Simultaneous Process Design and Control Optimization using Reinforcement Learning

Steven Sachio
Department of Chemical Engineering
Imperial College London
London, SW7 2AZ
steven.sachio19@imperial.ac.uk

Antonio Ehecatl del-Rio Chanona
Department of Chemical Engineering
Imperial College London
London, SW7 2AZ
a.del-rio-chanona@imperial.ac.uk

Panagiotis Petsagkourakis
Department of Chemical Engineering
University College London
London, WC1E 7JE
p.petsagkourakis@ucl.ac.uk

Abstract

With the ever-increasing numbers in population and quality in healthcare, it is inevitable for the demand of energy and natural resources to rise. Therefore, it is important to design highly efficient and sustainable chemical processes in the pursuit of sustainability. The performance of a chemical plant is highly affected by its design and control. A design cannot be evaluated without its controls and vice versa. To optimally address design and control simultaneously, one must formulate a bi-level mixed-integer nonlinear program with a dynamic optimization problem as the inner problem; this, is intractable. However, by computing an optimal policy using reinforcement learning, a controller with close-form expression can be found and embedded into the mathematical program. In this work, an approach using a policy gradient method along with mathematical programming to solve the problem simultaneously is proposed. The approach was tested in two case studies and the performance of the controller was evaluated. It was shown that the proposed approach outperforms current state-of-the-art control strategies. This opens a whole new range of possibilities to address the simultaneous design and control of engineering systems.

1 Introduction

As the human population grows, the demand of consumables and natural resources increases. Therefore, it is important to design sustainable and efficient processes. The performance of a chemical plant is substantially affected by its design and its ability to maintain the optimal operating conditions under operational uncertainty (Diangelakis et al., 2017). A design cannot be evaluated without the control and vice versa. Hence, it is essential to formulate simultaneous process design and control approaches to create energy and cost-efficient processes in the pursuit of sustainability.

The PSE (Process Systems Engineering) community has been challenging these problems since decades ago (Burnak et al., 2019). Several frameworks have been proposed in the past to tackle simultaneous process design and control problems such as formulation and solution of a bi-level mixed-integer dynamic optimization (MIDO) problem by transformation to mixed-integer nonlinear programming problem (Flores-Tlacuahuac and Biegler, 2007), model-based flow sheet and process group contribution method (Alvarado-Morales et al., 2010) and control structure selection and

Workshop on machine learning for engineering modeling, simulation and design @ NeurIPS 2020

design (Skogestad and Morari, 1987). The controller used in previous studies were mainly the PI and PID controllers as they are simple and robust. However, it comes with its shortcomings such as not being able to handle slow disturbances (Sung and Lee, 1996). Multi-parametric methods which depends on the solution of the bi-level optimization problem has been used previously to address these problems (Sakizlis et al., 2004; Diangelakis et al., 2017), however, these rely in linearizations which do not represent well nonlinear dynamics.

To solve this intractable bi-level optimization problem, it is proposed to take advantage of the closed form (explicit) nature of reinforcement learning controllers. Although the technology is not mature yet, reinforcement learning (RL) controllers (Sutton and Barto, 2018), has been proven to be able to handle disturbances, highly non-linear and complex processes (Petsagkourakis, Sandoval, Bradford, Zhang and del Rio-Chanona, 2020).

2 Method and Integration

2.1 Problem Statement

The approach proposed here builds upon the work done by Diangelakis et al. (2017). Given a simultaneous process design and control problem, first a bi-level optimization problem is formulated and then split into two, the design problem and the control problem. The design problem is formulated as a mixed integer dynamic optimization (MIDO) problem as shown in Equation 1

$$
\begin{aligned}
\min_{\boldsymbol{Y},\boldsymbol{des}} \; & J_{SDC} = \int_0^{T_F} \mathcal{P}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{Y}, \boldsymbol{d}, \boldsymbol{des}) d\tau \\
s.t. \; & \frac{dx}{d\tau} = f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{Y}, \boldsymbol{d}, \boldsymbol{des}) \\
& \boldsymbol{y}_{min} \leq \boldsymbol{y} = g(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{Y}, \boldsymbol{d}, \boldsymbol{des}) \leq \boldsymbol{y}_{max} \\
& \boldsymbol{u}_t = \pi_\theta(\boldsymbol{x}_t^{obs}, \boldsymbol{x}_{t-1}^{obs}, \boldsymbol{y}_t^{obs}, \boldsymbol{y}_{t-1}^{obs}, \boldsymbol{des}^{obs}) \; \forall t \in \{1, ...n_T - 1\} \\
& \boldsymbol{Y} \in \{0, 1\}, \; \boldsymbol{x}_t^{obs} \in \boldsymbol{x}_t, \; \boldsymbol{y}_t^{obs} \in \boldsymbol{y}_t, \; \boldsymbol{des}^{obs} \in \boldsymbol{des} \\
& [\boldsymbol{x}_{min}, \boldsymbol{d}_{min}] \leq [\boldsymbol{x}_\tau, \boldsymbol{d}_\tau] \leq [\boldsymbol{x}_{max}, \boldsymbol{d}_{max}] \; \forall \tau \in [0, T_F] \\
& \boldsymbol{des}_{min} \leq \boldsymbol{des} \leq \boldsymbol{des}_{max}
\end{aligned} \tag{1}
$$

where τ is the continuous variable time $\tau \in [0, T_F]$, T_F is the time horizon discretized in n_T time steps of size $\Delta\tau = T_F/n_T$. The subscript t is the time step number, it represents the value of a variable at time $\tau = t\Delta\tau$ where $t \in \{0,\ 1, ...,\ n_T - 1,\ n_T\}$. While $\boldsymbol{x}$ is the vector of states of the system, $\boldsymbol{y}$ is the vector of systems outputs, $\boldsymbol{u}$ is the control action represented discretely as a piece-wise constant variable, $\boldsymbol{d}$ is the disturbance vector, $\boldsymbol{des}$ is the design variable vector including the steady-state manipulated variables, $\mathcal{P}$ is the integral function of the objective function J_{SDC}, f is the differential equation describing the process dynamics and g is the constraints. The vectors with superscript obs are the values of the vector which is observable by the controller since it may not necessarily need to or able to see the whole state of the process. π_θ represents the controller, which can be a linearized one or an optimization problem like in Equation 2. This means that the problem turns into bi-level as there is an outer optimization for the design and an inner optimization for the control.

So, the goal of this problem is to find combinations of the design and binary variables (for example, reactor type, process layout, etc.), given an optimized controller in the form of an ANN (artificial neural network) to minimize an objective function J_{SDC} over a time horizon T_F discretized in n_T number of time steps of size $\Delta\tau = T_F/n_T$ while satisfying the constraints f and g. To get the optimized controller, an optimal control problem (OCP) is formulated in a way that the controller would also take in the observable design variable as its input. The formulation is shown in Equation 2.

$$
\begin{aligned}
\max_{\pi_\theta} \quad & \mathbb{E}[J_{OCP}(\boldsymbol{x}_t^{obs}, \boldsymbol{y}_t^{obs}, \boldsymbol{u}_t, \boldsymbol{des})] \\
s.t. \quad & \boldsymbol{x}(0) = \boldsymbol{x}_0 \\
& \boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{d}_t, \boldsymbol{des}) \; \forall t \in \{1, ..., n_T - 1\} \\
& \boldsymbol{u}_t = \pi(\theta, \boldsymbol{x}_t^{obs}, \boldsymbol{x}_{t-1}^{obs}, \boldsymbol{y}_t^{obs}, \boldsymbol{y}_{t-1}^{obs}, \boldsymbol{des}^{obs}) \\
& \boldsymbol{x}_t^{obs} \in \boldsymbol{x}_t, \; \boldsymbol{y}_t^{obs} \in \boldsymbol{y}_t, \; \boldsymbol{des}^{obs} \in \boldsymbol{des} \\
& [\boldsymbol{x}_{min}, \boldsymbol{u}_{min}] \leq [\boldsymbol{x}_t, \boldsymbol{u}_t] \leq [\boldsymbol{x}_{max}, \boldsymbol{u}_{max}] \; \forall t \in \{1, ..., n_T - 1\} \\
& \boldsymbol{y}_{min} \leq \boldsymbol{y}_t \leq \boldsymbol{y}_{max} \; \forall t \in \{1, ..., n_T - 1\}
\end{aligned}
\tag{2}
$$

where all of the notations have the same meaning as before except for the cost function J_{OCP} just to indicate that the cost function is not the same as the one in the MIDO formulation. The goal of this problem is to find the optimum parameters of the policy for which it maximizes the expectation of the objective function.

Using traditional approach would deem this bi-level optimization problem containing a MINLP on the outside and an OCP in the inner loop intractable, however, by using RL method this problem can be addressed. The bi-level optimization problem can be separated and formulated as presented in Equation 2. The OCP problem formulated in Equation 2 shows that the resulting controller also takes in the observable design variables to calculate the control output, this is why it can operate for a wide range of design. Furthermore, the final form of a policy gradient controller is an explicit function, this is why it can be used easily in the MIDO problem formulated in Equation 1.

2.2 Full Approach

Given a simultaneous process design and optimization problem, the approach shown in Algorithm 1 can be used. For training the controller, any policy optimization algorithm can be used (e.g. TRPO (Schulman et al., 2017), PPO (Schulman et al., 2017), DDPG (Lillicrap et al., 2019)) but for simplicity the Reinforce algorithm with baseline (Sutton and Barto, 2018) was used in this work.

Algorithm 1 Full approach for simultaneous design and control optimization using RL.

Input: Simultaneous process design and control problem.
Output: Optimal design and controller.

1. **State design-control simultaneous optimization**: Formulate the design problem as a mixed-integer dynamic optimization (MIDO) problem as presented in Equation 1.
2. **Optimal Control problem**: Pose the optimal control problem (OCP) based on the design problem as in Equation 2.
3. **OCP as RL**: Formulate the OCP to be solved via a RL policy optimization framework.
   (a) **Pre-training**: pre-train from an initial policy either via simulation or plant data, this can be done via supervised or apprenticeship learning, amongst others.
   (b) **Policy learning via policy gradients**:Begin full-training using a policy optimization algorithm.
4. **Embedded bi-level program**: Embed the final policy obtained from **step 3** into the design (MIDO) problem and solve using an appropriate solver.

It is important to highlight, that without a policy based method, the OCP is an optimization problem, hence the integrated formulation in Equation 1 is an intractable bi-level optimization problem.

3 Results and Discussions

There are two case studies considered in this work, a more complete description can be found in (Diangelakis et al., 2017).

3.1 Tank Design

In this simpler case study, it is desired to design a simple tank with continuous inlet and outlet flows. There are no reactions involved, the inlet flow is in the form of a sinusoidal signal and it has a nominal inlet flow of F_{nom} with a maximum deviation of F_{dev}. The outlet flow is the manipulated variable of the PG (policy gradient) controller. A setpoint which depends on the values of F_{nom} and F_{dev} is inputted into the controller.

The goal of the design problem is to determine the maximum deviation for which the controller is able to maintain the desired setpoint. While the goal of the OCP is setpoint tracking. The design problem was formulated as a MIDO problem presented in Equation 3 and the OCP in Equation 4.τ is the continuous time variable, $V(t)$ is the volume of liquid in the tank, $F_{in}(t)$ and $F_{out}(t)$ are the inlet and outlet flows in $\mathrm{m}^3\,\mathrm{s}^{-1}$, respectively, a is the valve position (0 is closed, 1 is open), and $freq$ is the frequency of the sinusoidal wave disturbance in s^{-1}. V_{SP} is the volume set point, V_{tank} is the volume of the tank, err_{π_θ} is the integral set point error, ε is the maximum allowable error and is equal to 1 %, a_t is the control action output from the policy applied to the system.

$$
\begin{aligned}
\max_{V_{tank},F_{dev},F_{nom},V(0)} \quad & J_{SDC} = \int_0^1 F_{dev} d\tau \\
s.t.\ & \frac{dV(\tau)}{d\tau} = F_{in}(\tau) - F_{out}(\tau) \\
& F_{out}(\tau) = \alpha_t V(\tau) \\
& F_{in}(\tau) = F_{nom} + F_{dev}\sin(\tau/freq) \\
& freq = \frac{1}{2\pi} \\
& V_{SP} = F_{nom} + F_{dev} \leq V_{tank} \\
& err_{\pi_\theta} = \int_0^1 \frac{||V(\tau) - V_{SP}||}{V_{SP}} d\tau \\
& \text{End-point constraints:} \\
& (1-\varepsilon/100)V(0) \leq V(T_F) \leq (1+\varepsilon/100)V(0) \\
& err_{\pi_\theta} \leq \varepsilon/100 \\
& \text{Interior-point constraints:} \\
& a_t = \pi_\theta(F_{in,t}, V_{SP}, V_t, V_{t-1})\ \forall t \in \{0,...,n_T - 1\}
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
\max_{\pi_\theta} \quad & J_{OCP} = -10\sum_{t=0}^{n_T}(V_t - V_{SP})^2 \\
s.t.\ & \frac{dV(\tau)}{d\tau} = F_{in}(\tau) - F_{out}(\tau) \\
& F_{out}(\tau) = \alpha_t V(\tau) \\
& F_{in}(\tau) = F_{nom} + F_{dev}\sin(\tau/freq) \\
& freq = \frac{1}{2\pi} \\
& V_{SP} = F_{nom} + F_{dev} \leq V_{tank} \\
& a_t = \pi_\theta(F_{in,t}, V_{SP}, V_t, V_{t-1})\ \forall t \in \{0,...,n_T - 1\}
\end{aligned} \tag{4}
$$

In this case study, a proportional derivative (PD) controller were also used to compare its performance with the policy gradient (PG) controller. The problem was solved using the approach proposed here and a solution to the bi-level optimization problem (intractable using traditional approaches) was found. The final results showed that the PG controller were able to handle deviation values of up to 3.84 $\mathrm{m}^3\,\mathrm{s}^{-1}$. The control performances are plotted in Figure 1.

Figure 1 shows the performance of the controllers in the tank case study at the MIDO solution averaged over 1000 simulations with 2 % measurement noise. It can be seen clearly that the PG controller outperforms the PD controller. The PD control action shows a large variance on its control action indicating very chaotic response. The mean setpoint error of the PD controller is two times larger than the PG controller mean error. Furthermore, the PG controller were able to perform better than the mpMPC used by Diangelakis et al. (2017) as the PG controller managed to handle a higher maximum deviation even with the presence of measurement noise.

3.2 Continuous Stirred Tank Reactor (CSTR) Design

The second case study considers the design of a CSTR (continuous stirred tank reactor). The reaction in the CSTR is a first-order endothermic reaction and a heating jacket is equipped. The temperature of the jacket is controlled by the PG controller. The feed flowrate and concentrations are disturbed by a sinusoidal wave similar to the tank case study. The controller is not able to see the value of inlet feed flowrate, however, in this case study, there is an option to have a settling tank before the CSTR to normalize the feed mass flowrate to its nominal value.

The goal of the design problem is to minimize a cost related objective function while satisfying the temperature constraint. While the goal of the OCP is to minimize the concentration of A by manipulating the heating jacket temperature while keeping the CSTR temperature below 450 K. The dynamic system can be found in (Diangelakis et al., 2017). The design problem formulation is shown in Equation 5 and the OCP in Equation 6.

$$
\begin{aligned}
&\min_{V,m_{dev},m_{nom},C_{A,dev},C_{A,nom},Y_S} J_{SDC} = Cost_{Total} \\
s.t.\ &\frac{dC_A}{d\tau} = \frac{m}{\rho V}(C_{A_0} - C_A) - k_0 C_A e^{-\frac{E_A}{RT}} \\
&\frac{dT}{d\tau} = \frac{mC_P(T_0 - T) + V\Delta H_{rxn} k_0 C_A e^{-\frac{E_A}{RT}} + UA(T_H - T)}{V\rho C_P} \\
&C_{A_0} = C_{A_{0,nom}} + C_{A_{0,nom}} + C_{A_{0,dev}} \sin(t/freq) \\
&m = m_{nom} + m_{dev}(1 - Y_{S,t}) \sin(t/freq) \\
&freq = 100/(2\pi) \\
&T_{H,t} = \pi_\theta(C_{A,t}, C_{A,t-1}, T_t, V, C_{A_0,t}, C_A^{SP})\ \forall t \in 0,...,n_T - 1 \\
&err_{\pi_\theta} = \int_0^{T_F} C_A - C_A^{SP} d\tau \\
&C_A^{SP} = 0 \\
&\text{Objective function:} \\
&Cost_{Total} = Cost_{Equipment} + Cost_{Operational} \\
&Cost_{Equipment} = 10((V - 750)/\pi) + 1000 + 400Y_{S,f} \\
&\frac{dCost_{Operational}}{d\tau} = -m(C_{A_0} - C_A) + 4Y_S \\
&\text{Endpoint constraints:} \\
&err_{\pi_0} \leq 100,\ C_{A_{0,dev}} \leq C_{A_{0,nom}},\ m_{dev} \leq m_{nom} \\
&\text{Interior point constraints: } \forall t \in \{0,...,n_T\} \\
&Y_{S,t} - Y_{S,f} \leq 0,\ Y_{S,t} \in \{0,1\},\ Y_{S,f} \in \{0,1\} \\
&T_t \leq 450
\end{aligned}
\tag{5}
$$

$$
\begin{aligned}
&\max_{\pi_\theta} J_{OCP} = -\sum_{t=0}^{n_T}(C_{A,t} - C_A^{SP}) - 100\sum_{t=0}^{n_T}\max(0, T_t - 450) \\
s.t.\ &\frac{dC_A}{d\tau} = \frac{m}{\rho V}(C_{A_0} - C_A) - k_0 C_A e^{-\frac{E_A}{RT}} \\
&\frac{dT}{d\tau} = \frac{mC_P(T_0 - T) + V\Delta H_{rxn} k_0 C_A e^{-\frac{E_A}{RT}} + UA(T_H - T)}{V\rho C_P} \\
&C_{A_0} = C_{A_{0,nom}} + C_{A_{0,nom}} + C_{A_{0,dev}} \sin(t/freq) \\
&m = m_{nom} + m_{dev} \sin(t/freq) \\
&freq = 100/(2\pi) \\
&T_{H,t} = \pi_\theta(C_{A,t}, C_{A,t-1}, T_t, V, C_{A_0,t}, C_A^{SP})\ \forall t \in 0,...,n_T - 1 \\
&C_A^{SP} = 0,\ T_t \leq 450
\end{aligned}
\tag{6}
$$

where C_A and C_{A_0} are the concentrations of A in the reactor and in the feed, respectively, m is the mass flowrate into and out of the reactor, ρ is the density of the liquid, V is the volume of the reactor, k_0 is the pre-exponential factor, E_A is the activation energy, R is the ideal gas constant, ΔH_{rxn} is the heat of reaction, T is the temperature inside the reactor, T_H is the heating utility temperature, C_P is the heat capacity, UA is the overall heat transfer coefficient multiplied by the heat transfer area, and $Y_{S,t}$ is a piece-wise constant binary variable which determines the utilization of the settling tank while $Y_{S,f}$ is a binary variable which denote the existence of the settling tank.

A PID controller was not implemented in this case study as it would need complex modifications for it to work with constraints. Again, the approach were able to solve this otherwise intractable problem for traditional approaches. The solution has final cost of 27,539 and the control performance is presented in Figure 1. It shows that the settling tank was only needed for a brief amount of time at the start of the process. At the start, the concentration of A in the CSTR is higher than the feed concentration, making it hard to satisfy the temperature constraint at initial time, therefore the settling tank was used. Overall, the controller is performing very well. It was able to minimize the concentration of A with very good performance (very close to zero almost all of the time) and the temperature constraint is satisfied at all times. This is also true for a wide range of design variables.

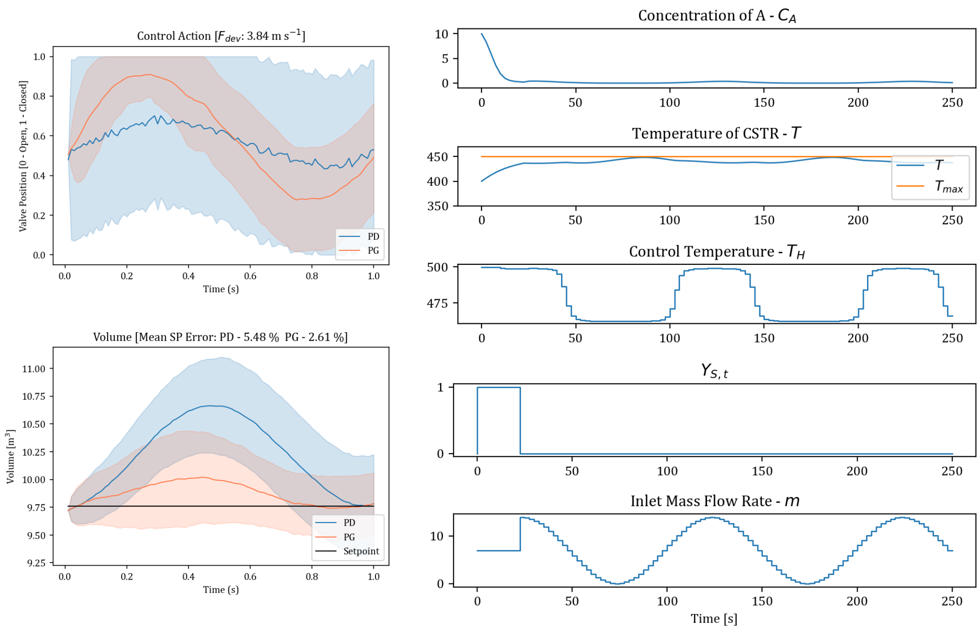

Figure 1: Case study results, LEFT: Tank, RIGHT: CSTR.

4 Conclusions

In this work we proposed the use of RL to address a long standing challenge for the process design and control community; the simultaneous design and control of sustainable chemical processes. It was shown that using RL it is possible to address the otherwise intractable nonlinear mixed integer dynamic bi-level optimization problem. Through case studies, we show that the PG-based controller was able to perform well in a wide range of design problems and design variables enabling the optimization of the process. Policy gradient-based controllers can handle measurement noise and stochastic systems naturally and the final form is an explicit function, which can be directly embedded into optimization problems. For future work, we will add the handling of constraints with high probability such as in (Petsagkourakis, Sandoval, Bradford, Galvanin, Zhang and del Rio-Chanona, 2020), and a further hyperparameter tuning scheme (e.g Snoek et al. (2012) Bayesian optimization or other expensive black box optimization techniques). We aim to address also harder and larger problems in the near future.

Broader Impact

With the ever-increasing numbers in population and quality in healthcare, it is inevitable for the demand for energy and consumer products to increase. Increased demand for energy, in turn, increases the consumption of fossil fuels which raises the strictness in environmental regulations due to the negative impacts associated with fossil fuels. Hence, it is important to design energy and cost-efficient processes in the pursuit of sustainability. The performance of a chemical plant is substantially affected by its design and its ability to maintain the optimal operating conditions under operational uncertainty (Diangelakis et al., 2017). A design cannot be evaluated without the control and vice versa. Hence, it is essential to formulate simultaneous process design and control approaches to create efficient processes in the pursuit of sustainability. Compared to the traditional approaches, the RL approach was able to solve bi-level optimization problems and show great control performance. Furthermore, this method can be applied to both renewable and non-renewable processes and from a methods point of view it could also be applied to other bi-level optimization problem.

Acknowledgments and Disclosure of Funding

This project has received funding from the EPSRC projects (EP/R032807/1) and (EP/P016650/1).

References

Alvarado-Morales, M., Hamid, M. K. A., Sin, G., Gernaey, K. V., Woodley, J. M. and Gani, R. (2010), 'A model-based methodology for simultaneous design and control of a bioethanol production process', *Computers & Chemical Engineering* **34**(12), 2043 – 2061. 10th International Symposium on Process Systems Engineering, Salvador, Bahia, Brasil, 16-20 August 2009.
URL: *http://www.sciencedirect.com/science/article/pii/S0098135410002516*

Burnak, B., Diangelakis, N. A. and Pistikopoulos, E. N. (2019), 'Towards the grand unification of process design, scheduling, and control—utopia or reality?', *Processes* **7**(7).
URL: *https://www.mdpi.com/2227-9717/7/7/461*

Diangelakis, N. A., Burnak, B., Katz, J. and Pistikopoulos, E. N. (2017), 'Process design and control optimization: A simultaneous approach by multi-parametric programming', *AIChE Journal* **63**(11), 4827–4846.
URL: *https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.15825*

Flores-Tlacuahuac, A. and Biegler, L. T. (2007), 'Simultaneous mixed-integer dynamic optimization for integrated design and control', *Computers & Chemical Engineering* **31**(5), 588 – 600. ESCAPE-15.
URL: *http://www.sciencedirect.com/science/article/pii/S0098135406002316*

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D. (2019), 'Continuous control with deep reinforcement learning'.

Petsagkourakis, P., Sandoval, I., Bradford, E., Zhang, D. and del Rio-Chanona, E. (2020), 'Reinforcement learning for batch bioprocess optimization', *Computers & Chemical Engineering* **133**, 106649.
URL: *http://www.sciencedirect.com/science/article/pii/S0098135419304168*

Petsagkourakis, P., Sandoval, I. O., Bradford, E., Galvanin, F., Zhang, D. and del Rio-Chanona, E. A. (2020), 'Chance constrained policy optimization for process control and optimization'.

Sakizlis, V., Perkins, J. D. and Pistikopoulos, E. N. (2004), 'Recent advances in optimization-based simultaneous process and control design', *Computers & Chemical Engineering* **28**(10), 2069 – 2086. Special Issue for Professor Arthur W. Westerberg.
URL: *http://www.sciencedirect.com/science/article/pii/S0098135404001115*

Schulman, J., Levine, S., Moritz, P., Jordan, M. I. and Abbeel, P. (2017), 'Trust region policy optimization'.

Skogestad, S. and Morari, M. (1987), 'Control configuration selection for distillation columns', *AIChE Journal* **33**(10), 1620–1635.
URL: *https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690331006*

Snoek, J., Larochelle, H. and Adams, R. P. (2012), 'Practical bayesian optimization of machine learning algorithms'.

Sung, S. W. and Lee, I.-B. (1996), 'Limitations and countermeasures of pid controllers', *Industrial & Engineering Chemistry Research* **35**(8), 2596–2610.
URL: *https://doi.org/10.1021/ie960090+*

Sutton, R. S. and Barto, A. G. (2018), *Reinforcement Learning: An Introduction Second Edition*, MIT Press.