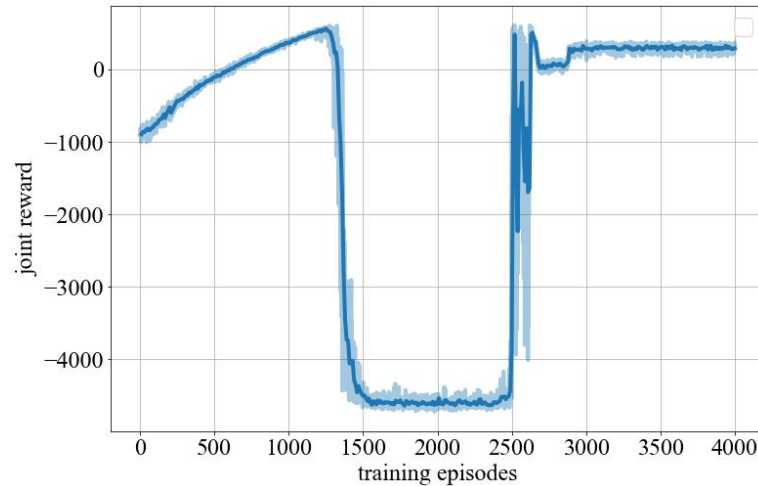# 1 Value-based MADRL

(1) The joint-reward curve increases gradually during the early episodes, but then it decreases precipitously. The following diagram shows.

Possible Reasons:

- The estimation of state-action values is not accurate.

Adjustment Methods:

- Increase the action exploration (to prevent action networks from converging to inaccurate values).

- Decrease the learning rate of action networks (to prevent action networks from converging to inaccurate values).

- Increase the update frequency of target networks (to prevent action networks from converging to inaccurate values).

- Increase networks' complexity (to enhance networks' fitting ability)