Adaptive Optimization for System Performance: Parameterized Differential Dynamic Programming

Alex Oshin*#, Matthew D. Houghton*, Michael J. Acheson*, Irene M. Gregory* and Evangelos Theodorou#

*NASA Langley Research Center**

Georgia Institute of Technology#

AIAA SciTech Forum and Exposition

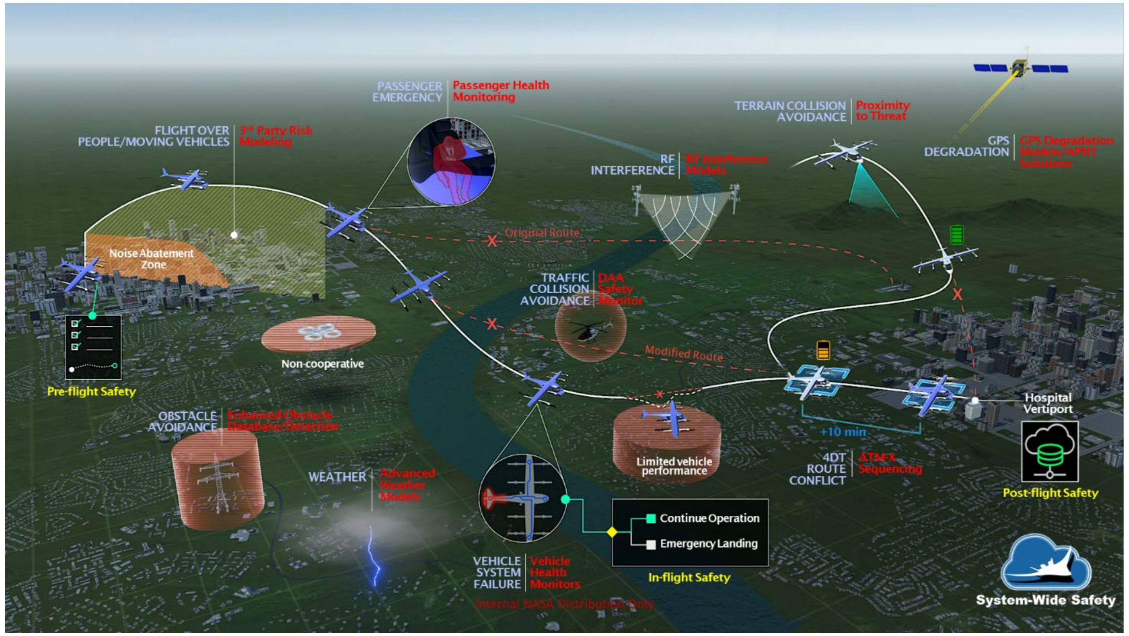
23-27 January 2023

National Harbor, MD

Motivation for Adaptive Optimization for System Performance

Emerging aerospace sectors – missions and vehicles

- Autonomous cargo delivery
- Urban Air Mobility (UAM)
- Complexity of the environment
- Unconventional configurations with multi-modal dynamics - rotor-borne vertical takeoff/landing, fixed-wing cruise, transition phase between the two
- Highly nonlinear flight dynamics
- Autonomous flight for scalability

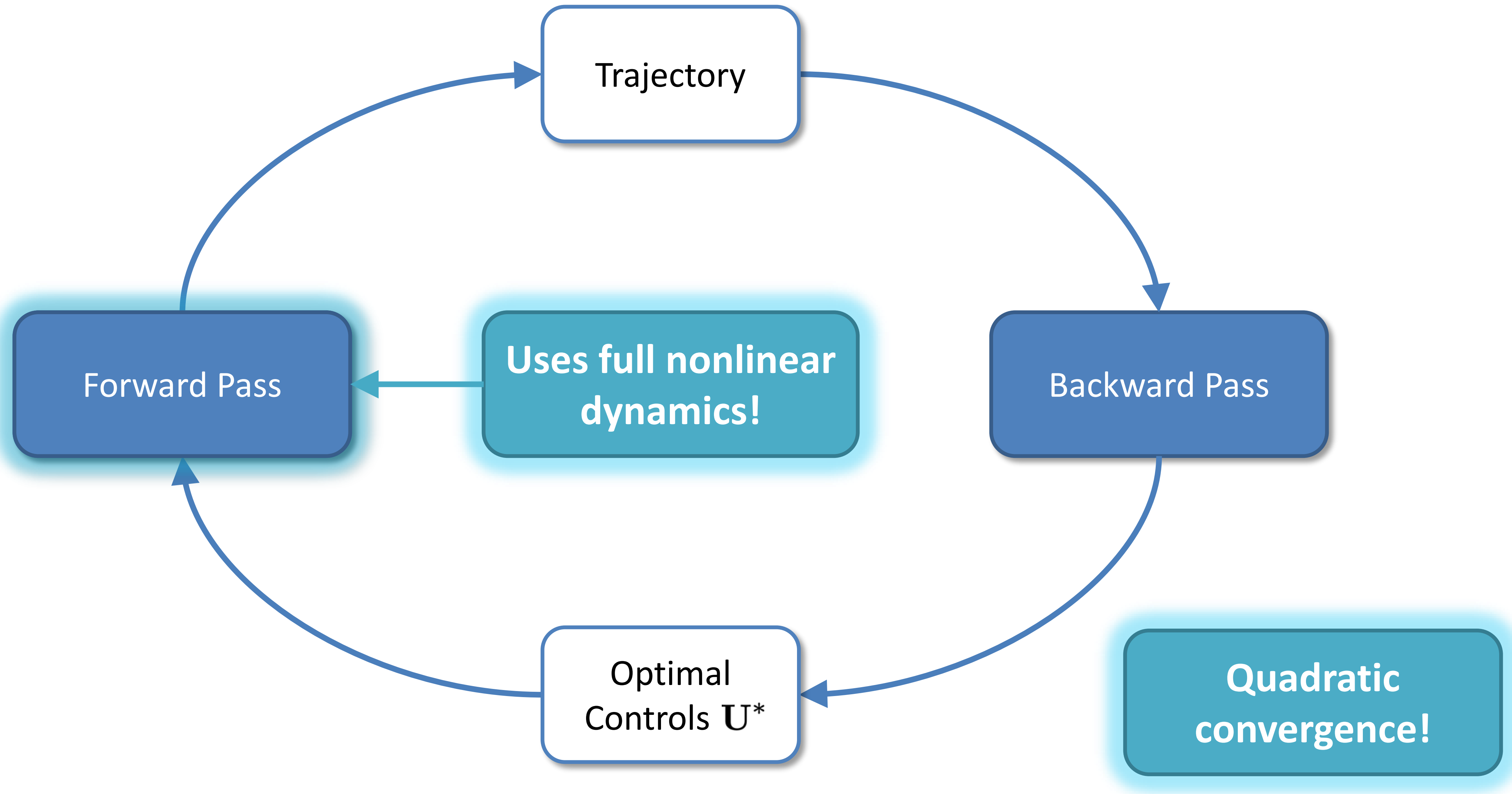


Planner challenges:

- Principled solutions/guarantees
- Accurate trajectory planning & replanning
- Epistemic uncertainty in model
- Multiple operational modes and flight regimes
- Transferability to different vehicles

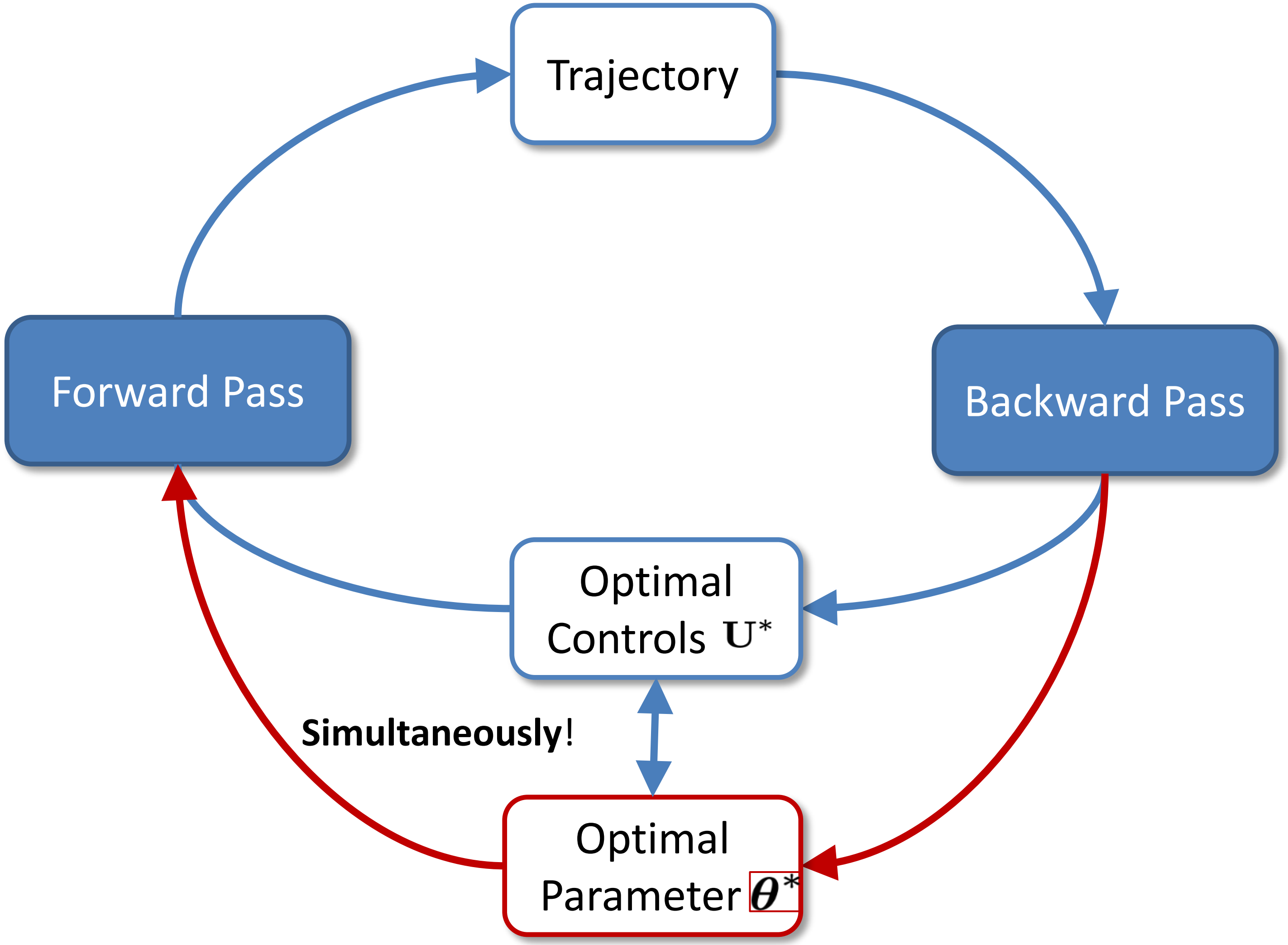
Irene.M.Gregory@nasa.gov

Differential Dynamic Programming (DDP)



Parameterized Differential Dynamic Programming (PDDP)*

- Second-order algorithm derived by extending classical optimal control (DDP)
- **Convergence guarantees** independent of initialization
- **Co-optimizes** for controls and parameters simultaneously
- **Generalizes** to multiple tasks, including adaptive MPC and switching time optimization
- Enables time-optimal trajectory planning for multimodal systems, including **UAM vehicles**



* Oshin, A., Houghton, M., Acheson, M., Gregory, I., and Theodorou, E., "Parameterized Differential Dynamic Programming," *Proceedings of Robotics: Science and Systems*, New York City, NY, USA, 2022. https://doi.org/10.15607/RSS.2022.XVIII.046.

PDDP Applications

Adaptive Model Predictive Control

Moving Horizon Estimation

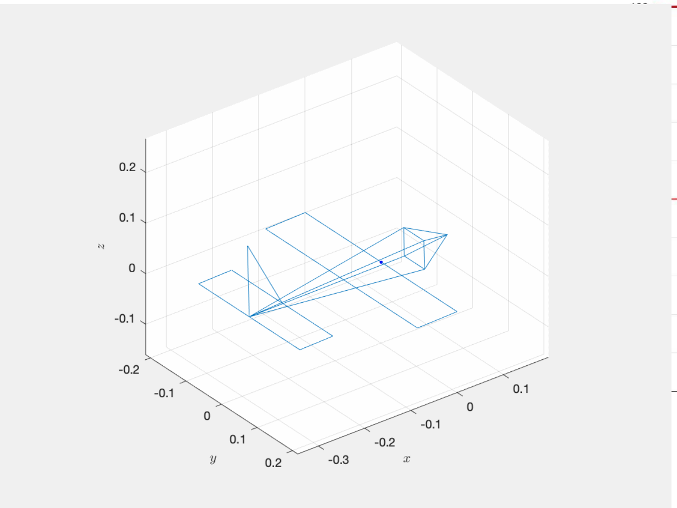
Maximize likelihood of observed states

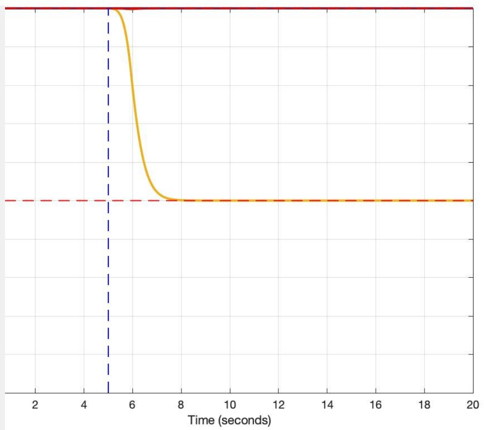
Model Predictive Control

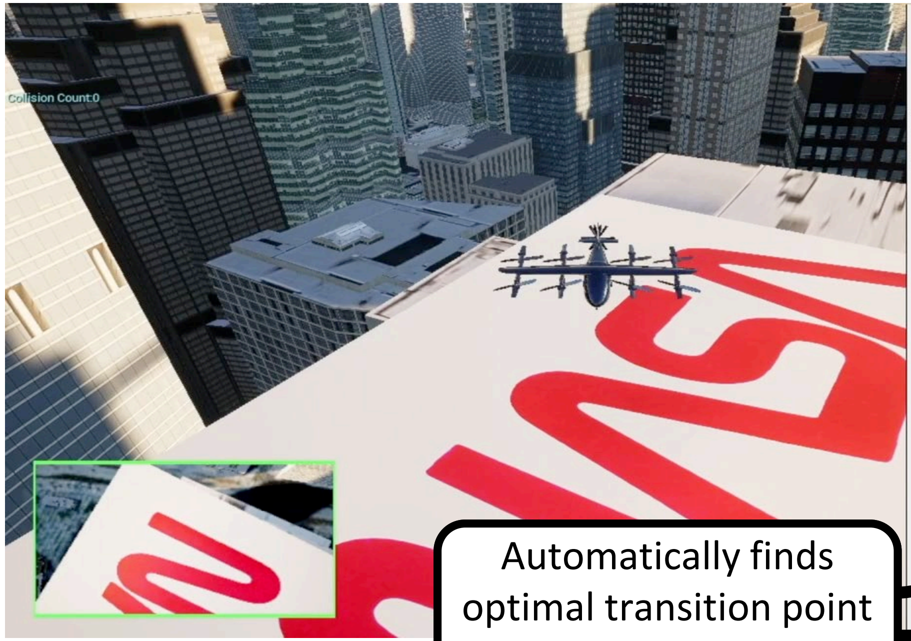
Plan future trajectory

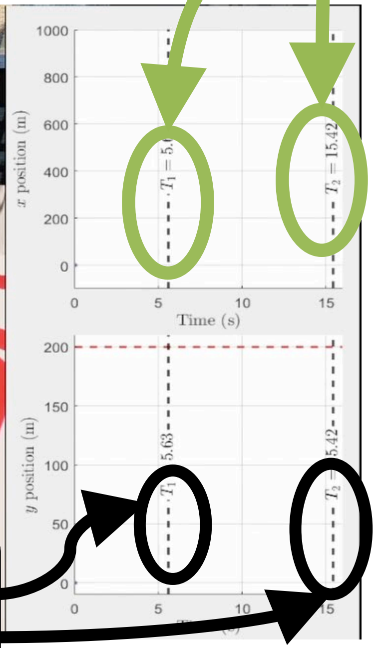
Switching Time Optimization

Avoids manual tuning of terminal times!









Automatically finds optimal transition point between modes!

PDDP Transition Optimization

- Long-term planning requires the L+C vehicle to change operating modes from hover to forward flight. Classical trajectory planning methods struggle with determining how to transition between modes.
- PDDP L+C experiment involves vertical takeoff into cruise transition with multiple target states
- Switching Time Optimization selects the optimal transition times between targets and flight regimes (without direct input from researchers)

Multimodal System

$$\dot{\mathbf{x}}(t) = \mathbf{f}^{(i)}(\mathbf{x}(t), \mathbf{u}(t))$$

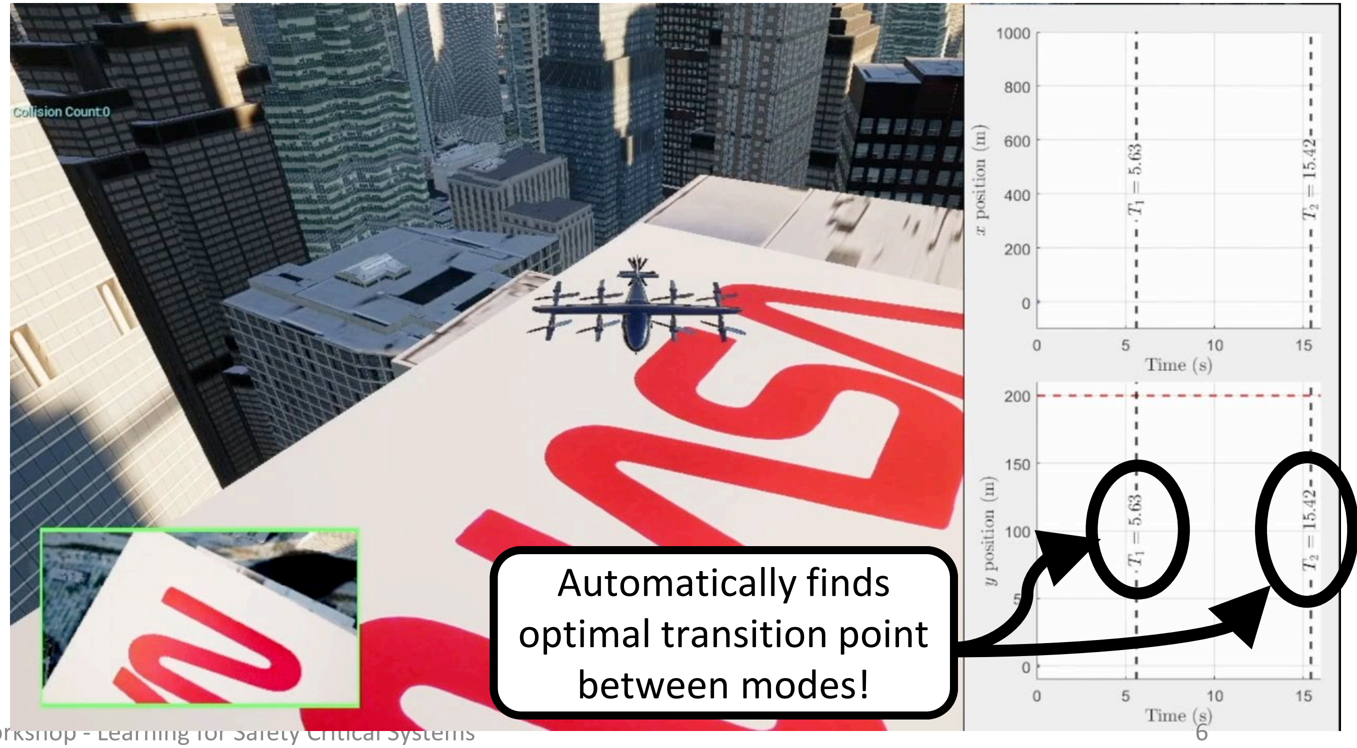
$$i = 1, 2, \ldots, N$$

Switching Time System

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \theta_i \mathbf{f}^{(i)}(\mathbf{x}_t, \mathbf{u}_t)\Delta t$$

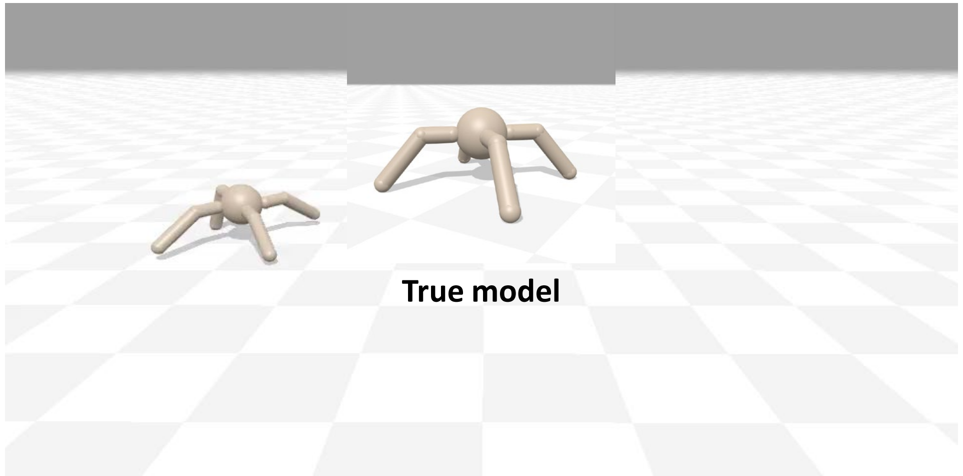
$$\sum_{t=T_{i-1}+1}^{T_i} \theta_i \mathcal{L}^{(i)}(\mathbf{x}_t, \mathbf{u}_t) + \phi^{(i)}(\mathbf{x}_{T_i+1})$$

$$i = 1, 2, \ldots, N$$



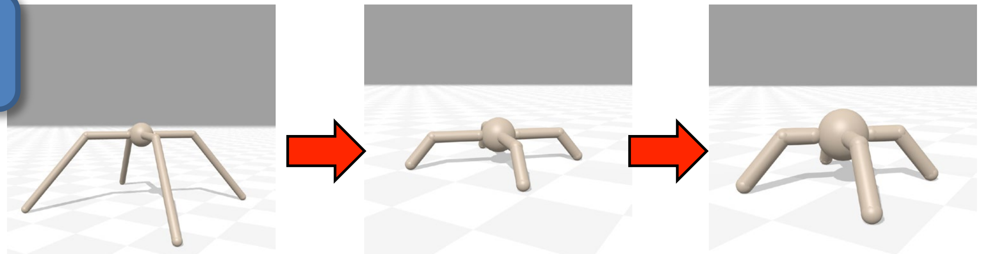
PDDP: Adaptive MPC

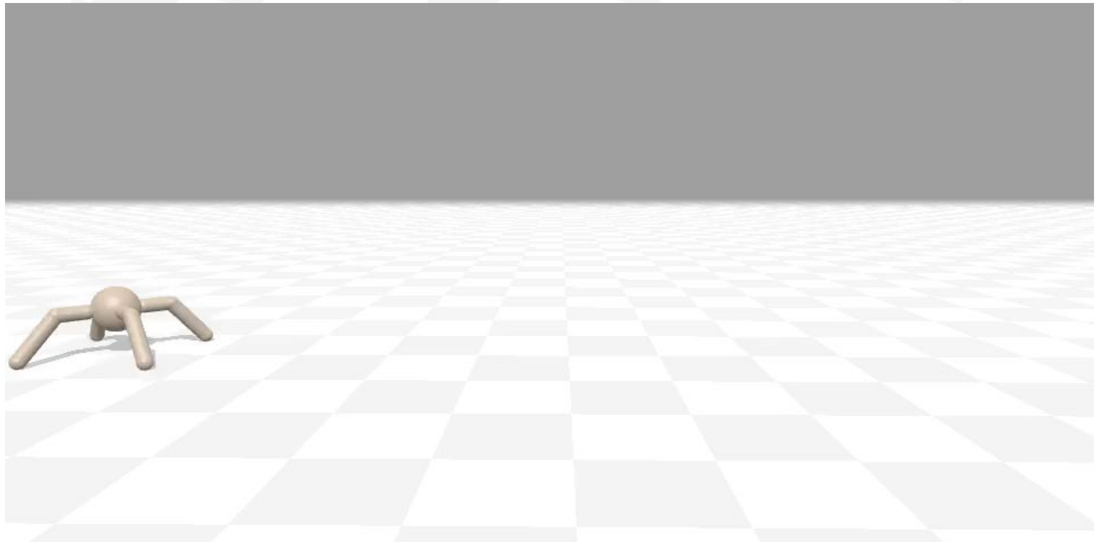
DDP planning on model with incorrect parameters



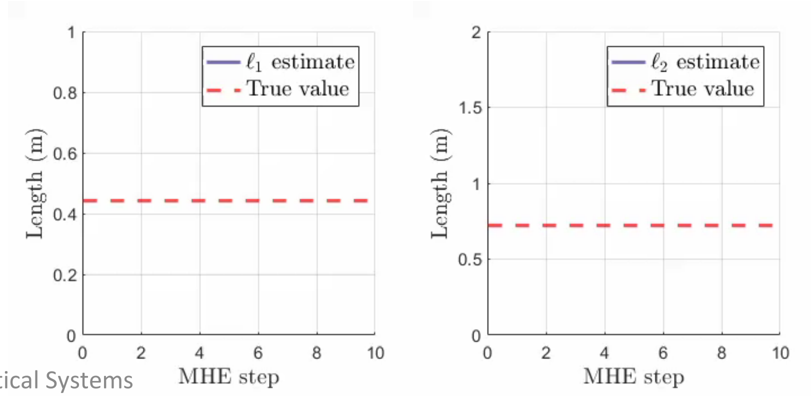
Executing plan on true model: Failure

Ant





PDDP with adaptive control: Success



Parameterized Differential Dynamic Programming (PDDP)

PDDP is a trajectory optimization algorithm that builds upon DDP

- Enables the **co-optimization** of a **trajectory** and time invariant **parameters** in the same process.
- Parameters can be extremely diverse and goal specific
- Experiments tested PDDP's ability to successfully **estimate vehicle dynamic parameters** while implementing **optimal trajectories**, resulting in Adaptive Model Predictive Control

Switching Time Optimization

- Calculation of **optimal transition time** between flight regimes (Difficult for highly nonlinear vehicles like L+C)
- **Decreases tuning** work for engineers when planning for common maneuvers that transition between flight regimes (Vertical takeoff into fixed-wing cruise)
- Allows for the input and optimization of multiple target states for long-term planning and replanning

Fault Detection

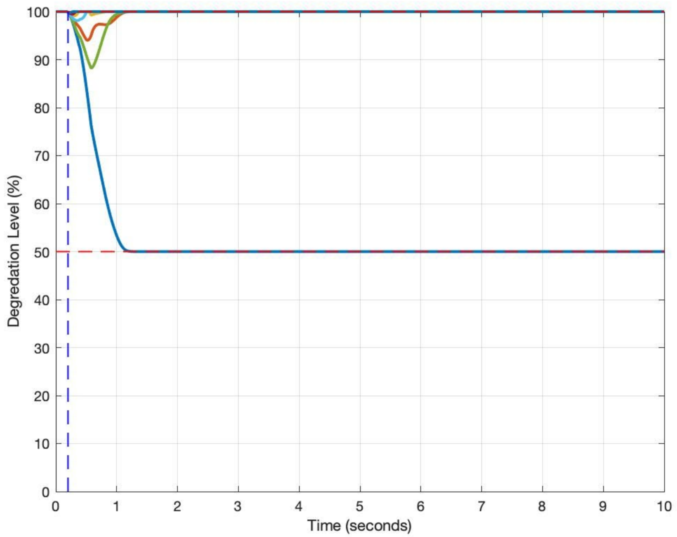
- Online estimation of vehicle **dynamic** parameters
- **Online estimation** of **degradation** level for effectors + rotors
- **Replan trajectory** based on new estimation of vehicle parameters
- Deviations in estimation from norms can alert system ID of vehicle to run further diagnostics of vehicle health

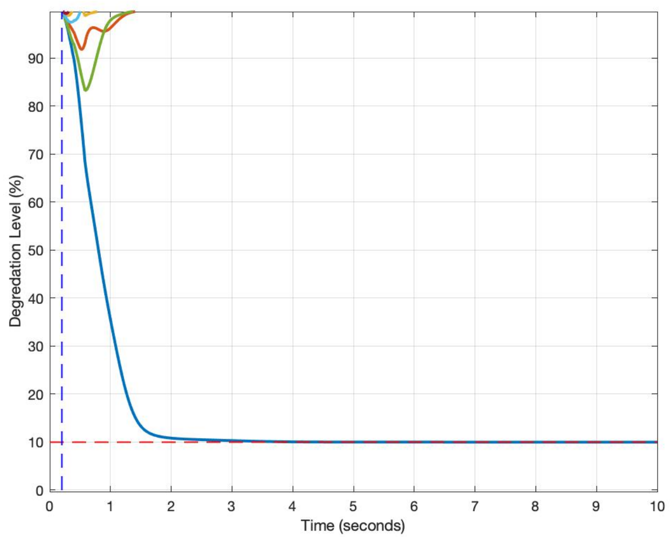
Fault Detection: Rotor Failure

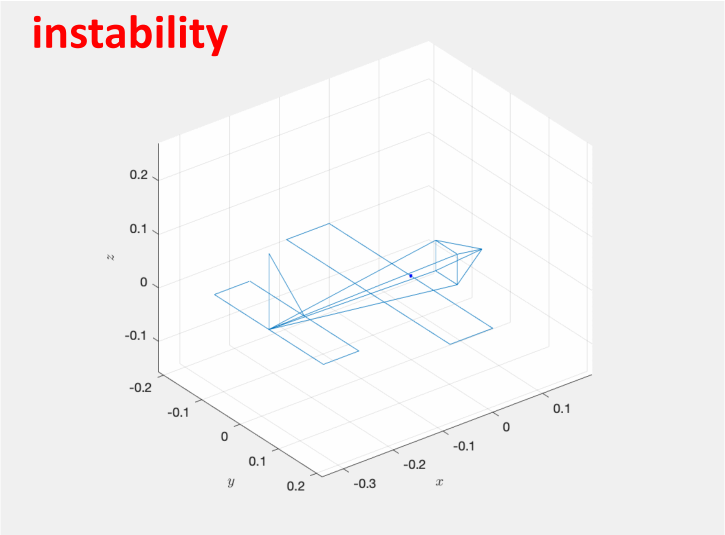
PDDP extends to Fault Detection of vehicle states (rotors and effectors)

<u>Experiment 1: Vertical Takeoff</u>

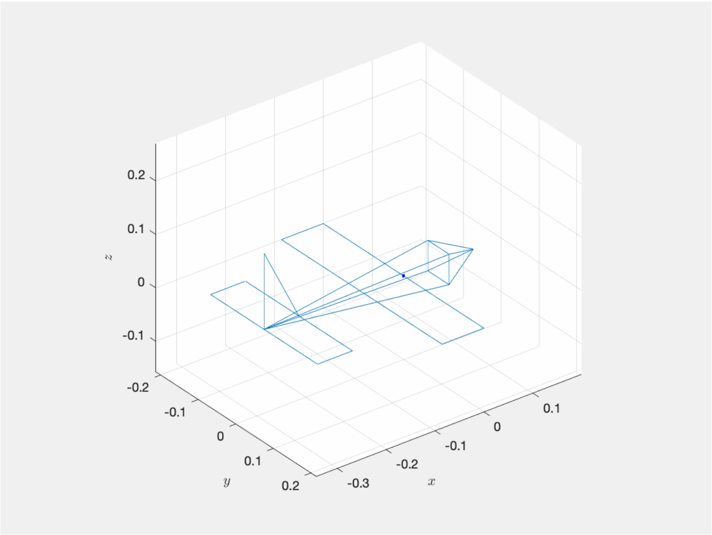
- Begin in hover
- Early Failure/Degradation
- Ascent to 200 ft
- Heavily utilizes rotors in VTOL flight regime



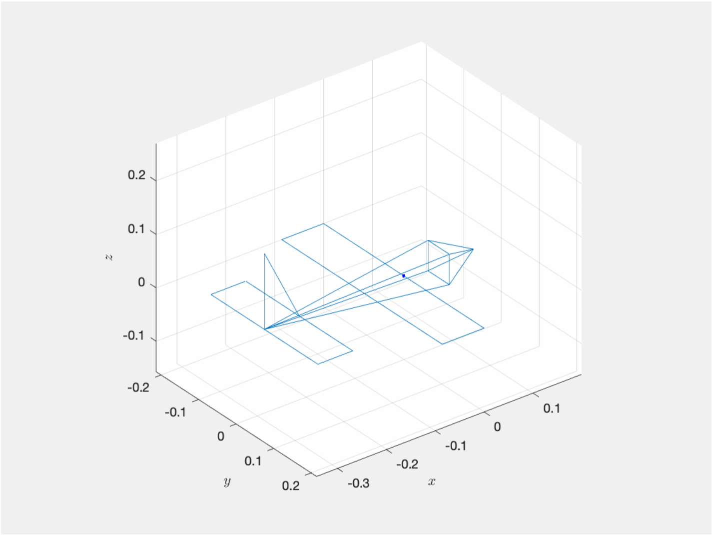




Takeoff Failure Without PDDP



50 % Rotor 1 Degradation

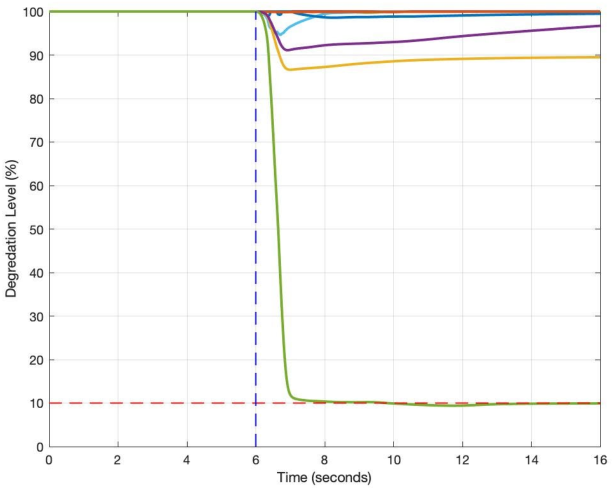


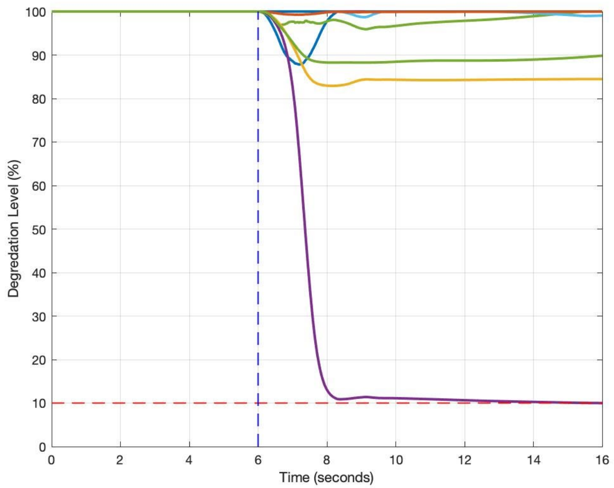
90 % Rotor 1 Degradation

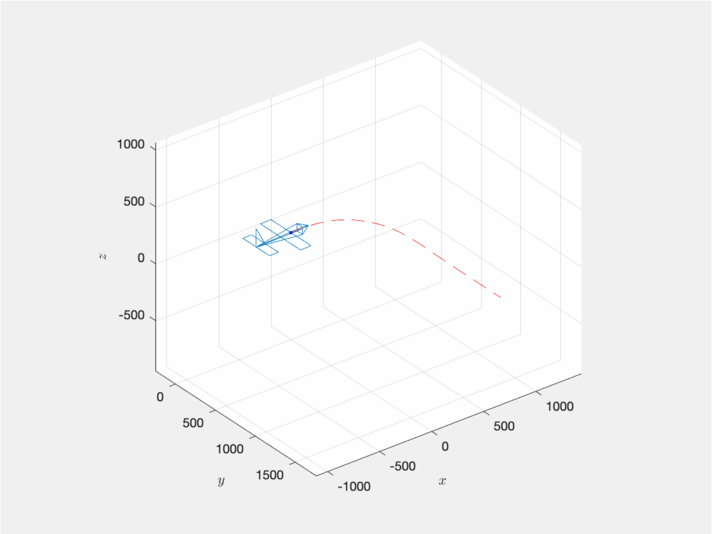
Fault Detection: Effectors

Experiment 2: Bank Right Turn

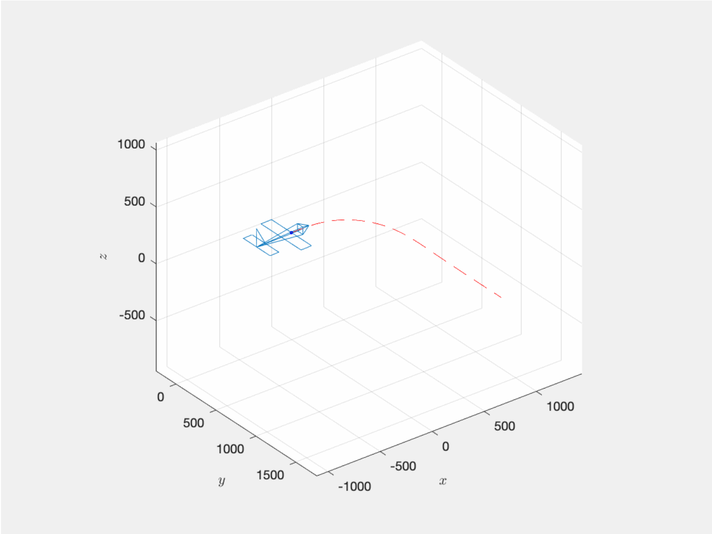
- Begin in fixed-wing cruise
- Failure/Degrad at 6 seconds
- Perform a right bank turn
- Heavily utilizes effectors in fixed-wing flight regime



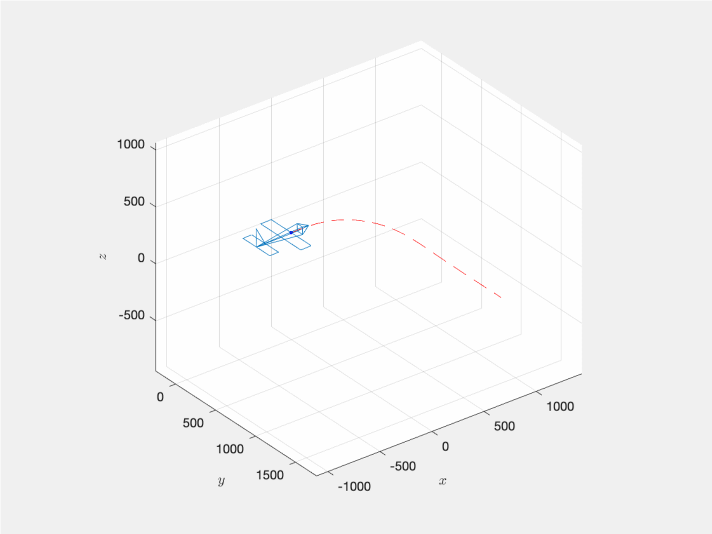




Failure Mid Bank Turn No PDDP



90 % Rudder Degradation



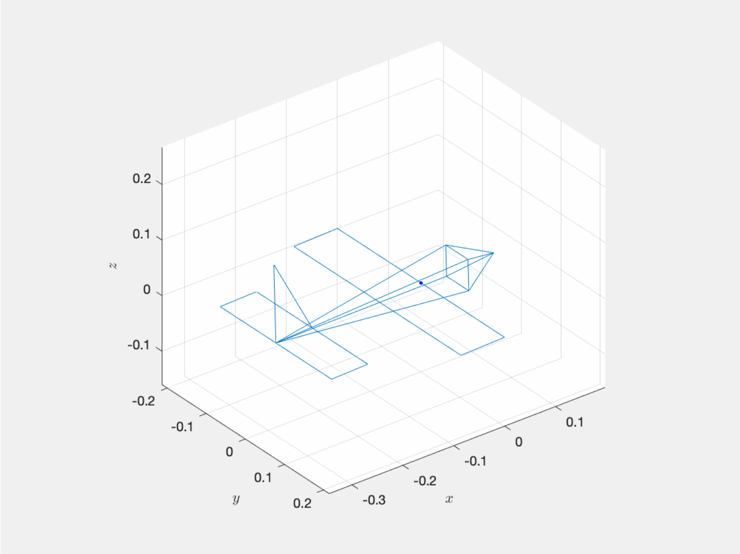
90 % Aileron Degradation

Fault Detection: Split Effector Failure

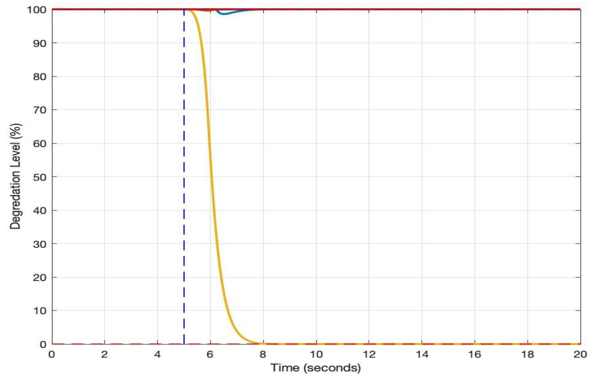
<u>Experiment 3: Split Effector Bank Right</u>

- Previous state configuration for L+C has used ganged effectors
- This experiment added state values of both the LEFT and RIGHT Ailerons
- Added states found to reduce the uncertainty of PDDP's parameter estimation even at small degradation values
- Failure/Degradation of Left Aileron ONLY at 5 seconds for bank right turn experiment
- All experiments capable of replanning a similar trajectory post failure

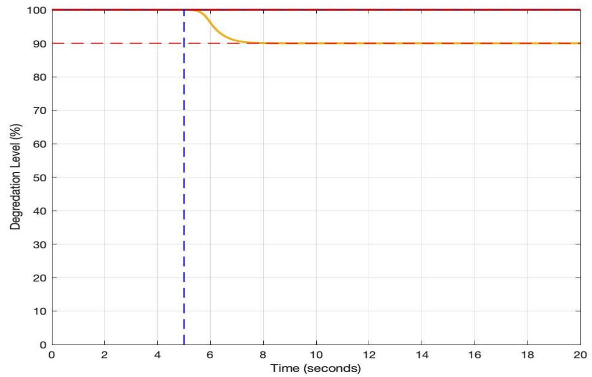
Left Aileron failure Bank Right



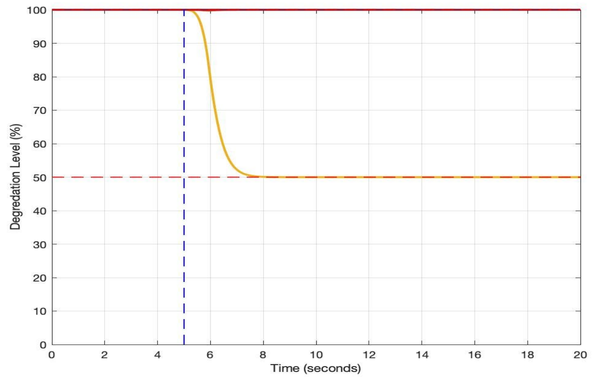
100% Degradation



10% Degradation



50% Degradation

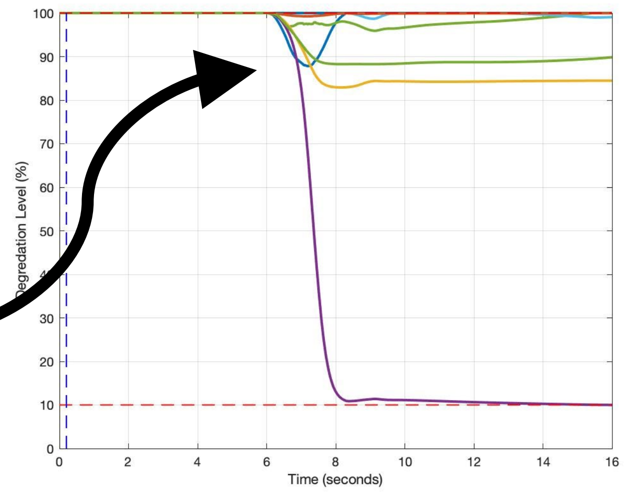


Fault Detection: Effect of Split Effector Failure

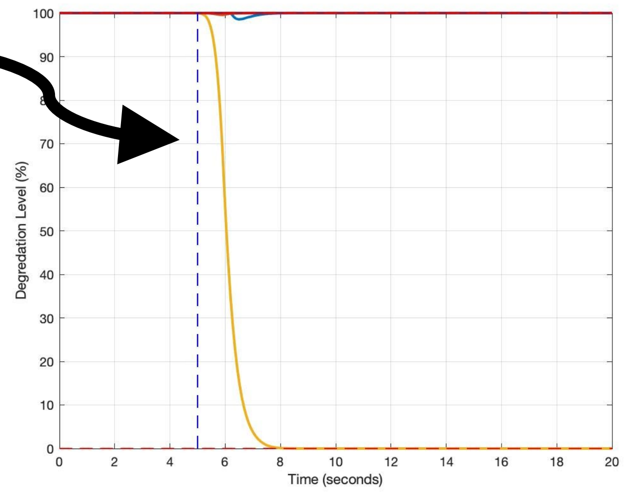
Results

- PDDP effectively utilizes state information to estimate both severe and minimal failures
- PDDP can replan using updated parameters in MPC fashion
- PDDP estimates are improved by utilization of state and the specificity of state information
- PDDP is sensitive enough to inform system ID to minor and major degradation/failures

Note:
Giving PDDP greater access to specific vehicle states improves the distinguishability of fault estimation



90 % Ganged Aileron Degradation



100 % Split Aileron Degradation

Summary - Parameterized Differential Dynamic Programming (PDDP)

- Second-order algorithm derived by extending classical optimal control
- **Convergence guarantees** independent of initialization
- **Co-optimizes** for controls and parameters simultaneously
- **Generalizes** to multiple tasks, including adaptive MPC and switching time optimization
- Enables time-optimal trajectory planning for multimodal systems, including **UAM vehicles**

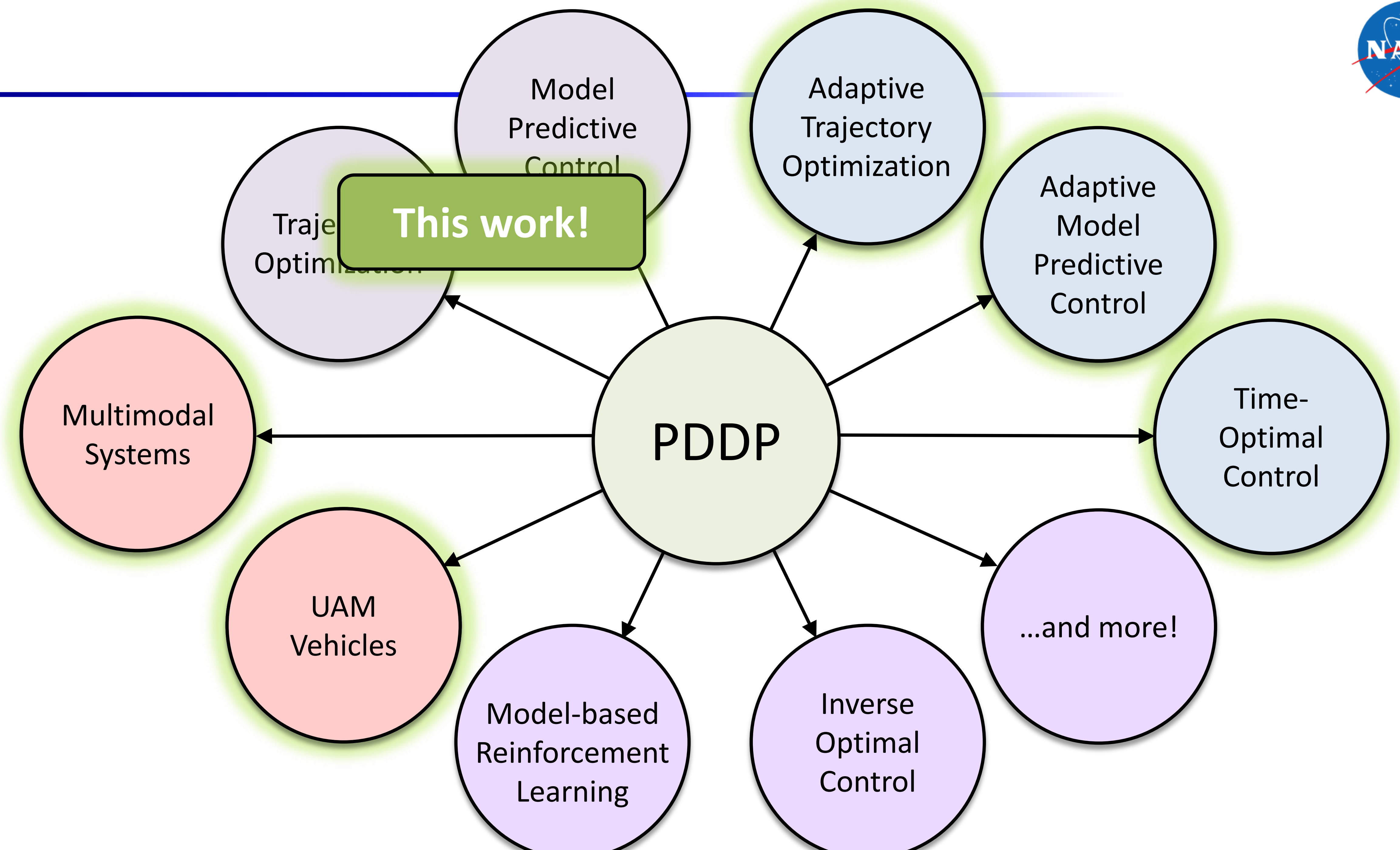
Application of PDDP – Current experimentation and directions

- **Fault detection** (parameter estimation)
  - Can run both as a full optimal control or strictly in the backward path to identify dynamic degradation
- **Adaptive MPC** - Replanning trajectory to accommodate new identified dynamics
  - Even when vehicle is incapable of following original trajectory new trajectory is planned to attain the original goal as closely as dynamically feasible
- **Switching Time Optimization**
  - **Optimal transition time** between flight regimes (Difficult for highly nonlinear vehicles like L+C)
  - **Decreases tuning** work for engineers when planning for common maneuvers that transition between flight regimes
  - Allows for the input and optimization of multiple target states for **long-term planning** and replanning

Questions?

Contact Information:

Irene.M.Gregory@nasa.gov


PDDP
Model Predictive Control
This work!
Traje
Optim
Adaptive Trajectory Optimization
Adaptive Model Predictive Control
Time-Optimal Control
Multimodal Systems
UAM Vehicles
Model-based Reinforcement Learning
Inverse Optimal Control
...and more!

Brief Overview of DDP and PDDP

Differential Dynamic Programming:

- Given nominal trajectory, use linear (or quadratic) approx. of system nonlinear dynamics and quadratic approx. of cost to yield updates to optimal controls that quadratically converge

Parametric Differential Dynamic Programming:

- Discrete system with nonlinear dynamics
- θ represents time-invariant system parameter(s)
- Goal is now to minimize the cost function with respect to both the controls, u and the parameters, θ
  - Estimation of unknown parameters and states of a dynamical system through Moving Horizon Estimation (MHE)
  - Initial parameters are set for a dynamical system, θ and for this example do not match the real system
  - The vehicle applies a portion of the trajectory given these initial parameters using a typical MPC cost
  - The resulting trajectory taken is fed into the estimation cost, which tries to find the correct parameters given the difference between the planned trajectory and what occurred on the real system
  - The new parameters are used to update the model of system. A combined cost can be derived over both task simultaneously using PDDP