Week 6

Advisor: Dr. Yue Wang

Student: David Tang

Project Recap

- **Motivation**
- Standard RL typically assumes a scalar reward
- Many problems have multiple objectives
- Trade-offs often need to be made

- **Question:**
- How to effectively navigate objective trade-offs in a balanced way?

Existing Approaches

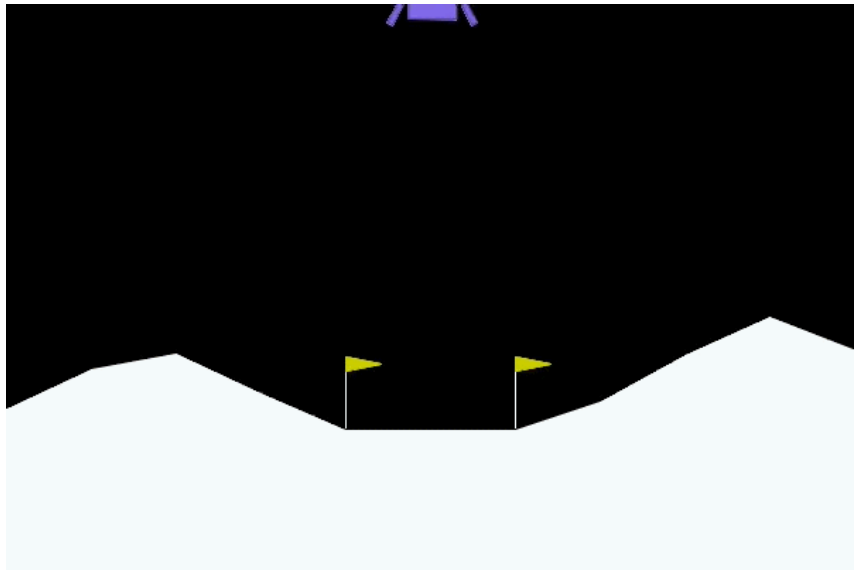
- Linear Scalarization
- Nash Bargaining Solution (Product Based)
- Chebyshev / Generalized Maximin (Floor Based)

Our Method:

- Nash + Chebyshev Convex Combination
- Compensate the weaknesses of each method

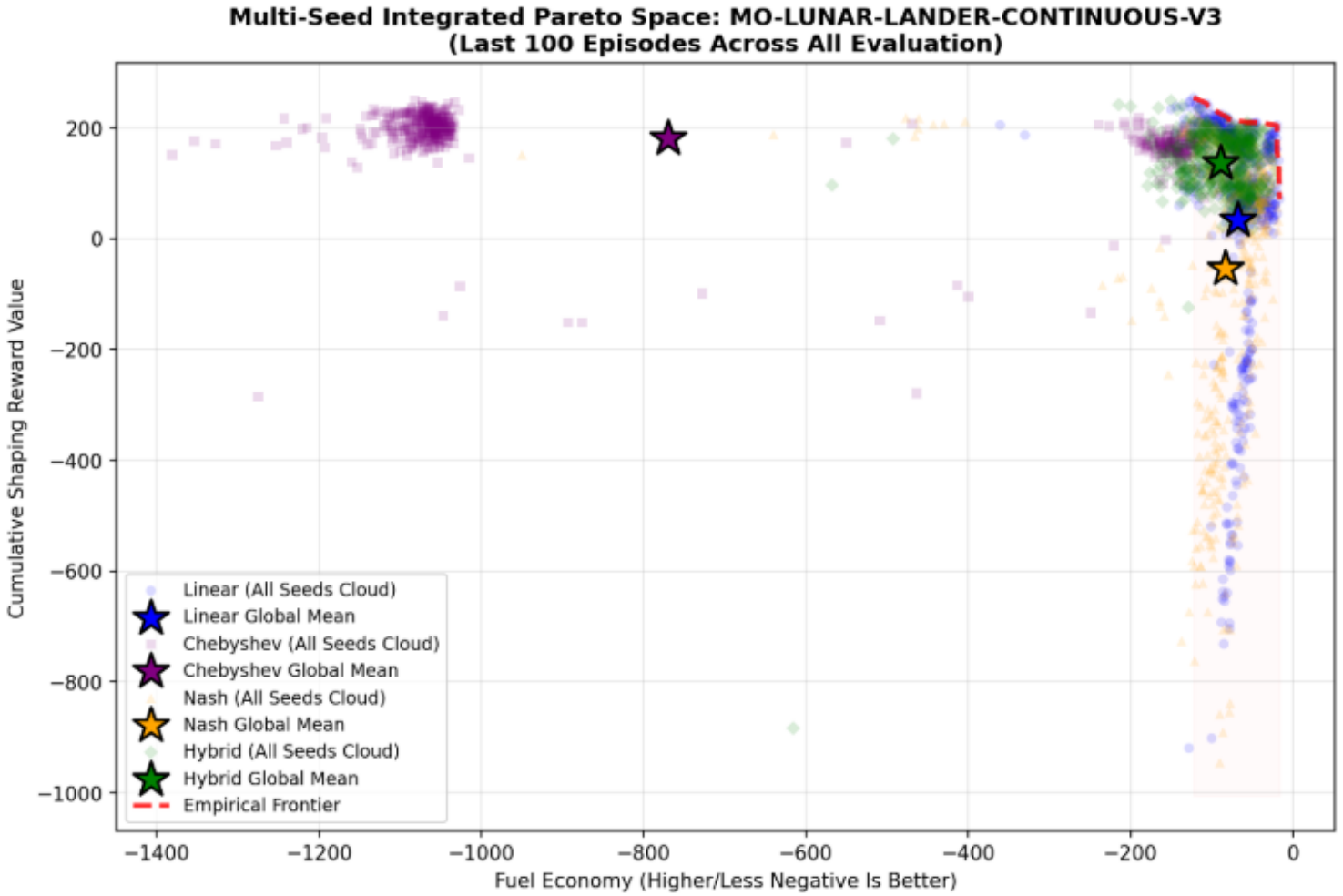
Environment

- Mo-lunar-lander-continuous-v3
- Same as mo-lunar-lander-v3 but continuous action space
- Ensure generalizable for discrete and continuous environments



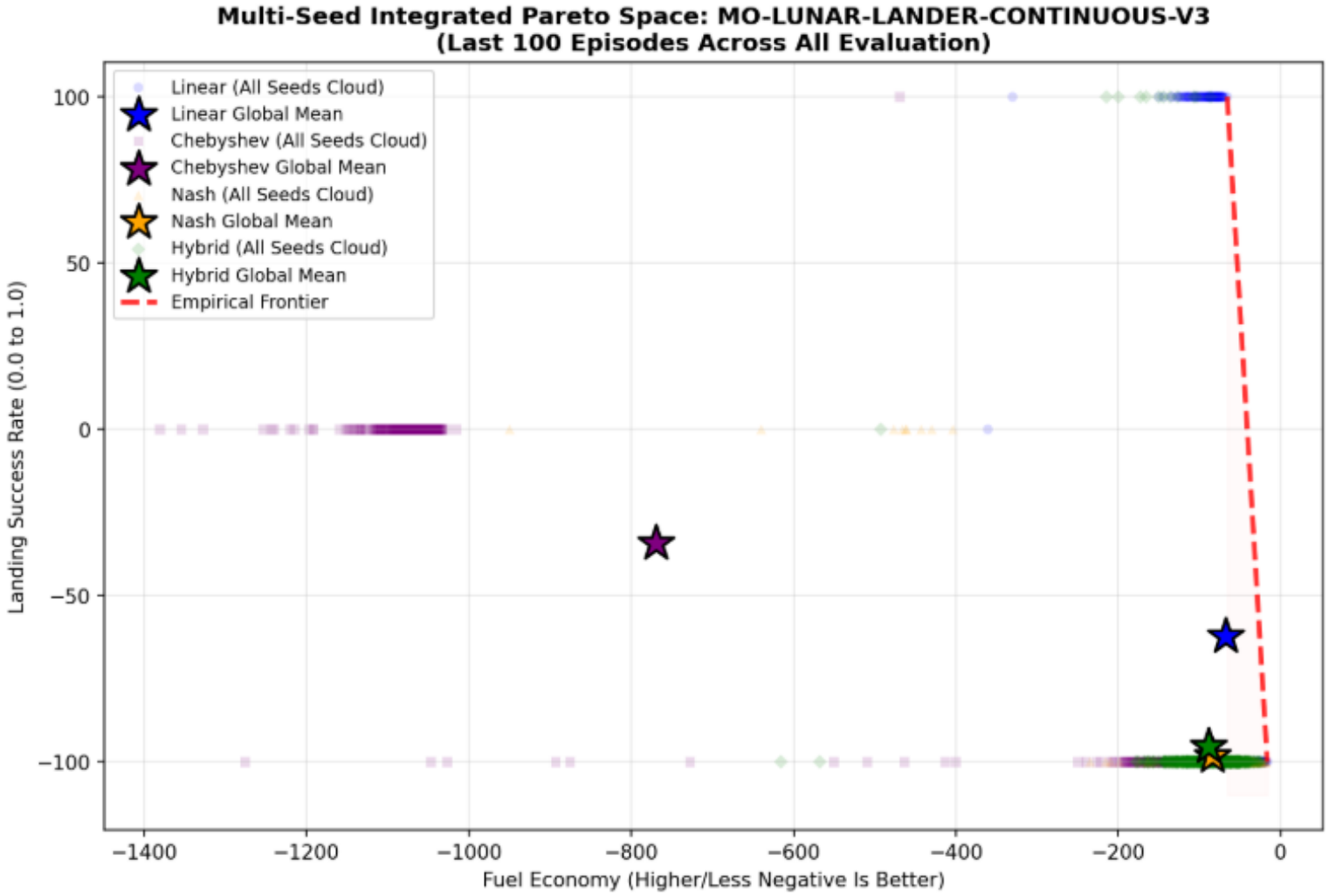
Results Tradeoff: Shaping vs Fuel

- 5 seeds
- Our method navigates to middle of pareto front

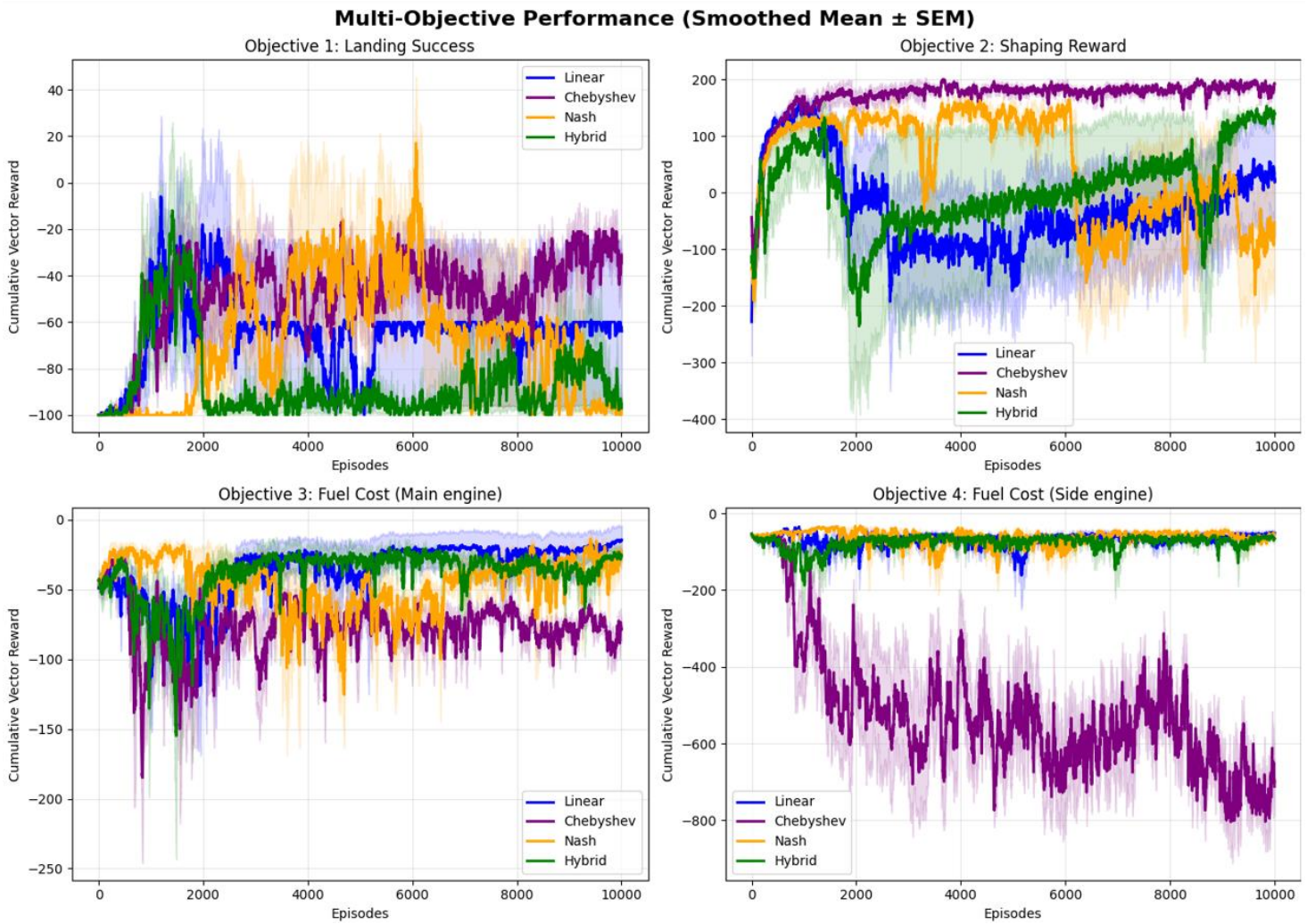


Results Tradeoff: Landing vs Fuel

- 10k episodes is early for multi objective stochastic continuous environment
- No method consistently on frontier

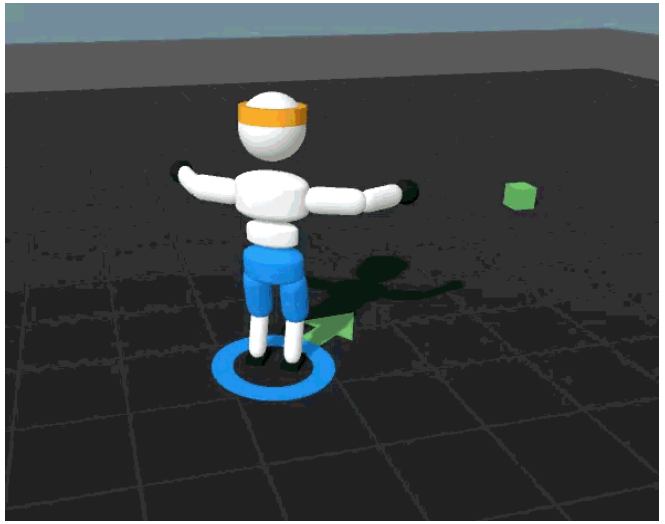


Training Curves



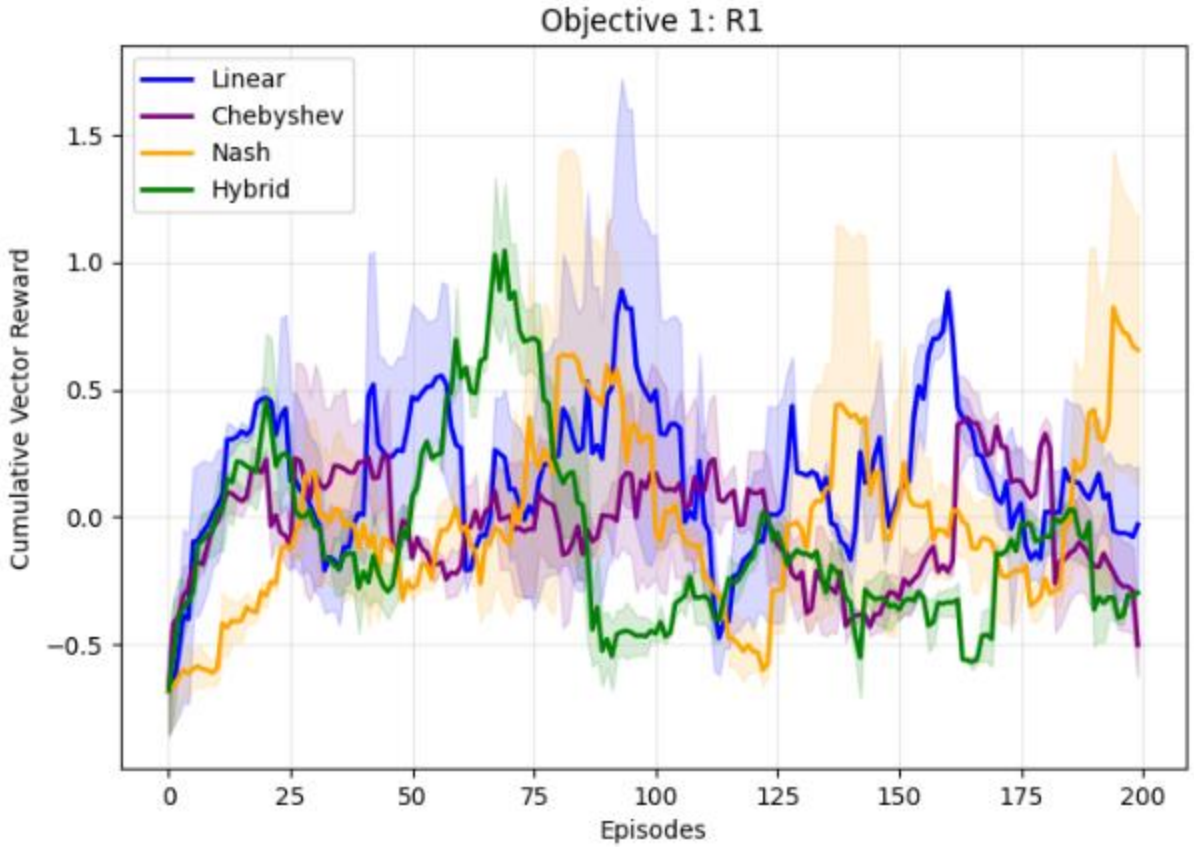
Other Environment

- Walker ML-Agents
- 4 Objectives
  - Body velocity in goal direction
  - Head y position
  - Body alignment with goal direction
  - Head velocity – body velocity
- Continuous action space
- Continuous state space



Other Environment

- Slower Training
- Shows feasible in complex environments
- Put on hold for now



Next Steps

- Develop extra metric for tradeoff frontier stability/smoothness
- Hyperparameter tuning and ablations
- Statistical Analysis and Paper Writing

Thank You

- Questions?