Emergent Communication - First steps

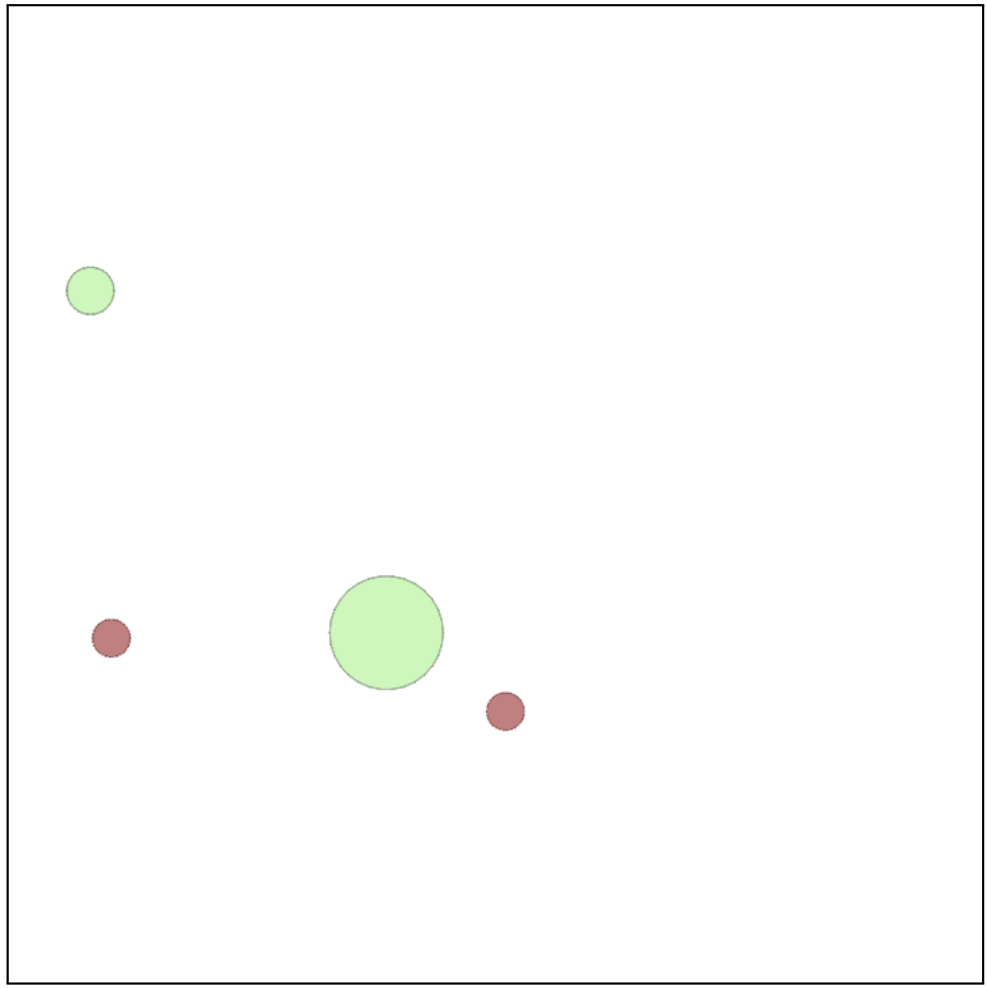
Maxime Toquebiau

December 2nd, 2021

Cooperative Push Environment

Fixing the environment:

- Bigger agents and object
- More steps (100 steps maximum)
- Reward:
  - *obj*: -1 x squared distance between object and landmark
  - *dist*: -1 x average distance between agents and object
  - + penalty for collision between agents

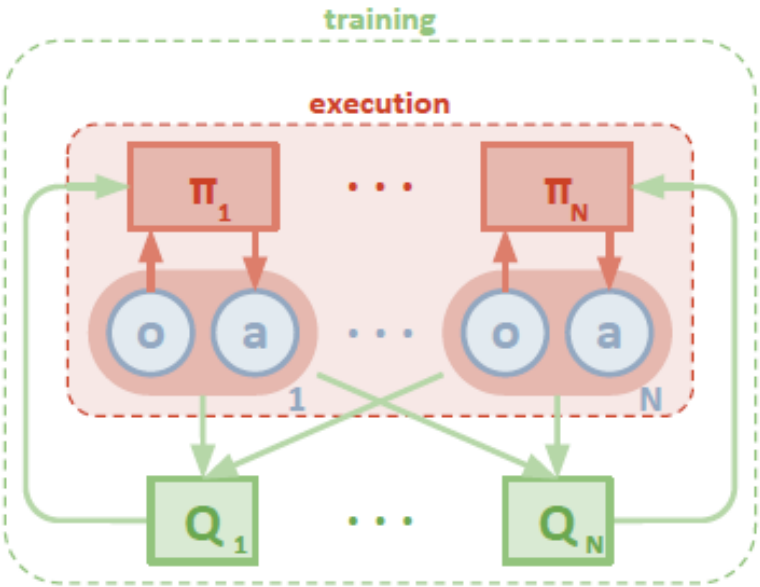


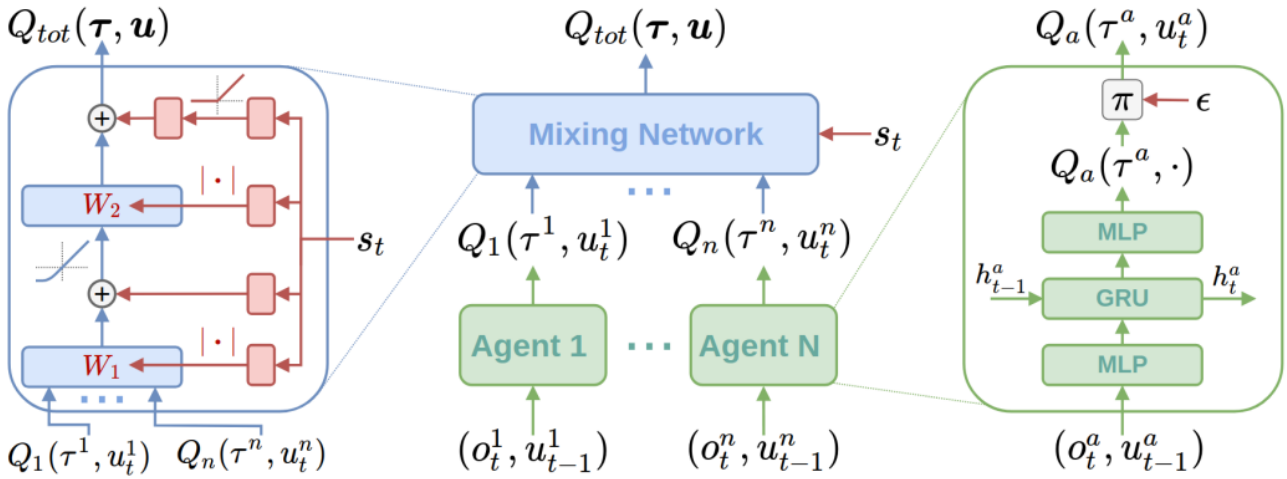
Closing the environment:

- Walls on each side of the frame
- Agents can't go beyond $[-1, 1]^2$
  $\Rightarrow$ Should be easier to solve

Model training

- MADDPG
- CMA-ES
- QMIX

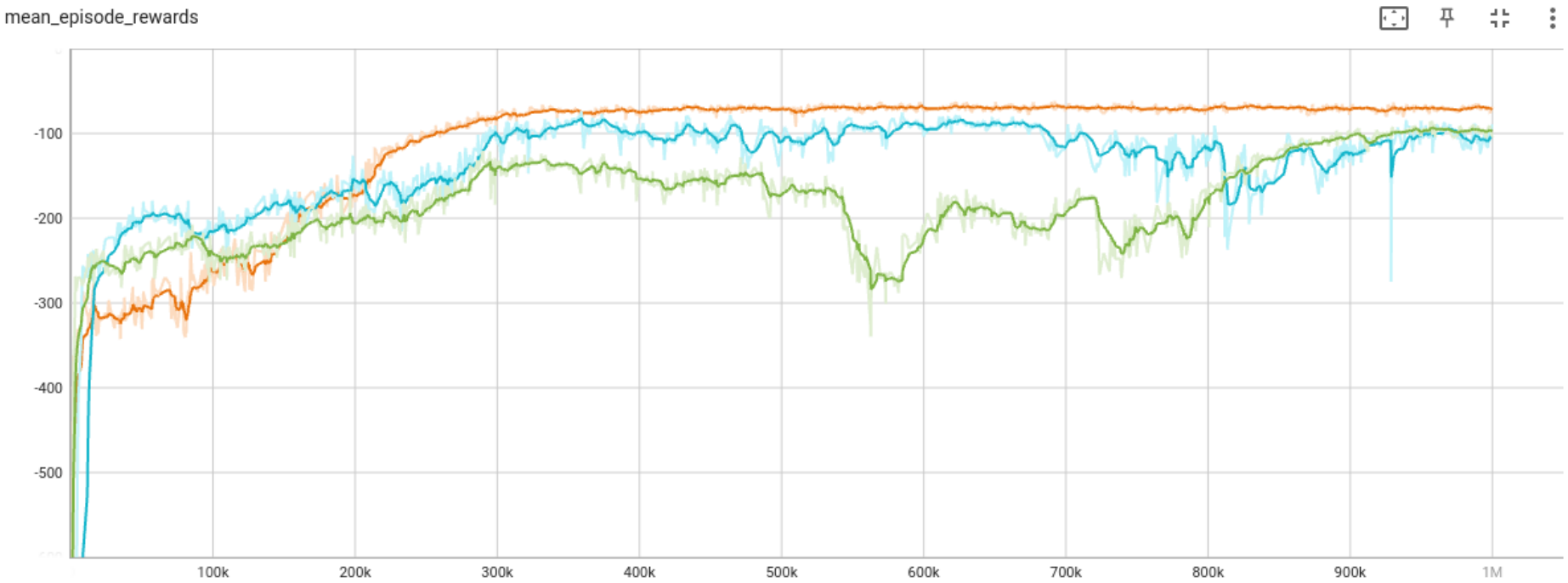




Opened environment:

Fully observable

- MADDPG



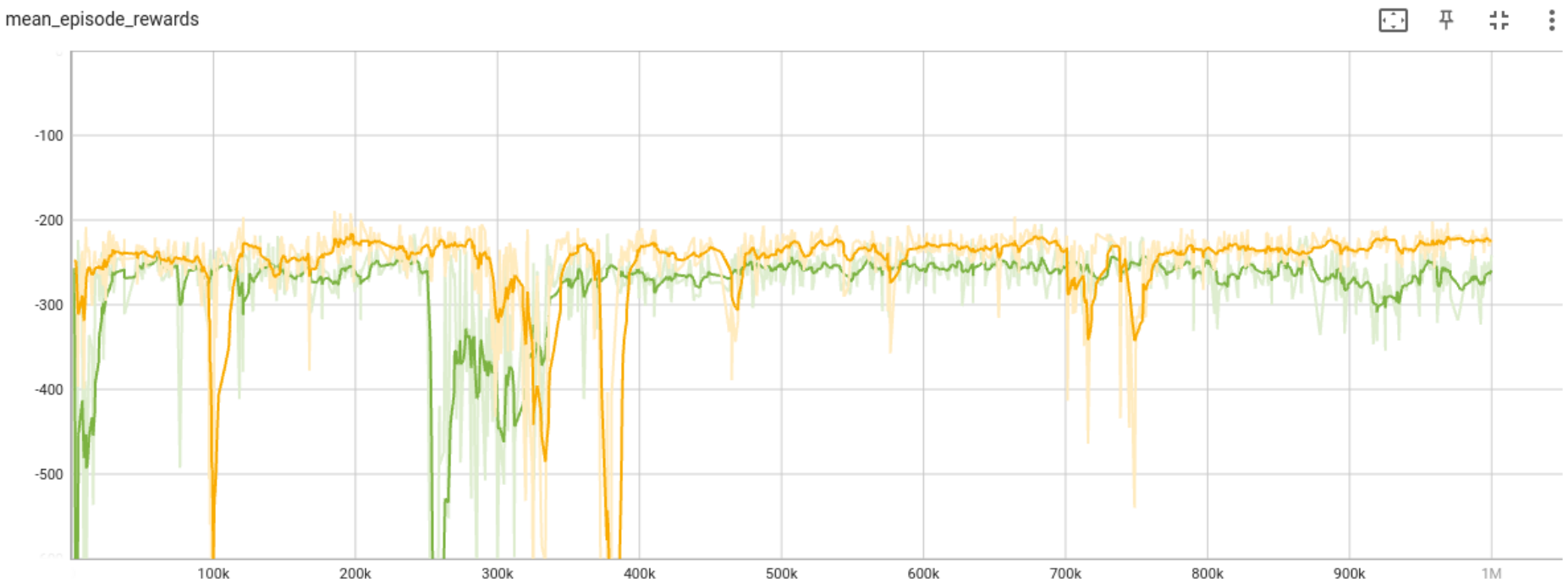
Best run: 2 agents, lr=0.005, discrete actions

Near optimal mean episode reward $\simeq$ -60

Opened environment:

Fully observable, 2 agents

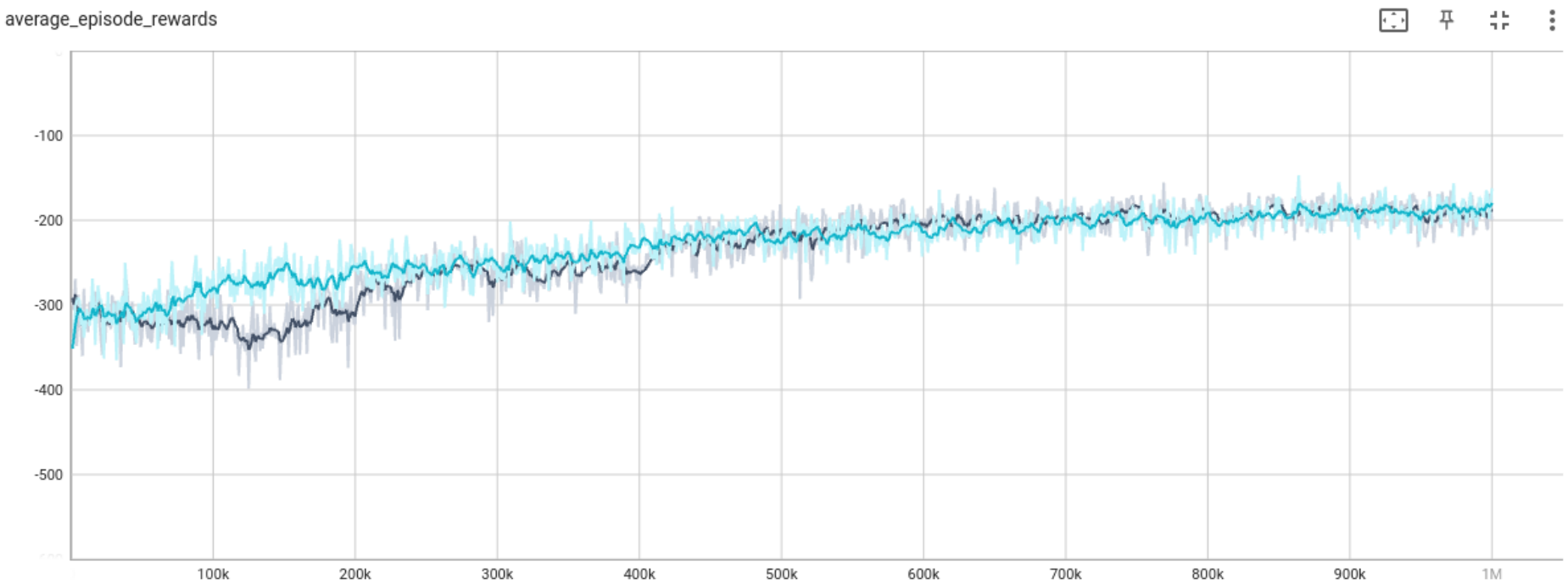
- CMA-ES (160 parameters, population = 19)



Opened environment:

Fully observable, 2 agents

- QMIX

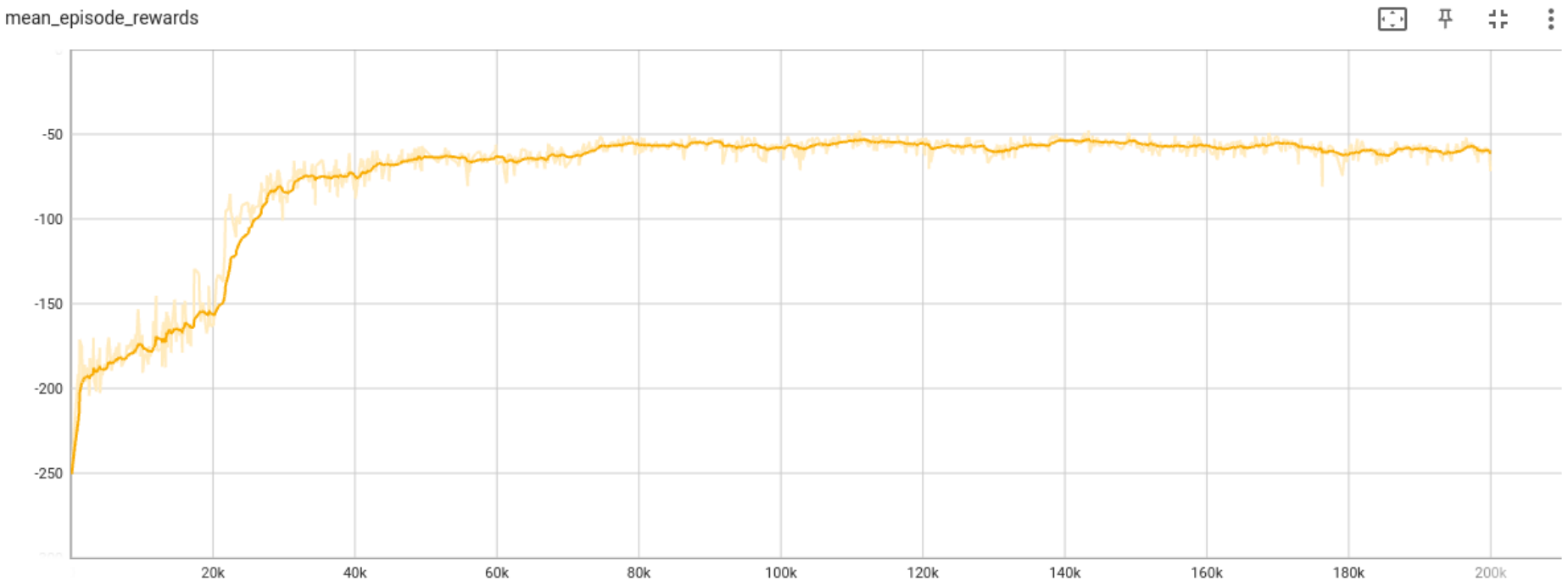


Some results

Closed environment:

Fully observable, 2 agents, no collision penalty

- MADDPG



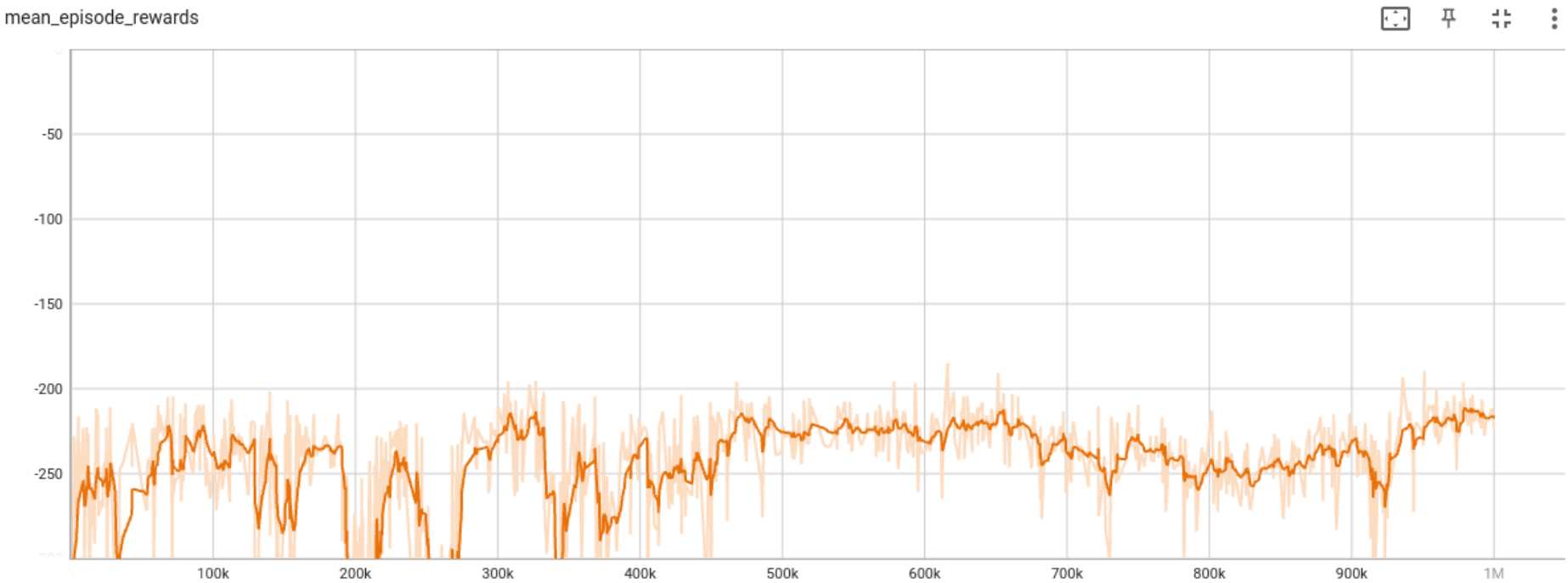
Best run so far: lr=0.005, continuous actions, train every 2500 steps

Some results

Closed environment:

Fully observable, 2 agents

- CMA-ES

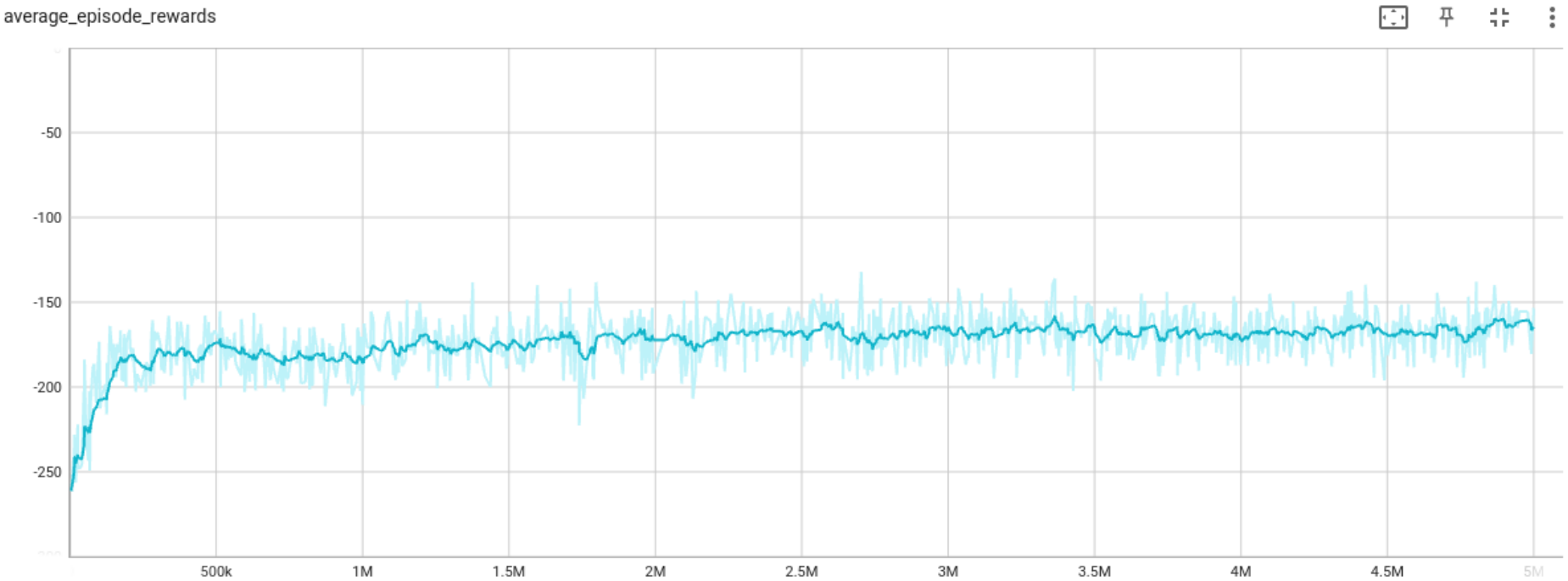


Some results

Closed environment:

Fully observable, 2 agents

- QMIX



Future works

Fixing issues:

- Fix CMA-ES
- Fix QMIX
- Find better reward function

Testing harder versions of the environment:

- Partial observability
- Heavier object
- More agents

Future works

Integrating communication

- How to learn emergent communication well ?
  - Sukhbaatar et al., Learning Multiagent Communication with Backpropagation, 2016
  - Das et al., TarMAC: Targeted Multi-Agent Communication, 2019
  - Learning to Ground Multi-Agent Communication with Autoencoders, 2021
- Use transformers ?
  - Parisotto et al., Stabilizing Transformers for Reinforcement Learning, 2020
  - Karch et al., Grounding Spatio-Temporal Language with Transformers, 2021
- Path to discrete communication
  - Language modeling
  - Language-Augmented Reinforcement Learning

Thank you!