Research Article

Optimal Control for a Linear System Subject to a General ARIMA Disturbance

Hongyan Xie¹ and Fangyi He²

¹School of Economics, Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, China

²School of Finance, Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, China

Correspondence should be addressed to Fangyi He; fangyi_he@swufe.edu.cn

Received 28 November 2014; Revised 24 January 2015; Accepted 16 February 2015

Academic Editor: Alain Vande Wouwer

Copyright © 2015 H. Xie and F. 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.

A novel run-to-run (R2R) control algorithm based on Kalman filter approach is proposed to deal with a linear system with a general ARIMA(p, d, q) process, in the presence of measurement error and adjustment error together with a random initial bias. We mathematically prove its optimality. The performance of the newly proposed controller and the proportional-integral (PI) controller is evaluated and compared under multiple scenarios through Monte Carlo simulations. Almost all the results reflect the new controller's superiority.

1. Introduction and Motivation

Optimizing a sequence of actions to obtain some future goal is the general topic of control theory (see [1]). Many optimal control methods have been widely used in actuarial science (e.g., [2, 3]), production management (e.g., [4–6]), and quality management (e.g., [7–9]). Specifically, disturbance rejection is one of the major concerns in quality control problems, such as machine setup adjustment problems and R2R process control problems in semiconductor manufacturing (see [10, 11]). Grubbs [12] originally studies the machine setup adjustment problem and develops an optimal adjustment rule for a linear system with normal disturbances, which is referred to by Trietsch [13] as the “harmonic rule.” Vander Wiel et al. [14] give an optimal control algorithm for a linear system with ARMA(1, 1) disturbances. For a linear system with IMA(1, 1) disturbances, Ingolfsson and Sachs [15] and Box et al. [16] introduce the exponential weighted moving average (EWMA) control algorithm and prove its optimality; He et al. [17] recently develop an optimal control algorithm, the ARMA controller, for a linear system with a general ARMA(p, q) disturbance.

In this paper, we will extend the results in [17] and develop an optimal control algorithm for a linear system with a general ARIMA(p, d, q) disturbance, where $d = 1, 2, \dots$. The ARIMA disturbance has been widely used to describe process

dynamics (see [18–21]). Since ARIMA(p, d, q) processes can be used to model a large class of nonstationary disturbances, many machine setup problems and R2R control problems can be solved under this new framework. Similar to [17], the newly proposed controller will consider a more realistic case such that both measurement error and adjustment error exist, and the initial bias of the process is a random variable. We present the problem as follows.

Supposing a process to be controlled y_t can be expressed as

$$y_t = g_t + n_t + e_t, \quad (1)$$

where g_t is the state of y_t at time t ; n_t is an ARIMA(p, d, q) process satisfying

$$\phi(\mathcal{B})(1 - \mathcal{B})^d n_t = \theta(\mathcal{B}) z_t, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2)$$

where $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$, $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$, and z_t is a white noise process with mean 0 and variance σ_z^2 , that is, $z_t \sim \text{WN}(0, \sigma_z^2)$; e_t is the measurement error for y_t and $e_t \sim \text{WN}(0, \sigma_e^2)$. $\mathcal{B}$ is the backward shift operator defined by

$$\mathcal{B}^j n_t = n_{t-j}, \quad j = 0, 1, 2, \dots \quad (3)$$

Neither g_t nor n_t can be measured or observed directly. At time $t \geq 1$, suppose we need to make an adjustment of magnitude Δg_t to g_t to bring the process output to target τ in the next run. That is,

$$g_{t+1} = g_t + \Delta g_t + w_t, \quad (4)$$

where w_t is the adjustment error and $w_t \sim \text{WN}(0, \sigma_w^2)$. In practice, the process adjustment Δg_t is assumed to be done by adjusting one controllable factor, u_t , via the following model:

$$\Delta g_t = \beta \Delta u_t, \quad \beta \neq 0, \quad (5)$$

where β is called the process gain and $\Delta u_t = u_{t+1} - u_t$. Assume $E(e_t w_t) = E(w_t z_t) = E(z_t e_t) = 0$, $t = 0, \pm 1, \pm 2, \dots$. The initial value g_0 is assumed to be a random variable with $E(g_0 z_t) = E(g_0 e_t) = E(g_0 w_t) = 0$, $t = 0, \pm 1, \pm 2, \dots$, and $E(g_0) = \hat{g}_0$, $\text{Var}(g_0) = \sigma_g^2$. Without loss of generality, the target τ is assumed to be 0 in the rest of this paper. For a positive time index T , we hope to determine the optimal Δu_t , $t = 0, 1, \dots, T-1$ that satisfy

$$\min_{\Delta u_t, t=0,1,\dots,T-1} E_0 \left(\sum_{t=0}^{T-1} y_{t+1}^2 \right), \quad (6)$$

where $E_t(\cdot)$ is the expectation operator conditional on all the information until time t . Problem (6) is a finite-horizon problem. When $T \rightarrow \infty$, we get the infinite-horizon

problem that searches the optimal Δu_t , $t = 0, 1, \dots$ satisfying

$$\min_{\Delta u_t, t=0,1,\dots} E_0 \left(\sum_{t=0}^{\infty} y_{t+1}^2 \right). \quad (7)$$

The purpose of this paper is to find the solutions to both problems (6) and (7).

The rest of this paper is organized as follows. In Section 2 we derive the state-space representation of the system and get the recursive estimation formulae of the system states using Kalman filter. In Section 3, we develop a control algorithm for the system and prove its optimality without normal distribution assumptions. We further give the implementation steps of the controller in practice. Simulation studies are done in Section 4 under multiple scenarios. Section 5 gives an illustrative example for the application of the control algorithm. Concluding remarks are included in Section 6.

2. The State-Space Representation

Let $m = \max(p, q)$ and $\psi_1, \dots, \psi_m$ be the coefficients of $z, z^2, \dots, z^m$ in the power series expansion of $\theta(z)/\phi(z)$, $|z| \leq 1$. Brockwell and Davis [22] gave a state-space representation for a general ARIMA(p, d, q) process. By extending their results we can get the following Theorem 1 easily.

Theorem 1. *The linear system (1)–(5) has a state-space representation as follows:*

$$y_t = \left[1 \quad G' \quad (-1)^{d+1} \binom{d}{d} \quad (-1)^d \binom{d}{d-1} \quad (-1)^{d-1} \binom{d}{d-2} \quad \cdots \quad d \right] \mathbf{R}_t + z_t + e_t, \quad t = 1, 2, \dots \quad (8)$$

$$\mathbf{R}_{t+1} = \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & F & \mathbf{0} \\ \mathbf{0} & AG' & B \end{bmatrix} \mathbf{R}_t + \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \beta \Delta u_t + \begin{bmatrix} w_t \\ Hz_t \\ Az_t \end{bmatrix}, \quad t = 1, 2, \dots, \quad (9)$$

where A and B are the $d \times 1$ and $d \times d$ matrices defined by $A = B = 1$ if $d = 1$ and

$$A = [0 \quad 0 \quad \cdots \quad 0 \quad 1]',$$

$$B = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ (-1)^{d+1} \binom{d}{d} & (-1)^d \binom{d}{d-1} & (-1)^{d-1} \binom{d}{d-2} & \cdots & d \end{bmatrix} \quad (10)$$

if $d > 1$; G and F are the $m \times 1$ and $m \times m$ matrices defined by

$G = 1$, $F = \phi_1$ if $m = 1$ and

$$G = [1 \quad 0 \quad 0 \quad \cdots \quad 0]',$$

$$F = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \phi_m & \phi_{m-1} & \phi_{m-2} & \cdots & \phi_1 \end{bmatrix} \quad (11)$$

if $m > 1$; $H = [\psi_1 \quad \psi_2 \quad \cdots \quad \psi_{m-1} \quad \psi_m]'$, $\mathbf{N}_{t-1} = [n_{t-d}, \dots, n_{t-1}]'$, $\mathbf{X}_t = [X_{t-m+1}, \dots, X_t]'$, $\mathbf{R}_t = [g_t \quad \mathbf{X}_t' \quad \mathbf{N}_{t-1}']'$.

A proof of Theorem 1 is presented in Appendix A. For simplicity, in the rest of this paper we denote

$$\mathbf{h} \equiv \left[1 \quad G' \quad (-1)^{d+1} \binom{d}{d} \quad (-1)^d \binom{d}{d-1} \quad (-1)^{d-1} \binom{d}{d-2} \quad \dots \quad d \right]',$$

$$L \equiv \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & F & \mathbf{0} \\ \mathbf{0} & AG' & B \end{bmatrix},$$

$$\boldsymbol{\ell} \equiv \begin{bmatrix} 1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{m} = \begin{bmatrix} 0 \\ H\sigma_z^2 \\ A\sigma_z^2 \end{bmatrix},$$

$$M = \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \text{diag}(H) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \text{diag}(A) \end{bmatrix},$$

$$\Sigma \equiv \begin{bmatrix} \sigma_w^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_m \mathbf{1}_m' \sigma_z^2 & \mathbf{1}_m \mathbf{1}_d' \sigma_z^2 \\ \mathbf{0} & \mathbf{1}_d \mathbf{1}_m' \sigma_z^2 & \mathbf{1}_d \mathbf{1}_d' \sigma_z^2 \end{bmatrix},$$
(12)

where $\mathbf{1}_k$ is defined to be a $k \times 1$ vector of ones.

That the problems of optimal control and state estimation can be decoupled in certain cases is one of the most fundamental principles in feedback control theory. This is known as the *separation principle* (see [23]). The Kalman filter produces the statistically optimal estimate of the state of system (8)-(9). Note the fact that the disturbances of (8) and (9) are correlated to each other, so we should use the Kalman filter formulae for correlated measurement and process noise. At any time t , let us define $\widehat{\mathbf{R}}_{t+1}^- = E_t(\mathbf{R}_{t+1})$ and $V_{t+1}^- = \text{Var}_t(\mathbf{R}_{t+1})$ and assume $\widehat{\mathbf{R}}_1^- = [\hat{g}_0 \quad \mathbf{0} \quad \mathbf{0}]'$ and $V_1^- = \text{diag}(\sigma_g^2, \mathbf{0}, \mathbf{0})$. Using the results on page 123 in Lewis' book [24], we can directly get the following Lemma 2.

Lemma 2. For the systems (8)-(9), we have the following recursive formulae:

$$\widehat{\mathbf{R}}_{t+1}^- = L\widehat{\mathbf{R}}_t^- + \boldsymbol{\ell}\beta\Delta u_t + K_t(y_t - \mathbf{h}'\widehat{\mathbf{R}}_t^-), \quad (13)$$

$$V_{t+1}^- = LV_t^-L' - K_t(\mathbf{h}'V_t^-\mathbf{h} + \sigma_z^2 + \sigma_e^2)K_t' + M\Sigma M', \quad (14)$$

where K_t is called the modified Kalman Gain that

$$K_t = (LV_t^-\mathbf{h} + \mathbf{m})(\mathbf{h}'V_t^-\mathbf{h} + \sigma_z^2 + \sigma_e^2)^{-1}. \quad (15)$$

3. The Optimal Control Algorithm

In this section, we will derive an optimal control algorithm without the normal distribution assumptions on $\mathbf{R}_1$, e_t , w_t , and z_t , $t = 0, \pm 1, \dots$. This optimal control algorithm can be applied to both the finite-horizon and infinite-horizon problems.

Theorem 3. The optimal control algorithm that solves both

$$\min_{\Delta u_t, t=0,1,\dots,T-1} E_0 \left(\sum_{t=0}^{T-1} y_{t+1}^2 \right), \quad \min_{\Delta u_t, t=0,1,\dots} E_0 \left(\sum_{t=0}^{\infty} y_{t+1}^2 \right) \quad (16)$$

for the system (1)-(5) is

$$\Delta u_t = -\frac{1}{\beta} \mathbf{h}' \left[L\widehat{\mathbf{R}}_t^- + K_t(y_t - \mathbf{h}'\widehat{\mathbf{R}}_t^-) \right], \quad t = 1, 2, \dots, \quad (17)$$

where $\widehat{\mathbf{R}}_t^-$ is updated according to

$$\widehat{\mathbf{R}}_{t+1}^- = (I - \boldsymbol{\ell}\mathbf{h}') \left[L\widehat{\mathbf{R}}_t^- + K_t(y_t - \mathbf{h}'\widehat{\mathbf{R}}_t^-) \right] \quad (18)$$

and (14), (15) are with initial values $\widehat{\mathbf{R}}_1^-$ and V_1^- .

Proof. For a given positive time index T , let $J_t = \min_{\Delta u_s, s=t, t+1, \dots, T-1} E_t(\sum_{s=t}^{T-1} y_{s+1}^2)$ and $J_T \equiv 0$. We can get the Bellman equation as

$$J_t = \min_{\Delta u_t} \left[E_t(y_{t+1}^2) + E_t(J_{t+1}) \right]. \quad (19)$$

From (8), it is evident that $E_t(y_{t+1}^2) = E_t(\mathbf{R}_{t+1}' \mathbf{h} \mathbf{h}' \mathbf{R}_{t+1}) + \sigma_z^2 + \sigma_e^2$. Using the fact that

$$\begin{aligned} E_t(\mathbf{R}_{t+1}' \mathbf{h} \mathbf{h}' \mathbf{R}_{t+1}) &= (\widehat{\mathbf{R}}_{t+1}^-)' \mathbf{h} \mathbf{h}' (\widehat{\mathbf{R}}_{t+1}^-) \\ &\quad + E_t \left[(\mathbf{R}_{t+1} - \widehat{\mathbf{R}}_{t+1}^-)' \mathbf{h} \mathbf{h}' (\mathbf{R}_{t+1} - \widehat{\mathbf{R}}_{t+1}^-) \right] \end{aligned} \quad (20)$$

and the property of trace operator, we obtain that

$$\begin{aligned} E_t(y_{t+1}^2) &= (\widehat{\mathbf{R}}_{t+1}^-)' \mathbf{h} \mathbf{h}' (\widehat{\mathbf{R}}_{t+1}^-) \\ &\quad + \text{trace} \left\{ \mathbf{h} \mathbf{h}' E_t \left[(\mathbf{R}_{t+1} - \widehat{\mathbf{R}}_{t+1}^-) (\mathbf{R}_{t+1} - \widehat{\mathbf{R}}_{t+1}^-)' \right] \right\} \\ &\quad + \sigma_z^2 + \sigma_e^2. \end{aligned} \quad (21)$$

From (13) and the fact that $\mathbf{h}'\boldsymbol{\ell} \equiv 1$, we have

$$\begin{aligned} E_t(y_{t+1}^2) &= \left[L\widehat{\mathbf{R}}_t^- + \boldsymbol{\ell}\beta\Delta u_t + K_t(y_t - \mathbf{h}'\widehat{\mathbf{R}}_t^-) \right]' \mathbf{h} \mathbf{h}' \\ &\quad \cdot \left[L\widehat{\mathbf{R}}_t^- + \boldsymbol{\ell}\beta\Delta u_t + K_t(y_t - \mathbf{h}'\widehat{\mathbf{R}}_t^-) \right] \\ &\quad + \text{trace}(\mathbf{h} \mathbf{h}' V_{t+1}^-) + \sigma_z^2 + \sigma_e^2 \\ &= [\beta\Delta u_t + \mathbf{h}'L\widehat{\mathbf{R}}_t^- + \mathbf{h}'K_t(y_t - \mathbf{h}'\widehat{\mathbf{R}}_t^-)]^2 \\ &\quad + \mathbf{h}'V_{t+1}^-\mathbf{h} + \sigma_z^2 + \sigma_e^2. \end{aligned} \quad (22)$$

Let $S_t \equiv \mathbf{h}'V_{t+1}^-\mathbf{h} + \sigma_z^2 + \sigma_e^2$. It is evident that S_t is unrelated to Δu_t since V_{t+1}^- is updated based on (14) and (15). Now (19) changes to

$$\begin{aligned} J_t &= \min_{\Delta u_t} \left\{ [\beta\Delta u_t + \mathbf{h}'L\widehat{\mathbf{R}}_t^- + \mathbf{h}'K_t(y_t - \mathbf{h}'\widehat{\mathbf{R}}_t^-)]^2 + E_t(J_{t+1}) \right\} \\ &\quad + S_t. \end{aligned} \quad (23)$$

When $t = T - 1$, we get

$$J_{T-1} = \min_{\Delta u_{T-1}} \left[\beta \Delta u_{T-1} + \mathbf{h}' L \widehat{\mathbf{R}}_{T-1}^- + \mathbf{h}' K_{T-1} (y_{T-1} - \mathbf{h}' \widehat{\mathbf{R}}_{T-1}^-) \right]^2 + S_{T-1}. \quad (24)$$

Solving the first order condition, we get that

$$\Delta u_{T-1}^* = -\frac{1}{\beta} \mathbf{h}' \left[L \widehat{\mathbf{R}}_{T-1}^- + K_{T-1} (y_{T-1} - \mathbf{h}' \widehat{\mathbf{R}}_{T-1}^-) \right]. \quad (25)$$

Putting (25) into (24), we get $J_{T-1} = S_{T-1}$. When $t = T - 2$, we have

$$J_{T-2} = \min_{\Delta u_{T-2}} \left[\beta \Delta u_{T-2} + \mathbf{h}' L \widehat{\mathbf{R}}_{T-2}^- + \mathbf{h}' K_{T-2} (y_{T-2} - \mathbf{h}' \widehat{\mathbf{R}}_{T-2}^-) \right]^2 + S_{T-2} + S_{T-1}. \quad (26)$$

Solving the first order condition, we get

$$\Delta u_{T-2}^* = -\frac{1}{\beta} \mathbf{h}' \left[L \widehat{\mathbf{R}}_{T-2}^- + K_{T-2} (y_{T-2} - \mathbf{h}' \widehat{\mathbf{R}}_{T-2}^-) \right] \quad (27)$$

and $J_{T-2} = S_{T-2} + S_{T-1}$. Repeating the procedures above, we can prove that (17) is the optimal adjustment strategy at time t and get

$$\begin{aligned} J_0 &= \min_{\Delta u_t, t=0,1,\dots,T-1} E_0 \left(\sum_{t=0}^{T-1} y_{t+1}^2 \right) \\ &= \sum_{t=0}^{T-1} S_t = \mathbf{h}' \left(\sum_{t=0}^{T-1} V_{t+1}^- \right) \mathbf{h} + T (\sigma_z^2 + \sigma_e^2). \end{aligned} \quad (28)$$

At last, we can derive (18) by putting (17) into (13). Since (17) is unrelated to the given time index T , the control strategy also solves the infinite-horizon problem. Then we finish the proof. $\square$

In practice, the steps to implement the proposed control algorithm are as follows.

Step 0. Setup the algorithm's parameters and initial values, such as $\widehat{\beta}$, $\widehat{g}_0$, $\widehat{\phi}_i$ ($i = 1, 2, \dots, p$), $\widehat{\theta}_j$ ($j = 1, 2, \dots, q$), $\widehat{\mathbf{R}}_1^-$, V_1^- , based on experience or history data.

Step 1. Compute K_t based on (15).

Step 2. Collect the new observation y_t and compute the adjustment Δu_t based on (17).

Step 3. Update $\widehat{\mathbf{R}}_{t+1}^-$ based on (18).

Step 4. Update V_{t+1}^- based on (14); let $t = t + 1$ and go back to Step 1.

Note that Step 0 is an off-line procedure and Step 1 to Step 4 are on-line procedures. As the newly proposed control algorithm is specially designed for adjusting any ARIMA disturbances, we call it ARIMA controller in the rest of this paper.

4. Simulation Study

In this section, we will study the performance of the ARIMA controller under multiple scenarios through Monte Carlo simulations. For simplicity, we only focus on the system (1)–(5) with n_t being an ARIMA(1, 1, 1) disturbance, although the ARIMA controller can be applied to any general ARIMA(p, d, q) disturbance. Without loss of generality, we set the parameters in the ARIMA(1, 1, 1) disturbance as $\phi_1 = 0.7$, $\theta_1 = -0.3$, and $\sigma_z = 1$. For comparison, a *PI* controller's performance is also evaluated. The *PI* controller is widely used in feedback control, and it involves two separate constant parameters: the proportional and the integral values, denoted by k_p and k_I . The steps to implement a *PI* controller can be found in [25].

Tuning a *PI* controller could be a challenging task. There are many ways of tuning, such as Ziegler-Nichols tuning, lambda tuning, robust loop shaping, optimization methods, and others (see [26]). Among them, optimization methods are powerful and direct ways. In order to do fair comparisons with the ARIMA controller, we need to choose k_p and k_I to make the *PI* controller have the best performance for controlling the specified ARIMA(1, 1, 1) disturbance. To the best of our knowledge, there are no closed-form expressions for k_p and k_I to optimally control a ARIMA(1, 1, 1) process; then experimental tuning of the *PI* parameters has to be performed. The procedure of optimally choosing k_p and k_I in this paper is described in Appendix B. For the ARIMA(1, 1, 1) disturbance with $\phi_1 = 0.7$ and $\theta_1 = -0.3$, we get $k_p = 1.223$ and $k_I = 0.190$ according to the optimization procedure.

Four scenarios are examined in the following simulations. In Scenario 1, the effect of the measurement error and the adjustment error on the ARIMA controller and the *PI* controller is investigated; in Scenario 2, the effect of the process initial bias g_0 on the two controllers is explored; in Scenario 3, we study how the estimate of the process gain $\widehat{\beta}$ affects the two controllers; in Scenario 4, we investigate both controllers' performance when the disturbance parameters are not estimated accurately.

We do 1000 replications for each case and run 100 steps for y_t in each replication. The first 100-run mean square error (MSE) of y_t is computed. The Average MSE (AMSE) of the 1000 replications is reported in Tables 1–4. Also the standard error in the AMSE (SEAMSE) is computed, which is

$$\text{SEAMSE} = \frac{\text{SDMSE}}{\sqrt{\text{Number of replicates}}}, \quad (29)$$

where SDMSE is the standard deviation of mean square errors, and the number of replicates here is 1000. AMSE measures the performance of the controllers, and SEAMSE reflects the variability of the AMSE. The smaller the AMSE is, the better performance the controller has.

4.1. Effects of Measurement Error and Adjustment Error. In order to focus on the effects of the measurement error e_t and the adjustment error w_t on the controllers, we set the initial process bias $g_0 = 0$, that is, $\widehat{g}_0 = \sigma_g = 0$, and set the estimate of the process gain $\widehat{\beta}$ to the true value β . We also assume all

TABLE 1: AMSE when both types of error exist. n_t is the ARIMA(1, 1, 1) disturbance with $\phi_1 = 0.7$ and $\theta_1 = -0.3$.

ARIMA controller $\hat{\phi}_1 = 0.7 \quad \hat{\theta}_1 = -0.3$					
σ_e	$\sigma_w = 0$	0.5	1.0	1.5	2.0
0	1.014 (.005)	1.298 (.006)	2.104 (.009)	3.392 (.015)	5.155 (.023)
0.5	1.657 (.008)	1.922 (.009)	2.695 (.012)	3.959 (.018)	5.709 (.026)
1.0	3.165 (.014)	3.431 (.016)	4.205 (.019)	5.471 (.025)	7.225 (.032)
1.5	5.286 (.024)	5.569 (.025)	6.381 (.029)	7.698 (.035)	9.503 (.043)
2.0	7.966 (.037)	8.269 (.038)	9.136 (.041)	10.530 (.047)	12.420 (.056)
PI controller $k_p = 1.223 \quad k_I = 0.190$					
σ_e	$\sigma_w = 0$	0.5	1.0	1.5	2.0
0	1.084 (.005)	1.379 (.006)	2.265 (.010)	3.742 (.017)	5.809 (.027)
0.5	1.872 (.009)	2.172 (.010)	3.063 (.015)	4.544 (.022)	6.616 (.032)
1.0	4.247 (.023)	4.552 (.025)	5.448 (.029)	6.934 (.036)	9.010 (.047)
1.5	8.211 (.048)	8.520 (.049)	9.420 (.053)	10.911 (.060)	12.992 (.071)
2.0	13.761 (.082)	14.075 (.083)	14.980 (.088)	16.475 (.095)	18.561 (.105)

The corresponding SEAMSE's are enclosed in the parentheses.

TABLE 2: AMSE when the initial bias' mean and standard deviation vary. n_t is the ARIMA(1, 1, 1) disturbance with $\phi_1 = 0.7$ and $\theta_1 = -0.3$.

ARIMA controller $\hat{\phi}_1 = 0.7 \quad \hat{\theta}_1 = -0.3$					
$\hat{g}_0$	$\sigma_g = 0$	1	2	3	4
-4	1.176 (.006)	1.180 (.007)	1.203 (.008)	1.244 (.010)	1.304 (.013)
-3	1.106 (.006)	1.110 (.006)	1.134 (.007)	1.176 (.009)	1.237 (.011)
-2	1.055 (.005)	1.061 (.005)	1.085 (.006)	1.128 (.008)	1.190 (.010)
-1	1.025 (.005)	1.031 (.005)	1.057 (.005)	1.101 (.007)	1.163 (.009)
0	1.014 (.005)	1.021 (.005)	1.048 (.005)	1.093 (.006)	1.156 (.008)
1	1.024 (.005)	1.032 (.005)	1.059 (.005)	1.105 (.007)	1.169 (.009)
2	1.053 (.005)	1.062 (.005)	1.091 (.006)	1.137 (.008)	1.202 (.010)
3	1.103 (.006)	1.113 (.006)	1.142 (.007)	1.189 (.009)	1.255 (.011)
4	1.172 (.007)	1.183 (.007)	1.213 (.008)	1.261 (.011)	1.328 (.013)
PI controller $k_p = 1.223 \quad k_I = 0.190$					
$\hat{g}_0$	$\sigma_g = 0$	1	2	3	4
-4	1.279 (.007)	1.284 (.008)	1.311 (.010)	1.360 (.012)	1.432 (.016)
-3	1.194 (.006)	1.200 (.007)	1.228 (.008)	1.279 (.010)	1.351 (.013)
-2	1.133 (.006)	1.140 (.006)	1.170 (.007)	1.221 (.009)	1.294 (.011)
-1	1.096 (.005)	1.105 (.005)	1.135 (.006)	1.187 (.008)	1.261 (.010)
0	1.084 (.005)	1.093 (.005)	1.124 (.006)	1.177 (.007)	1.253 (.010)
1	1.095 (.005)	1.105 (.005)	1.137 (.006)	1.192 (.008)	1.268 (.010)
2	1.130 (.006)	1.141 (.006)	1.174 (.007)	1.230 (.009)	1.307 (.011)
3	1.189 (.007)	1.201 (.007)	1.235 (.008)	1.292 (.010)	1.370 (.013)
4	1.272 (.007)	1.285 (.008)	1.320 (.009)	1.378 (.012)	1.458 (.016)

The corresponding SEAMSE's are enclosed in the parentheses. Neither measurement errors nor adjustment errors are considered in the simulations.

the ARIMA(1, 1, 1) parameters can be accurately estimated, that is, $\hat{\phi}_1 = \phi_1$ and $\hat{\theta}_1 = \theta_1$.

Table 1 and Figure 1 show the performance of the two controllers affected by the measurement errors and adjustment errors. We set the values of σ_e and σ_w to be 0, 0.5, 1.0, 1.5 and 2.0, respectively, so there are 25 pairs of (σ_e, σ_w) . For each pair of (σ_e, σ_w) , we repeat the simulations. We can observe that for both controllers the AMSE of y_t increases when σ_e or σ_w increases. However, the AMSE of y_t using the ARIMA controller is always smaller than that using the PI controller at a significant level for all the pairs of (σ_e, σ_w) .

4.2. Effects of Initial Bias. For varying sizes of $\hat{g}_0$ and σ_g in the prior distribution of the process initial bias, Table 2 and Figure 2 show the results of the AMSE of y_t using the ARIMA controller or the PI controller for the ARIMA(1, 1, 1) disturbance. In order to focus on the effects of the initial bias, we set all the e_t and w_t to 0, $\hat{\beta} = \beta$, $\hat{\phi}_1 = \phi_1$, and $\hat{\theta}_1 = \theta_1$.

It can be seen that for both controllers, the AMSE increases with σ_g when $\hat{g}_0$ is fixed; and the AMSE increases with $|\hat{g}_0|$ when σ_g is fixed. For the same pair of $(\hat{g}_0, \sigma_g)$, the AMSE of y_t using the ARIMA controller is always smaller than that using the PI controller at a significant level.

TABLE 3: AMSE with different estimates of the process gain. The true process gain $\beta = 2.5$.

n_t is ARIMA(1, 1, 1) $\phi = 0.7$ $\theta = -0.3$					
$\hat{\beta}$	ARIMA controller ($\hat{\phi} = 0.7$ $\hat{\theta} = -0.3$)	PI controller ($k_p = 1.223$ $k_I = 0.190$)	$\hat{\beta}$	ARIMA controller ($\hat{\phi} = 0.7$ $\hat{\theta} = -0.3$)	PI controller ($k_p = 1.223$ $k_I = 0.190$)
	AMSE (SEAMSE)	AMSE (SEAMSE)		AMSE (SEAMSE)	AMSE (SEAMSE)
1.7	2.207 (.020)	5.113 (.084)	2.6	1.017 (.005)	1.088 (.005)
1.8	1.583 (.011)	2.213 (.021)	2.7	1.024 (.005)	1.099 (.005)
1.9	1.320 (.008)	1.597 (.011)	2.8	1.036 (.005)	1.115 (.006)
2.0	1.183 (.006)	1.343 (.008)	2.9	1.051 (.005)	1.135 (.006)
2.1	1.104 (.006)	1.215 (.006)	3.0	1.068 (.005)	1.159 (.006)
2.2	1.058 (.005)	1.145 (.006)	3.1	1.087 (.005)	1.186 (.006)
2.3	1.032 (.005)	1.107 (.005)	3.2	1.108 (.005)	1.215 (.007)
2.4	1.018 (.005)	1.089 (.005)	3.3	1.131 (.006)	1.246 (.007)
2.5	1.014 (.005)	1.084 (.005)	3.4	1.155 (.006)	1.279 (.007)

TABLE 4: AMSE with uncertainties in the ARIMA parameters.

ARIMA controller $\hat{\phi}_1 = 0.7$ $\hat{\theta}_1 = -0.3$					
ϕ_1	$\theta_1 = -0.5$	-0.4	-0.3	-0.2	-0.1
0.5	1.177 (.006)	1.109 (.005)	1.061 (.005)	1.034 (.005)	1.027 (.005)
0.6	1.106 (.005)	1.056 (.005)	1.027 (.005)	1.018 (.005)	1.031 (.005)
0.7	1.052 (.005)	1.022 (.005)	1.014 (.005)	1.028 (.005)	1.064 (.005)
0.8	1.024 (.005)	1.021 (.005)	1.044 (.005)	1.091 (.006)	1.162 (.006)
0.9	1.073 (.006)	1.129 (.007)	1.218 (.008)	1.340 (.010)	1.495 (.012)
0.95	1.248 (.010)	1.398 (.013)	1.596 (.018)	1.843 (.023)	2.138 (.029)

PI controller $k_p = 1.223$ $k_I = 0.190$					
ϕ_1	$\theta_1 = -0.5$	-0.4	-0.3	-0.2	-0.1
0.5	1.199 (.006)	1.137 (.006)	1.097 (.005)	1.078 (.005)	1.081 (.005)
0.6	1.136 (.006)	1.095 (.005)	1.077 (.005)	1.081 (.005)	1.109 (.005)
0.7	1.092 (.005)	1.076 (.005)	1.084 (.005)	1.116 (.006)	1.172 (.006)
0.8	1.074 (.005)	1.087 (.005)	1.128 (.006)	1.197 (.007)	1.293 (.008)
0.9	1.094 (.005)	1.151 (.006)	1.242 (.007)	1.366 (.009)	1.523 (.011)
0.95	1.129 (.006)	1.220 (.007)	1.348 (.009)	1.515 (.011)	1.720 (.014)

The corresponding SEAMSE's are included in the parentheses.

FIGURE 1: The performance of the ARIMA controller and the PI controller when both measurement error (σ_e) and adjustment error (σ_w) exist and vary.

FIGURE 2: The performance of the ARIMA controller and the PI controller when the process initial bias varies.

4.3. Effect of Estimation Uncertainties in the Process Gain. In order to focus on the effect of the estimate of the process gain brought to the controller's performance, we assume both

measurement error and adjustment error are absent, that is, $e_t = w_t = 0$; the initial bias g_0 is 0, that is, $\hat{g}_0 = \sigma_g = 0$; and all the ARIMA parameters are accurately estimated, that is, $\hat{\phi}_1 = \phi_1$ and $\hat{\theta}_1 = \theta_1$.

FIGURE 3: The performance of the ARIMA controller and the PI controller when the process gain is not accurately estimated.

FIGURE 4: Contour plots of AMSE for the ARIMA controller and the PI controller when the disturbance parameters (ϕ_1, θ_1) vary.

We set the process gain β to 2.5 and assume $\hat{\beta} \in [1.7, 3.4]$. Table 3 and Figure 3 present the performance of the two controllers when $\hat{\beta}$ varies. It is shown that the ARIMA controller always outperforms the PI controller for all the $\hat{\beta}$ s we examined at a significant level. Additionally, underestimated β (i.e., $\hat{\beta}/\beta < 1$) will hurt both controllers' performance more than overestimated β (i.e., $\hat{\beta}/\beta > 1$). It is shown that if the process gain β is underestimated to some degree, both controllers would completely fail.

4.4. Effect of Estimation Uncertainties in the Disturbance.

Again, in order to focus on the effect of the estimation uncertainties in the disturbance, we assume $e_t = w_t = \hat{g}_0 = \sigma_g = 0$ and $\hat{\beta} = \beta$. Suppose the disturbance parameters are estimated as $\hat{\phi}_1 = 0.7$ and $\hat{\theta}_1 = -0.3$. However, the disturbance's true parameters may not be equal to the estimates, that is, $\phi_1 \neq \hat{\phi}_1$ and/or $\theta_1 \neq \hat{\theta}_1$. We choose 5 values of ϕ_1 and 5 values of θ_1 as the true parameters for the ARIMA(1, 1, 1) disturbance, so there are a total of 25 disturbance ARIMA(1, 1, 1) models that are studied to investigate both controllers' robustness.

Table 4 reports the results of the AMSE of y_t using the ARIMA controller and the PI controller for different ARIMA(1, 1, 1) disturbances with varying ϕ_1 and θ_1 . It is shown that the ARIMA controller's performance is better than the PI controller's in most of the cases at a significant level. Only when $\phi_1 = 0.95$ is the PI controller's performance better than the ARIMA controller's. This is because a ARIMA(1, 1, 1) process will become a ARIMA(0, 2, 1)

process when $\phi_1 = 1$. As this ARIMA controller is purposely designed to control a ARIMA(1, 1, 1) disturbance, its performance will be contaminated when ϕ_1 is close to 1. The results also imply that if a ARIMA(0, 2, 1) disturbance is misidentified to be a ARIMA(1, 1, 1) disturbance, the ARIMA controller would have a bad performance.

For illustration, Figure 4 further presents the contour plots for the AMSE as a function of ϕ_1 and θ_1 by using the ARIMA controller and the PI controller, respectively. It can be generally observed that both controllers that are designed based on underestimated disturbance parameters (i.e., $\hat{\phi}_1 < \phi_1$, $\hat{\theta}_1 < \theta_1$) will have worse performance than the ones designed based on overestimated disturbance parameters (i.e., $\hat{\phi}_1 > \phi_1$, $\hat{\theta}_1 > \theta_1$). Specially, as can be seen from Figure 4(a), when ϕ_1 is not greater than 0.9, the AMSE of y_t by using the ARIMA controller changes very slowly with varying ϕ_1 and θ_1 . However, when ϕ_1 is greater than 0.9, the AMSE increases very quickly with ϕ_1 increasing. The parameter θ_1 has less effects than ϕ_1 on the performance of the ARIMA controller. In comparison with the ARIMA controller, the PI controller's performance is affected by both ϕ_1 and θ_1 in almost the same degree as shown in Figure 4(b). In all the change area of ϕ_1 and θ_1 except the area $\phi_1 > 0.9$, the PI controller's performance is worse than the ARIMA controller's. However, the maximum AMSE (the worst case) used by the PI controller in the whole change area of ϕ_1 and θ_1 is smaller than that used by the ARIMA controller. In this sense, we can say that the PI controller's performance is more robust than the ARIMA controller's when uncertainties exist in the ARIMA parameters.

5. An Illustrative Example

The studies in the last section are carried out based on an ARIMA(1, 1, 1) disturbance. As aforementioned, the proposed ARIMA controller can be applied to any general ARIMA(p, d, q) disturbance. For illustration purpose, we assume a process as follows:

$$y_t = x_{t-1} + n_t \quad (30)$$

and $(1 - 0.2\mathcal{B})(1 - \mathcal{B})n_t = (1 - 0.1\mathcal{B} - 0.41\mathcal{B}^2 + 0.105\mathcal{B}^3)\varepsilon_t$, $\varepsilon_t \sim \text{WN}(0, 0.8^2)$; y_t is the deviation from target. The disturbance model n_t is the same as one of the ARIMA(1, 1, 3) disturbances presented in [21]. For the sake of simplicity, we further set $\sigma_e = \sigma_w = 0$, $y_0 = 0$, and assume that all the disturbance parameters are known. That is, we ignore measurement errors, adjustment errors, initial bias uncertainties, and disturbance parameter uncertainties and only show the ARIMA controller's superiority in controlling higher order ARIMA disturbance models.

For the above ARIMA(1, 1, 3) disturbance, we can get $m = \max(1, 3) = 3$ and

$$\begin{aligned} \frac{\theta(z)}{\phi(z)} &= \frac{1 - 0.1z - 0.41z^2 + 0.105z^3}{1 - 0.2z} \\ &= 1 + 0.1z - 0.39z^2 + 0.027z^3 + \dots \end{aligned} \quad (31)$$

So, in the offline procedure, the ARIMA controller's parameters are given as follows:

$$\begin{aligned} \mathbf{h} &= \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & -0.39 & 0 & 0 \\ 0 & 0 & 0 & 0.027 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ L &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.2 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \\ \Sigma &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0.64 & 0.64 & 0.64 & 0.64 \\ 0 & 0.64 & 0.64 & 0.64 & 0.64 \\ 0 & 0.64 & 0.64 & 0.64 & 0.64 \\ 0 & 0.64 & 0.64 & 0.64 & 0.64 \end{bmatrix} \quad \mathbf{m} = \begin{bmatrix} 0 \\ 0.1 \\ -0.39 \\ 0.027 \\ 1 \end{bmatrix}. \end{aligned} \quad (32)$$

The initial values $\widehat{\mathbf{R}}_1^- = \mathbf{V}_1^- = 0$. Using the same optimization tuning techniques described in Appendix B, we get $k_p = 1.128$ and $k_I = -0.008$.

We randomly draw one simulation and show the paths of the process output y_t in Figure 5 when the two control

FIGURE 5: One simulated path by using the ARIMA and PI controllers, respectively, for the ARIMA(1,1,3) disturbance that $(1 - 0.2\mathcal{B})(1 - \mathcal{B})n_t = (1 - 0.1\mathcal{B} - 0.41\mathcal{B}^2 + 0.105\mathcal{B}^3)\varepsilon_t$ with $\sigma_e = 0.8$.

algorithms are implemented, respectively. The MSE of y_t implemented by the PI controller is 0.819 while that implemented by the ARIMA controller is 0.646. The paths suggest that the ARIMA controller maintains the process output y_t closer to the target than the PI controller in much more places in the first 100 simulated runs.

6. Concluding Remarks

In this paper, we have developed an optimal control algorithm, the ARIMA controller, for a linear system with a general ARIMA(p, d, q) disturbance, in the presence of measurement errors and adjustment errors together with a random initial bias. We theoretically prove that the ARIMA controller is optimal to both the finite-horizon and infinite-horizon problems. The performances of the ARIMA and PI controllers have been evaluated and compared via Monte Carlo simulations under multiple scenarios. In almost all the analyzed scenarios, the ARIMA controller outperforms the PI controller. Only when uncertainties exist in the ARIMA parameters does the PI controller have more robust performance than the ARIMA controller. Our simulation studies also show that when designing the two controllers, underestimating the process parameters, including the process gain and the ARIMA parameters, will have more negative impact on the controllers' performance than overestimating the parameters. Although the steps to implement the PI controller are simple, a tuning process is always needed to determine the PI parameters k_p and k_I for each given disturbance. It is usually time-consuming. Conversely, the ARIMA controller can automatically determine its algorithm parameters for each given disturbance.

The ARIMA controller is a complement to the ARMA controller which is designed to optimally control any general ARMA(p, q) disturbance. The ARMA controller can be used to control a large class of weakly stationary disturbances while the ARIMA controller can handle a large class of non-stationary or periodic disturbances. Such disturbances are common in rapid thermal processing, reactive ion etching, I-line lithography, and lapping processes in the semiconductor manufacturing.

Appendices

A. Proof of Theorem 1

In the following proof, all the notations are the same as defined in Theorem 1. For the general ARIMA(p, d, q) pro-

cess n_t in (2), Brockwell and Davis [22] already provided its state-space representation. Referring to Example 12.1.7 on page 471 in the book [22], we can get n_t 's state-space representation as follows:

$$n_t = \begin{bmatrix} G' & (-1)^{d+1} \binom{d}{d} & (-1)^d \binom{d}{d-1} & (-1)^{d-1} \binom{d}{d-2} & \cdots & d \end{bmatrix} \begin{bmatrix} \mathbf{X}_t \\ \mathbf{N}_{t-1} \end{bmatrix} + z_t, \quad t = 1, 2, \dots, \quad (\text{A.1})$$

$$\begin{bmatrix} \mathbf{X}_{t+1} \\ \mathbf{N}_t \end{bmatrix} = \begin{bmatrix} F & \mathbf{0} \\ AG' & B \end{bmatrix} \begin{bmatrix} \mathbf{X}_t \\ \mathbf{N}_{t-1} \end{bmatrix} + \begin{bmatrix} Hz_t \\ Az_t \end{bmatrix}, \quad t = 1, 2, \dots \quad (\text{A.2})$$

Now, we will give the state-space representation of the linear system (1)–(5). Based on (1) and (A.1), we can get

$$\begin{aligned} y_t &= g_t + n_t + e_t \\ &= g_t + \begin{bmatrix} G' & (-1)^{d+1} \binom{d}{d} & (-1)^d \binom{d}{d-1} & (-1)^{d-1} \binom{d}{d-2} & \cdots & d \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} \mathbf{X}_t \\ \mathbf{N}_{t-1} \end{bmatrix} + z_t + e_t \\ &= \begin{bmatrix} 1 & G' & (-1)^{d+1} \binom{d}{d} & (-1)^d \binom{d}{d-1} & (-1)^{d-1} \binom{d}{d-2} & \cdots & d \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} g_t \\ \mathbf{X}_t \\ \mathbf{N}_{t-1} \end{bmatrix} + z_t + e_t \\ &= \begin{bmatrix} 1 & G' & (-1)^{d+1} \binom{d}{d} & (-1)^d \binom{d}{d-1} & (-1)^{d-1} \binom{d}{d-2} & \cdots & d \end{bmatrix} \\ &\quad \cdot \mathbf{R}_t + z_t + e_t, \quad t = 1, 2, \dots \end{aligned} \quad (\text{A.3})$$

So, we proved (8) in Theorem 1. Then, based on (4) and (5), we can get

$$g_{t+1} = g_t + \beta \Delta u_t + w_t. \quad (\text{A.4})$$

Combining (A.4) with (A.2), we can obtain the state equation:

$$\begin{bmatrix} g_{t+1} \\ \mathbf{X}_{t+1} \\ \mathbf{N}_t \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & F & \mathbf{0} \\ \mathbf{0} & AG' & B \end{bmatrix} \begin{bmatrix} g_t \\ \mathbf{X}_t \\ \mathbf{N}_{t-1} \end{bmatrix} + \begin{bmatrix} \beta \Delta u_t \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} w_t \\ Hz_t \\ Az_t \end{bmatrix}, \quad t = 1, 2, \dots \quad (\text{A.5})$$

Thus, we proved (9) in Theorem 1. Then, we finish the proof of Theorem 1.

B. Tuning Procedure for PI Parameters

Given each pair of k_p and k_I , the MSE of the first 100 outputs $y_1, y_2, \dots, y_{100}$ is computed. We do 1000 replicates of this procedure and use the AMSE as the performance of the PI controller. Obviously, AMSE is a function of k_p and k_I . Our way to determine k_p and k_I is to solve the following optimization problem:

$$\min_{k_p, k_I} \text{AMSE}(k_p, k_I). \quad (\text{B.1})$$

To solve (B.1), we use the function *optim* in the base package of R software (see [27]). Two types of numerical optimization techniques, the Nelder-Mead method (see [28]) and the Broyden-Fletcher-Goldfarb-Shanno algorithm (see [29–32]), are used, respectively. For the two different numerical optimization techniques, we get the same optimal solution that $k_p = 1.223$ and $k_I = 0.190$ for the ARIMA(1, 1, 1) disturbance with $\phi_1 = 0.7$ and $\theta_1 = -0.3$ investigated in Section 4; and the same optimal solution that $k_p = 1.128$ and $k_I = -0.008$ for the ARIMA(1, 1, 3) disturbance with $\phi_1 = 0.2, \theta_1 = -0.1, \theta_2 = -0.41$, and $\theta_3 = 0.105$ is investigated in Section 5.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 71102145) and the Fundamental Research Funds for the Central Universities of China (Grant no. JBK150144). The authors are grateful to the editors and the anonymous reviewers for their valuable comments and suggestions that have greatly improved the quality of this paper.

References

- [1] R. F. Stengel, *Optimal Control and Estimation*, Dover Publications, New York, NY, USA, 1993.

- [2] A. Cadenillas, T. Choulli, M. Taksar, and L. Zhang, "Classical and impulse stochastic control for the optimization of the dividend and risk policies of an insurance firm," *Mathematical Finance*, vol. 16, no. 1, pp. 181–202, 2006.
- [3] N. Kulenko and H. Schmidli, "Optimal dividend strategies in a Cramer-Lundberg model with capital injections," *Insurance: Mathematics and Economics*, vol. 43, no. 2, pp. 270–278, 2008.
- [4] K. Nakashima, H. Arimitsu, T. Nose, and S. Kuriyama, "Optimal control of a remanufacturing system," *International Journal of Production Research*, vol. 42, no. 17, pp. 3619–3625, 2004.
- [5] H. Liu, Z. B. Zabinsky, and W. Kohn, "Rule-based forecasting and production control system design utilizing a feedback control architecture," *IIE Transactions*, vol. 43, no. 2, pp. 143–152, 2011.
- [6] G. Singer and E. Khmelnitsky, "A finite-horizon, stochastic optimal control policy for a production-inventory system with backlog-dependent lost sales," *IIE Transactions*, vol. 42, no. 12, pp. 855–864, 2010.
- [7] F. He, K. Wang, and W. Jiang, "A general harmonic rule controller for run-to-run process control," *IEEE Transactions on Semiconductor Manufacturing*, vol. 22, no. 2, pp. 232–244, 2009.
- [8] Z. K. Nagy and R. D. Braatz, "Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis," *Journal of Process Control*, vol. 14, no. 4, pp. 411–422, 2004.
- [9] J. Shi and S. Zhou, "Quality control and improvement for multistage systems: a survey," *IIE Transactions (Institute of Industrial Engineers)*, vol. 41, no. 9, pp. 744–753, 2009.
- [10] E. del Castillo, "Statistical process adjustment: a brief retrospective, current status, and some opportunities for further work," *Statistica Neerlandica*, vol. 60, no. 3, pp. 309–326, 2006.
- [11] A.-J. Su, J.-C. Jeng, H.-P. Huang, C.-C. Yu, S.-Y. Hung, and C.-K. Chao, "Control relevant issues in semiconductor manufacturing: overview with some new results," *Control Engineering Practice*, vol. 15, no. 10, pp. 1268–1279, 2007.
- [12] F. E. Grubbs, "An optimum procedure for setting machines or adjusting processes," *Journal of Quality Technology*, vol. 15, pp. 186–189, 1983.
- [13] D. Trietsch, "The harmonic rule for process setup adjustment with quadratic loss," *Journal of Quality Technology*, vol. 30, no. 1, pp. 75–84, 1998.
- [14] S. A. Vander Wiel, W. T. Tucker, F. W. Faltin, and N. Doganaksoy, "Algorithmic statistical process control. Concepts and an application," *Technometrics*, vol. 34, no. 3, pp. 286–291, 1992.
- [15] A. Ingolfsson and E. Sachs, "Stability and sensitivity of an EWMA controller," *Journal of Quality Technology*, vol. 25, pp. 271–287, 1993.
- [16] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition, 1994.
- [17] F. He, H. Xie, and K. Wang, "Optimal setup adjustment and control of a process under ARMA disturbances," *IIE Transactions*, vol. 47, no. 3, pp. 230–244, 2015.
- [18] J. F. MacGregor, T. J. Harris, and J. D. Wright, "Duality between the control of processes subject to randomly occurring deterministic disturbances and ARIMA stochastic disturbances," *Technometrics*, vol. 26, no. 4, pp. 389–397, 1984.
- [19] F. Tsung, H. Wu, and V. N. Nair, "On the efficiency and robustness of discrete proportional-integral control schemes," *Technometrics*, vol. 40, no. 3, pp. 214–222, 1998.
- [20] D. W. Apley, "A cautious minimum variance controller with ARIMA disturbances," *IIE Transactions*, vol. 36, no. 5, pp. 417–432, 2004.
- [21] M.-D. Ma, C.-C. Chang, S.-S. Jang, and D. S.-H. Wong, "Mixed product run-to-run process control—an ANOVA model with ARIMA disturbance approach," *Journal of Process Control*, vol. 19, no. 4, pp. 604–614, 2009.
- [22] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer, New York, NY, USA, 2009.
- [23] K. J. Åström, *Introduction To Stochastic Control Theory*, Mathematics in Science and Engineering, Vol. 70, Academic Press, New York, NY, USA, 1970.
- [24] F. L. Lewis, *Optimal Estimation With An Introduction To Stochastic Control Theory*, John Wiley & Sons, New York, NY, USA, 1986.
- [25] E. del Castillo, *Statistical Process Adjustment for Quality Control*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, USA, 2002.
- [26] K. J. Åström and T. Hägglund, *Advanced PID Control*, ISA, Research Triangle Park, NC, USA, 2006.
- [27] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2014, <http://www.R-project.org/>.
- [28] J. A. Nelder and R. Mead, "A simplex algorithm for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.
- [29] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms," *Journal of the Institute of Mathematics and Its Applications*, vol. 6, pp. 76–90, 1970.
- [30] R. Fletcher, "New approach to variable metric algorithms," *Computer Journal*, vol. 13, no. 3, pp. 317–322, 1970.
- [31] D. Goldfarb, "A family of variable-metric methods derived by variational means," *Mathematics of Computation*, vol. 24, pp. 23–26, 1970.
- [32] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Mathematics of Computation*, vol. 24, pp. 647–656, 1970.