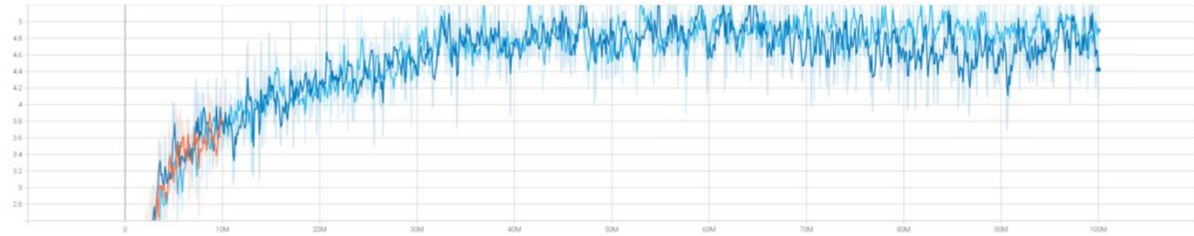
# Reinforcement Learning for Limit Order Book Dynamics

We ran some first experiments with the PPO algorithm and replicating the experiments from the Github.

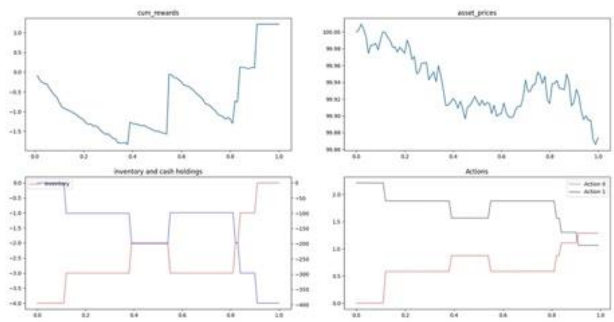
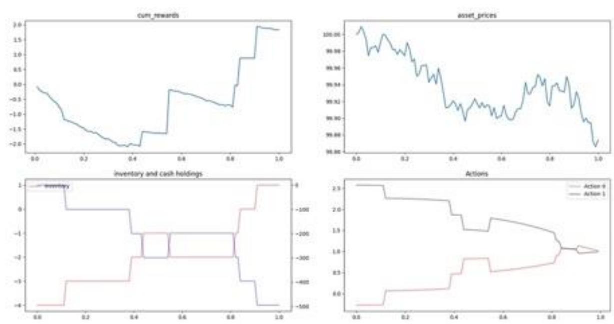
## Convergence

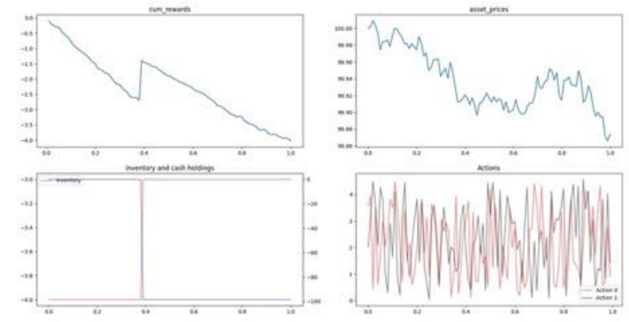
The figure below shows the convergence of the PPO agent in terms of average reward (different colors correspond to different runs, with different seeds and total numbers of iterations).



## Comparison to Optimal and Random Startegy

We can test the performance of the agent on a sample asset path. We compare it to the optimal strategy and a random strategy. For this simple setting, the optimal strategy is known, see (Cartea, 2015).

| | |
|---|---|
| RL |  |
| Optimal |  |

| | |
|---|---|
| Random |  |

# Bibliography

Cartea, Á. S. (2015). *Algorithmic and high-frequency trading.* Cambridge University Press.