Week 8

Student: David Tang

Mentor: Dr. Yue Wang

Project Recap

Motivation

- Standard RL typically assumes a scalar reward
- Many problems have multiple objectives
- Tradeoffs need to be made, but real-world frontiers are noisy

Question

How to effectively navigate tradeoffs in a balanced way and determine their stability?

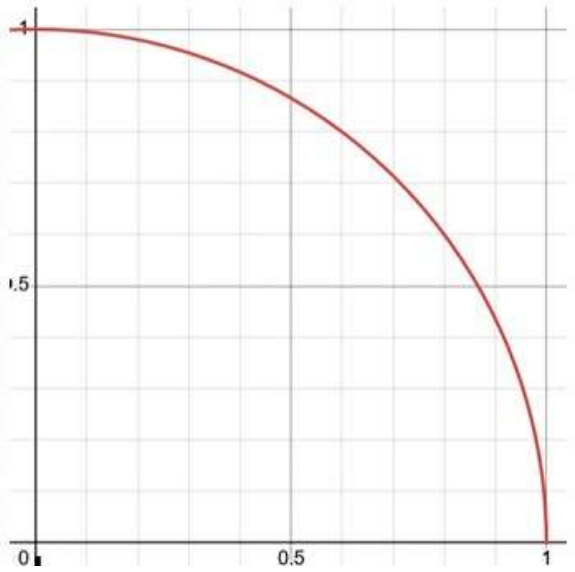
Project Recap

Existing Approaches

- Linear, Nash, Chebyshev
- Arbitrary Point Evaluation
- Artificial Instability Added

Our Method

- Nash + Chebyshev
- Integrated Sensitivity Score



Sensitivity Metric Recap

A chosen multi-objective selection rule $\mathcal{D}$ (e.g., Nash, Chebyshev, Linear)

Optimal Vector $\mathbf{v}^* = \mathcal{D}(\mathbf{x}, \mathbf{y}) = [x^*, y^*]^T$ and $\mathbf{v}_E^* = \mathcal{D}(\hat{\mathbf{x}}_E, \hat{\mathbf{y}}_E) = [x_E^*, y_E^*]^T$

Sensitivity Curve $S(E) = \|\boldsymbol{v}^* - \boldsymbol{v}_E^*\|_2$ Retained signal energy

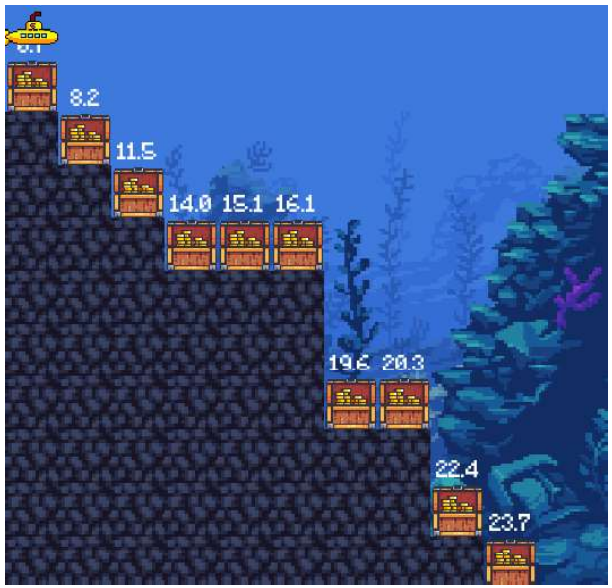
$E \in [0, 1]$

Integrated Sensitivity Score $A = \int_0^1 S(E)\, dE$

Intuition: How stable is our selection rule under simplification?

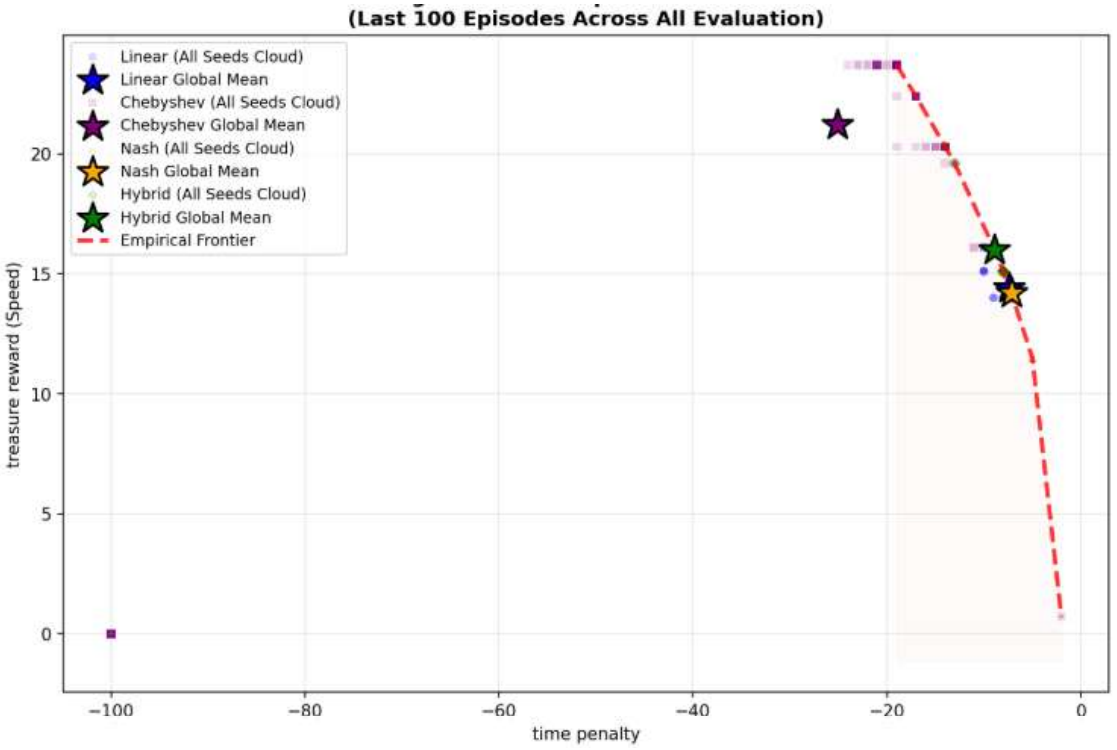
Environment 1

- deep-sea-treasure-v0
- 2 Objectives
  - Treasure Reward
  - Time Penalty
- Discrete action space
- Discrete state space



Deep-Sea-Treasure-V0 Frontier

- Convex
- Shows the discovered empirical frontier



Deep-Sea-Treasure-V0 Sensitivity

Integrated Scores
(Lower is better)

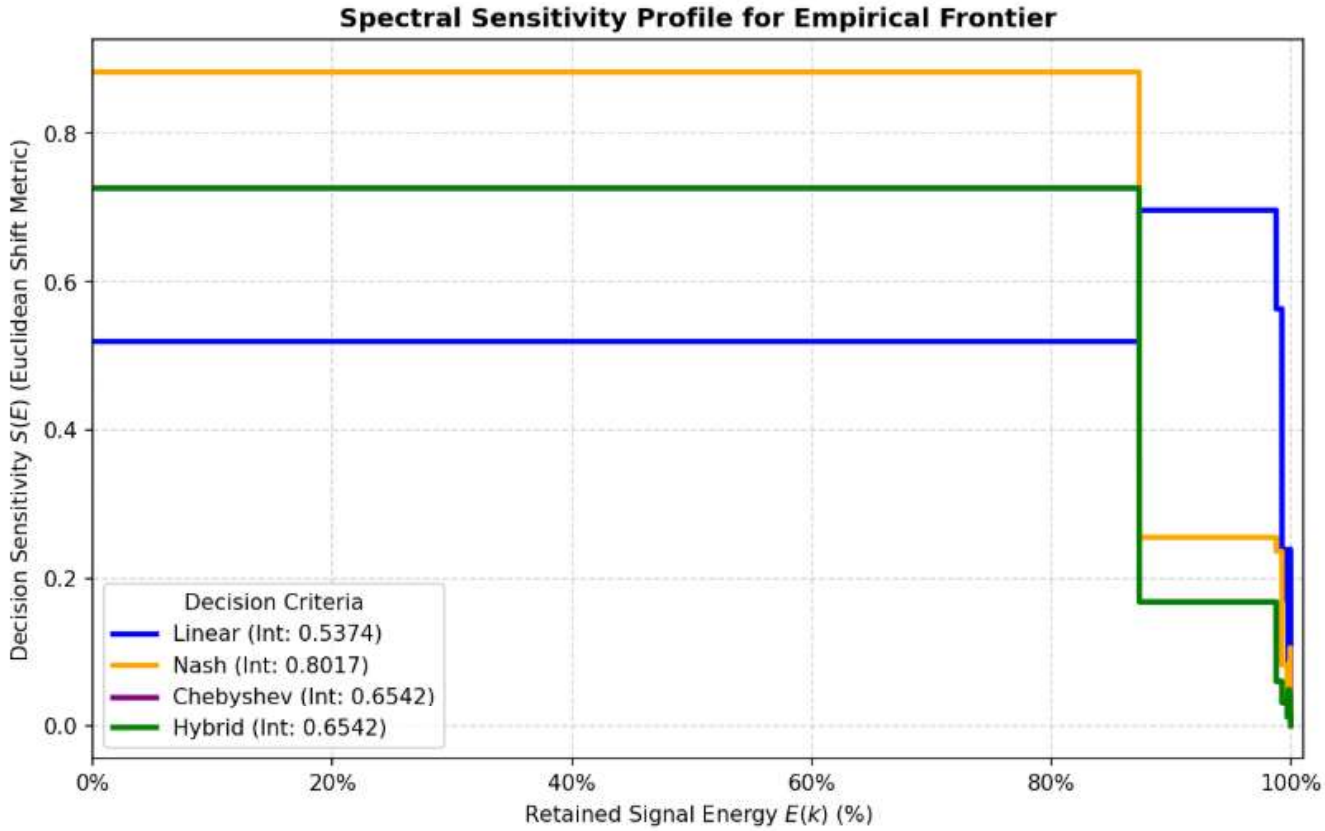
1st Linear (0.53)

2nd Chebyshev (0.65)

2nd Hybrid (0.65)

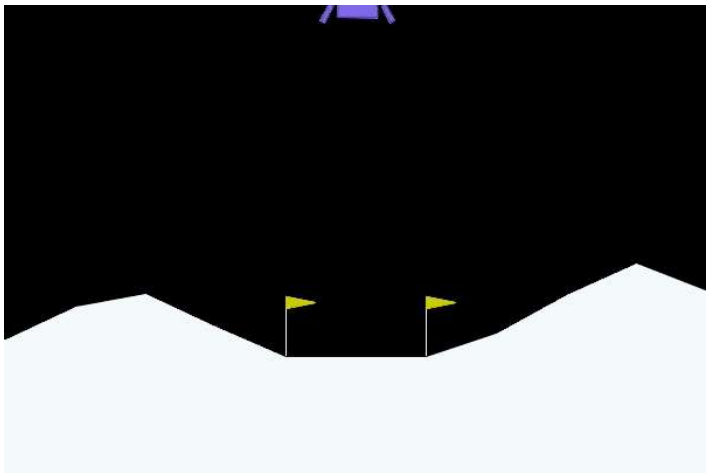
3rd Nash (0.80)

Linear instability
spikes at ~E=0.9



Environment 2

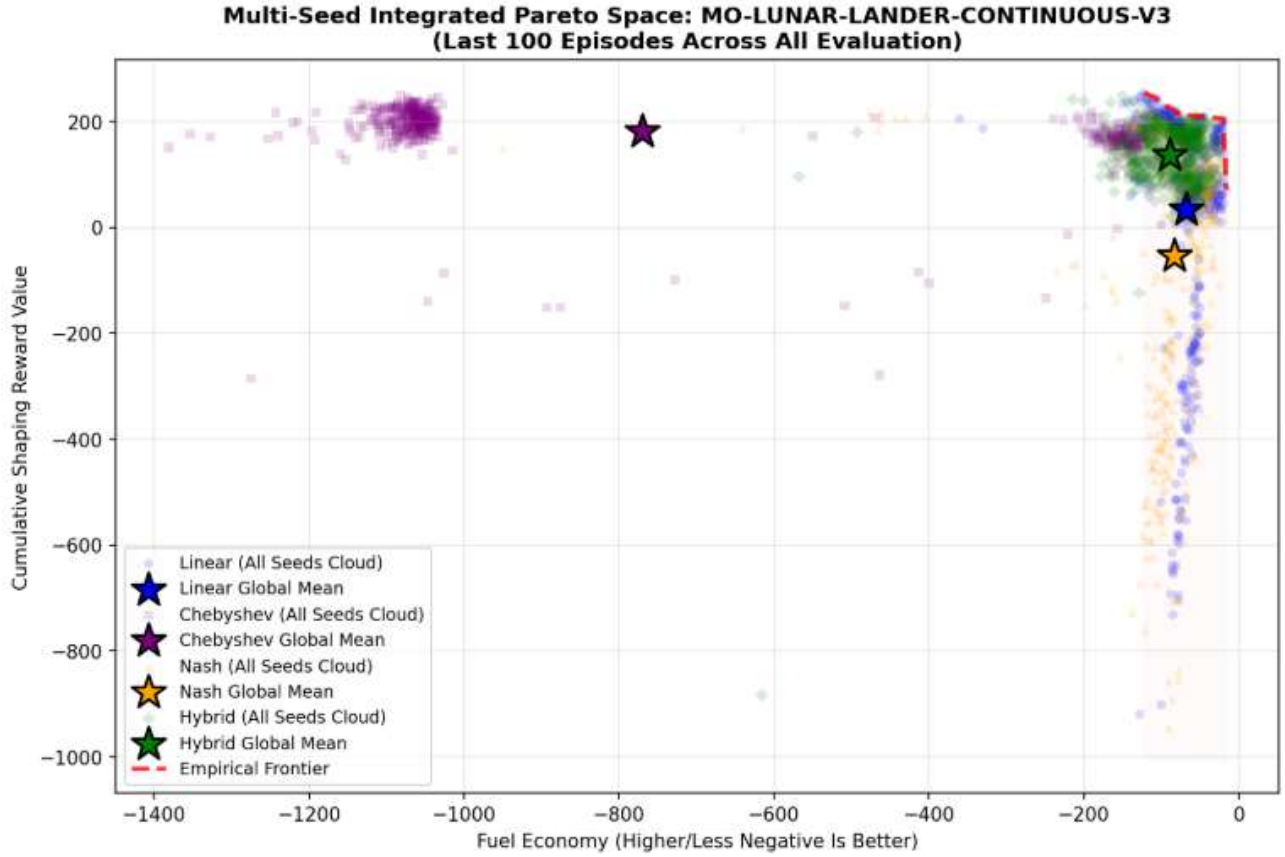
- mo-lunar-lander-continuous-v3
- 4 Objectives
  - Landing Success
  - Shaping Reward
  - Main Engine Fuel
  - Side Engine Fuel
- Continuous action space
- Continuous state space



Mo-Lunar-Lander-Continuous-V3 Frontier 1

Shaping vs Fuel

- Knee present
- Sharp cliff
- Concave



Mo-Lunar-Lander-Continuous-V3 Sensitivity 1

Shaping vs Fuel

Integrated Scores (Lower is better)

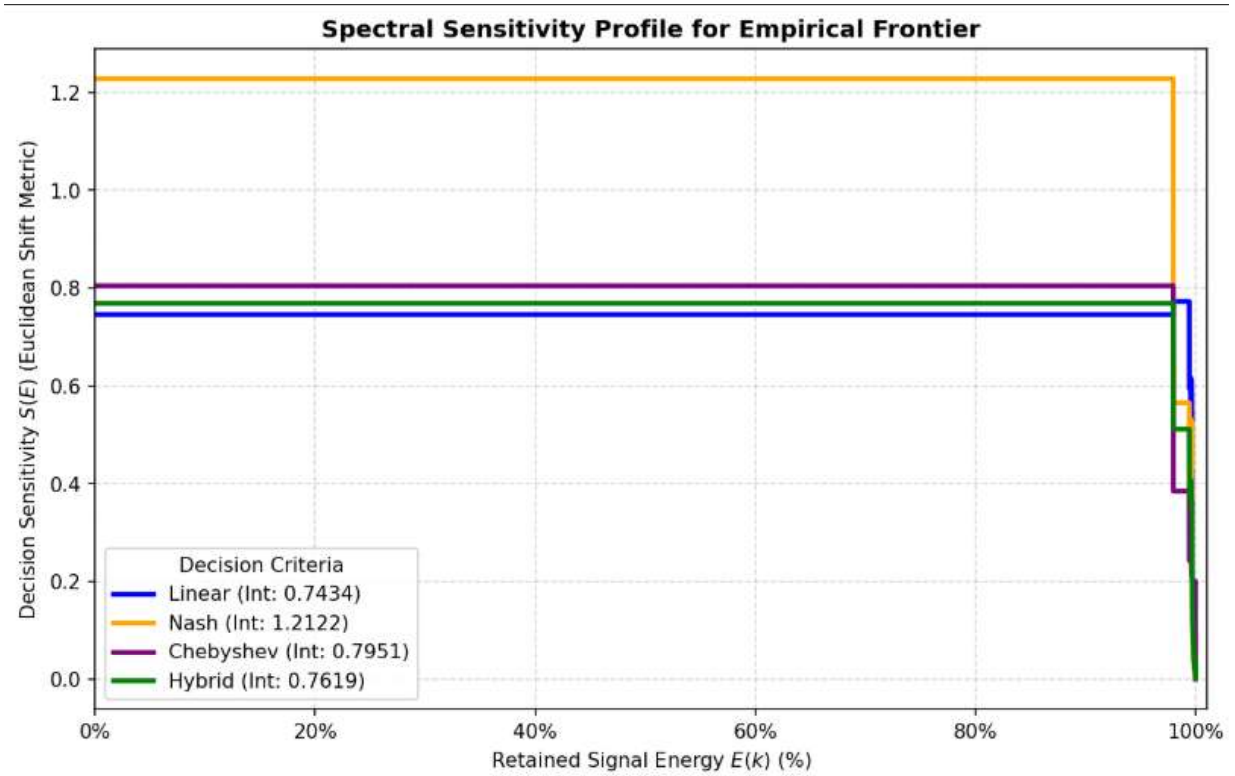
1st Linear (0.74)

2nd Hybrid (0.76)

3rd Chebyshev (0.79)

4th Nash (1.2)

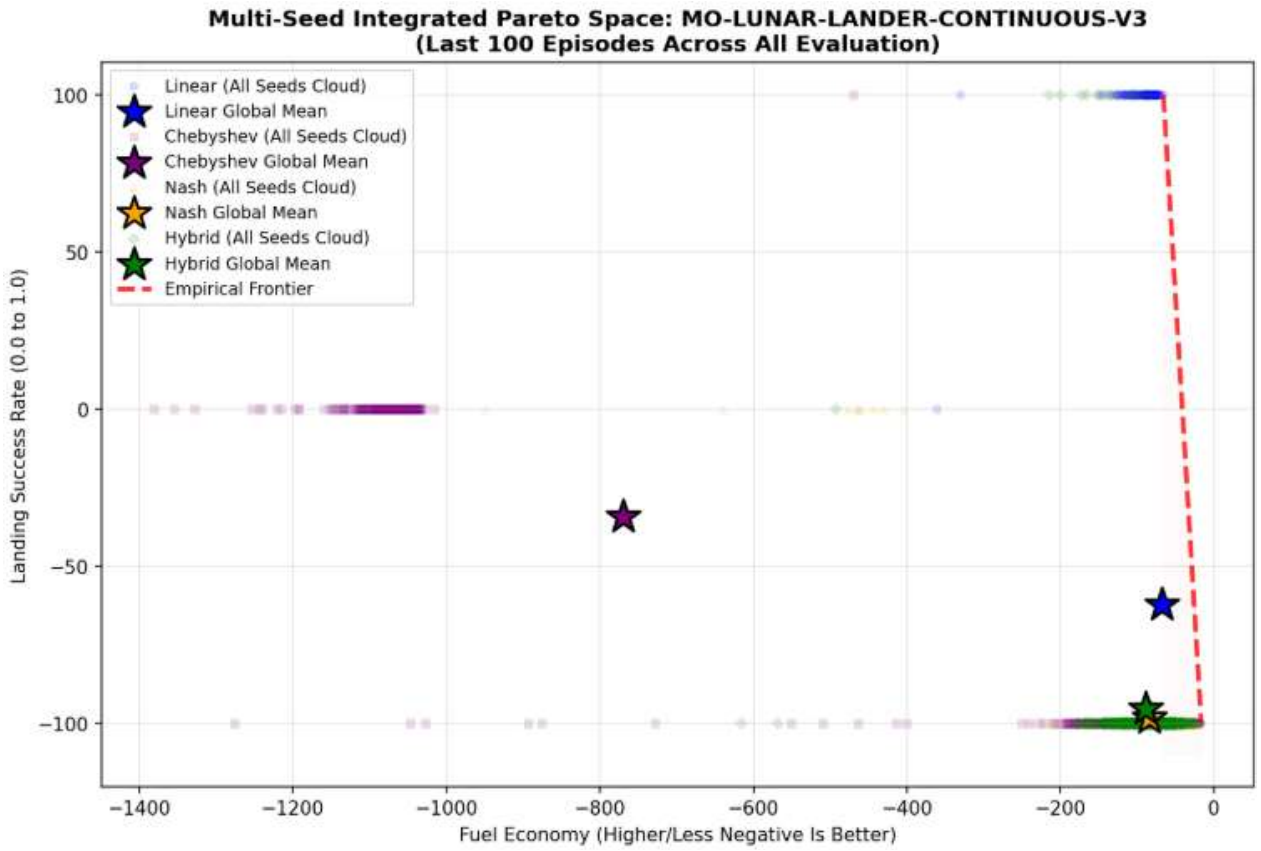
Linear small spike at ~E=0.99



Mo-Lunar-Lander-Continuous-V3 Frontier 2

Landing vs Fuel

Linear Frontier



Mo-Lunar-Lander-Continuous-V3 Sensitivity 2

Landing vs Fuel

Integrated Scores

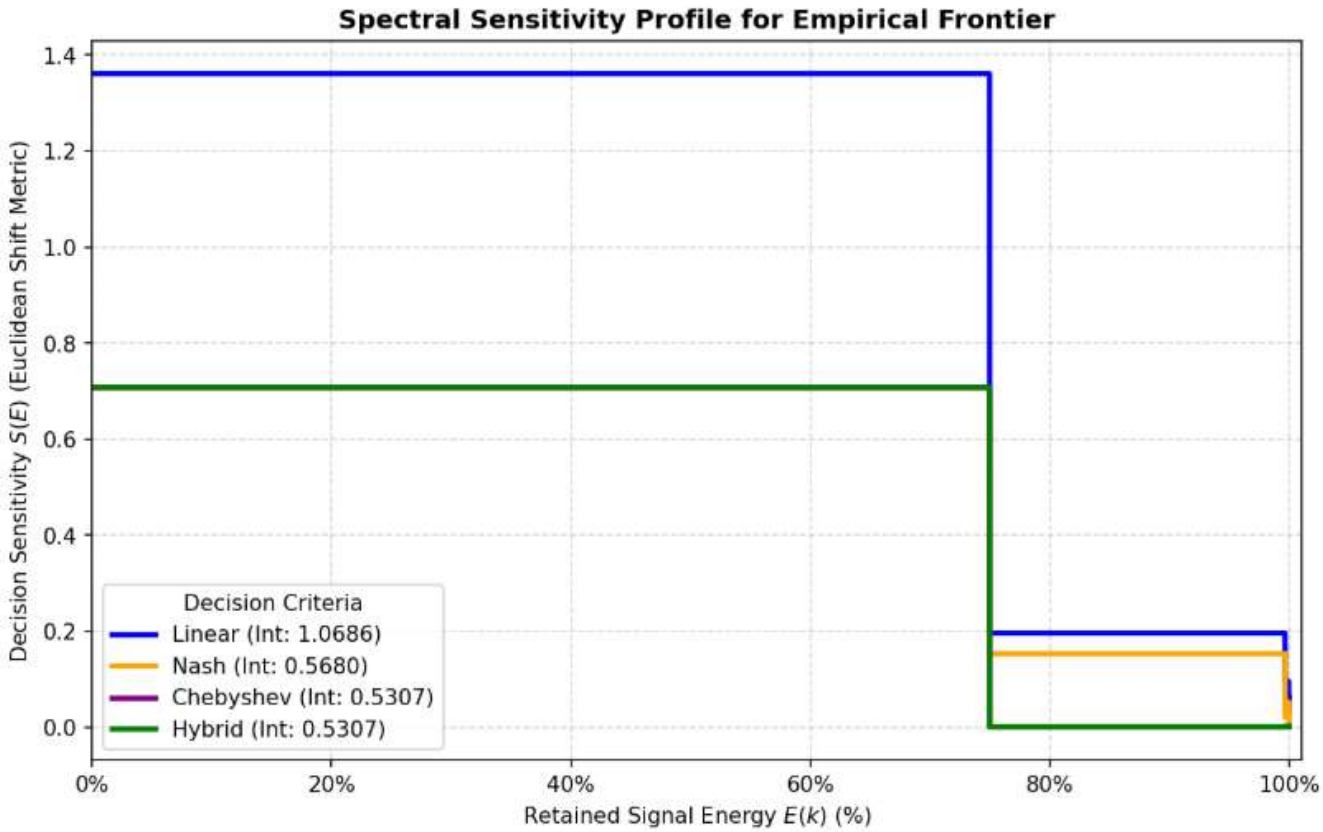
(Lower is better)

1st Chebyshev (0.53)

1st Hybrid (0.53)

2nd Nash (0.56)

3rd Linear (1.06)



Takeaways

- More complicated frontier -> higher the scores
- Linear instability can spike even when energy increases
- Nash / Chebyshev decrease monotonically as energy increases

Next Steps

- Clean up codebase
- Statistical Analysis
- Continue Paper Writing

Extras: Synthetic Examples (Analytical)

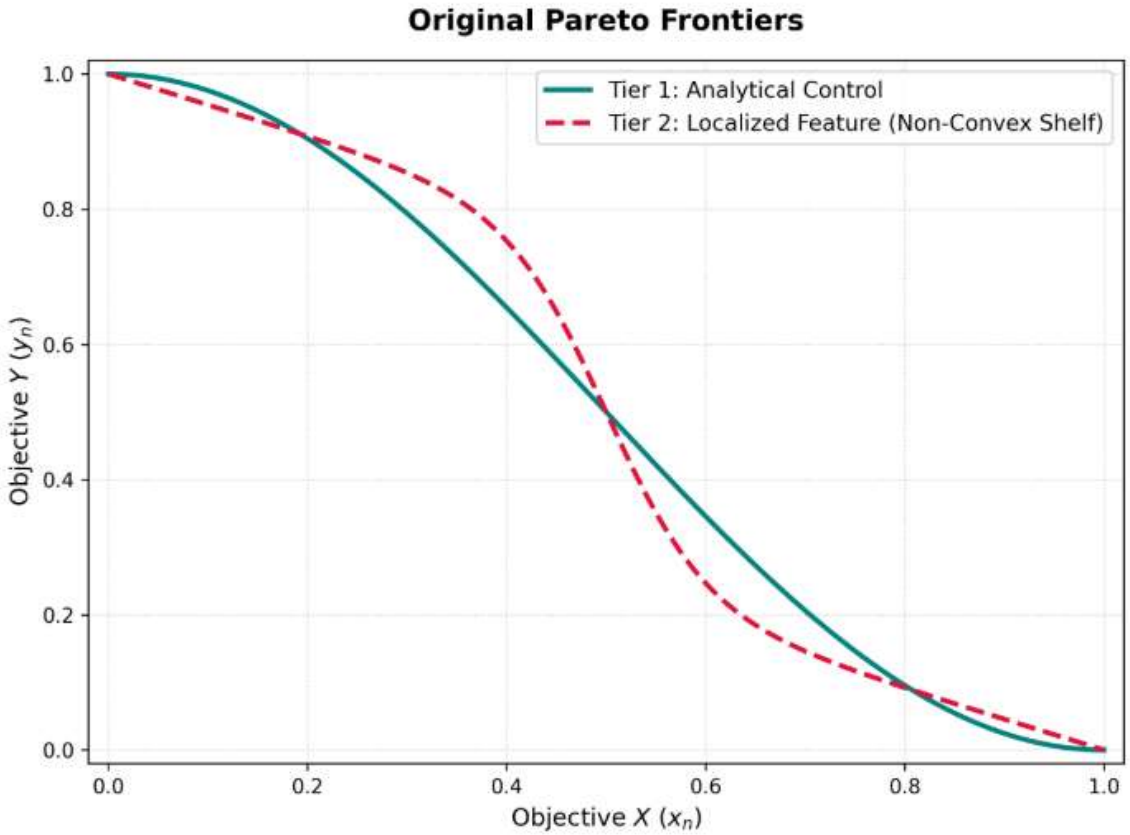
- Tier 1

$$y_n = \frac{1}{2}\left(1 + \cos(\pi x_n)\right)$$

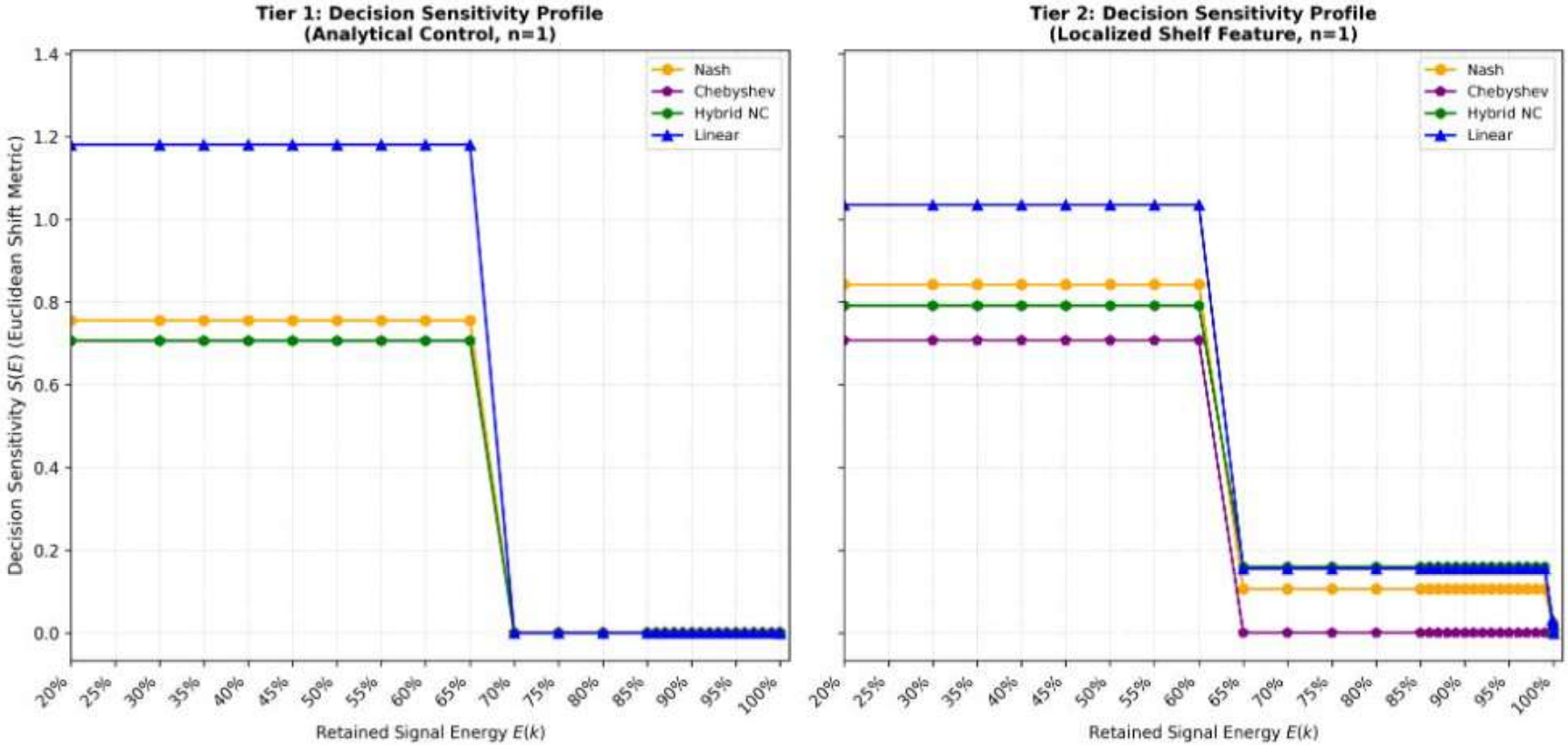
- Tier 2

$$y_n = 1.0 - 0.5x_n - 0.3\tanh\left(10(x_n - 0.5)\right)$$

$$y_n = \frac{y_n - y_{N-1}}{y_0 - y_{N-1}}$$



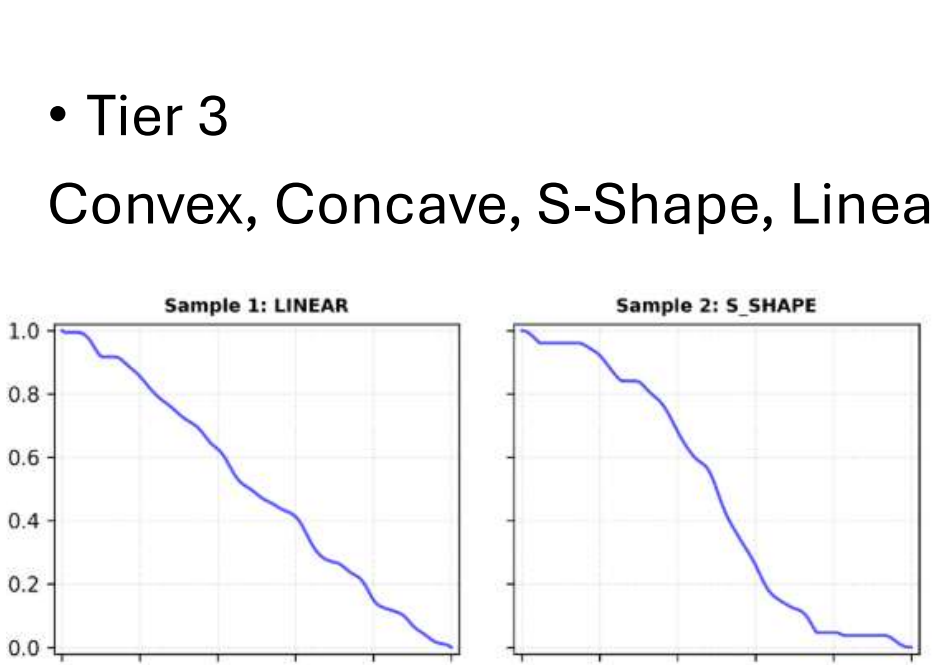
Extras: Synthetic Examples (Analytical)

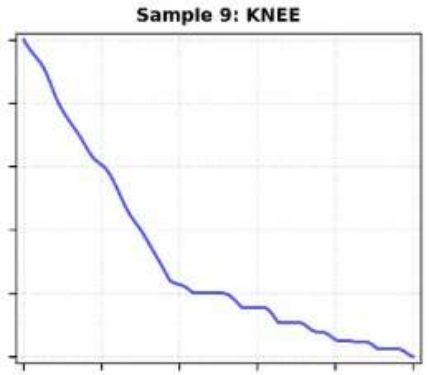


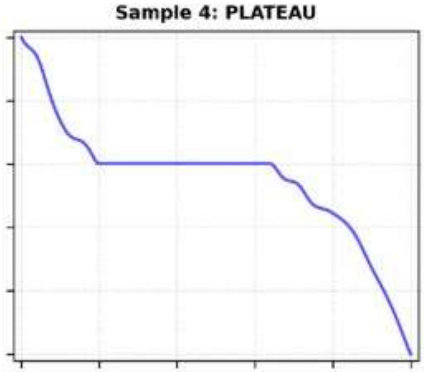
Extras: Synthetic Examples (Stochastic)

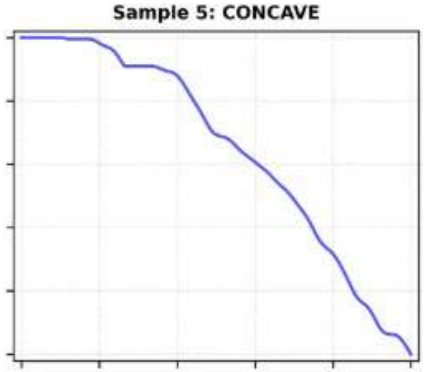
- Tier 3

Convex, Concave, S-Shape, Linear, Knee, Plateau









Extras: Synthetic Examples (Stochastic)

Shows consistency with known behavior of the different aggregation methods

