Learning To Solve Differential Equations Across Initial Conditions

Shehryar Malik[*^1], Usman Anwar[*1], Ali Ahmed[1],
Alireza Aghasi[2]

[1] Information Technology University, Lahore, Pakistan.
[2]J. Mack Robinson College of Business, Georgia State University, Atlanta, GA, USA.

*Equal contribution.
^Corresponding author: shehryar.malik@itu.edu.pk

Motivation

- **Traditional solvers**
  - Discretize solution by creating a mesh and approximate solution on mesh using numerical methods (finite difference, spectral methods etc.)
  - Limitations:
    - Curse Of Dimensionality: Mesh size grows exponentially in number of dimensions
    - Limitations On Resolution: Do not permit arbitrarily high resolution

- **Can learned PDE solvers do better?**

Machine Learning Approach To Solving PDEs

- **Physics Informed Neural Network [Raissi et al. 2019]**
  - Parametrize neural network as the solution to PDE i.e. u = NN(x,t) where u is solution of u_t - f(x,t,u_x,..) and train under the constraints imposed by partial differential equation model.
  - No discretization or mesh like structure needed
  - Data free training
  - Neural networks better at handling curse of dimensionality

Machine Learning Approach To Solving PDEs

- **Physics Informed Neural Network [Raissi et al. 2019]**
  - Parametrize neural network as the solution to PDE i.e. u = NN(x,t) where u is solution of u_t - f(x,t,u_x,..) and train under the constraints imposed by partial differential equation model.
  - No discretization or mesh like structure needed
  - Data free training
  - Neural networks better at handling curse of dimensionality

- **What if your initial conditions change slightly?**
  - Need to retrain the NN
  - Can be very time consuming when solutions of a PDE across large number of conditions are required.

- **Can a NN solution be made to generalize over initial conditions?**

Problem Setup

We consider stiff PDEs having the form

$$\frac{\partial u(\mathbf{x},t;i)}{\partial t} = f\left(\mathbf{x}, t, u(\mathbf{x}, t; i), \frac{\partial u(\mathbf{x},t;i)}{\partial \mathbf{x}}, \ldots, \frac{\partial^d u(\mathbf{x},t;i)}{\partial \mathbf{x}^d}\right),$$

where $\mathbf{x} \in \mathbb{R}^p, t \in \mathbb{R}, u : \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}$ is the solution of the PDE and f is a known function. Since the solution u depends upon the initial conditions i we make this dependence explicit by writing $u(x, t; i)$.

$\mathcal{I}$ will denote the distribution of initial conditions.

We are interested in training a single model that approximates $u(x, t; i)$ for any $i \sim \mathcal{I}$ when trained on a subset of them.

Proposed Methodology

Use a Generative Adversarial Network (GAN) [Goodfellow et al., 2014].

Idea: Feed in the initial conditions to the generator along with z..

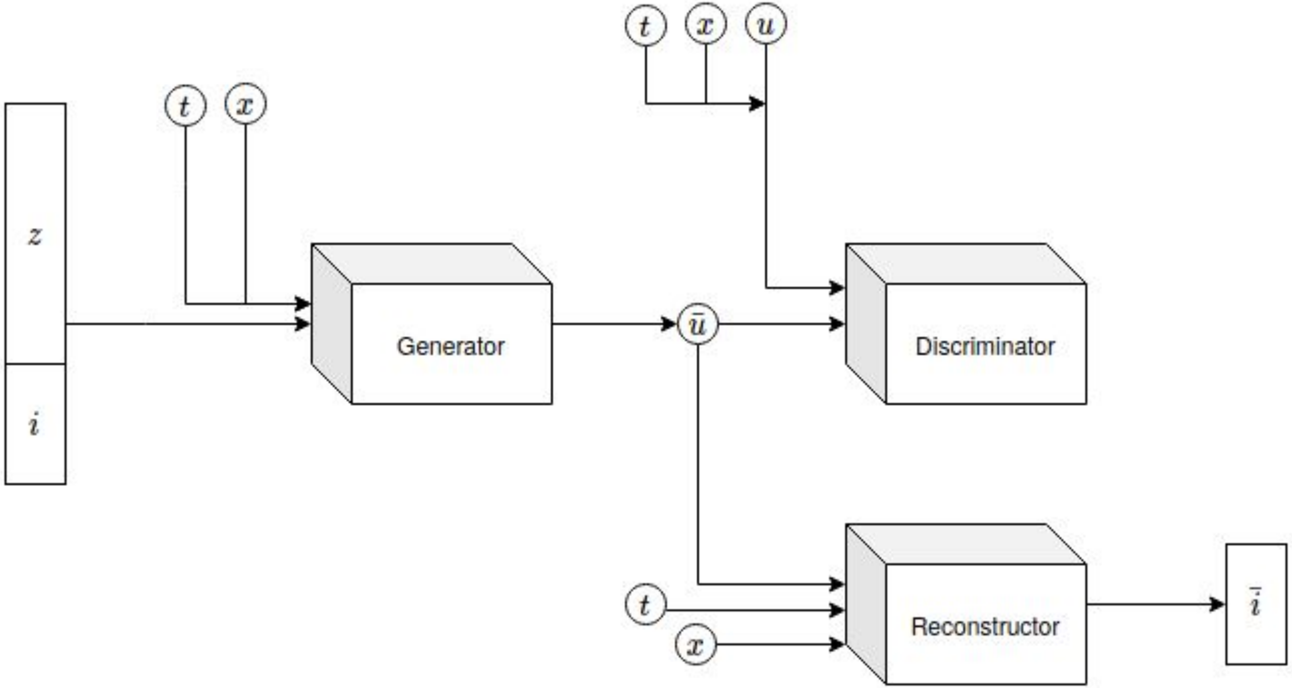
Problem: How to force the generator to condition on these initial conditions?

Use an InfoGan [Chen et al., 2016] style-inspired reconstructor to condition on the initial conditions.

3 components:

1. Generator: Takes in x, t and i and produces $\bar{u}$.
2. Reconstructor: Takes in $\bar{u}$ and tries to reconstruct i.
3. Discriminator: Takes in $\bar{u}$ and u and tries to discriminate between them.

Proposed Methodology - Setup



Proposed Methodology - Details

Train the generator and discriminator by minimizing

$$\mathcal{L} = \mathcal{L}_{GAN} + \alpha \mathcal{L}_{PDE} + \beta \mathcal{L}_{IC} + \gamma \mathcal{L}_{BC}$$

where

$$\mathcal{L}_{GAN} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}} \left[\log D_{\psi}\left(\mathbf{x}, t, u(\mathbf{x}, t; i), i\right)\right] + \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}} \left[1 - \log(D_{\psi}\left(\mathbf{x}, t, G_{\theta}(\mathbf{x}, t, i, \mathbf{z}), i\right))\right],$$

$$\mathcal{L}_{PDE} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}} \left[\left|\left| \frac{\partial G_{\theta}(\mathbf{x},t,i,\mathbf{z})}{\partial t} - f\left(\mathbf{x}, t, G_{\theta}(\mathbf{x}, t, i, \mathbf{z}), \frac{\partial G_{\theta}(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}}, \ldots, \frac{\partial^d G_{\theta}(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}^d}\right) \right|\right|_2^2 \right],$$

$$\mathcal{L}_{IC} = \mathbb{E}_{i,\mathbf{x},\mathbf{z}} \left[||u(\mathbf{x}, t=0; i) - G(\mathbf{x}, t=0, i, \mathbf{z})||_2^2 \right],$$

$$\mathcal{L}_{BC} = \mathbb{E}_{t,i,\mathbf{z}} \left[||G(\mathbf{x}_{UB}, t, i, \mathbf{z}) - G(\mathbf{x}_{LB}, t, i, \mathbf{z})||_2^2 \right] + \mathbb{E}_{t,i,\mathbf{z}} \left[\left|\left| \frac{\partial G(\mathbf{x}_{UB},t,i,\mathbf{z})}{\partial \mathbf{x}} - \frac{\partial G(\mathbf{x}_{LB},t,i,\mathbf{z})}{\partial \mathbf{x}} \right|\right|_2^2 \right].$$

Train the reconstructor to minimize

$$\mathcal{L}_r = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}} \left[||i - R(G(x, t, i, z))||_2^2 \right]$$

Proposed Methodology - Details

Train the generator and discriminator by minimizing

$$\mathcal{L} = \mathcal{L}_{GAN} + \alpha\mathcal{L}_{PDE} + \beta\mathcal{L}_{IC} + \gamma\mathcal{L}_{BC}$$

where

$$\mathcal{L}_{GAN} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\log D_{\psi}\left(\mathbf{x}, t, u(\mathbf{x}, t; i), i\right)\right] + \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[1 - \log(D_{\psi}\left(\mathbf{x}, t, G_{\theta}(\mathbf{x}, t, i, \mathbf{z}), i\right))\right],$$

$$\mathcal{L}_{PDE} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\left|\left|\frac{\partial G_{\theta}(\mathbf{x},t,i,\mathbf{z})}{\partial t} - f\left(\mathbf{x}, t, G_{\theta}(\mathbf{x}, t, i, \mathbf{z}), \frac{\partial G_{\theta}(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}}, \ldots, \frac{\partial^{d} G_{\theta}(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}^{d}}\right)\right|\right|_2^2\right],$$

$$\mathcal{L}_{IC} = \mathbb{E}_{i,\mathbf{x},\mathbf{z}}\left[||u(\mathbf{x}, t = 0; i) - G(\mathbf{x}, t = 0, i, \mathbf{z})||_2^2\right],$$

$$\mathcal{L}_{BC} = \mathbb{E}_{t,i,\mathbf{z}}\left[||G(\mathbf{x}_{UB}, t, i, \mathbf{z}) - G(\mathbf{x}_{LB}, t, i, \mathbf{z})||_2^2\right] + \mathbb{E}_{t,i,\mathbf{z}}\left[\left|\left|\frac{\partial G(\mathbf{x}_{UB},t,i,\mathbf{z})}{\partial \mathbf{x}} - \frac{\partial G(\mathbf{x}_{LB},t,i,\mathbf{z})}{\partial \mathbf{x}}\right|\right|_2^2\right].$$

Train the reconstructor to minimize

$$\mathcal{L}_r = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[||i - R(G(x, t, i, z))||_2^2\right]$$

Proposed Methodology - Details

Train the generator and discriminator by minimizing

$$\mathcal{L} = \mathcal{L}_{GAN} + \alpha\mathcal{L}_{PDE} + \beta\mathcal{L}_{IC} + \gamma\mathcal{L}_{BC}$$

where

$$\mathcal{L}_{GAN} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\log D_\psi\left(\mathbf{x}, t, u(\mathbf{x}, t; i), i\right)\right] + \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[1 - \log(D_\psi\left(\mathbf{x}, t, G_\theta(\mathbf{x}, t, i, \mathbf{z}), i\right))\right],$$

$$\mathcal{L}_{PDE} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\left|\left|\frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial t} - f\left(\mathbf{x}, t, G_\theta(\mathbf{x}, t, i, \mathbf{z}), \frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}}, \ldots, \frac{\partial^d G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}^d}\right)\right|\right|_2^2\right],$$

$$\mathcal{L}_{IC} = \mathbb{E}_{i,\mathbf{x},\mathbf{z}}\left[||u(\mathbf{x}, t = 0; i) - G(\mathbf{x}, t = 0, i, \mathbf{z})||_2^2\right],$$

$$\mathcal{L}_{BC} = \mathbb{E}_{t,i,\mathbf{z}}\left[||G(\mathbf{x}_{UB}, t, i, \mathbf{z}) - G(\mathbf{x}_{LB}, t, i, \mathbf{z})||_2^2\right] + \mathbb{E}_{t,i,\mathbf{z}}\left[\left|\left|\frac{\partial G(\mathbf{x}_{UB},t,i,\mathbf{z})}{\partial \mathbf{x}} - \frac{\partial G(\mathbf{x}_{LB},t,i,\mathbf{z})}{\partial \mathbf{x}}\right|\right|_2^2\right].$$

Train the reconstructor to minimize

$$\mathcal{L}_r = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[||i - R(G(x, t, i, z))||_2^2\right]$$

Proposed Methodology - Details

Train the generator and discriminator by minimizing

$$\mathcal{L} = \mathcal{L}_{GAN} + \alpha\mathcal{L}_{PDE} + \beta\mathcal{L}_{IC} + \gamma\mathcal{L}_{BC}$$

where

$$\mathcal{L}_{GAN} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\log D_\psi\left(\mathbf{x}, t, u(\mathbf{x}, t; i), i\right)\right] + \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[1 - \log(D_\psi\left(\mathbf{x}, t, G_\theta(\mathbf{x}, t, i, \mathbf{z}), i\right))\right],$$

$$\mathcal{L}_{PDE} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\left|\left|\frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial t} - f\left(\mathbf{x}, t, G_\theta(\mathbf{x}, t, i, \mathbf{z}), \frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}}, \ldots, \frac{\partial^d G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}^d}\right)\right|\right|_2^2\right],$$

$$\mathcal{L}_{IC} = \mathbb{E}_{i,\mathbf{x},\mathbf{z}}\left[||u(\mathbf{x}, t = 0; i) - G(\mathbf{x}, t = 0, i, \mathbf{z})||_2^2\right],$$

$$\mathcal{L}_{BC} = \mathbb{E}_{t,i,\mathbf{z}}\left[||G(\mathbf{x}_{UB}, t, i, \mathbf{z}) - G(\mathbf{x}_{LB}, t, i, \mathbf{z})||_2^2\right] + \mathbb{E}_{t,i,\mathbf{z}}\left[\left|\left|\frac{\partial G(\mathbf{x}_{UB},t,i,\mathbf{z})}{\partial \mathbf{x}} - \frac{\partial G(\mathbf{x}_{LB},t,i,\mathbf{z})}{\partial \mathbf{x}}\right|\right|_2^2\right].$$

Train the reconstructor to minimize

$$\mathcal{L}_r = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[||i - R(G(x, t, i, z))||_2^2\right]$$

Proposed Methodology - Details

Train the generator and discriminator by minimizing

$$\mathcal{L} = \mathcal{L}_{GAN} + \alpha\mathcal{L}_{PDE} + \beta\mathcal{L}_{IC} + \gamma\mathcal{L}_{BC}$$

where

$$\mathcal{L}_{GAN} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\log D_\psi\left(\mathbf{x}, t, u(\mathbf{x}, t; i), i\right)\right] + \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[1 - \log(D_\psi\left(\mathbf{x}, t, G_\theta(\mathbf{x}, t, i, \mathbf{z}), i\right))\right],$$

$$\mathcal{L}_{PDE} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\left|\left|\frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial t} - f\left(\mathbf{x}, t, G_\theta(\mathbf{x}, t, i, \mathbf{z}), \frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}}, \ldots, \frac{\partial^d G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}^d}\right)\right|\right|_2^2\right],$$

$$\mathcal{L}_{IC} = \mathbb{E}_{i,\mathbf{x},\mathbf{z}}\left[||u(\mathbf{x}, t = 0; i) - G(\mathbf{x}, t = 0, i, \mathbf{z})||_2^2\right],$$

$$\mathcal{L}_{BC} = \mathbb{E}_{t,i,\mathbf{z}}\left[||G(\mathbf{x}_{UB}, t, i, \mathbf{z}) - G(\mathbf{x}_{LB}, t, i, \mathbf{z})||_2^2\right] + \mathbb{E}_{t,i,\mathbf{z}}\left[\left|\left|\frac{\partial G(\mathbf{x}_{UB},t,i,\mathbf{z})}{\partial \mathbf{x}} - \frac{\partial G(\mathbf{x}_{LB},t,i,\mathbf{z})}{\partial \mathbf{x}}\right|\right|_2^2\right].$$

Train the reconstructor to minimize

$$\mathcal{L}_r = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[||i - R(G(x, t, i, z))||_2^2\right]$$

Proposed Methodology - Details

Train the generator and discriminator by minimizing

$$\mathcal{L} = \mathcal{L}_{GAN} + \alpha\mathcal{L}_{PDE} + \beta\mathcal{L}_{IC} + \gamma\mathcal{L}_{BC}$$

where

$$\mathcal{L}_{GAN} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\log D_\psi\left(\mathbf{x},t,u(\mathbf{x},t;i),i\right)\right] + \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[1-\log(D_\psi\left(\mathbf{x},t,G_\theta(\mathbf{x},t,i,\mathbf{z}),i\right))\right],$$

$$\mathcal{L}_{PDE} = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[\left|\left|\frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial t} - f\left(\mathbf{x},t,G_\theta(\mathbf{x},t,i,\mathbf{z}),\frac{\partial G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}},\ldots,\frac{\partial^d G_\theta(\mathbf{x},t,i,\mathbf{z})}{\partial \mathbf{x}^d}\right)\right|\right|_2^2\right],$$

$$\mathcal{L}_{IC} = \mathbb{E}_{i,\mathbf{x},\mathbf{z}}\left[||u(\mathbf{x},t=0;i) - G(\mathbf{x},t=0,i,\mathbf{z})||_2^2\right],$$

$$\mathcal{L}_{BC} = \mathbb{E}_{t,i,\mathbf{z}}\left[||G(\mathbf{x}_{UB},t,i,\mathbf{z}) - G(\mathbf{x}_{LB},t,i,\mathbf{z})||_2^2\right] + \mathbb{E}_{t,i,\mathbf{z}}\left[\left|\left|\frac{\partial G(\mathbf{x}_{UB},t,i,\mathbf{z})}{\partial \mathbf{x}} - \frac{\partial G(\mathbf{x}_{LB},t,i,\mathbf{z})}{\partial \mathbf{x}}\right|\right|_2^2\right].$$

Train the reconstructor to minimize

$$\mathcal{L}_r = \mathbb{E}_{i,\mathbf{x},t,\mathbf{z}}\left[||i - R(G(x,t,i,z))||_2^2\right]$$

Results and Discussion

Burger's Equation: $\frac{\partial u}{\partial t} = 0.1\frac{\partial^2 u}{\partial x^2} - u\frac{\partial u}{\partial x}$

Initial conditions: $a\sin(bx + c\pi) + .d$

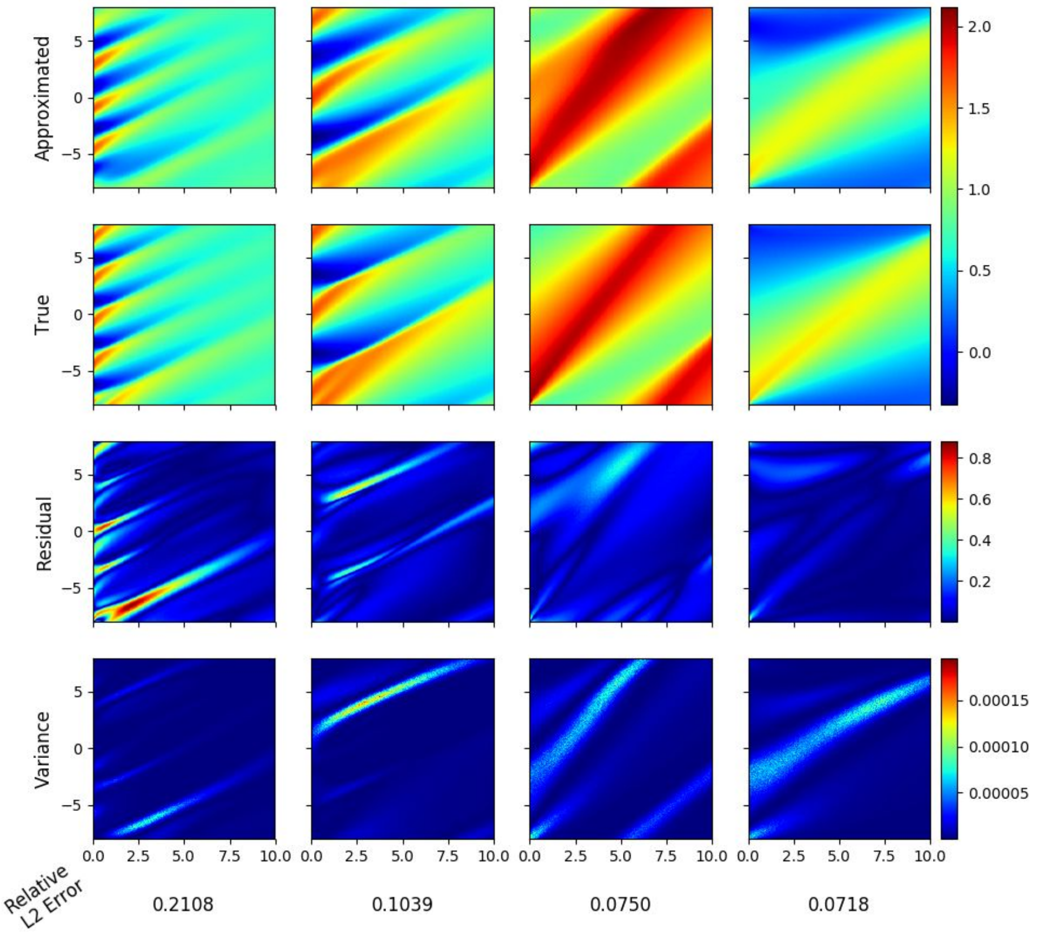
Generate 120 fields. Train and validate on 100 and test on 20.

1. Generalizes across initial conditions.

2. Most of the error is concentrated at the discontinuities and boundaries.

3. Low error elsewhere.

4. Variance corresponds with error loosely.



Conclusion And Future Work

- **General Differential Equation solver**
  - First step towards general learned differential equation solver
    - General across types of PDEs, initial and boundary conditions
    - Handle high dimensions and higher order derivatives
    - Low time and space complexity

- Learning based methods have promise but still a lot needs to be done.

Thank You!