

## Q-MPC: Integrated Algorithm of Model-free Reinforcement Learning and Model Predictive Control

오태훈, 이종민<sup>†</sup>

서울대학교

(jongmin@snu.ac.kr<sup>†</sup>)

Optimizing the substrate feeding strategy of a bioreactor is one of the challenging tasks. In this work, we propose an integrated algorithm of model-free RL and model predictive control (MPC) that improves the initial control policy only with small number of data points. Similar to MPC, the proposed algorithm adopts the receding horizon principle and assigns the action-value function, which learns from the plant data, as the terminal cost. In this way, the adaptation of system dynamics can be achieved without modifying the model. On the other hand, the learning of the action-value function is performed with the conventional deep reinforcement learning with double Q-learning (DDQN) algorithm in off-policy fashion. The proposed method is the generalization of the DDQN and MPC. For the simulation study, the proposed method is applied to the penicillin product semi-batch bioprocess where the system dynamics are structurally different from the model used in MPC. For the comparison, DDQN, deep deterministic policy gradient (DDPG), and differential dynamic programming (DDP) algorithms are applied to the bioprocess with same conditions.