Scientific
Research
Publishing

Discrete-Time Nonlinear Stochastic Optimal Control Problem Based on Stochastic Approximation Approach

Sie Long Kek[1], Sy Yi Sim[2], Wah June Leong[3], Kok Lay Teo[4]

[1]Department of Mathematics and Statistics, Universiti Tun Hussein Onn Malaysia, Pagoh, Malaysia
[2]Department of Electrical Engineering Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Malaysia
[3]Department of Mathematics, Universiti Putra Malaysia, Serdang, Malaysia
[4]Department of Mathematics and Statistics, Curtin University of Technology, Perth, Australia
Email: slkek@uthm.edu.my, sysim@uthm.edu.my, leongwj@upm.edu.my, K.L.Teo@curtin.edu.au

Abstract

In this paper, a computational approach is proposed for solving the discrete-time nonlinear optimal control problem, which is disturbed by a sequence of random noises. Because of the exact solution of such optimal control problem is impossible to be obtained, estimating the state dynamics is currently required. Here, it is assumed that the output can be measured from the real plant process. In our approach, the state mean propagation is applied in order to construct a linear model-based optimal control problem, where the model output is measureable. On this basis, an output error, which takes into account the differences between the real output and the model output, is defined. Then, this output error is minimized by applying the stochastic approximation approach. During the computation procedure, the stochastic gradient is established, so as the optimal solution of the model used can be updated iteratively. Once the convergence is achieved, the iterative solution approximates to the true optimal solution of the original optimal control problem, in spite of model-reality differences. For illustration, an example on a continuous stirred-tank reactor problem is studied, and the result obtained shows the applicability of the approach proposed. Hence, the efficiency of the approach proposed is highly recommended.

1. Introduction

Nonlinear optimal control problem, which is disturbed by random noises, is an interesting research topic. In presence of the random noises, the entire state trajectory could not be measured exactly. Due to the nonlinear structure and the fluctuation behavior of the dynamical system, an efficient computational approach is, therefore, necessarily required to estimate the state dynamics. Further from this, the state estimate shall be used to optimize and control the dynamical system, where the optimal control policy is drawn apparently [1] [2] [3] [4] [5]. From literatures, the applications of the nonlinear stochastic optimal control are widely studied, see for examples, vehicle trajectory planning [6], portfolio selection problem [7], building structural system [8], investment in insurance [9], switching system [10], machine maintenance problem [11], nonlinear differential game problem [12], and viscoelastic systems [13].

In recent years, using the linear optimal control model with model-reality differences in solving the nonlinear optimal control problem, especially for discrete-time nonlinear stochastic optimal control problem, is proposed [14] [15] [16] [17]. Such method is known as the integrated optimal control and parameter estimation (IOCPE) algorithm. In this approach, the adjusted parameters are introduced into the model used, so as the differences between the real plant and the model used can be calculated repeatedly. This algorithm is an iterative procedure, where system optimization and parameter estimation are integrated interactively. During the computation procedure, the optimal solution of the model used is updated iteratively. Once the convergence is achieved, the iterative solution of the model used approximates to the true optimal solution of the original optimal control problem, in spite of model-reality differences.

Besides, the applications of the IOCPE algorithm in providing the expectation solution as well as the filtering solution of the discrete-time nonlinear stochastic optimal control problem have been well-demonstrated [14] [15]. In addition, the optimal output solution obtained from the IOCPE algorithm has been improved by using the weighted output residual [16], which is introduced into the model cost function, and the output matching scheme [17], where the adjusted parameter is introduced into the model output. Moreover, the application of the approaches on the least-square and the Gauss-Newton with the principle of model-reality differences, which omits from using the adjusted parameters, enhance the practical usage of the IOCPE algorithm for delivering the optimal solution of the original optimal control problem [18] [19].

By virtue of the improvement done, it is simply seen that the efficiency of the IOCPE algorithm for solving the discrete-time nonlinear stochastic optimal control problem is shown. However, we find that the output residual from the Kalman filtering theory could be further reduced, in turn, having an efficient output solution for representing the original output. Hence, in this paper, we aim to improve the accuracy of the output solution of the model used. In our approach, the stochastic approximation approach, which is an iterative stochastic optimization algorithm [20] [21] [22] [23], is applied. The advantage of the

stochastic approximation algorithm is to find the optimum of a function, which cannot be computed directly, but only be estimated from noisy observations [24] [25] [26] [27], and its applications to control systems have been well-defined [28] [29] [30] [31] [32]. This advantage motivates us on applying the stochastic approximation algorithm into the IOCPE algorithm can significantly reduce the output residual compared to those output residual from the Kalman filtering theory. Here, the optimal control law, which is based on the state mean propagation, is constructed. At the end of iteration, the trajectories of state and control, which are in expectation manner, are obtained, while the output trajectory could track the real output closely. Hence, the efficiency of the approach proposed is highly recommended.

The rest of the paper is organized as follows. In Section 2, a general discrete-time nonlinear stochastic optimal control problem is described. In Section 3, the stochastic approximation scheme, which is combined with the principle of model-reality differences, is discussed. The calculation procedure is then formulated as an iterative algorithm. In Section 4, an illustrative example on a continuous stirred-tank reactor problem is studied and the applicability of the approach proposed is presented. Finally, some concluding remarks are made.

2. Problem Statement

Consider a general discrete-time nonlinear stochastic optimal control problem given by

$$\min_{u(k)} J_0(u) = E\left[\varphi(x(N), N) + \sum_{k=0}^{k-1} L(x(k), u(k), k) \right]$$

$$\text{subject to} \quad x(k+1) = f(x(k), u(k), k) + G\omega(k) \tag{1}$$

$$y(k) = h(x(k), k) + \eta(k)$$

where $u(k) \in \mathfrak{R}^m, k = 0, 1, \cdots, N-1$, $x(k) \in \mathfrak{R}^n, k = 0, 1, \cdots, N$ and $y(k) \in \mathfrak{R}^p$, $k = 0, 1, \cdots, N$ are, respectively, control sequence, state sequence and output sequence. The process noise sequence $\omega(k) \in \mathfrak{R}^q, k = 0, 1, \cdots, N-1$ and the measurement noise sequence $\eta(k) \in \mathfrak{R}^p, k = 0, 1, \cdots, N$ are the stationary Gaussian white noise sequences with zero mean and their covariance matrices are, respectively, given by $Q_\omega \in \mathfrak{R}^{q \times q}$ and $R_\eta \in \mathfrak{R}^{p \times p}$, which both are positive definite matrices. While, $G \in \mathfrak{R}^{n \times q}$ is a process noise coefficient matrix, $f : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \to \mathfrak{R}^n$ represents the real plant, and $h : \mathfrak{R}^n \times \mathfrak{R} \to \mathfrak{R}^p$ is the output measurement, whereas $\varphi : \mathfrak{R}^n \times \mathfrak{R} \to \mathfrak{R}$ is the terminal cost and $L : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \to \mathfrak{R}$ is the cost under summation. Here, J_0 is the scalar cost function and $E[\cdot]$ is the expectation operator. It is assumed that all functions in (1) are continuously differentiable with respect to their respective arguments.

The initial state

$$x(0) = x_0,$$

where $x_0 \in \mathfrak{R}^n$ is a random vector with mean and covariance are, respectively, given by

$$E[x_0] = \overline{x}_0 \text{ and } E\left[(x_0 - \overline{x}_0)(x_0 - \overline{x}_0)^{\mathrm{T}}\right] = M_0.$$

Here, $M_0 \in \mathfrak{R}^{n \times n}$ is a positive definite matrix. It is assumed that initial state, process noise and measurement noise are statistically independent.

This problem, which is regarded as the discrete-time stochastic optimal control problem, is referred to as Problem (P). Notice that the exact solution of Problem (P) is, in general, unable to be obtained. Moreover, applying the nonlinear filtering theory to estimate the state of the real plant is computationally demanding. Nevertheless, the output can be measured from the real plant process.

In view of these weaknesses, a linear model-based optimal control problem, which is referred to as Problem (M), is constructed, given by

$$\min_{u(k)} J_1(u) = \frac{1}{2}\overline{x}(N)^{\mathrm{T}} S(N)\overline{x}(N) + \sum_{k=0}^{N-1} \frac{1}{2}\left(\overline{x}(k)^{\mathrm{T}} Q\overline{x}(k) + u(k)^{\mathrm{T}} Ru(k)\right)$$

$$\text{subject to} \quad \overline{x}(k+1) = A\overline{x}(k) + Bu(k) \tag{2}$$

$$\overline{y}(k) = C\overline{x}(k)$$

$$\overline{x}(0) = \overline{x}_0$$

where $\overline{x}(k) \in \mathfrak{R}^n, k = 0, 1, \cdots, N$ and $\overline{y}(k) \in \mathfrak{R}^p, k = 0, 1, \cdots, N$ are, respectively, the expected state sequence and the expected output sequence; $A \in \mathfrak{R}^{n \times n}$ is a state transition matrix, $B \in \mathfrak{R}^{n \times m}$ is a control coefficient matrix, and $C \in \mathfrak{R}^{p \times n}$ is an output coefficient matrix, while $S(N) \in \mathfrak{R}^{n \times n}$ and $Q \in \mathfrak{R}^{n \times n}$ are positive semi-definite matrices and $R \in \mathfrak{R}^{m \times m}$ is a positive definite matrix. Here, J_1 is the scalar cost function.

It is emphasized that only solving Problem (M) would not give the optimal solution of Problem (P). However, by establishing an efficient matching scheme based on the output error, which is the differences between the real output and the model output, to Problem (M), it is possible to obtain the optimal solution of Problem (P) as solving Problem (M) iteratively. In this point of view, we are motivated to look into the possibility of constructing an expanded optimal control model with the output error. This model formulation is for obtaining the true optimal solution of Problem (P) despite model-reality differences.

3. Optimal Control with Stochastic Approximation

Now, let us define the expanded optimal control problem, which is referred to as Problem (E), is formulated by

$$\min_{\alpha(k)} J_2(\alpha) = \min_{u(k)} J_1(u) + \frac{1}{2}\sum_{k=0}^{N} \alpha(k)^{\mathrm{T}} \alpha(k)$$

$$\text{subject to} \quad \overline{x}(k+1) = A\overline{x}(k) + Bu(k)$$

$$\overline{y}(k) = C\overline{x}(k) + \alpha(k)$$

$$\overline{x}(0) = \overline{x}_0 \tag{3}$$

$$\hat{y}(k) + \alpha(k) = y(k)$$

$$\hat{y}(k) = \overline{y}(k)$$

where $\hat{y}(k) \in \Re^p, k = 0,1,\cdots,N$ is introduced to separate the output sequence from the respective signals in the output error problem. It is important to note that the algorithm is to be designed such that the constraint $\hat{y}(k) = \overline{y}(k)$ will be satisfied at the end of the iterations. In this situation, the output $\hat{y}(k)$ will be used for the output error problem and the establishment of the matching scheme, whereas the corresponding output $\overline{y}(k)$ will be reserved for the model output after optimizing the model-based optimal control problem. Here, the output error is defined as

$$\alpha(k) = y(k) - \hat{y}(k), k = 0,1,\cdots,N. \tag{4}$$

3.1. Necessary Optimality Conditions

Define the Hamiltonian function as follows

$$\begin{aligned}
H_e(k) &= \frac{1}{2}\left(\overline{x}(k)^{\mathrm{T}} Q\overline{x}(k) + u(k)^{\mathrm{T}} Ru(k)\right) + \frac{1}{2}\alpha(k)^{\mathrm{T}} \alpha(k) \\
&\quad + p(k+1)^{\mathrm{T}}\left(A\overline{x}(k) + Bu(k)\right) - r(k)^{\mathrm{T}} \overline{y}(k) \\
&\quad + q(k)^{\mathrm{T}}\left(C\overline{x}(k) + \alpha(k) - \overline{y}(k)\right)
\end{aligned} \tag{5}$$

then, the augmented cost function becomes

$$\begin{aligned}
J_2'(\alpha) &= \frac{1}{2}\overline{x}(N)^{\mathrm{T}} S(N)\overline{x}(N) + \frac{1}{2}\alpha(N)^{\mathrm{T}} \alpha(N) + p(0)^{\mathrm{T}} \overline{x}(0) \\
&\quad - p(N)^{\mathrm{T}} \overline{x}(N) + \sum_{k=0}^{N-1} H_e(k) - p(k)^{\mathrm{T}} \overline{x}(k) + r(k)^{\mathrm{T}} \hat{y}(k) \\
&\quad + s(k)^{\mathrm{T}}\left(y(k) - \hat{y}(k) - \alpha(k)\right)
\end{aligned} \tag{6}$$

where $p(k), q(k), r(k)$ and $s(k)$ are the appropriate multipliers to be determined later.

Applying the calculus of variation [2] [14] [33] to the augmented cost function (6), the following necessary optimality conditions are obtained:

1) Stationary condition:

$$Ru(k) + B^{\mathrm{T}} p(k+1) = 0 \tag{7a}$$

2) Co-state equation:

$$p(k) = Q\overline{x}(k) + A^{\mathrm{T}} p(k+1) \tag{7b}$$

3) State equation:

$$\overline{x}(k+1) = A\overline{x}(k) + Bu(k) \tag{7c}$$

with the boundary conditions $\overline{x}(0) = \overline{x}_0$ and $p(N) = 0$.

4) Output equation:

$$\overline{y}(k) = C\overline{x}(k) + \alpha(k) \tag{7d}$$

5) Separable variables:

$$\hat{y}(k) = \overline{y}(k), \ \hat{p}(k) = p(k) \tag{7e}$$

with the multipliers $r(k) = -q(k)$, $s(k) = q(k)$ and $q(k) = 0$.

In view of these necessary optimality conditions, the conditions (7a), (7b) and (7c) are the necessary conditions for Problem (M), while the necessary condition (7d) is an adjustable output measurement. Notice that with this adjustable output, the real output could be tracked by the model output as closely as possible once the output residual is significantly minimized.

3.2. Feedback Optimal Control Law

From (7a), the feedback optimal control law can be calculated from

$$u(k) = -K(k)\overline{x}(k), k = 0,1,\cdots,N-1 \tag{8}$$

where

$$K(k) = \left(R + B^{\mathrm{T}}S(k+1)B\right)^{-1}B^{\mathrm{T}}S(k+1)A, \tag{9a}$$

$$S(k) = Q + A^{\mathrm{T}}S(k+1)\left(A - BK(k)\right). \tag{9b}$$

For more detail, see [14] [18] [19] [33] for the proof of the derivation on this feedback optimal control law.

Applying (8), the state equation is written as

$$\overline{x}(k+1) = \left(A - BK(k)\right)\overline{x}(k), k = 0,1,\cdots,N-1 \tag{10}$$

and the co-state equation is given by

$$p(k) = S(k)\overline{x}(k), k = 0,1,\cdots,N. \tag{11}$$

3.3. Stochastic Approximation Scheme

In general, the recursive equation for the stochastic approximation (SA) algorithm [28] [30] [31] [32] is defined by

$$\theta(k+1) = \theta(k) - a(k)g(\theta(k),k) \tag{12}$$

where $\theta(k)$ is the set of the parameters to be estimated, $g(\theta(k),k)$ is the stochastic gradient, and $a(k)$ is the gain sequence. On this basis, refer to Problem (E), let us define $\theta(k) = (u(k),\overline{x}(k),\hat{y}(k))^{\mathrm{T}}$ and the stochastic gradient, which is assumed to be measurable for the objective function given in (3), is introduced as

$$g(\theta(k),k) = \left(\frac{\partial J_2(k)}{\partial u(k)}, \frac{\partial J_2(k)}{\partial \overline{x}(k)}, \frac{\partial J_2(k)}{\partial \hat{y}(k)}\right)^{\mathrm{T}}.$$

Refer to the SA algorithm (12), it leads to the following iterative equations:

$$u(k)^{i+1} = u(k)^i - a(k)\frac{\partial J_2(\alpha)}{\partial u(k)^i} \tag{13a}$$

$$\overline{x}(k)^{i+1} = \overline{x}(k)^i - a(k)\frac{\partial J_2(\alpha)}{\partial \overline{x}(k)^i} \tag{13b}$$

$$\hat{y}(k)^{i+1} = \hat{y}(k)^i - a(k)\frac{\partial J_2(\alpha)}{\partial \hat{y}(k)^i}. \tag{13c}$$

These equations would be used to update the optimal solution of Problem (E), in turn, to approximate the optimal solution of Problem (P), in spite of model-reality differences.

Consequently, to evaluate the stochastic gradient, rewrite the output error defined in (4), for $k = k+1$, as

$$\alpha(k+1) = y(k+1) - \hat{y}(k+1) = y(k+1) - \overline{y}(k+1) \tag{14}$$

where the separable variable in (7e) is satisfied. After that, taking the expected output measured (7d) for $k = k+1$, and substituting $\overline{x}(k+1)$ by the state equation (10), we have

$$\alpha(k+1) = \frac{1}{2}\left(y(k+1) - C\left(A\overline{x}(k) + Bu(k)\right)\right). \tag{15}$$

Hence, from the objective function (3) in Problem (E), the stochastic gradient, which the chain rule differentiation is applied, is calculated from

$$\frac{\partial J_2(\alpha)}{\partial u(k)} = \left[\frac{\partial \alpha(k)}{\partial u(k)}\right]^{\mathrm{T}} \left[\frac{\mathrm{d}J_2(k)}{\mathrm{d}\alpha(k)}\right] = -(CB)^{\mathrm{T}}\alpha(k) \tag{16a}$$

$$\frac{\partial J_2(\alpha)}{\partial \overline{x}(k)} = \left[\frac{\partial \alpha(k)}{\partial \overline{x}(k)}\right]^{\mathrm{T}} \left[\frac{\mathrm{d}J_2(k)}{\mathrm{d}\alpha(k)}\right] = -(CA)^{\mathrm{T}}\alpha(k) \tag{16b}$$

$$\frac{\partial J_2(\alpha)}{\partial \hat{y}(k)} = \left[\frac{\partial \alpha(k)}{\partial \hat{y}(k)}\right]^{\mathrm{T}} \left[\frac{\mathrm{d}J_2(k)}{\mathrm{d}\alpha(k)}\right] = -\alpha(k) \tag{16c}$$

On the other hand, the gain sequence $a(k)$, which is given in (12), has the asymptotic normality and its convergence property has been well-defined [20] [24] [26] [30] [31]. In particular, the formulation form of the gain sequence $a(k)$ is given from

$$a(k) = \frac{a}{(k+1+A)^b} \tag{17}$$

where a and b are strictly positive and the stability constant $A \geq 0$. The practical value of b is 0.602, which provides the generally more desirable slowly decaying gain (17).

3.4. Computational Algorithm

From the discussion above, the resulting algorithm provides the optimal solution of the linear model-based optimal control problem. This optimal solution is then updated based on the stochastic approximation algorithm to approximate the true optimal solution of the original optimal control problem. As a result, the computation procedure of the iterative algorithm is summarized as follows.

Iterative algorithm with SA scheme

Data: Given $A, B, C, G, Q, R, Q_\omega, R_\eta, S(N), M_0, \overline{x}_0, N, f, L, h, \varphi$.

Step 0: Compute a nominal solution. Calculate $K(k)$ and $S(k)$ from (9a) and (9b), respectively, Then, solve Problem (M) defined by (2) to obtain

$u(k), k = 0,1, \cdots, N-1$, $\overline{x}(k), k = 0,1, \cdots, N$ and $\overline{y}(k), k = 0,1, \cdots, N$. Set $i = 0$, $u(k)^0 = u(k)$, $\overline{x}(k)^0 = \overline{x}(k)$ and $\overline{y}(k)^0 = \overline{y}(k)$.

Step 1: Compute the output error $\alpha(k)^i, k = 0,1, \cdots, N$ from (4).

Step 2: With the determined $\alpha(k)^i$, solve Problem (E) defined by (3) to obtain the new $u(k)^i, k = 0,1, \cdots, N-1$, the new $\overline{x}(k)^i, k = 0,1, \cdots, N$, and the new $\overline{y}(k)^i, k = 0,1, \cdots, N$, respectively, from (8), (10) and (7d).

Step 3: Update the optimal solution given, respectively, by (13a), (13b) and (13c). If $u(k)^{i+1} = u(k)^i, k = 0,1, \cdots, N-1$, $\overline{x}(k)^{i+1} = \overline{x}(k)^i, k = 0,1, \cdots, N$ and $\overline{y}(k)^{i+1} = \overline{y}(k)^i, k = 0,1, \cdots, N$, within a given tolerance, stop; else set $i = i+1$ and repeat from Step 1.

Remarks

1) The off-line computation is done, as stated in Step 0, to calculate $K(k), k = 0,1, \cdots, N-1$ and $S(k), k = 0,1, \cdots, N$, for the control law design. Then, these parameters are used for solving Problem (M) in Step 0 and for solving Problem (E) in Step 2, respectively.

2) The variable $\alpha(k)^i$ is zero in Step 0 and the calculated value of $\alpha(k)^i$ changes from iteration to iteration.

3) Problem (P) is not necessary to be linear or to have a quadratic cost function.

4) The conditions $u(k)^{i+1} = u(k)^i$ and $\overline{x}(k)^{i+1} = \overline{x}(k)^i$ are required to be satisfied for the converged optimal control sequence and the converged state estimate sequence. The following averaged 2-norms are computed and then they are compared with a given tolerance to verify the convergence of $u(k)$ and $\overline{x}(k)$:

$$\left\| u^{i+1} - u^i \right\|_2 = \left(\frac{1}{N-1} \sum_{k=0}^{N-1} \left\| u(k)^{i+1} - u(k)^i \right\| \right)^{1/2} \tag{18a}$$

$$\left\| \overline{x}^{i+1} - \overline{x}^i \right\|_2 = \left(\frac{1}{N} \sum_{k=0}^{N} \left\| \overline{x}(k)^{i+1} - \overline{x}(k)^i \right\| \right)^{1/2} \tag{18b}$$

5) The gain sequence $a(k)$, which is considered in the algorithm proposed, is

$$a(k) = \frac{2}{(1+k)^{0.602}} \tag{19}$$

where $A = 0$ from (17).

4. Illustrative Example

Consider the optimal control of a continuous stirred-tank reactor problem [34]:

$$\min_{u(k)} J_0(u) = 0.01 \sum_{k=0}^{76} E \left[(x_1(k))^2 + (x_2(k))^2 + 0.1(u(k))^2 \right]$$

subject to

$$x_1(k+1) = x_1(k) - 0.02(x_1(k)+0.25) + 0.01(x_2(k)+0.5)\exp\left[\frac{25x_1(k)}{x_1(k)+2}\right]$$

$$-0.01(x_1(k)+0.25)u(k) + \omega_1(k)$$

$$x_2(k+1) = 0.99x_2(k) - 0.005 - 0.01(x_2(k)+0.5)\exp\left[\frac{25x_1(k)}{x_1(k)+2}\right] + \omega_2(k)$$

$$y(k) = x_1(k) + \eta(k)$$

with the initial condition

$$x_1(0) = 0.05, \quad x_2(0) = 0.$$

Here, $\omega(k) = \begin{bmatrix} \omega_1(k) & \omega_2(k) \end{bmatrix}^{\mathrm{T}}$ and $\eta(k)$ are Gaussian white noise sequences with their respective covariance given by $Q_\omega = 10^{-3} I_2$ and $R_\eta = 10^{-3}$.

This problem is referred to as Problem (P).

The linear model-based optimal control problem, which is simplified from Problem (P) and is referred to as Problem (M), is defined by

$$\min_{u(k)} J_1(u) = \frac{1}{2}\sum_{k=0}^{76}\left[(\bar{x}_1(k))^2 + (\bar{x}_2(k))^2 + 0.1(u(k))^2\right]$$

subject to

$$\begin{bmatrix} \bar{x}_1(k+1) \\ \bar{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 1.0895 & 0.0184 \\ -0.1095 & 0.9716 \end{bmatrix}\begin{bmatrix} \bar{x}_1(k) \\ \bar{x}_2(k) \end{bmatrix} + \begin{bmatrix} -0.003 \\ 0.000 \end{bmatrix}u(k)$$

$$\bar{y}(k) = \bar{x}_1(k) + \alpha(k)$$

with the initial condition

$$\bar{x}_1(0) = 0.05, \quad \bar{x}_2(0) = 0$$

and the adjusted parameter $\alpha(k)$ is added into the output measurement channel.

By running the approach proposed, the simulation result is shown in Table 1, where it is compared to the result of the filtering solution [15]. It can be seen that the iteration number of the approach proposed is more than the iteration number of filtering model, and the final cost of the approach proposed is greater than the final cost of filtering model. But, it is found that the output residual of the approach proposed is dramatically reduced to 0.000216 unit, which is a 99 percent reduction. This percentage shows that the model output solution obtained by the approach proposed is significantly closely to the real output trajectory. Hence, this indicates that the approach proposed is practically useful in obtaining the real output solution.

The trajectories of control, state, and output are, respectively, shown in Figures 1-3. It is noticed that the trajectories of control and state are smoothly freely from the disturbance of random noise sequences. This is because of they are an ideal deterministic optimal solution to the nonlinear model-based optimal control problem. However, the real output that is disturbed by the random noise sequences is really fluctuated. By applying the approach proposed, the model

Table 1. Simulation result.

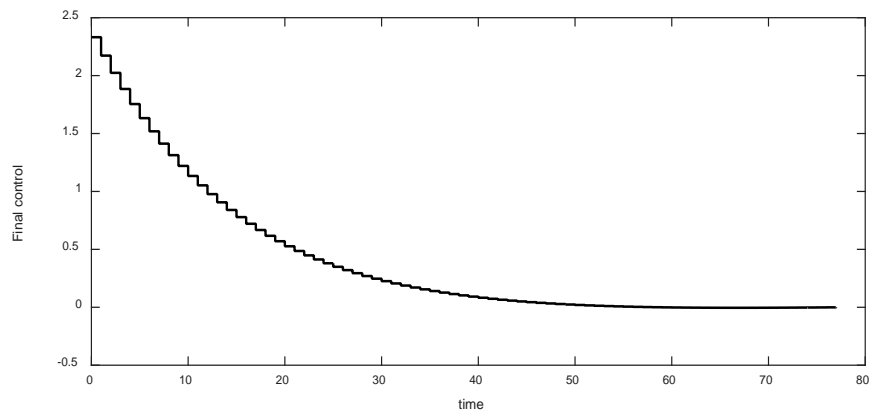| Model	Iteration Number	Elapsed Time	Initial Cost	Final Cost	Output Residual
Filtering	6	0.811	3.7910	0.0215	0.034731
Proposed	19	0.079	2.0973	2.0849	0.000216

Figure 1. Control trajectory.
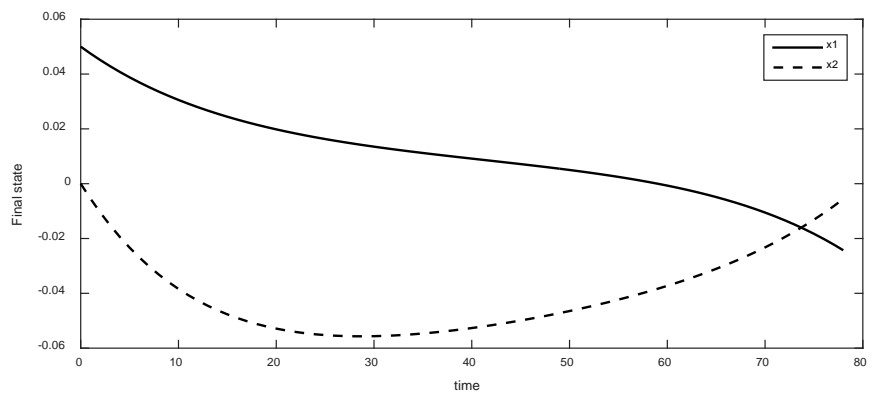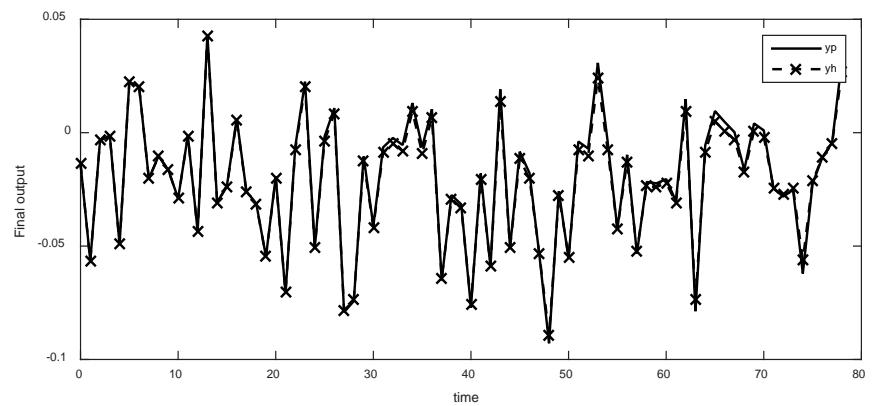

Figure 2. State trajectories.

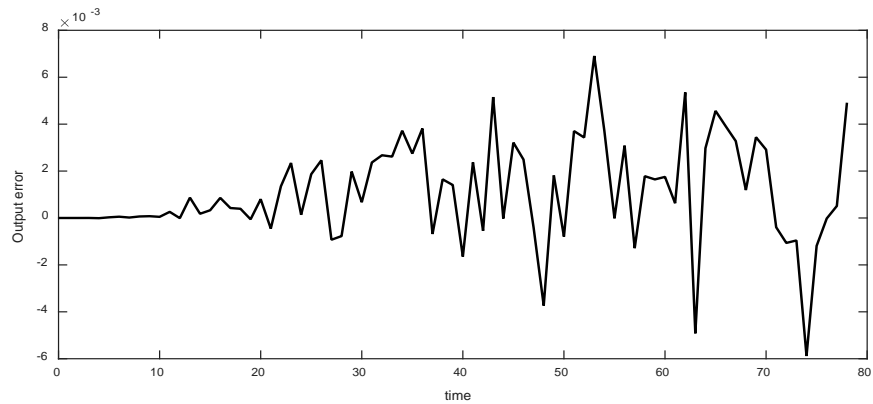**Figure 3.** Output trajectories.

241

Figure 4. Output error.

output trajectory could follow the real output trajectory as closely as possible. Additionally, the output error, which is presenting the differences between the real output and the model output, is shown in **Figure 4**. As a result of this, it is concluded that the approach proposed is efficient and its applicability is demonstrated.

5. Concluding Remarks

Applying the stochastic approximation scheme into the IOCPE algorithm was discussed in this paper. The aim is to improve the output solution of the model used. From previous studies, the IOCPE algorithm is for solving the discrete-time nonlinear stochastic optimal control problem, while the stochastic approximation is for the stochastic optimization. In combining these two approaches, the state mean propagation is constructed, where the adjusted parameter is added into the model output used. During the calculation procedure, the differences between the real plant and the model used are taken into account for updating the iterative solution repeatedly. On the other hand, the least square output error is established such that the stochastic gradient is derived. Consequently, the iterative solution approximates to the optimal solution of the original optimal control problem, in spite of model-reality differences. For illustration, an example on a continuous stirred-tank reactor problem was studied to show the applicability of the approach proposed. In conclusion, the efficiency of the approach proposed is highly recommended.

For the future research direction, it is suggested to apply the SA algorithm to solve the linear model-based optimal control problem, without calculating the adjusted parameter, in order to obtain the true optimal solution of the nonlinear optimal control problem. The result would be compared to the result which is obtained by using the Gauss-Newton method [18] [19]. Hence, the calculation procedure in the IOCPE could be simplified.

References

[1] Kalman, R.E. (1960) Contributions to the Theory of Optimal Control. *Boletín de la*

Sociedad Matemática Mexicana, **5**, 102-119.

[2] Bryson, A.E. and Ho, Y.C. (1975) Applied Optimal Control. Hemisphere, Washington, DC.

[3] Bagchi, A. (1993) Optimal Control of Stochastic Systems. Prentice-Hall, New York.

[4] Ahmed, N.U. (1999) Linear and Nonlinear Filtering for Scientists and Engineers. World Scientific Publishers, Singapore. https://doi.org/10.1142/3911

[5] Simon, D. (2006) Optimal State Estimation: Kalman, H-Infinity and Nonlinear Approaches. John Wiley & Sons, Hoboken, NJ. https://doi.org/10.1002/0470045345

[6] Liu, H.F., Zhang, Y., Chen, S.F. and Chen, J. (2012) Autonomous Vehicle Trajectory Planning under Uncertainty Using Stochastic Collocation. *Advanced Materials Research*, **580**, 175-179. https://doi.org/10.4028/www.scientific.net/AMR.580.175

[7] Zhou, Y. and Wu, Z. (2013) Mean-Variance Portfolio Selection with Margin Requirements. *Journal of Mathematics*, **2013**, Article ID 726297.

[8] Li, X.P., Yu, C., Zhang, J.Y., Zhou, J.J. and Zhang, L.M. (2013) Instantaneous Stochastic Optimal Control of Seismically Excited Structures Based on Time Domain Explicit Method. *Advanced Materials Research*, **790**, 215-218. https://doi.org/10.4028/www.scientific.net/AMR.790.215

[9] Liu, J., Yiu, K.F.C., Loxton, R. and Teo, K.L. (2013) Optimal Investment and Proportional Reinsurance with Risk Constraint. *Journal of Mathematical Finance*, **3**, 437-447. https://doi.org/10.4236/jmf.2013.34046

[10] Abushov, Q. and Aghayeva, C. (2014) Stochastic Maximum Principle for Nonlinear Optimal Control Problem of Switching Systems. *Journal of Computational and Applied Mathematics Part B*, **259**, 371-376. https://doi.org/10.1016/j.cam.2013.06.010

[11] Sun, Y., Aw, G., Loxton, R. and Teo, K.L. (2014) An Optimal Machine Maintenance Problem with Probabilistic State Constraints. *Information Sciences*, **281**, 386-398. https://doi.org/10.1016/j.ins.2014.05.051

[12] Basimanebotlhe, O. and Xue, X. (2014) Stochastic Optimal Control to a Nonlinear Differential Game. *Advances in Difference Equations*, **2014**, 1-14. https://doi.org/10.1186/1687-1847-2014-266

[13] Xiong, H. and Zhu, W. (2015) Nonlinear Stochastic Optimal Control of Viscoelastic Systems. *Journal of Vibration and Control*, **21**, 1029-1040. https://doi.org/10.1177/1077546313489589

[14] Kek, S.L., Teo, K.L. and Ismail, A.A.M. (2010) An Integrated Optimal Control Algorithm for Discrete-Time Nonlinear Stochastic System. *International Journal of Control*, **83**, 2536-2545. https://doi.org/10.1080/00207179.2010.531766

[15] Kek, S.L., Teo, K.L. and Ismail, A.A.M. (2012) Filtering Solution of Nonlinear Stochastic Optimal Control Problem in Discrete-Time with Model-Reality Differences. *Numerical Algebra, Control and Optimization*, **2**, 207-222. https://doi.org/10.3934/naco.2012.2.207

[16] Kek, S.L., Ismail, A.A.M., Teo, K.L. and Rohanin, A. (2013) An Iterative Algorithm Based on Model-Reality Differences for Discrete-Time Nonlinear Stochastic Optimal Control Problems. *Numerical Algebra, Control and Optimization*, **3**, 109-125. https://doi.org/10.3934/naco.2013.3.109

[17] Kek, S.L., Teo, K.L. and Ismail, A.A.M. (2014) Efficient Output Solution for Nonlinear Stochastic Optimal Control Problem with Model-Reality Differences. *Mathematical Problems in Engineering*, **2014**, Article ID 659506.

[18] Kek, S.L., Li, J. and Teo, K.L. (2017) Least Squares Solution for Discrete Time Nonlinear Stochastic Optimal Control Problem with Model-Reality Differences. *Applied*

Mathematics, **8**, 1-14. https://doi.org/10.4236/am.2017.81001

[19] Kek, S.L., Li, J., Leong, W.J. and Ismail, A.A.M. (2017), A Gauss-Newton Approach for Nonlinear Optimal Control Problem with Model-Reality Differences. *Open Journal of Optimization* (OJOp), **6**, 85-100. https://doi.org/10.4236/ojop.2017.63007

[20] Robbins, H. and Monro, S. (1951) A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, **22**, 400-407. https://doi.org/10.1214/aoms/1177729586

[21] Kiefer, J. and Wolfowitz, J. (1952) Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, **23**, 462-466. https://doi.org/10.1214/aoms/1177729392

[22] Sacks, J. (1958) Asymptotic Distribution of Stochastic Approximation Procedures. *The Annals of Mathematical Statistics*, **29**, 373-405. https://doi.org/10.1214/aoms/1177706619

[23] Martin, R. and Masreliez, C. (1975) Robust Estimation via Stochastic Approximation. *IEEE Transactions on Information Theory*, **21**, 263-271. https://doi.org/10.1109/TIT.1975.1055386

[24] Nemirovski, A. and Yudin, D. (1983) Problem Complexity and Method Efficiency in Optimization. John Wiley, New York.

[25] Polyak, B.T. and Juditsky, A.B. (1992) Acceleration of Stochastic Approximation by Averaging. *SIAM Journal on Control and Optimization*, **30**, 838-855. https://doi.org/10.1137/0330046

[26] Kushner, H.J. and Yin, G.G. (1997) Stochastic Approximation Algorithms and Applications. Springer, New York. https://doi.org/10.1007/978-1-4899-2696-8

[27] Nemirovski, A., Juditsky, A., Lan, G. and Shapiro, A. (2009) Robust Stochastic Approximation Approach to Stochastic Programming. *SIAM Journal on Optimization*, **19**, 1574-1609. https://doi.org/10.1137/070704277

[28] Sin, K.S. and Goodwin, G.C. (1982) Stochastic Adaptive Control Using a Modified Least Squares Algorithm. *Automatica*, **18**, 315-321. https://doi.org/10.1016/0005-1098(82)90091-7

[29] Spall, J.C. and Cristion, J.A. (1998) Model-Free Control of Nonlinear Stochastic Systems with Discrete-Time Measurements. *IEEE Transactions on Automatic Control*, **43**, 1198-1210. https://doi.org/10.1109/9.718605

[30] Spall, J.C. (2000) Adaptive Stochastic Approximation by the Simultaneous Perturbation Method. *IEEE Transactions on Automatic Control*, **45**, 1839-1853. https://doi.org/10.1109/TAC.2000.880982

[31] Spall, J.C. (2003) Introduction to Stochastic Search and Optimization: Estimation, Simulation and Control. John Wiley & Sons, Inc, New York. https://doi.org/10.1002/0471722138

[32] Aksakalli, V. and Ursu, D. (2006) Control of Nonlinear Stochastic Systems: Model-Free Controllers versus Linear Quadratic Regulators. *Proceedings of the* 45*th IEEE Conference on Decision and Control* (*CDC* '06), San Diego, CA, December **2006**, 4145-4150. https://doi.org/10.1109/CDC.2006.377721

[33] Lewis, F.L., Vrabie, V. and Symos, V.L. (20012) Optimal Control. 3rd Edition, John Wiley & Sons, Inc, New York.

[34] Kirk, D.E. (2004) Optimal Control Theory: An Introduction. Dover Publications, Mineola, NY.