# Application of Physics-Informed Neural Networks to Neutron Star Magnetospheres

## International Conference on Machine Learning for Astrophysics

**Petros Stefanou**

Co-authors:
Jorge Urbán
Jose Pons

July 2024

Universitat d'Alacant

BLADES

GENERALITAT VALENCIANA
Conselleria de Educación, Universidades y Empleo

MINISTERIO DE CIENCIA E INNOVACIÓN
Financiado por la Unión Europea
NextGenerationEU
Plan de Recuperación, Transformación y Resiliencia
AGENCIA ESTATAL DE INVESTIGACIÓN

MACHINE LEARNING FOR ASTROPHYSICS
2ND EDITION
CATANIA, 8-12 JULY, 2024

# Table of Contents


- Introduction
- Motivation
- Optimisation method
- Application to neutron star magnetospheres
- Summary

# Table of Contents


- Introduction
- Motivation
- Optimisation method
- Application to neutron star magnetospheres
- Summary

# Physics-informed neural networks

## 1 Introduction

## Basics

- Solve PDEs using neural networks
- Input: coordinates in some domain
- Output: an approximate solution
- Loss: the residuals of the PDE describing the system
- Training set: Large number of random points
- Boundary conditions imposed through hard-enforcement

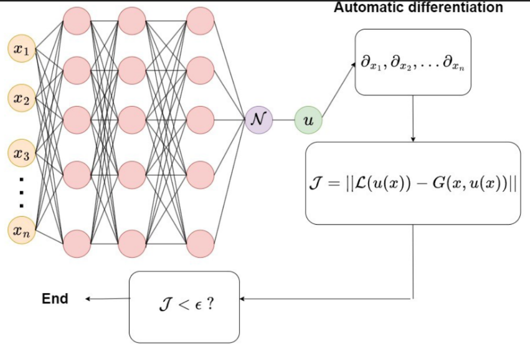


Figure 1: A physics-informed neural network

# Optimisation process

# Optimisation process

## Correction step

$$\mathbf{\Theta}_{k+1} = \mathbf{\Theta}_k + \alpha_k \boldsymbol{p}_k$$
$$\boldsymbol{p}_k = -H_k \nabla \mathcal{J}(\mathbf{\Theta}_k)$$

## Line search methods

- Trainable parameters $\mathbf{\Theta}$
- Loss function $\mathcal{J}$
- $\mathbf{\Theta}$ are adjusted so that $\mathcal{J}(\mathbf{\Theta}) \to 0$
- $\alpha_k$ determines the step size
- $\boldsymbol{p}_k$ determines the direction

# Table of Contents

2 Motivation


- Introduction
- Motivation
- Optimisation method
- Application to neutron star magnetospheres
- Summary

# Improve convergence

# Improve convergence



**PINNs vs. classical methods**

- Precision
- Efficiency

# Improve convergence

## PINNs vs. classical methods

- Precision
- Efficiency

## Ill-conditioning

- Loss function in PINNs is poorly-scaled
- Broad eigenvalue spectrum of $hess(\mathcal{J})$ close to minimum
- Condition number $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}} \gg 1$
- Some directions are flat
- Gradient-based methods **fail miserably**

# Improve convergence



## PINNs vs. classical methods

- Precision
- Efficiency

## Ill-conditioning

- Loss function in PINNs is poorly-scaled
- Broad eigenvalue spectrum of $hess(\mathcal{J})$ close to minimum
- Condition number $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}} \gg 1$
- Some directions are flat
- Gradient-based methods **fail miserably**

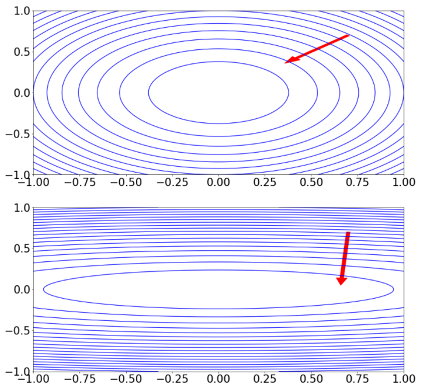


Figure 2: Contours of a loss function

# Table of Contents

3 Optimisation method


- Introduction
- Motivation
- Optimisation method
- Application to neutron star magnetospheres
- Summary

# Quasi-Newton methods

## Broyden class

$$H_{k+1} = \frac{1}{\tau_k}\left[H_k - \frac{H_k \boldsymbol{y}_k \otimes H_k \boldsymbol{y}_k}{\boldsymbol{y}_k \cdot H_k \boldsymbol{y}_k} + \phi_k \boldsymbol{v}_k \otimes \boldsymbol{v}_k\right] + \frac{\boldsymbol{s}_k \otimes \boldsymbol{s}_k}{\boldsymbol{y}_k \cdot \boldsymbol{s}_k}$$

- $H_k$ approximates the inverse Hessian of the loss
- Depends on $\boldsymbol{\Theta}_k$ and $\nabla \mathcal{J}(\boldsymbol{\Theta}_k)$
- $\tau_k = 1, \phi_k = 1$ define the BFGS optimizer
- Different $\tau_k, \phi_k$ define different methods

# The role of $H_k$

# The role of $H_k$



**Curvature information**

- $H_k$ is associated with curvature
- Step direction points towards minimum
- Less iterations to converge

## The role of $H_k$



### Curvature information

- $H_k$ is associated with curvature
- Step direction points towards minimum
- Less iterations to converge

### Preconditioning

- Define new variables $\boldsymbol{z} = H_k^{-1/2}\boldsymbol{\Theta}$
- Update rule becomes

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \alpha_k \nabla \mathcal{J}(\mathbf{z}_k)$$

- Hessian matrix in $\boldsymbol{z}$-space

$$\text{hess}\left(\mathcal{J}\left(\mathbf{z}\right)\right) = H_k^{1/2}\text{hess}\left(\mathcal{J}\left(\boldsymbol{\Theta}\right)\right) H_k^{1/2}$$

- Spectrum concentrated around $\sim 1$

# Optimisation algorithm choices

# Optimisation algorithm choices

3 Optimisation method

### Self-scaled BFGS

$$\tau_k^{(1)} = \min\left\{1, \frac{\boldsymbol{y}_k \cdot \boldsymbol{s}_k}{\boldsymbol{s}_k \cdot H_k^{-1}\boldsymbol{s}_k}\right\}$$

$$\phi_k^{(1)} = 1$$

### Self-scaled Broyden

$$\tau_k^{(2)} = \begin{cases} \tau_k^{(1)} \min\left(\sigma_k^{-1/(n-1)}, \frac{1}{\theta_k^{(1)}}\right) & \text{if } \theta_k > 0 \\ \min\left(\tau_k^{(1)}\sigma_k^{-1/(n-1)}, \sigma_k\right) & \text{if } \theta_k \leq 0 \end{cases}$$

$$\phi_k^{(2)} = \frac{1 - \theta_k}{1 + a_k\theta_k}$$

# Table of Contents




- ▶ Introduction
- ▶ Motivation
- ▶ Optimisation method
- ▶ Application to neutron star magnetospheres
- ▶ Summary

# Neutron star magnetospheres

# Neutron star magnetospheres

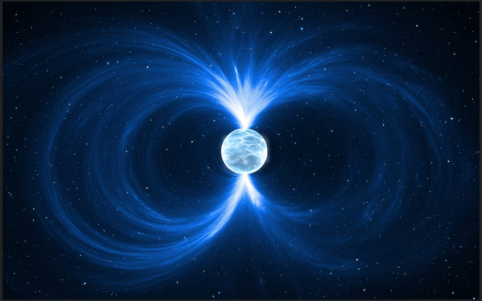

Figure 3: An artist's conception of a NS

**Features**

- Plasma flowing along the star's magnetic field
- Large scale field is dipolar
- Toroidal fields
- Force-free regime
- Non-rotating limit
- Axisymmetry

$$q^2 \frac{\partial}{\partial q}\left(q^2 \frac{\partial \mathcal{P}}{\partial q}\right) + \left(1-\mu^2\right) q^2 \frac{\partial^2 \mathcal{P}}{\partial \mu^2} + \mathcal{T}\frac{d\mathcal{T}}{d\mathcal{P}} = 0$$

# Results

4 Application to neutron star magnetospheres

# Results

## 4 Application to neutron star magnetospheres

### Current-free case

- $\mathcal{T}\frac{\mathcal{T}}{\mathcal{P}} = 0$
- Simple problem to check our method
- Analytical solution for comparison

# Results

## Current-free case

- $\mathcal{T}\frac{\mathcal{T}}{\mathcal{P}} = 0$
- Simple problem to check our method
- Analytical solution for comparison

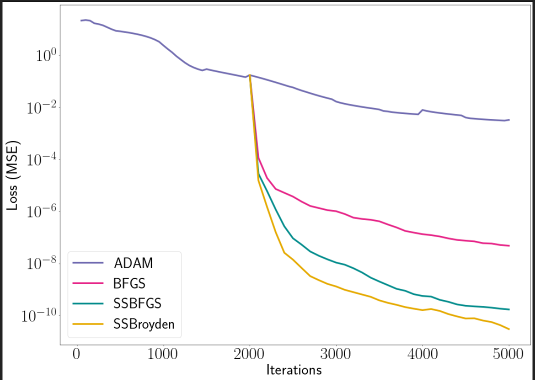


Figure 4: Loss function vs iterations

# Results

## Current-free case

- $\mathcal{T}\frac{\mathcal{T}}{\mathcal{P}} = 0$
- Simple problem to check our method
- Analytical solution for comparison

## Main points

- Quasi-Newton methods are vastly superior to gradient-based methods
- New algorithms further improve convergence compared to BFGS
- Relative error wrt analytical $\sim 10^{-7}$

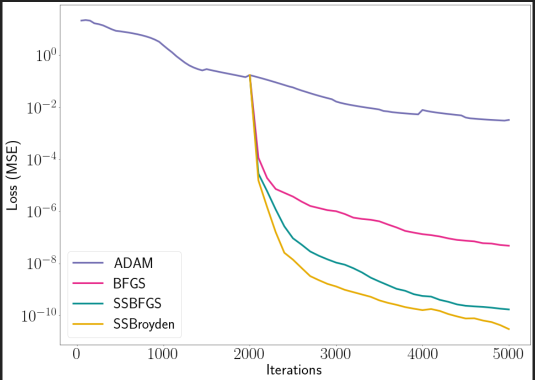


Figure 4: Loss function vs iterations

# Results

## 4 Application to neutron star magnetospheres

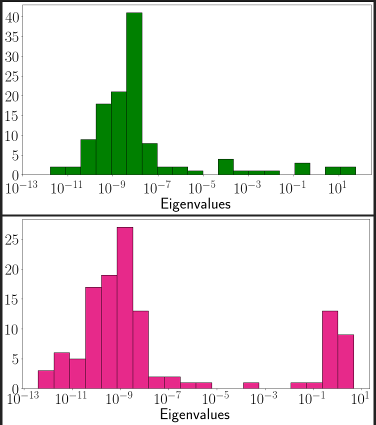


Figure 5: Upper:Adam. Lower:BFGS

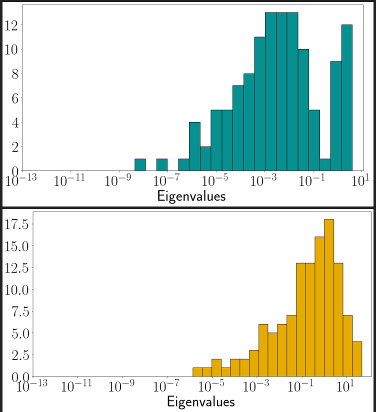


Figure 6: Upper: SSBFGS. Lower:SSBroyden

# Results

4 Application to neutron star magnetospheres

## Force-free case

- Multipolar content
- Non-linear source term

$$\mathcal{T}(\mathcal{P}) = \begin{cases} s\left(|\mathcal{P}| - \mathcal{P}_c\right)^{\sigma} & \text{if } |\mathcal{P}| > \mathcal{P}_c \\ 0 & \text{if } |\mathcal{P}| < \mathcal{P}_c, \end{cases}$$

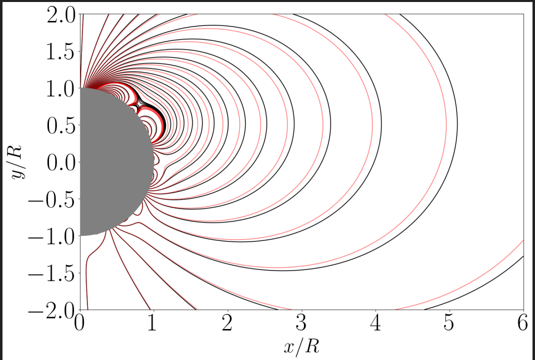


Figure 7: A force-free solution

# Results

4 Application to neutron star magnetospheres

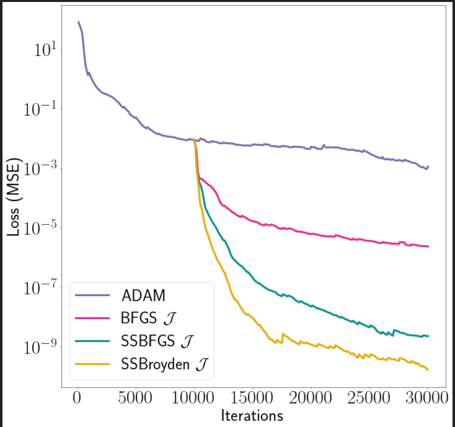


Figure 8: Loss function vs iterations

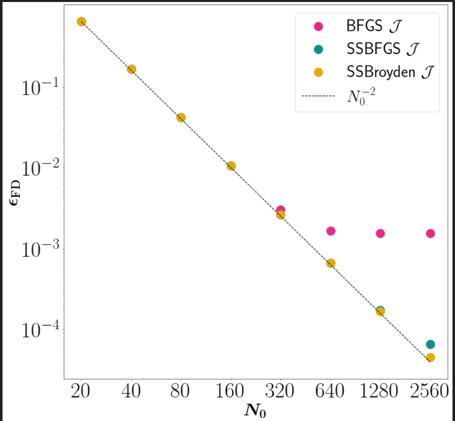


Figure 9: PINN error estimation

# Table of Contents

5 Summary


- ▶ Introduction
- ▶ Motivation
- ▶ Optimisation method
- ▶ Application to neutron star magnetospheres
- ▶ Summary

# Key points



- Optimisation process plays a crucial role
- Suitable choice of optimisation algorithm leads to state-of-the-art precision
- Smaller networks can be employed leading to improved efficiency
- Results extend to other problems
- Preprint for more details:

Urbán, Jorge F., Petros Stefanou, and José A. Pons (May 2024). "Unveiling the optimization process of Physics Informed Neural Networks: How accurate and competitive can PINNs be?" In: *arXiv e-prints*. DOI: 10.48550/arXiv.2405.04230.

# Application of Physics-Informed Neural Networks to Neutron Star Magnetospheres

*Thank you*