

Optimal Dynamic Regime for CO Oxidation Reaction discovered by Policy-Gradient Reinforcement Learning Algorithm

Mikhail Lifar, Alexander Guda, Andrei Tereshchenko, Aleksei Bulgakov
The Southern Federal University
Rostov-on-Don, Russia
{lifar, guda}@sfedu.ru

Abstract—Metal nanoparticles are widely used as heterogeneous catalysts to activate adsorbed molecules and reduce the energy barrier of the reaction. Reaction product yield depends on the interplay between elementary processes - adsorption, activation, reaction, desorption. These processes in turn depend on the inlet feed concentrations, temperature, and pressure. At stationary conditions, the active surface sites may be poisoned by reaction byproducts or blocked by thermodynamically adsorbed gaseous reagents. Thus, the yield of reaction products can significantly drop. On the contrary, the dynamic control accounts for the changes in the surface properties and adjusts reaction parameters accordingly. Therefore dynamic control may be more efficient than stationary control. In this work, a reinforcement learning algorithm has been applied to control the simulation of CO oxidation on a catalyst. The policy gradient algorithm learned to maximize the total yield of CO₂ through dynamic control of the CO and O₂ flows. Reaction models based on differential equations with and without deactivation of a catalyst were investigated and nonstationary solutions were found for the model with surface deactivation. The maximal product yield was achieved for periodic variations of the gas flows ensuring a balance between available adsorption sites and the concentration of activated intermediates. This methodology opens a perspective for the optimization of catalytic reactions under nonstationary conditions.