ASIPP

Fast control of plasma vertical displacement based on robust adversarial reinforcement learning

B.N.Liu[1,2,*], Y.H.Wang[1], W.H.Hu[1], Q.P.Yuan[1,3], B.J.Xiao[1,2,3]

[1] *Institute of Plasma Physics, Chinese Academy of Sciences, Hefei, China*
[2] *University of Science & Technology of China, Hefei, China*
[3] *Institute of Energy, Hefei Comprehensive National Science Center*
*Email: binnuo.liu@ipp.ac.cn

Outline

Background

- **Vertical displacement is unstable and must be feedback controlled**

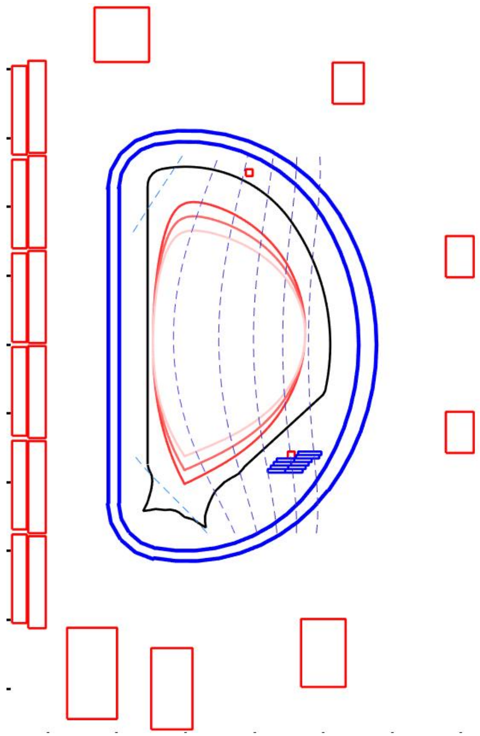
  Loss of Control -> Disruption

- **Vertical displacement is affected by passive structures, power delays, etc ...**

  higher order system ＆ Response complexity

- **System capacity is limited and perturbations are complex and varied**

  High demand for robustness



Plasma with elongated shapes has vertical displacement instability

Problem：Traditional PID controllers expose the IC coil to overcurrent risk when facing PF coil current perturbation

Background

- **Solution to Traditional Vertical Displacement Control**
  - **Filter PD Control**

    Avoid overcurrent due to slow disturbance of PF coil current by high pass filtering

    The system is still unstable and has the possibility of slow overcurrent

    Control parameters are adjusted based on simulations or experiments

    It is hard to get the optimal parameters
  - **Velocity feedback control**

    Feedback velocity and Ic coil current to avoid overcurrent due to slow perturbation

    Parameter range is narrow, difficult to adapt the response to different shapes

    Adding a compensator to adjust the dynamic response

    Compensators cause additional jitter

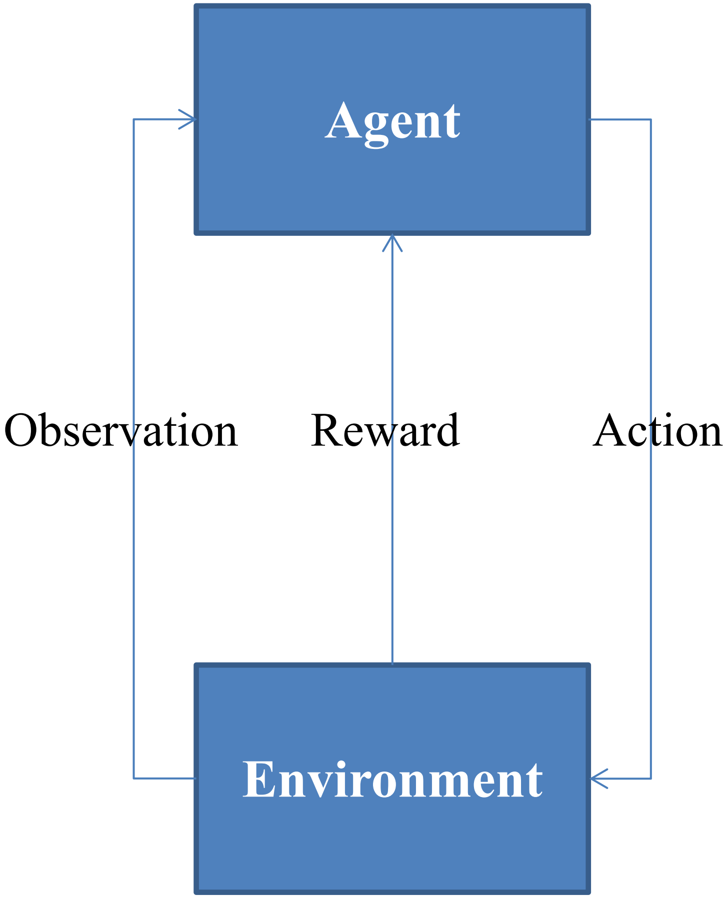
Background

- **Reinforcement learning(RL) can probably solve this problem**

  The strategy of RL agent is to choose the action that maximizes the reward based on the observations

  Add IC current to reward function **so that the agent's strategy will tend to prevent its overcurrent**

  Agent are able to extract information related to the equilibrium position of the horizontal field from the observations, **thus avoiding the need to stabilize z with a large IC current**



Outline

ASIPP

Robust Adversarial Reinforcement Learning

- **Goal:**

  The controller needs to avoid IC coil overcurrent

  under the perturbation of PF current

  while controlling z

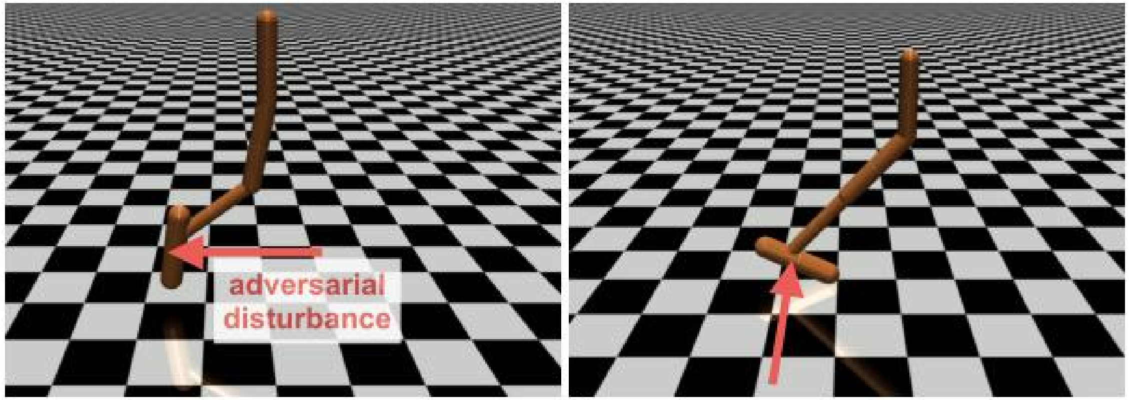
- **Robust Adversarial Reinforcement Learning（RARL）**

  RARL adds an adversary to the agent, who is also an agent

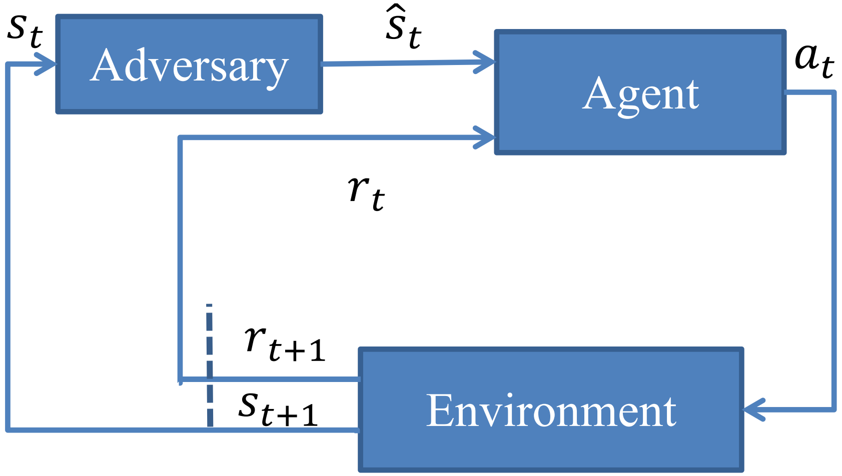
  The adversary will attack the agent's weaknesses

  The agent need to find optimal strategies in worst case

  Add an adversary to attack I_{PF}



Perturbation of control objects



Perturbation of observations

Physical Model

- Physical Model

In this study, the RZIP rigid response model is used to simulate the EAST system, and according to the RZIP model the plasma linear response matrix can be expressed by the state space equation in the following form:

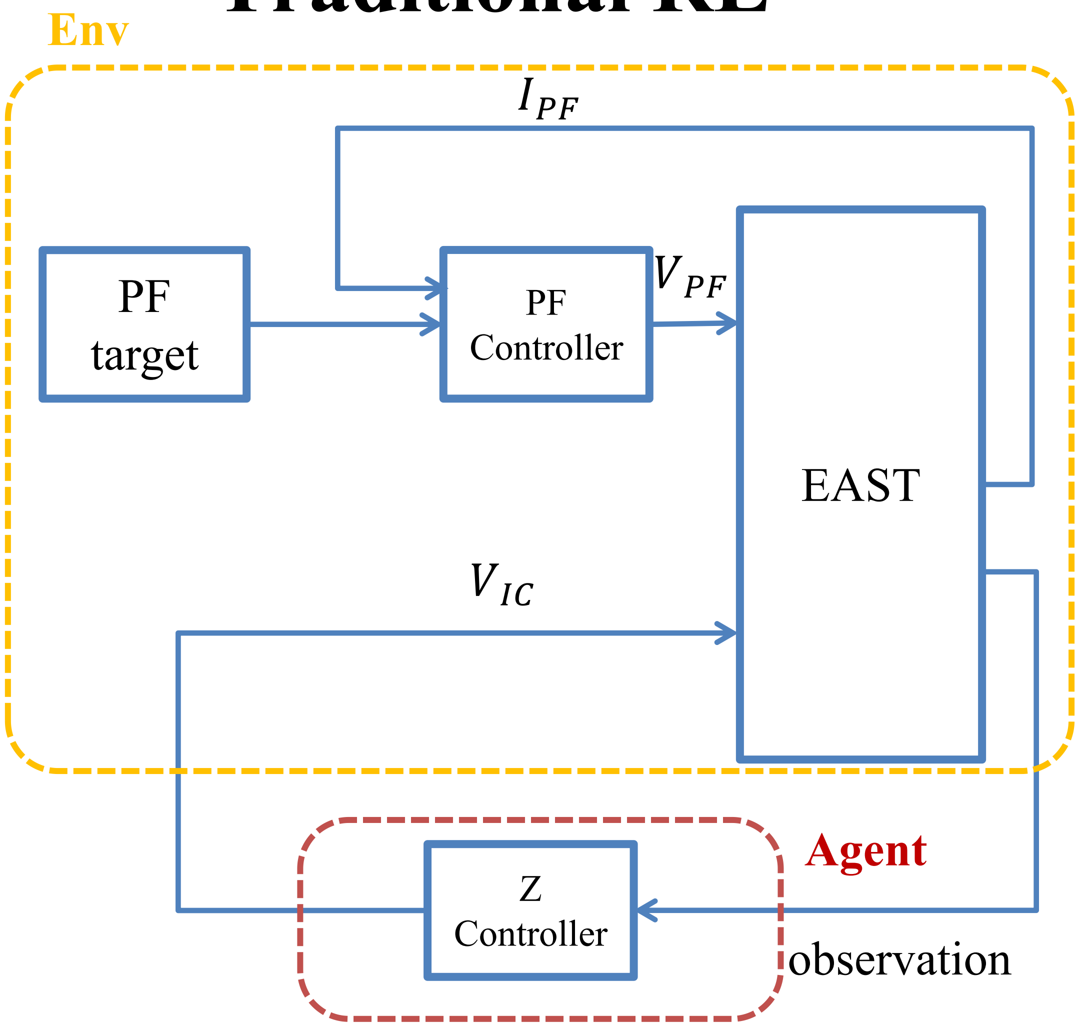
$$x(t+1) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

where x is the current on the coil and passive structure, u is the IC coil voltage, y is the plasma vertical displacement and IC coil current, and the control frequency is 10 kHZ

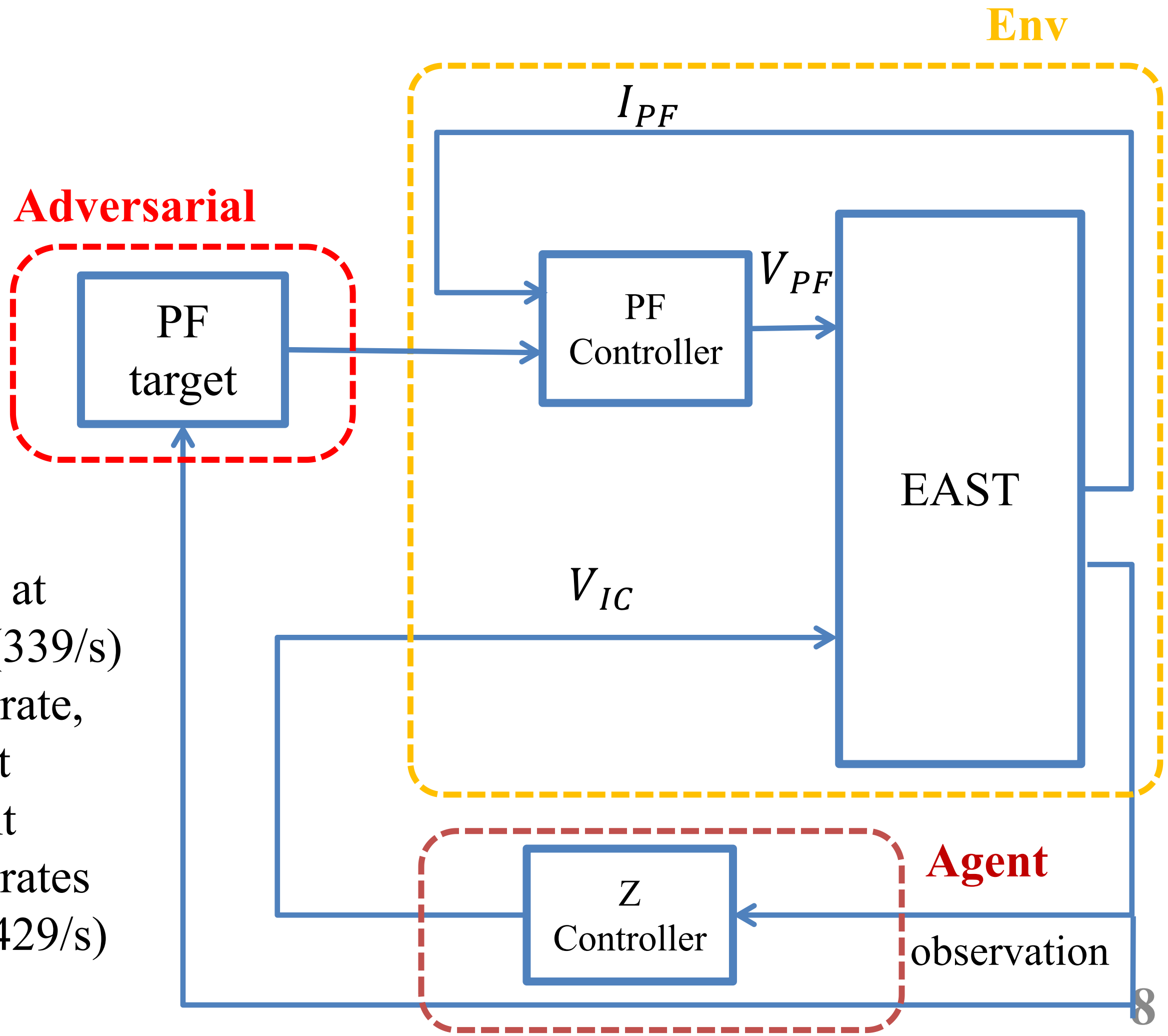
Environment Configuration

Traditional RL



Trained at higher (339/s) growth rate, tested at different growth rates (184/s-429/s)

RARL



Algorithm Configuration

ASIPP

- **Controller Configuration**
  - Control Frequency： 10kHZ
  - Training Algorithm: TD3
  - Fully connected neural network
  - Power Delays： 300 μs
  - $O_{1_t} = [z_{t-3}: z_t, I_{t-3}: I_t, V_{t-3}: V_{t-1}]$
  - $A_{1_t} = V_t \in [-1200V, 1200V]$
- **Adversary Configuration**
  - $O_{2_t} = I_t$
  - $A_{2_t} = \Delta_{PF\ target} \in [-400A, 400A]$
  - $\boldsymbol{R_{Adversary}(s, a) =- R_{Controller}(s, a)}$

- **Termination conditions**

  When $|I_{ic}| > I_{ic_{max}} = 9000A$， the current episode is terminated and a large negative reward is feedback to the agent, resetting the environment.

- **Maximum duration of each episode**

  In order to fully reflect the effect of shape-control coupling in the simulation, the maximum duration of t=0.5s is set.

Reward Function

- $R_1(s,a) \leftarrow +$
  - $-k_1 \frac{|z|}{|z_{max}|}$ Easy to train
  - $-k_2 \frac{|I_{ic}|}{|I_{ic_{max}}|}$ Avoid overcurrent
  - $-k_3 \frac{|V_{ic}|}{|V_{ic_{max}}|}$ Reducing the jitter of z

Use R_1 for pre-training

- $R_2(s,a) \leftarrow +$
  - $-k_2 \frac{|I_{ic}|}{|I_{ic_{max}}|}$ Avoid overcurrent
  - $-k_3 \frac{|V_{ic}|}{|V_{ic_{max}}|}$ Reducing the jitter of z

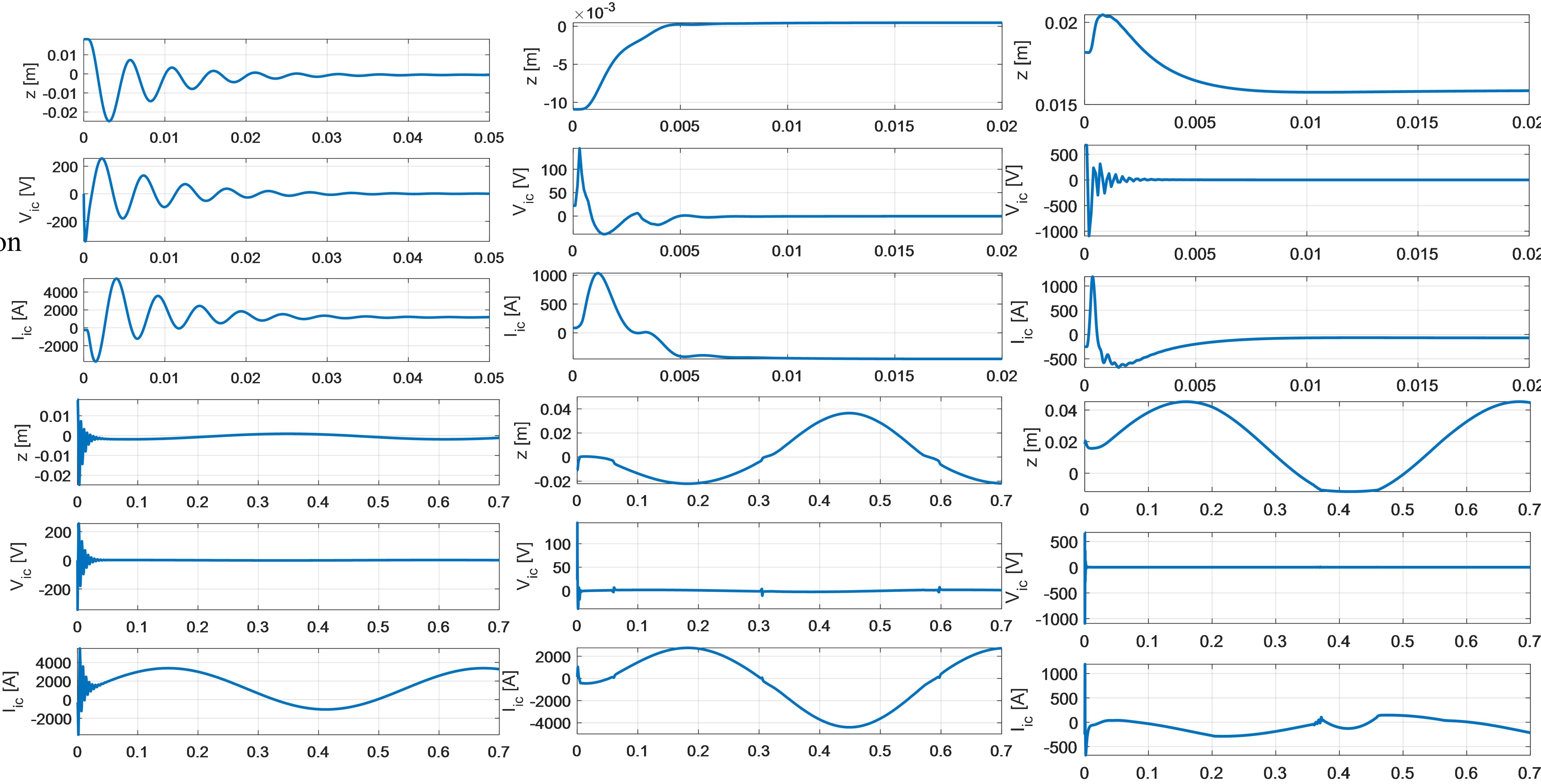
Initialize the model with pre-trained agent and adversary and then train the model with R_2

Outline

Simulation Test Results

Dynamic response without perturbation

Applying a sine wave perturbation of 800A to a PF coil



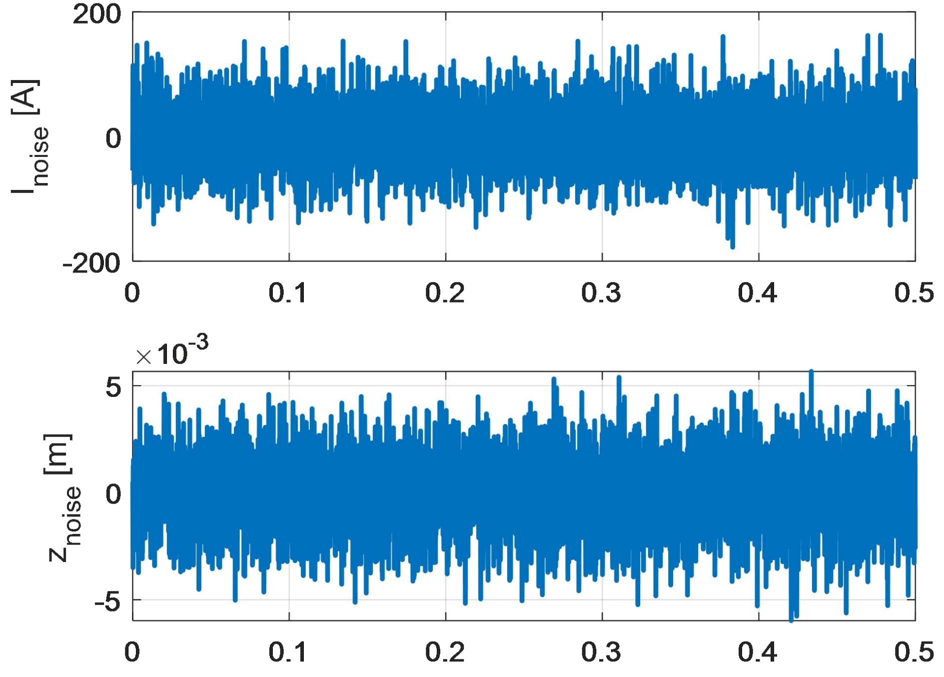
Traditional PID Controller

Traditional RL Controller

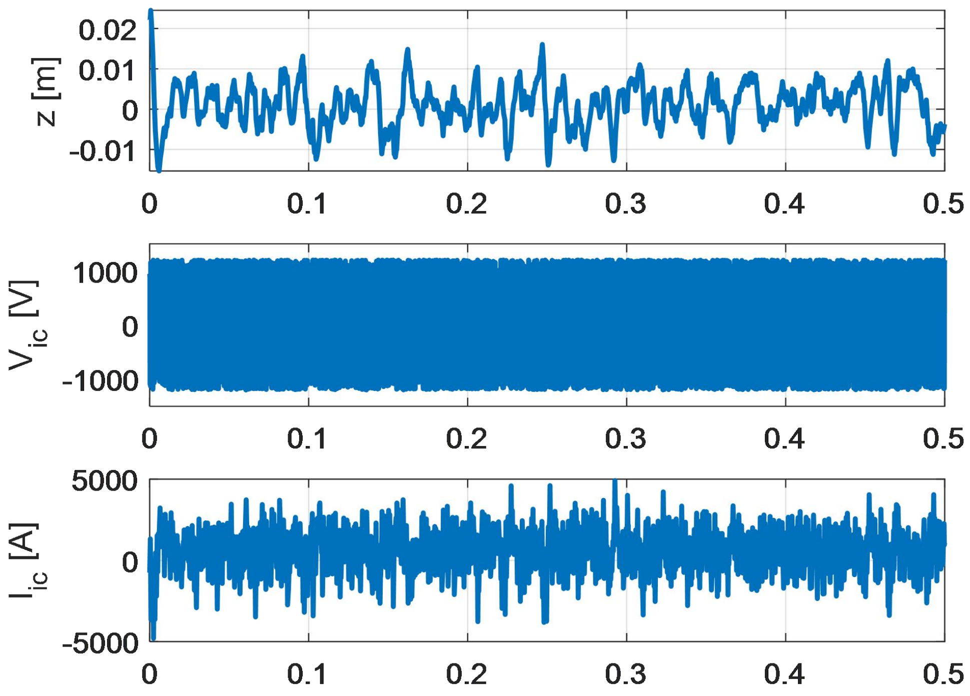
RARL Controller

Simulation Test Results

- Noise Resistance Test



Noise

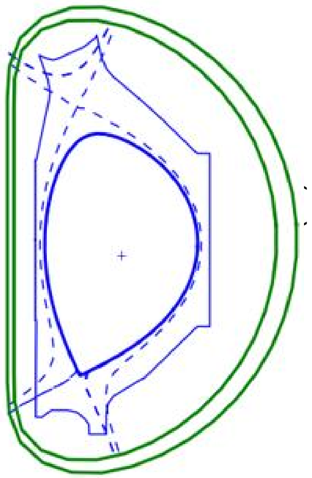


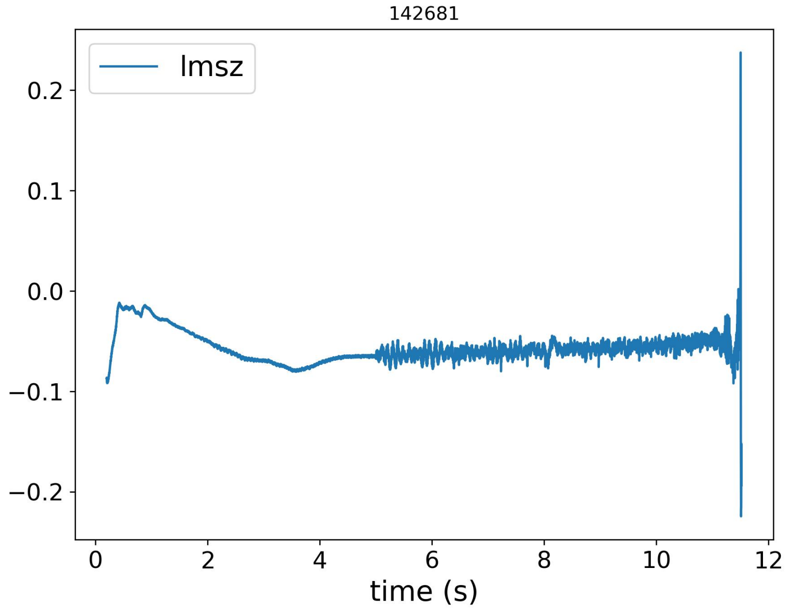
Test results with noise

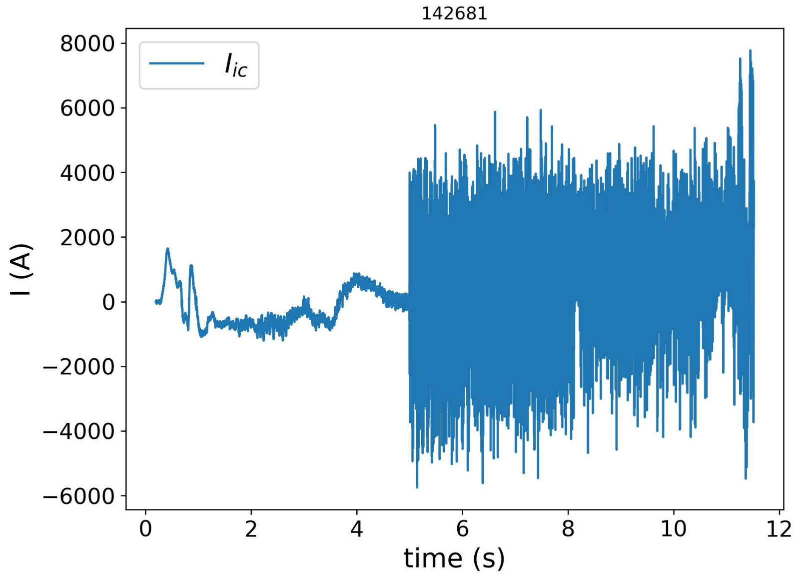
The controller is too sensitive to noise

Experiment Results

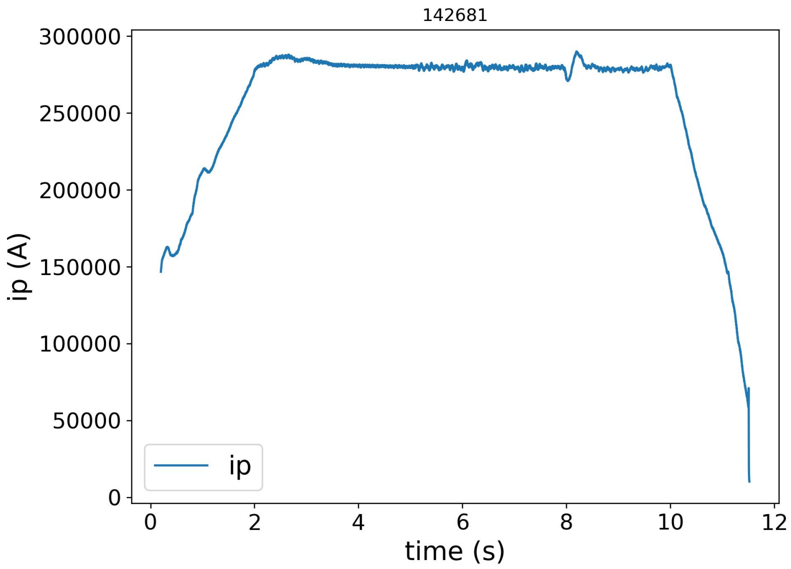
- **#142681 switch to RARL agent @5s, full discharge**
  - **ip：280kA**
  - **γ：200/s**
  - **Iic offset 3kA**
  - **close slow z filter**

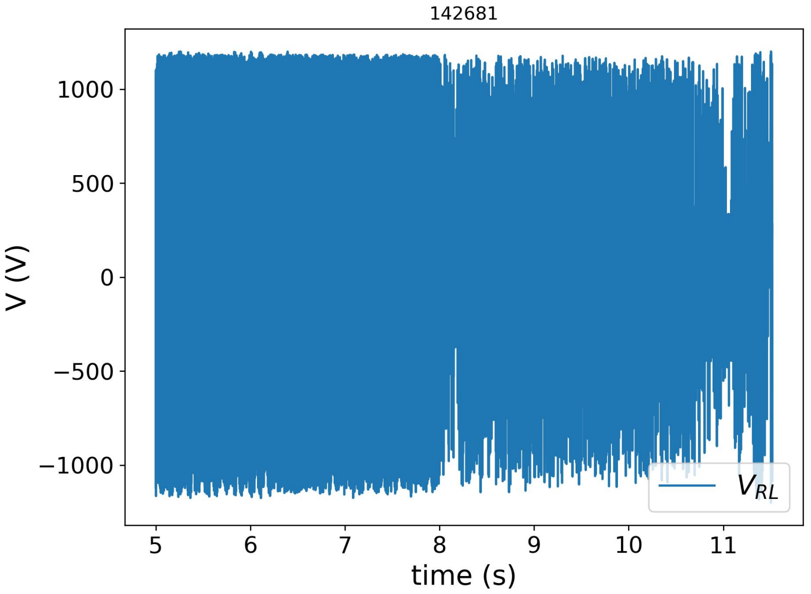






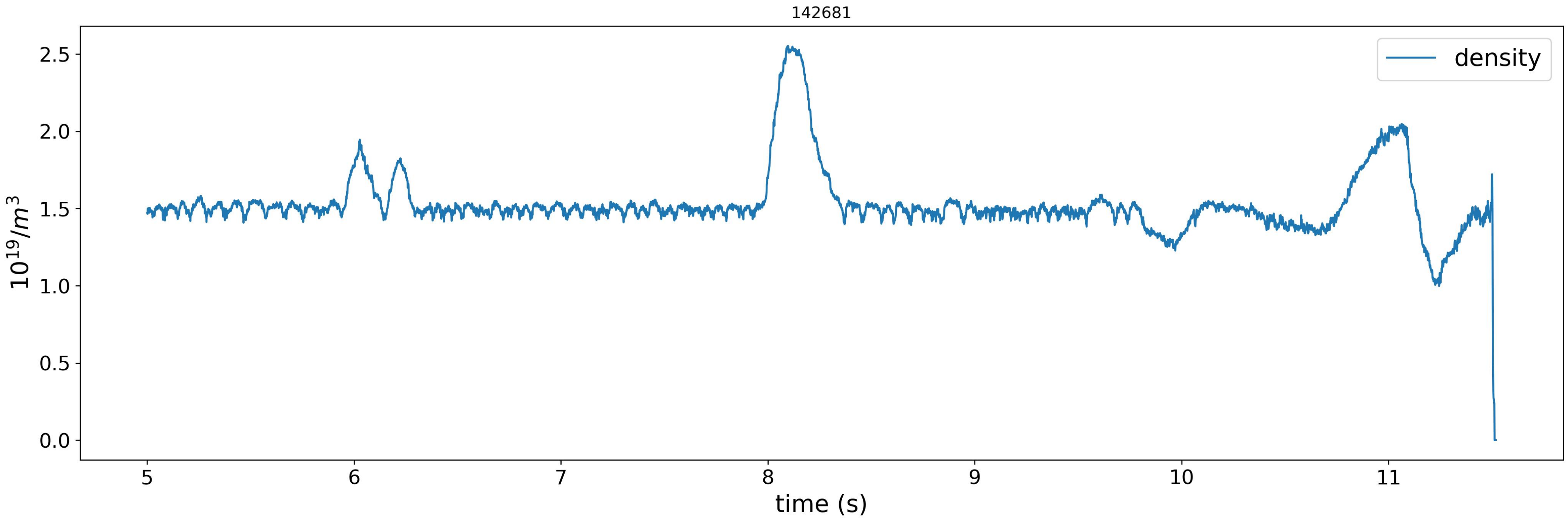
- **The performance of the controller is similar to that of the simulation with the addition of noise, which shows that the main problem at present is the poor resistance of the controller to noise**





Experiment Results

The density of #142681 suddenly increases dramatically at 8 seconds due to some foreign material falling into the running plasma. The RARL controller is able to control the plasma after this minor disruption, indicating that the model is resistant to perturbation to some degree.



Reasons and Methods

ASIPP

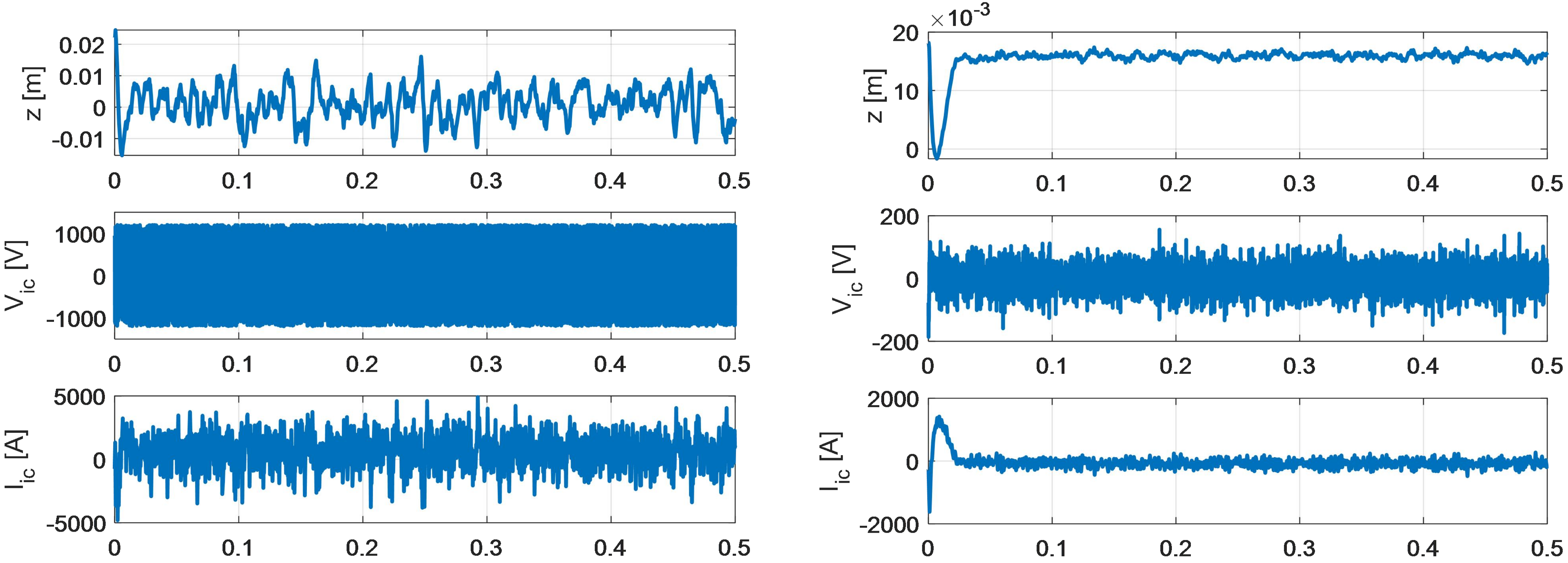
Reason for sensitivity to noise: The inputs to the controller include historical information, from which the controller extracts the critical information of the rate of change. Even a small noise can have a large effect on the rate of change, which can give the controller a danger signal. So the controller tries to stabilize z as quickly as possible, generating a large control command. Since there is no noise in the training, generating large control commands is often effective. However, in real experiments, this strategy is often bad due to the presence of noise.

Increase the model's resistance to noise:

1. Apply noise to the controller's observations during training, this noise can be white noise or noise applied by another adversary.

2. Do a feature extraction artificially . Replace the historical information in the observations with information that has been filtered by different feature time filters to reduce the effect of noise on the features extracted by the agent.

New Results



Previous results

New results

Outline

ASIPP

Future Plans

- Increase the model's resistance to noise
- Getting better results in the EAST experiment

Thanks !