# Understanding the impact of entropy on policy optimization

Zafarali Ahmed, Nicolas Le Roux,
Mohammad Norouzi, Dale Schuurmans

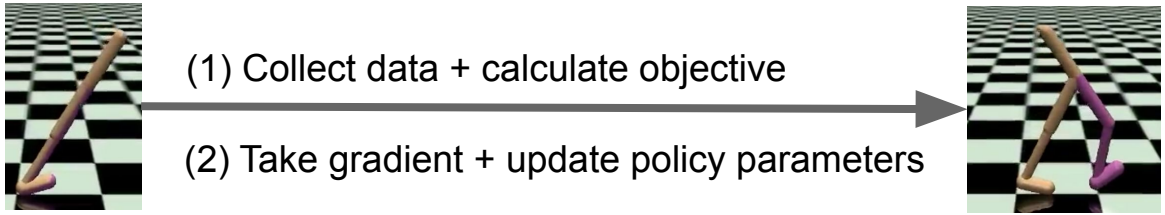

bit.ly/2HQvGoQ

*zafarali.ahmed@mail.mcgill.ca*

# Why should we understand policy optimization?

## What is policy optimization?

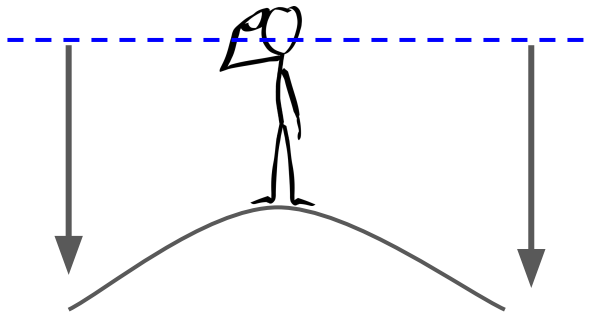
Find parameterized policy that maximizes rewards.



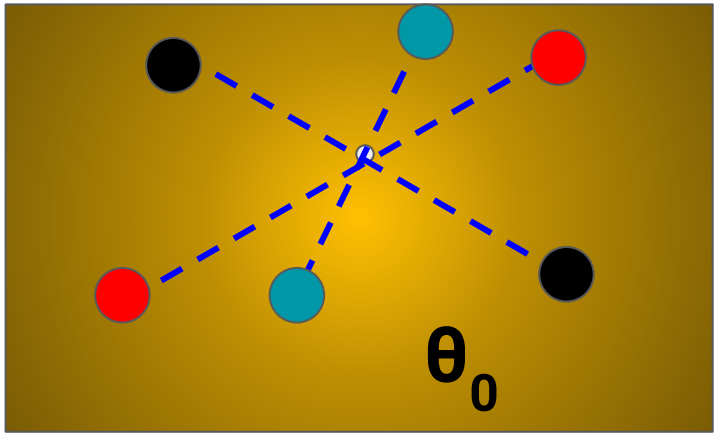

## Why is it difficult?

- Bad gradient estimates?
- Difficult geometry?
- Poor conditioning?
- Not enough “Exploration”?

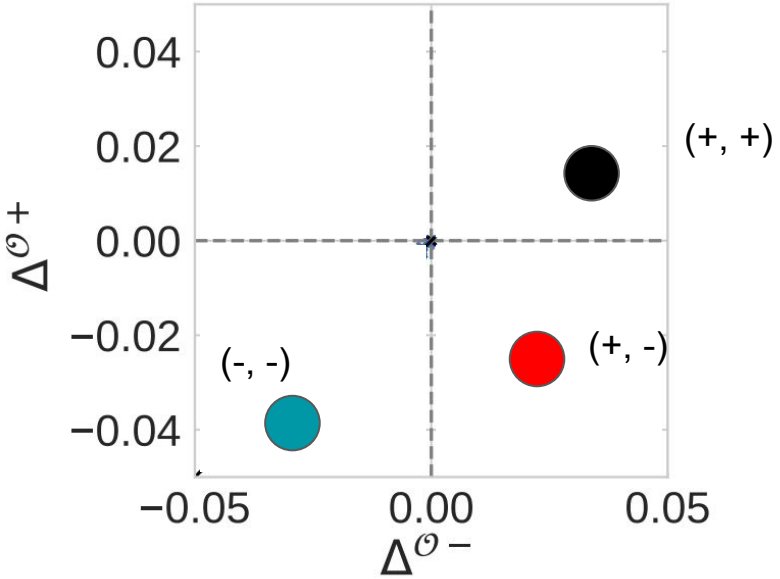
# Contribution 1: How do we study high dim objective functions?



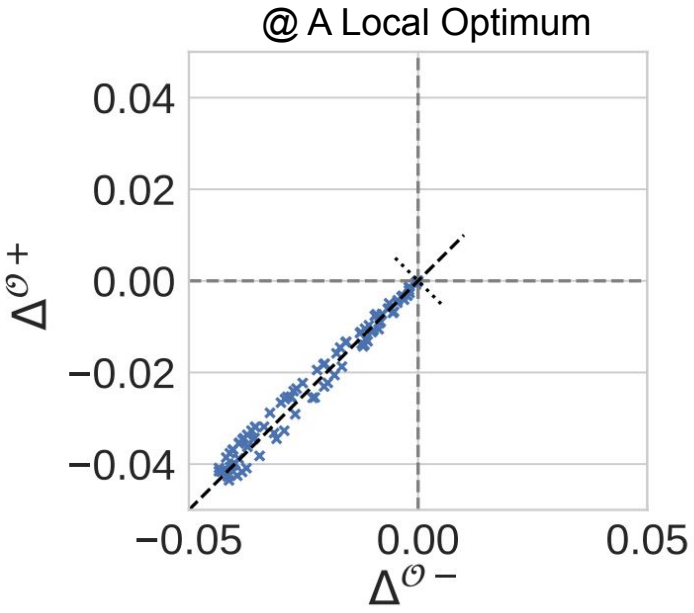
**STEP 1:** Collect **<u>random perturbations</u>** of objective

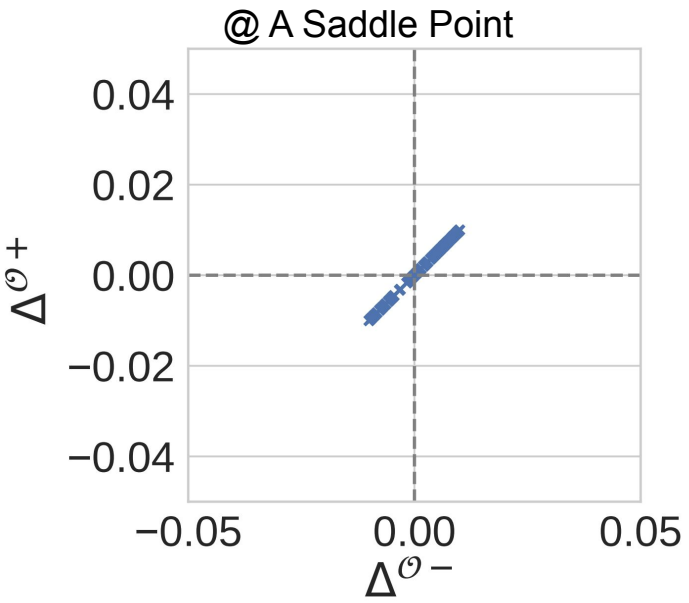



**STEP 2:** How does objective change along random perturbations?



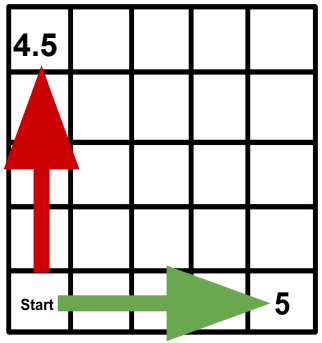
# Contribution 1: How do we study high dim objective functions?
## *Examples*

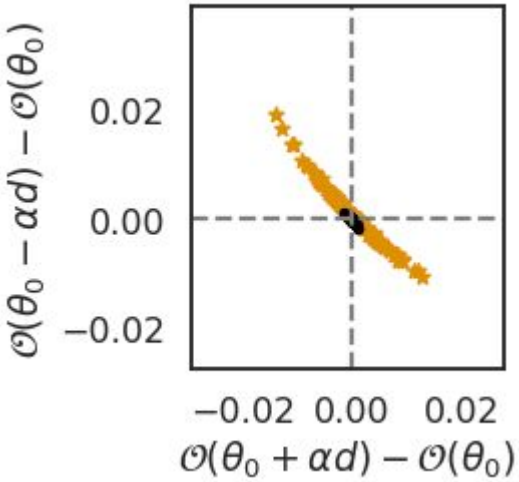


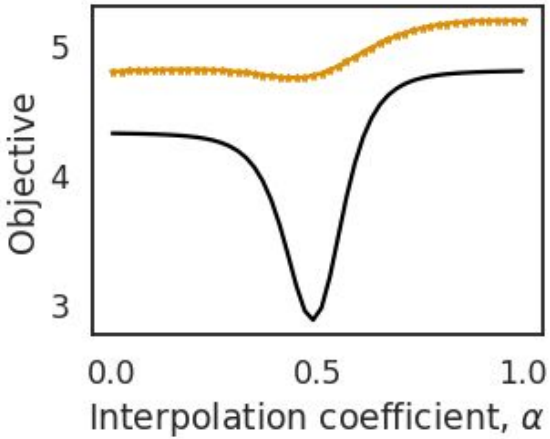



# Contribution 2: Why does entropy regularization help?

*Experiments on exact grid worlds and Mujoco*




**Conclusion:** Even the absence of gradient estimation error, policy entropy helps by smoothing the objective function:

# Understanding the impact of entropy on policy optimization

Read the paper!

bit.ly/2HQvGoQ

Come see poster!

## Poster # 29

## TODAY - 6.30 PM Pacific Ballroom

Chat with me!

**zafarali.ahmed@mail.mcgill.ca**