# Training Dynamics and Decision Stability in Multi-Objective Reinforcement Learning via Spectral Sensitivity Analysis

*David Tang, Mentor: Dr. Yue Wang*

# Introduction

- Pareto Frontiers are used to evaluate in multi-objective cases

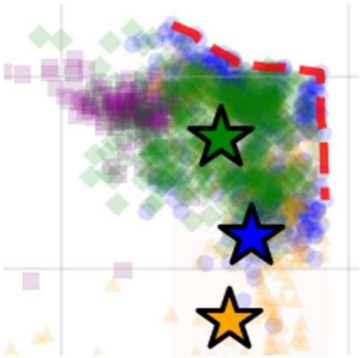
# Introduction

- Pareto Frontiers are used to evaluate in multi-objective cases
- Tradeoffs between objectives (e.g. Speed v Safety) must be made

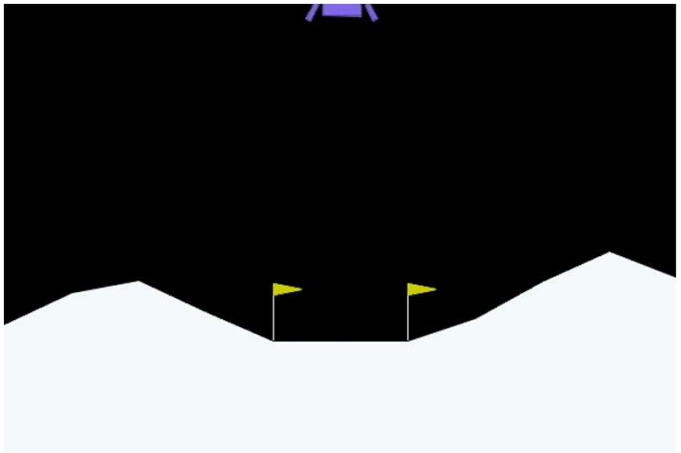
# Introduction

- Pareto Frontiers are used to evaluate in multi-objective cases
- Tradeoffs between objectives (e.g. Speed v Safety) must be made
- Optimal tradeoff point should not deviate given slight changes

# Pareto Frontier



Discovered Pareto Frontier of mo-lunar-lander-v3-continuous Shaping (y) vs fuel (x)



Lunar Lander Agent balancing fuel and pathing

# Method

Convex Combination of Nash and Chebyshev

$$f_{\text{hybrid}}(\mathbf{u};\alpha) = \alpha f_{\text{Nash}}(\mathbf{u}) + (1-\alpha) f_{\text{Chebyshev}}(\mathbf{u}), \quad \alpha \in (0,1).$$

Sensitivity Evaluation Pipeline

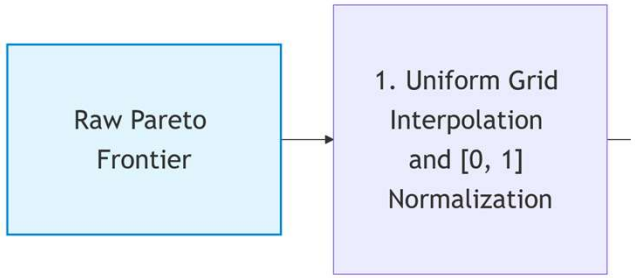


# Method

Convex Combination of Nash and Chebyshev

$$f_{\text{hybrid}}(\mathbf{u};\alpha) = \alpha f_{\text{Nash}}(\mathbf{u}) + (1-\alpha) f_{\text{Chebyshev}}(\mathbf{u}), \quad \alpha \in (0,1).$$

Sensitivity Evaluation Pipeline

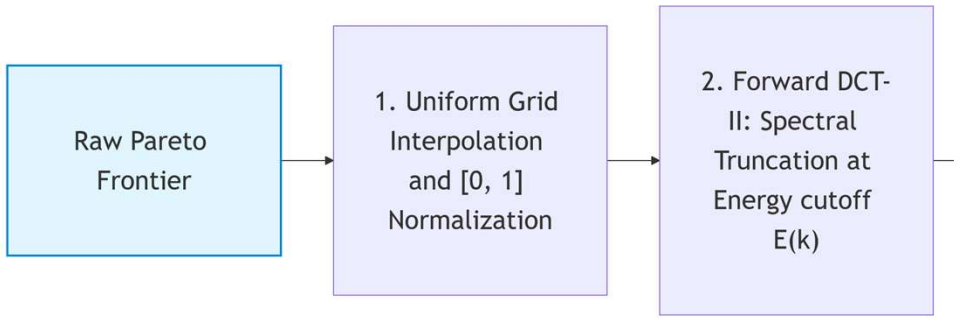


# Method

Convex Combination of Nash and Chebyshev

$$f_{\text{hybrid}}(\mathbf{u};\alpha) = \alpha f_{\text{Nash}}(\mathbf{u}) + (1-\alpha) f_{\text{Chebyshev}}(\mathbf{u}), \quad \alpha \in (0,1).$$

Sensitivity Evaluation Pipeline

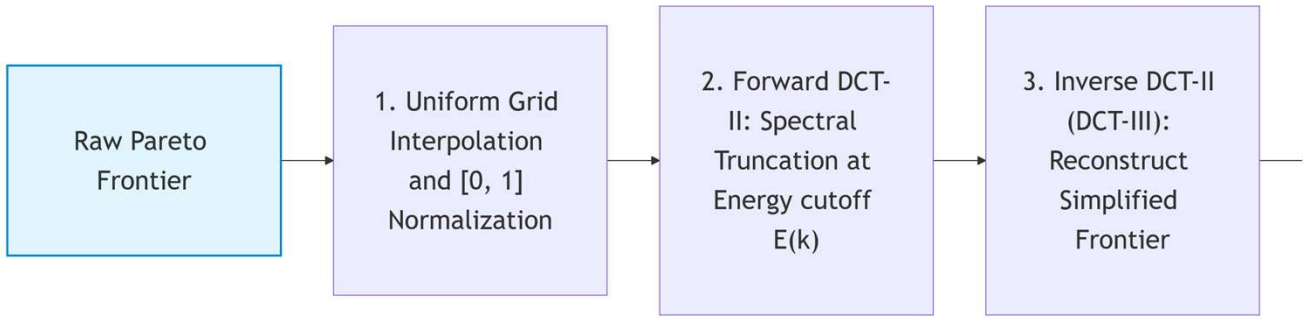


# Method

Convex Combination of Nash and Chebyshev

$$f_{\text{hybrid}}(\mathbf{u};\alpha) = \alpha f_{\text{Nash}}(\mathbf{u}) + (1-\alpha) f_{\text{Chebyshev}}(\mathbf{u}), \quad \alpha \in (0,1).$$

Sensitivity Evaluation Pipeline

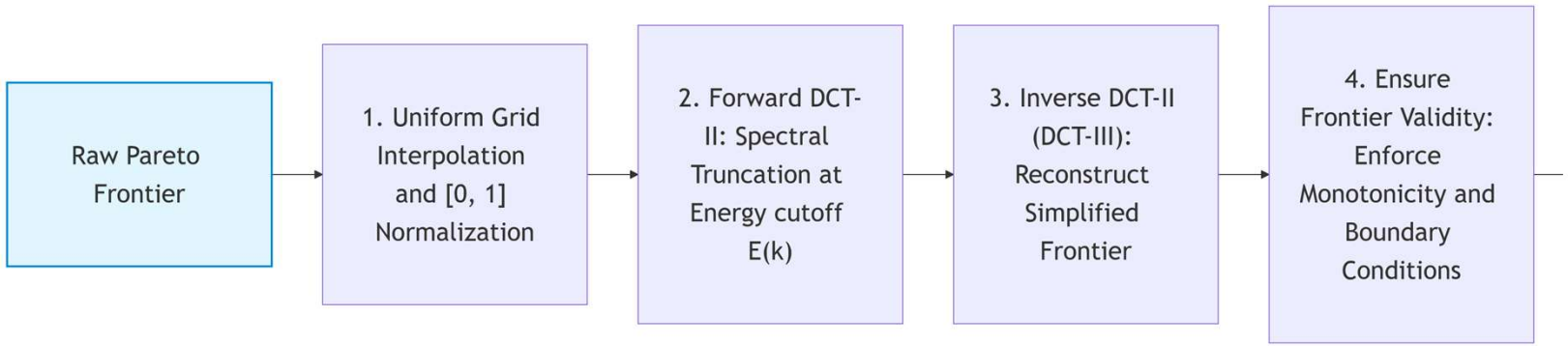


# Method

Convex Combination of Nash and Chebyshev

$$f_{\text{hybrid}}(\mathbf{u};\alpha) = \alpha f_{\text{Nash}}(\mathbf{u}) + (1-\alpha) f_{\text{Chebyshev}}(\mathbf{u}), \quad \alpha \in (0,1).$$

Sensitivity Evaluation Pipeline

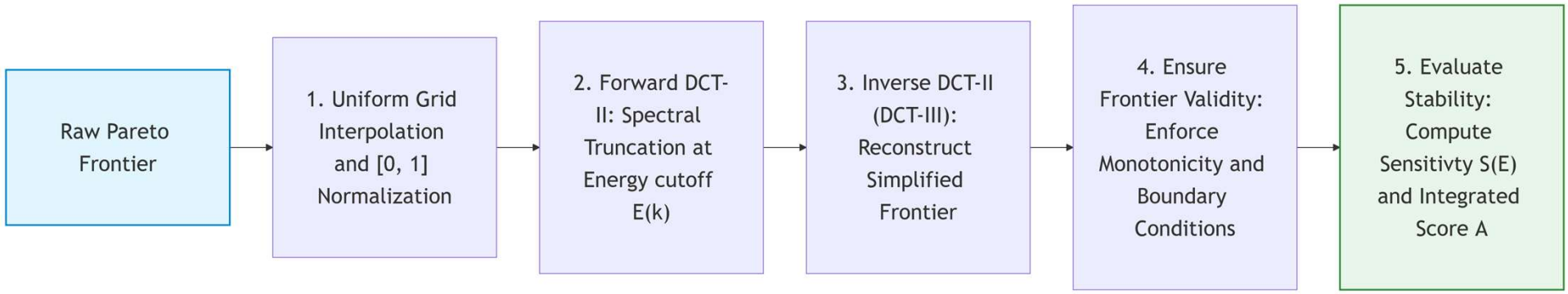


# Method

Convex Combination of Nash and Chebyshev

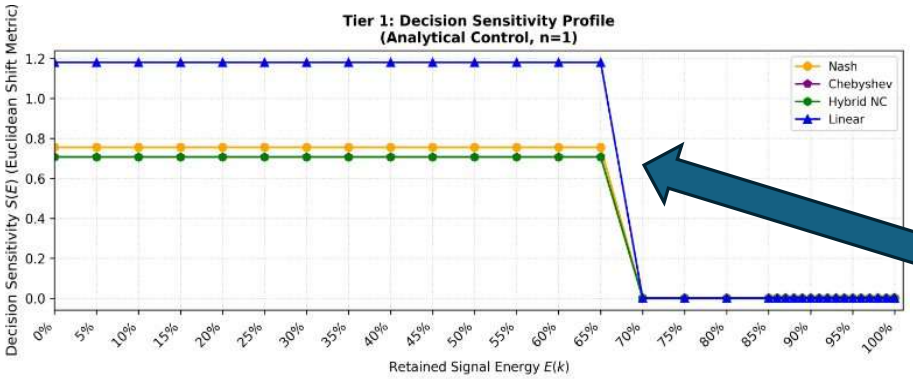
$$f_{\text{hybrid}}(\mathbf{u};\alpha) = \alpha f_{\text{Nash}}(\mathbf{u}) + (1-\alpha) f_{\text{Chebyshev}}(\mathbf{u}), \quad \alpha \in (0,1).$$

Sensitivity Evaluation Pipeline

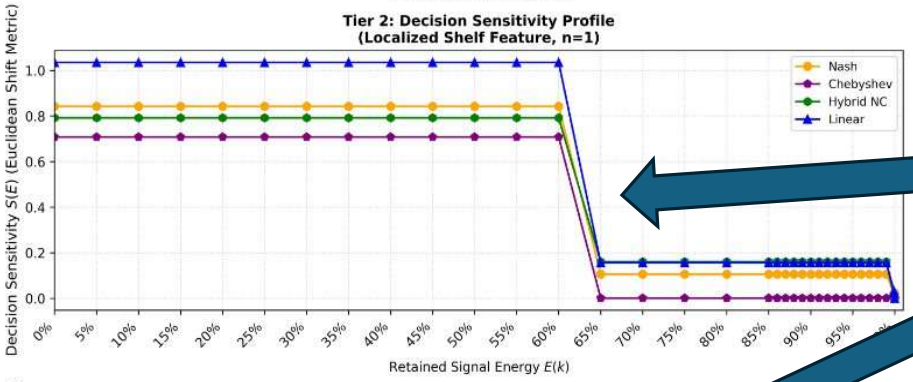


# Experiment

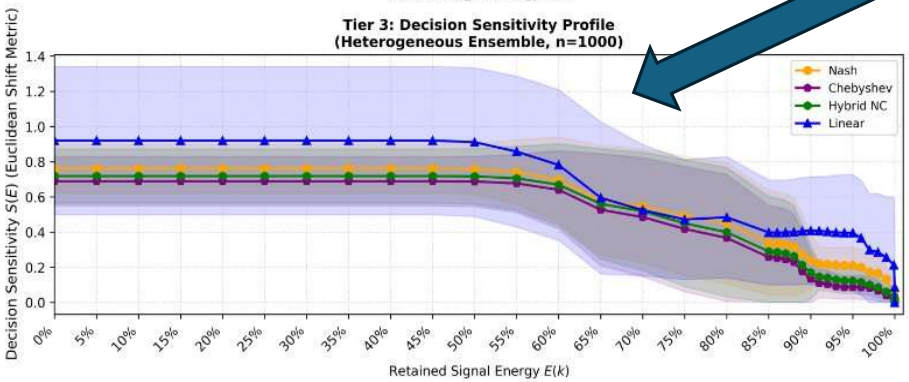
Intuition: Less Area is More Stable




Linear(Blue) highest area




Chebyshev (Purple) and Hybrid (Green) have lower area



# Experiment Continued

Table 1: Integrated Sensitivity Score Summary

| **Decision Metric** | **Tier 1** ($n = 1$) Mean ($\pm$ Std) | **Tier 2** ($n = 1$) Mean ($\pm$ Std) | **Tier 3 Ensemble** ($n = 1000$) Mean ($\pm$ Std) |
|---|---|---|---|
| MaxMin | 0.4774 ($\pm$0.0000) | 0.4427 ($\pm$0.0000) | 0.5430 ($\pm$0.1549) |
| Chebyshev | 0.4774 ($\pm$0.0000) | 0.4427 ($\pm$0.0000) | 0.5430 ($\pm$0.1549) |
| Hybrid NMM | 0.4774 ($\pm$0.0000) | 0.4675 ($\pm$0.0000) | 0.5598 ($\pm$0.1716) |
| Hybrid NC | 0.4774 ($\pm$0.0000) | 0.5543 ($\pm$0.0000) | 0.5736 ($\pm$0.1696) |
| Nash | 0.5107 ($\pm$0.0000) | 0.5657 ($\pm$0.0000) | 0.6175 ($\pm$0.1535) |
| Linear | 0.7972 ($\pm$0.0000) | 0.7047 ($\pm$0.0000) | 0.7336 ($\pm$0.1867) |

Maximin/Chebyshev is the most stable and the Hybrid variants are second most stable

# Takeaways

- Our Hybrid Method combines strengths of Chebyshev and Nash and increase stability

- We introduce a framework for determining stability of decision criteria on the Pareto Frontier

Thank You!