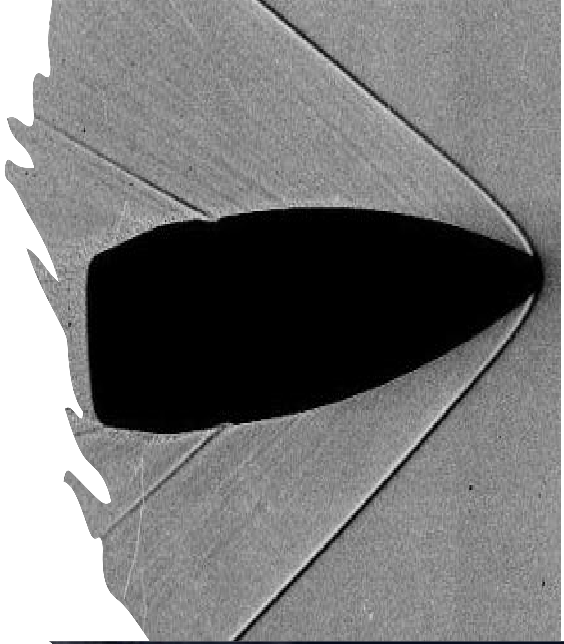
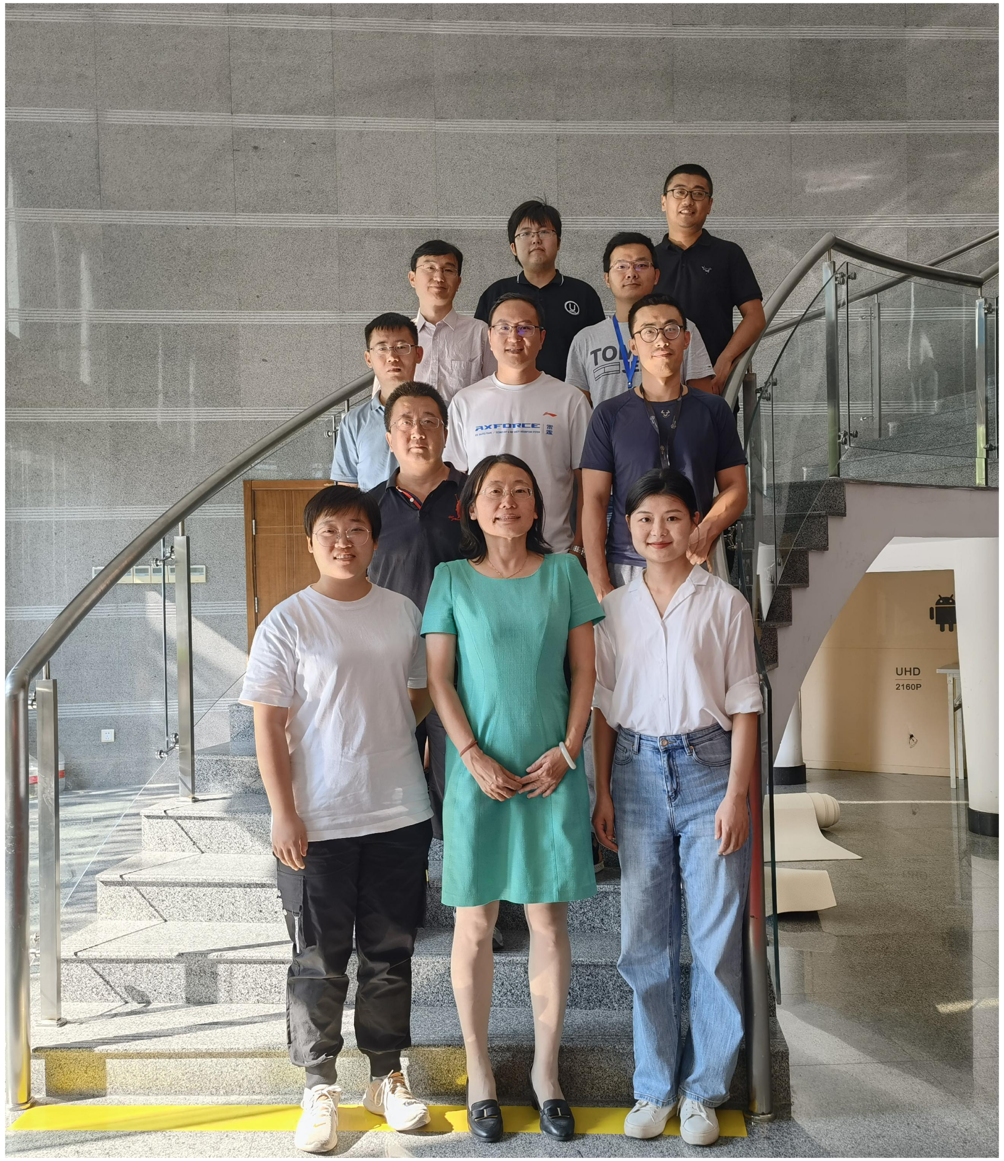
北京应用物理与计算数学研究所
Institute of Applied Physics and Computational Mathematics

Discontinuity computing using physics-informed neural networks (PINNs)

Li Liu



AI++ Team



- **Tengchao Yu**(于腾超）, **Lufeng Liu**（刘鲁峰）
- **Zehua Hu**（胡泽华）, **Shengping Liu**（刘升平
- **Shuai Wang**（王帅）, **Hui Xie**（谢辉）, **Li Liu**（刘利）
- **Danhua Shangguan** （上官丹骅）
- **Ruiyao Ma**（马锐垚）, **Heng Yong**（勇珩）, **Mengjuan Xiao**（肖梦娟）

Interests:
Fluid dynamics, turbulence and instability, radiation transport, Monte Carlo, mesh optimization and nuclear parameters, et al

Contents

Scientific Machine Learning

Partial differential equation (PDE) solving

$$f\left(\mathbf{x};\frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_d};\frac{\partial^2 u}{\partial x_1 \partial x_1},\ldots,\frac{\partial^2 u}{\partial x_1 \partial x_d};\ldots;\boldsymbol{\lambda}\right)=0 \qquad \mathcal{B}(u,\mathbf{x})=0 \quad \text{on} \quad \partial\Omega.$$

Difficults in traditional numerical methods：

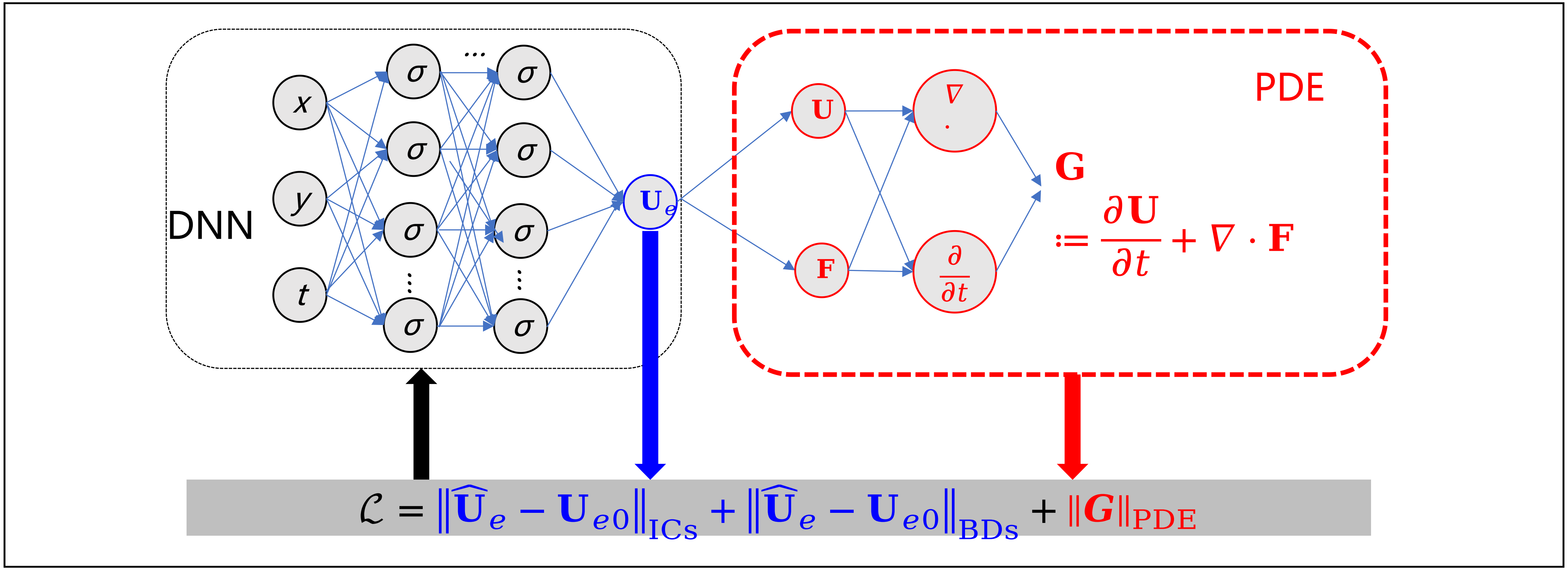
- **Mesh generation**
- **Curse of dimensionality**： 1.complexity increases exponentially 2. Lack of high - dimensional theory
- **Data combination**： The equations contain undetermined parameters
- **Data compression**： The complex flow field data has a large storage volume

Use machine learning methods to solve difficult problems in scientific computing

Physics - Informed Neural Networks (PINNs)

Solving conservative convection PDEs:

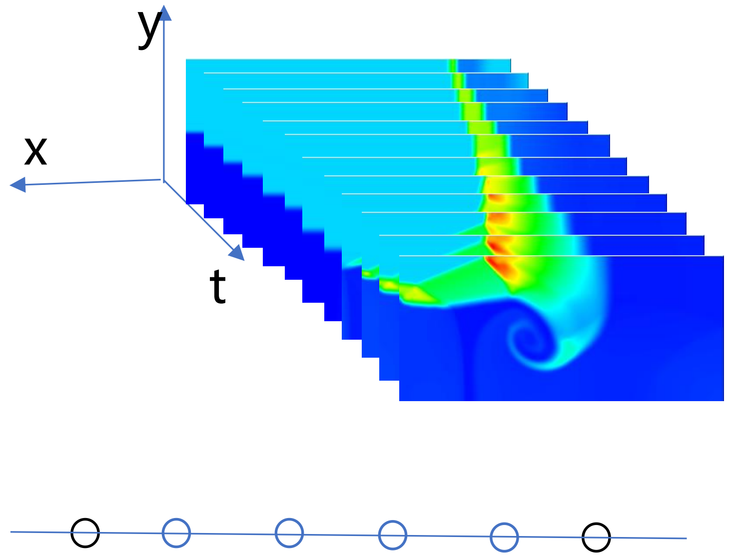
$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F} = 0$$ Initial and boundary conditions



Raissi, Yazdani & Karniadakis, JCP, 2019

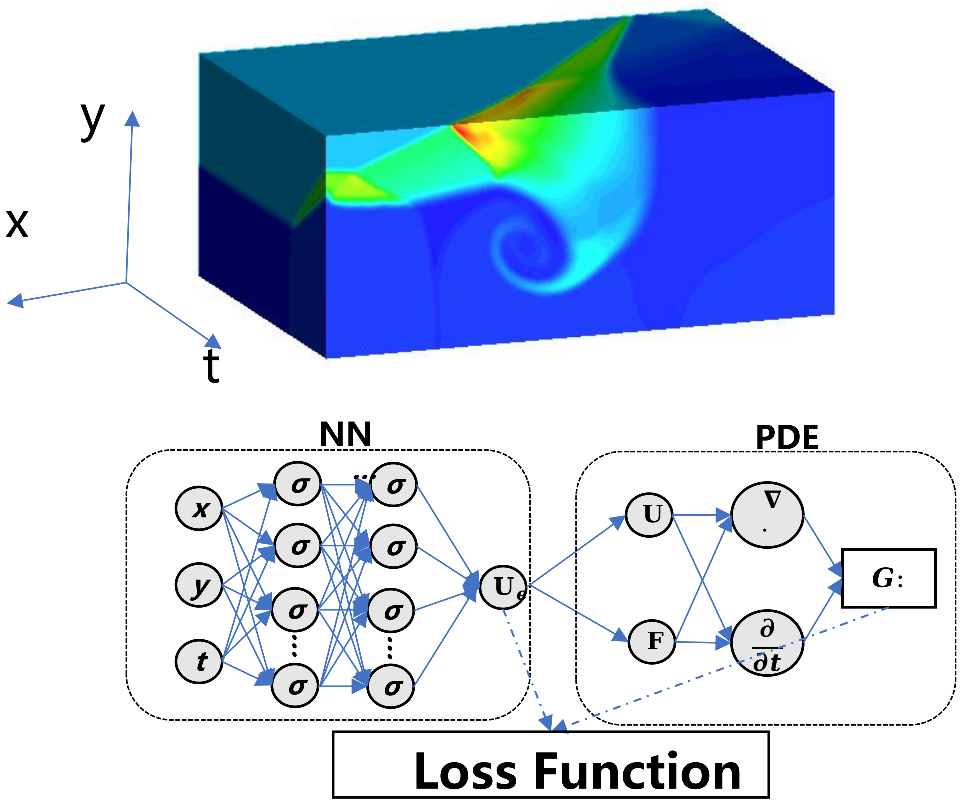
PINNs compared with traditional methods:

Tradition Methods (FDM)



$$\frac{\partial u}{\partial x}\Big|_i = \frac{u_{i+1} - u_{i-1}}{2\Delta x} + O(\Delta x^2)$$

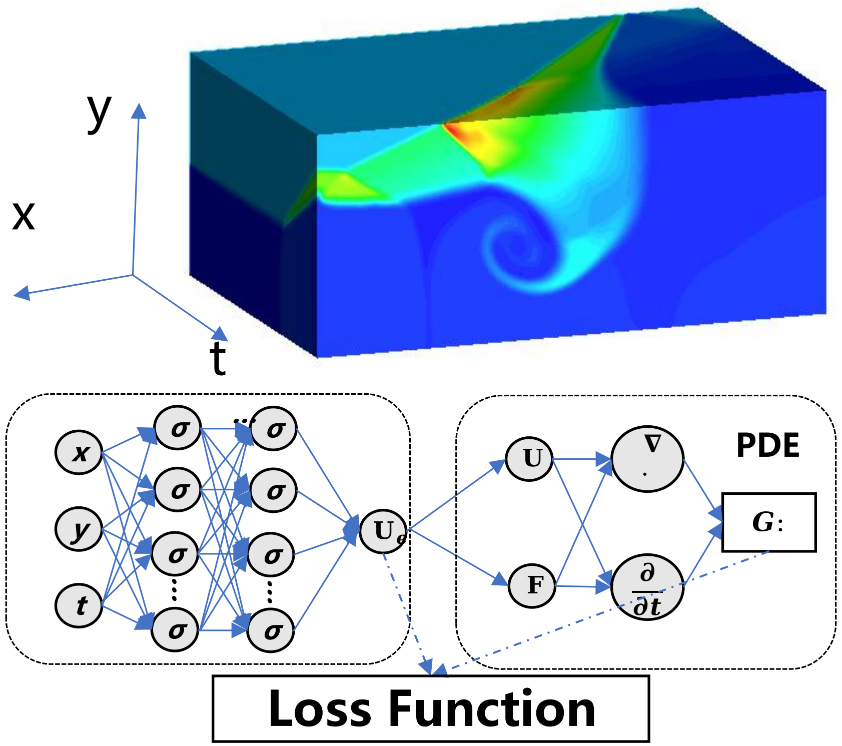
PINNs



Equation stepping solving： U.A.T → **Constrained optimization problem**

Local → Global　　　**Global → Local**

Advantages of PINNs

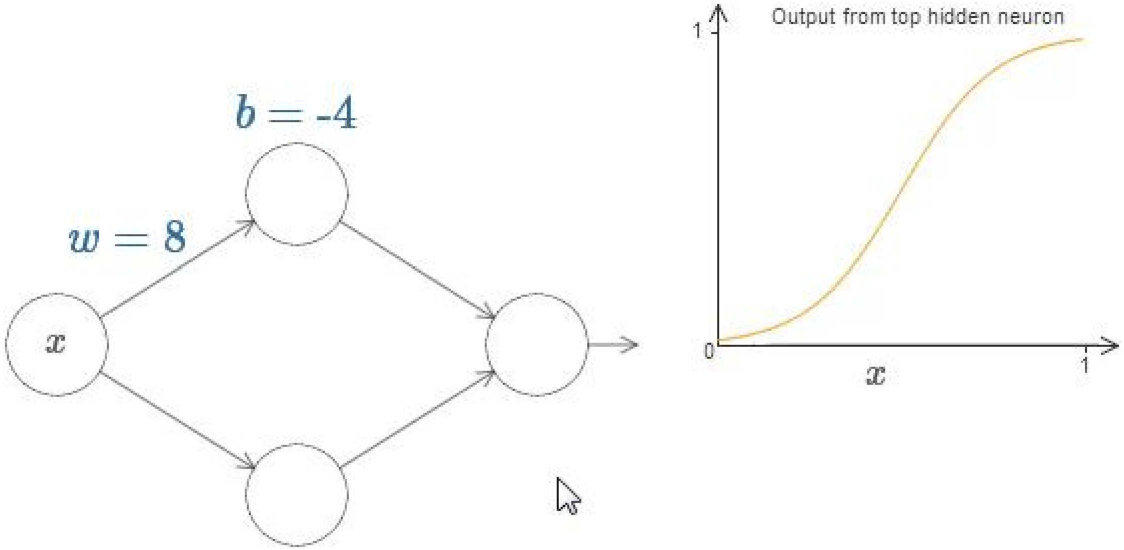


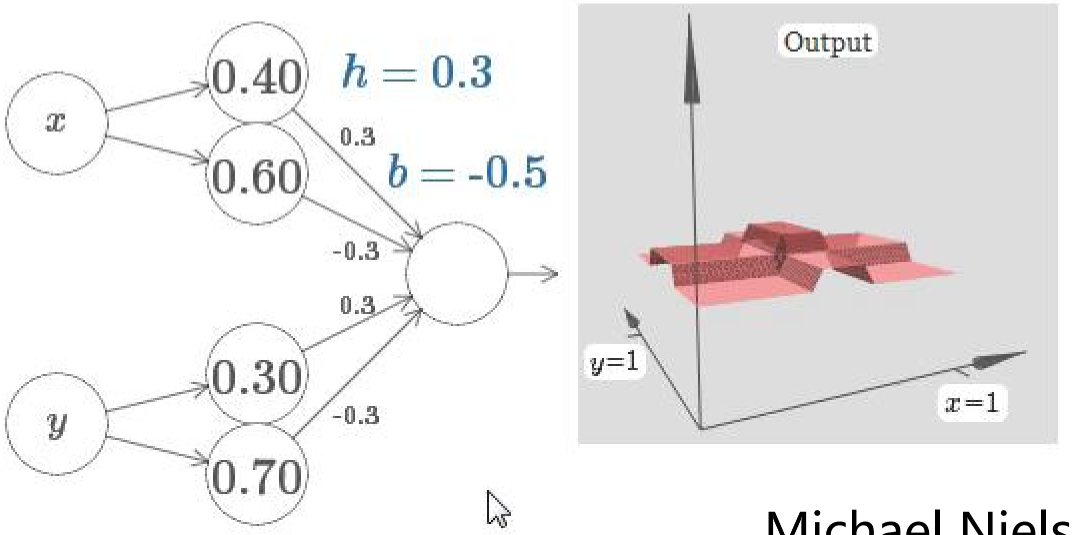
- Mesh - free and particle - free
- Scheme Free
- **Directly solve the initial and boundary value problems (forward problems)**
- **Easy to solve the data combination problems (inverse problems)**
- **Alleviate the curse of dimensionality**
- compression is completed during training
- ***Discontinuous solution has the potential for higher resolution***

Z. Cai, et.al., JCP, 2021, Li Liu, et,al., JSC 2024

Advantages of discontinuous solution

Universal approximation theorem (Hornik et al., 1989; Cybenko, 1989):
A feed - forward neural network with a linear output layer and at least one hidden layer with an activation function having any' squeezing' property (such as the logistic sigmoid activation function) can approximate any Borel measurable function from one finite- dimensional space to another finite - dimensional space with arbitrary accuracy as long as the network is given a sufficient number of hidden units.



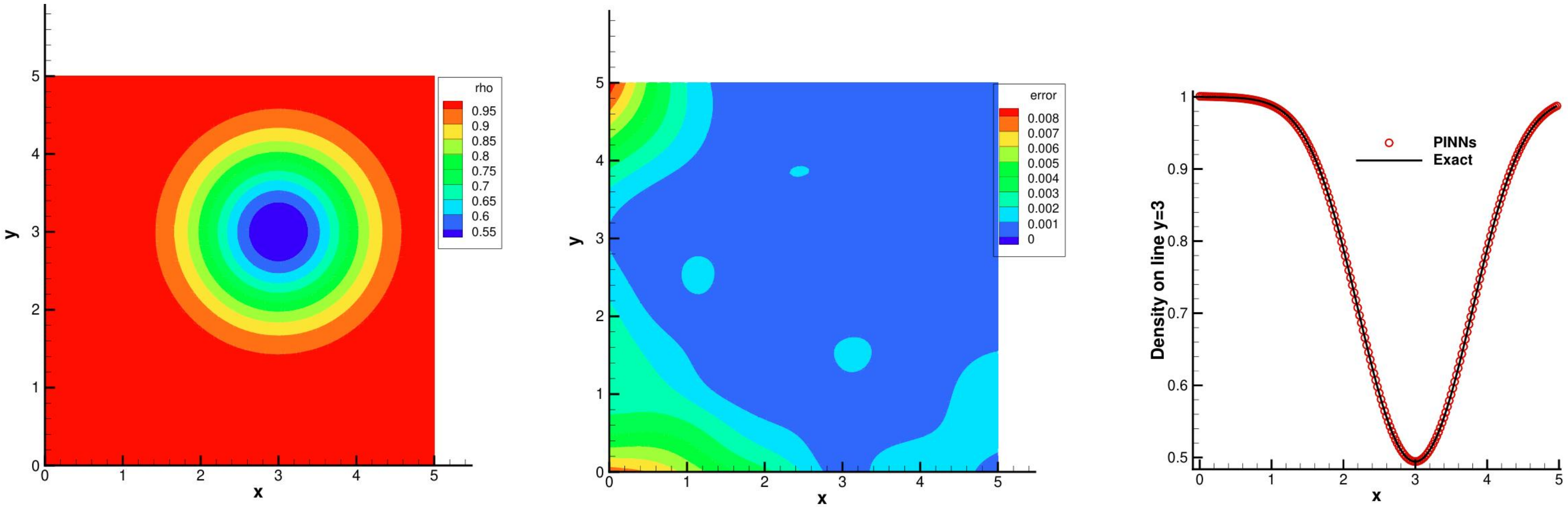


Michael Nielsen

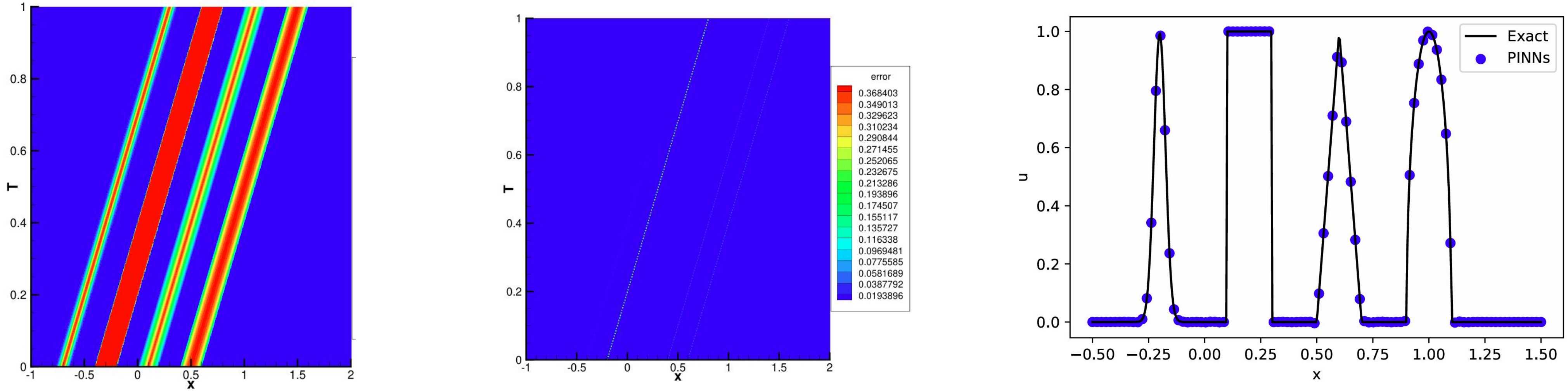
Tanh activation function，Two Neurons

Tanh activation function，Four Neurons

Smooth problems, linear and weakly discontinuous problems

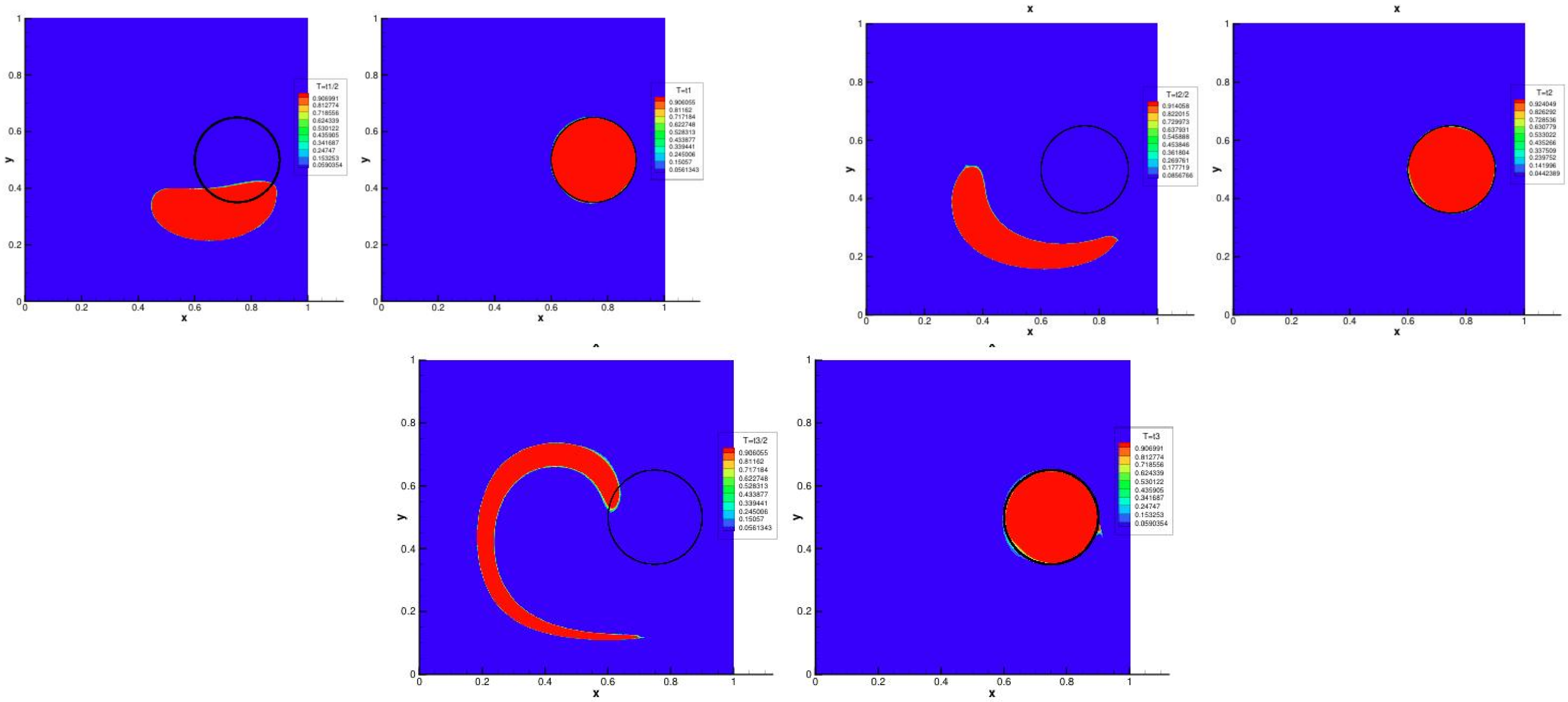


Vortex convection problem



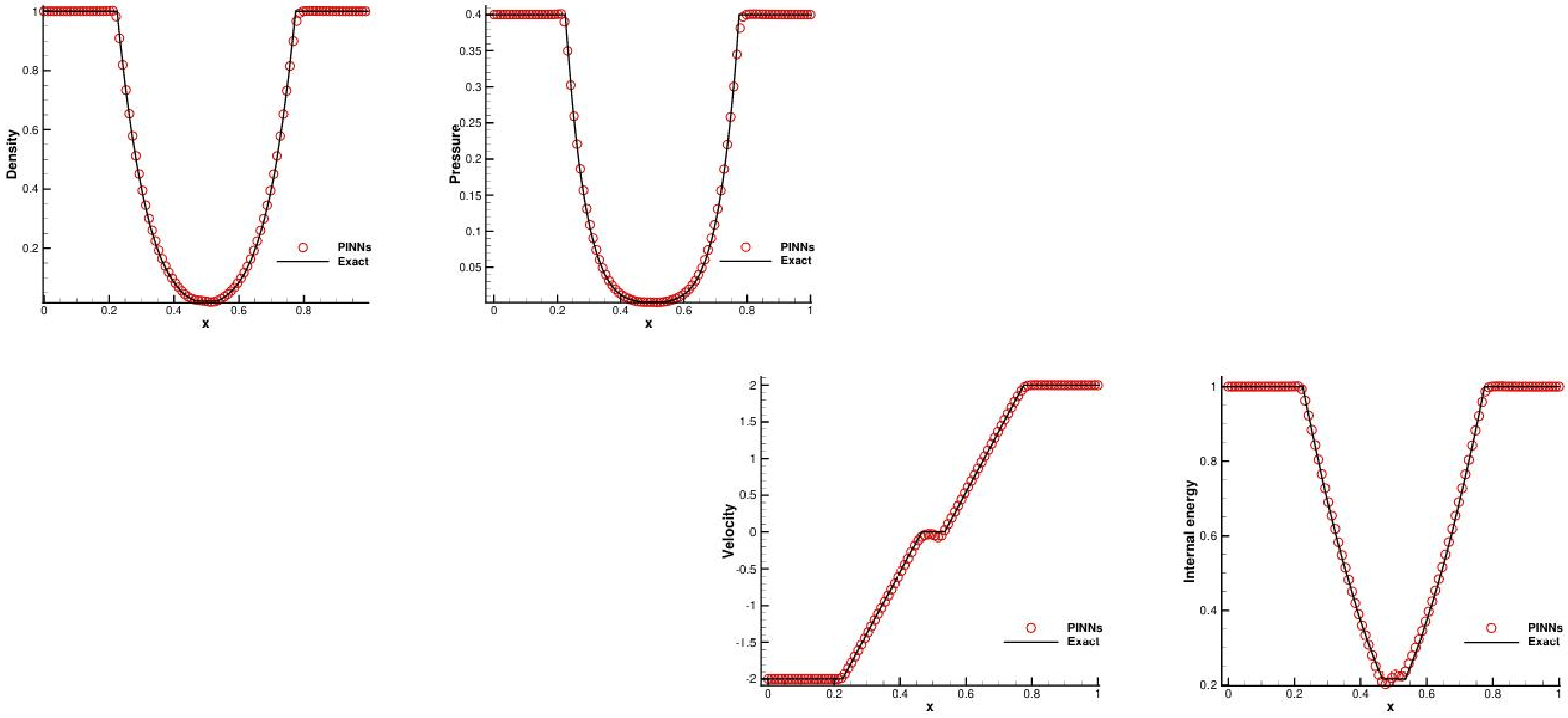
Linear Convection Combined Wave Problem

Smooth problems, linear and weakly discontinuous problems



2D Interface Problem

Smooth problems, linear and weakly discontinuous problems



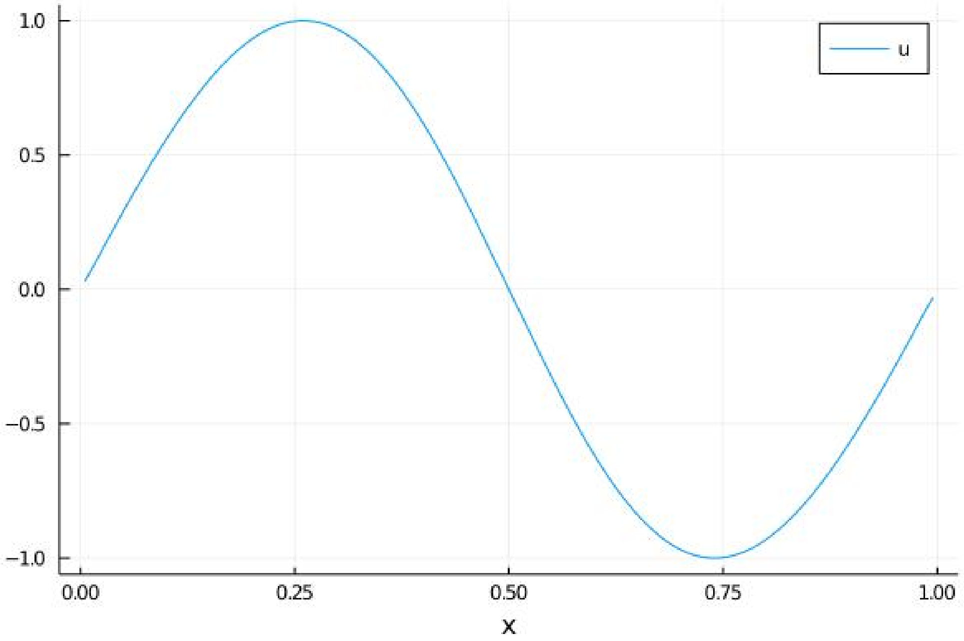
Euler Equation 123problem

Contents

Problem Analysis With Strong Discontinuity

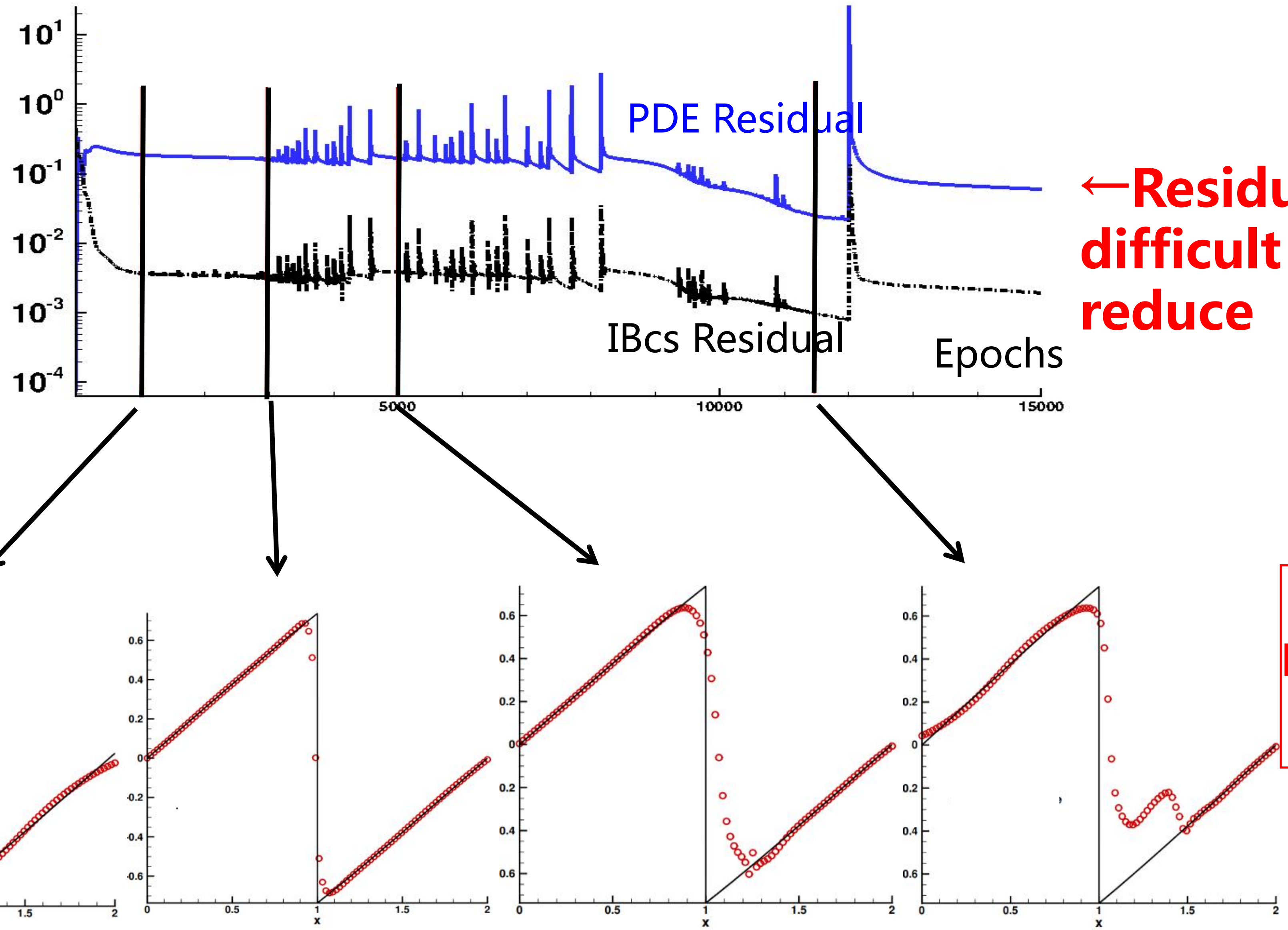
Burgers Problem:

Eq: $\frac{\partial u}{\partial t}+\frac{\partial (u^2/2)}{\partial x}=0$ IC: $u(0,x)=\sin(\pi x)$ BC: $u(t,0)=u(t,1)=0$



A typical discontinuous solution example

Residual

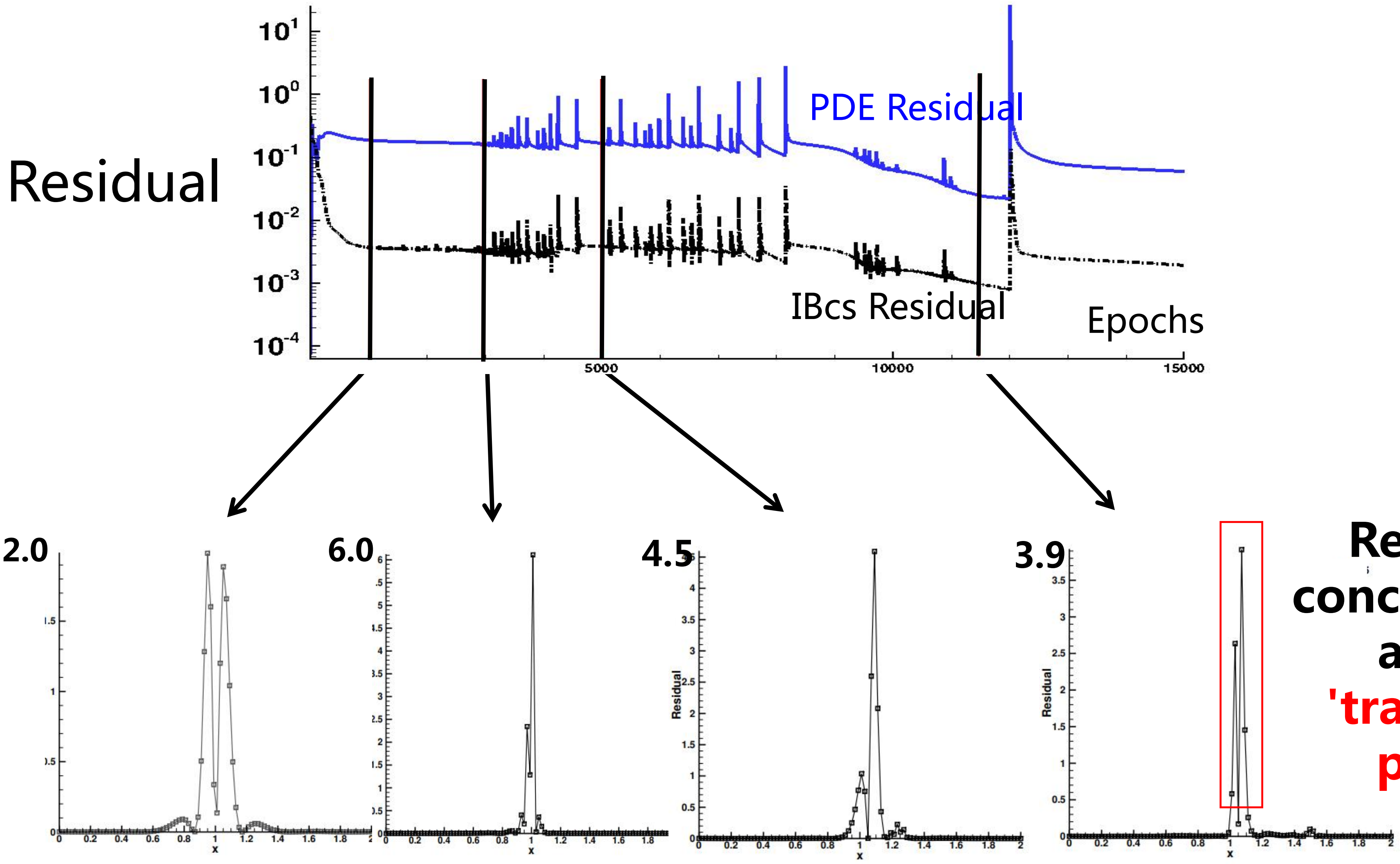


←Residual difficult to reduce

large Prediction error

Prediction of u at time t = 1 for different training epochs of the model

PINNs Analysis



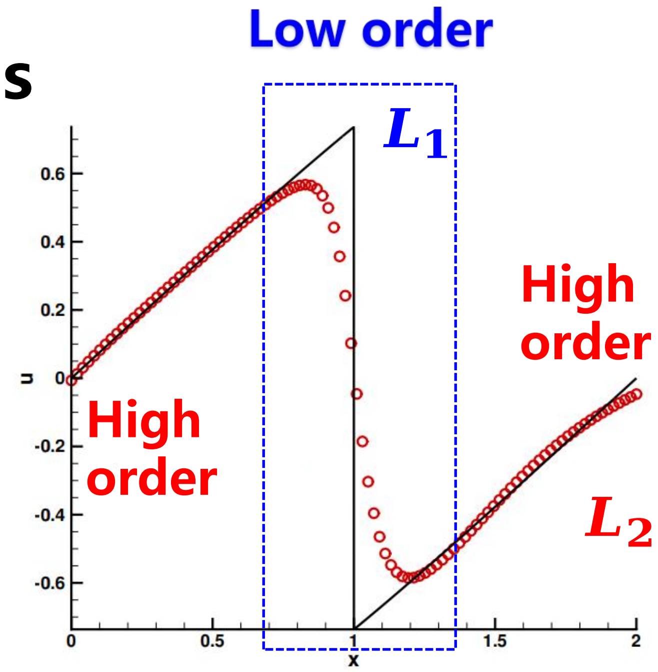
Residual concentrated at the 'transition point'

Residual distribution at time t=1 for different model training steps

Numerical Dissipation & PINNs

Traditional discontinuous capturing method：

1. Introduce scheme/artificially viscosity to suppress large gradients
2. Reduce the accuracy (order) in the Transition region



PINNs：

$$L_{PDE} = L_1 + L_2$$

If $\boldsymbol{L_1 = O(0.1)}$
and $\boldsymbol{L_2 < O(0.1)}$
Then $\boldsymbol{L_{PDE} = O(0.1)}$

loses control in the smooth region，unless L_1 is effectively reduced，

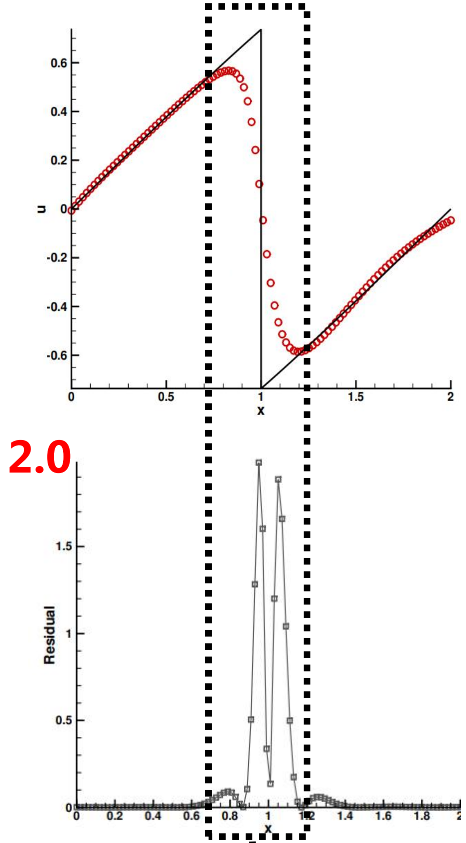
PINNs（By Dissipation）：
→Only transition region is sufficiently smooth, accuracy can be improved
→**difficult to surpass the traditional method**

New idea:

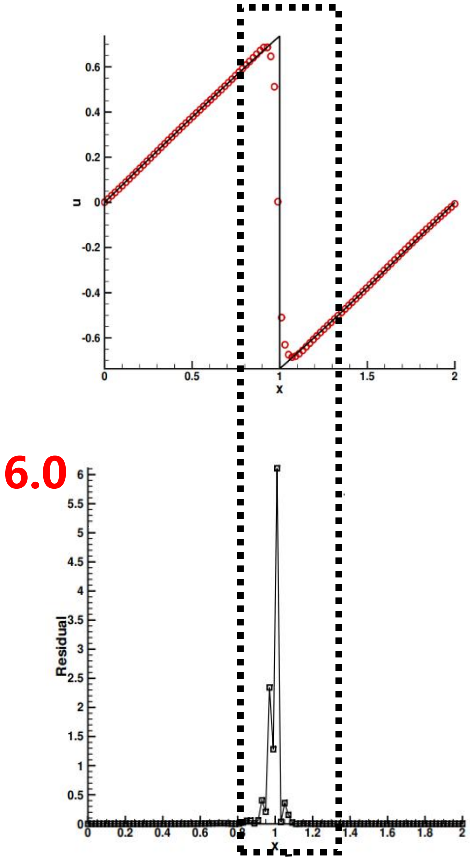
Eliminate the transition point

Analysis at the transition point

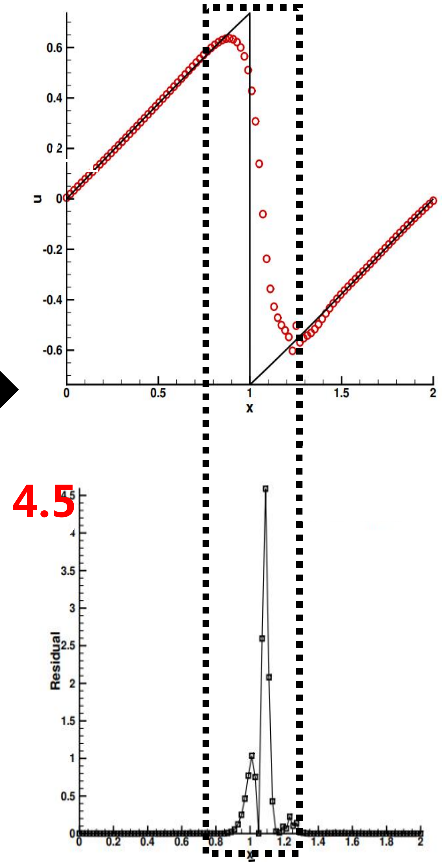
There is a paradoxical state at the transition point



Carry most of the residual:
Concentrate
It does not satisfy the strong - form PDE



Increase the gradient:
Approaching solution, the residual increases and the the ” direction ” is wrong



Decrease the gradient:
Getting away from the solution, the residual increases the ” direction ” is wrong

Reason Analysis and The new Idea

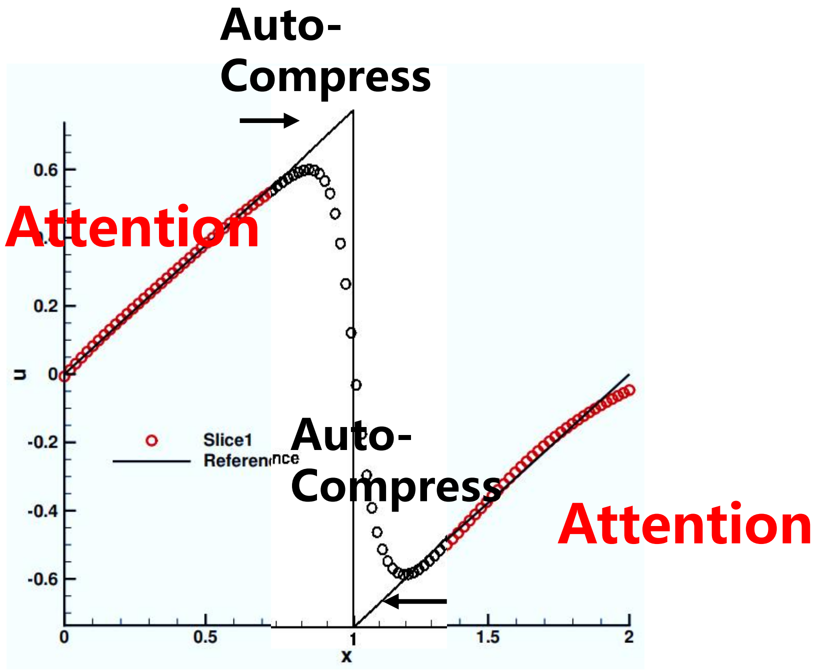
Reason：

Decrease the gradient： large gradient brings larger equation loss

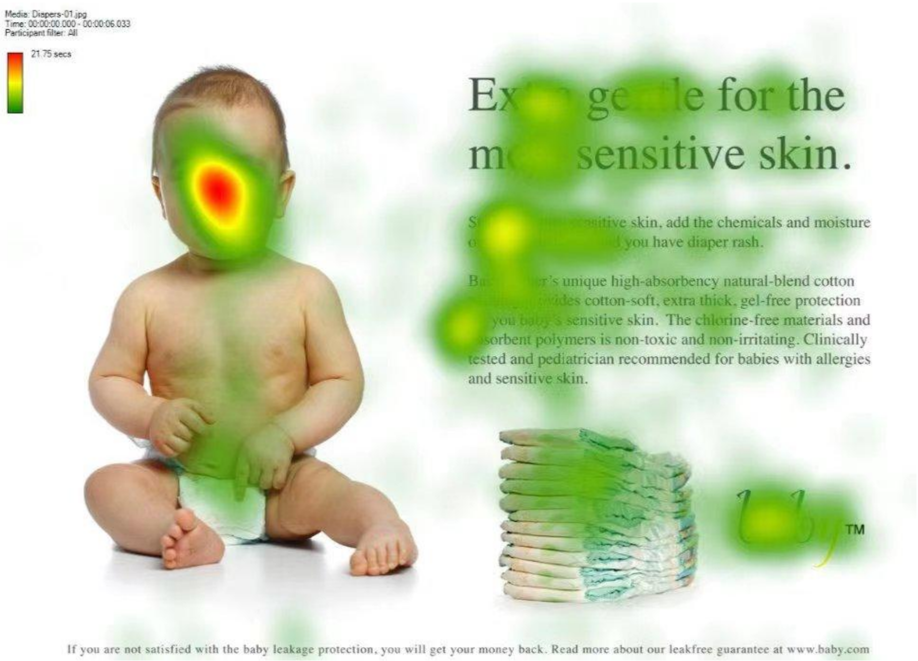
Confrontation

Increase the gradient： shock wave is in a physical compression state

Idea： **Weaken the neural network’s attention to the transition point→Eliminate the confrontation**



Weaken the neural network’s attention



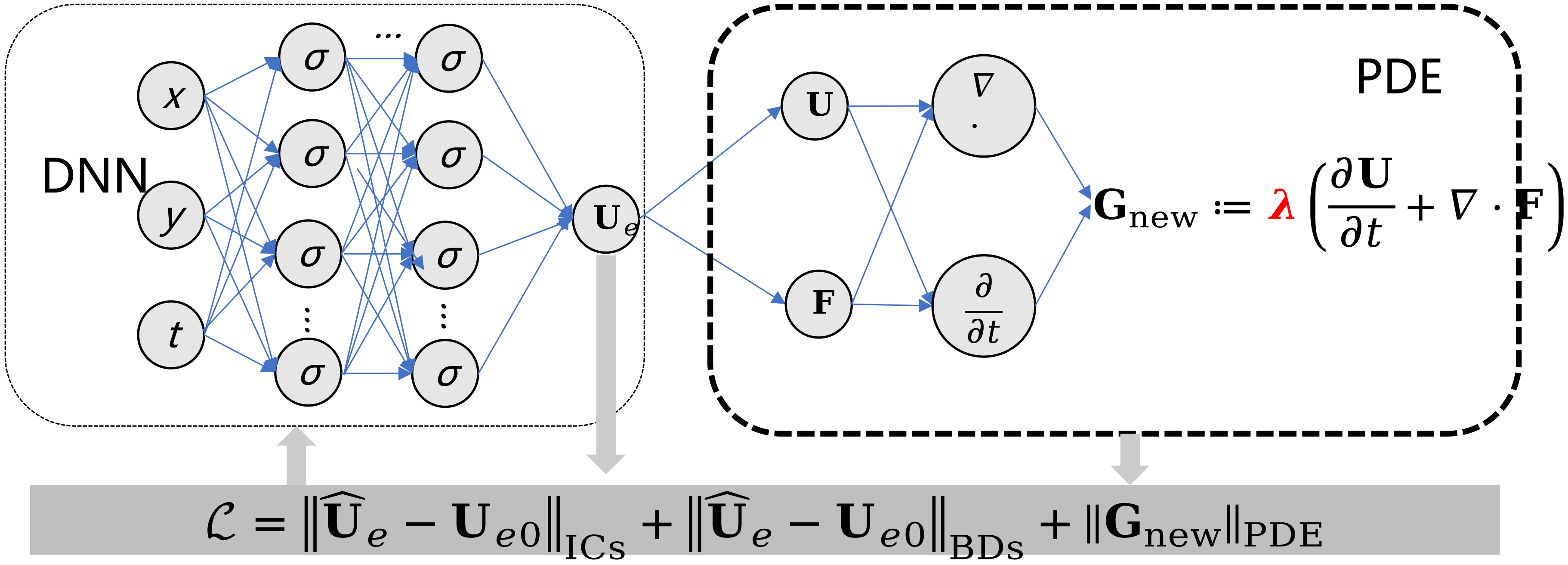
- The attention mechanism is a biological mechanism
- There are two methods to adjust the network attention in PINNs:

1. Adjust the sampling ×

2.**Adjust the weight** √

Weighted equation method (WE)

PINNs-WE



$$\lambda = \frac{1}{\varepsilon_2(|\nabla \cdot \vec{u}| - \nabla \cdot \vec{u}) + 1}$$

Compress: $\lambda \rightarrow 0$
Other Regions: $\lambda \rightarrow 1$

where: $\varepsilon_2 = 0.1$

von Neumann artificial viscosity

Weighted equation method (WE)

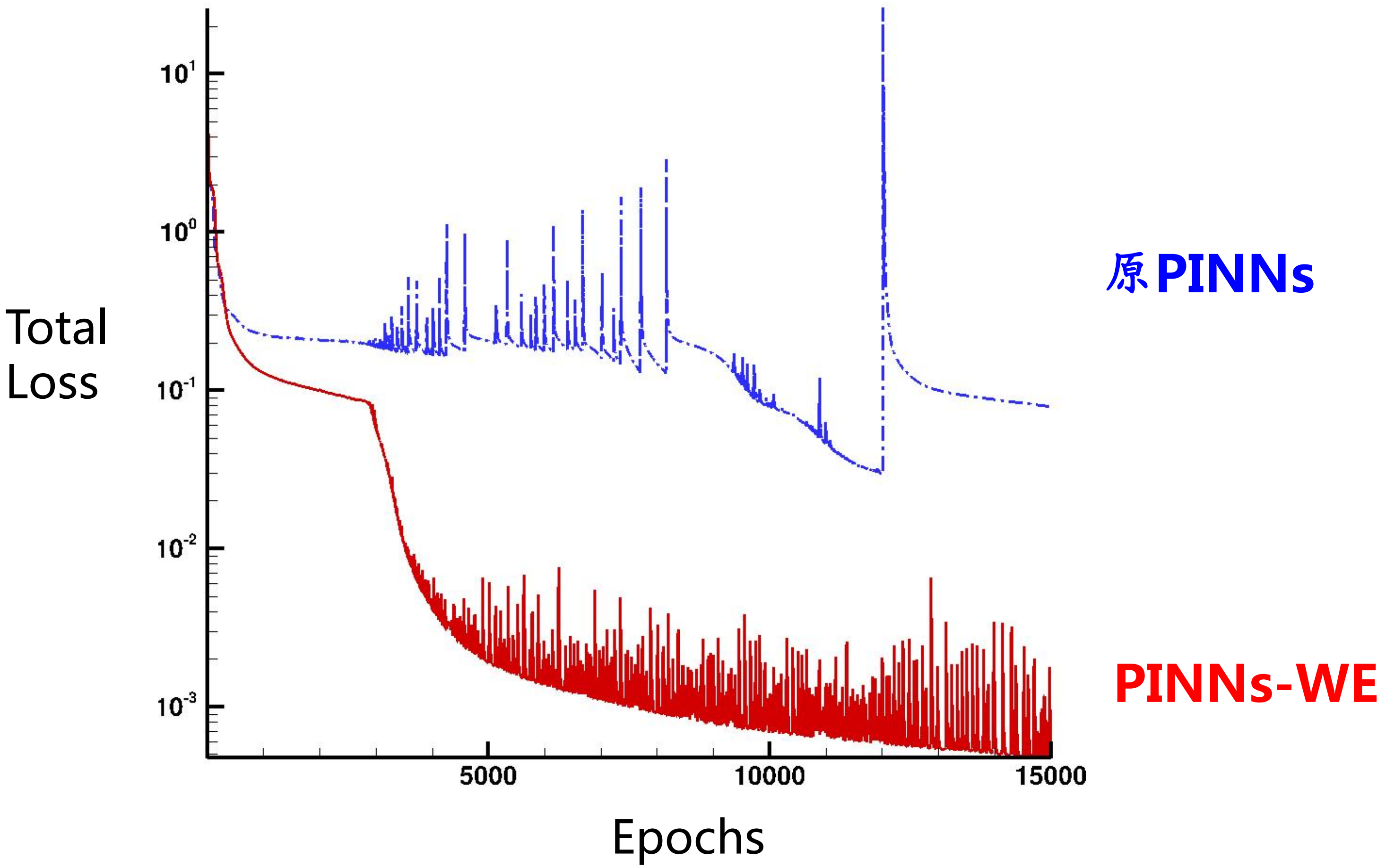
PINNs-WE

$$\mathbf{G}_{\text{new}} := \lambda\left(\frac{\partial \mathbf{U}}{\partial t} + \nabla\cdot\mathbf{F}\right) \quad \lambda = \frac{1}{\varepsilon_2(|\nabla\cdot\vec{u}| - \nabla\cdot\vec{u}) + 1}$$

$$\lambda > 0$$

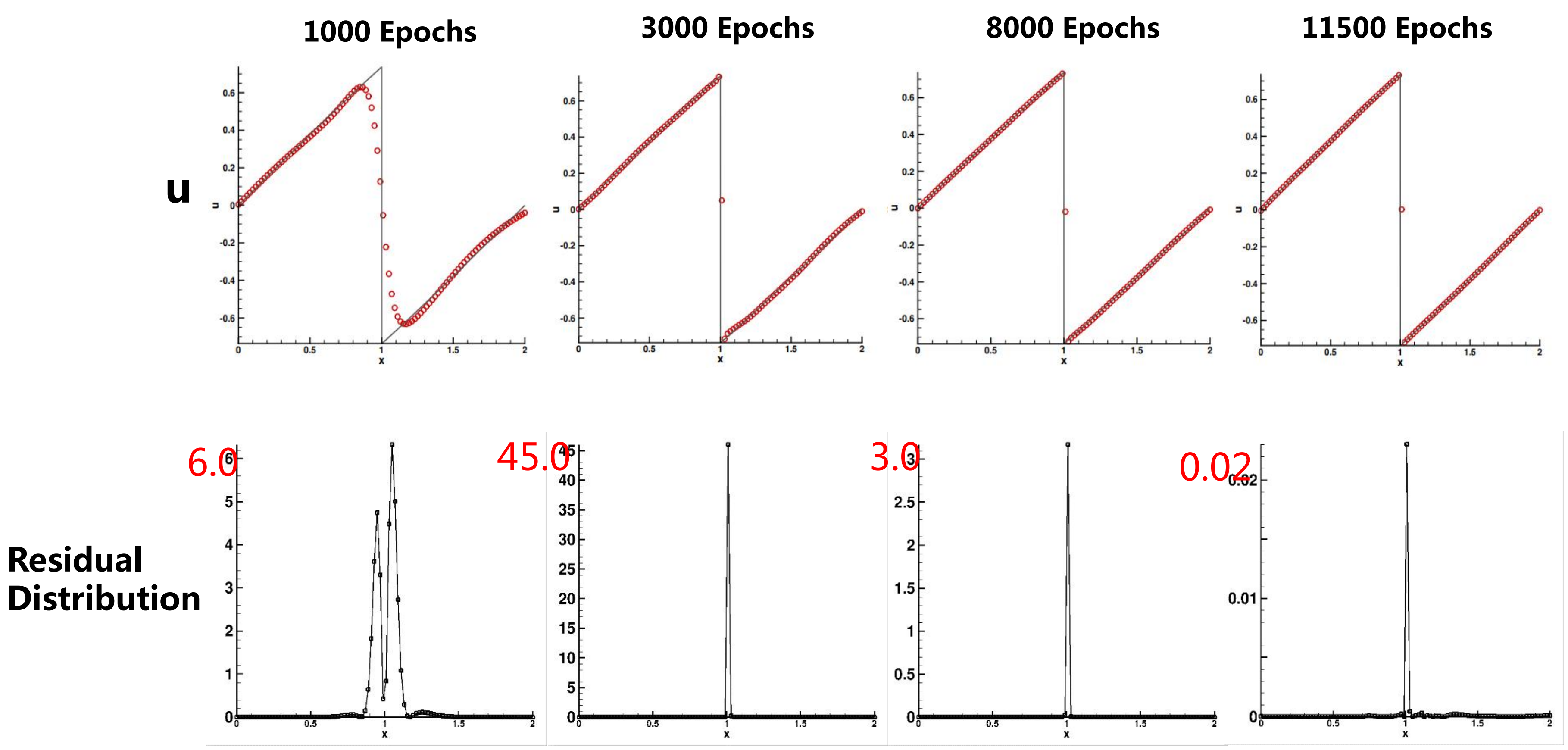
$$\lambda\left(\frac{\partial \mathbf{U}}{\partial t} + \nabla\cdot\mathbf{F}\right) = 0 \; eq. \quad \frac{\partial \mathbf{U}}{\partial t} + \nabla\cdot\mathbf{F} = 0$$

Weighted equation method (WE)



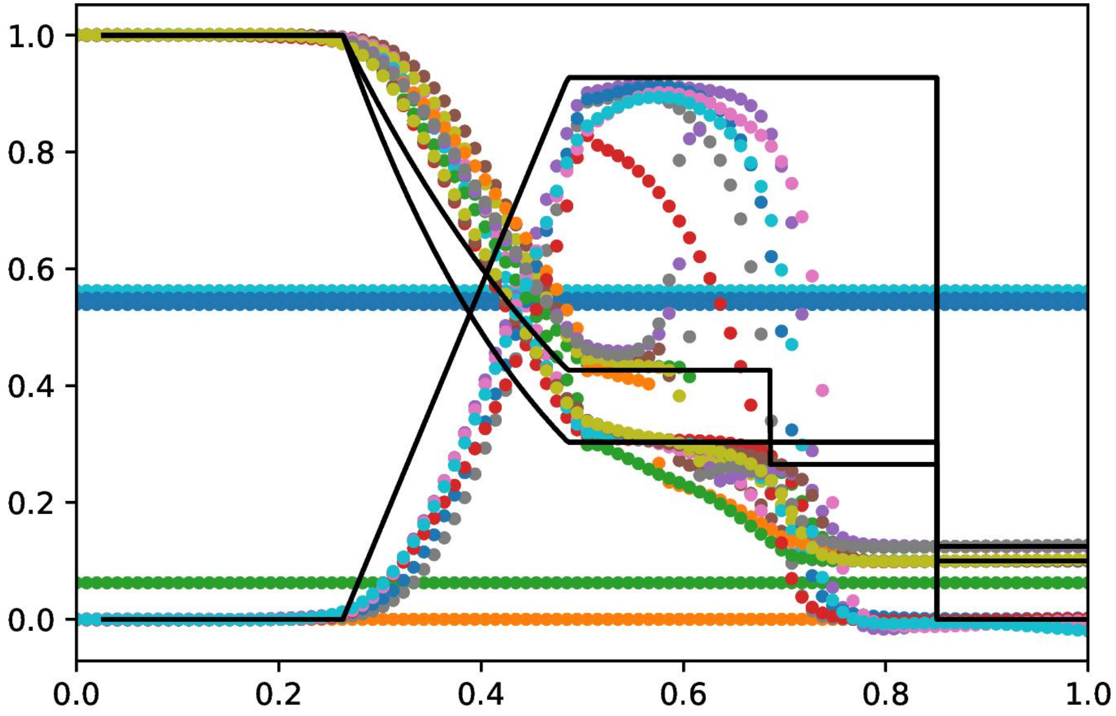
The new method can effectively reduce the total residual.

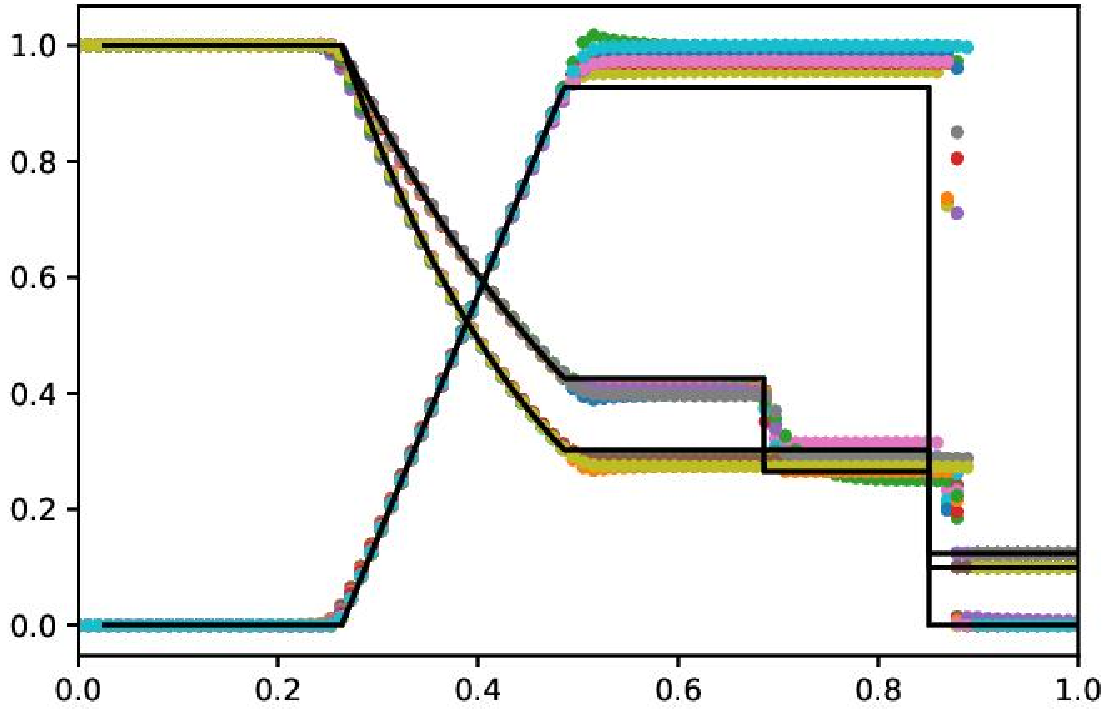
Weighted equation method (WE)



The new method can effectively eliminate transition points and reduce the function prediction error

Sod Problem





Original PINNs

PINNs -WE

- Although it converges to a discontinuous solution, the random perturbation is large and the **under - determination** is severe

Math.

PDE : $\boxed{\dfrac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F} = \mathbf{0} \not\Rightarrow}$ **Nonlinear Discontinuity**

Phys.

$\Rightarrow$

NN

Hard to Convergence

Method

Equation Weight

Math.

PDE : $\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F} = \mathbf{0}$

Lack of reasonable discontinuity constraints

Phys.

Nonlinear Discontinuity

ill - posed

NN

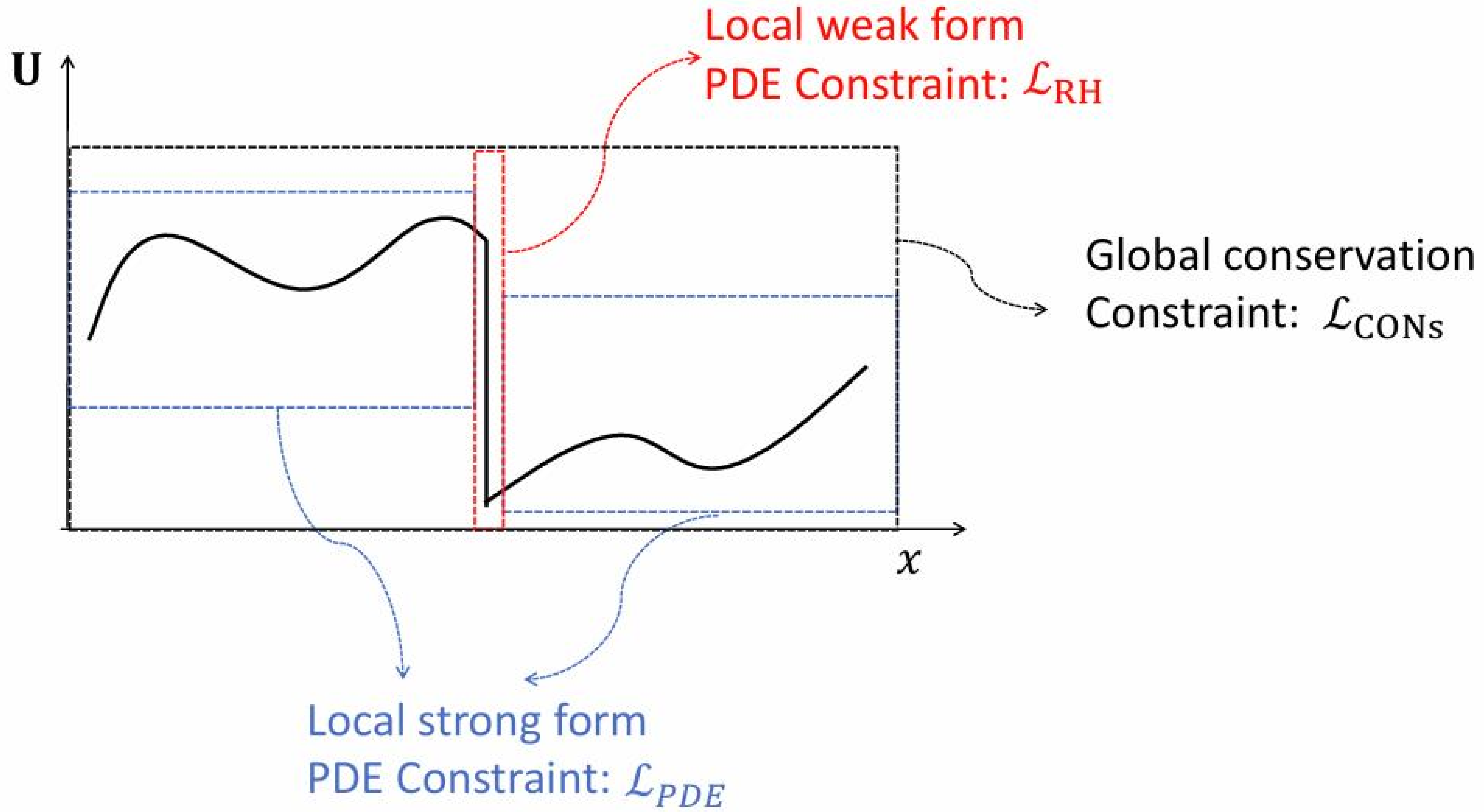
Hard to Convergence

multiple solutions

Method

Equation Weight

weak solutions: local and global conservation constraints


U
x
Local weak form
PDE Constraint: $\mathcal{L}_{\mathrm{RH}}$
Global conservation
Constraint: $\mathcal{L}_{\mathrm{CONs}}$
Local strong form
PDE Constraint: $\mathcal{L}_{PDE}$

Flow Conservation Constraint

$$s \bullet [[\mathbf{U}]] = [[\mathbf{F}]]$$

2D Euler equations RH Condition :

$$s[\![\rho]\!] - [\![\rho\vec{u}\cdot\vec{n}]\!] = 0,$$

$$s[\![\rho u]\!] - [\![(\rho\vec{u}\cdot\vec{n})\,u + pn_x]\!] = 0,$$

$$s[\![\rho v]\!] - [\![(\rho\vec{u}\cdot\vec{n})\,v + pn_y]\!] = 0,$$

$$s[\![E]\!] - [\![(E+p)\vec{u}\cdot\vec{n}]\!] = 0.$$

$$d\rho = \rho - \rho_0 \to 0$$
$$du = u - u_0 \to 0$$
$$dp = p - p_0 \to 0$$

Under differentiability, the Rankine-Hugoniot conditions can be transformed into rarefaction wave conditions.

Aerodynamic relations: :

$$\rho_1\rho_2\left[(u_1-u_2)^2 + (v_1-v_2)^2\right] = (p_1-p_2)(\rho_1-\rho_2),$$

$$\rho_1\rho_2(e_1-e_2) = \frac{1}{2}(p_1+p_2)(\rho_1-\rho_2).$$

Limiter 2 :

$$\lambda_2(\mathbf{U}_1,\mathbf{U}_2) = \begin{cases} |(p_1-p_2)(\vec{u}_1-\vec{u}_2)| & \text{if} \quad |p_1-p_2| > \varepsilon_1 \quad \text{and} \quad |u_1-u_2| > \varepsilon_2, \\ 0 & \text{elsewhere.} \end{cases}$$

Global Conservation Constraints

Global Conservation :

$$\int_V \rho \mathrm{d}V|_{t=t_2} - \int_V \rho \mathrm{d}V|_{t=t_1} = \int_{t_1}^{t_2} \oint_{\partial V} \rho(\vec{u} \cdot \vec{n}_{\partial V}) \mathrm{d}t\mathrm{d}A,$$

$$\int_V \rho\vec{u} \mathrm{d}V|_{t=t_2} - \int_V \rho\vec{u} \mathrm{d}V|_{t=t_1} = \int_{t_1}^{t_2} \oint_{\partial V} \rho\vec{u}(\vec{u} \cdot \vec{n}_{\partial V}) + p\vec{n}_{\partial V} \mathrm{d}t\mathrm{d}A,$$

$$\int_V E \mathrm{d}V|_{t=t_2} - \int_V E \mathrm{d}V|_{t=t_1} = \int_{t_1}^{t_2} \oint_{\partial V} (E+p)(\vec{u} \cdot \vec{n}_{\partial V}) \mathrm{d}t\mathrm{d}A,$$

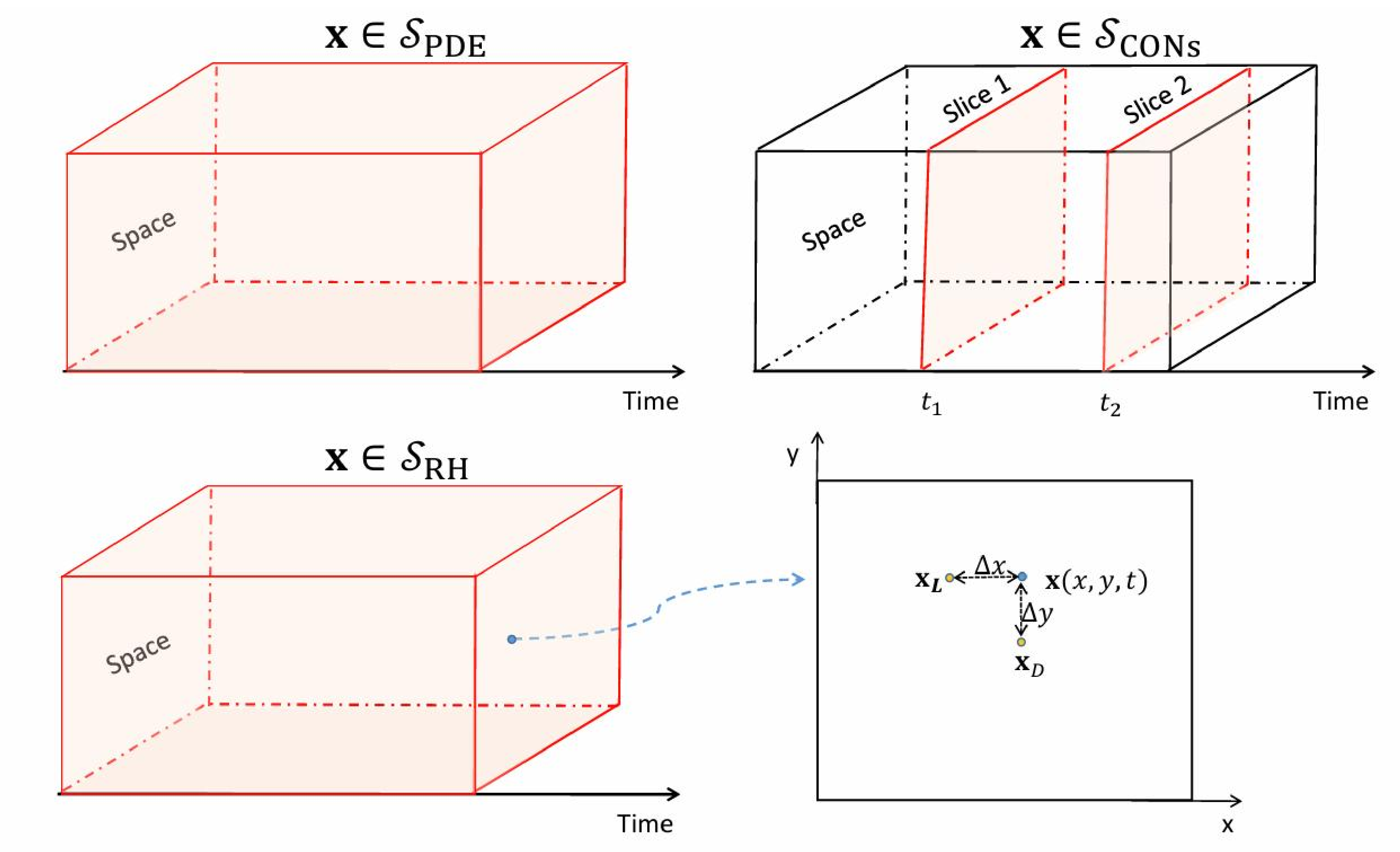
Approximation :

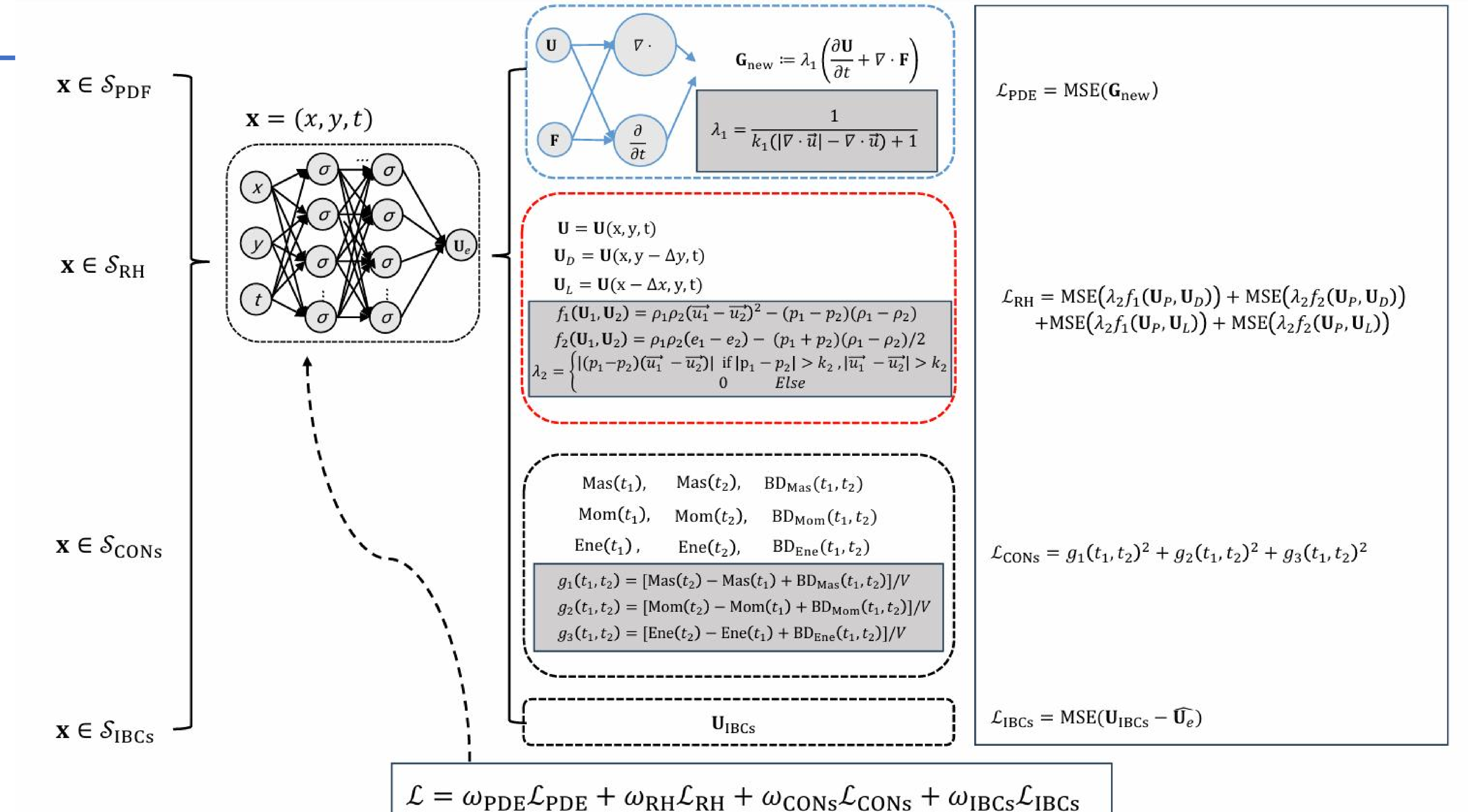
$$\mathrm{Mas}(t_k) = \frac{V}{|\mathcal{S}_{\mathrm{Con}}(t_k)|} \sum_{\mathbf{x} \in \mathcal{S}_{\mathrm{Con}}(t_k)} \rho(\mathbf{x}) \qquad \mathrm{BD}_{\mathrm{Mas}}(t_1, t_2) = \frac{(t_2 - t_1)A}{|\mathcal{S}_{\mathrm{BD}}|} \sum_{\mathbf{x} \in \mathcal{S}_{\mathrm{BD}}} \rho(\mathbf{x})(\vec{u}(\mathbf{x}) \cdot \vec{n}_{\partial V}(\mathbf{x})),$$

$$\mathrm{Mom}(t_k) = \frac{V}{|\mathcal{S}_{\mathrm{Con}}(t_k)|} \sum_{\mathbf{x} \in \mathcal{S}_{\mathrm{Con}}(t_k)} \rho(\mathbf{x})\vec{u}(\mathbf{x}) \qquad \mathrm{BD}_{\mathrm{Mom}}(t_1, t_2) = \frac{(t_2 - t_1)A}{|\mathcal{S}_{\mathrm{BD}}|} \sum_{\mathbf{x} \in \mathcal{S}_{\mathrm{BD}}} \left[\rho(\mathbf{x})(\vec{u}(\mathbf{x}))^2 + p(\mathbf{x})\right] \cdot \vec{n}_{\partial V}(\mathbf{x}),$$

$$\mathrm{Ene}(t_k) = \frac{V}{|\mathcal{S}_{\mathrm{Con}}(t_k)|} \sum_{\mathbf{x} \in \mathcal{S}_{\mathrm{Con}}(t_k)} E(\mathbf{x}) \qquad \mathrm{BD}_{\mathrm{Ene}}(t_1, t_2) = \frac{(t_2 - t_1)A}{|\mathcal{S}_{\mathrm{BD}}|} \sum_{\mathbf{x} \in \mathcal{S}_{\mathrm{BD}}} \left[E(\mathbf{x}) + p(\mathbf{x})\right] \vec{u}(\mathbf{x}) \cdot \vec{n}_{\partial V}(\mathbf{x}),$$

Sampling space

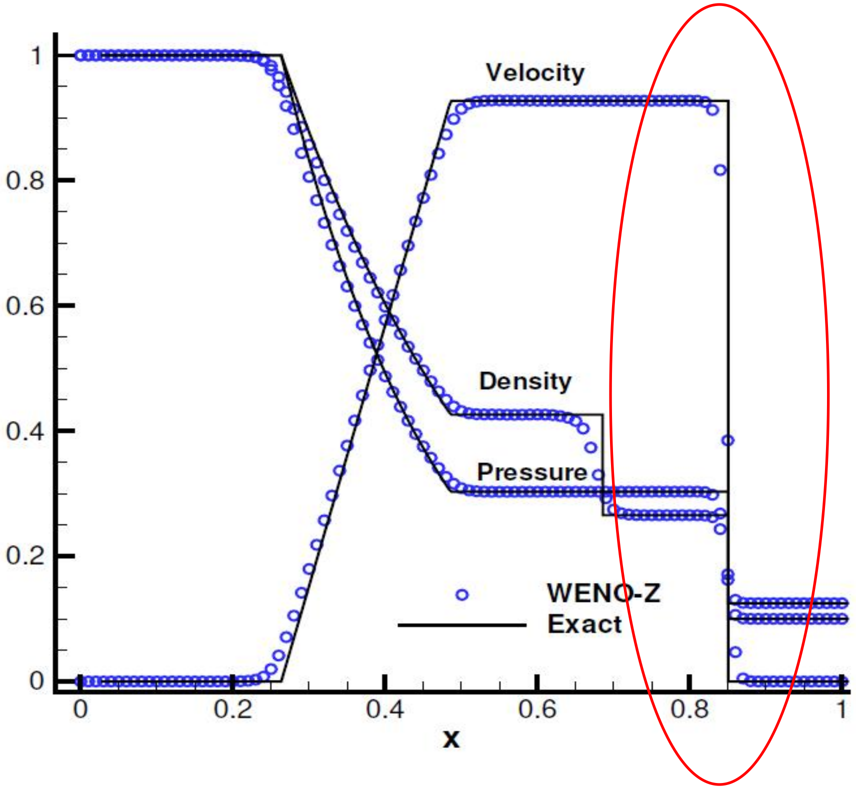



$\mathbf{x} \in \mathcal{S}_{\mathrm{PDF}}$
$\mathbf{x} \in \mathcal{S}_{\mathrm{RH}}$
$\mathbf{x} \in \mathcal{S}_{\mathrm{CONs}}$
$\mathbf{x} \in \mathcal{S}_{\mathrm{IBCs}}$
$\mathbf{x} = (x, y, t)$
x
y
t
σ
$\mathbf{U}_e$
$\mathbf{U}$
$\mathbf{F}$
$\nabla \cdot$
$\frac{\partial}{\partial t}$
$\mathbf{G}_{\mathrm{new}} := \lambda_1 \left(\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F} \right)$
$\lambda_1 = \frac{1}{k_1(|\nabla \cdot \vec{u}| - \nabla \cdot \vec{u}) + 1}$
$\mathcal{L}_{\mathrm{PDE}} = \mathrm{MSE}(\mathbf{G}_{\mathrm{new}})$
$\mathbf{U} = \mathbf{U}(\mathrm{x}, \mathrm{y}, \mathrm{t})$
$\mathbf{U}_D = \mathbf{U}(\mathrm{x}, \mathrm{y} - \Delta y, \mathrm{t})$
$\mathbf{U}_L = \mathbf{U}(\mathrm{x} - \Delta x, \mathrm{y}, \mathrm{t})$
$f_1(\mathbf{U}_1, \mathbf{U}_2) = \rho_1\rho_2(\overrightarrow{u_1} - \overrightarrow{u_2})^2 - (p_1 - p_2)(\rho_1 - \rho_2)$
$f_2(\mathbf{U}_1, \mathbf{U}_2) = \rho_1\rho_2(e_1 - e_2) - (p_1 + p_2)(\rho_1 - \rho_2)/2$
$\lambda_2 = \begin{cases} |(p_1 - p_2)(\overrightarrow{u_1} - \overrightarrow{u_2})| & \text{if } |\mathrm{p}_1 - p_2| > k_2, |\overrightarrow{u_1} - \overrightarrow{u_2}| > k_2 \\ 0 & Else \end{cases}$
$\mathcal{L}_{\mathrm{RH}} = \mathrm{MSE}(\lambda_2 f_1(\mathbf{U}_P, \mathbf{U}_D)) + \mathrm{MSE}(\lambda_2 f_2(\mathbf{U}_P, \mathbf{U}_D)) + \mathrm{MSE}(\lambda_2 f_1(\mathbf{U}_P, \mathbf{U}_L)) + \mathrm{MSE}(\lambda_2 f_2(\mathbf{U}_P, \mathbf{U}_L))$
$\mathrm{Mas}(t_1)$, $\mathrm{Mas}(t_2)$, $\mathrm{BD}_{\mathrm{Mas}}(t_1, t_2)$
$\mathrm{Mom}(t_1)$, $\mathrm{Mom}(t_2)$, $\mathrm{BD}_{\mathrm{Mom}}(t_1, t_2)$
$\mathrm{Ene}(t_1)$, $\mathrm{Ene}(t_2)$, $\mathrm{BD}_{\mathrm{Ene}}(t_1, t_2)$
$g_1(t_1, t_2) = [\mathrm{Mas}(t_2) - \mathrm{Mas}(t_1) + \mathrm{BD}_{\mathrm{Mas}}(t_1, t_2)]/V$
$g_2(t_1, t_2) = [\mathrm{Mom}(t_2) - \mathrm{Mom}(t_1) + \mathrm{BD}_{\mathrm{Mom}}(t_1, t_2)]/V$
$g_3(t_1, t_2) = [\mathrm{Ene}(t_2) - \mathrm{Ene}(t_1) + \mathrm{BD}_{\mathrm{Ene}}(t_1, t_2)]/V$
$\mathcal{L}_{\mathrm{CONs}} = g_1(t_1, t_2)^2 + g_2(t_1, t_2)^2 + g_3(t_1, t_2)^2$
$\mathbf{U}_{\mathrm{IBCs}}$
$\mathcal{L}_{\mathrm{IBCs}} = \mathrm{MSE}(\mathbf{U}_{\mathrm{IBCs}} - \widehat{\mathbf{U}_e})$
$\mathcal{L} = \omega_{\mathrm{PDE}}\mathcal{L}_{\mathrm{PDE}} + \omega_{\mathrm{RH}}\mathcal{L}_{\mathrm{RH}} + \omega_{\mathrm{CONs}}\mathcal{L}_{\mathrm{CONs}} + \omega_{\mathrm{IBCs}}\mathcal{L}_{\mathrm{IBCs}}$

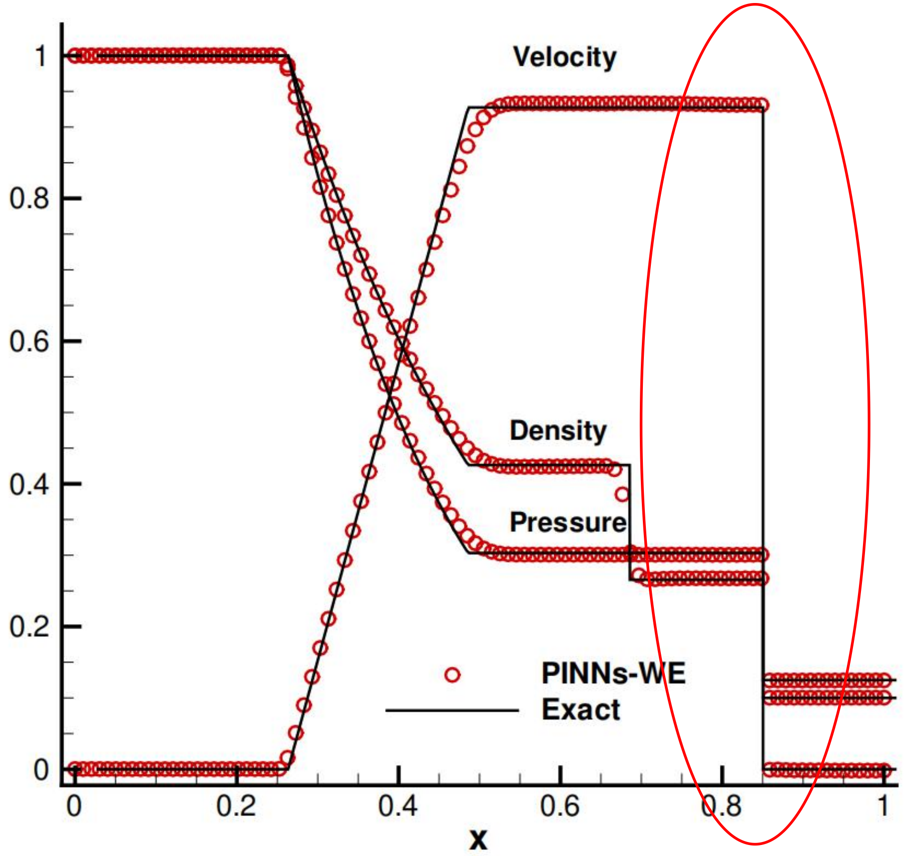
Contents

Sod Problem

- Network setting: 5 layers with 50 neurons
- Training residual points: $X * T = 100 * 100$
- Test points: Based on a grid of 100
- Comparison methods: 5th-WENO-Z, 3rd-RK-scheme, and characteristic reconstruction

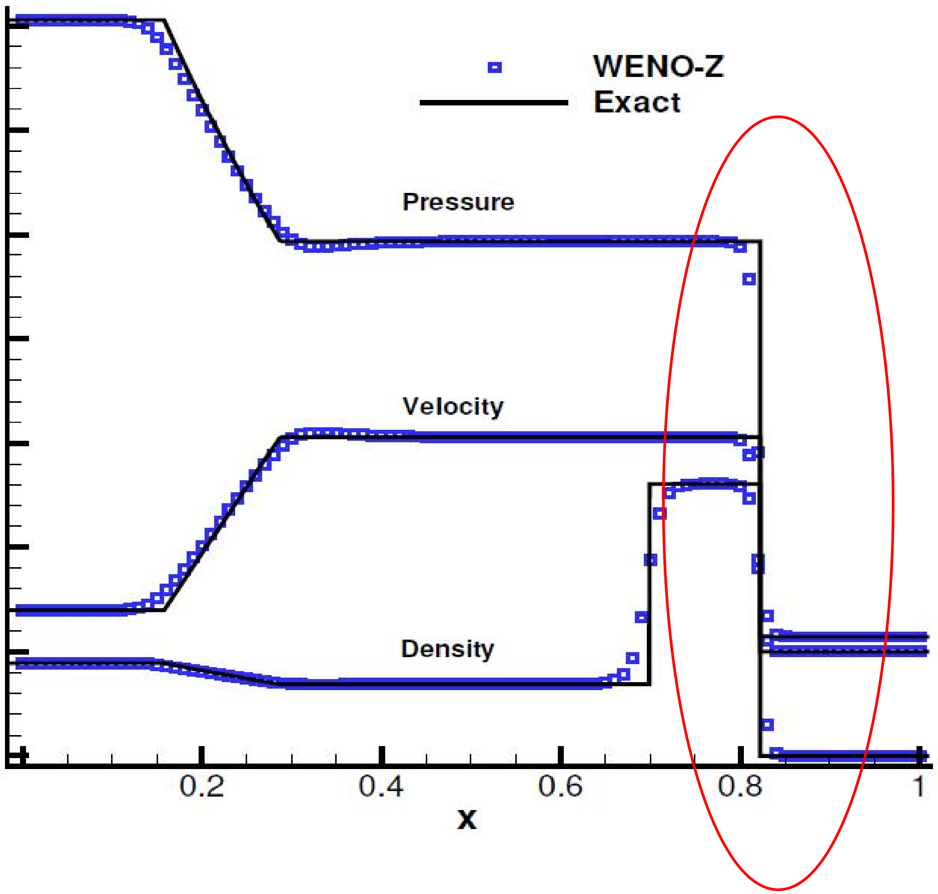


WENO-Z

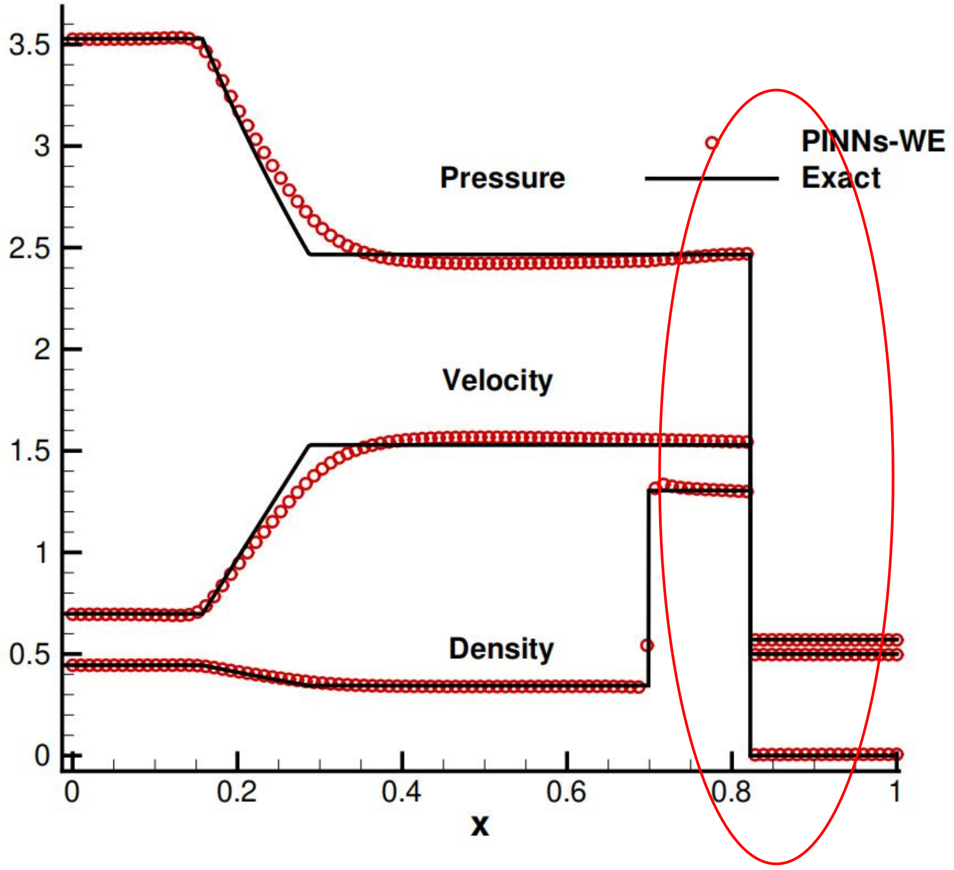


PINNs-WE

Lax Problem



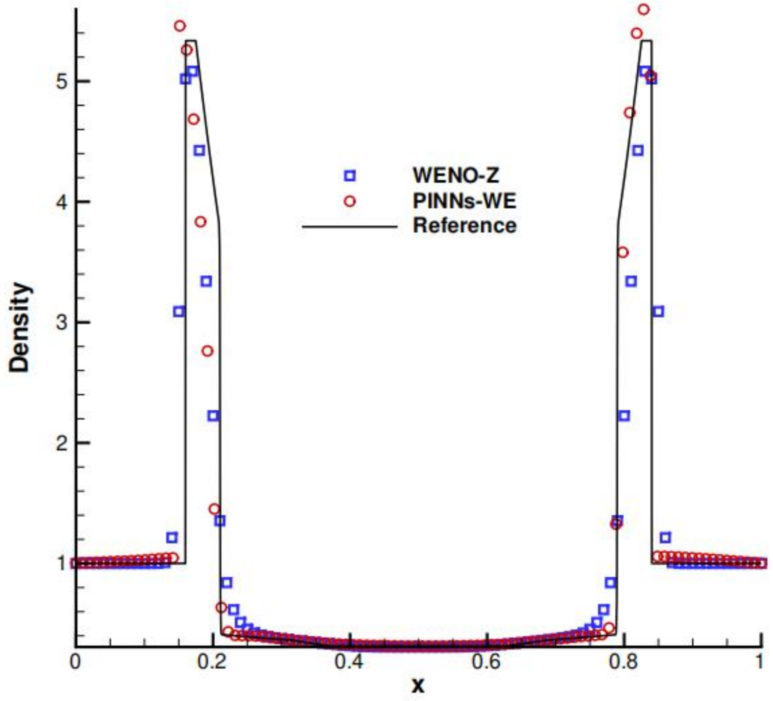
WENO-Z

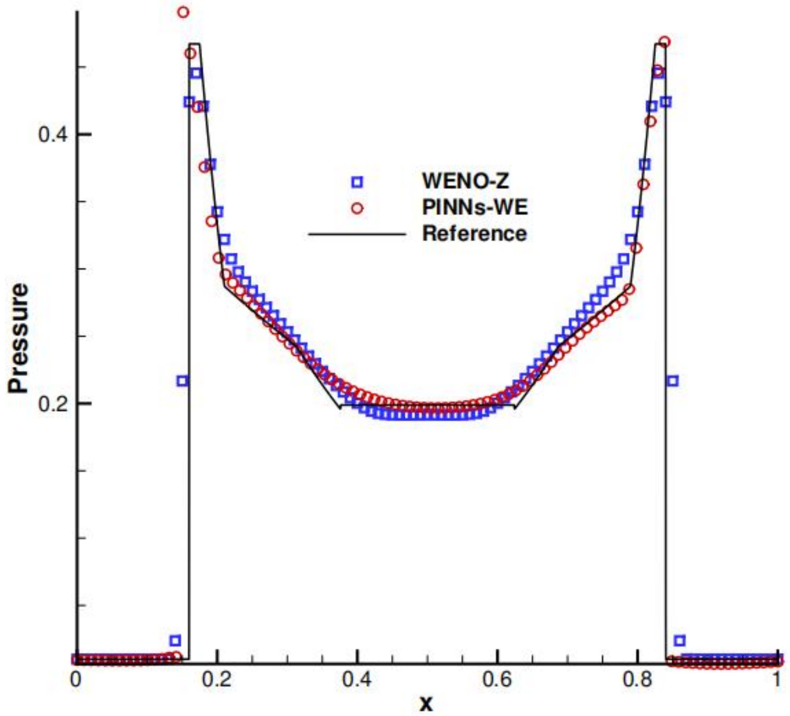


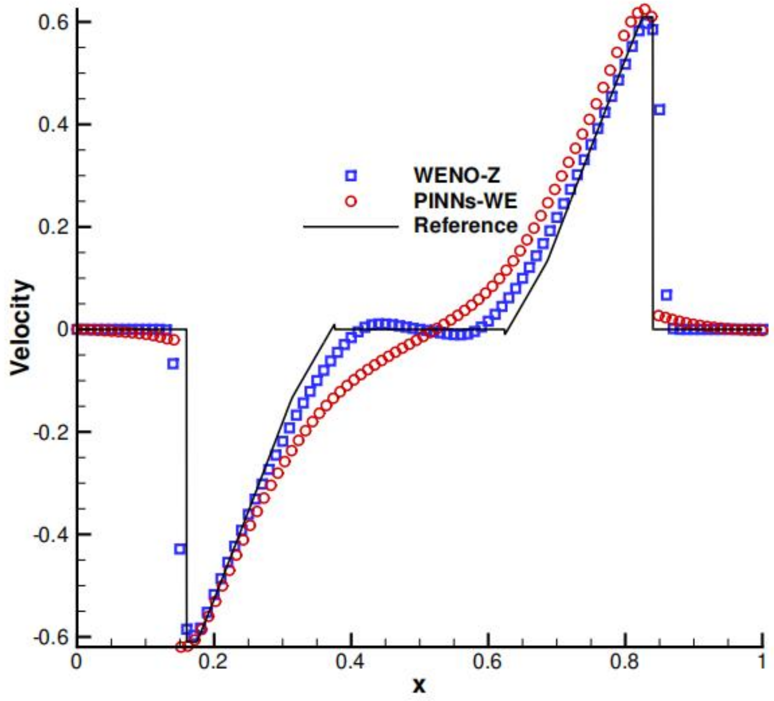
PINNs-WE

Without introducing numerical dissipation → , the shock wave can be captured more clearly

Double shock wave problem



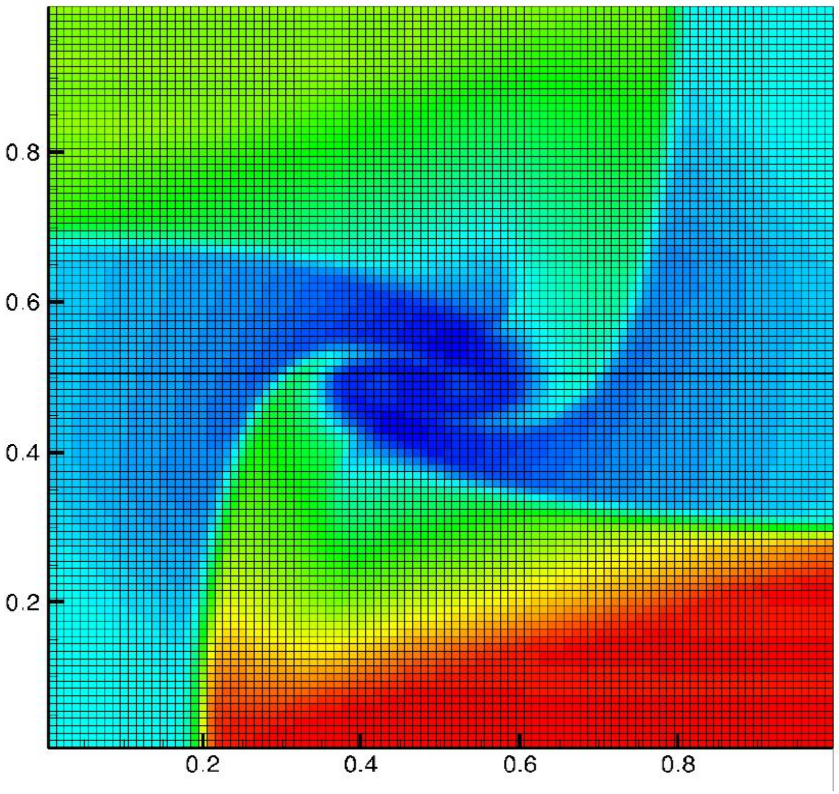




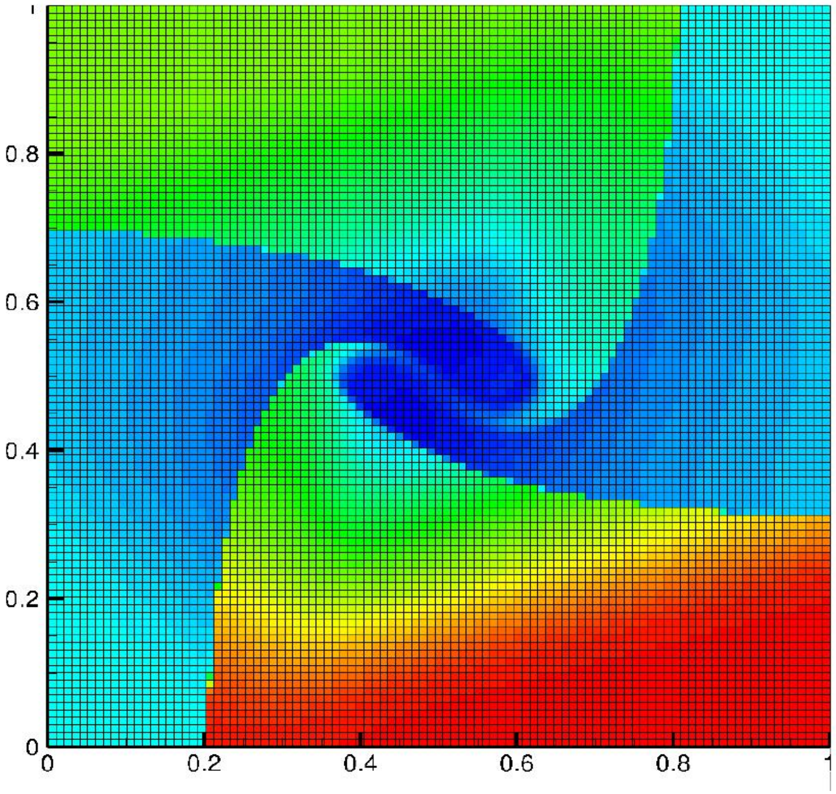
Accurate simulation can also be achieved for stronger shock waves

Two-dimensional Riemann problem

Residual points: T ∗ X
∗ Y = 100000 Latin hypercube sampling,
About 50 points per dimension
Test: Grid-based 100*100
WENO-Z: X*Y = 100*100



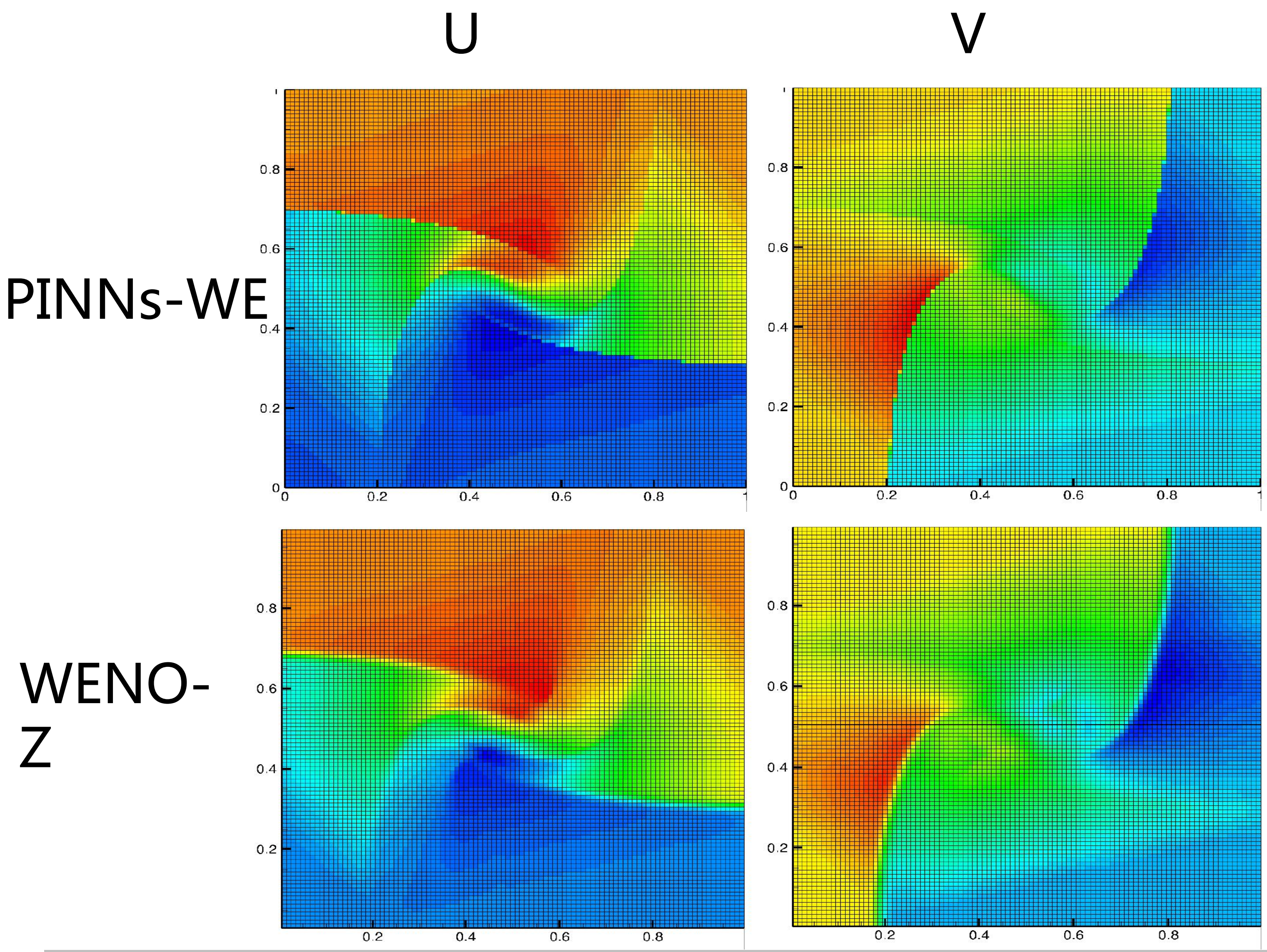
Density，WENO-Z



Density，PINNs-

Less data, clearer discontinuities

Two-dimensional Riemann problem



Less data, clearer discontinuities

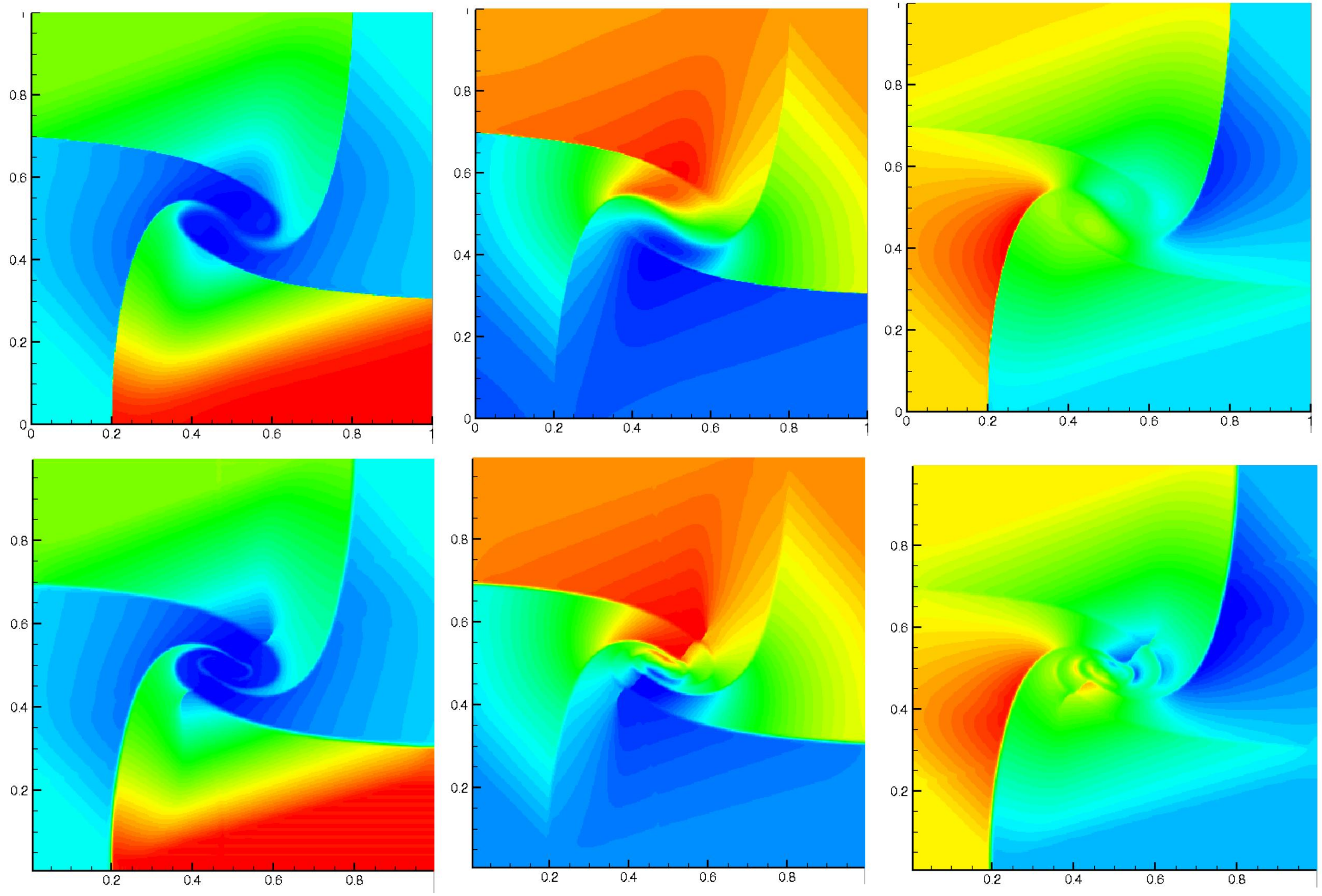
Comparison 2

Residual points: T*X*Y = 100000 Latin hypercube sampling, about 50 points per dimension

Test: Grid-based 400*400 WENO-Z: X*Y = 400*400

PINNs-WE

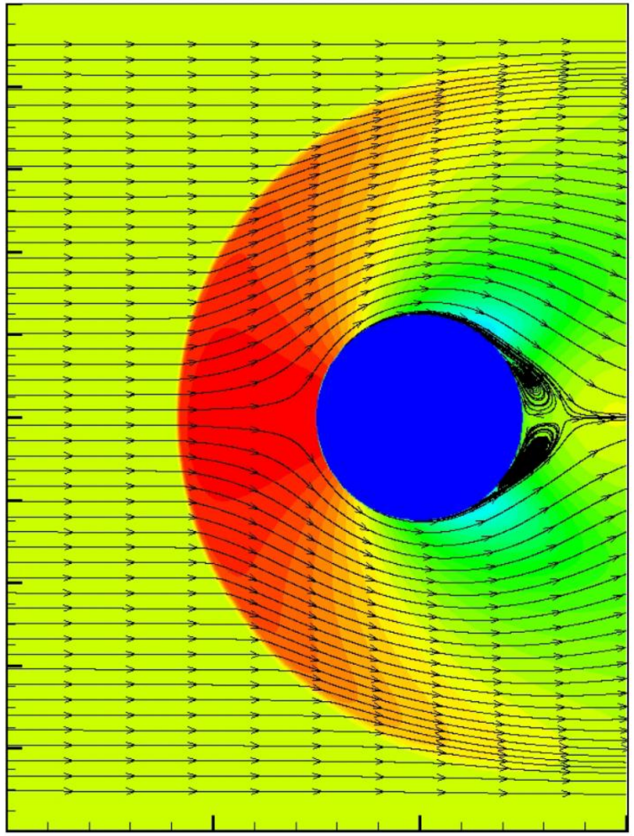
WENO-Z



Density U V

With insufficient data, large structures can be clearly solved, super-convergence

Transonic flow around a circular cylinder example



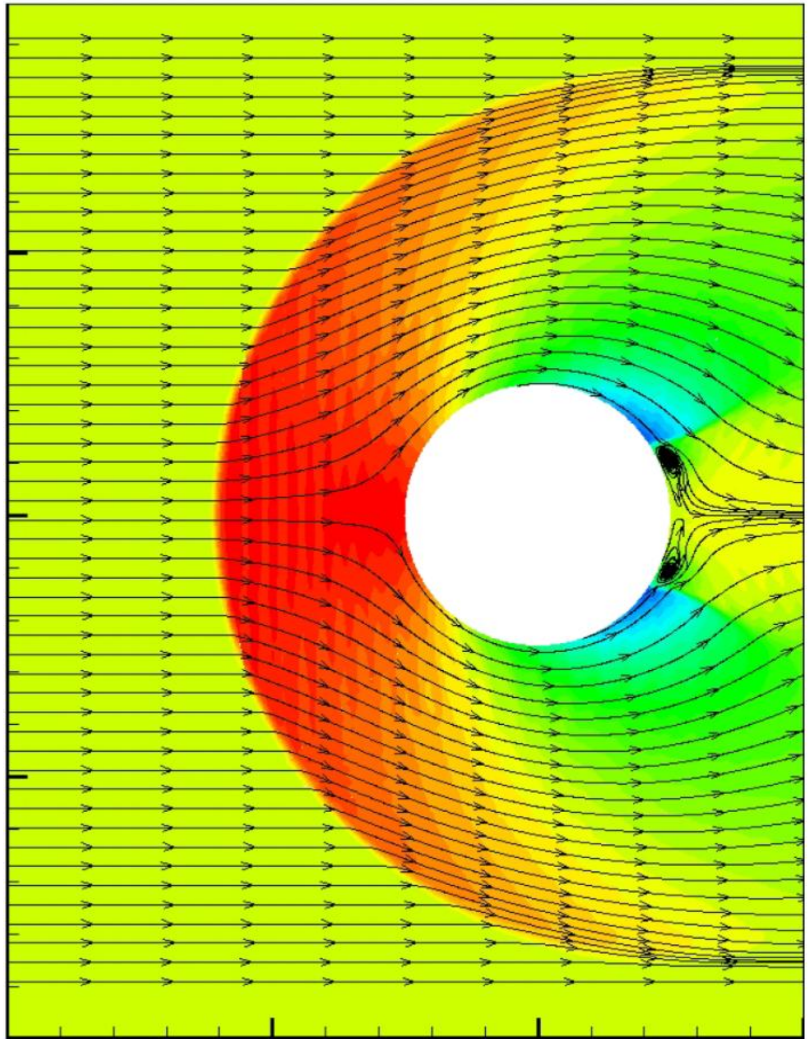
Ma=0.72 **Transonic problem with shock waves**

- **Computational domain :** $T \in [0,0.4], X \in [0,1.5], Y \in [0.2]$
- **Network setting :** 7*90*90 (All space-time: 110,000 data）
- **Sampling :** 300000，Latin hypercube (about 67 per dimension)
- **IBCs :** 15000，Latin hypercube
- **WENO- Z**：200*200 Mesh（160,000 data points at one time）

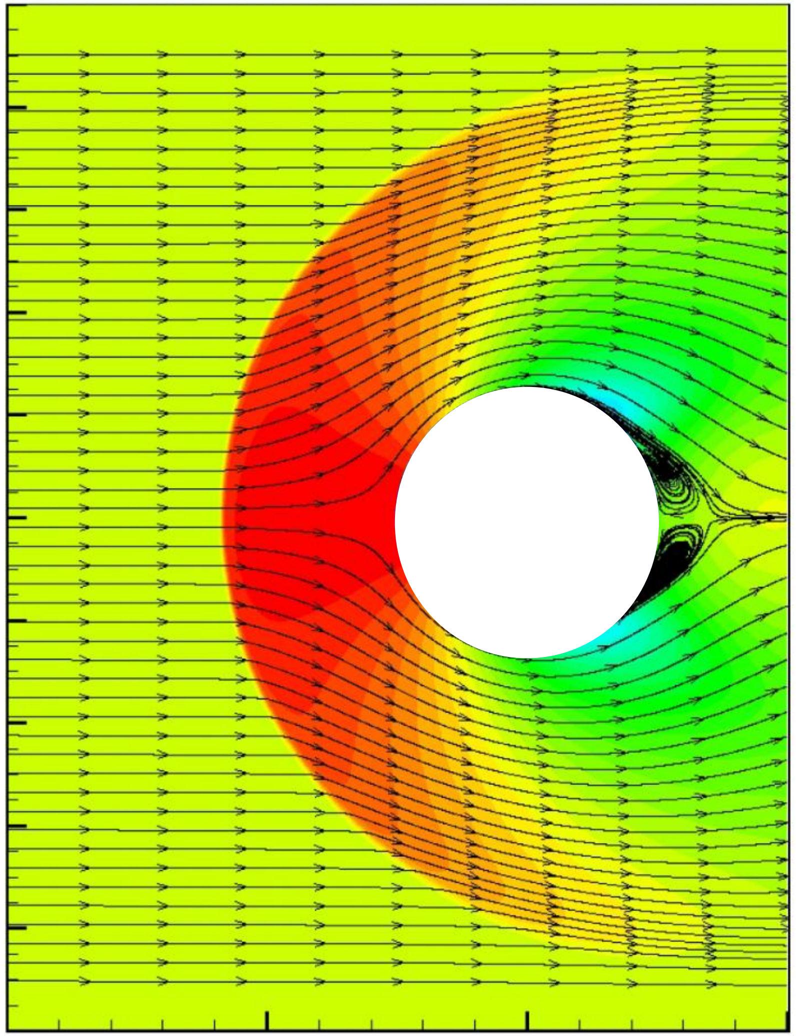
Data compression

Transonic flow around a circular cylinder example

Density/ Streamlines



WENO-Z

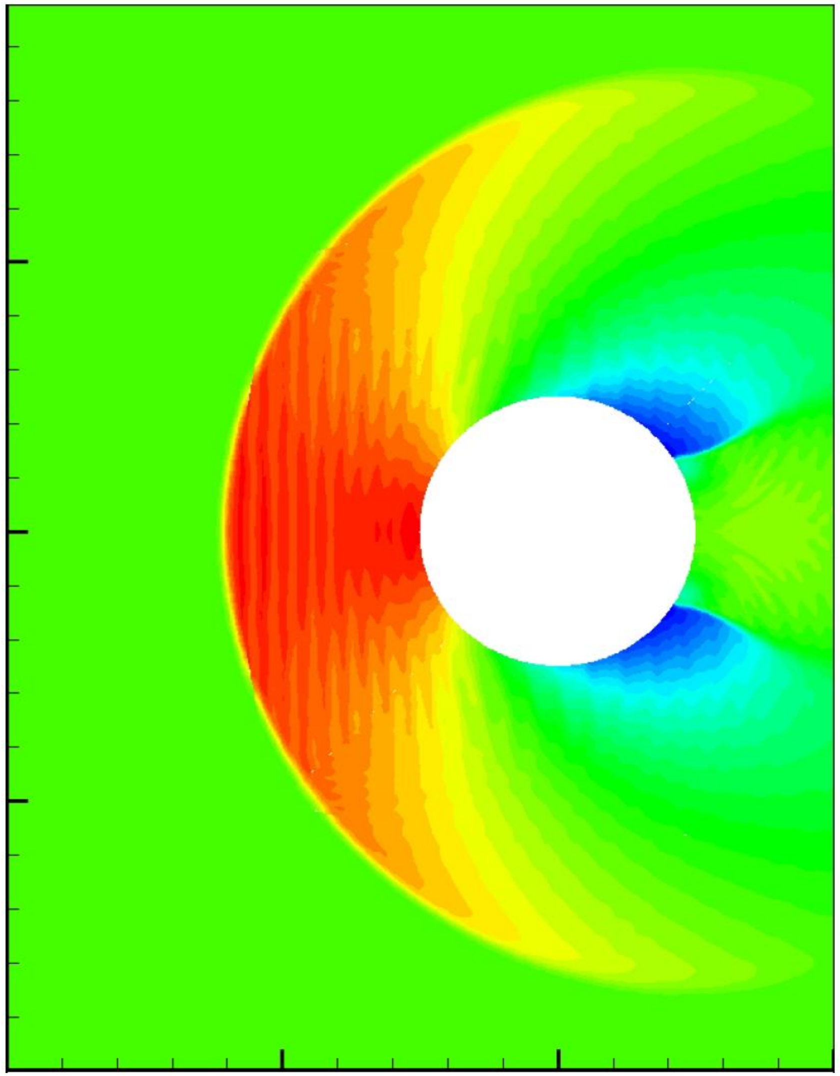


PINNs-WE

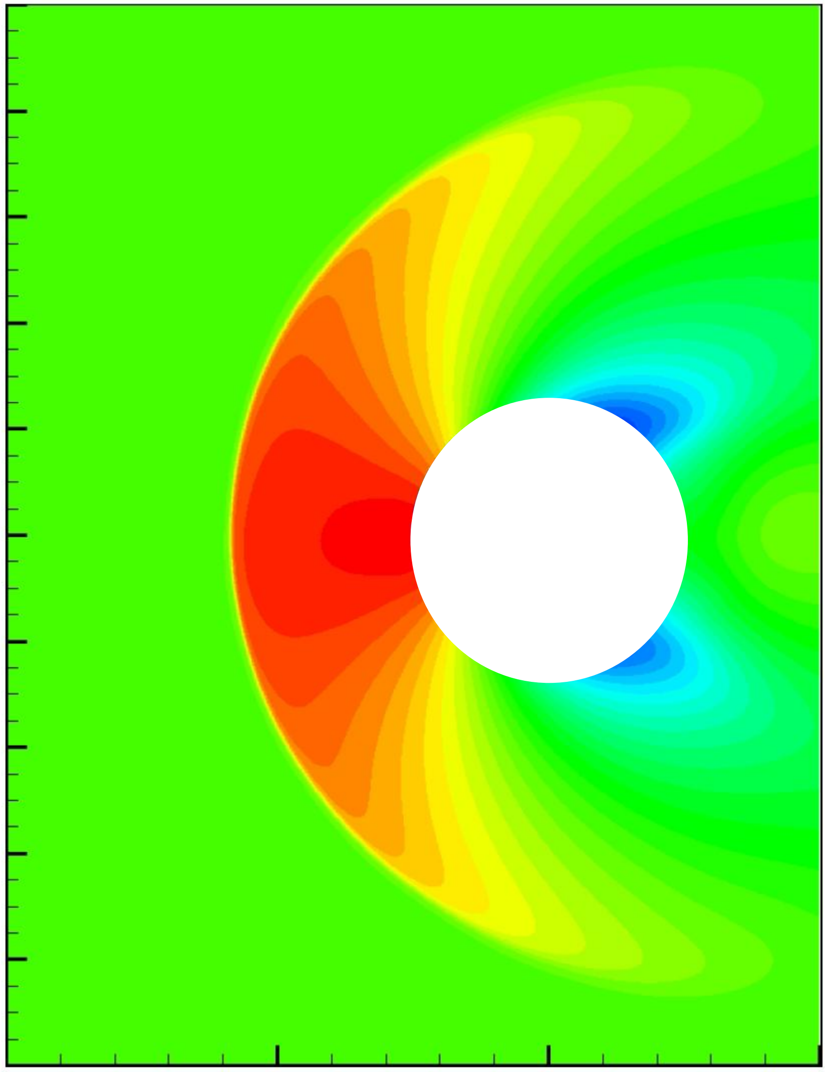
The flow field structure and density distribution are very similar.

Transonic flow around a circular cylinder example

Pressure



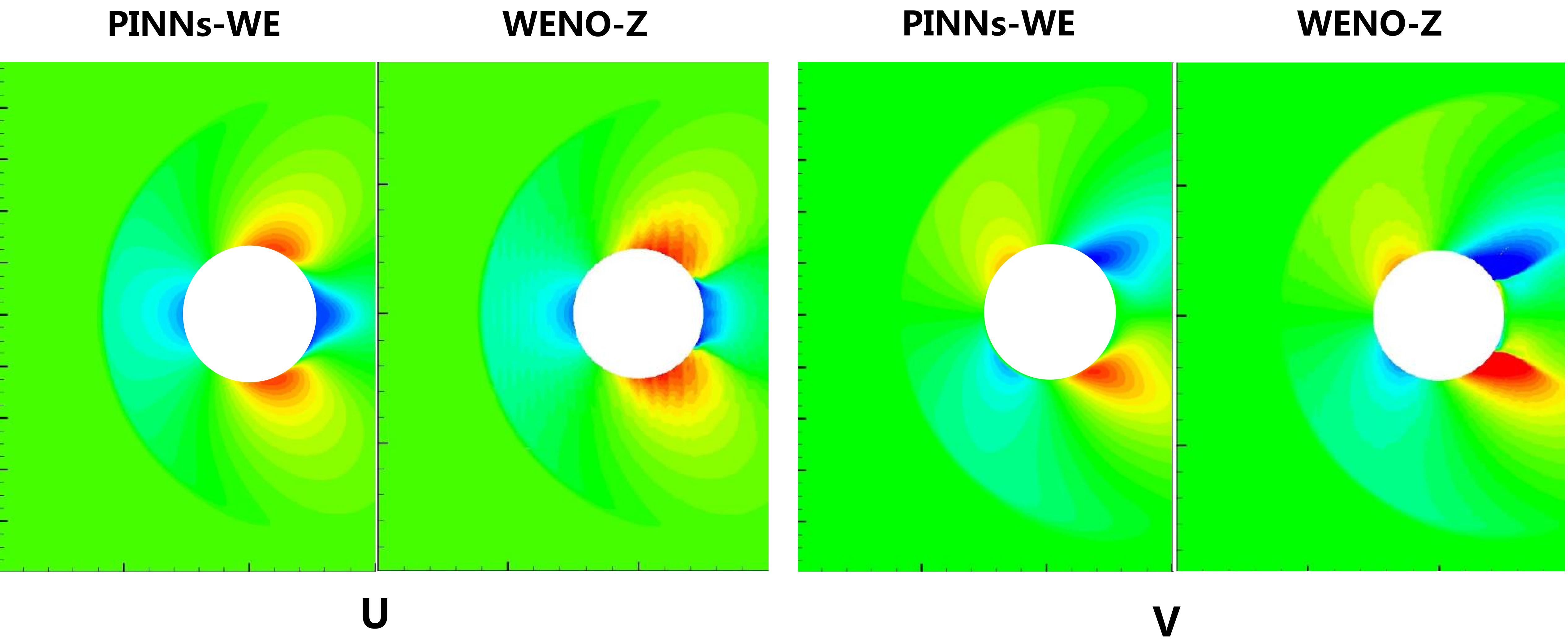
WENO-Z



PINNs-WE

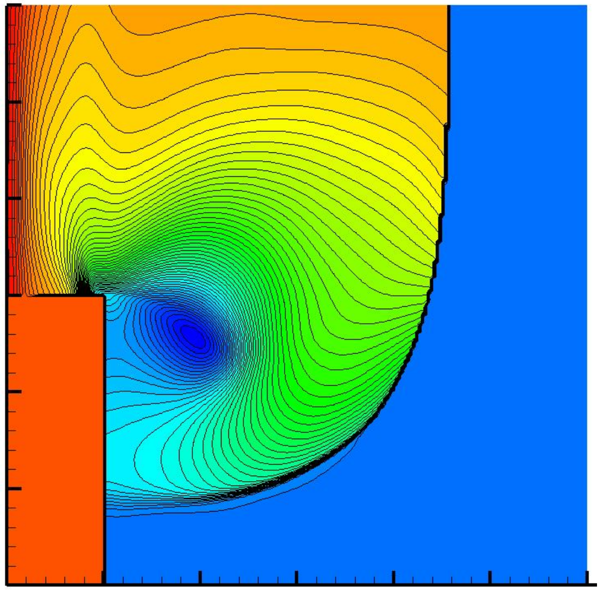
The discontinuities are sharper and the flow field is smoother.

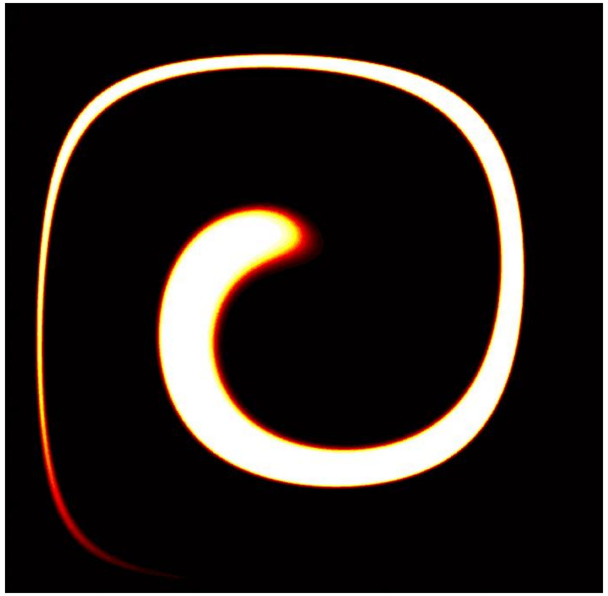
Transonic flow around a circular cylinder example

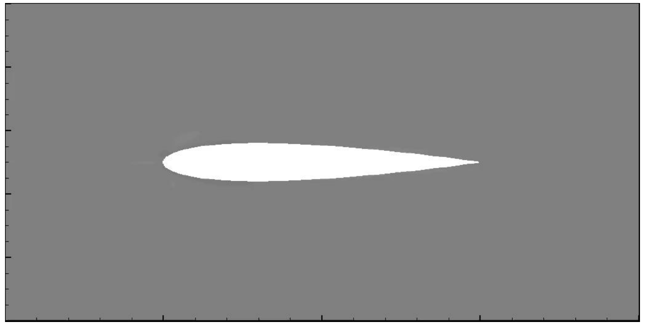


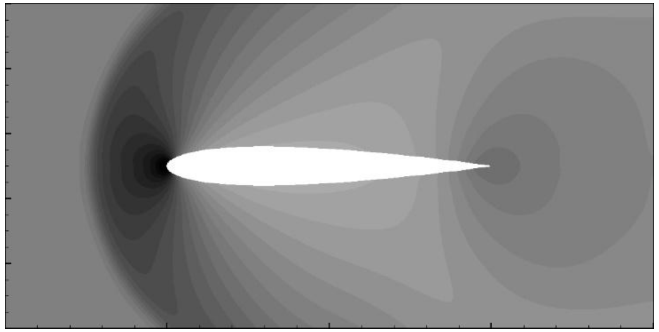
The large structures are similar and the flow field is smoother

Other cases









NACA0012

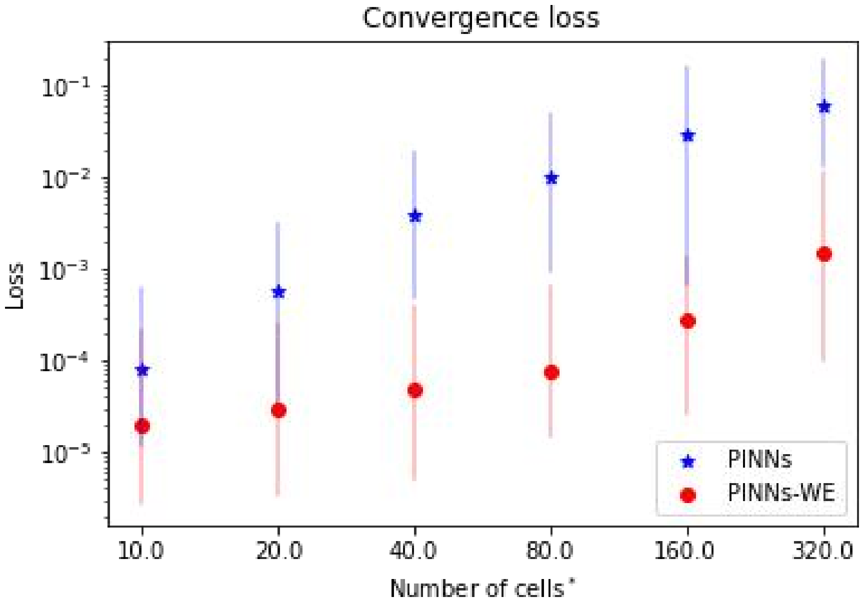
引用总数 被引用次数：61

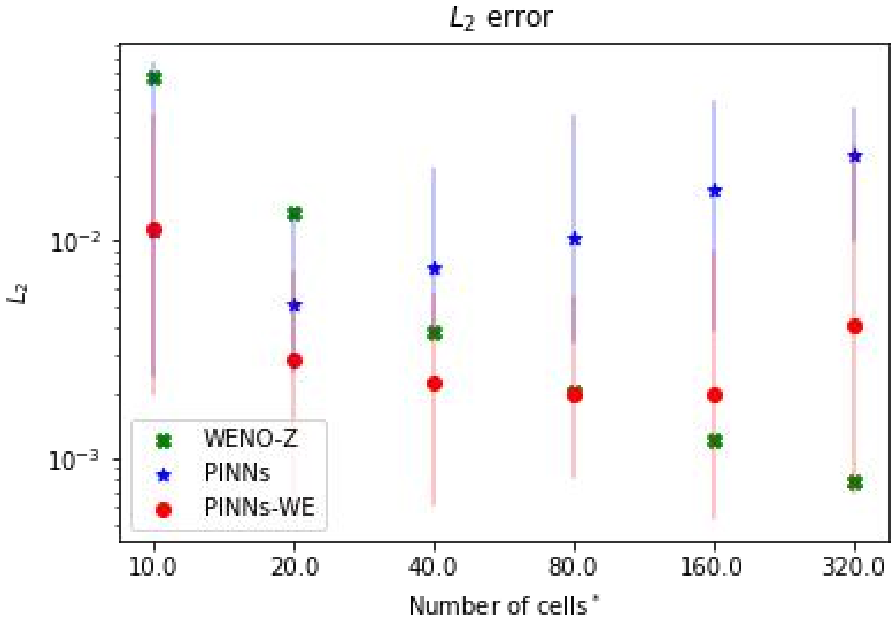
学术搜索中的文章

Discontinuity computing using physics-informed neural networks
L Liu, S Liu, H Xie, F Xiong, T Yu, M Xiao, L Liu, H Yong - Journal of Scientific Computing, 2024
被引用次数：50 相关文章 所有 9 个版本

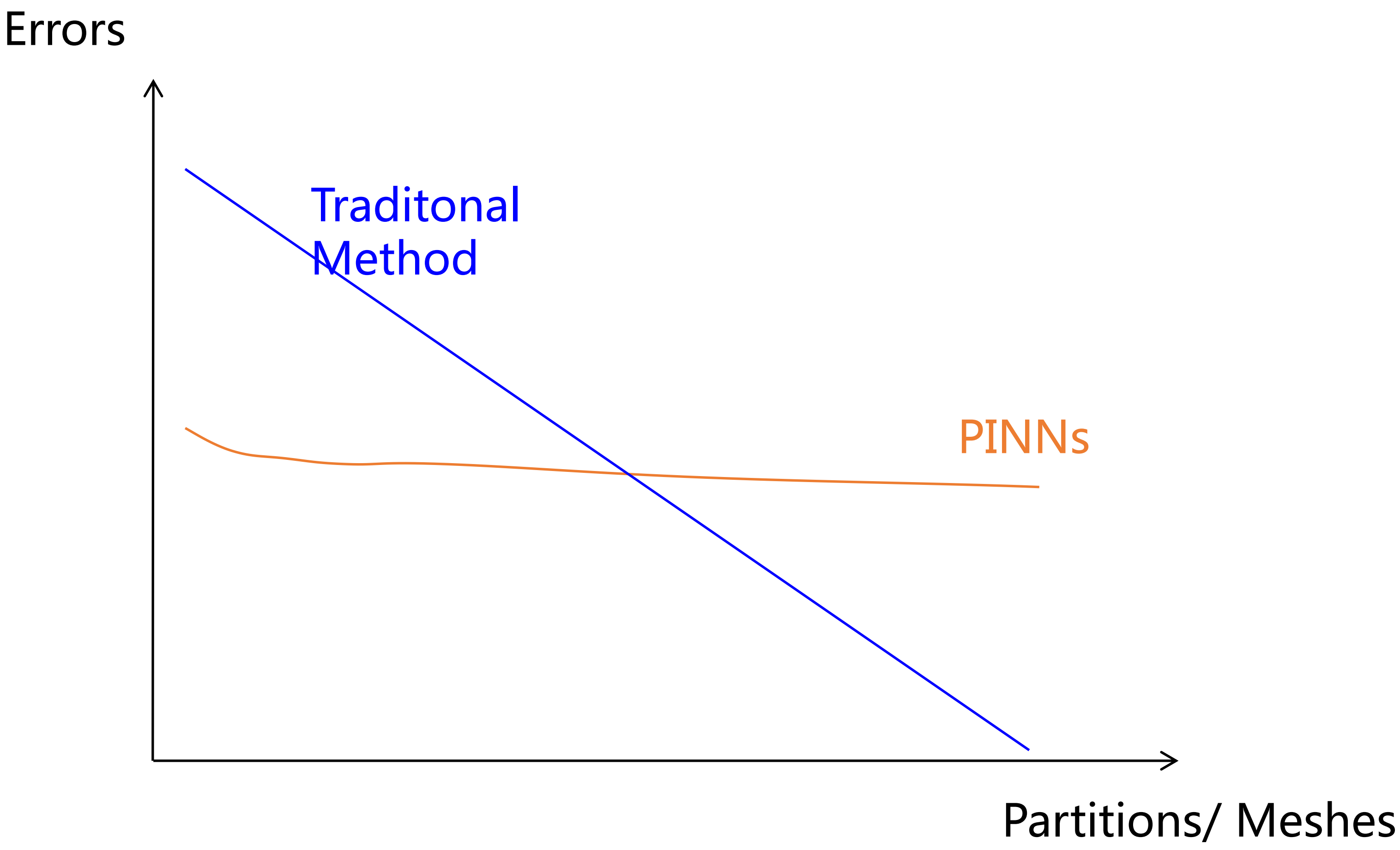
Discontinuity computing with physics-informed neural network *
L Liu, S Liu, H Yong, F Xiong, T Yu - arXiv preprint arXiv:2206.03864, 2022
被引用次数：13 相关文章 所有 2 个版本

Conclusions





- Limited by the network's expressive ability and optimization methods, there is no mesh convergence;
- -The advantages are obvious under insufficient sampling


Errors
Traditonal
Method
PINNs
Partitions/ Meshes

Advantages :

- Mesh-Free
- Sparse sampling points (for high - dimensional problems)
- Sharpness in discontinuities
- Data fusion

Disadvantages :

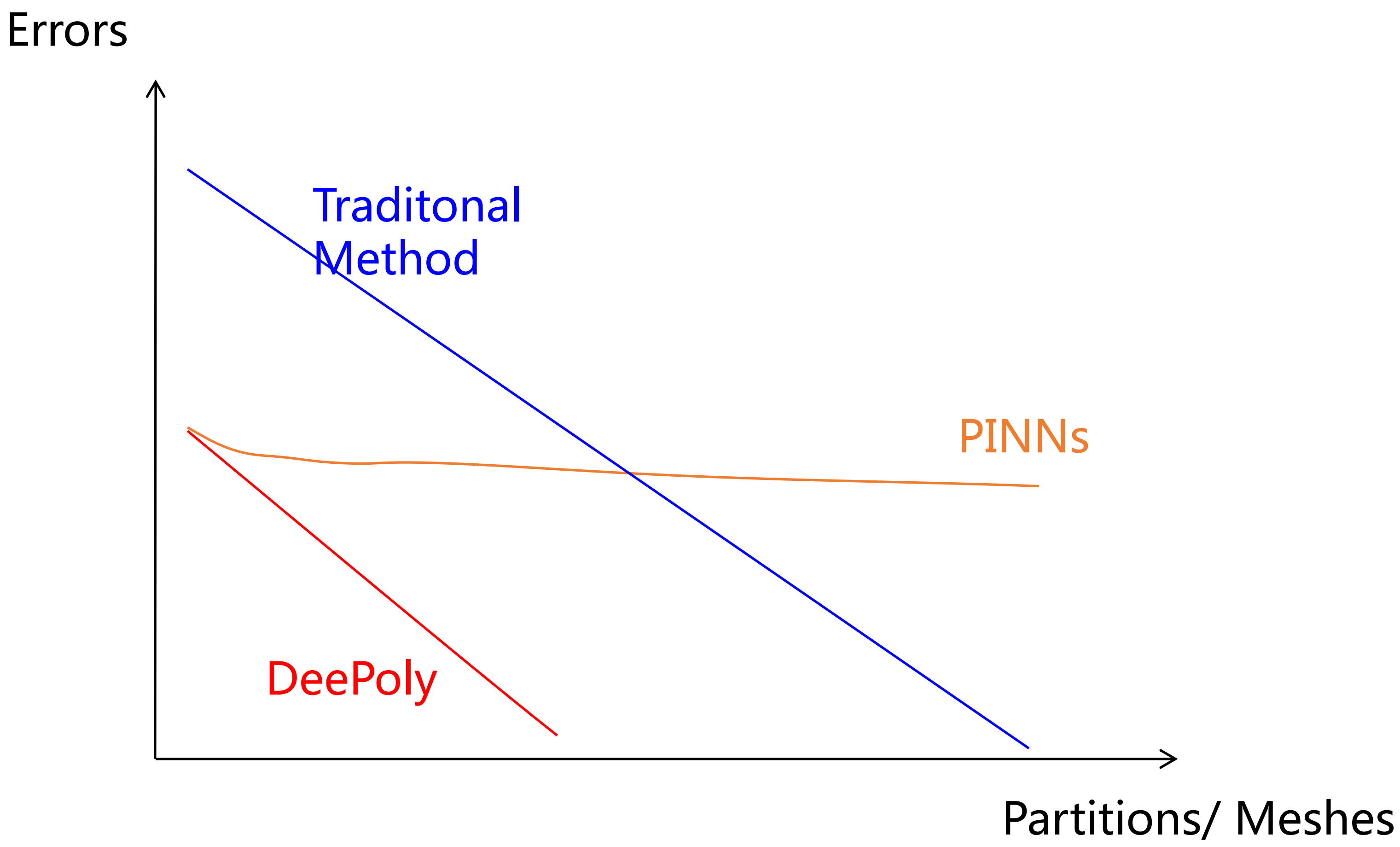
- Efficiency of forward problem solving
- Grid convergence

→ Complex problems (high frequency, multi - scale)

Outlook:

- Inverse problem:
  - Flow field inversion, filling, equation parameter calibration, equation modeling
- offline
- Localization and New framework

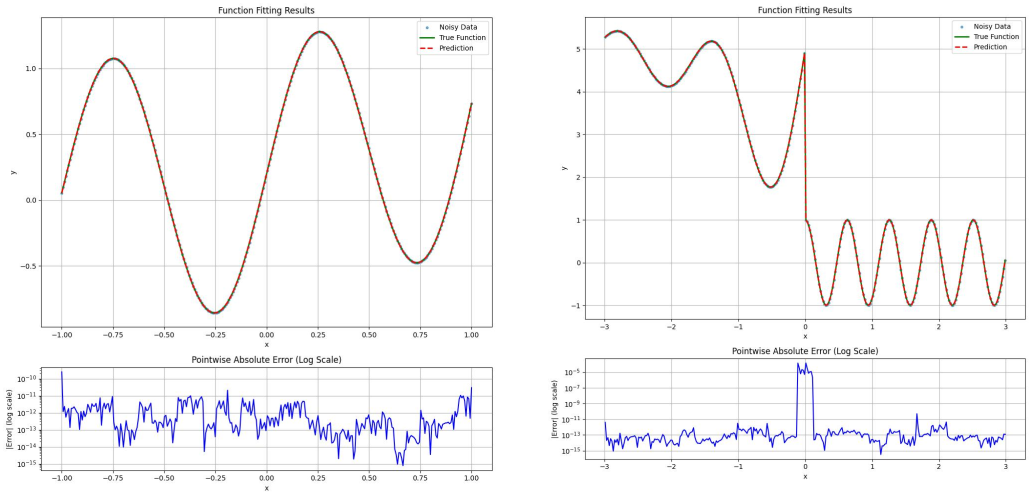
Recent Work

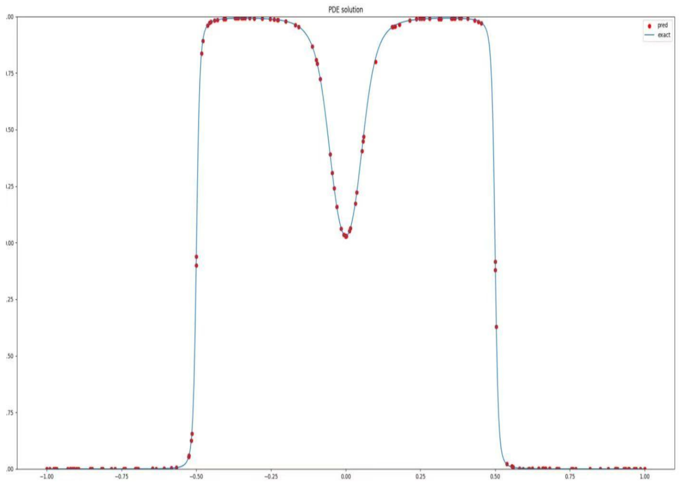


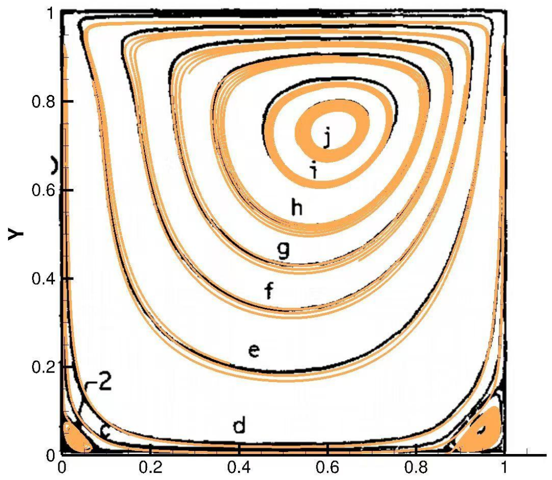
DeePoly: A High-Order Accuracy and Efficiency Deep-Polynomial Framework for Scientific Machine Learning

DeePoly

github.com/bfly123/DeePoly







Machine Error

AC equation

NS equations

- Meshfree, Schemefree based on Auto-Differential
- High-order Accuracy and Higher efficiency than PINNs
- Fit for Discontinous/Smooth Problem
- Fit for all the PDEs

Thanks！

Tel： 18911939563
email： liu_li@iapcm.ac.cn
RG: www.researchgate.net/profile/Li-Liu-72