Operational Control of Mineral Grinding Processes Using Adaptive Dynamic Programming and Reference Governor

Xinglong Lu, *Student Member, IEEE*, Bahare Kiumarsi, *Member, IEEE*, Tianyou Chai, *Fellow, IEEE*, Yi Jiang, *Student Member, IEEE*, and Frank L. Lewis, *Fellow, IEEE*

***Abstract*—Operation performance of mineral grinding processes is measured by the grinding product particle size and the circulating load, as two of the most crucial operational indices that measure the product quality and operation efficiency, respectively. In this paper, a data-driven method is proposed for the operational control design of mineral grinding processes with input constraints. A reference governor is introduced to take into account the input constraints and the infeasible setpoint issue. The reference governor generates feasible setpoints that keep control inputs within allowed regions. The lookup table embedded in the reference governor mapping steady-state outputs to inputs provides feasible setpoints for output regulation and baseline for inputs. An *ad hoc* optimization guarantees that the input constraints are not violated, with the priority of regulating the grinding product particle size if regulation of both indices is not feasible. Since the dynamic model of the controlled plant is complicated because of the strongly nonlinear and intricately coupled nature of ball mills and hydrocyclones, a novel policy iteration algorithm is proposed for optimal regulator design without system modeling. Simulation results comparing performances of a mineral grinding process with and without the reference governor show the effectiveness of the proposed method.**

***Index Terms*—Adaptive dynamic programming (ADP), input constraints, mineral grinding processes, operational control, reference governor.**

I. Introduction

MINERAL grinding is one of the most energy consuming and costly procedures in the mineral processing industry, and hence, improving product quality and operation efficiency with limited consumption of energy has attracted significant interest in industrial process control community [1]–[5]. Regulation of the grinding product particle size (GPPS) and circulating load (CL), as two crucial operational indices in mineral grinding processes, is the key to achieve favorable mineral concentrate grade, metal recovery rate, and operation efficiency, which is known as operational control [6].

Because of the scale of equipment in mineral grinding processes, multiple control actuations and control loops have been used. Cascade control is widely used in mineral grinding processes, which allocates different control tasks for two loops. In the inner loop, the control objective is to design regulators for fresh ore feed rate (FOFR) and sump water flow rate (SWFR), respectively. In the outer loop, operational control design should be performed to eliminate tracking errors of the operational indices with respect to setpoints of FOFR and SWFR. The operational control loop performance has a strong influence on the economic indices of mineral processing plants. Compared with economic loss caused by deteriorated control performance in the inner loop, operating at wrong setpoints of FOFR and SWFR could lead to much more negative consequences [3], such as poor product quality, low efficiency, and operational failures in the worst-case scenario.

With the ever-fierce global competition, more stringent requirements on concentrate grade, consumption of energy and fresh ore, and operation efficiency have to be satisfied. Due to actuator saturation and operation safety in practice, constraints are imposed on the inputs, *namely* setpoints of FOFR and SWFR, and the outputs, *namely* GPPS and CL. Constraints on the outputs show the demands of satisfactory quality of product and high operation efficiency. Actuator saturation is commonly seen in industrial practical operations and the control inputs that violate those constraints are not implementable due to physical limitations. More specifically, tracking errors can diverge or become disproportionately large and persist for a long time which, for example, happens in the integral windup situation.

Design approaches have been explored for input-constrained operational control in mineral grinding processes, such as real-time optimization (RTO) [7], model predictive control (MPC) [8], [9], and expert system (ES) [10], [11]. In [7], a setpoint updating method with real-time linear programming is

Manuscript received July 6, 2018; accepted August 14, 2018. Date of publication September 3, 2018; date of current version April 3, 2019. This work was supported in part by the China NSFC Grant 61633007. Paper no. TII-18-1749. *(Corresponding author: Tianyou Chai.)*

X. Lu, T. Chai, and Y. Jiang are with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China (e-mail: xinglonglu@stumail.neu.edu.cn; tychai@mail.neu.edu.cn; JY369356904@163.com).

B. Kiumarsi is with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: b_kiomarsi@yahoo.com).

F. L. Lewis is with UTA Research Institute, University of Texas at Arlington, Fort Worth, TX 76118 USA, and also with the State Key Laboratory of Synthetical Automation or Process Industries, Northeastern University, Shenyang 110819, China (e-mail: lewis@uta.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2018.2868473

1551-3203 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

proposed based on a linearized dynamic model of mineral grinding processes. The proposed optimization cannot be performed until a steady state is achieved, which typically leads to time-consuming convergence. MPC is a widely used control framework for mineral grinding processes, since it provides optimization-wise solution and the constraints are well dealt with. In [8], a nonlinear MPC design is reported to deal with parameter perturbation and input constraints. A velocity-form nonlinear MPC paradigm is proposed in [9] to solve the reference tracking problem of grinding mill circuits. However, establishing a sufficiently accurate nonlinear model and identifying its parameters for mineral grinding processes are challenging tasks when RTO and MPC are used. In [10], if–then rule statements are developed to modify the setpoints of the inner loop. In [11], fuzzy logic and online optimization techniques are integrated to form a supervisory control method. But it is difficult to establish analytical properties and assess the optimality of the ES-based designs.

The challenge with the model-based approaches is the lack of explicit model to describe the input–output relationship. Compromises have to be made to perform model-based designs, e.g., using steady-state static model or linearized model. Therefore, data-driven control approaches without system modeling have received considerable attention. Recent works on data-driven control and its industrial applications have been reviewed in [12] and[13]. Adaptive dynamic programming (ADP), which provides a systematic frame for model-free adaptive optimal designs, is a promising candidate to solve input-constrained operational control for complex industrial processes [14]–[19]. In [14], a data-driven optimization method is proposed for hematite grinding processes based on Q-learning and if–then rules to satisfy input constraints. In [15], a data-driven control method based on heuristic dynamic programming is developed to solve the setpoint tracking problem of a class of industrial processes that do not possess a static inner loop. Barrier functions are used to prevent the input constraints from being violated. In [16], a penalty function is included in the performance index of the optimization problem of the flotation industrial process that guarantees the boundedness of the control inputs. In [17], a data-driven fault-tolerant control method is proposed for time-delay Markov jump systems with Ito stochastic process and output disturbance based on the sliding mode observer. A data-driven control approach based on action-dependent heuristic dynamic programming is proposed for the air-breathing hypersonic vehicle tracking problem in [18]. And in [19], a stable iterative ADP is developed to solve optimal temperature control problems for water–gas shift reaction systems.

The aforementioned works only develop bounds for inputs such that the input constraints are not violated. However, a given setpoint may be unable to be followed by the controlled plant through constrained inputs, in which case the setpoint is called an infeasible setpoint. To the best knowledge of the authors, there is no method that takes into consideration of the infeasible setpoint issue in a model-free design. Therefore, a data-driven model-free method is proposed in this paper to address the input constraints considering infeasible setpoints.

In this paper, with an enormous amount of historical operation data, a lookup table is developed, mapping steady-state outputs onto corresponding steady-state inputs. Based on the lookup table, a reference governor is designed to solve for the feasible setpoints that are as follows: 1) closest to prescribed setpoints and 2) unable to violate the input constraints given fixed control policy. Instead of following the prescribed setpoints that may lead to input constraint violation, the controlled plant is regulated to follow the feasible setpoints. Introducing the lookup table becomes advantageous since it avoids the need of an explicit system model, which is typically used to calculate the maximal output admissible set [20], [21] in standard reference governor designs.

After the feasible setpoints are determined, the corresponding steady-state inputs can be obtained through the lookup table. Since the references for inputs and outputs are known, a properly defined optimal control problem can be formulated to solve the setpoint following problem, of which solution automatically satisfies the internal model principle [22]. To solve the optimal control problem in a model-free manner without involving a system identification neural network (NN) model as in [16],[23], and[24], a new policy iteration called approximate policy iteration is proposed to relax the requirement of system modeling. Future state is predicted taking advantage of the fact that the error dynamics are affine in control.

The contributions of this paper are as follows. First, a reference governor based on a lookup table mapping steady-state outputs onto corresponding steady-state inputs is proposed to modify the reference signal to satisfy the input constraints without system modeling. Then, a data-driven strategy based on the reference governor and ADP is proposed giving references of the inputs and outputs. Last but not least, a novel approximate policy iteration algorithm is proposed to relax the standard policy iteration algorithm from requiring a system model, thereby achieving a model-free control design.

II. Problem Formulation

In this section, one-stage closed-circuit mineral grinding processes are first introduced. Then, the input constrained operational control problem is formulated.

A. Description of Mineral Grinding Processes

The one-stage closed-circuit mineral grinding process considered in this paper is shown in Fig. 1. Fresh ore from ore bin are fed into the ball mill through a conveyor. Size reduction of the fresh ore is done by impact as fresh ore and grinding media fall from near the top of the cylindrical shell. The slurry is discharged into sump and pumped to the hydrocyclone. Coarse ore with higher settling velocity are deposited on the bottom and transported back to the ball mill for further grinding. Fine ore with lower settling velocity are carried away by overflow for follow-up mineral processing procedures.

Two key operational indices that are crucial to the operation of the mineral grinding process include: 1) GPPS (denoted by r_1), which represents the particle size distribution of the fine ore

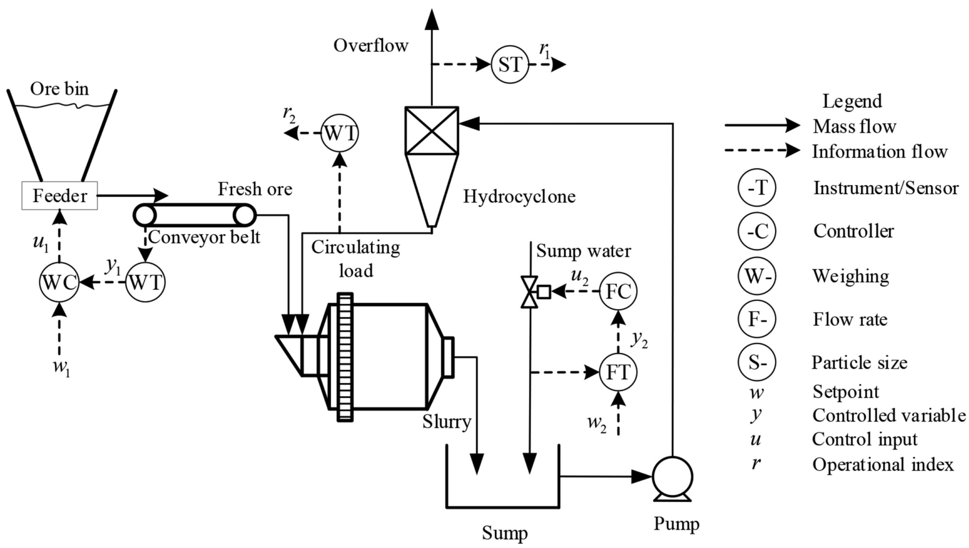


Fig. 1. One-stage closed-circuit mineral grinding process.

in the overflow and 2) CL (denoted by r_2), which represents the discharge rate of the coarse ore. These two variables dictate the product quality and the operation efficiency of mineral grinding processes, respectively. These two operational indices are mostly influenced by the fresh ore feed rate (denoted by y_1) and the sump water feed rate (denoted by y_2). A vibrating feeder and a control valve are used as actuators to regulate the FOFR and SWFR. The inner loop controllers for the actuators, typically PI controllers, are used to regulate the FOFR and SWFR to track their setpoints denoted by w_1 and w_2, respectively. The objective in mineral grinding processes is to regulate the operational indices to track prescribed setpoints (denoted by r_1^* and r_2^*, respectively).

B. Operational Control Problem Formulation for Mineral Grinding Processes

Since the time scales of the inner and outer loops are separable and the inner loop dynamics are much faster than the outer loop, PI controllers in the inner loop can guarantee that the inner closed loop reaches its steady state quickly within a single sample time of the outer loop, namely, $y_1 = w_1, y_2 = w_2$. Therefore, in the outer loop control design, the dynamics of the inner closed loop can be omitted, and w_1 and w_2 become the new inputs. Then, the dynamic model in the outer loop is

$$r(k+1) = f_1(r(k), \ldots, r(k-n+1) \quad w(k-d+1), \ldots, w(k-n+1)) \tag{1}$$

where n and d are the order and the delay of the model, respectively, $r(k) = (r_1(k), r_2(k))^{\mathrm{T}}$, $w(k) = (w_1(k), w_2(k))^{\mathrm{T}}$, $f_1(\cdot)$ is an unknown nonlinear function that represents the input–output dynamic relationship of the controlled plant, and k is the discrete-time index of the outer loop.

The constraints are imposed on both inputs and outputs of the controlled plant. The operational control problem of mineral grinding processes can be formulated as follows: consider the controlled plant (1) and design the control input $w(k)$ to assure that

$$|r_i^* - r_i(k)| < \epsilon_i, k \to +\infty, i = 1, 2$$

while satisfying the constraints

$$r_i(k) \in [r_i^{\min}, r_i^{\max}], w_i(k) \in [w_i^{\min}, w_i^{\max}], i = 1, 2$$

where ϵ_1 and ϵ_2 are tolerance errors in a steady state.

III. Reference Governor Design for Mineral Grinding Processes

In this section, a reference governor-based control strategy is first described, followed by the reference governor design.

The proposed operational control strategy for mineral grinding processes with input constraints is presented and shown in Fig. 2. The control strategy is composed of two control loops, i.e., the operational control for operational indices setpoint tracking, and the inner loop control for actuator regulation. The actuators in the inner loop are controlled by PI controllers. In the outer loop, a reference governor is used to determine feasible setpoints $r_g(k)$ from the prescribed setpoints $r^*(k)$ that keep the control inputs within allowed regions. At the same time, steady-state inputs $w_s(k)$ corresponding to the feasible setpoints are provided by the lookup table embedded therein. Then, the actor-critic structure consisting of two neural networks is used to implement the optimal regulator $w_{\mathrm{opt}}(k)$ based on a novel approximate policy iteration algorithm.

The control inputs are determined by

$$w(k) = w_s(k) + w_{\mathrm{opt}}(k) \tag{2}$$

where $w_s(k)$ and $w_{\mathrm{opt}}(k)$ will be designed specifically.

A. Reference Governor-Based Design of $w_s(k)$

To deal with the input constraints, a lookup table is first developed to map the steady-state output $r(k)$ onto the corresponding

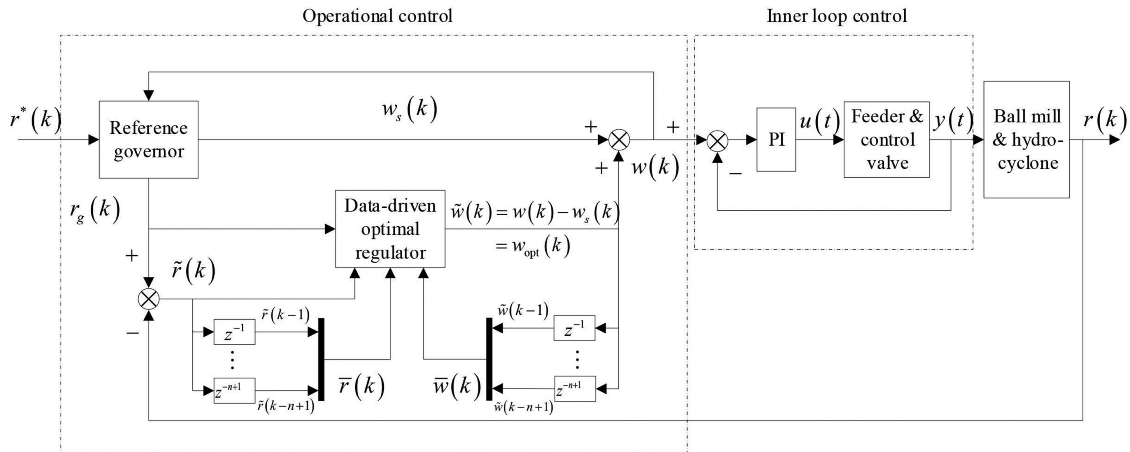


Fig. 2. Operational control strategy.

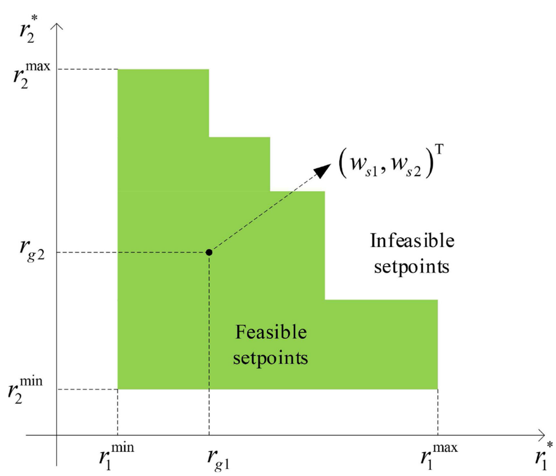


Fig. 3. Lookup table embedded in the reference governor.

steady-state input $w(k)$ as

$$w(k) = \phi(r(k)). \tag{3}$$

The lookup table is shown in Fig. 3. Through the steady-state output–input pairing, it is easy to divide the prescribed setpoints $r^*(k)$ into two categories with respect to the input and output constraints—feasible setpoints and infeasible ones. Feasible setpoints are the setpoints of operational indices in a steady-state operation of mineral grinding processes, that are within allowed output constraints and reachable without violating the input constraints. Otherwise, the setpoints are infeasible. The set composed of feasible setpoints is a closed polygon whose edges are determined by the input and output constraints. The reason why the lookup table can be developed is that a massive amount of historical operation data are stored. In the case that no explicit models exist for describing the input–output relationship, historical data of steady-state operation become vitally important. Pretreatments like low-pass filtering should be carried out before the lookup table is developed to improve the usefulness of the data.

Remark 1: Since the lookup table provides the references for inputs, the lookup table works as the internal model in sense of solving an output regulation problem. Therefore, if the mineral grinding process deviates from the operating point on which the lookup table is developed, the lookup table should be recalibrated accordingly.

The feasible setpoints are determined by

$$r_g(k) = (\beta_1(k) r_1^*(k), \beta_2(k) r_2^*(k))^{\mathrm{T}}, W_s = \phi(R_s) \tag{4}$$

where $r_g(k)$ are the feasible setpoints, W_s and R_s are the output and input subspaces that are composed of feasible setpoints and the corresponding steady-state inputs, respectively, $\phi(\cdot)$ is the output-to-input mapping, $\beta_{1,2}$ are the scalar gains. Optimization can be carried out to determine the feasible setpoints $r_g(k)$ that are closest to the original setpoints $r^*(k)$. At the same time, the corresponding inputs should be given through the lookup table with $w_s(k) = \phi(r_g(k))$.

Once the feasible setpoints are determined, the inputs can be obtained immediately with fixed control policy $w_{\text{opt}}(\cdot)$. Optimization of the scalar gains should not only be about searching for the closest feasible setpoints, but also be carried out with one-step prediction of the inputs such that the input constraints will not be violated. On the other hand, grinding product quality is more important than the operation efficiency when a compromise has to be made between these two operational indices. Therefore, the *ad hoc* optimization to determine feasible setpoints is defined as

$$\begin{aligned}\beta_1(k) = \max\Big\{&\alpha_1 \in [0,1] \,|\, (\alpha_1 r_1^*(k), \alpha_2 r_2^*(k))^{\mathrm{T}} \in R_s\\ &\phi\left((\alpha_1 r_1^*(k), \alpha_2 r_2^*(k))^{\mathrm{T}}\right) + w_{\text{opt}}(k)\\ &\in W_s, \alpha_2 \in [0,1]\Big\}\end{aligned} \tag{5}$$

$$\begin{aligned}\beta_2(k) = \max\Big\{&\alpha_2 \in [0,1] \,|\, (\beta_1(k) r_1^*(k), \alpha_2 r_2^*(k))^{\mathrm{T}} \in R_s\\ &\phi\left((\beta_1(k) r_1^*(k), \alpha_2 r_2^*(k))^{\mathrm{T}}\right) + w_{\text{opt}}(k) \in W_s\Big\}.\end{aligned} \tag{6}$$

The optimizations should be carried out strictly in the following order: first, (5), and then, (6). The optimization of scalar gain $\beta_1(k)$ is carried out prior to the one of scalar gain $\beta_2(k)$ such that $r_{g1}(k)$ is closest to the its setpoint $r_1^*(k)$ with priority. Then, $w_s(k)$ is obtained as

$$\begin{aligned} r_g(k) &= \left(\beta_1(k) r_1^*(k), \beta_2(k) r_2^*(k)\right)^{\mathrm{T}} \\ w_s(k) &= \phi\left(r_g(k)\right). \end{aligned} \tag{7}$$

B. New Formulation With Embedded Reference Governor

The feasible setpoints $r_g(k)$ and their corresponding steady-state inputs $w_s(k)$ are given by the reference governor. The tracking errors of the outputs and inputs can be represented as

$$\tilde{r}(k) = r(k) - r_g(k), \tilde{w}(k) = w(k) - w_s(k). \tag{8}$$

Then, by defining

$$\begin{aligned} \bar{r}(k) &= \left[\tilde{r}^{\mathrm{T}}(k-1), \ldots, \tilde{r}^{\mathrm{T}}(k-n+1)\right]^{\mathrm{T}} \\ \bar{w}(k) &= \left[\tilde{w}^{\mathrm{T}}(k-1), \ldots, \tilde{w}^{\mathrm{T}}(k-n+1)\right]^{\mathrm{T}} \\ x(k) &= \left[\tilde{r}^{\mathrm{T}}(k), \bar{r}^{\mathrm{T}}(k), \bar{w}^{\mathrm{T}}(k)\right]^{\mathrm{T}} \end{aligned}$$

the error dynamics state-space representation is derived as

$$\begin{aligned} x(k+1) &= \begin{bmatrix} f_2\left(\tilde{r}(k), \bar{r}(k), H\bar{w}(k)\right) \\ F_0\bar{r}(k) + G_0\tilde{r}(k) \\ F_0\bar{w}(k) \end{bmatrix} + \begin{bmatrix} 0_{2\times 2} \\ 0_{2(n-1)\times 2} \\ G_0 \end{bmatrix} \tilde{w}(k) \\ &\triangleq f(x(k)) + G\tilde{w}(k) \end{aligned} \tag{9}$$

where $x(k) \in \mathbb{R}^{2(2n-1)}$ is the new state, $\bar{r}(k)$ is the vector of the tracking errors of operational indices, $f_2(\cdot)$ represent the nonlinear functions $f_1(\cdot)$ with $\tilde{r}(k), \bar{r}(k),$ and $\bar{w}(k)$ as new arguments, and

$$F_0 = \begin{bmatrix} 0_{2\times 2} & & \cdots & & 0_{2\times 2} \\ I_{2\times 2} & 0_{2\times 2} & & & \\ 0_{2\times 2} & I_{2\times 2} & \ddots & & \vdots \\ \vdots & & \ddots & & \\ 0_{2\times 2} & \cdots & 0_{2\times 2} & I_{2\times 2} & 0_{2\times 2} \end{bmatrix}_{2(n-1)\times 2(n-1)}$$

$$\begin{aligned} G_0 &= \begin{bmatrix} I_{2\times 2} & 0_{2\times 2(n-2)} \end{bmatrix}^{\mathrm{T}}_{2(n-1)\times 2} \\ G &= \begin{bmatrix} 0_{2\times 2n} & G_0^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}_{2(2n-1)\times 2} \\ H &= \begin{bmatrix} 0_{2(n-d+1)\times 2(d-2)} & I_{2(n-d+1)\times 2(n-d+1)} \end{bmatrix}. \end{aligned}$$

Remark 2: Note that H is ill-defined if $d = 1$. Consequently, the state-space representation (9) is valid only when $d \geq 2$. If $d = 1$, a unit time delay can be introduced between the controller and the controlled plant during control design to put the system in the representation (9). And this extra time delay is removed when the controller is implemented that does not influence the controlled plant.

IV. Data-Driven Optimal Regulator Design for Mineral Grinding Processes

In this section, a new policy iteration algorithm is proposed to design a data-driven optimal regulator without system modeling of mineral grinding processes. Then, online implementation with NNs is described.

A. Data-Driven Optimal Design of $w_{\mathrm{opt}}(k)$

To design the optimal regulator $w_{\mathrm{opt}}(k)$, the following optimal control problem can be formulated:

$$\min_{\tilde{w}(\cdot)} V(x(k)) = \min_{\tilde{w}(\cdot)} \sum_{i=k}^{\infty} \left(x^{\mathrm{T}}(i) Q x(i) + \tilde{w}^{\mathrm{T}}(i) R \tilde{w}(i)\right) \tag{10}$$

$$\text{s.t. } x(k+1) = f(x(k)) + G\tilde{w}(k) \tag{11}$$

where Q and R are positive definite. The cost function (10) can be rewritten in a backwards-in-time manner as

$$V(x(k)) = \rho(k) + V(x(k+1)) \tag{12}$$

where $\rho(k) = x^{\mathrm{T}}(k) Q x(k) + \tilde{w}^{\mathrm{T}}(k) R \tilde{w}(k)$ is the cost at discrete time k and $V(0) = 0$. From Bellman's principle of optimality, the optimal cost function $V^*(x(k))$ satisfies the following Bellman equation:

$$V^*(x(k)) = \min_{\tilde{w}(k)} \left\{\rho(k) + V^*(x(k+1))\right\}. \tag{13}$$

The optimal control should satisfy the first-order necessary condition of optimality and can be obtained as

$$\tilde{w}^*(x(k)) = -\frac{1}{2} R^{-1} G^{\mathrm{T}} \frac{\partial V^*(x(k+1))}{\partial x(k+1)}. \tag{14}$$

Thus, $w_{\mathrm{opt}}(k)$ is designed as $w_{\mathrm{opt}}(k) = \tilde{w}^*(x(k))$. The following policy iteration algorithm 1 [25], [26] can be used to find $V^*(\cdot)$ and $w_{\mathrm{opt}}(\cdot)$ iteratively.

Algorithm 1: Policy Iteration.

Initialization. Select any initial stabilizing policy $w_{\mathrm{opt}}^0(\cdot)$.
Policy Evaluation. Determine the value function of the current control policy using

$$V^{j+1}(x(k)) = \rho(k) + V^{j+1}(x(k+1)). \tag{15}$$

Policy Improvement. Determine an improved policy using

$$w_{\mathrm{opt}}^{j+1}(x(k)) = -\frac{1}{2} R^{-1} G^{\mathrm{T}} \frac{\partial V^{j+1}(x(k+1))}{\partial x(k+1)}. \tag{16}$$

Stop. Stop when $\left\|V^{j+1}(x(k)) - V^j(x(k))\right\|$ is less than a prescribed tolerance.

It has been proven in [26] that for discrete-time nonlinear systems, the stability of the system can be guaranteed with policy iteration algorithm. However, the policy iteration algorithm 1 needs a system model for the state prediction from $x(k)$ to $x(k+1)$. This is typically done by introducing a system identification NN [16], [23], [24]. Note that the error dynamics are affine in $\tilde{w}(k)$ and G is assumed known. By examining the following equations:

$$x^0(k+1) = f(x(k)) + Gw^0_{\text{opt}}(k)$$
$$x^{j+1}(k+1) = f(x(k)) + Gw^{j+1}_{\text{opt}}(k)$$

where $x^j(k+1)$ denotes the future state with $w^j_{\text{opt}}(k)$ implemented, the following state prediction approach can be obtained without requiring a system model:

$$x^{j+1}(k+1) = x^0(k+1) + G\left(w^{j+1}_{\text{opt}}(k) - w^0_{\text{opt}}(k)\right).$$

To avoid the requirement of the system dynamics, the following approximate policy iteration algorithm 2 is proposed.

Algorithm 2: Approximate Policy Iteration.

Initialization

1) Select any initial stabilizing policy $w^0_{\text{opt}}(\cdot)$ and apply $w^0_{\text{opt}}(x(k))$ to the controlled plant;
2) For each k, set $j=0$, $x^0(k+1) = x(k+1)$, $w^0_{\text{opt}}(x(k)) = w_{\text{opt}}(k)$;

Policy Evaluation

3) Determine the value function of control policy $w^j_{\text{opt}}(\cdot)$

$$V^{j+1}(x(k)) = x^{\mathrm{T}}(k)Qx(k) + w^{j\mathrm{T}}_{\text{opt}}(x(k))Rw^j_{\text{opt}}(x(k)) + V^{j+1}\left(x^j(k+1)\right). \tag{17}$$

Policy Improvement

4) Determine an improved policy $w^{j+1}_{\text{opt}}(\cdot)$

$$w^{j+1}_{\text{opt}}(x(k)) = -\frac{1}{2}R^{-1}G^{\mathrm{T}}\frac{\partial V^{j+1}\left(x^j(k+1)\right)}{\partial x^j(k+1)}. \tag{18}$$

State Prediction

5) Update the state under control policy $w^{j+1}_{\text{opt}}(\cdot)$

$$x^{j+1}(k+1) = x^0(k+1) + G\left(w^{j+1}_{\text{opt}}(x(k)) - w^0_{\text{opt}}(x(k))\right). \tag{19}$$

Check Convergence

6) Check $\left\|V^{j+1}(x(k)) - V^j(x(k))\right\| < \epsilon^c$, where ϵ^c is a small positive tolerance. If it is untrue, set $j=j+1$ and then go to step 3). If true, set $w^{j+1}_{\text{opt}}(\cdot) \to w^{N_k}_{\text{opt}}(\cdot)$ and apply $w^{N_k}_{\text{opt}}(x(k+1))$ to the controlled plant. Set $k=k+1$ and then go back to step 2).

Remark 3: In contrast to the standard policy iteration algorithm, the iterations are run after the initial policy is implemented in approximate policy iteration. In other words, the policy $w^{N_{k+1}}_{\text{opt}}(\cdot)$ that should be implemented at discrete-time $k+1$ is approximated by $w^{N_k}_{\text{opt}}(\cdot)$.

B. Online Implementation With NNs

To implement Algorithm 2, actor–critic learning structure [25], [27] composed of two NNs is used to solve functionals (17) and (18). The critic and actor NNs are used to approximate the optimal cost function $V^*(\cdot)$ and the optimal control policy $w_{\text{opt}}(\cdot)$, respectively. Let the numbers of the hidden layer neurons, the input-to-hidden and hidden-to-output weights, and the activation functions of critic and actor NNs be denoted by $n^c, n^a, V^c, V^a, W^c, W^a, \sigma^c(\cdot)$, and $\sigma^a(\cdot)$, respectively. Then, the output of the critic NN is given by

$$\hat{V}(x(k)) = \sum_{j=1}^{n^c} W^c_j \sigma^c_j\left(\sum_{h=1}^{4n-2} V^c_{jh}x_h(k)\right) \triangleq \sum_{j=1}^{n^c} W^c_j\sigma^c_j(V^c x(k)) \triangleq W^c\sigma^c(k) \tag{20}$$

and the output of the actor NN is given by

$$\hat{w}_{\text{opt}(l)}(x(k)) = \sum_{j=1}^{n^a} W^a_{lj}\sigma^a_j\left(\sum_{h=1}^{4n-2} V^a_{jh}x_h(k)\right) \triangleq \sum_{j=1}^{n^a} W^a_{lj}\sigma^a_j(V^a x(k)) \triangleq W^a_{l\cdot}\sigma^a(k), l=1,2 \tag{21}$$

where the subscript j of a vector denotes the jth element of the vector, the subscript jh denotes the element in jth row and hth column of the matrix, the subscript $l\cdot$ denotes the l row of the matrix.

The weights of the critic NN are updated to minimize the approximation error of the optimal cost function. The approximation error of the critic NN is defined with a temporal difference error [27],

$$e^c = x^{\mathrm{T}}(k)Qx(k) + w^{j\mathrm{T}}(k)Rw^j(k) + \hat{V}\left(x^j(k+1)\right) - \hat{V}(x(k)). \tag{22}$$

Therefore, weights of the critic NN are updated to minimize the following squared temporal difference error:

$$E^c = \frac{1}{2}(e^c)^2. \tag{23}$$

An improved policy should minimize the value function based on the Bellman equation (13). Therefore, the weights of the actor NN are tuned to minimize $V^{j+1}\left(x^j(k+1)\right)$.

It has been shown in [28] and[29] that NNs still possess the approximation ability when only the hidden-to-output weights W^c and W^a are tuned during the training. Therefore, the input-to-hidden weights V^c and V^a are assigned randomly and kept constant. Then, the tuning laws for the weights of the critic and actor NNs are given as follows with denominator layout notation:

$$\Delta W^c = -l^c\frac{\partial E^c}{\partial W^{c\mathrm{T}}} \tag{24}$$

$$\Delta W^a_{l\cdot} = -l^a\partial\left(x^{\mathrm{T}}(k)Qx(k) + w^{j\mathrm{T}}(k)Rw^j(k) + V^{j+1}\left(x^j(k+1)\right)\right)/\partial W^{a\mathrm{T}}_{l\cdot} \tag{25}$$

Algorithm 3: Operational Control of Mineral Grinding Processes With Input Constraints.

Lookup Table Creation

1) Establish the lookup table with historical steady-state operation data;

Initialization

2) Determine the order n and the delay d of the controlled plant (1).
3) Calculate G with (9). Determine whether a unit time delay should be introduced with d. Choose Q and R based on the grinding technology. Initialize the structures of the critic and actor NNs.
4) Obtain a stabilizing control policy $w_{\text{opt}}^{0}(\cdot)$ with offline training based on field experience.

Online Training

5) Choose any constant feasible setpoints r^* and determine $w_s = \phi(r^*)$.
6) Set $j = 0$.
7) Determine $w_{\text{opt}}^{j}(k)$ using (21).
8) Determine $w(k)$ with (2) and (4) and apply it to the controlled plant to obtain $x^0(k+1)$.
9) Train the weights of the critic NN using (22)–(24) and (26) until convergence.
10) Update the weights of the actor NN using (25) and (27).
11) Set $j = j+1$, update $x^{j+1}(k+1)$ with (19), and return to step 9) until $\|V^{j+1}(x(k)) - V^{j}(x(k))\| < \epsilon^c$.
12) Set $w_{\text{opt}}^{j+1}(\cdot) \to w_{\text{opt}}^{N_k}(\cdot)$ and apply $w_{\text{opt}}^{N_k}(x(k+1))$ to the controlled plant. Set $k = k+1$, and then, return to step 6) until tracking is achieved;

Online Implementation

13) Fix the weights of both NNs;
14) Determine $\beta_{1,2}$ using (5) and (6) if r^* is changed or the input constraints will be violated.

where l^c and l^a are the learning rates of the critic and actor NNs. From the chain rule, it follows that

$$\begin{aligned}\Delta W^c &= -l^c \frac{\partial E^c}{\partial e^c}\frac{\partial e^c}{\partial W^c}\\ &= -l^c e^c\left(\sigma^c\left(V^c x^j(k+1)\right) - \sigma^c\left(V^c x(k)\right)\right) \end{aligned} \tag{26}$$

$$\begin{aligned}\Delta W_{l\cdot}^a &= -l^a\left(\frac{\partial \hat{V}(x(k))}{\partial w^j(k)}\right)^{\mathrm{T}} \frac{\partial w^j(x(k))}{\partial w_l^j(x(k))}\frac{w_l^j(k)}{\partial W_{l\cdot}^a}\\ &= -l^a\left(2Rw^j(x(k)) + G^{\mathrm{T}}V^{c\mathrm{T}}\frac{\partial \sigma^c\left(V^c x^j(k+1)\right)}{\partial\left(V^c x^j(k+1)\right)}W^{c\mathrm{T}}\right)^{\mathrm{T}}\\ &\quad\times \frac{\partial w^j(x(k))}{\partial w_l^j(x(k))}\sigma^a\left(V^a x(k)\right). \end{aligned} \tag{27}$$

Then, the proposed data-driven input-constrained operational control method for mineral grinding processes is summarized as shown in Algorithm 3.

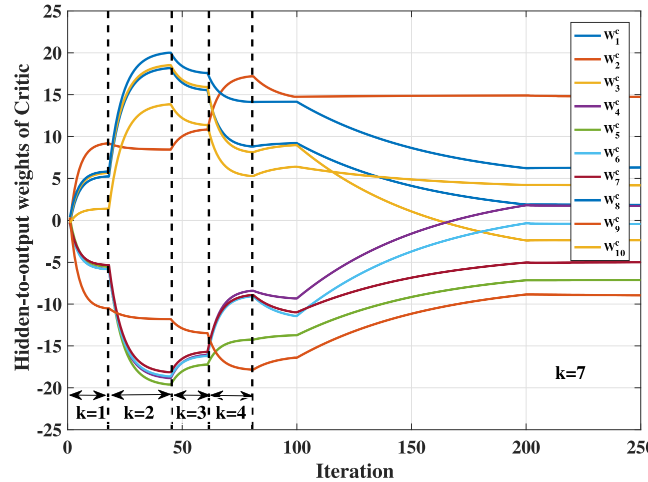


Fig. 4. Weights W^c of the critic NN.

V. Simulation Results

In this section, simulation results are shown to compare the performances with and without the reference governor. To emulate the dynamics of mineral grinding processes, a simulation model that characterizes the physical mechanisms of the ball mill and the hydrocyclone is used as testbed based on the modeling work in [30]–[32]. The simulation model is discussed in the Appendix and more detailed discussions can be found in [30]–[32] and the references therein.

The input and output constraints are specified according to the grinding technology as

$$\begin{aligned}&r_1^{\min} = 55, r_1^{\max} = 61, r_2^{\min} = 0.1, r_2^{\max} = 0.5\\ &w_1^{\min} = 0.1, w_1^{\max} = 0.4, w_2^{\min} = 0.1, \text{and } w_2^{\max} = 1.\end{aligned}$$

The lookup table is developed from historical operation data. To simplify the lookup table and lower the work load, the space of W_s is discretized. The GPPS is sampled every 0.1 and the CL is sampled every 0.01. A total of 2400 discretized points are contained in the lookup table.

In the simulations, sample time is 1 min. $n = 2, d = 1$ are determined with the time behavior of unit step response. Using (9), G is obtained as

$$G = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}^{\mathrm{T}}.$$

The parameters of the algorithm are chosen as $Q = I_{6\times 6}, R = I_{2\times 2}, l^c = l^a = 0.05, n^c = n^a = 10, \sigma^c(\cdot) = \sigma^a(\cdot) = \tanh(\cdot)$. The weights of critic and actor NNs are randomly set in the intervals $(-0.5, 0.5)$ and $(-0.1, 0.1)$, respectively, and the input-to-hidden weights are kept constant.

The online training of the critic and actor NNs are shown in Figs. 4–6. As can be seen, the hidden-to-output weights of both the critic and actor NNs converge to constant values within 250 iterations and 7 sample times.

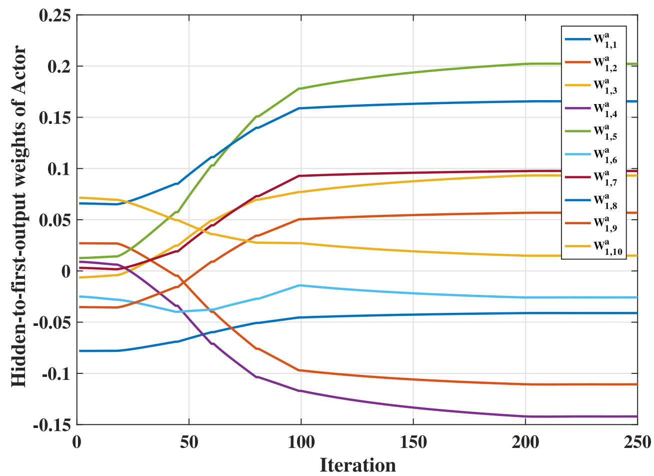


Fig. 5. Weights W_1^a of the actor NN.

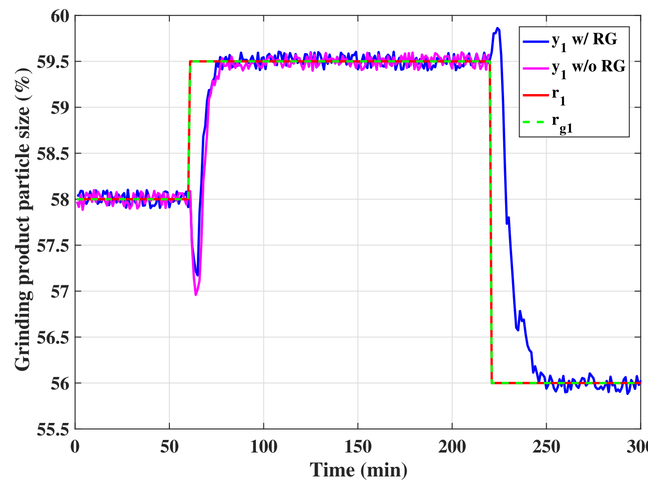


Fig. 7. Trajectories of GPPS, setpoint, and RG's feasible setpoint.

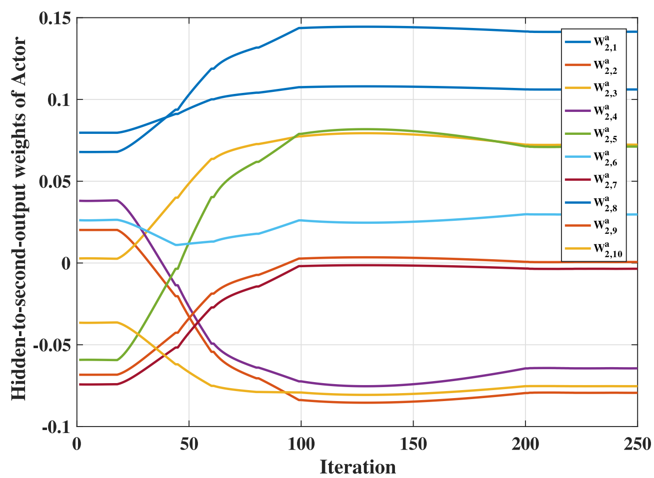


Fig. 6. Weights W_2^a of the actor NN.

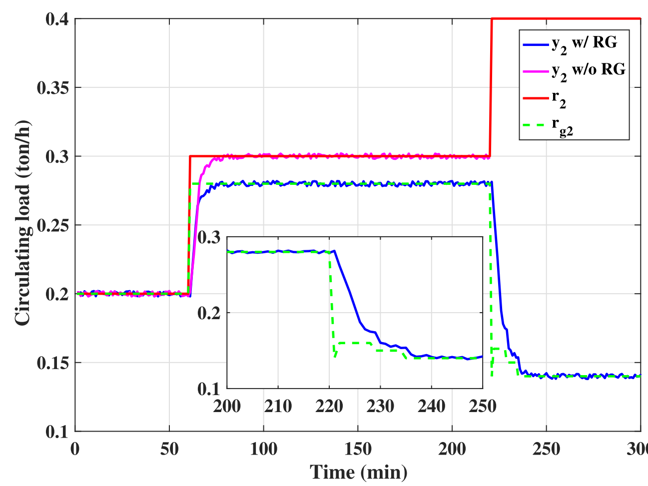


Fig. 8. Trajectories of CL, setpoint, and RG's feasible setpoint.

After obtaining the optimal policy $w_{\text{opt}}(\cdot)$ with NNs, the mineral grinding process starts working in a steady state at $r = r^* = [58; 0.2]$. After 60 min, the setpoints of the GPPS and CL are changed according to the demands of the mineral process plant to $r^* = [59.5; 0.3]$. The corresponding steady-state inputs are $w_s = [0.1204; 1.0987]$, which clearly violate the input constraints. Consequently, the setpoints $r^* = [59.5; 0.3]$ are infeasible. To further challenge the reference governor, at 220 min, the setpoints are changed again to $r^* = [56; 0.4]$, which are unreachable through manipulating the inputs of the simulation model. To simulate the limited acceleration of the actuators in practice, the absolute changes of w within unit sample time are bounded by 0.01 and 0.08, respectively.

Meanwhile, a comparative experiment without the reference governor is carried out in which the same optimal control policy is applied and steady-state inputs of the references are given. The simulation results are shown in Figs. 7–11.

Figs. 7 and 8 present the trajectories of the operational indices, and the feasible setpoints that are actually tracked. It is shown in Fig. 7 and 8 that when the setpoints are infeasible, the proposed method with the reference governor can guarantee that the GPPS can track its prescribed setpoint with priority over CL. The CL is compromised to track its feasible setpoint r_{g2} instead of its actual setpoint r_2^*.

When there is no reference governor applied to the controller, constraint satisfaction is no longer guaranteed, which can be clearly seen in Fig. 10. Since the setpoints $r^* = [56; 0.4]$ are unreachable, the simulation fails to proceed without a reference governor searching for feasible setpoints after 220 min, as shown in Figs. 7 and 8.

The input constraints are not violated during the whole time with the proposed method as can be seen in Figs. 9 and 10. The reference governor is activated when w approaches near the constraints at around 225 min. This is the reason why w_1 is very close to its lower bound but the constraints are not violated by the overshoot of w_1.

The trajectories of w_{opt} are shown in Fig. 11. w_{opt}, as a data-driven optimal regulator, can stabilize and regulate the controlled plant to its feasible setpoints. w_{opt} converge to 0 as the tracking errors decrease to 0.

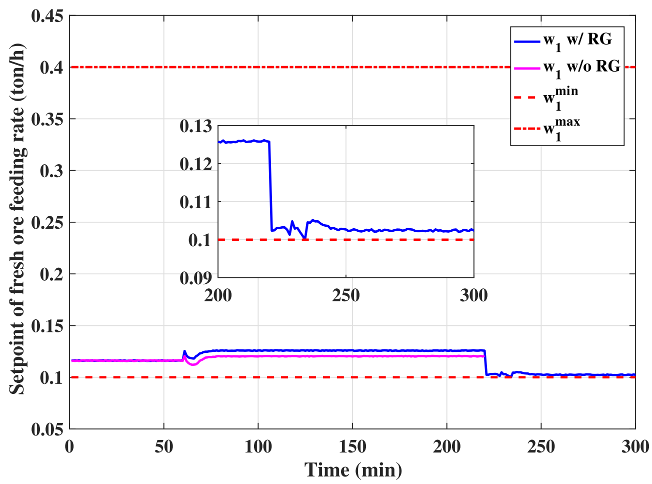


Fig. 9. Trajectories of setpoint of the FOFR and its constraints.

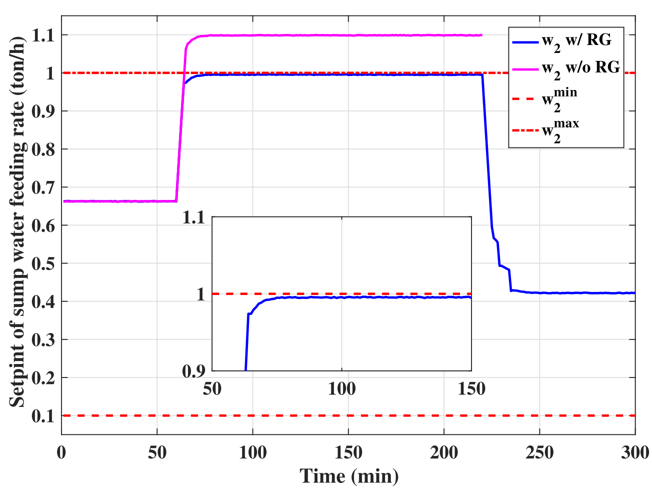


Fig. 10. Trajectories of setpoint of the SWFR and its constraints.

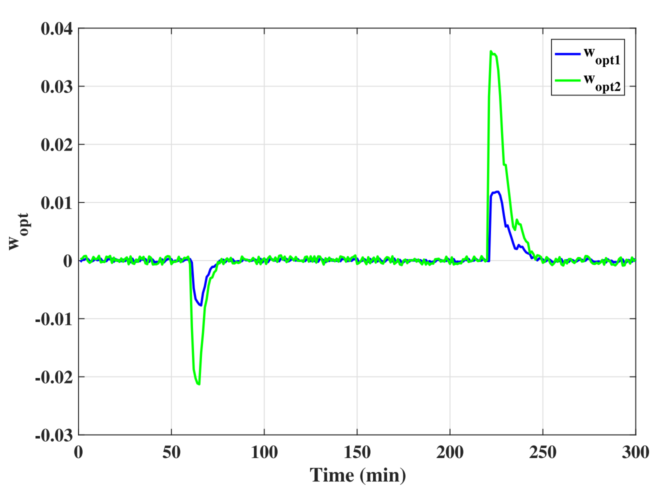


Fig. 11. Trajectories of $w_{\rm opt}$.

VI. Conclusion

In this paper, a data-driven input-constrained operational control method for mineral grinding processes is proposed. The reference governor is designed and used to deal with the infeasible setpoint problem and prevent the input constraints from being violated. A new policy iteration algorithm is proposed to fulfill the model-free design of the regulator with an ADP technique. The outputs can be regulated to the steady-state feasible setpoint without modeling the dynamic relationship between GPPS, CL, and setpoints of the FOFR and SWFR. Simulation results have shown the effectiveness of the proposed method.

Appendix

For a brief description of the simulation model used in this study, we list the main equations to represent the ball mill and hydrocyclone models in the Appendix. All the variables and subscripts of the model for ball mill and hydrocyclone are summarized in Table I.

A. Model for Ball Mill

The model for a ball mill is developed based on the principle of the mass balance with the assumption that materials in the mill are perfectly mixed. By the principle of mass balance, the concentration of the slurry out of the ball mill is

$$\dot{C}_{YB} = \frac{1}{V_B}\left(Q_{XB}C_{XB} - Q_{YB}C_{YB}\right). \tag{28}$$

The volume of the slurry in the ball mill is given by

$$\dot{V}_B = Q_{XB} - Q_{YB}. \tag{29}$$

The dynamics of the particle size distribution can be described as

$$\dot{M}_{YBi} = \frac{Q_{XB}C_{XB}}{V_B C_{YB}}\left(M_{XBi} - M_{YBi}\right) - S_i M_{YBi} + \sum_{j=1}^{i-1} b_{i,j} S_j M_{YBj} \tag{30}$$

where S_i is the selection function of ith grade particles, representing the percentage of ith grade particles that are crushed and broken into other grades, and $b_{i,j}$ is a breakage function that denotes the percentage of the transformation from jth grade to ith grade. Breakage function is defined using accumulative breakage function B

$$b_{i,j} = B_{i,j} - B_{i+1,j}. \tag{31}$$

Empirical models are used for identification of S_i and $B_{i,j}$. The empirical model for accumulative breakage function[30], [33] is given as

$$B_{i,j} = \alpha_1\left(\frac{\bar{d}_i}{\bar{d}_j}\right)^{\alpha_2} + (1-\alpha_1)\left(\frac{\bar{d}_i}{\bar{d}_j}\right)^{\alpha_3} \tag{32}$$

where $\bar{d}_i = \sqrt{d_i d_{i+1}}$ is geometric mean of d_i and d_{i+1}, and parameters α_{1-3} should be determined experimentally. The

TABLE I
NOMENCLATURE FOR MODELING BALL MILL AND HYDROCYCLONE

Variables	
b	Breakage function $(\%)$
B	Cumulative breakage function $(\%)$
C	Concentration of solid in the slurry $\left(t/m^3\right)$
d	Particle size (μm)
E	Actual classification efficiency $(\%)$
E^c	Corrected classification efficiency $(\%)$
M	Particle size distribution $(\%)$
N	Number of ore grades
Q	Flow rate of the slurry (t/h)
V	Volume $\left(m^3\right)$
S	Selection function $(\%)$
ρ	Density $\left(t/m^3\right)$
Subscripts	
B	Ball mill
C	Hydrocyclone
i	i-th grade
i,j	from j-th to i-th grade
O	Overflow
U	Underflow
X	Input stream
Y	Output stream

empirical model for the selection function [34], [35] is given as

$$\ln S_i = \gamma_1 \left(d_i\right)^{\gamma_2} P\left(d_i\right) = \gamma_1 \left(d_i\right)^{\gamma_2} \frac{1}{1+\left(d_i/\gamma_3\right)^{\gamma_4}} \tag{33}$$

where $P\left(d_i\right)$ is a correction factor for the abnormal breakage of large particles [35], and γ_{1-4} are parameters that need identification through experiments.

B. Model for Hydrocyclone

A dynamic model is unnecessary because the response of the hydrocyclone is virtually instantaneous[32]. From the mass balance principle and definition of classification efficiency, we have

$$Q_{\text{YCU}} C_{\text{YCU}} M_{\text{YCUi}} = Q_{XC} C_{XC} M_{\text{XCi}} E_i \tag{34}$$

$$Q_{\text{YCO}} C_{\text{YCO}} M_{\text{YCOi}} = Q_{XC} C_{XC} M_{\text{XCi}} \left(1 - E_i\right) \tag{35}$$

where E_i denotes the ratio of the ith grade ore classified into underflow.

The empirical model for the hydrocyclone [32] is given as

$$Q_{\text{YCO}} \left(1 - C_{\text{YCO}}\right) = \beta_1 Q_{XC} \left(1 - C_{XC}\right) + \beta_2 \tag{36}$$

$$E_i = E_i^c \left(1 - R_f\right) + R_f \tag{37}$$

$$E_i^c = 1 - \exp\left(-0.693 \left(\frac{\bar{d}_i}{d_{50}}\right)^{\beta_3}\right) \tag{38}$$

$$d_{50} = \exp\left(\beta_4 + \beta_5 \ln Q_{XC} + \beta_6 C_{XC}\right) \tag{39}$$

$$R_f = \beta_7 + \beta_8 \frac{Q_{\text{YCO}} \left(1 - C_{\text{YCO}}\right)}{Q_{XC} \left(1 - C_{XC}\right)} \tag{40}$$

where R_f is the ratio of the fine ore in the underflow, d_{50} is the particle size at which 50% of the particles are separated to the underflow, and β_{1-8} are model parameters that should be determined experimentally.

From (35) and (36), it follows that

$$Q_{\text{YCO}} C_{\text{YCO}} = Q_{XC} C_{XC} \left(1 - \sum_{i=1}^{N} M_{\text{XCi}} E_i\right). \tag{41}$$

Therefore, the outputs of the model can be represented as

$$r_1 = \sum_{i=n_{74}}^{N} M_{\text{YCOi}} \tag{42}$$

$$r_2 = Q_{XC} C_{XC} \left(1 - \sum_{i=1}^{N} M_{\text{XCi}} E_i\right) \tag{43}$$

where n_{74} is the grade corresponding to particle size of 74 μm. In this case, $N = 25$ and $n_{74} = 14$.

To sum up, there are $(N+2)$ differential equations (28)–(30) in the simulation model. Dynamics of the sump are omitted because of the perfect mixing assumption. Parameters of the model for the ball mill and hydrocyclone are chosen according to the simulation model verification and validation in [31], and the modeling work in [30] as follows:

$$\alpha_{1-3} = [0.1, 0.12, 1.1]$$

$$\gamma_{1-4} = [0.3994, 0.5, 10000, 2.513]$$

$$\beta_{1-8} = [1.363, -10.75, 1.6, 3.616, -0.15, 2.3, 0.818, -0.7932].$$

For simplicity of control evaluation, the parameters are kept constant. Uniformly distributed random noises in the interval $(-0.1, 0.1)$ and $(-0.002, 0.002)$ are added in the outputs r_1 and r_2, respectively.

REFERENCES

[1] D. Hodouin, S. L. Jämsä-Jounela, M. Carvalho, and L. Bergh, “State of the art and challenges in mineral processing control,” *Control Eng. Practice*, vol. 9, no. 9, pp. 995–1005, 2001.

[2] D. Wei and I. K. Craig, “Grinding mill circuitsła survey of control and economic concerns,” *Int. J. Mineral Process.*, vol. 90, no. 1, pp. 56–66, 2009.

[3] D. Hodouin, “Methods for automatic control, observation, and optimization in mineral processing plants,” *J. Process Control*, vol. 21, no. 2, pp. 211–225, 2011.

[4] P. Zhou, S. Lu, M. Yuan, and T. Chai, “Survey on higher-level advanced control for grinding circuits operation,” *Powder Technol.*, vol. 288, pp. 324–338, 2016.

[5] M. Ramasamy, S. Narayanan, and C. D. Rao, “Control of ball mill grinding circuit using model predictive control scheme,” *J. Process Control*, vol. 15, no. 3, pp. 273–283, 2005.

[6] T. Chai, S. J. Qin, and H. Wang, “Optimal operational control for complex industrial processes,” *Annu. Rev. Control*, vol. 38, no. 1, pp. 81–92, 2014.

[7] R. Lestage, A. Pomerleau, and D. Hodouin, “Constrained real-time optimization of a grinding circuit using steady-state linear programming supervisory control,” *Powder Technol.*, vol. 124, no. 3, pp. 254–263, 2002.

[8] L. C. Coetzee, I. K. Craig, and E. C. Kerrigan, “Robust nonlinear model predictive control of a run-of-mine ore milling circuit,” *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 1, pp. 222–229, Jan. 2010.

[9] J. D. le Roux, R. Padhi, and I. K. Craig, “Optimal control of grinding mill circuit using model predictive static programming: A new nonlinear MPC paradigm,” *J. Process Control*, vol. 24, no. 12, pp. 29–40, 2014.

[10] X. Chen, Q. Li, and S. Fei, “Supervisory expert control for ball mill grinding circuits,” *Expert Syst. Appl.*, vol. 34, no. 3, pp. 1877–1885, 2008.

[11] P. Zhou, T. Chai, and J. Sun, "Intelligence-based supervisory control for optimal operation of a DCS-controlled grinding system," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 1, pp. 162–175, Jan. 2013.

[12] S. Yin, H. Gao, and O. Kaynak, "Data-driven control and process monitoring for industrial applications—Part I," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6356–6359, Nov. 2014.

[13] S. Yin, X. Li, H. Gao, and O. Kaynak, "Data-based techniques focused on modern industry: An overview," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 657–667, Jan. 2015.

[14] W. Dai, T. Chai, and S. X. Yang, "Data-driven optimization control for safety operation of hematite grinding process," *IEEE Trans. Ind. Electron.*, vol. 62, no. 5, pp. 2930–2941, May 2015.

[15] X. Lu, B. Kiumarsi, T. Chai, and F. L. Lewis, "Data-driven optimal control of operational indices for a class of industrial processes," *IET Control Theory A*, vol. 10, no. 12, pp. 1348–1356, 2016.

[16] Y. Jiang, J. Fan, T. Chai, J. Li, and F. L. Lewis, "Data-driven flotation industrial process operational optimal control based on reinforcement learning," *IEEE Trans. Ind. Informat.*, vol. 14, no. 5, pp. 1974–1989, May 2018.

[17] H. Yang, Y. Jiang, and S. Yin, "Fault-tolerant control of time-delay Markov jump systems with Ito stochastic process and output disturbance based on sliding mode observer," *IEEE Trans. Ind. Informat.*, vol. 14, no. 12, pp. 5299–5307, Dec. 2018, doi: 10.1109/TII.2018.2812754.

[18] C. Mu, Z. Ni, C. Sun, and H. He, "Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 584–598, Mar. 2017.

[19] Q. Wei *et al.*, "Data-driven neuro-optimal temperature control of water-gas shift reaction using stable iterative adaptive dynamic programming," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6399–6408, Nov. 2014.

[20] E. G. Gilbert and K. T. Tan, "Linear systems with state and control constraints: The theory and application of maximal output admissible sets," *IEEE Trans. Autom. control*, vol. 36, no. 9, pp. 1008–1020, Sep. 1991.

[21] E. G. Gilbert, I. Kolmanovsky, and K. T. Tan, "Discrete-time reference governors and the nonlinear control of systems with state and control constraints," *Int. J. Robust Nonlin.*, vol. 5, no. 5, pp. 487–504, 1995.

[22] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, vol. 12, no. 5, pp. 457–465, 1976.

[23] H. Zhang, L. Cui, X. Zhang, and Y. Luo, "Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2226–2236, Dec. 2011.

[24] H. Modares, F. L. Lewis, and M. B. Naghibi-Sistani, "Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 10, pp. 1513–1525, Oct. 2013.

[25] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE Circuits Syst. Mag.*, vol. 9, no. 3, pp. 32–50, Third Quarter 2009.

[26] D. Liu and Q. Wei, "Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 3, pp. 621–634, Mar. 2014.

[27] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. Cambridge, MA, USA: MIT Press, 1998.

[28] B. Igelnik and Y.-H. Pao, "Stochastic choice of basis functions in adaptive function approximation and the functional-link net," *IEEE Trans. Neural Netw.*, vol. 6, no. 6, pp. 1320–1329, Nov. 1995.

[29] T. Dierks and S. Jagannathan, "Online optimal control of affine nonlinear discrete-time systems with unknown internal dynamics by using time-based policy update," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1118–1129, Jul. 2012.

[30] W. Dai, "Optimal operational control software system for hematite grinding process," Ph.D. dissertation, Northeastern Univ., Shenyang, China, 2015.

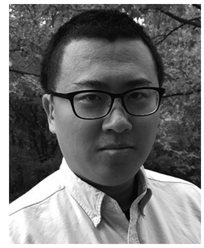
[31] S. Lu, P. Zhou, T. Chai, and W. Dai, "Modeling and simulation of whole ball mill grinding plant for integrated control," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 4, pp. 1004–1019, Oct. 2014.

[32] R. K. Rajamani and J. A. Herbst, "Optimal control of a ball mill grinding circuit—I. Grinding circuit modeling and dynamic simulation," *Chem. Eng. Sci.*, vol. 46, no. 3, pp. 861–870, 1991.

[33] J. Herbst, K. Rajamani, and D. Kinneberg, *Estimill—A Program for Grinding Simulation and Parameter Estimation With Linear Models*, University of Utah, Salt Lake City, UT, USA, 1977.

[34] L. G. Austin, R. R. Klimpel, and P. T. Luckie, *Process Engineering of Size Reduction: Ball Milling*, American Institute of Mining, Metallurgical, and Petroleum Engineers, Englewood, CO, USA, 1984.

[35] M. Boulvin, A. V. Wouwer, R. Lepore, C. Renotte, and M. Remy, "Modeling and control of cement grinding processes," *IEEE Trans. Control Syst. Technol.*, vol. 11, no. 5, pp. 715–725, Sep. 2003.



Xinglong Lu (S'18) received the bachelor's degree in automation and the master's degree in control engineering from Northeastern University, Shenyang, China, 2011 and 2013, respectively, where he is currently working toward the Ph.D. degree in control theory and control engineering with the State Key Laboratory of Synthetical Automation for Process Industries.

From 2015 to 2016, he was a Visiting Scholar with the Naval Architecture and Marine Engineering Department, University of Michigan, Ann Arbor, MI, USA. His current research interests include nonlinear dynamics and control, reinforcement learning, extremum seeking, statistical process monitoring, and industrial process control.

Bahare Kiumarsi (M'17) received the B.S. degree from the Shahrood University of Technology, Shahrud, Iran, in 2009, the M.S. degree from the Ferdowsi University of Mashhad, Mashhad, Iran, in 2013, and the Ph.D. degree from the University of Texas at Arlington, Arlington, TX, USA, in August 2017, all in electrical engineering.

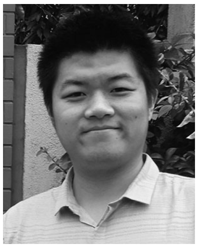
She is a Postdoctoral Research Associate with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Champaign, IL, USA. Her current research interests include machine learning in control, security of cyber-physical systems, game theory, and distributed control.

Dr. Kiumarsi was the recipient of the UT-Arlington N.M. Stelmakh Outstanding Student Research Award and the UT Arlington Graduate Dissertation Fellowship for summer 2017.

Tianyou Chai (M'90–SM'97–F'08) received the Ph.D. degree in control theory and engineering from Northeastern University, Shenyang, China, in 1985.

He became a Professor with Northeastern University in 1988. He is the Founder and the Director of the Center of Automation, which became a National Engineering and Technology Research Center and a State Key Laboratory. He is the Director with the Department of Information Science, National Natural Science Foundation of China, Beijing, China. He has authored and co-authored more than 200 peer reviewed international journal papers. He has developed control technologies with applications to various industrial processes. His current research interests include modeling, control, optimization, and integrated automation of complex industrial processes.

Dr. Chai is a Member of the Chinese Academy of Engineering and a Fellow of the International Federation of Automatic Control. His paper titled "Hybrid Intelligent Control For Optimal Operation of Shaft Furnace Roasting Process" was selected as one of three best papers for the Control Engineering Practice Paper Prize for 2011–2013. For his contributions, he received four prestigious awards of the National Science and Technology Progress and National Technological Innovation, the 2007 Industry Award for Excellence in Transitional Control Research from the IEEE Multiple-conference on Systems and Control, and the 2017 Wook Hyun Kwon Education Award from Asian Control Association.



Yi Jiang (S'14) was born in Hubei Province, China. He received the B.Eng. degree in automation and the M.S. degree in control theory and control engineering from the Information Science and Engineering College and the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China, in 2014 and 2016, respectively, where he is currently working toward the Ph.D. degree in control theory and control engineering.

From January to July, 2017, he was a Visiting Scholar with the UTA Research Institute, University of Texas at Arlington, Arlington, TX, USA, and since March 2018, he has been a Research Assistant with the University of Alberta, Edmonton, Canada. His research interests include networked control systems, industrial process operational control, reinforcement learning, and event-triggered estimation and control.

Frank L. Lewis (S'70–M'81–SM'86–F'94) received the bachelor's degree in physics/electrical engineering and the M.S.E.E. degree from Rice University, Houston, TX, USA, in 1971, the M.S. degree in aeronautical engineering from the University of West Florida, Pensacola, FL, USA, in 1977, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, GA, USA, in 1981.

He is a Distinguished Scholar Professor, a Distinguished Teaching Professor, and the Moncrief-O'Donnell Chair with the UTA Research Institute, University of Texas at Arlington, Arlington, TX. He is also a Qian Ren Thousand Talents Consulting Professor and a Project 111 Professor with Northeastern University, Shenyang, China. He is Distinguished Visiting Professor with the Nanjing University of Science & Technology, Nanjing, China. His research interests include feedback control, intelligent systems, cooperative control systems, and nonlinear systems. He is an author of seven U.S. patents, numerous journal special issues, 380 journal papers, and 20 books, including "*Optimal Control, Aircraft Control, Optimal Estimation*" and "*Robot Manipulator Control*," which are used as university textbooks worldwide.

Dr. Lewis is a Member of the National Academy of Inventors, a Fellow of the International Federation of Automatic Control, a Fellow of the American Association for the Advancement of Science, a Fellow of the U.K. Institute of Measurement & Control, a Fellow of the Texas Board of Professional Engineers (PE Texas), and a U.K. Chartered Engineer. He ranked at position 85 worldwide, 63 in the USA, and 3 in Texas of all scientists in Computer Science and Electronics, by Guide2Research (May 2018). He was the recipient of the China Liaoning Friendship Award, the Fulbright Research Award, the National Science Foundation (NSF) Research Initiation Grant, the American Society for Engineering Education (ASEE) Terman Award, the International Neural Network Society Gabor Award, the U.K. Institute Measurement and Control Honeywell Field Engineering Medal, the IEEE Computational Intelligence Society Neural Networks Pioneer Award, the American Institute of Aeronautics and Astronautics Intelligent Systems Award, the American Automatic Control Council (AACC) Ragazzini Award, the Outstanding Service Award from Dallas IEEE Section, and the Texas Regents Outstanding Teaching Award 2013. He was selected as the Engineer of the year by Ft. Worth IEEE Section, and was listed in Ft. Worth Business Press Top 200 Leaders in Manufacturing. He is a Founding Member of the Board of Governors of the Mediterranean Control Association.