Research Article

Data-Driven Approximated Optimal Control of Sulfur Flotation Process

Mingfang He

School of Computer and Information Engineering, Central South University of Forestry and Technology, Changsha 410004, China

Correspondence should be addressed to Mingfang He; t20162306@csuft.edu.cn

Received 30 March 2019; Revised 6 June 2019; Accepted 3 July 2019; Published 16 July 2019

Academic Editor: Yongping Pan

Copyright © 2019 Mingfang He. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sulfur flotation process is a typical industry process with complex dynamics. For a sulfur flotation cell, the structure of the system model could be derived using first-principles and reaction kinetics. However, the model parameters cannot be obtained under certain working conditions. In this paper, by using adaptive dynamic programming (ADP), we establish a data-driven optimal control approach for the operation of a sulfur flotation cell without knowing the model parameters. By learning from the online production data, an initial admissible control policy iteratively converges to an approximated optimal control law, and the dependence of optimal control design on the full model knowledge is eliminated. A simulation environment of sulfur flotation process is constructed based on phenomenological model and industrial data. Some practical problems in the implementation of ADP, i.e., selection of basis functions, how to use the model structural information in the ADP-based control design, are investigated. The feasibility and performance of the proposed data-driven optimal control are tested in the simulation environment. The results indicate the potential of applying bioinspired control methods in flotation process.

1. Introduction

Sulfur is an important element with wide applications in pharmacy industry, chemical engineering, light industry, and food industry, to name a few. In most cases, sulfur exists in the form of sulfide or sulfate minerals. So sulfur flotation process is usually integrated into mineral processing and hydrometallurgy plants to recover the valuable sulfur from the mineral residue as a secondary product. On the other hand, by using sulfur flotation, more advanced and economic technologies can be applied in metallurgy plants, e.g., direct leaching technology. The grade and recovery are two main technical indexes which define the overall performance of a sulfur flotation process. The operation objective of a sulfur flotation process is to recover as much as high-graded sulfur as possible from the mineral residue by finding the optimal combination of the manipulated variables in a real-time manner.

Sulfur flotation is a typical three-phase (solid-liquid-gas) process. Generally, the operation of this type of process is nontrivial due to its complex and nonlinear dynamics which is a synthetical outcome of the intricate interactions between the reactants, inlet conditions, manipulated variables, and

disturbances. Therefore, the economical operation of flotation processes has aroused the interest from both mineral processing community and control community [1–3].

The flotation modeling approaches can be categorized into first-principle modeling [4] and soft-computing-based modeling [5]. Lynch et al. [6] and Gorain et al. [7] developed empirical models for the entire flotation plant and subprocesses in a flotation system. Yianatos proposed a first-order kinetic flotation model [8]. Koh and Schwarz studied the probabilistic modeling of the microprocesses, e.g., adhesion of particles to the bubbles in the pulp, in the flotation [9]. The flotation rate constant in the model varies with feeding conditions and working conditions. Therefore, to precisely determine the flotation rate constant remains an active research topic [4]. Bascur proposed a phenomenological model which considers both the particle population balance and hydraulic behavior [10]. The phenomenological model combined population balance, transfer of particles between phases and liquid balance, and is able to simulate the effects of manipulated variables (e.g., tailing and air flow rate) on a flotation process [2].

In flotation plants, the operation of a flotation process is conducted in a “setting and control” manner. The

FIGURE 1: Flow-diagram of a sulphur flotation circuit.

basic control loops maintain the primary variables at set-points [2], which is achieved by conventional PID control. On a higher level, machine vision-based control systems, model-based control methods, e.g., Model Predictive Control (MPC), adaptive control, expert systems, neural network, fuzzy-logic, and their commercial software are developed and applied [11]. The performance of the above model-based controllers relies on the model quality. Although the structure of a mathematical model could be obtained by applying physical/chemical principles and reaction kinetics, the model parameters vary with the working conditions and are not available under certain working conditions. Robust control approaches are the natural choice for these working conditions with modeling uncertainties. However, in some cases, the conservative feature of the robust control approaches keeps the derived controller away from optimal which in turn leads to economical loss.

In recent years, bioinspired control approaches have been wide-spreading. Learning is the basic feature of bioinspired control methods. It obtains the optimal or nearly optimal controller by learning from the “excitation-response” interactions with the environment [12], of which “try and error” is a typical example. Reinforcement learning (RL) is a bioinspired machine learning method. It is originally observed from the learning behavior in mammals [13]. RL represents how an agent/animal interacts with the unknown environment and modifies its action to obtain a maximized reward. Adaptive dynamic programming is an integration of adaptive control, dynamic programming, and reinforcement learning [14, 15]. By using an “actor-critic” structure, it can approximate the solution of the Hamilton-Jacobi-Bellman (HJB) equation online without knowing the full knowledge of the system model, or even when the system model is unknown [16, 17]. Thus, if an initial admissible control policy is available, an approximated optimal control can be obtained by solving the HJB equation iteratively. The industrial applications of ADP include oil production [18], coal gasification [19], and energy scheduling [20], to name a few. In recent years, learning-based control has been applied in the operation and control of froth flotation process. In [21, 22], a dual-rate and a two-layer control strategy based on Q-learning and off-policy reinforcement learning were proposed and tested via experimental study, respectively.

The aim of this study is to design an alternative optimal control approach for a sulfur flotation cell when the model parameters are unknown and the manipulated variables are constrained. The rest of this paper is organized as follows. In Section 2, the classical flotation kinetic model developed by Bascur [10] is introduced and used in the analysis of the optimal control problem in Section 3. Then the model is adopted to simulate the real dynamics of a flotation cell in Section 4. An optimal control based on ADP is designed to iteratively improve the performance of an initial admissible controller by learning from the online operational data generated from the simulation environment. The performance of the proposed optimal control is compared with the traditional PI controller. The simulation results, advantages, and shortcomings as well as future extensions of the proposed control are analyzed in Section 6.

2. Process Description

Typically, a sulfur flotation process is composed of three stages with specific functions, namely, roughing, scavenging, and cleaning (Figure 1). The roughing stage performs the initial separation of sulfur from gangue. Recovery rate is the main concern of roughing stage. The tailing of roughing stage is delivered to scavenging stage to seek for the remaining sulfur. The tailing of scavenging stage is also the final tailing of the entire flotation process. The froth of roughing stage overflows to the cleaning stage. The scavenging froth as well as the cleaning tails are returned to the roughing stage. The overflow of the cleaning stage is the final product, i.e., the sulfur concentrate. The three stages work collaboratively to maximize the sulfur recovery and concentrate grade and minimize the grade of final tails.

The economical production of a flotation process relies on the optimal control of a single flotation cell which is the basic flotation unit. As shown in Figure 2, the operation of sulfur flotation in a flotation cell is based on the natural hydrophobicity of sulfur. First, slurry containing both valuable sulfur and base gangue is fed into the flotation cell. Then, by blowing air into the flotation cell, the sulfur particles in the slurry are attached to the bubbles. The bubbles float upward and carry the particles to the froth layer. Finally, the sulfur particles are collected in the concentrate launder via the natural overflow

FIGURE 2: Structure of a sulphur flotation cell.

FIGURE 3: Particle movement in a flotation cell.

of the froth. Focused on a single flotation cell, the rest of this section presents the model structure which is then used along with the flotation mechanism to analyze the optimal control problem.

2.1. Flotation Behavior Analysis. The physical activities in a sulfur flotation cell involve the movement of solid particles driven by the flow of air and liquid. As illustrated in Figures 2 and 3 [10], considering the different roles of pulp and froth in determining the cell dynamics, the flotation cell is divided into two sections: froth volume and pulp volume. The froth volume and pulp volume are then subdivided in the liquid and air phases with which the mineral particles are associated. The states and movements of particles in a flotation cell mainly include the following:

- (i) S_{LP} : free in the pulp (particles in the liquid phase of the pulp volume)

- (ii) S_{BP} : attached in the pulp (particles attached to the bubbles in the pulp volume)
- (iii) S_{LF} : free in the froth (particles in the liquid of the froth volume)
- (iv) S_{BF} : attached in the froth (particles attached to the bubbles of the froth volume)

and

- (i) $S_{LP} \rightarrow S_{BP}$: free particles in the pulp attached to the bubbles in the pulp volume with the attachment rate constant k_{ij}^{PAT}
- (ii) $S_{BP} \rightarrow S_{LP}$: particles associated with the bubbles detached and enter the liquid phase in the pulp volume with the detachment rate constant k_{ij}^{PDT}
- (iii) $S_{LF} \rightarrow S_{BF}$: free particles in the froth layer attached to the bubbles in the froth volume with the attachment rate constant k_{ij}^{FAT}

- (iv) $S_{BF} \rightarrow S_{LF}$: particles associated with the bubbles detached and enter the liquid phase in the froth volume with the detachment rate constant k_{ij}^{FDT}
- (v) $Feed \rightarrow S_{LP}$: the particles in the input slurry (with liquid volume flow rate Q_{Feed} and Ψ_{ij}^{Feed} the number of particles of size i and mineralogical species j per unit volume) feed into the flotation cell
- (vi) $S_{LF} \rightarrow S_{LP}$: particles draining from the liquid phase of the froth volume (with the flow rate of Q_R and Ψ_{ij}^{LF} the number of particles of size i and mineralogical species j per unit volume of liquid in the froth volume, k_{ij}^R a constant indicating the segregation mechanism in the draining)
- (vii) $S_{LP} \rightarrow S_{LF}$: particles entrained by the liquid flow from pulp volume to the froth volume (with the flow rate of Q_E)
- (viii) $S_{BP} \rightarrow S_{BF}$: transport of particles attached on the bubbles in the pulp volume to the froth volume (with the air flow rate of Q'_A)
- (ix) $S_{LP} \rightarrow Tailings$: the particles in the pulp leave through the tailing port (with the flow rate of Q_T and Ψ_{ij}^{LP} the number of particles of size i and mineralogical species j per unit volume of liquid in the pulp)
- (x) $S_{BP} \rightarrow Tailings$: exit of particles attached on the bubbles in the pulp volume to the tailings (with the aeration rate of Q_{AT})
- (xi) $S_{LF} \rightarrow Concentrate$: the particles free in the froth overflow to the concentrate lander (with the flow rate of Q_C)
- (xii) $S_{BF} \rightarrow Concentrate$: the particles attached to the bubbles in the froth overflow to the concentrate lander (with the flow rate of Q_{AC})

2.2. Phenomenological Model of a Flotation Cell. Based on the above simplifications, consider the hydraulic behavior, and apply the following assumptions:

Assumption 1. Each phase in the pulp and froth volume of the flotation cell is perfectly mixed.

Assumption 2. The particle size and shape are uniform.

Assumption 3. The changes in the air volumes occur much faster than the liquid volumes, and steady state relations are used; i.e., $Q_A = Q'_A + Q_{AT}$, $Q'_A = Q_{AC}$ [10, 23].

then a process model, which is a combination of the particle population balance in the liquid phase and the air phase for the pulp and froth volume, is built to describe the dynamics of a flotation cell:

$$\frac{dC_j^P}{dt} = \frac{Q_{Feed}}{V_{LP}} C_j^{Feed} - \frac{Q_E + Q_A \alpha_j^P (V_{LP}/V_{BP})}{(1 + \alpha_j^P) V_{LP}} C_j^P \quad (1a)$$

$$+ \frac{Q_R k_j^R}{(1 + \alpha_j^F) V_{LP}} C_j^F - \frac{Q_T}{(1 + \alpha_j^P) V_{LP}} C_j^P$$

$$\frac{dC_j^F}{dt} = \frac{Q_E + Q_A \alpha_j^P (V_{LP}/V_{BP})}{(1 + \alpha_j^P) V_{LF}} C_j^P$$

$$- \frac{Q_R k_j^R}{(1 + \alpha_j^F) V_{LF}} C_j^F - \frac{Q_T \alpha_j^P}{(1 + \alpha_j^P) V_{LF}} C_j^P$$

$$- \frac{Q_A}{V_{BF}} C_j^F + \frac{Q_T V_{BP}}{V_{BF} V_{LP}} C_j^F \quad (1b)$$

$$\frac{dV_{LP}}{dt} = Q_{Feed} - Q_T - Q_E + Q_R \quad (1c)$$

$$\frac{dV_{LF}}{dt} = Q_E - Q_R - Q_C \quad (1d)$$

$$\frac{dh}{dt} = \frac{V_{LP}}{S(2V_{LP} - Sh)} (Q_{Feed} - Q_T - Q_E + Q_R) \quad (1e)$$

where (1a) and (1b) present the mass balance for mineralogical species j , while (1c) and (1d) describe the liquid volume balance at the pulp volume and froth volume, and

$$V_{BP} = Sh - V_{LP} \quad (2)$$

$$V_{BF} = V - Sh - V_{LF} \quad (3)$$

$$Q_C = \frac{V_{LF}}{V_{BF}} \left(Q_A - \frac{V_{BP}}{V_{LP}} Q_T \right) \quad (4)$$

The model (1a), (1b), (1c), (1d), and (1e) is a combination of the froth flotation models proposed in [10, 24]. It keeps the phenomenological behavior in [10], while it also approximates the dynamics of the pulp level in [24]. The physical meanings of the model parameters and the variables in Figure 3 are listed in Table 1. In Section 5, this model is applied to simulate the real flotation cell.

3. Optimal Control Problem Analysis

To start with the optimal control problem analysis, the model is reformulated first to investigate how the manipulated variables affect the flotation performance.

As indicated in [10], Q_E and Q_R can be formulated as

$$Q_E = \frac{6\delta}{d_{BP}} Q_A = r_1 Q_A \quad (5)$$

$$Q_R = \frac{1 - \mu_F}{\mu_F} \frac{Q_A}{S} = r_2 Q_A \quad (6)$$

where δ is the thickness of film surrounding the bubble, d_{BP} is the mean size bubble diameter in the pulp region, μ_F is the average air holdup in the plateau border of froth volume, and r_1 and r_2 are two variables changing with working conditions.

TABLE 1: List of model parameters and process variables.

Parm	Unit	Physical meaning
Q_{Feed}	m^3/h	Liquid volume flow rate of the slurry feed
Q_{C}	m^3/h	Liquid volume flow rate of the concentrate
Q_{T}	m^3/h	Liquid volume flow rate of the tailings
Q_{E}	m^3/h	Liquid volume flow rate of the entrainment from pulp to froth volume
Q_{R}	m^3/h	Liquid volume flow rate of the drainage from froth to pulp volume
Q_{A}	m^3/h	Air flow rate of the aeration
$Q_{\text{A}'}$	m^3/h	Air flow rate that leaves the pulp volume into the froth volume
Q_{AC}	m^3/h	Air flow rate that leaves the froth volume into the concentrate launder
Q_{AT}	m^3/h	Air flow rate that leaves the pulp volume into the tailings
V	m^3/h	Volume of the flotation cell
V_{TF}	m^3/h	Volume of the froth layer in the flotation cell
V_{LF}	m^3/h	Volume of liquid phase in the froth layer
V_{BF}	m^3/h	Volume of air phase in the froth layer
V_{TP}	m^3/h	Volume of the pulp layer in the flotation cell
V_{LP}	m^3/h	Volume of liquid phase in the pulp layer
V_{BP}	m^3/h	Volume of air phase in the pulp layer
C_j^{Feed}	t/m^3	Concentration of species j in the liquid volume of the feed slurry
C_j^{F}	t/m^3	Concentration of species j in the liquid volume of the froth layer
C_j^{P}	t/m^3	Concentration of species j in the liquid volume of the pulp layer
Ψ_{ij}^{Feed}	-	Number of particles of size i and mineralogical species j per unit volume in the feed
Ψ_{ij}^{LF}	-	Number of particles of size i and mineralogical species j per unit volume of liquid phase in the froth layer
Ψ_{ij}^{BF}	-	Number of particles of size i and mineralogical species j per unit volume of air phase in the froth layer
Ψ_{ij}^{LP}	-	Number of particles of size i and mineralogical species j per unit volume of liquid phase in the pulp layer
Ψ_{ij}^{BP}	-	Number of particles of size i and mineralogical species j per unit volume of air phase in the pulp layer
k_{ij}^{PAT}	-	Rate constant of particles of size i and mineralogical species j attach from the pulp to the bubble in the pulp volume
k_{ij}^{PDT}	-	Rate constant of particles of size i and mineralogical species j detach from the bubble to the pulp in the pulp volume
k_{ij}^{FAT}	-	Rate constant of particles of size i and mineralogical species j attach from the pulp to the bubble in the froth volume
k_{ij}^{FDT}	-	Rate constant of particles of size i and mineralogical species j attach from the bubble to the pulp in the froth volume
α_j^{F}	-	Rate constant of the attachment and detachment for species j in the froth
α_j^{P}	-	Rate constant of the attachment and detachment for species j in the pulp
k_j^{R}	-	Rate constant of draining
S	m^2	Cross-sectional area of the flotation cell
h	m	Pulp level

Combining (1a), (1b), (1c), (1d), and (1e), (5), (6), assume there are two mineralogical classes in the pulp, i.e., sulfur rich mineral (class 1) and sulfur poor mineral (class 2); the original model can be reformulated as

$$\frac{dC_1^{\text{P}}}{dt} = \frac{1}{V_{\text{LP}}} \left(Q_{\text{Feed}} C_1^{\text{Feed}} - \beta_1 Q_{\text{A}} - \frac{Q_{\text{T}}}{1 + \alpha_1^{\text{P}}} C_1^{\text{P}} \right) \quad (7a)$$

$$\frac{dC_1^{\text{F}}}{dt} = \frac{1}{V_{\text{LF}}} \left[\beta_1 Q_{\text{A}} - \frac{V_{\text{LF}}}{V_{\text{BF}}} Q_{\text{A}} C_1^{\text{F}} - \frac{\alpha_1^{\text{P}}}{1 + \alpha_1^{\text{P}}} Q_{\text{T}} C_1^{\text{P}} + \frac{V_{\text{LF}} V_{\text{BP}}}{V_{\text{BF}} V_{\text{LP}}} Q_{\text{T}} C_1^{\text{F}} \right] \quad (7b)$$

$$\frac{dC_2^P}{dt} = \frac{1}{V_{LP}} \left(Q_{Feed} C_2^{Feed} - \beta_2 Q_A - \frac{Q_T}{1 + \alpha_2^P} C_2^P \right) \quad (7c)$$

$$\frac{dC_2^F}{dt} = \frac{1}{V_{LF}} \left[\beta_2 Q_A - \frac{V_{LF}}{V_{BF}} Q_A C_2^F - \frac{\alpha_2^P}{1 + \alpha_2^P} Q_T C_2^P + \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} Q_T C_2^F \right] \quad (7d)$$

$$\frac{dV_{LP}}{dt} = Q_{Feed} - Q_T - r_3 Q_A \quad (7e)$$

$$\frac{dV_{LF}}{dt} = r_3 Q_A - \frac{V_{LF}}{V_{BF}} Q_A + \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} Q_T \quad (7f)$$

$$\frac{dh}{dt} = \frac{V_{LP}}{S(2V_{LP} - Sh)} (Q_{Feed} - Q_T - r_3 Q_A) \quad (7g)$$

where

$$r_3 = r_1 - r_2 \quad (8)$$

and for the mineralogical class i

$$\beta_i = \frac{1}{1 + \alpha_i^P} \left(r_1 + \alpha_i^P \frac{V_{LP}}{V_{BP}} \right) C_i^P - r_2 \alpha_i^{kF} C_i^F \quad (9)$$

where

$$\alpha_i^{kF} = \frac{k_i^R}{1 + \alpha_i^F} \quad (10)$$

Q_A and Q_T are constrained in reasonable operation ranges and rationally adjusted to avoid too high or too low pulp level.

The reformulated model (7a), (7b), (7c), (7d), (7e), (7f), and (7g) indicates that air flow rate Q_A , tailing flow rate Q_T , and pulp level h are important factors that determine the flotation efficiency. This conclusion is consistent with the site experience. In a flotation plant, the operators adjust these variables to obtain the desired recovery γ^* and concentrate grade η^* which are determined by a higher-level decision maker. The concentrate grade η and recovery γ are defined by

$$\eta = \frac{V_{LF} C_1^F g_1 + V_{LF} C_2^F g_2}{V_{LF} C_1^F + V_{LF} C_2^F} \quad (11a)$$

$$\gamma = \frac{(\eta_{in} - \eta_{tail}) \eta}{(\eta - \eta_{tail}) \eta_{in}} \quad (11b)$$

where C^F is the density of the final concentrate (t/m^3), g_1 and g_2 are the sulfur grade of the two mineralogical classes, η_{in} is the grade of the feed, and

$$\eta_{tail} = \frac{V_{LP} C_1^P g_1 + V_{LP} C_2^P g_2}{V_{LP} C_1^P + V_{LP} C_2^P} \quad (12)$$

is the grade of the tailings. In practice, the optimal control of the entire flotation process can be transformed to the setting and control of V_{LP} and V_{LF} (determine the pulp level) and C^F (determines the concentrate grade) of each cell to assure that, in each cell, a sufficient part of the froth is collected while undesired situations, e.g., pulp tuning, are avoided.

4. ADP-Based Optimal Control Design

The dynamic of a flotation cell can be expressed in a state-space format:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \mathbf{u} \quad (13)$$

in which, the states and the controlled and manipulated variables are

$$\mathbf{x} = [C_1^P, C_1^F, C_2^P, C_2^F, V_{LP}, V_{LF}, h]^T \quad (14)$$

$$\mathbf{u} = [Q_T, Q_A]^T$$

and

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \frac{C_1^{Feed}}{V_{LP}} Q_{Feed} \\ 0 \\ \frac{C_2^{Feed}}{V_{LP}} Q_{Feed} \\ 0 \\ Q_{Feed} \\ 0 \\ \frac{V_{LP}}{S(2V_{LP} - Sh)} Q_{Feed} \end{bmatrix} \quad (15)$$

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} -\frac{C_1^P}{(1 + \alpha_1^P) V_{LP}} & -\frac{\beta_1}{V_{LP}} \\ -\frac{\alpha_1^P C_1^P}{(1 + \alpha_1^P) V_{LF}} + \frac{V_{BP} C_1^F}{V_{BF} V_{LP}} & \frac{\beta_1}{V_{LF}} - \frac{C_1^F}{V_{BF}} \\ -\frac{C_2^P}{(1 + \alpha_2^P) V_{LP}} & -\frac{\beta_2}{V_{LP}} \\ -\frac{\alpha_2^P C_2^P}{(1 + \alpha_2^P) V_{LF}} + \frac{V_{BP} C_2^F}{V_{BF} V_{LP}} & \frac{\beta_2}{V_{LF}} - \frac{C_2^F}{V_{BF}} \\ -1 & -r_3 \\ \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} & r_3 - \frac{V_{LF}}{V_{BF}} \\ -\frac{V_{LP}}{S(2V_{LP} - Sh)} & -\frac{r_3 V_{LP}}{S(2V_{LP} - Sh)} \end{bmatrix} \quad (16)$$

If the model parameters $[r_1, r_2, \alpha_1^P, \alpha_1^F, k_1^R, \alpha_2^P, \alpha_2^F, k_2^R]$ are unknown under certain working conditions, then the model is only partially known and the performance of traditional optimal control method cannot be guaranteed. This situation also exists in other industry processes. To provide an optimal control alternative when the model parameters are unknown, this section studies the ADP-based optimal control of a flotation cell. ADP does not rely on a precise system model. It can, starting from an initial admissible controller, iteratively improve the controller via learning from the online operational data.

4.1. Preliminaries on Policy Iteration. Consider the following objective function:

$$J(\cdot) = \int_{t_0}^{t_f} [Q(\mathbf{x}) + W(\mathbf{u}(\tau))] dt \quad (17)$$

where $Q(\mathbf{x})$ is the performance integrand of the system states, and $W(\mathbf{u})$ is the nonquadratic performance integrand of the input variables:

$$W(\mathbf{u}) = 2 \int_0^{\mathbf{u}} (\mathbf{H}^{-1}(\mathbf{v}))^T \mathbf{R} d\mathbf{v} \quad (18)$$

where

$$\mathbf{H}(\mathbf{v}) = [h_1(v_1) \ h_2(v_2) \ \cdots \ h_m(v_m)]^T \in \mathbb{R}^m \quad (19)$$

$$\mathbf{H}^{-1}(\mathbf{v}) = [h_1^{-1}(v_1) \ h_2^{-1}(v_2) \ \cdots \ h_m^{-1}(v_m)]^T \in \mathbb{R}^m$$

$h_i(\cdot)$ is a bounded, monotonic, odd $C^p(p \geq 1)$ function ($i = 1, 2, \dots, m$) and maps the control bounds imposed, $h_i^{-1}(\cdot)$ is the inverse function of $h_i(\cdot)$, $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_m]^T$, and $\mathbf{R} \in \mathbb{R}^{m \times m}$ is a diagonal matrix with its diagonal elements all positive.

The optimal control policy is

$$\mathbf{u}^*(\mathbf{x}) = -\mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} g^T(\mathbf{x}) \nabla V^*(\mathbf{x}) \right) \quad (20)$$

where $V^*(\mathbf{x})$ is the solution of a Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} & \nabla V(\mathbf{x})^T \left[f(\mathbf{x}) - g(\mathbf{x}) \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} g^T(\mathbf{x}) \nabla V(\mathbf{x}) \right) \right] \\ & + Q(\mathbf{x}) + 2 \int_0^{-\mathbf{H}((1/2)\mathbf{R}^{-1}g^T(\mathbf{x})\nabla V(\mathbf{x}))} (\mathbf{H}^{-1}(\mathbf{v}))^T \mathbf{R} d\mathbf{v} \quad (21) \\ & = 0 \end{aligned}$$

with $V(0) = 0$.

The HJB equation is generally difficult to solve analytically. Therefore, Police Iteration (PI) approach is required:

- (1) Apply an initial stabilizing admissible controller $u_0(\mathbf{x})$ to the system
- (2) For $i \geq 0$, solve the following Lyapunov equation for $V(i)$:

$$\begin{aligned} & \nabla V_i(\mathbf{x})^T [f(\mathbf{x}) + g(\mathbf{x}) \mathbf{u}_i(\mathbf{x})] + Q(\mathbf{x}) \\ & + 2 \int_0^{\mathbf{u}_i(\mathbf{x})} (\mathbf{H}^{-1}(\mathbf{v}))^T \mathbf{R} d\mathbf{v} = 0 \quad (22) \end{aligned}$$

- (3) Update the control law

$$\mathbf{u}_{i+1}(\mathbf{x}) = -\mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} g^T(\mathbf{x}) \nabla V_i(\mathbf{x}) \right) \quad (23)$$

- (4) if stop criterion is met, then stop; otherwise, $i = i + 1$; go to Step (2).

4.2. ADP-Based Input-Constrained Optimal Control. In the police iteration, full knowledge of $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$ is required. In this study, a model-free PI technique is extended to the input-constrained case to eliminate this requirement.

To start with, reformulate $\mathbf{u}$ as

$$\mathbf{u} = \mathbf{u}^{(i)} + \mathbf{v}^{(i)} \quad (24)$$

Then

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \mathbf{u}^{(i)} + \mathbf{g}(\mathbf{x}) \mathbf{v}^{(i)} \quad (25)$$

where $\mathbf{u}$ is the control forced actually on the system, and $\mathbf{u}^{(i)}$ is the control policy to be improved iteratively.

Consider a cost function $V^{(i)}$ satisfying (22):

$$\begin{aligned} \dot{V}^{(i)} &= V^{(i)T} \left\{ [f(\mathbf{x}) + g(\mathbf{x}) \mathbf{u}^{(i)}] + g(\mathbf{x}) \mathbf{v}^{(i)} \right\} \\ &= -Q(\mathbf{x}) - W(\mathbf{u}^{(i)}) - 2\mathbf{H}^{-T}(\mathbf{u}^{(i+1)}) \mathbf{R} \mathbf{v}^{(i)} \quad (26) \end{aligned}$$

Integrate $\dot{V}^{(i)}$ on a time interval $[t, t + \Delta t]$:

$$\begin{aligned} & V^{(i)}(t + \Delta t) - V^{(i)}(t) \\ &= \int_t^{t+\Delta t} [-Q(\mathbf{x}) - W(\mathbf{u}^{(i)}) - 2\mathbf{H}^{-T}(\mathbf{u}^{(i+1)}) \mathbf{R} \mathbf{v}^{(i)}] d\tau \quad (27) \end{aligned}$$

Assume $\{\phi_j(\mathbf{x})\}_{j=1}^{\infty}$ and $\{\psi_j(\mathbf{x})\}_{j=1}^{\infty}$ are two infinite sequences of linearly independent smooth basis functions on $\mathbb{X}$, where $\phi_j(0) = 0$ and $\psi_j(0) = 0$ for all $j = 1, 2, \dots$. Approximate the cost function and the control policy as

$$\widehat{V}^{(i)} = \sum_{j=1}^{N_1} \widehat{c}_j^{(i)} \phi_j(\mathbf{x}) = \Phi(\mathbf{x})^T \widehat{\mathbf{c}}^{(i)} \quad (28)$$

where $\widehat{\mathbf{c}}^{(i)} = [\widehat{c}_1^{(i)} \ \widehat{c}_2^{(i)} \ \cdots \ \widehat{c}_{N_1}^{(i)}]^T$, $\Phi(\mathbf{x}) = [\phi_1(\mathbf{x}) \ \phi_2(\mathbf{x}) \ \cdots \ \phi_{N_1}(\mathbf{x})]^T$, and

$$\widehat{\mathbf{u}}^{(i)} = -\mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-T} \widehat{\mathbf{Y}}_i(\mathbf{x}) \right) \quad (29)$$

where

$$\widehat{\mathbf{Y}}_i(\mathbf{x}) = \widehat{\mathbf{k}}^{(i)} \Psi(\mathbf{x}) \quad (30)$$

with

$$\widehat{\mathbf{k}}^{(i)} = \begin{bmatrix} \widehat{k}_{11}^{(i)} & \widehat{k}_{12}^{(i)} & \cdots & \widehat{k}_{1N_2}^{(i)} \\ \widehat{k}_{21}^{(i)} & \widehat{k}_{22}^{(i)} & \cdots & \widehat{k}_{2N_2}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{k}_{m1}^{(i)} & \widehat{k}_{m2}^{(i)} & \cdots & \widehat{k}_{mN_2}^{(i)} \end{bmatrix} \quad (31)$$

$$\Psi(\mathbf{x}) = [\psi_1(\mathbf{x}) \ \psi_2(\mathbf{x}) \ \cdots \ \psi_{N_2}(\mathbf{x})]^T \quad (32)$$

Note that the input is approximated by (29). This eliminates the requirement on the knowledge on $\mathbf{g}(\mathbf{x})$, as indicated by

(23). This problem can also be handled by constructing a NN-based approximation system, such as the method proposed in [25].

In (28)-(32), $N_1 > 0$ and $N_2 > 0$ are two sufficiently large integers. Substitute Eqs. (28)-(32) into (27), and denote t and $t + \Delta t$ as t_k and t_{k+1} :

$$\begin{aligned} & \sum_{j=1}^{N_1} \hat{c}_j^{(i)} [\phi_j(\mathbf{x}(t_{k+1})) - \phi_j(\mathbf{x}(t_k))] \\ & - \int_{t_k}^{t_{k+1}} \hat{\mathbf{k}}^{(i+1)} \Psi(\mathbf{x}) (u_j(\mathbf{x}) - u_j^{(i)}(\mathbf{x})) dt \\ & = - \int_{t_k}^{t_{k+1}} [Q(\mathbf{x}) + W(\mathbf{u}^{(i)})] dt + \varepsilon_{i,k} \end{aligned} \quad (33)$$

In (33), there are $N_1 + N_2$ variables to be determined in $\hat{\mathbf{c}}^{(i)}$ and $\hat{\mathbf{k}}^{(i+1)}$ ($j = 1, 2, \dots, m$), and consider a sufficient long time sequence $\{t_k\}_{k=0}^L$ with $L \geq N_1 + N_2$; then according to (33)

$$\begin{aligned} & [\mathbf{I}_{\Phi\mathbf{x}} - (\mathbf{I}_{\Psi\mathbf{u}} + \mathbf{I}_{\Psi\mathbf{u}^{(i)}})] \begin{bmatrix} \hat{\mathbf{c}}^{(i)} \\ \text{vec}(\hat{\mathbf{k}}^{(i+1)\top}) \end{bmatrix} \\ & = -\mathbf{I}_{Q\mathbf{x}} - \mathbf{I}_{W\mathbf{u}^{(i)}} + \Delta_i \end{aligned} \quad (34)$$

where

$$\mathbf{I}_{\Phi\mathbf{x}} = \begin{bmatrix} \Phi(\mathbf{x}(1))^\top - \Phi(\mathbf{x}(0))^\top \\ \Phi(\mathbf{x}(2))^\top - \Phi(\mathbf{x}(1))^\top \\ \vdots \\ \Phi(\mathbf{x}(L))^\top - \Phi(\mathbf{x}(L-1))^\top \end{bmatrix}_{L \times N_1} \quad (35)$$

$$\mathbf{I}_{\Psi\mathbf{u}} = \begin{bmatrix} \int_{t_0}^{t_1} \mathbf{u}^\top [I_m \otimes \Psi(\mathbf{x})^\top] dt \\ \int_{t_1}^{t_2} \mathbf{u}^\top [I_m \otimes \Psi(\mathbf{x})^\top] dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} \mathbf{u}^\top [I_m \otimes \Psi(\mathbf{x})^\top] dt \end{bmatrix}_{L \times m N_2} \quad (36)$$

$$\begin{aligned} & \mathbf{I}_{\Psi\mathbf{u}^{(i)}} \\ & = \begin{bmatrix} \int_{t_0}^{t_1} \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} \hat{\mathbf{k}}^{(i)} \Psi(\mathbf{x}) \right)^\top [I_m \otimes \Psi(\mathbf{x})^\top] dt \\ \int_{t_1}^{t_2} \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} \hat{\mathbf{k}}^{(i)} \Psi(\mathbf{x}) \right)^\top [I_m \otimes \Psi(\mathbf{x})^\top] dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} \hat{\mathbf{k}}^{(i)} \Psi(\mathbf{x}) \right)^\top [I_m \otimes \Psi(\mathbf{x})^\top] dt \end{bmatrix}_{L \times m N_2} \end{aligned} \quad (37)$$

$$\mathbf{I}_{Q\mathbf{x}} = \begin{bmatrix} \int_{t_0}^{t_1} Q(\mathbf{x}) dt \\ \int_{t_1}^{t_2} Q(\mathbf{x}) dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} Q(\mathbf{x}) dt \end{bmatrix}_{L \times 1} \quad (38)$$

$$\mathbf{I}_{W\mathbf{u}^{(i)}} = \begin{bmatrix} \int_{t_0}^{t_1} W(\mathbf{u}^{(i)}) dt \\ \int_{t_1}^{t_2} W(\mathbf{u}^{(i)}) dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} W(\mathbf{u}^{(i)}) dt \end{bmatrix}_{L \times 1} \quad (39)$$

and $\Delta_i = [\varepsilon_{i,1} \ \varepsilon_{i,1} \ \dots \ \varepsilon_{i,L}]^\top \in \mathbb{R}^L$ is the approximation error.

From (34), if

$$\text{rank}(\Theta(\hat{\mathbf{k}}^{(i)})) = N_1 + m N_2 \quad (40)$$

where $\Theta(\hat{\mathbf{k}}^{(i)}) = [\mathbf{I}_{\Phi\mathbf{x}} \ -(\mathbf{I}_{\Psi\mathbf{u}} + \mathbf{I}_{\Psi\mathbf{u}^{(i)}})]$, then $\hat{\mathbf{c}}^{(i)}$ and $\hat{\mathbf{k}}^{(i+1)}$ can be obtained as

$$\begin{aligned} & \begin{bmatrix} \hat{\mathbf{c}}^{(i)} \\ \text{vec}(\hat{\mathbf{k}}^{(i+1)\top}) \end{bmatrix} \\ & = [\Theta(\hat{\mathbf{k}}^{(i)})^\top \Theta(\hat{\mathbf{k}}^{(i)})]^{-1} \Theta(\hat{\mathbf{k}}^{(i)})^\top \Pi(\hat{\mathbf{k}}^{(i)}) \end{aligned} \quad (41)$$

where $\Pi(\hat{\mathbf{k}}^{(i)}) = [-\mathbf{I}_{Q\mathbf{x}} - \mathbf{I}_{W\mathbf{u}^{(i)}}]$. (41) along with (28)-(32) serves as a computational approach to approximate $V^{(i)}$ and generate an improved controller $\mathbf{u}^{(i+1)}$ when an admissible control policy $\mathbf{u}^{(i)}$ is given. By using this approach iteratively, starting from an admissible bounded control $\mathbf{u}^{(0)}$, a control sequence $\{\mathbf{u}^{(i)}(\mathbf{x})\}_{i=0}^\infty$ could be obtained and satisfies the following theorem. As the iteration is conducted after the online stage when the information needed for the calculation of $\Theta(\hat{\mathbf{k}}^{(i)})$ and $\Pi(\hat{\mathbf{k}}^{(i)})$ is collected, this method belongs to off-policy reinforcement learning [26, 27].

Theorem 4. Consider system (13) with input constraints; if there exists an optimal controller and Assumptions 1 and 2 hold, then

$$\mathbf{u}^{(i+1)} = -\mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} \hat{\mathbf{k}}^{(i)} \Psi(\mathbf{x}) \right) \quad (42)$$

which is generated by the PI ((41), (28), and (29)) is an admissible control sequence for (13) on $\mathbb{X}$. The cost function $V^{(i)} = J(t_0)|_{\mathbf{u}=\mathbf{u}^{(i)}}$ satisfies the following Lyapunov equation (LE):

$$V^{(i)\top} [\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \mathbf{u}^{(i)}] + [Q(\mathbf{x}) + W(\mathbf{u}^{(i)})] = 0 \quad (43)$$

Then $V^*(\mathbf{x}) \leq V^{(i+1)}(\mathbf{x}) \leq V^{(i)}(\mathbf{x})$ for $\forall \mathbf{x} \in \mathbb{X}$. In addition, $V^{(i)} \rightarrow V^*$, $\mathbf{u}^{(i)} \rightarrow \mathbf{u}^*$ uniformly on $\mathbb{X}$.

Proof. The proof of this theorem follows the same lines of reasoning as in the proof of Theorem 1 and Theorem 3.1 in [28, 29], and it is omitted here for brevity. $\square$

Assumption 5. There exist $L_0 > 0$ and $\delta > 0$, such that for all $L > L_0$

$$\frac{1}{L} \sum_{k=1}^L \theta_{i,k}^T \theta_{i,k} \geq I_{N_1+N_2} \quad (44)$$

where $\theta_{i,k}$ is the k th row of $\Theta(\hat{\mathbf{k}}^{(i)})$.

Assumption 6. The closed loop system is composed of (16) and

$$\mathbf{u} = \mathbf{u}_0 + \mathbf{e} \quad (45)$$

is ISS (Input-to-State Stable) [30] when the exploration noise $\mathbf{e}$ is considered as the input.

5. Experimental Study

5.1. Simulation Setup. In order to test the feasibility and performance of the proposed optimal control, a testing environment based on the model introduced in Section 2 was constructed to simulate a real flotation cell. Consider the available instrumentation in the real flotation plant; the model (7a), (7b), (7c), (7d), (7e), (7f), and (7g) was reformulated to (46a), (46b), (46c), (46d), (46e), (46f), and (46g), which makes the simulation more close to practical. In the reformulated model, new system states $M_1^P = V_{LP} C_1^P$, $M_1^F = V_{LF} C_1^F$, $M_2^P = V_{LP} C_2^P$, and $M_2^F = V_{LF} C_2^F$ are introduced. So $C_1^P = M_1^P / V_{LP}$, $C_1^F = M_1^F / V_{LF}$, $C_2^P = M_2^P / V_{LP}$, and $C_2^F = M_2^F / V_{LF}$.

$$\begin{aligned} \frac{dM_1^P}{dt} = & C_1^P (Q_{\text{Feed}} - Q_T - r_3 Q_A) + \left(Q_{\text{Feed}} C_1^{\text{Feed}} \right. \\ & \left. - \beta_1 Q_A - \frac{Q_T}{1 + \alpha_1^P} C_1^P \right) \end{aligned} \quad (46a)$$

$$\begin{aligned} \frac{dM_1^F}{dt} = & C_1^F \left(r_3 Q_A - \frac{V_{LF}}{V_{BF}} Q_A + \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} Q_T \right) \\ & + \left[\beta_1 Q_A - \frac{V_{LF}}{V_{BF}} Q_A C_1^F - \frac{\alpha_1^P}{1 + \alpha_1^P} Q_T C_1^P \right. \\ & \left. + \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} Q_T C_1^F \right] \end{aligned} \quad (46b)$$

$$\begin{aligned} \frac{dM_2^P}{dt} = & C_2^P (Q_{\text{Feed}} - Q_T - r_3 Q_A) + \left(Q_{\text{Feed}} C_2^{\text{Feed}} \right. \\ & \left. - \beta_2 Q_A - \frac{Q_T}{1 + \alpha_2^P} C_2^P \right) \end{aligned} \quad (46c)$$

$$\begin{aligned} \frac{dM_2^F}{dt} = & C_2^F \left(r_3 Q_A - \frac{V_{LF}}{V_{BF}} Q_A + \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} Q_T \right) \\ & + \left[\beta_2 Q_A - \frac{V_{LF}}{V_{BF}} Q_A C_2^F - \frac{\alpha_2^P}{1 + \alpha_2^P} Q_T C_2^P \right. \\ & \left. + \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} Q_T C_2^F \right] \end{aligned} \quad (46d)$$

$$\frac{dV_{LP}}{dt} = Q_{\text{Feed}} - Q_T - r_3 Q_A \quad (46e)$$

$$\frac{dV_{LF}}{dt} = r_3 Q_A - \frac{V_{LF}}{V_{BF}} Q_A + \frac{V_{LF} V_{BP}}{V_{BF} V_{LP}} Q_T \quad (46f)$$

$$\frac{dh}{dt} = \frac{V_{LP}}{S(2V_{LP} - Sh)} (Q_{\text{Feed}} - Q_T - r_3 Q_A) \quad (46g)$$

The parameters of the model $[r_1, r_2, \alpha_1^P, \alpha_1^F, k_1^R, \alpha_2^P, \alpha_2^F, k_2^R]$ were identified from industrial data collected from a real plant [10]. The initial value of system states is

$$\begin{aligned} [M_1^P(0), M_1^F(0), M_2^P(0), M_2^F(0), V_{LP}(0), V_{LF}(0), h(0)] \\ = [4.4531, 2.1558, 8.3847, 0.6074, 33.0105, 0.8893, 3.0] \end{aligned} \quad (47)$$

The model parameter set is

$$\begin{aligned} [r_1, r_2, r_3, \alpha_1^P, \alpha_2^P, \alpha_1^{kF}, \alpha_2^{kF}] \\ = [0.1667, 0.0417, 0.125, 4.2060, 0.2921, 0.5335, 4.6816] \end{aligned} \quad (48)$$

the concentration and sulfur grade of the two mineralogical classes are:

$$\begin{aligned} [C_1^{\text{Feed}}, C_2^{\text{Feed}}] &= [0.22, 0.27] \\ [g_1, g_2] &= [12\%, 0.5\%] \end{aligned} \quad (49)$$

In order to test two different ways of control, i.e., setting the recovery rate and grade and setting the concentrations of each mineralogical species in both pulp and froth layer, the experimental study includes two tasks:

(i) Task 1: the control objective is to obtain the desired recovery rate and grade:

$$[\gamma^{\text{obj}}, \eta^{\text{obj}}] = [41\%, 0.09\%] \quad (50)$$

(ii) Task 2: the control objective is to obtain desired concentrations of each mineralogical species in both pulp and froth layer:

$$\begin{aligned} C^* &= [C_1^{P*}, C_1^{F*}, C_2^{P*}, C_2^{F*}] \\ &= [0.1290, 2.5000, 0.2545, 0.6600] \end{aligned} \quad (51)$$

In Task 1 and Task 2, the performance integrand $Q(\mathbf{x})$ was chosen as

$$Q(\mathbf{x}) = q_1^{\text{obj}} (\gamma - \gamma^{\text{obj}})^2 + q_2^{\text{obj}} (\eta - \eta^{\text{obj}})^2 \quad (52)$$

and

$$\mathbf{Q}(\mathbf{x}) = (\mathbf{C} - \mathbf{C}^*)^T \mathbf{Q}(\mathbf{C} - \mathbf{C}^*) \quad (53)$$

respectively, where

$$\begin{aligned} q_1^{\text{obj}} &= 2 \times 10^6 \\ q_2^{\text{obj}} &= 1 \\ \mathbf{Q} &= \text{diag}([1, 1, 1, 1]) \\ \mathbf{C} &= [\mathbf{C}_1^P, \mathbf{C}_1^F, \mathbf{C}_2^P, \mathbf{C}_2^F] \end{aligned} \quad (54)$$

To handle the input constraint, $h(\cdot)$ was selected as hyperbolic tangent function, and $\mathbf{R} = \mathbf{I}_2$.

The basis functions for the approximation of the control input were selected as

$$\begin{aligned} Q_T &= \frac{Q_T^{\max} + Q_T^{\min}}{2} + \frac{Q_T^{\max} - Q_T^{\min}}{2} \tanh(\hat{k}_{11}x_1 \\ &\quad + \hat{k}_{12}x_2 + \hat{k}_{13}x_3 + \hat{k}_{14}x_4 + \hat{k}_{15}x_5 + \hat{k}_{16}x_6 + e_1) \end{aligned} \quad (55)$$

$$\begin{aligned} Q_A &= \frac{Q_A^{\max} + Q_A^{\min}}{2} + \frac{Q_A^{\max} - Q_A^{\min}}{2} \tanh(\hat{k}_{21}x_1 \\ &\quad + \hat{k}_{22}x_2 + \hat{k}_{23}x_3 + \hat{k}_{24}x_4 + \hat{k}_{25}x_5 + \hat{k}_{26}x_6 + e_2) \end{aligned} \quad (56)$$

where e_1 and e_2 are the explosion noises, and

$$\begin{aligned} e_1 &= 100(\text{rand}_{11} - 0.5)[(\text{rand}_{12} - 0.5)\sin(t) \\ &\quad + (\text{rand}_{13} - 0.5)\sin(10t) \\ &\quad + (\text{rand}_{14} - 0.5)\sin(100t)] \\ e_2 &= 100(\text{rand}_{21} - 0.5)[(\text{rand}_{22} - 0.5)\sin(t) \\ &\quad + (\text{rand}_{23} - 0.5)\sin(10t) \\ &\quad + (\text{rand}_{24} - 0.5)\sin(100t)] \end{aligned} \quad (57)$$

$\{\text{rand}_{ij}\}$ are random variables, $i = 1, 2, j = 1, 2, 3, 4$.

The basis functions for the approximation of the value function were selected as:

$$\begin{aligned} \hat{V} &= \hat{c}_1x_1^2 + \hat{c}_2x_2^2 + \hat{c}_3x_3^2 + \hat{c}_4x_4^2 + \hat{c}_5x_5^2 + \hat{c}_6x_6^2 + \hat{c}_7x_7^2 \\ &\quad + \hat{c}_8x_1^2x_2^2 + \hat{c}_9x_1^2x_3^2 + \hat{c}_{10}x_1^2x_4^2 + \hat{c}_{11}x_1^2x_5^2 \\ &\quad + \hat{c}_{12}x_1^2x_6^2 + \hat{c}_{13}x_1^2x_7^2 + \hat{c}_{14}x_2^2x_3^2 + \hat{c}_{15}x_2^2x_4^2 \\ &\quad + \hat{c}_{16}x_2^2x_5^2 + \hat{c}_{17}x_2^2x_6^2 + \hat{c}_{18}x_2^2x_7^2 + \hat{c}_{19}x_3^2x_4^2 \\ &\quad + \hat{c}_{20}x_3^2x_5^2 + \hat{c}_{21}x_3^2x_6^2 + \hat{c}_{22}x_3^2x_7^2 + \hat{c}_{23}x_4^2x_5^2 \\ &\quad + \hat{c}_{24}x_4^2x_6^2 + \hat{c}_{25}x_4^2x_7^2 + \hat{c}_{26}x_5^2x_6^2 + \hat{c}_{27}x_5^2x_7^2 \\ &\quad + \hat{c}_{28}x_6^2x_7^2 \end{aligned} \quad (58)$$

for Task 1, and

$$\begin{aligned} \hat{V} &= \hat{c}_1z_1^2 + \hat{c}_2z_2^2 + \hat{c}_3z_3^2 + \hat{c}_4z_4^2 + \hat{c}_5z_1^4 + \hat{c}_6z_2^4 + \hat{c}_7z_3^4 \\ &\quad + \hat{c}_8z_4^4 + \hat{c}_9z_1^2z_2^2 + \hat{c}_{10}z_1^2z_3^2 + \hat{c}_{11}z_1^2z_4^2 + \hat{c}_{12}z_2^2z_3^2 \\ &\quad + \hat{c}_{13}z_2^2z_4^2 + \hat{c}_{14}z_3^2z_4^2 + \hat{c}_{15}z_1^6 + \hat{c}_{16}z_2^6 + \hat{c}_{17}z_3^6 \\ &\quad + \hat{c}_{18}z_4^6 + \hat{c}_{19}z_1^2z_2^4 + \hat{c}_{20}z_1^2z_3^4 + \hat{c}_{21}z_1^2z_4^4 + \hat{c}_{22}z_2^2z_3^4 \\ &\quad + \hat{c}_{23}z_2^2z_4^4 + \hat{c}_{24}z_3^2z_4^4 + \hat{c}_{25}z_1^4z_2^2 + \hat{c}_{26}z_1^4z_3^2 \\ &\quad + \hat{c}_{27}z_1^4z_4^2 + \hat{c}_{28}z_2^4z_3^2 + \hat{c}_{29}z_2^4z_4^2 + \hat{c}_{30}z_3^4z_4^2 \end{aligned} \quad (59)$$

for Task 2, respectively, where

$$z_1 = \frac{x_1}{x_5} - \frac{x_1^{\text{obj}}}{x_5} \quad (60)$$

The learning phase is 30 minutes. After the operational data was collected, an offline policy iteration was conducted until the stop criterion was met.

5.2. Simulation Results and Discussion. In Task 1, the initial admissible controller was selected as:

$$\begin{aligned} &[\hat{k}_{11}, \hat{k}_{12}, \hat{k}_{13}, \hat{k}_{14}, \hat{k}_{15}, \hat{k}_{16}] \\ &= [-0.01, -0.01, -0.01, -0.01, 0, 0] \\ &[\hat{k}_{21}, \hat{k}_{22}, \hat{k}_{23}, \hat{k}_{24}, \hat{k}_{25}, \hat{k}_{26}] \\ &= [-0.01, -0.01, -0.01, -0.01, 0, 0] \end{aligned} \quad (61)$$

The resulting controller was

$$\begin{aligned} Q_T &= 965 + 10 \tanh(133.8804x_1 - 13.3021x_2 \\ &\quad - 21.5489x_3 + 19.2795x_4 - 58.2083x_5 \\ &\quad - 4.5132x_6) \\ Q_A &= 240 + 10 \tanh(-283.1446x_1 + 32.4985x_2 \\ &\quad + 51.4569x_3 - 31.8602x_4 + 96.6250x_5 \\ &\quad + 9.2052x_6) \end{aligned} \quad (62)$$

The simulation results of Task 1 are shown in Figures 4–14. Figure 12 shows that the recovery raised to 41% after 1 hour and kept stable since then, while the variation of concentrate grade was small. Figures 13 and 14 demonstrate that the weighting variables for the approximation of control input gradually converge. Figures 4–10 present the variations of C_{P1} , C_{F1} , C_{P2} , C_{F2} , V_{LP} , V_{LF} , and pulp level, respectively. Figure 11 indicates the tailing flow rate and the air flow rate are within their limitations; see [955, 975] and [230, 250]. From these figures, it was found that the increase of recovery is mainly caused by the increase of C_{F1} (the concentration of the rich mineralogical class in the froth) and decrease of pulp level h , which is consistent with site experience.

FIGURE 4: C_{P1} trajectory in Task 1. FIGURE 7: C_{F2} trajectory in Task 1. FIGURE 5: C_{F1} trajectory in Task 1. FIGURE 8: V_{LP} trajectory in Task 1. FIGURE 6: C_{P2} trajectory in Task 1. FIGURE 9: V_{LF} trajectory in Task 1.

FIGURE 10: Pulp level trajectory in Task 1.

FIGURE 11: Tailing flow rate and air flow rate trajectories in Task 1.

In Task 2, the initial admissible controller was selected as

$$\begin{aligned}
 & [\hat{k}_{11}, \hat{k}_{12}, \hat{k}_{13}, \hat{k}_{14}, \hat{k}_{15}, \hat{k}_{16}] \\
 & = [-0.01, -0.01, -0.01, -0.01, 0, 0] \\
 & [\hat{k}_{21}, \hat{k}_{22}, \hat{k}_{23}, \hat{k}_{24}, \hat{k}_{25}, \hat{k}_{26}] \\
 & = [-0.01, -0.01, -0.01, -0.01, 0, 0]
 \end{aligned} \tag{63}$$

The resulting controller is

$$\begin{aligned}
 Q_T &= 965 + 10 \tanh(-74.4174x_1 - 12.8180x_2 \\
 &\quad + 221.4781x_3 - 5.0377x_4 - 111.5485x_5 \\
 &\quad - 45.4909x_6)
 \end{aligned}$$

FIGURE 12: Recovery and grade trajectories in Task 1.

FIGURE 13: Iteration of the weighting variables for tailing flow rate in Task 1.

$$\begin{aligned}
 Q_A &= 240 + 10 \tanh(-23.1681x_1 - 6.5684x_2 \\
 &\quad - 63.9545x_3 - 4.6639x_4 + 65.0465x_5 \\
 &\quad + 28.1158x_6)
 \end{aligned} \tag{64}$$

The simulation results of Task 2 are shown in Figures 15–25. Figures 15 and 16 show that the concentrations of the rich mineralogical species in the pulp and froth layer converge to their desired setting values with small steady-state bias caused by approximation error, while the recovery increased to 42% (Figure 23). This indicates that the desired recovery rate can be achieved by the setting and control of

FIGURE 14: Iteration of the weighting variables for air flow rate in Task 1.

FIGURE 15: C_{P1} trajectory in Task 2.

concentrations of each mineralogical species in the flotation cell. This is a useful rule in the optimal control of a flotation circuit composed of multiple cells. Figures 17, 18, 19, 20, and 21 demonstrate the variations of the concentration of poor mineralogical species in the pulp and froth layer, V_{LP} , V_{LF} , pulp level, respectively. Figure 22 indicates that the flow rate of tailing and air input are within their limitations; see [955, 975] and [230, 250]. Figures 24 and 25 present the convergence of the weighting variables for the approximation of control input, respectively.

In this study, it was observed that the selection of basis function affects the approximation accuracy, which could in turn affect the algorithm convergence and controller performance. Theoretically, more basis functions for the

FIGURE 16: C_{F1} trajectory in Task 2.

FIGURE 17: C_{P2} trajectory in Task 2.

value function and control input could decrease the approximation error. In the two tasks, different basis functions for the approximation of value function were used due to different optimization objectives. The selection of basis functions for the approximation of control input should have certain physical sense. The model structural information could provide more insights in the basis function selection. The magnitude of explosion noise is also an important factor. A small explosion cannot guarantee the rank condition, while an overlarge explosion noise could cause overexcitation.

6. Conclusions

An ADP-based optimal control approach for a sulfur flotation cell is proposed as an alternative when the parameters of

FIGURE 18: C_{F2} trajectory in Task 2. FIGURE 20: V_{LF} trajectory in Task 2. FIGURE 19: V_{LP} trajectory in Task 2.

FIGURE 21: Pulp level trajectory in Task 2.

the system model are unknown. The experimental study verified the feasibility and performance of the approach, which indicates the potential of applying bioinspired control methods in complex industrial processes. The advantage of ADP-based optimal control is that it does not rely on the full knowledge of the process model. The difficulties in the implementation of an ADP-based control are in the design of an initial admissible controller and the selection of basis functions. This is problem-specific and could benefit from the structural information of the process model and site experience. In addition, in this study, the system states are assumed available online which is not always the case, so the integration of state-observer [31] or machine-vision

[32] and ADP-based control for flotation process should be studied in future. Other future research topics include ADP robust to approximation error caused by limited learning time [33], the optimal control of a single flotation cell in more complicated cases, e.g., with state constraints, the optimal control of a flotation circuit composed of multiple flotation cells, the integration of ADP and classic model-based optimal control (e.g., MPC) and its application in flotation processes.

FIGURE 22: Tailing flow rate and air flow rate trajectories in Task 2.

FIGURE 23: Recovery and grade trajectories in Task 2.

Data Availability

The data supporting the conclusion is available via reader's request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

FIGURE 24: Iteration of the weighting variables for tailing flow rate in Task 2.

FIGURE 25: Iteration of the weighting variables for air flow rate in Task 2.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant number 61703441].

References

- [1] J. J. Liu and J. F. MacGregor, "Froth-based modeling and control of flotation processes," *Minerals Engineering*, vol. 21, no. 9, pp. 642–651, 2008.
- [2] B. J. Shean and J. J. Cilliers, "A review of froth flotation control," *International Journal of Mineral Processing*, vol. 100, no. 3–4, pp. 57–71, 2011.
- [3] M. Gharai and R. Venugopal, "Modeling of flotation process – an overview of different approaches," *Mineral Processing and Extractive Metallurgy*, 2015.
- [4] I. Jovanović and I. Miljanović, "Modelling of flotation processes by classical mathematical methods – a review," *Archives of Mining Sciences*, vol. 60, no. 4, pp. 905–919, 2015.
- [5] I. Jovanović, I. Miljanović, and T. Jovanović, "Soft computing-based modeling of flotation processes – A review," *Minerals Engineering*, vol. 84, pp. 34–63, 2015.
- [6] M. P. Jones, N. Johnson, E. Manlapig, and C. Thorne, *Mineral and Coal Flotation Circuits - Their Simulation and Control (Developments in Mineral Processing)*, vol. 3, Elsevier Scientific Publishing Co, 1981.
- [7] B. Gorain, J. Franzidis, and E. Manlapig, "The empirical prediction of bubble surface area flux in mechanical flotation cells from cell design and operating data," *Minerals Engineering*, vol. 12, no. 3, pp. 309–322, 1999.
- [8] J. Yianatos and F. Contreras, "Particle entrainment model for industrial flotation cells," *Powder Technology*, vol. 197, no. 3, pp. 260–267, 2010.
- [9] P. Koh and M. Schwarz, "CFD modelling of bubble - Particle attachments in flotation cells," *Minerals Engineering*, vol. 19, no. 6–8, pp. 619–626, 2006.
- [10] O. A. Bascur, *Modelling and Computer Control of a Flotation Cell*, University of Utah, Salt Lake City, Utah, USA, 1982.
- [11] I. Jovanović and I. Miljanović, "Contemporary advanced control techniques for flotation plants with mechanical flotation cells – A review," *Minerals Engineering*, vol. 70, pp. 228–249, 2015.
- [12] J. Na and G. Herrmann, "Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuous-time unknown nonlinear systems," *IEEE/CAA Journal of Automatica Sinica*, vol. 1, no. 4, pp. 412–422, 2014.
- [13] Y. Jiang, *Robust adaptive dynamic programming for continuous-time linear and nonlinear systems*, Polytechnic Institute of New York University, New York, NY, US, 2014.
- [14] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE Circuits and Systems Magazine*, vol. 9, no. 3, pp. 32–50, 2009.
- [15] Y. Lv, X. Ren, and J. Na, "Online optimal solutions for multiplayer nonzero-sum game with completely unknown dynamics," *Neurocomputing*, vol. 283, pp. 87–97, 2018.
- [16] Y. Jiang and Z.-P. Jiang, "Global adaptive dynamic programming for continuous-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2917–2929, 2015.
- [17] Y. Lv and X. Ren, "Approximate nash solutions for multiplayer mixed-zero-sum game with reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–12, 2018.
- [18] Z. Wen, L. J. Durllofsky, B. Van Roy, and K. Aziz, *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*, Wiley Online Library, 2013.
- [19] Q. Wei and D. Liu, "Adaptive dynamic programming for optimal tracking control of unknown nonlinear systems with application to coal gasification," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 4, pp. 1020–1036, 2014.
- [20] S. Xie, W. Zhong, K. Xie, R. Yu, and Y. Zhang, "Fair energy scheduling for vehicle-to-grid networks using adaptive dynamic programming," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1697–1707, 2016.
- [21] J. Li, T. Chai, F. L. Lewis, J. Fan, Z. Ding, and J. Ding, "Off-policy q-learning: Set-point design for optimizing dual-rate rougher flotation," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, pp. 4092–4102, 2018.
- [22] J. Li, B. Kiumarsi, T. Chai, F. L. Lewis, and J. Fan, "Off-policy reinforcement learning: optimal operational control for two-time-scale industrial processes," *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4547–4558, 2017.
- [23] J. Yianatos, F. Contreras, and F. Díaz, "GAS holdup and RTD measurement in an industrial flotation cell," *Minerals Engineering*, vol. 23, no. 2, pp. 125–130, 2010.
- [24] A. Casali, G. Gonzalez, H. Agosto, and G. Vallebuona, "Dynamic simulator of a rougher flotation circuit for a copper sulphide ore," *Minerals Engineering*, vol. 15, no. 4, pp. 253–262, 2002.
- [25] Y. Lv, J. Na, and X. Ren, "Online H_∞ control for completely unknown nonlinear systems via an identifier-critic-based ADP structure," *International Journal of Control*, vol. 92, no. 1, pp. 100–111, 2016.
- [26] B. Luo, H. Wu, T. Huang, and D. Liu, "Data-based approximate policy iteration for affine nonlinear continuous-time optimal control design," *Automatica*, vol. 50, no. 12, pp. 3281–3290, 2014.
- [27] J. Li, H. Modares, T. Chai, F. L. Lewis, and L. Xie, "Off-policy reinforcement learning for synchronization in multi-agent graphical games," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2434–2445, 2017.
- [28] M. Abu-Khalaf and F. L. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779–791, 2005.
- [29] Y. Jiang and Z. P. Jiang, "Robust adaptive dynamic programming and feedback stabilization of nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 882–893, 2014.
- [30] H. K. Khalil, *Nonlinear Systems*, Prentice Hall, New Jersey, NJ, USA, 3rd edition, 2002.
- [31] S. Wang, X. Ren, J. Na, and T. Zeng, "Extended-state-observer-based funnel control for nonlinear servomechanisms with prescribed tracking performance," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 1, pp. 98–108, 2017.
- [32] W. Gui, C. Yang, D. Xu, M. Lu, and Y. Xie, "Machine-vision-based online measuring and controlling technologies for mineral flotation—a review," *Acta Automatica Sinica*, vol. 39, no. 11, pp. 1879–1888, 2013.
- [33] S. Wang, J. Na, and X. Ren, "RISE-based asymptotic prescribed performance tracking control of nonlinear servo mechanisms," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 12, pp. 2359–2370, 2018.

Submit your manuscripts at
www.hindawi.com