APPLICATION OF MODEL-BASED AND MODEL-FREE CONTROL TECHNIQUES FOR PROCESS CONTROL

NIKITA GUPTA

DEPARTMENT OF CHEMICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI

APRIL 2025

Application of Model-based and Model-free Control Techniques for Process Control

by

Nikita Gupta
Department of Chemical Engineering

Submitted
in fulfilment of the requirements of the degree of
Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI
April 2025

THESIS CERTIFICATE

This is to certify that the thesis titled **Application of Model-based and Model-free Control Techniques for Process Control**, submitted by **Nikita Gupta (2019CHZ8455)**, to the Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy**, is a bonafide record of the research work done by her under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Prof. Hariprasad Kodamana
Research Supervisor
Dept. of Chemical Engineering &
Yardi School of Artificial Intelligence, Indian Institute of Technology-Delhi, Hauzkhas, Delhi-110016

Date: April 23, 2025

ACKNOWLEDGEMENTS

First and foremost, with the utmost respect, I wish to express my sincere gratitude to my esteemed supervisor, Prof. Hariprasad Kodamana, for his invaluable guidance, unwavering support, and constant encouragement throughout the course of my research. His insightful feedback and profound expertise have greatly enriched this work and contributed significantly to my growth as a researcher. I feel truly honored and privileged to have had the opportunity to learn under his mentorship. I am especially thankful to him for making me understand the value of patience, humility, and wisdom in professional life-which really enhanced my academic endeavors and kept me grounded plus focused on my goals.

I am deeply thankful to all the revered collaborators who contributed to this work. My Special thanks to the senior professors namely Prof. Sharad Bhartiya, Prof. Manojkumar C Ramteke, and Prof. Harikumar Kandath for their invaluable help and guidance, which greatly enhanced the quality of this research work. I also extend my gratitude to Dr. Riju, Dr. Tanuja, Deepak, and Shikhar for their contributions and insightful discussions that enriched my understanding and enhanced the quality of this thesis.

I am deeply grateful to my student review committee (SRC) members Prof. Manojkumar C Ramteke, Prof. Jayati Sarkar, and Prof. Shubhendu Bhasin for their invaluable insights, encouragement, and academic guidance throughout the course of my Ph.D. research.

I would also like to acknowledge my fellow lab mates, Dr. Tanuja, Dr. Reena, Dr. Avan, Dr. Jyoti, Umang, Deepak, Ahtesham, Nasre, Vinayak, Arjun, Anirudh, Shobhita, Robin, and Amal whose camaraderie and collaborative spirit made the research journey enjoyable and rewarding. The stimulating discussions augmented by mutual support created an inspiring and motivating environment for me which I always cherish.

I extend my heartfelt thanks to Sreshta, Sukhi, Pialy, Tina, and Parna Di, for being a constant source of strength and positivity throughout this journey. Your encouragement, understanding, and the countless light-hearted moments we shared brought balance and joy, even during the most challenging times. I am profoundly grateful to Mayuna, Somya, Sheri,

and Sandy for their unwavering support and motivation, which inspired me to persevere and overcome every hurdle. Also, a special thanks to Prof. A.D.Rao Uncle and Aunty for their kindness, unconditional care, and constant encouragement, which added warmth and support to this journey. I am deeply grateful to each of you for making this journey truly meaningful and memorable.

A special note of gratitude goes to my father (Prof. K.D. Gupta) for his unwavering belief in me and his constant motivation, which has been my driving force, that has provided me with the confidence and determination to persevere and achieve my goals. I am also grateful to my mother (late Mamta Gupta) and brother (Nitish Gupta), for their love and encouragement.

Above all, I am profoundly grateful to the Almighty God for granting me the strength, patience, and guidance to overcome challenges and accomplish this work.

Nikita Gupta

ABSTRACT

Chemical process reactors are essential to industrial processes and demand precise control to ensure stability, efficiency, and product quality. Although widely utilized, traditional process control methods like proportional-integral-derivative (PID) controllers often lack the flexibility to handle the complexity and nonlinearity inherent in modern reactors. In this thesis work, an attempt has been made to integrate model-based and model-free control techniques for process control. Effective controllers for systems with well-understood dynamics are model-based controllers, such as Iterative Learning Control (ILC) and Model Predictive Control (MPC), which use comprehensive process models to forecast future behavior and optimize control actions over a predetermined horizon. Conversely, model-free methods such as Reinforcement Learning (RL) and its extensions, such as Multi-Agent RL (MARL) and Inverse Reinforcement Learning (IRL), do not depend on an explicit process model; rather, they learn optimal control strategies through interactions with the environment. MARL extends RL to scenarios with multiple agents, allowing for coordinated control in complex, interconnected systems, while IRL focuses on inferring the underlying reward structure from expert behavior. This work aims to improve control performance and robustness in chemical process systems by investigating the synergistic potential of different approaches.

In the first part of the thesis, sophisticated model-based controllers have been developed, such as ILC, and Model Predictive Control (MPC), which use mathematical models to forecast system behavior and maximize control operations. Explicit MPC (eMPC) extends the capabilities of MPC by reducing computational demand for real-time applications, making it more suitable for high-speed processes. The recurring nature of batch operations, incidentally, aids in optimizing the control policy for the subsequent batch based on knowledge from past batch runs. Consequently, batch-to-batch ILC is frequently implemented to control batch processes. This suggested algorithm integrates explicit MPC with ILC since it facilitates the rectification of disturbances both within and between batches.

The drawbacks of these model-based techniques originate from their need for precise system models, which could be challenging to establish in real-world scenarios. Therefore, model-free reinforcement learning (RL) techniques have become strong substitutes. Deep Deterministic Policy Gradient (DDPG), Twin-Delayed DDPG (TD3), and Proximal Policy Optimization (PPO) algorithms have demonstrated potential in optimizing control policies directly from data, without a comprehensive process model. In the subsequent part of the work, PPO and its multi-actor form are among the further advancements in policy optimization that have given robust solutions to the problems associated with continuous control in chemical reactors. Even with these developments, multi-component systems with intricate interactions provide challenges for single-agent RL techniques. By allowing several agents to develop coordinated control techniques, multi-agent RL (MARL) provides a framework to address these limitations. Therefore, a twin-agent RL framework has been proposed, that integrates deterministic and stochastic agents within a multi-agent setup, enhancing performance in both deterministic and stochastic environments.

Regardless of these developments, reward design remains a major difficulty for RL techniques, since sparse or poorly specified incentives can reduce learning efficiency and result in poor control policies. This constraint is mitigated by inverse reinforcement learning (IRL), which eliminates the requirement for explicit reward design by extracting reward functions from expert demonstrations. IRL optimizes policies more precisely and efficiently by deriving the underlying objectives from expert behavior, particularly in complicated chemical process control. This method is improved by further modifications such as Adversarial IRL (AIRL), which offers a strong framework for rewards design. To create more stable and broadly applicable reward functions, AIRL combines the advantages of adversarial networks with IRL, enabling the controller to respond more resiliently to fluctuating process dynamics. Model-based controllers, such as MPC, provide optimal trajectories that can be leveraged as expert trajectories in AIRL, facilitating the integration of model-based precision with the adaptability of model-free control techniques to handle complex and nonlinear systems.

The efficacies of these proposed control strategies have been proven by applying these techniques to bio-process like transesterification process, mAb production in bioreactors, and Propylene Glycol (PG) in continuous stirred tank reactors (CSTRs). This thesis aims to examine the application of these advanced control techniques in chemical process reactors

to achieve resilient, adaptive, and optimal control by bridging the gap between model-free and model-based approaches.

KEYWORDS: Model Predicitve Control (MPC); explicit MPC; Reinforcement Learning, Deep Deterministic Policy Gradient (DDPG); Twin Delayed DDPG (TD3), Proximal Policy Optimization (PPO); Inverse Reinforcement Learning (IRL); Adversarial IRL (AIRL)

सारांश

रासायनिक प्रक्रिया रिएक्टर औद्योगिक प्रक्रियाओं के लिए आवश्यक होते हैं और स्थिरता, दक्षता, और उत्पाद गुणवत्ता सुनिश्चित करने के लिए सटीक नियंत्रण की आवश्यकता होती है। हालांकि पारंपरिक प्रक्रिया नियंत्रण विधियाँ जैसे कि प्रोपोर्शनल-इंटीग्रल-डेरिवेटिव (PID) नियंत्रक व्यापक रूप से उपयोग की जाती हैं, वे अक्सर आधुनिक रिएक्टरों में निहित जटिलता और गैर-रेखीयता को संभालने के लिए लचीलापन की कमी होती हैं। इस थीसिस में प्रक्रिया नियंत्रण के लिए मॉडल-आधारित और मॉडल-रहित नियंत्रण तकनीकों को एकीकृत करने का प्रयास किया गया है। सुस्पष्ट डाइनामिक्स वाले सिस्टम के लिए प्रभावी कंट्रोलर्स मॉडल-आधारित कंट्रोलर्स होते हैं, जैसे कि इटरेटिव लर्निंग कंट्रोल (ILC) और मॉडल प्रेडिक्टिव कंट्रोल (MPC), जो भविष्य के व्यवहार का पूर्वानुमान करने और पूर्व निर्धारित सीमा पर कंट्रोल क्रियाओं को ऑप्टिमाइज़ करने के लिए व्यापक प्रोसेस मॉडल्स का उपयोग करते हैं। इसके विपरीत, मॉडल-फ्री विधियाँ जैसे कि रिइन्फोर्समेंट लर्निंग (RL) और इसके विस्तार, जैसे मल्टी-एजेंट RL (MARL) और इनवर्स रिइन्फोर्समेंट लर्निंग (IRL), एक स्पष्ट प्रक्रिया मॉडल पर निर्भर नहीं होतीं; बल्कि, ये पर्यावरण के साथ इंटरएक्शन के माध्यम से सर्वोत्तम नियंत्रण रणनीतियाँ सीखती हैं। MARL, RL को कई एजेंटों के साथ परिदृश्यों में विस्तारित करता है, जिससे जटिल, आपस में जुड़े हुए सिस्टमों में समन्वित नियंत्रण संभव होता है, जबकि IRL विशेषज्ञ के व्यवहार से अंतर्निहित पुरस्कार संरचना का अनुमान लगाने पर केंद्रित है।

थीसिस के पहले भाग में, परिष्कृत मॉडल-आधारित नियंत्रक विकसित किए गए हैं, जैसे कि ILC और मॉडल प्रेडिक्टिव कंट्रोल (MPC), जो सिस्टम के व्यवहार की पूर्वानुमान करने और नियंत्रण संचालन को अधिकतम करने के लिए गणितीय मॉडलों का उपयोग करते हैं। स्पष्ट MPC (eMPC) MPC की क्षमताओं का विस्तार करता है, जिससे वास्तविक समय अनुप्रयोगों के लिए संगणनात्मक मांग को कम किया जाता है, जिससे यह उच्च गति वाली प्रक्रियाओं के लिए अधिक उपयुक्त बनता है। बैच संचालन की आवर्ती प्रकृति, संयोगवश, पिछले बैच संचालन से प्राप्त ज्ञान के आधार पर अगले बैच के लिए नियंत्रण नीति को अनुकूलित करने में सहायता करती है।इसके परिणामस्वरूप, बैच-टू-बैच ILC (इंटरएक्टिव लर्निंग कंट्रोल) को अक्सर बैच प्रक्रियाओं को नियंत्रित करने के लिए लागू किया जाता है। यह सुझाया गया एल्गोरिदम स्पष्ट MPC को ILC के साथ एकीकृत करता है क्योंकि यह बैचों के भीतर और उनके बीच अवरोधों को सुधारने में सहायक होता है।

इन मॉडल-आधारित तकनीकों की कमियाँ उनके लिए सटीक प्रणाली मॉडलों की आवश्यकता से उत्पन्न होती हैं, जिन्हें वास्तविक दुनिया के परिदृश्यों में स्थापित करना चुनौतीपूर्ण हो सकता है। इसलिए, मॉडल-फ्री रिइंफोर्समेंट लर्निंग (RL) तकनीकें मजबूत विकल्प बन गई हैं। डीप डिटर्मिनिस्टिक पॉलिसी ग्रेडियंट (DDPG), ट्विन-डिलेयड DDPG (TD3), और प्रोक्षिमल पॉलिसी ऑप्टिमाइजेशन (PPO) एल्गोरिदम ने डेटा से सीधे नियंत्रण नीतियों को ऑप्टिमाइज़ करने में क्षमता दिखाई है, बिना किसी विस्तृत प्रक्रिया मॉडल के। कार्य के अगले हिस्से में, PPO और इसका मल्टी-एक्टर रूप नीति अनुकूलन में आगे की प्रगति के रूप में हैं, जिन्होंने रासायनिक रिएक्टरों में निरंतर नियंत्रण से जुड़ी समस्याओं के लिए मजबूत समाधान प्रदान किए हैं। इन विकासों के बावजूद, जटिल इंटरएक्शन वाली बहु-तत्त्व प्रणालियाँ सिंगल-एजेंट RL तकनीकों के लिए चुनौतियाँ उत्पन्न करती हैं। कई एजेंट्स को समन्वित नियंत्रण तकनीकों को विकसित करने की अनुमति देकर, मल्टी-एजेंट RL (MARL) इन सीमाओं को संबोधित करने के लिए एक रूपरेखा प्रदान करता है। इसलिए, एक ट्विन-एजेंट RL फ्रेमवर्क प्रस्तावित किया गया है, जो एक

मल्टी-एजेंट सेटअप में निर्धारक और यादृच्छिक एजेंट्स को एकीकृत करता है, जिससे निर्धारक और यादृच्छिक पर्यावरणों दोनों में प्रदर्शन को बेहतर बनाया जा सकता है।

इन विकासों के बावजूद, रिवॉर्ड डिज़ाइन RL तकनीकों के लिए एक प्रमुख कठिनाई बनी रहती है, क्योंकि विरल या खराब तरीके से निर्दिष्ट प्रोत्साहन लर्निंग की क्षमता को कम कर सकते हैं और खराब नियंत्रण नीतियों का परिणाम हो सकता है। यह प्रतिबंध इनवर्स रिइंफोर्समेंट लर्निंग (IRL) द्वारा कम किया जाता है, जो विशेषज्ञों के प्रदर्शन से रिवॉर्ड फंक्शंस निकालकर स्पष्ट रिवॉर्ड डिज़ाइन की आवश्यकता को समाप्त कर देता है। IRL विशेषज्ञ के व्यवहार से अंतर्निहित उद्देश्यों को निकालकर नीतियों को अधिक सटीक और प्रभावी तरीके से अनुकूलित करता है, विशेष रूप से जटिल रासायनिक प्रक्रिया नियंत्रण में। यह विधि आगे के सुधारों द्वारा सुधारित की जाती है, जैसे कि एडवर्सेरियल इन्बर्स रिइन्फोर्समेंट लर्निंग (AIRL), जो रिवॉर्ड डिज़ाइन के लिए एक मजबूत ढांचा प्रदान करता है। अधिक स्थिर और व्यापक रूप से लागू होने योग्य इनाम कार्यों को बनाने के लिए, AIRL विरोधी नेटवर्क्स के लाभों को IRL के साथ जोड़ता है, जिससे नियंत्रक को परिवर्तनीय प्रक्रिया डायनेमिक्स के प्रति अधिक लचीलापन से प्रतिक्रिया देने की अनुमति मिलती है। मॉडल-आधारित कंट्रोलर्स, जैसे कि MPC, ऑप्टिमल ट्रेजेक्टोरीज़ प्रदान करते हैं जिन्हें AIRL में एक्सपर्ट ट्रेजेक्टोरीज़ के रूप में उपयोग किया जा सकता है, जिससे मॉडल-आधारित सटीकता को मॉडल-फ्री कंट्रोल तकनीकों की अनुकूलता के साथ एकीकृत करना संभव होता है, जो जटिल और गैर-रैखिक प्रणालियों को संभालने में मदद करता है।

इन प्रस्तावित नियंत्रण रणनीतियों की प्रभावशीलता को जैव-प्रक्रियाओं जैसे ट्रांसएस्टरिफिकेशन प्रक्रिया, बायोरिएक्टर्स में mAb उत्पादन, और निरंतर मिश्रण टैंक रिएक्टर्स (CSTRs) में प्रोपलीन ग्लाइकोल (PG) पर इन तकनीकों को लागू करके सिद्ध किया गया है। यह थीसिस प्रबंध रासायनिक प्रक्रिया रिएक्टरों में इन उन्नत नियंत्रण तकनीकों के अनुप्रयोग की जांच करने का उद्देश्य रखता है, ताकि मॉडल-फ्री और मॉडल-आधारित दृष्टिकोणों के बीच अंतर को पाटते हुए लचीला, अनुकूलनशील और सर्वोत्तम नियंत्रण प्राप्त किया जा सके।

कीवर्ड्स: मॉडल प्रेडिक्टिव कंट्रोल (MPC); एक्सप्लिसिट MPC; रिइन्फोर्समेंट लर्निंग, डीप डिटर्मिनिस्टिक पॉलिसी ग्रेडियंट (DDPG); ट्विन डिलेयड DDPG (TD3), प्रोक्षिमल पॉलिसी ऑप्टिमाइजेशन (PPO); इनवर्स रिइन्फोर्समेंट लर्निंग (IRL); एडवर्सेरियल IRL (AIRL)

Contents

List of Figures

List of Tables

ABBREVIATIONS

MPC	Model Predicitve Control
ILC	Iterative Learning Control
RL	Reinforcement Learning
DDPG	Deep Deterministic Policy Gradient
TD3	Twin-Delayed Deep Deterministic Policy Gradient
PPO	Proximal Policy optimization
MARL	Multi-agent Reinforcement Learning
GAN	Generative Adversarial Networks
IRL	Inverse Reinforcement Learning
AIRL	Adversarial Inverse Reinforcement Learning
RMSE	Root Mean Square Error
IAE	Integral Absolute Error