offers a systematic approach to quantify uncertainties in estimating unknown parameters for the Navier-Stokes equation (5). Similar to the framework described in equations [2-3] **update this**, apart from a surrogate for $\theta$, we also assign a prior distribution for $\lambda$, which can be independent of the prior for $\theta$. The likelihood is then defined as $P(\mathcal{D}|\theta, \lambda)$, and we then calculate the joint posterior of $[\theta, \lambda]$:

$$p(\theta, \lambda|\mathcal{D}) = \frac{P(\mathcal{D}|\theta, \lambda)P(\theta, \lambda)}{P(\mathcal{D})} \cong P(\mathcal{D}|\theta, \lambda)P(\theta, \lambda) = P(\mathcal{D}|\theta, \lambda)P(\theta)P(\lambda) \tag{6}$$

| | HMC | DE | MCD |
|---|---|---|---|
| $\lambda_1$ (mean $\pm$ std) | $0.758 \pm 0.0$ | $0.957 \pm 0.024$ | $0.843 \pm 0.075$ |
| $\lambda_2$ (mean $\pm$ std) | $0.017 \pm 2.13\text{e}{-06}$ | $0.014 \pm 0.001$ | $0.015 \pm 0.058$ |

Table 1: Navier Stokes equation - inverse problem : Predictions for $\lambda_1, \lambda_2$ using HMC, DE, MCD; actual values for $\lambda_1 = 1.0, \lambda_2 = 0.01$

The primary objective is to estimate $\lambda = \{\lambda_1, \lambda_2\}$ and associated uncertainty based on the limited measurements of $f$ and $\mathbf{u} = [u, v]$ outlined in section 2.2. To do so, we employ the MLP model of ten hidden layers with 40 neurons in each layer and tanh non-linearity. The predicted values of $\lambda$'s are displayed in Table 1. The DE method has provided relatively precise estimates, reflecting a good degree of certainty in its predictions. This suggests that ensemble techniques effectively capture these parameters' underlying distributions. While HMC provides high confidence in its estimates, the absence of uncertainty is unrealistic, and this overconfidence could be a sign of the model not capturing all sources of uncertainty. MCD provides a broader uncertainty estimation, which might be capturing more sources of uncertainties, but it could also be overestimating the uncertainty in the parameters. The wider confidence intervals for MCD could either mean that MCD is being more cautious or it's not as effective in pinpointing the true parameter values. These findings underscore the effectiveness of the DE approach in not only identifying the unknown parameters but also quantifying the uncertainty arising from the sparse and noisy sensor measurements.

## 4 Summary

This study compares and evaluates various UQ approaches, particularly Bayesian and Deep Ensemble (DE) techniques. While all approaches, including DE, HMC, MCD, effectively reconstruct flow systems and predict unknown parameters for the two examples considered, Bayesian methods demonstrate higher certainty in predictions but may underestimate the total uncertainty, thereby appearing overly confident. In contrast, while offering more conservative certainty estimates, the DE method is computationally taxing. The study underscores the need for balancing predictive certainty, computational efficiency, and accuracy when using Bayesian or DE approaches for flow system modeling and parameter prediction.

## References

Maziar Raissi, Paris Perdikaris, and George E 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.

Wenchong He and Zhe Jiang. A survey on uncertainty quantification methods for deep neural networks: An uncertainty source perspective. *arXiv preprint arXiv:2302.13425*, 2023.

R Rico-Martinez, JS Anderson, and IG Kevrekidis. Continuous-time nonlinear signal processing: a neural network based approach for gray box identification. In *Proceedings of IEEE Workshop on Neural Networks for Signal Processing*, pages 596–605. IEEE, 1994.