Residual Aware RL-MPC

EXPLICIT RL POLICY ON MODEL RESIDUALS:
A SELF-CORRECTING CONTROL FRAMEWORK

ME 418/518 – DATA-BASED CONTROL SYSTEMS

ERİNÇ ADA CEYLAN 22101844
INSTRUCTOR: **ASSOC. PROF. YILDIRAY YILDIZ**

Problem Definition

The success of **model based control** is negatively affected by the **mismatch** between real world physics and system identification.

- System dynamics are generally highly **nonlinear** and **underactuated**
- Linear identification models are only **locally** valid
- MPC **fully trusts** its prediction model
- Fixed MPC parameters **fail** when the model becomes unreliable

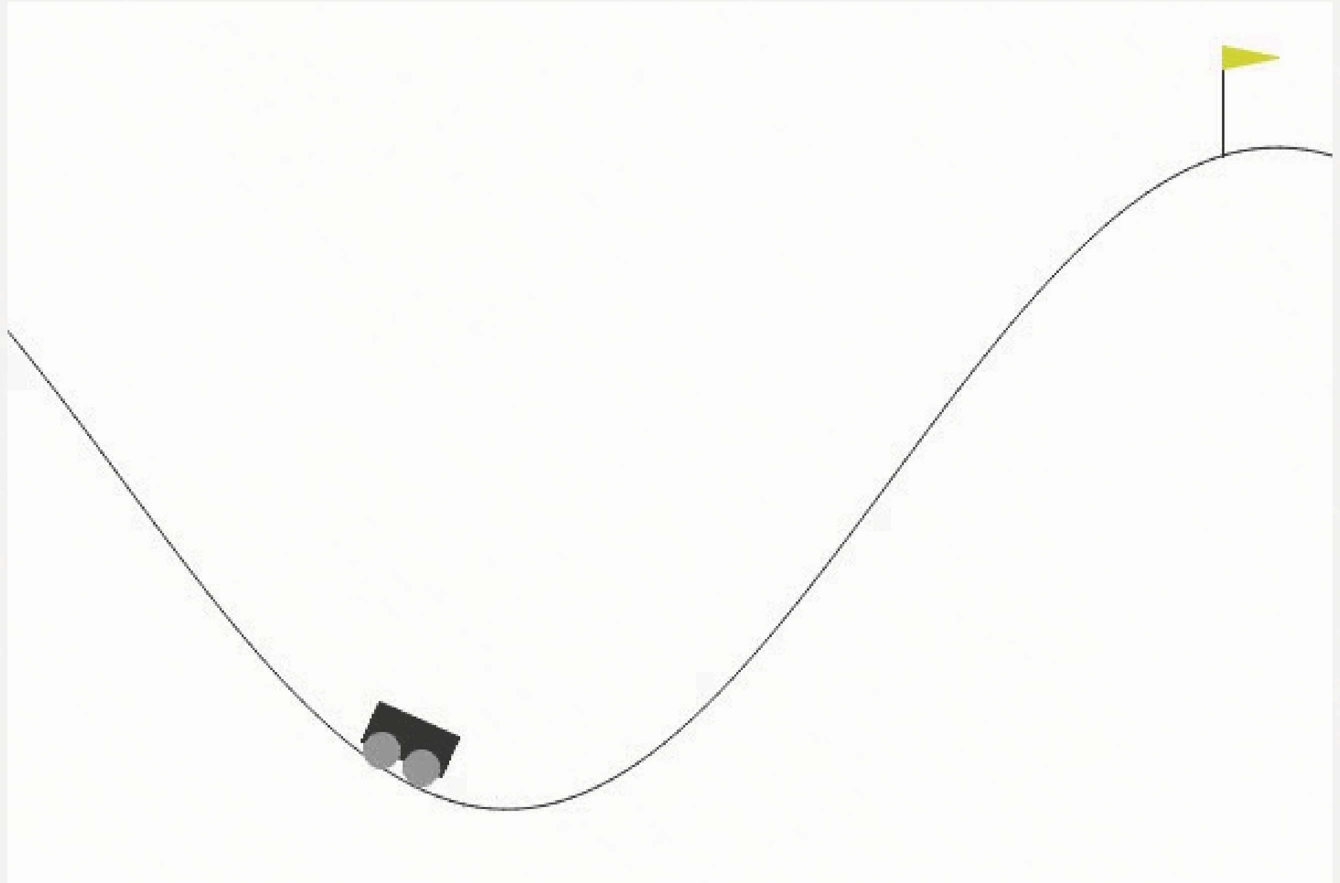

Figure 1: Baseline MPC MountainCar Model Simulaiton [1]

Existing Approaches & Novelty

Common Assumption in Existing Methods

- Prediction model is treated as **accurate**
- Model mismatch handled **beforehands**
- No explicit signal for model reliability
- Performance degradation when model becomes **invalid**

Infulential Work (TU Delft)- RL Weight Tuning in MPC [2]

- MPC weights (Q,R,N) tuned using RL
- **RL policy** depends on **state, time**, or **performance metrics**
- Model dynamics assumed to be reliable
- Focus on tracking performance

Novelty
Residual-Aware RL–MPC

- RL policy conditioned on **model residual**
- Residual = **real time model reliability indicator**
- MPC parameters **adapt** when the model becomes unreliable

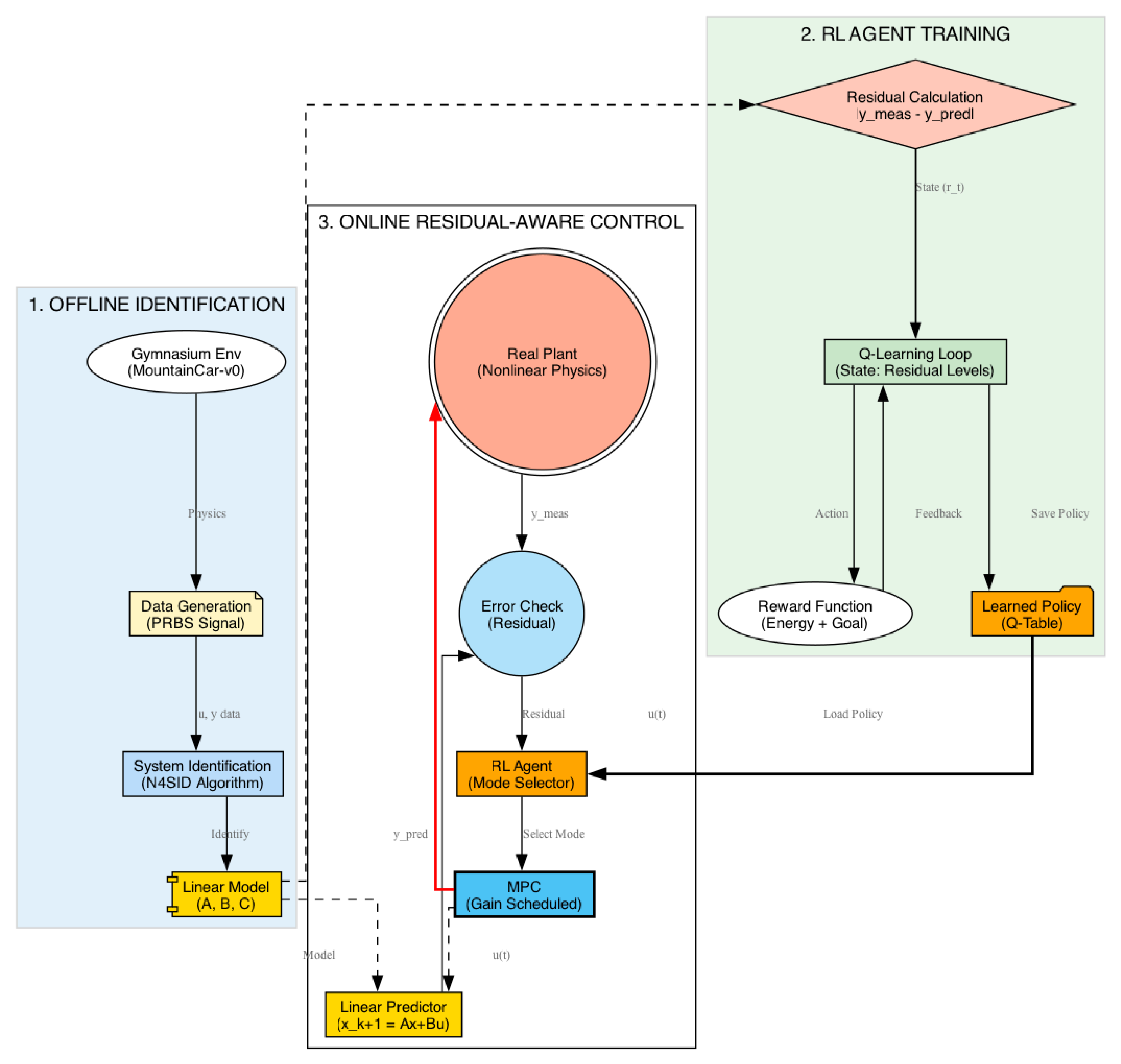

Figure 2: RL-MPC Pipeline

Overall Pipeline

Data Generation & Identification

Data Generation

- Covers the **full** gym position & velocity range
- **Nonlinear** MountainCar physics
- **Noise** added to measurements

Model Prediction (System ID) Quality

- N4SID provides the state-space model
- Best compromise between accuracy and MPC compatibility

Residual as an Information Signal

- **Residual magnitude differs** across models
- Captures **local model validity**
- Used as **state input** to RL agent

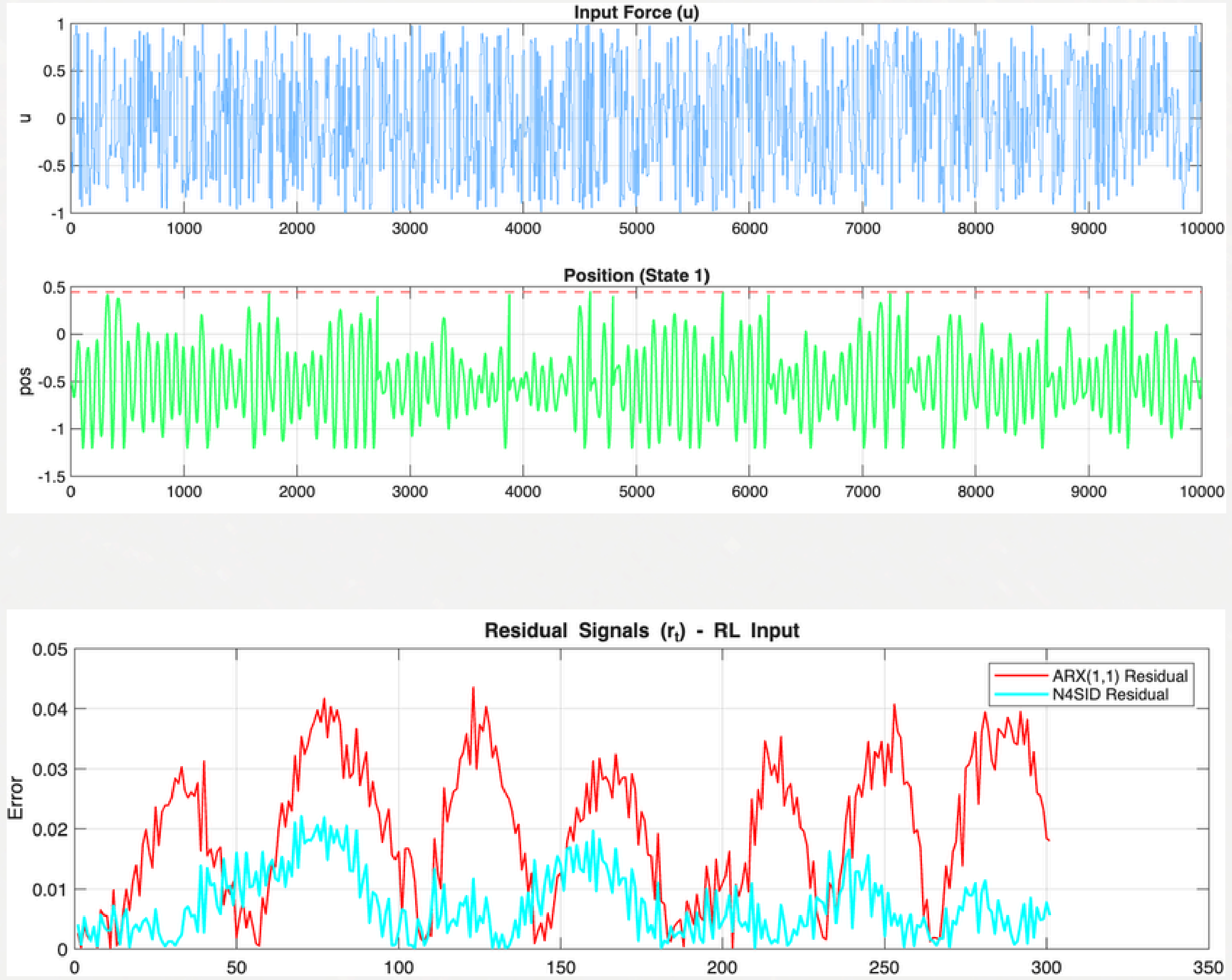

Figure 3-4 : Data Generation Process & Residuals

Baseline Linear MPC Performance

Fixed parameter linear MPC fails under strong model **mismatch**

- MPC is built on a **linear N4SID model**
- Model is only valid **locally**
- Nonlinear gravity term dominates outside local region
- Controller attempts **direct hill climbing**
- **No mechanism** to detect model failure

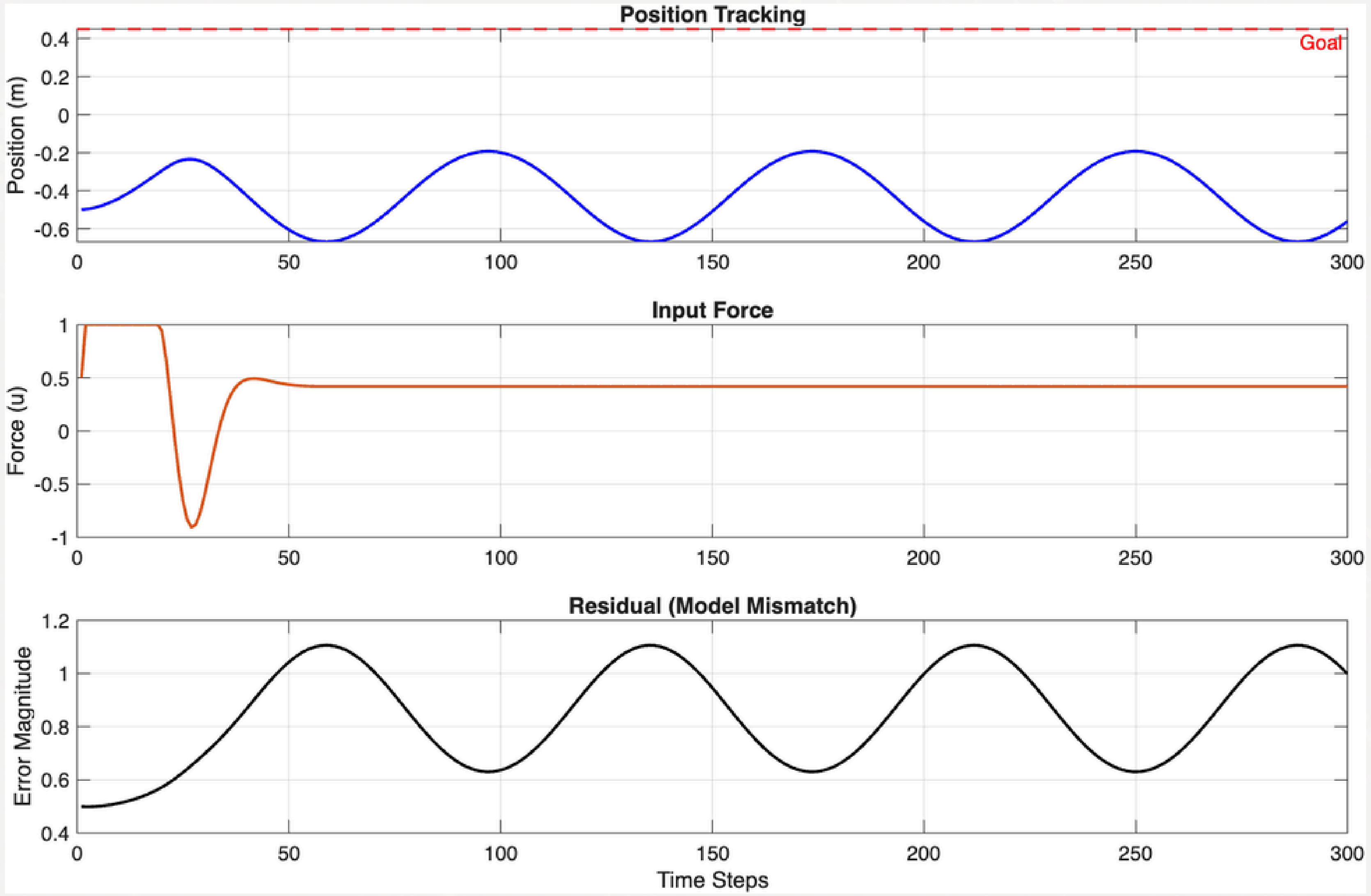

Figure 5: MPC Baseline Results

Residual-Aware RL-MPC Logic

RL tunes MPC parameters based on model reliability. RL does not control the system directly, observes only residual levels. Actions for the agent are thje selection of MPC mode:

- Recovery
- Tracking
- Stabilize

```
% Action 1: High Residual Recovery
Modes(1).Name = 'Recovery';
Modes(1).Q = 10; % Lowered trust to identification model
Modes(1).R = 0.01; % Ideal power for motor to overcome the recovery
Modes(1).Ref = -1.2; % Bias setpoint to the leftwards wall induce negative force

% Action 2: Low-Residual Tracking
Modes(2).Name = 'Tracking';
Modes(2).Q = 100; % Full trust to identifaction model
Modes(2).R = 0.01; % Ideal power for motor to track optimal climbing of the identifaction model
Modes(2).Ref = 0.45; % Standard Goal

% Action 3: Conservative Hold
Modes(3).Name = 'Stabilize';
Modes(3).Q = 1; % Minimized trust to identification model
Modes(3).R = 10; % High input penalty to compansate the controller. Lowered agressiveness.
Modes(3).Ref = 0.45; % Standard Goal
```

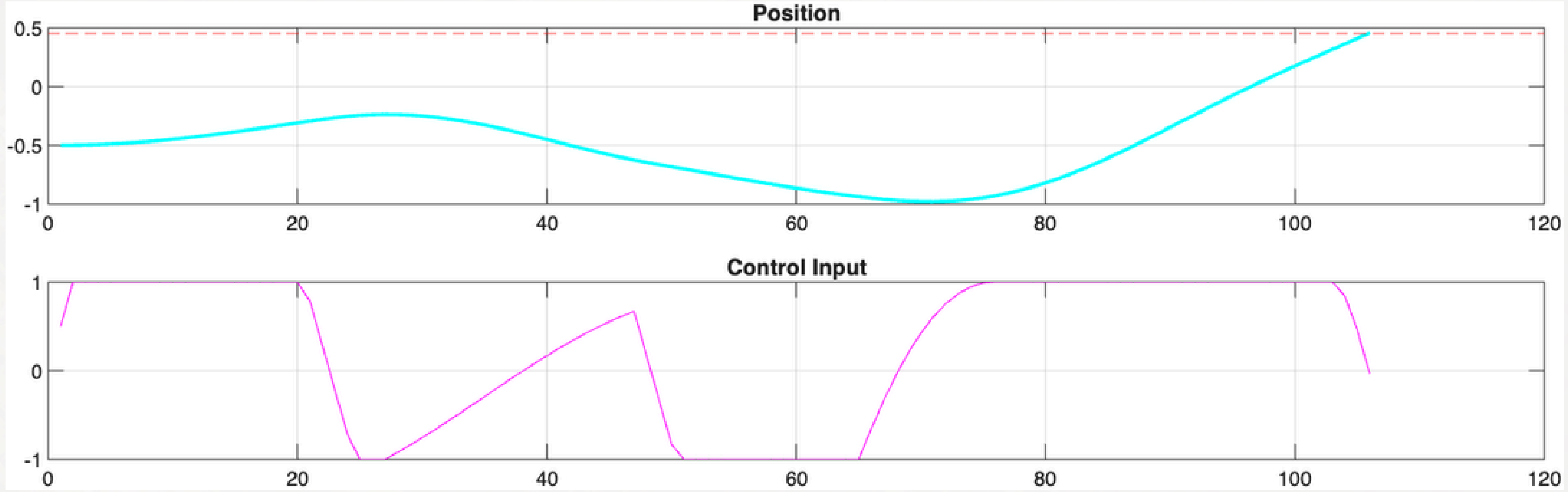

Figure 6: RL-MPC Parameter Tuning, Achieving Goal

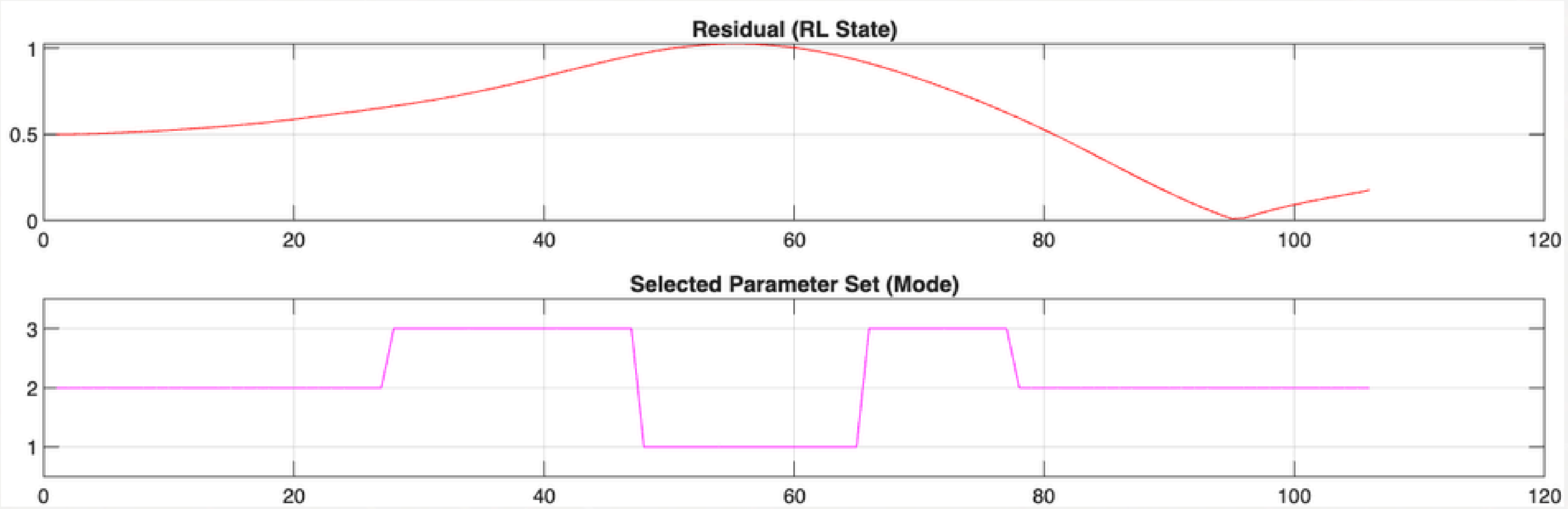

Figure 7: RL-MPC Parameter Tuning, Mode Switching

Residual-Aware RL-MPC Results & Conclusions

Key Results

- **Fixed parameter MPC** fails under complex dynamics
- **Prediction residual** acts as a real time model reliability signal
- **RL** learns to adapt **MPC parameters** when the model becomes **unreliable**

Novelty Mark

- **Existing RL-MPC** tunes weights based on **physics**
- **This work** conditions **RL** directly on **model accuracy**
- Delievering a **lightweit** and highly **transfarable** logic

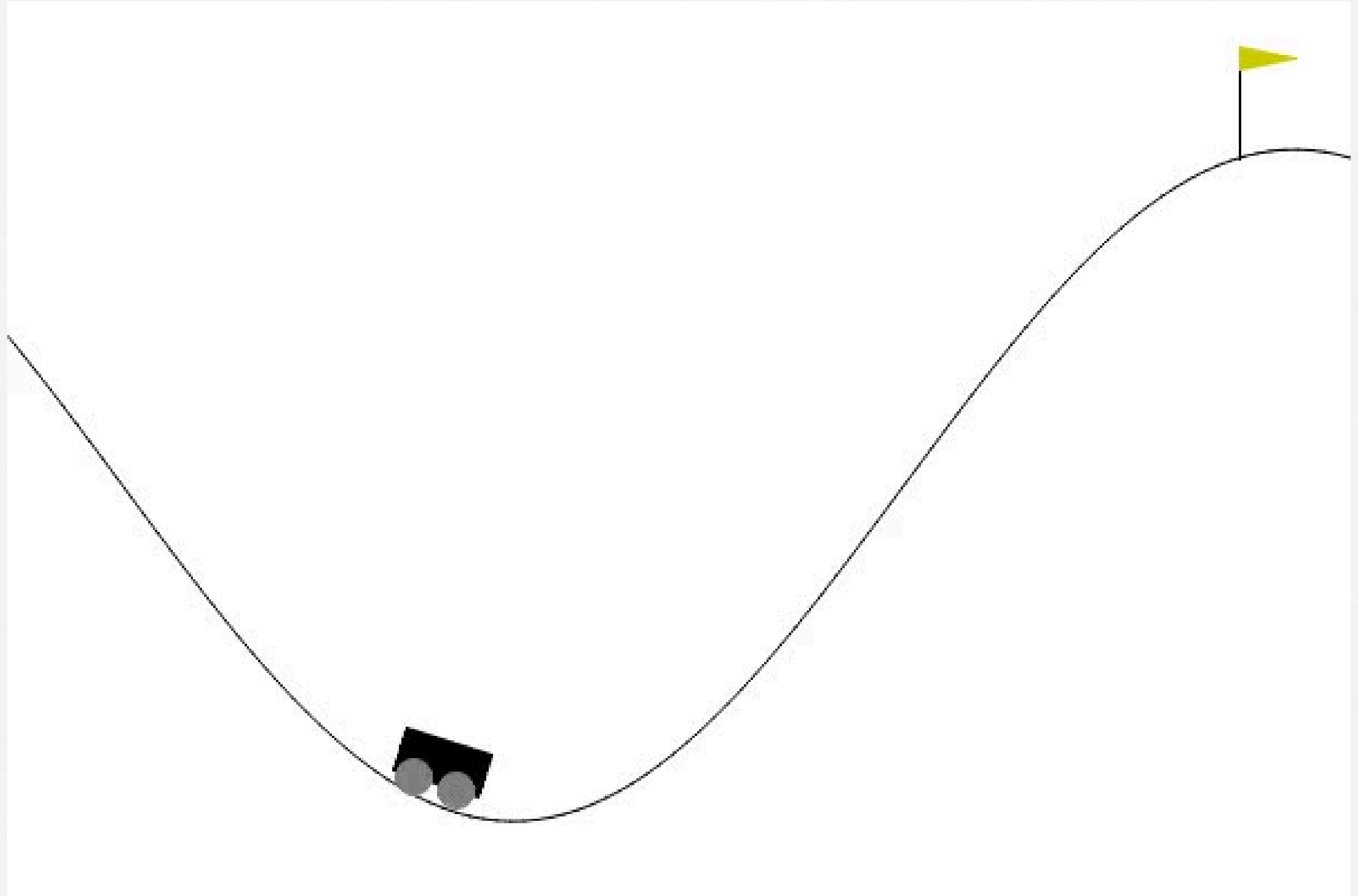

Figure 8: RL- MPC MountainCar Model Simulaiton [1]

References

[1] **MountainCar Continuous Environment**
Farama Foundation, Gymnasium: MountainCarContinuous-v0.
https://gymnasium.farama.org/environments/classic_control/mountain_car_continuous/
[2] **RL-based Weight Tuning for MPC (TU Delft)**
F. Airaldi, mpcrl: Reinforcement Learning with Model Predictive Control, TU Delft, 2024.
https://github.com/TUDelft-DataDrivenControl/mpcrl
[3] **Safe RL-driven Weights-Varying MPC**
B. Zarrouki, M. Spanakakis, and J. Betz, "A Safe Reinforcement Learning-driven Weights-Varying Model Predictive Control for Autonomous Vehicle Motion Control," arXiv preprint arXiv:2402.02624, 2024.

THANK YOU FOR LISTENING

Any Questions?

GITHUB github.com/erincada/Data-Based-Control-Systems
E-MAIL ada.ceylan@ug.bilkent.edu.tr