# ODE indentification Without time label

BY ZHIJUN ZENG

2023年5月8日

## 1 Method

1. Using Clustering Method to segment data into short trajectories with initial condition and determined time length.

2. Using Neural Network $\boldsymbol{y}_\theta(t)$ to learn the solution function $\boldsymbol{y}(t)$ of ODEs $\frac{d\boldsymbol{y}}{\mathrm{dt}} = \boldsymbol{f}(\boldsymbol{y}, t)$ via Sliced Wasserstein Loss.

3. Using STRidge Regression Method (or PINN Inverse method) to identify the hidden dynamic system of $\boldsymbol{y}_\theta(t)$.

### 1.1 PINN Augmented

Here we introduce an additional loss to enforce regularity of Neural Network $\boldsymbol{y}_\theta(t)$ in step2.

1. After Sliced Wasserstein Loss smaller than threshold, Training a Representation source $\boldsymbol{f}_\phi(\boldsymbol{y}, t)$ using PINN as inverse problem.

2. Add regularity loss $L_{\mathrm{reg}} = \left\| \frac{d\boldsymbol{y}_\theta}{\mathrm{dt}} - \boldsymbol{f}_\phi(\boldsymbol{y}_\theta, \{t_j\}_{j=1}^{N_{\mathrm{reg}}}) \right\|_2^2$ in training stage and train $\{\theta, \phi\}$ at the same time.

### 1.2 Alternating Direction Optimization

By using delayed updated, we can introduce Alternating Direction Optimization in our task:

1. After Sliced Wasserstein Loss smaller than threshold, Training a Representation source $\boldsymbol{f}_\phi(\boldsymbol{y}, t)$ using STRidge Regression.

2. Add regularity loss $L_{\mathrm{reg}} = \left\| \frac{d\boldsymbol{y}_\theta}{\mathrm{dt}} - \boldsymbol{f}_\phi(\boldsymbol{y}_\theta, \{t_j\}_{j=1}^{N_{\mathrm{reg}}}) \right\|_2^2$ in training stage and only train $\{\theta\}$ at training stage.

3. Update $\phi$ every $\mathrm{Iter}_{\mathrm{reg}}$ time.

### 1.3 Shooting Method

After learning with ADO or PINN, we obtain high-quality initial guess of Representation $\boldsymbol{f}_\phi(\boldsymbol{y}, t)$.We can use shooting method to obtain parameters with higher accuracy by introducing forward information.

# 2 Experiment

In our experiment , noise level means the gaussian noise of observation with sigma = noise leval * $L_\infty$ norm of data.

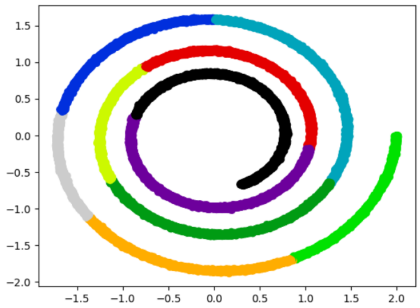
## 2.1 Low Dimensional Experiment

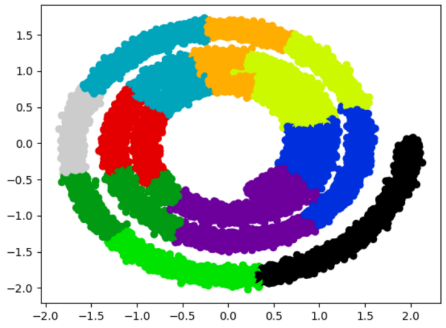
### 2.1.1 Linear2D

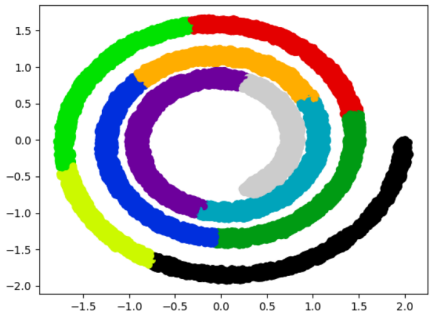
$$\frac{d}{\mathrm{dt}}\begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} -0.1 & 2 \\ -2 & -0.1 \end{bmatrix}\begin{bmatrix} q \\ p \end{bmatrix}$$

Setting:

| Validate num | T | Initial | Particle | Cluster |
|---|---|---|---|---|
| 2500 | 10 | [2.0,0.0] | 50000 | 10 |







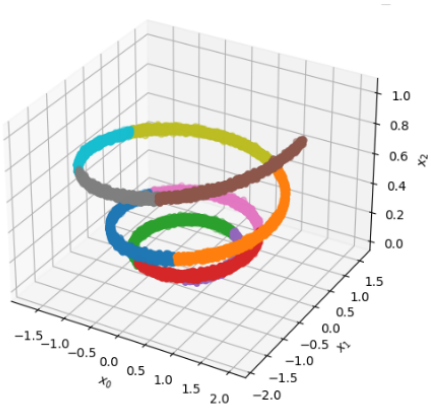
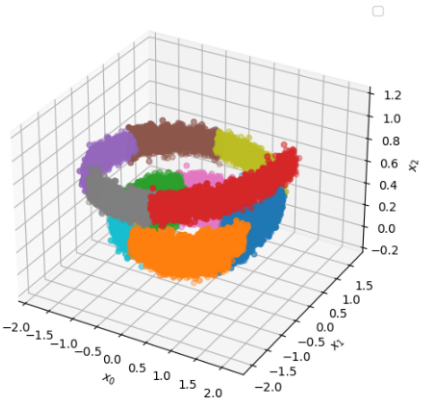
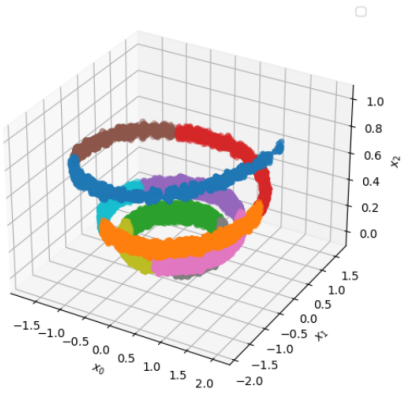
| | MAE(Parameters) | MAE(Traj) | MAE(Time) |
|---|---|---|---|
| 0%Noise | 0.01/0.1 and 1.99/2 | 0.8% | 0.65% |
| 1%Noise | 0.02/0.1 and 1.95/2 | 1.2% | 1% |
| 5%Noise | 0.03/0.1 and 1.95/2 | 1.7% | 3.% |

### 2.1.2 Linear3D

$$\frac{d}{\mathrm{dt}}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -0.1 & -2 & 0 \\ 2 & -0.1 & 0 \\ 0 & 0 & -0.3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{1}$$

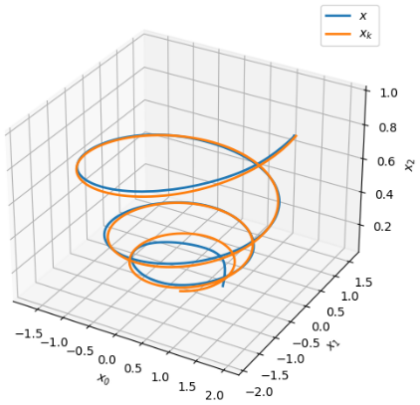
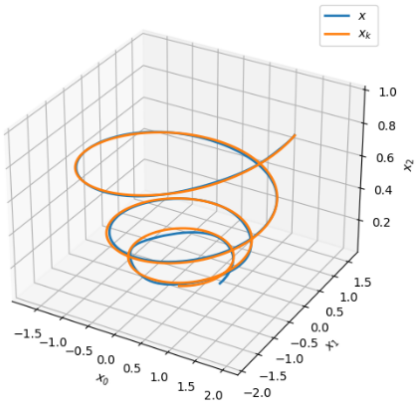
Setting:

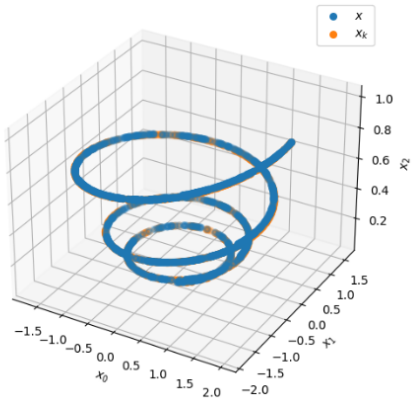
| Validate num | T | Initial | Particle | Cluster |
|---|---|---|---|---|
| 2500 | 10 | [2.0,0.0,1.0] | 50000 | 10 |

| | MAE(Parameters) | MAE(Traj) | MAE(Time) |
|---|---|---|---|
| 0%Noise | 0.01/0.1 1.99/2 0.02/0.3 | 0.6% | 1.1% |
| 1%Noise | 0.02/0.1 and 1.97/2 0.03/0.3 | 1.8% | 2.1% |
| 5%Noise | 0.02/0.1 and 1.94/2 0.04/0.3 | 2.3% | 2.9% |

Before ADO(1500 epoch)  and After ADO(3000epoch) and shooting

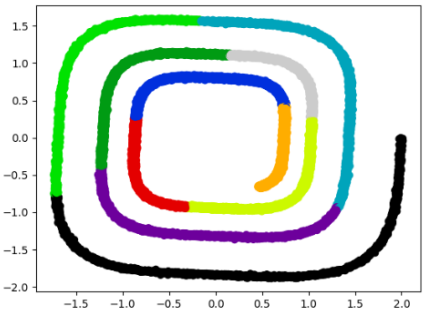



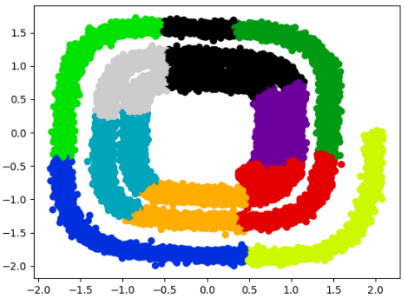


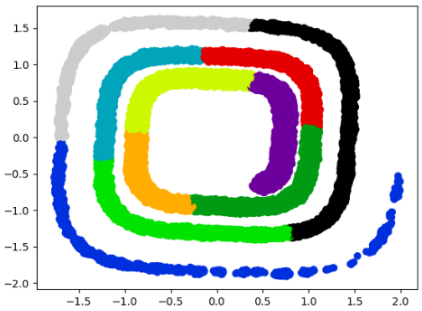

### 2.1.3 Cubic2D

$$\frac{d}{\mathrm{dt}}\begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} -0.1 & 2 \\ -2 & -0.1 \end{bmatrix}\begin{bmatrix} q^3 \\ p^3 \end{bmatrix}$$

| Validate num | T | Initial | Particle | Cluster |
|---|---|---|---|---|
| 2500 | 10 | [2.0,0.0] | 50000 | 10 |







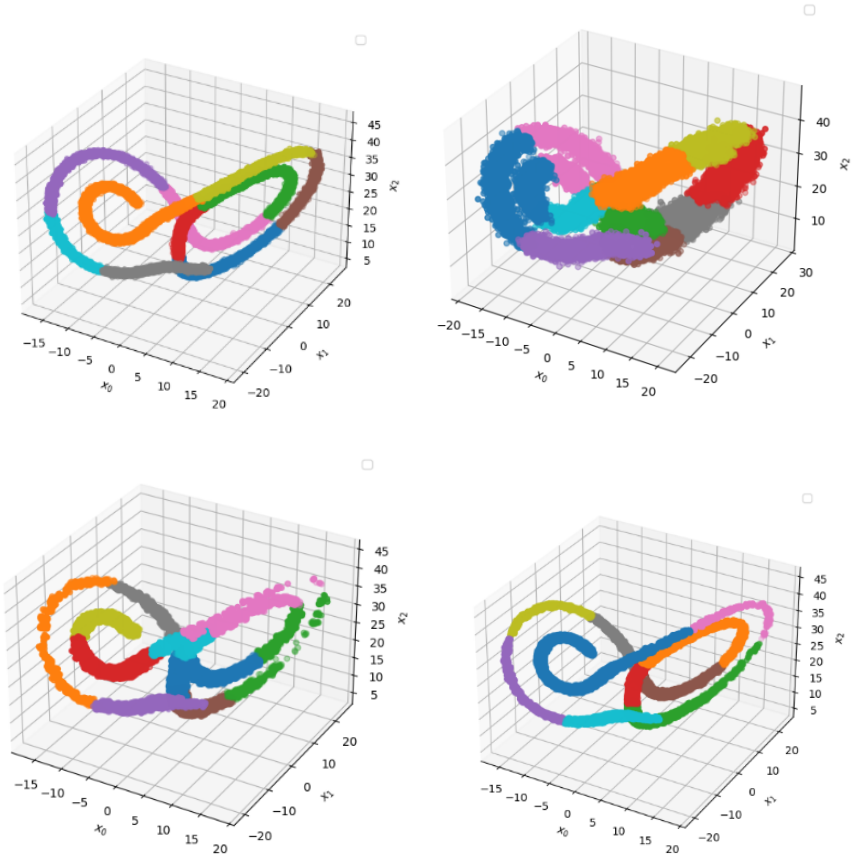
| | MAE(Parameters) | MAE(Traj) | MAE(Time) |
|---|---|---|---|
| 0%Noise | 0.02/0.1 1.98/2 | 1.% | 1.2% |
| 1%Noise | 0.02/0.1 and 1.98/2 | 2.2% | 2.2% |
| 5%Noise | 0.03/0.1 and 1.95/2 | 2.8% | 3.1% |

### 2.1.4 Lorentz

$$\begin{aligned} \dot{x} &= \sigma(y-x) \\ \dot{y} &= x(\rho-z)-y \\ \dot{z} &= \mathrm{xy}-\beta z \end{aligned} \tag{2}$$

Setting:

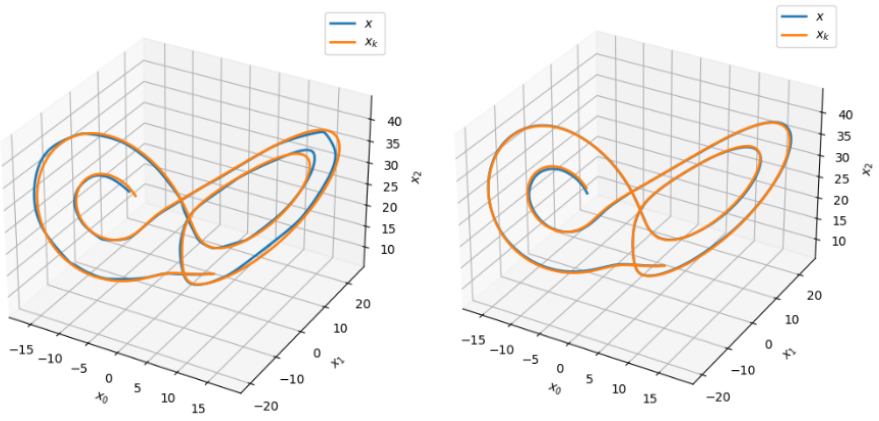
| Validate num | T | Initial | Particle | Cluster | Parameter |
|---|---|---|---|---|---|
| 2500 | 3 | [10,-10,20] | 50000 | 10 | $\sigma = 10, \beta = \frac{8}{3}, \rho = 28$ |



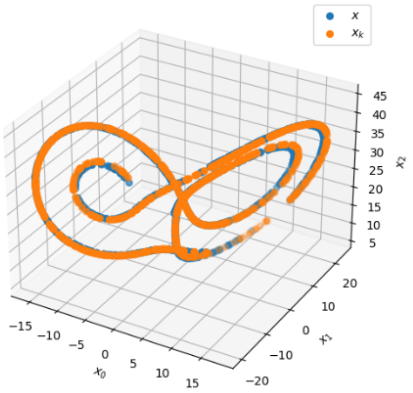
| | MAE(Parameters) | MAE(Traj) | MAE(Time) |
|---|---|---|---|
| 0%Noise | 0.5/10 0.07/1 0.15/2.6 0.93/28 | 2.% | 1.3% |
| 1%Noise | 1./10 0.09/1 0.13/2.6 1./28 | 3.1% | 2.1% |
| 5%Noise | 1.3/10 0.12/1 0.2/2.6 1.3/28 | 4.6% | 6.% 15% |

**Analysis**

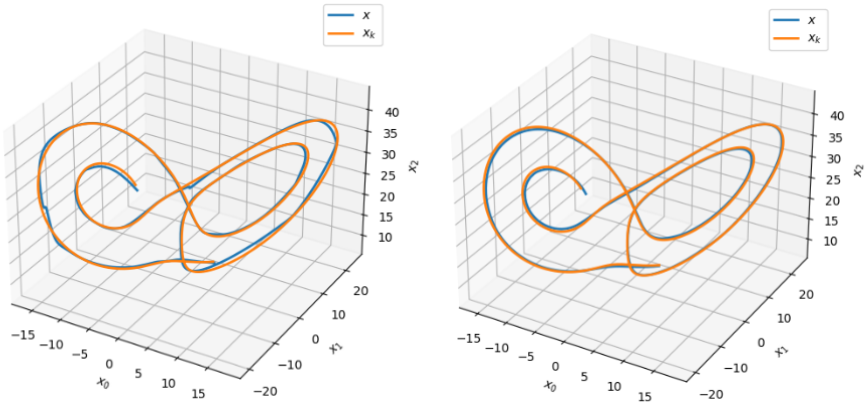
Before ADO (Pretrain 12000epoch) and After ADO (2000epoch)



ODE inference




Data with noise: (Pretrain 12000epoch) and After ADO (4000epoch)

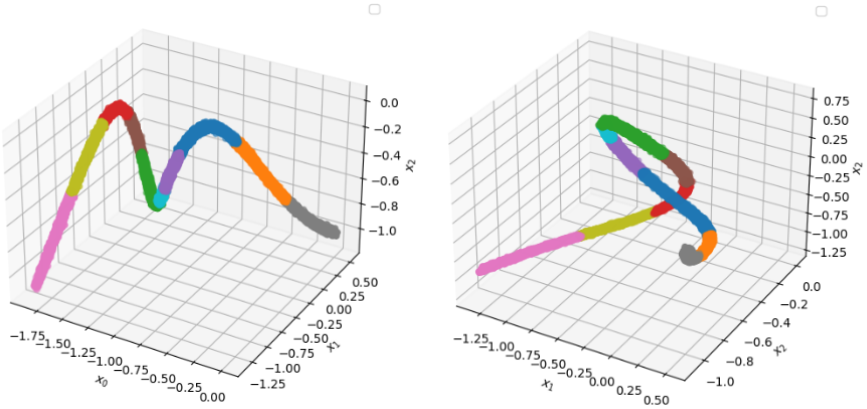


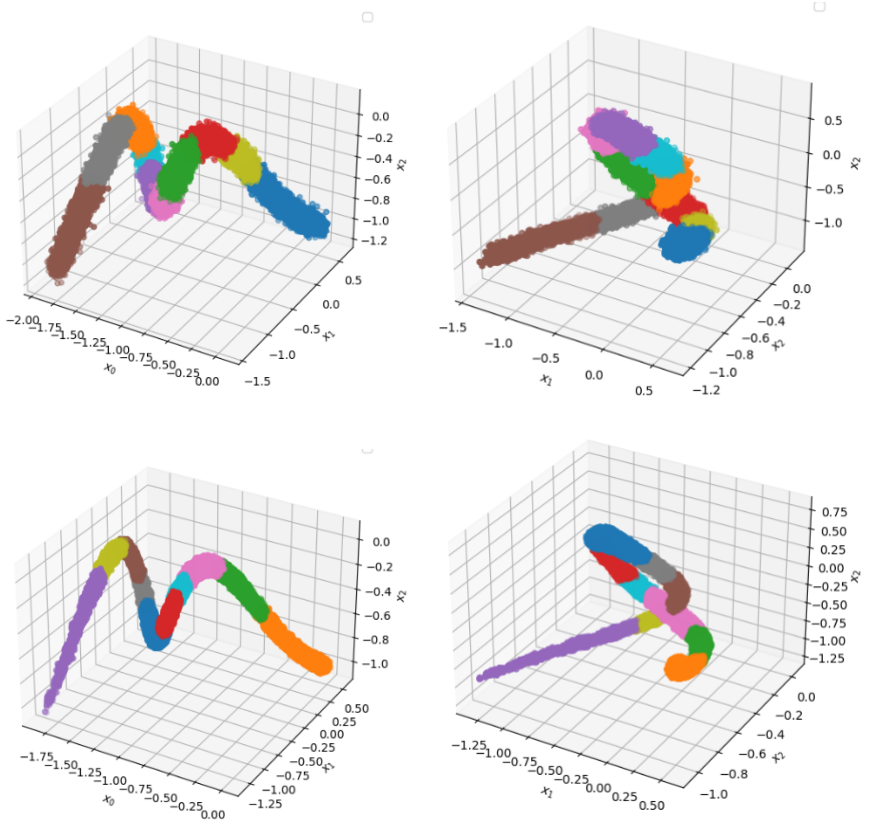

## 2.2 High Dimensional Experiment

### 2.2.1 HH model 4D

$$\frac{d}{dt}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$$
$$\frac{d}{dt}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} -q_1 - 2q_1q_2 \\ -q_2 - q_1^2 + q_2^2 \end{pmatrix}$$

| Validate num | T | Initial | Particle | Cluster |
|---|---|---|---|---|
| 2500 | 5 | [0,0.5,-1.0,0] | 50000 | 10 |





| | MAE(Parameters) | MAE(Traj) | MAE(Time) |
|---|---|---|---|
| 0%Noise | 0.03/1 | 1.2% | 1.% |
| 1%Noise | 0.035/1 | 2.1% | 2.4% |
| 5%Noise | 0.08/1 | 4.9% | 5.2% |