Adaptive Tuning of an Aircraft Pitch Controller with Reinforcement Learning

Adyasha Debahuti Mohanty—madyasha@stanford.edu Emma P. Schneider—epschnei@stanford.edu

Overview

Proportional-Integral-Derivative (PID) controller controls an aircraft pitch angle rate through deflection of the elevator. Traditional PID tuning does not produce optimal gains to unknown external disturbances. Adaptive tuning combines Reinforcement Learning (RL) and deep neural net to create a PID controller that reacts to varying uncertain observations. The RL agent takes the pitch angle error, derivative, and integral of error to output the three PID gains. Adaptive controllers perform better than existing fixed observation based PID controllers in steady state error, low overshoot, and fast rise time.

Model

- Uncouple aircraft pitch from roll dynamics and linearize model
  - Stability coefficients taken from a Boeing Aircraft
- Input– Elevator deflection angle, Output– Pitch Angle, Pitch Rate, Angle of Attack

State Space Model

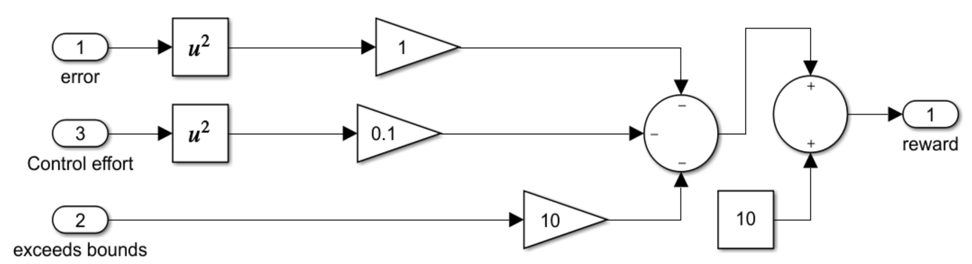
$$\begin{bmatrix}\dot{\alpha}\\ \dot{q}\\ \dot{\theta}\end{bmatrix} = \begin{bmatrix}-0.313 & 56.7 & 0\\ -0.0139 & -0.426 & 0\\ 0 & 56.7 & 0\end{bmatrix}\begin{bmatrix}\alpha\\ q\\ \theta\end{bmatrix} + \begin{bmatrix}0.232\\ 0.0203\\ 0\end{bmatrix}[\delta] \qquad y = [\,0 \;\; 0 \;\; 1\,]\begin{bmatrix}\alpha\\ q\\ \theta\end{bmatrix}$$

- Reward function-Penalize control effort (u) and pitch error , check if pitch angle lies outside allowed range

Control Effort

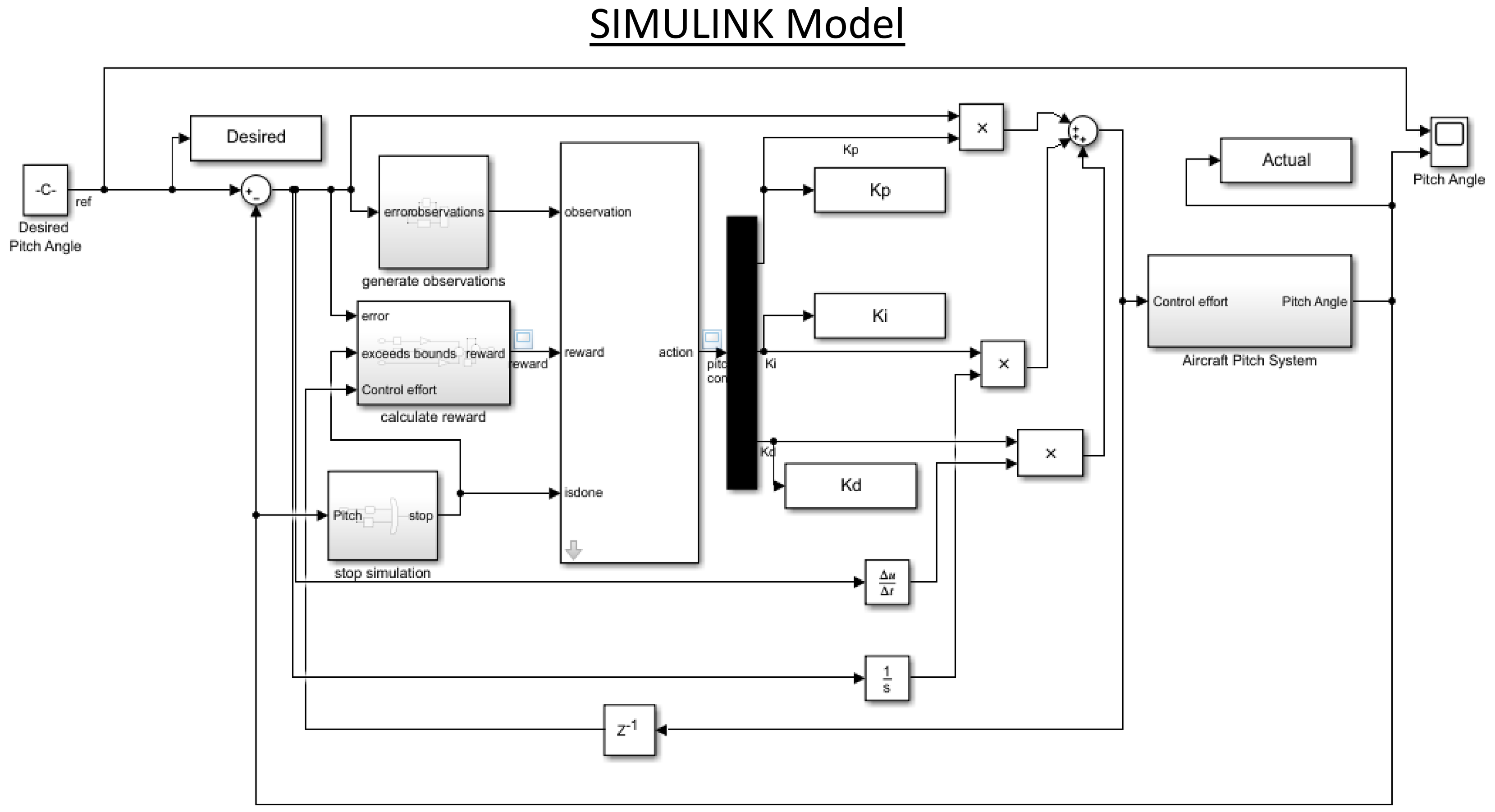
$$u = \underbrace{K_p e}_{\text{Proportional Term}} + \underbrace{K_i \int_0^t e\,dt}_{\text{Integral Term}} + \underbrace{K_d \frac{d}{dt} e}_{\text{Differential Term}}$$

Reward Function



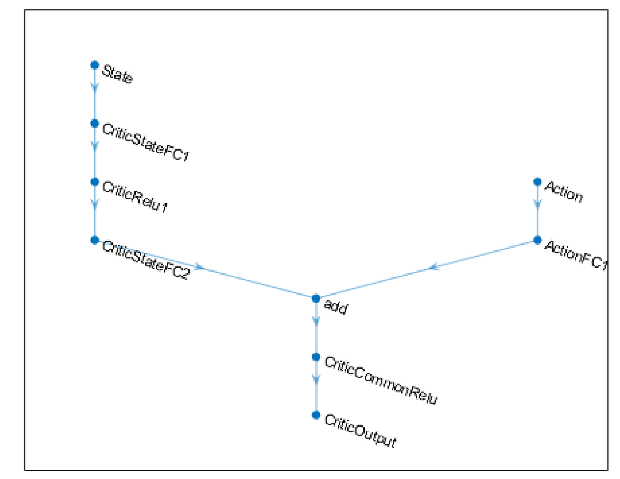
- Design net control effort from optimal gains
- Integrate input, output, observations, aircraft model, reward, action in SIMULINK:

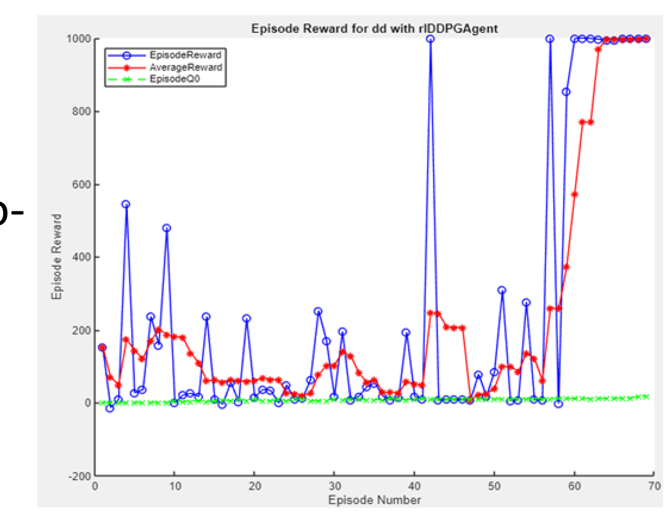
SIMULINK Model



Policy and Training

- Use Deep Deterministic Policy Gradient (DDPG) critic (state+ action) and actor framework
- Two fully connected and one ReLU activation layer, Hidden layer size-500
- Train until agent can successfully tune PID gains to follow desired input pitch value
- Hyperparameters– Gradient threshold: 1, Learn Rates: 1E-05 and 1E-04 (actor, critic), Mini-batch size: 64, Experience buffer: 1M
- Add variance of 0.3 initially during training to promote exploration. Value decays off towards end to promote exploration of optimal policy.

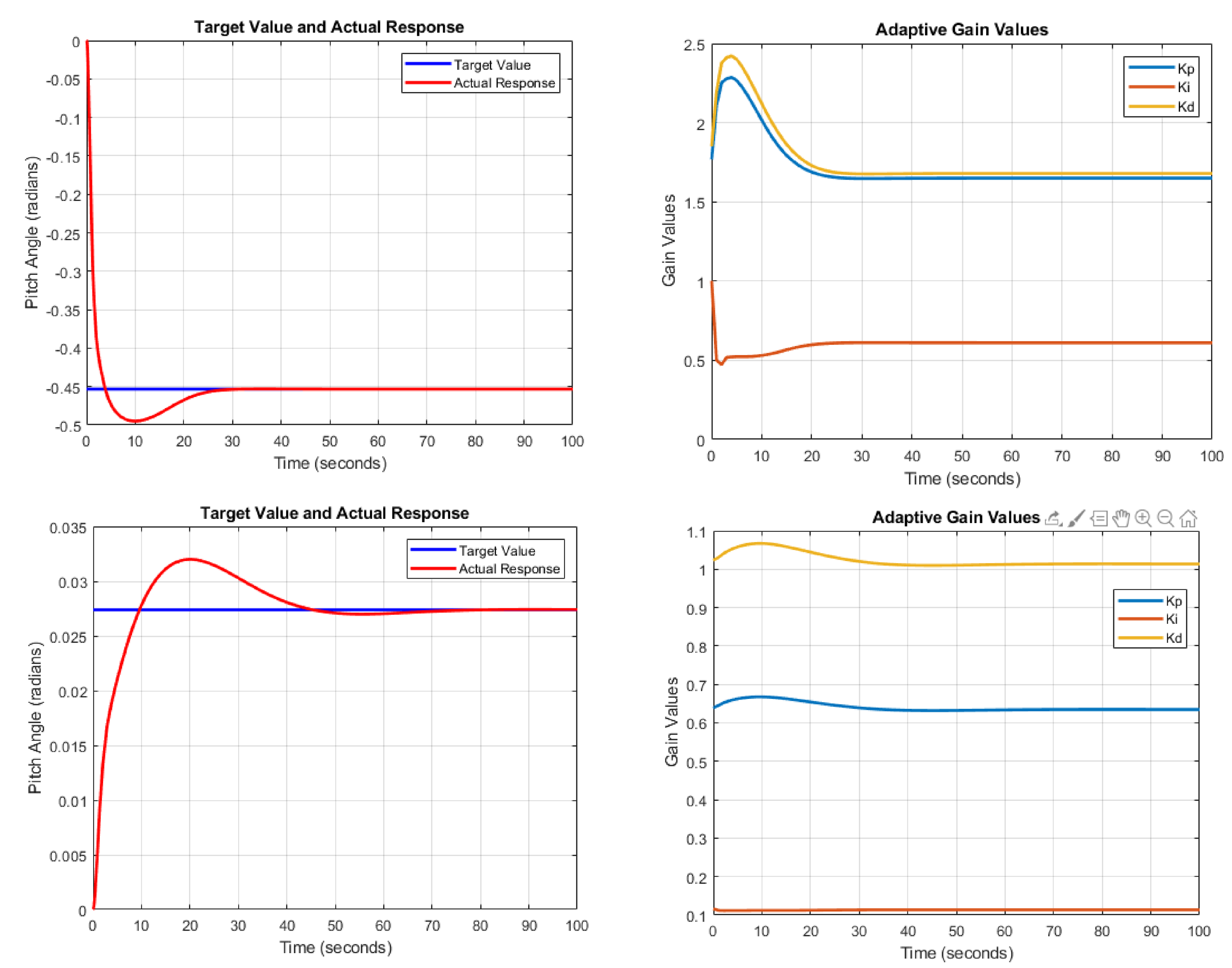




Results and Discussion

- Validate controller with random target pitch values between –1.5 to 1.5 rad
- Fast rise time, < 10% overshoot, 0 steady state error for all validation cases

Response curves/Gains for a negative and positive pitch input value



Results and Discussion

- Adaptive gains outperform MATLAB's PID Tuner controller in every aspect and produce least overshoot among all baseline controllers
- Rise time is comparable with other baseline controllers but settling time is higher
- Transient responses show performance due to adaptive gains is less susceptible to stochasticity and external disturbances

Parameter Name	Baseline 1	Baseline 2	MATLAB	RL-Adaptive
Rise Time (s)	0.20	2.90	6.38	2.25
Settling Time (s)	21.40	6.54	36.97	24.00
Overshoot (%)	30.20	14.00	15.53	9.50
Steady State Error	0.00	0.00	0.00	0.00

Bias/Variance Equivalence Tradeoff

- High bias present if hidden layer size < 100, high variance present if hidden layer size > 500 or no of layers in actor/critic > 4
- Very high or very low learning rates led to greater overshoot, settling time and rise time. Can cause training to not converge.
- Variation in Mini Batch Size had least effect on transient responses

Future Work

- Reduce settling and rise time
  - By tuning reward function
  - By changing initial weights and bias in the neural net
  - By adding a hyperbolic tangent layer to limit the gains
- Train with Proximal Policy Optimization (PPO) and other continuous-action agents
- Develop non-linear model with coupled roll-pitch dynamics
- Implement on physical system (i.e. RC aircraft)

References

A. K. Singh and R. Dahiya, "Design and Modeling of Controllers for Aircraft Pitch Contol Movement," *International Journal Of Engineering And Computer Science*, vol. 5, no. 11, pp. 1–14, Dec. 2016.

Mathworks, "Train Reinforcement Learning Agents," *Mathworks*, 2019. [Online]. Available: https://www.mathworks.com/help/reinforcement-learning/ug/train-reinforcement-learning-agents.html. [Accessed: 11-Dec-2019].

T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "CONTINUOUS CONTROL WITH DEEP REINFORCEMENT LEARNING," *International Conference on Learning Representations*, pp. 1–14, 2016.

Z. M. Motea, H. Wahid, J. Zahid, S. H. Lwin, and A. M. Hassan, "A Comparative Analysis of Intelligent and PID Controllers for an Aircraft Pitch Control System," *Applications of Modelling and Simulation*, vol. 2, no. 17-25, pp. 1–9, 2018.