We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Smoothing Solution for Discrete-Time Nonlinear Stochastic Optimal Control Problem with Model-Reality Differences

Sie Long Kek , Kok Lay Teo and
Mohd Ismail Abd Aziz

Additional information is available at the end of the chapter

Abstract

In this chapter, the performance of the integrated optimal control and parameter estimation (IOCPE) algorithm is improved using a modified fixed-interval smoothing scheme in order to solve the discrete-time nonlinear stochastic optimal control problem. In our approach, a linear model-based optimal control problem with adding the adjustable parameters into the model used is solved iteratively. The aim is to obtain the optimal solution of the original optimal control problem. In the presence of the random noise sequences in process plant and measurement channel, the state dynamics, which is estimated using Kalman filtering theory, is smoothed in a fixed interval. With such smoothed state estimate sequence that reduces the output residual, the feedback optimal control law is then designed. During the computation procedure, the optimal solution of the modified model-based optimal control problem can be updated at each iteration step. When convergence is achieved, the iterative solution approaches to the correct optimal solution of the original optimal control problem, in spite of model-reality differences. Moreover, the convergence of the resulting algorithm is also given. For illustration, optimal control of a continuous stirred-tank reactor problem is studied and the result obtained shows the efficiency of the approach proposed.

Keywords: fixed-interval smoothing, Kalman filtering theory, model-reality differences, adjustable parameters, iterative solution

1. Introduction

Optimal control approach provides the solution in solving dynamic real-world practical problems. Particularly, the linear problems, which are disturbed by the random noise sequence, have been well-defined with application of the optimal state estimate in designing the optimal feedback control law. In such situation, the optimal state estimator and the optimal controller are designed separately to optimize and control the dynamical systems. This is called the separation principle [1–4]. By virtue of this principle, the research works on stochastic optimal control and applications are growing widely, see for examples, linear systems [5, 6], fleet composition problem [7], optimal parameter selection problems [8], Markov jump process [9], power management [10], multiagent systems [11], portfolio selection model [12], 2-DOF vehicle model [13], sensorimotor system [14], and advertising model [15].

In fact, the exact solution of stochastic optimal control problems is impossible to be obtained, especially for the problems involving nonlinear system dynamics. To obtain an optimal solution of the discrete-time nonlinear stochastic optimal control problem, the integrated optimal control and parameter estimation (IOCPE) algorithm has been proposed to solve this kind of the problem iteratively [16–18]. In this algorithm, the linear quadratic Gaussian (LQG) model is applied to a model-based optimal control problem, where the state estimation procedure is done using the Kalman filtering theory. Based on this model, the adjusted parameters are added into the model so as system optimization and parameter estimation are integrated interactively. On this basis, the differences between the real plant and the model used are measured repeatedly in order to update the optimal solution of the model used. On the other hand, the output that is measured from the real plant is fed back into the model used for the state estimator design. When the convergence is achieved, the iterative solution approaches to the true optimal solution of the original optimal control problem despite model-reality differences. This optimal solution is the optimal filtering solution, which is obtained using the IOCPE algorithm. The efficiency of the IOCPE algorithm has been proven in Refs. [16–18].

However, the output trajectory of the model, which is obtained from the IOCPE algorithm, is less accurate in estimating the exact output measurement of the original optimal control problem. In this chapter, our aim is to improve the IOCPE algorithm using the fixed-interval smoothing approach, where the output residual shall be reduced within an appropriate tolerance to generate a better output trajectory. In our model, the state dynamics, which is disturbed by Gaussian noise sequences, is estimated by using the Kalman filtering theory, and then it is smoothed in a fixed-interval estimation. With such state estimation procedure, we modify the estimation procedure so that a smoothed state estimate is predicted backward in time and is used in designing the feedback optimal control law. It is noticed that the output residual of this smoothed state estimate is smaller than the output residual that is obtained by using the Kalman filtering theory, see [17]. The procedure of the solution method discussed in this chapter is almost the same as that was presented in the study of Kek et al. [17], but the accuracy of the optimal solution with the modified fixed-interval smoothing would be definitely increased.

The structure of the chapter is outlined as follows. In Section 2, the description of a general discrete-time nonlinear stochastic optimal control problem and its simplified model-based optimal control problem is made. In Section 3, an expanded optimal control model is introduced, where system optimization and parameter estimation are integrated mutually. The feedback control law, which is incorporated with the Kalman filtering theory and the fixed-interval smoothing, is designed. Then, the iterative algorithm based on principle of model-reality differences is derived so that discrete-time nonlinear stochastic optimal control problem could be solved. In Section 4, a convergence result for the algorithm proposed is provided. In Section 5, an example of optimal control of a continuous stirred-tank reactor problem is illustrated. Finally, some concluding remarks are made.

2. Problem description

Consider a general class of the dynamical system given below:

$$x(k+1) = f(x(k), u(k), k) + G\omega(k) \tag{1a}$$

$$y(k) = h(x(k), k) + \eta(k) \tag{1b}$$

where $u(k) \in \Re^m, k = 0, 1, ..., N-1, x(k) \in \Re^n, k = 0, 1, ..., N$, and $y(k) \in \Re^p, k = 0, 1, ..., N$ are the control sequence, the state sequence, and the output sequence, respectively. $\omega(k) \in \Re^q, k = 0, 1, ..., N-1$, which is the process noise sequence, and $\eta(k) \in \Re^p, k = 0, 1, ..., N$, which is the measurement noise sequence, are stationary Gaussian white noise sequences with zero mean, and their covariance matrices are given by $Q_\omega \in \Re^{q \times q}$ and $R_\eta \in \Re^{p \times p}$, respectively. Here, both of these covariance matrices are positive definite matrices. In addition, $f: \Re^n \times \Re^m \times \Re \to \Re^n$ represents the real plant and $h: \Re^n \times \Re \to \Re^p$ is the real output measurement, which both are assumed to be continuously differentiable with respect to their respective arguments, whereas $G \in \Re^{n \times q}$ is a process coefficient matrix.

The initial state is

$$x(0) = x_0$$

where $x_0 \in \Re^n$ is a random vector with mean and covariance given, respectively, by

$$E[x(0)] = \bar{x}_0 \text{ and } E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = M_0.$$

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

Therefore, our aim is to find an admissible control sequence $u(k) \in \mathfrak{R}^m, k = 0, 1, ..., N - 1$ subject to the dynamical system given in Eq. (1) such that the scalar cost function

$$J_0(u) = E[\varphi(x(N), N) + \sum_{k=0}^{N-1} L(x(k), u(k), k)] \tag{2}$$

is minimized, where $\varphi : \mathfrak{R}^n \times \mathfrak{R} \to \mathfrak{R}$ is the terminal cost and $L : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \to \mathfrak{R}$ is the cost under summation. It is assumed that these functions are continuously differentiable with respect to their respective arguments.

This problem is regarded as the discrete-time nonlinear stochastic optimal control problem and is referred to as Problem (P).

Notice that, in general, the exact solution of Problem (P) is unable to be obtained and estimating the state of the real plant by applying the nonlinear filtering theory is computationally demanding. Due to these reasons, a smoothing model-based optimal control problem, which is referred to as Problem (M), is proposed by

$$\min_{u(k)} J_m(u) = \frac{1}{2} \hat{x}_s(N)^{\mathrm{T}} S(N) \hat{x}_s(N) + \gamma(N)$$

$$+ \sum_{k=0}^{N-1} (\frac{1}{2} (\hat{x}_s(k)^{\mathrm{T}} Q \hat{x}_s(k) + u(k)^{\mathrm{T}} R u(k)) + \gamma(k)) \tag{3}$$

subject to

$$\hat{x}_s(k) = \hat{x}(k) + K_s(k)(\hat{x}_s(k+1) - \overline{x}(k+1))$$
$$\hat{y}_s(k) = C \hat{x}_s(k)$$

with the following state estimation procedure

$$\overline{x}(k+1) = A\hat{x}(k) + Bu(k) + \alpha_1(k) \tag{4a}$$

$$\hat{x}(k) = \overline{x}(k) + K_f(k)(y(k) - \overline{y}(k)) \tag{4b}$$

$$\overline{y}(k) = C\overline{x}(k) + \alpha_2(k) \tag{4c}$$

where $\hat{x}_s(k) \in \Re^n, k = 0, 1, ..., N$ and $\hat{y}_s(k) \in \Re^p, k = 0, 1, ..., N$ are, respectively, the smoothed state sequence and the smoothed output sequence. The matrices involved are given as follow: A is an $n \times n$ state transition matrix, B is an $n \times n$ control coefficient matrix, C is a $p \times n$ output coefficient matrix, $S(N)$ and Q are $n \times n$ positive semidefinite matrices, and R is a $m \times m$ positive definite matrix. The extra parameters $\alpha_1(k), k = 0, 1, ..., N-1, \alpha_2(k), k = 0, 1, ..., N$, and $\gamma(k), k = 0, 1, ..., N$ are introduced as adjustable parameters.

The state estimation procedure, which is given in (4a), (4b), and (4c), is obviously from the Kalman filtering theory, where $\hat{x}(k) \in \Re^n, k = 0, 1, ..., N-1$ and $\bar{x}(k) \in \Re^n, k = 0, 1, ..., N$ are, respectively, the filtered state sequence and the predicted state sequence, whereas $\bar{y}(k) \in \Re^p, k = 0, 1, ..., N$ is the expected output sequence. The filter and smoother gains, which are $K_f(k) \in \Re^{n \times p}$ and $K_s(k) \in \Re^{n \times n}$, are, respectively, given by

$$K_f(k) = M_x(k)C^{\mathrm{T}}M_y(k)^{-1} \tag{5a}$$

$$K_s(k) = P(k)A^{\mathrm{T}}M_x(k+1)^{-1} \tag{5b}$$

whereas the state error covariance matrices are

$$P(k) = M_x(k) - M_x(k)C^{\mathrm{T}}M_y(k)^{-1}CM_x(k) \tag{6a}$$

$$M_x(k+1) = AP(k)A^{\mathrm{T}} + GQ_\omega G^{\mathrm{T}} \tag{6b}$$

$$P_s(k) = P(k) + K_s(k)(P_s(k+1) - M_x(k+1))K_s(k)^{\mathrm{T}} \tag{6c}$$

and the output error covariance matrix is

$$M_y(k) = CM_x(k)C^{\mathrm{T}} + R_\eta \tag{6d}$$

with the boundary conditions $M_x(0) = M_0$ and $P_s(N) = M_x(N)$ The filtered state error covariance $P(k) \in \Re^{n \times n}$, the predicted state error covariance $M_x(k) \in \Re^{n \times n}$, the smoothed state error covariance $P_s(k) \in \Re^{n \times n}$, and the output error covariance $M_y(k) \in \Re^{p \times p}$ are positive definite matrices.

Here, the cost function given in Eq. (3) is evaluated from the expectation of the quadratic forms [2], both for random and deterministic terms with trace matrix $tr(\cdot)$, which is simplified by

a. $E\left[x(N)^{\mathrm{T}}S(N)x(N)\right] = tr\left(S(N)M_x(N)\right) + \bar{x}(N)^{\mathrm{T}}S(N)\bar{x}(N)$

b. $E\left[x(k)^{\mathrm{T}}Qx(k)\right] = tr\left(QM_x(k)\right) + \bar{x}(k)^{\mathrm{T}}Q\bar{x}(k)$

c. $E\left[u(k)^{\mathrm{T}}Ru(k)\right] = u(k)^{\mathrm{T}}Ru(k)$

d. $E[\gamma(k)] = \gamma(k)$, $E\left[\alpha_1(k)\right] = \alpha_1(k)$, and $E\left[\alpha_2(k)\right] = \alpha_2(k)$.

Follow from this simplification, the trace matrix terms that are depend on the state error covariance matrix are ignored in the model used since they are constant values. In such a way, the cost function of the linear model-based optimal control model could be evaluated.

Notice that the separation principle [1–4] is applied to solving Problem (M), where the optimal feedback control law and the optimal state estimate are designed separately as discussed in [16–18]. Further from this, the accuracy of the optimal state estimate is increased by smoothing the state estimate in the fixed interval [2, 4]. Then, based on this smoothed state estimate, the smoothing optimal control law is designed. On the other hand, the output measured from the real plant is fed back into the model used, in turn, to improve the state estimation procedure and to update the solution of the model used. Moreover, only solving Problem (M) without adding the adjusted parameters into the model used would not approximate to the optimal solution of Problem (P). Hence, by taking the adjusted parameters into the model used and solving Problem (M) iteratively, the correct optimal solution of the original optimal control problem could be obtained, in spite of model-reality differences.

3. Modified smoothing with model-reality differences

Now, let us introduce an expanded optimal control problem with smoothing state estimate, which is referred to as Problem (E), given below:

$$\min_{u(k)} J_e(u) = \tfrac{1}{2}\hat{x}_s(N)^{\mathrm{T}} S(N)\hat{x}_s(N) + \gamma(N)$$

$$+ \sum_{k=0}^{N-1}\left(\tfrac{1}{2}(\hat{x}_s(k)^{\mathrm{T}}Q\hat{x}_s(k) + u(k)^{\mathrm{T}}Ru(k)\right) + \gamma(k)) \tag{7}$$

$$+ \tfrac{1}{2}r_1 \parallel v(k) - u(k) \parallel^2 + \tfrac{1}{2}r_2 \parallel z(k) - \hat{x}_s(k) \parallel^2$$

subject to

$$\hat{x}_s(k) = \hat{x}(k) + K_s(k)(\hat{x}_s(k+1) - \bar{x}(k+1))$$

$$\hat{y}_s(k) = C\hat{x}_s(k)$$

$$\frac{1}{2}z(N)^{\mathrm{T}}S(N)z(N) + \gamma(N) = \varphi(z(N), N)$$

$$\frac{1}{2}(z(k)^{\mathrm{T}}Qz(k) + v(k)^{\mathrm{T}}Rv(k)) + \gamma(k) = L(z(k), v(k), k)$$

$$Az(k) + Bv(k) + \alpha_1(k) = f(z(k), v(k), k)$$

$$Cz(k) + \alpha_2(k) = h(z(k), k)$$

$$v(k) = u(k)$$

$$z(k) = \hat{x}_s(k)$$

where $v(k) \in \Re^m, k = 0, 1, ..., N - 1$ and $z(k) \in \Re^n, k = 0, 1, ..., N$ are introduced to separate the control and the smoothed state from the respective signals in the parameter estimation problem and $\| \cdot \|$ denotes the usual Euclidean norm. The terms $\frac{1}{2}r_1 \| u(k) - v(k) \|^2$ and $\frac{1}{2}r_2 \| \hat{x}_s(k) - z(k) \|^2$ are introduced such that the convexity is improved and the convergence of the iterative algorithm is enhanced. The main purpose of designing the algorithm in this way is to ensure that satisfying of the constraints $v(k) = u(k)$ and $z(k) = \hat{x}_s(k)$ is fulfilled at the end of the iterations. More specifically, applying the state estimate $z(k)$ and the control $v(k)$ for the computation in the parameter estimation and the matching schemes will increase the practical usage of the algorithm. Moreover, implementing the relevant smoothed state $\hat{x}_s(k)$ and control $u(k)$ that will be reserved for optimizing the model-based optimal control problem leads the iterative solution toward to the true optimal solution of the original optimal control problem.

Figure 1 shows the block diagram of the approach proposed. The methodology of the approach proposed is further discussed in the following sections.

From the block diagram in **Figure 1**, the definition of the principle of model-reality differences could be given.

Definition 3.1: Principle of model-reality differences is a unified framework, which integrates system optimization and parameter estimation interactively to define an expanded optimal

control problem, aims to give the correct optimal solution of the original optimal control problem by solving the model-based optimal control problem iteratively.
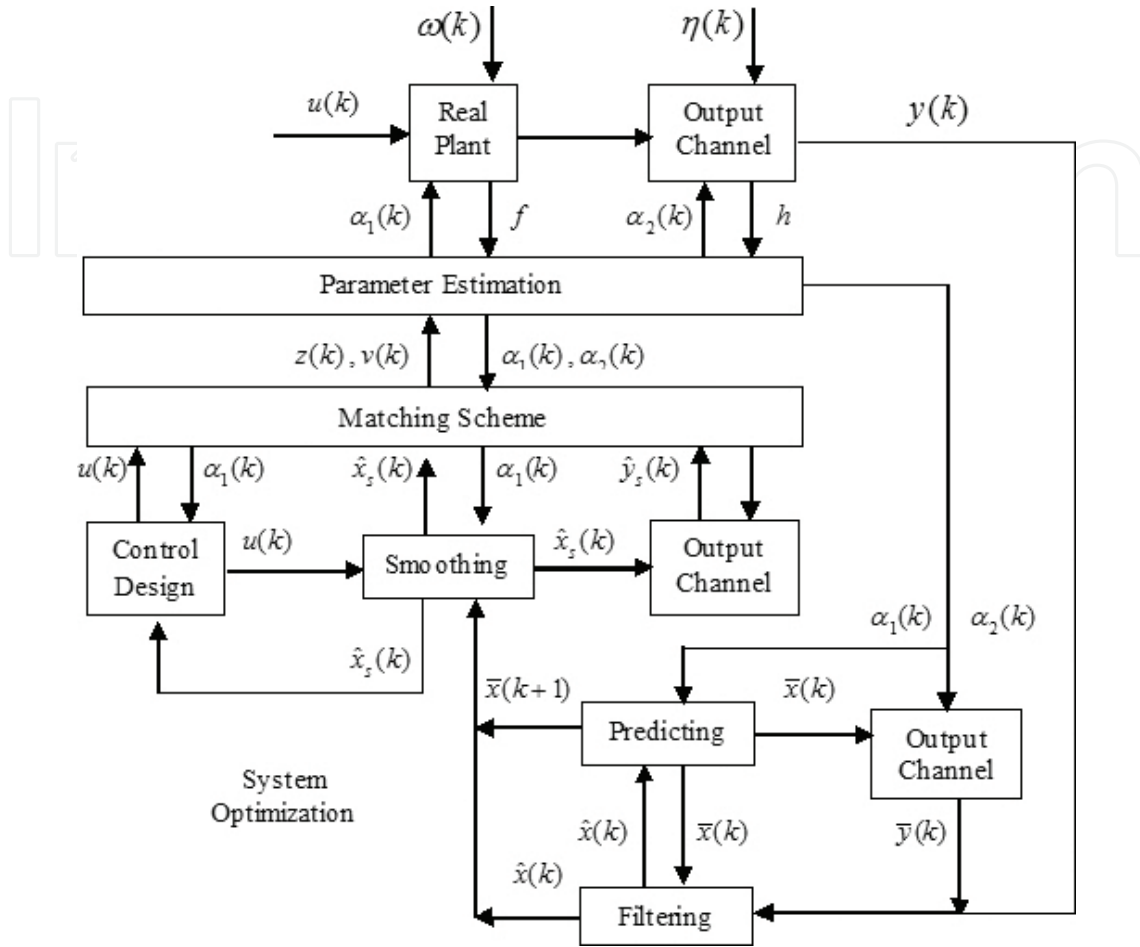

Figure 1. Block diagram of the approach proposed.

3.1. Optimality conditions

Define the Hamiltonian function for Problem (E) as follows:

$$
\begin{aligned}
H_e(k) = &\tfrac{1}{2}(\hat{x}_s(k)^\mathrm{T} Q\hat{x}_s(k) + u(k)^\mathrm{T} Ru(k)) + \gamma(k) \\
&+ \tfrac{1}{2}r_1 \| v(k) - u(k) \|^2 + \tfrac{1}{2}r_2 \| z(k) - \hat{x}_s(k) \|^2 \\
&- \lambda(k)^\mathrm{T} u(k) - \beta(k)^\mathrm{T} \hat{x}_s(k) \\
&+ q(k)^\mathrm{T} (C\hat{x}_s(k) - \hat{y}_s(k)) \\
&+ p(k+1)^\mathrm{T} (\hat{x}_s(k) - \hat{x}(k) - K_s(k)(\hat{x}_s(k+1) - \bar{x}(k+1))).
\end{aligned}
\tag{8}
$$

Then, the augmented cost function becomes

$$
\begin{aligned}
J_e'(k) = {} & \tfrac{1}{2}\hat{x}_s(N)^{\mathrm{T}}S(N)\hat{x}_s(N) + \gamma(N) + \Gamma^{\mathrm{T}}(\hat{x}_s(N) - z(N)) \\
& + \xi(N)(\varphi(z(N),N) - \tfrac{1}{2}z(N)^{\mathrm{T}}S(N)z(N) - \gamma(N)) \\
& + \sum_{k=0}^{N-1} H_e(k) + \lambda(k)^{\mathrm{T}}v(k) + \beta(k)^{\mathrm{T}}z(k) \\
& + \xi(k)(L(z(k),v(k),k) - \tfrac{1}{2}(z(k)^{\mathrm{T}}Qz(k) + v(k)^{\mathrm{T}}Rv(k)) - \gamma(k)) \\
& + \mu(k)^{\mathrm{T}}(f(z(k),v(k),k) - Az(k) - Bv(k) - \alpha_1(k)) \\
& + \pi(k)^{\mathrm{T}}(h(z(k),k) - Cz(k) - \alpha_2(k))
\end{aligned}
\tag{9}
$$

where $p(k), q(k), \mu(k), \xi(k), \pi(k), \Gamma, \beta(k)$, and $\lambda(k)$ are the proper multipliers to be judged the value later.

The following necessary conditions for optimality are resulted when applying the calculus of variation [2, 4, 17] to the augmented cost function given in Eq. (9):

(a) Stationary condition:

$$
Ru(k) + B^{\mathrm{T}}K_s(k)p(k+1) - \lambda(k) - r_1(v(k) - u(k)) = 0.
\tag{10a}
$$

(b) Smoothed costate equation:

$$
p(k) = Q\hat{x}_s(k) + p(k+1) - \beta(k) - r_2(z(k) - \hat{x}_s(k)).
\tag{10b}
$$

(c) Smoothed state equation:

$$
\hat{x}_s(k) = \hat{x}(k) + K_s(k)(\hat{x}_s(k+1) - \bar{x}(k+1))
\tag{10c}
$$

with the boundary conditions $\hat{x}_s(N) = \bar{x}(N)$ and $p(N) = \Gamma$.

(d) Adjustable parameter equations:

$$
\varphi(z(N),N) = \tfrac{1}{2}z(N)^{\mathrm{T}}S(N)z(N) + \gamma(N)
\tag{11a}
$$

$$
L(z(k),v(k),k) = \tfrac{1}{2}(z(k)^{\mathrm{T}}Qz(k) + v(k)^{\mathrm{T}}Rv(k)) + \gamma(k)
\tag{11b}
$$

$$
f(z(k),v(k),k) = Az(k) + Bv(k) + \alpha_1(k)
\tag{11c}
$$

$$h(z(k),k) = Cz(k) + \alpha_2(k).$$ (11d)

(e) Multiplier equations:

$$\Gamma - \nabla_{z(k)}\varphi + S(N)z(N) = 0$$ (12a)

$$\lambda(k) + (\nabla_{v(k)}L - Rv(k)) + \left(\frac{\partial f}{\partial v(k)} - B\right)^{\mathrm{T}} \hat{p}(k+1) = 0$$ (12b)

$$\beta(k) + (\nabla_{z(k)}L - Qz(k)) + \left(\frac{\partial f}{\partial z(k)} - A\right)^{\mathrm{T}} \hat{p}(k+1) = 0$$ (12c)

with $\xi(k) = 1, \mu(k) = \hat{p}(k+1)$ and $\pi(k) = q(k) = 0$.

(f) Separable variables:

$$v(k) = u(k), z(k) = \hat{x}_s(k), \hat{p}(k) = p(k).$$ (13)

In view of these necessary optimality conditions, the conditions (10a), (10b), and (10c) define the modified model-based optimal control problem, the conditions (11a), (11b), (11c), and (11d) define the parameter estimation problem and the conditions (12a), (12b), and (12c) are used to compute the multipliers. They are further discussed as follows.

3.2. Modified model-based optimal control problem

The modified model-based optimal control problem, which is referred to as Problem (MM), is given below:

$$
\begin{aligned}
\min_{u(k)} J_{mm}(u) = {}& \tfrac{1}{2}\hat{x}_s(N)^{\mathrm{T}} S(N)\hat{x}_s(N) + \gamma(N) + \Gamma^{\mathrm{T}}\hat{x}_s(N) \\
& + \sum_{k=0}^{N-1} \tfrac{1}{2}(\hat{x}_s(k)^{\mathrm{T}} Q\hat{x}_s(k) + u(k)^{\mathrm{T}} Ru(k)) + \gamma(k) \\
& + \tfrac{1}{2}r_1 \| v(k) - u(k) \|^2 + \tfrac{1}{2}r_2 \| z(k) - \hat{x}_s(k) \|^2 \\
& - \lambda(k)^{\mathrm{T}} u(k) - \beta(k)^{\mathrm{T}} \hat{x}_s(k)
\end{aligned}
$$ (14)

subject to

$$\hat{x}_s(k) = \hat{x}(k) + K_s(k)(\hat{x}_s(k+1) - \bar{x}(k+1))$$

$$\hat{y}_s(k) = C\hat{x}_s(k).$$

From the outcome of Problem (E) and Problem (MM), the theorem of the smoothed optimal control law which is applied to solve Problem (MM) is described.

Theorem 3.1: Suppose the expanded optimal control law for Problem (E) exists. Then, this control law is the smoothed feedback control law for Problem (MM) given by

$$u(k) = -K(k)\hat{x}_s(k) + u_{ff}(k) \tag{15}$$

where

$$
\begin{aligned}
u_{ff}(k) = &-(R_a + B^{\mathrm{T}}K_s(k)S(k+1)B)^{-1}(B^{\mathrm{T}}K_s(k)s(k+1) - \lambda_a(k) \\
&+ B^{\mathrm{T}}K_s(k)S(k+1)((A - K_s(k)^{-1})\hat{x}(k) + \alpha_1(k)))
\end{aligned} \tag{16a}
$$

$$K(k) = (R_a + B^{\mathrm{T}}K_s(k)S(k+1)B)^{-1}B^{\mathrm{T}}K_s(k)S(k+1)K_s(k)^{-1} \tag{16b}$$

$$S(k) = Q_a + S(k+1)(K_s(k)^{-1} - BK(k)) \tag{16c}$$

$$s(k) = S(k+1)((A - K_s(k)^{-1})\hat{x}(k) + Bu_{ff}(k) + \alpha_1(k)) + s(k+1) - \beta_a(k) \tag{16d}$$

with the boundary conditions $S(N)$ given and $s(N) = 0$, and

$$R_a = R + r_1 I_m; \quad Q_a = Q + r_2 I_n;$$

$$\lambda_a(k) = \lambda(k) + r_1 v(k); \quad \beta_a(k) = \beta(k) + r_2 z(k).$$

Proof: From the necessary optimality condition (10a), we have

$$R_a u(k) = -B^{\mathrm{T}}K_s(k)p(k+1) + \lambda_a(k). \tag{17}$$

Applying sweep method [2, 4],

$$p(k) = S(k)\hat{x}_s(k) + s(k) \tag{18}$$

we substitute Eq. (18) for $k = k + 1$ into Eq. (17), which yields

$$R_a u(k) = -B^{\mathrm{T}} K_s(k) S(k+1) x_s(k+1) - B^{\mathrm{T}} K_s(k) s(k+1) + \lambda_a(k). \tag{19}$$

Rewrite the smoothed state equation from Eq. (10c),

$$\hat{x}_s(k+1) = \bar{x}(k+1) + (K_s(k))^{-1} (\hat{x}_s(k) - \hat{x}(k)). \tag{20}$$

Then, substitute Eq. (20) into Eq. (19). After some algebraic manipulations, the smoothed control law (15) is obtained, where Eqs. (16a) and (16b) are satisfied.

From the smoothed costate equation (10b), we substitute Eq. (18) for $k = k + 1$ to give

$$p(k) = Q_a \hat{x}_s(k) + S(k+1)\hat{x}_s(k+1) + s(k+1) - \beta_a(k) \tag{21}$$

Consider Eq. (20) in Eq. (21), we obtain

$$p(k) = Q_a \hat{x}_s(k) + S(k+1)(\bar{x}(k+1) + (K_s(k))^{-1}(\hat{x}_s(k) - \hat{x}(k))) + s(k+1) - \beta_a(k). \tag{22}$$

By doing some algebraic manipulations, it is found that Eqs. (16c) and (16d) are satisfied after comparing to Eq. (18). This completes the proof.

From Eqs. (4a), (10c), and (15), the smoothed state equation becomes

$$\begin{aligned} \hat{x}_s(k) = (I_n - K_s(k)BK(k))^{-1}((I_n - K_s(k)A)\hat{x}(k) \\ + K_s(k)(\hat{x}_s(k+1) - Bu_{f\!f}(k) - \alpha_1(k))) \end{aligned} \tag{23}$$

and the smoothed output is measured from

$$\hat{y}_s(k) = C\hat{x}_s(k) \tag{24}$$

with the boundary condition $\hat{x}_s(N) = \bar{x}(k)$.

3.3. Parameter estimation

After solving Problem (MM), the defined separable variables given in Eq. (13) are used for the further computations. Particularly, in the parameter estimation problem, the differences between the real plant and the model used are taken into account in which the matching schemes are established. In view of this, the adjusted parameters, which are resulted from parameter estimation problem defined by Eq. (11), are calculated from

$$\alpha_1(k) = f(z(k), v(k), k) - Az(k) - Bv(k) \tag{25a}$$

$$\alpha_2(k) = h(z(k), k) - Cz(k) \tag{25b}$$

$$\gamma(N) = \varphi(z(N), N) - \tfrac{1}{2} z(N)^{\mathrm{T}} S(N) z(N) \tag{25c}$$

$$\gamma(k) = L(z(k), v(k), k) - \tfrac{1}{2} (z(k)^{\mathrm{T}} Qz(k) + v(k)^{\mathrm{T}} Rv(k)) \tag{25d}$$

3.4. Computation of multipliers

The multipliers, which are related to the Jacobian matrix of the functions f and L with respect to $v(k)$ and $z(k)$, are computed from

$$\Gamma = \nabla_{z(k)} \varphi - S(N) z(N) \tag{26a}$$

$$\lambda(k) = -(\nabla_{v(k)} L - Rv(k)) - \left(\frac{\partial f}{\partial v(k)} - B \right)^{\mathrm{T}} \hat{p}(k+1) \tag{26b}$$

$$\beta(k) = -(\nabla_{z(k)} L - Qz(k)) - \left(\frac{\partial f}{\partial z(k)} - A \right)^{\mathrm{T}} \hat{p}(k+1) \tag{26c}$$

3.5. Iterative algorithm

From the previous sections, the derivation of equations and the formulation of the resulting algorithm are clearly discussed. Following from these discussions, a summary on this iterative algorithm is delivered as follows:

Data $Q, R, S(N), A, B, C, G, Q_\omega, R_\eta, M_0, \bar{x}_0, N, r_1, r_2, k_v, k_z, k_p, f, L, h, \varphi$. Note that A and B may be chosen through the linearization of f, and C is obtained from the linearization of h.

Step 0: Compute a nominal solution. Assume $\alpha_1(k) = 0, k = 0, 1, ..., N-1, \alpha_2(k) = 0, k = 0, 1, ..., N$, and $r_1 = r_2 = 0$. Calculate $K_f(k)$ and $K_s(k)$ from Eqs. (5a) and (5b), $P(k), M_x(k), P_s(k)$ and $M_y(k)$ from Eqs. (6a), (6b), (6c), and (6d) for the state estimation, and solve Problem (M) defined by Eq. (3) to obtain $u(k)^0, k = 0, 1, ..., N-1$, and $\hat{x}_s(k)^0, \hat{y}_s(k)^0, p(k)^0, k = 0, 1, ..., N$. Then, with $\alpha_1(k) = 0, k = 0, 1, ..., N-1, \alpha_2(k) = 0, k = 0, 1, ..., N$, and r_1, r_2 from data, calculate $K(k)$ and $S(k)$, respectively, from Eqs. (16b) and (16c). Set $i = 0, z(k)^0 = \hat{x}_s(k)^0, v(k)^0 = u(k)^0$ and $\hat{p}(k)^0 = p(k)^0$.

Step 1: Calculate the adjustable parameters $\alpha_1(k)^i$, $k = 0, 1, ..., N-1, \alpha_2(k)^i, k = 0, 1, ..., N, \gamma(k)^i, k = 0, 1, ..., N$, from Eq. (25). This is called the *parameter estimation* step.

Step 2: Compute the modifiers $\Gamma^i, \lambda(k)^i$ and $\beta(k)^i, k = 0, 1, ..., N-1$, from Eq. (26). This requires the partial derivatives of f, h and L with respect to $v(k)^i$ and $z(k)^i$.

Step 3: With the determined $\alpha_1(k)^i, \alpha_2(k)^i, \gamma(k)^i, \Gamma^i, \lambda(k)^i, \beta(k)^i, v(k)^i$, and $z(k)^i$, solve Problem (MM) defined by Eq. (14) using the result in Theorem 3.1. This is called the *system optimization* step.

a. Obtain $s(k)^i, k = 0, 1, ..., N$ by solving Eq. (16d) backward, and obtain $u_{ff}(k)^i, k = 0, 1, ..., N-1$ by solving Eq. (16a), either backward or forward.

b. Calculate the new control $u(k)^i, k = 0, 1, ..., N-1$ using Eq. (15).

c. Calculate the new state $\hat{x}_s(k)^i, k = 0, 1, ..., N$, using Eq. (23).

d. Calculate the new costate $p(k)^i, k = 0, 1, ..., N$, using Eq. (18).

e. Calculate the new output $\hat{y}_s(k)^i, k = 0, 1, ..., N$, using Eq. (24).

Step 4: Update the optimal smoothing solution of Problem (P) and test the convergence of the algorithm. For regulating convergence, a mechanism, which is a simple relaxation method, shall be provided and given by:

$$z(k)^{i+1} = z(k)^i + k_z(\hat{x}_s(k)^i - z(k)^i) \tag{27a}$$

$$v(k)^{i+1} = v(k)^i + k_v(u(k)^i - v(k)^i) \tag{27b}$$

$$\hat{p}(k)^{i+1} = \hat{p}(k)^i + k_p(p(k)^i - \hat{p}(k)^i) \tag{27c}$$

where k_v, k_z, k_p, range in the interval of $(0, 1]$, are scalar gains. If $z(k)^{i+1} = z(k)^i, k = 0, 1, ..., N$, and $v(k)^{i+1} = v(k)^i, k = 0, 1, ..., N - 1$, within a given tolerance, stop; else repeat from Step 1 by setting $i = i + 1$.

Remarks:

a. The off-line computation, which is mentioned in Step 0, is done for the state estimator design, where $K_f(k), K_s(k), k = 0, 1, ..., N - 1, M_x(k), M_y(k), k = 0, 1, ..., N, P(k), P_s(k),$ $k = 0, 1, ..., N - 1$ are computed, and for the control law design, where $K(k), k = 0, 1, ..., N - 1, S(k), k = 0, 1, ..., N$ are calculated. In fact, these parameters are used for solving Problem (M) in Step 0 and for solving Problem (MM) in Step 3, respectively.

b. The variables $\gamma(k)^i, \alpha_1(k)^i, \alpha_2(k)^i, \Gamma^i, \lambda(k)^i, \beta(k)^i$, and $s(k)^i$ are initially zero in Step 0. Their computed values, where $\gamma(k)^i, \alpha_1(k)^i, \alpha_2(k)^i$ in Step 1, $\Gamma^i, \lambda(k)^i, \beta(k)^i$ in Step 2, and $s(k)^i$ in Step 3, would be changed from iteration to iteration.

c. The driving input $u_{ff}(k)$ in Eq. (16a) corrects the differences between the real plant and the model used, and it also drives the controller given in Eq. (15).

d. The state estimation without the control is done forward using the Kalman filtering, and then it is followed by the fixed-interval smoothing backward in order to design the feedback control law.

e. Problem (P) is not necessary to have a cost function in quadratic criterion or to be a linear problem.

f. The equations $z(k)^{i+1} = z(k)^i$ and $v(k)^{i+1} = v(k)^i$ can be definitely required to satisfy for the converged state estimate sequence and the converged optimal control sequence. On this point of view, the following averaged 2-norms are computed and, then, they are compared with a given tolerance to verify the convergence of $v(k)$ and $z(k)$:

$$\| v^{i+1} - v^i \|_2 = \left(\frac{1}{N-1} \sum_{k=0}^{N-1} \| v(k)^{i+1} - v(k)^i \| \right)^{1/2} \tag{28a}$$

$$\| z^{i+1} - z^i \|_2 = \left(\frac{1}{N} \sum_{k=0}^{N} \| z(k)^{i+1} - z(k)^i \| \right)^{1/2} \tag{28b}$$

g. The relaxation scalars (k_v, k_z, k_p) are the step-sizes in regulating the convergence mechanism. These scalars could be normally chosen as a certain value in the range of $(0, 1]$, but this choice may not provide the optimal number of iterations. Hence, it is important to note that the optimal choice of these scalars $k_v, k_z, k_p \in (0, 1]$ would be problem dependent.

As a rule of this case, the algorithm (from Step 1 to Step 4) is required to run few times. Initially, for first run of the algorithm (from Step 1 to Step 4), these scalars are set at $k_v = k_z = k_p = 1$, and then, with different values chosen from 0.1 to 0.9, the algorithm is run again. The value with the optimal number of iterations can be determined after that. Applying the parameters r_1 and r_2 is to enhance the convexity such that the convergence of the algorithm can be improved.

4. Convergence analysis

In this section, the convergence of the algorithm is discussed. The following assumptions are needed:

The derivatives of f, L and h exist.

The solution (u^*, x^*, y^*) is the optimal solution to Problem (P). That is, the optimal smoothing solution.

The convergence result is presented in Theorem 4.1, while the accuracy of the smoothed state in term of state error covariance is proven in Corollary 4.1.

Theorem 4.1: The converged solution of Problem (M) is the correct optimal smoothing solution of Problem (P).

Proof: Consider the real plant and the output measurement of Problem (P) with the exact optimal smoothing solution (u^*, x^*, y^*) as given below:

$$x^*(k+1) = f(x^*(k), u^*(k), k) \text{ and } y^*(k) = h(x^*(k), k) \tag{29}$$

In Problem (M), the model used consists of

$$\hat{x}^c(k) = \bar{x}^c(k) + K_f(k)(y(k) - \bar{y}^c(k)) \tag{30a}$$

$$\bar{x}^c(k+1) = A\hat{x}^c(k) + Bu^c(k) + \alpha_1(k) \tag{30b}$$

$$\bar{y}^c(k) = C\bar{x}^c(k) + \alpha_2(k) \tag{30c}$$

$$\hat{x}_s^c(k) = \hat{x}^c(k) + K_s(k)(\hat{x}_s^c(k+1) - \bar{x}^c(k+1)) \tag{30d}$$

$$\hat{y}_s^c(k) = C\hat{x}_s^c(k) \tag{30e}$$

where $u^c(k), \hat{x}_s^c(k), \hat{x}^c(k), \bar{x}^c(k), \hat{y}_s^c(k)$, and $\bar{y}^c(k)$ are, respectively, the converged sequences for control law, smoothed state estimate, filtered state estimate, expected state estimate, smoothed output, and expected output. Here, $y(k)$ is the output measured from the real plant.

Applying the adjusted parameters $\alpha_1(k)$ and $\alpha_2(k)$, which are given by

$$\alpha_1(k) = f(z(k), v(k), k) - Az(k) - Bv(k) \text{and}$$

$$\alpha_2(k) = h(z(k), k) - Cz(k),$$

into the model used given by Eq. (30b) and (30c), the differences between the real plant and the model used can be measured at each iteration. Moreover, at the end of iteration, from Eqs. (29) and (30a) – (30e) yields

$$\hat{x}_s^c(k+1) = f(z(k), v(k), k) \text{and} \hat{y}_s^c(k) = h(z(k), k)$$

which $v(k) = u^c(k)$ and $z(k) = \hat{x}_s^c(k) = \hat{x}^c(k)$ are satisfied. Hence, this implies that

$$u^c(k) = u^*(k), \hat{x}_s^c(k) = x^*(k), \hat{y}_s^c(k) = y^*(k)$$

This completes the proof.

Corollary 4.1: The smoothed state error covariance is the smallest among the values of state error covariance.

Proof: From Eq. (6), it is clear that the filtered state error covariance $P(k)$ is less than the predicted state error covariance $M_x(k)$. That is, $P(k) < M_x(k)$. Now, to prove $P_s(k) < P(k)$,, we shall show that $P_s(k+1) < M_x(k+1)$. Consider the boundary condition $P_s(N) = M_x(N)$ and taking $k = N - 1$, we have

$$P_s(N-1) = P(N-1) < M_x(N-1).$$

For $k = N - 2$, it shows that

$$P_s(N-2) < P(N-2) < M_x(N-2).$$

This statement can be deduced that

$$P_s(k+1) - M_x(k+1) < 0 \text{ for } k = k+1.$$

Thus, we conclude that

$$P_s(k) < P(k) < M_x(k), \quad k = 0,1,...,N-2,$$

which shows the accuracy of the smoothed state estimate. This completes the proof.

5. Illustrative example

Consider a continuous stirred-tank reactor problem [19], which consists of the state difference equations

$$x_1(k+1) = x_1(k) - 0.02(x_1(k) + 0.25) + 0.01(x_2(k) + 0.5) \exp\left[\frac{25x_1(k)}{x_1(k) + 2}\right]$$
$$- 0.01(x_1(k) + 0.25)u(k) + \omega_1(k)$$

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

for $k = 0, ..., 77$, and the output measurement $y(k) = x_1(k) + \eta(k)$. The initial state $x(0) = x_0$ is a random vector with mean and covariance given, respectively, by $\bar{x}_1(0) = 0.05$, $\bar{x}_2(0) = 0$, and $M_0 = 10^{-2}I_2$.

Here, $\omega(k) = [\omega_1(k) \ \omega_2(k)]^T$ and $\eta(k)$ are Gaussian white noise sequences with their respective covariance given by $Q_\omega = 10^{-3}I_2$ and $R_\eta = 10^{-3}$. The expected cost function

$$J_0(u) = 0.5 \sum_{k=0}^{N-1} E[(x_1(k))^2 + (x_2(k))^2 + 0.1(u(k))^2]$$

is to be minimized over the state difference equations and the output measurement.

This problem is referred to as Problem (P).

To obtain the optimal smoothing solution of Problem (P), we simplify the plant dynamics of Problem (P) and refer it as Problem (M), given by

$$\min_{u(k)} J_m(u) = \frac{1}{2} \sum_{k=0}^{N-1} [(\hat{x}_s(k))^2 + 0.1(u(k))^2 + 2\gamma(k)]$$

subject to

$$\hat{x}_s(k) = \hat{x}(k) + K_s(k)(\hat{x}_s(k+1) - \overline{x}(k+1))$$

$$\hat{y}_s(k) = C\hat{x}_s(k)$$

with

$$\hat{x}(k) = \overline{x}(k) + K_f(k)(y(k) - \overline{y}(k))$$

$$\begin{bmatrix} \overline{x}_1(k+1) \\ \overline{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 1.0895 & 0.0184 \\ -0.1095 & 0.9716 \end{bmatrix} \begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \end{bmatrix} + \begin{bmatrix} -0.003 \\ 0.000 \end{bmatrix} u(k) + \begin{bmatrix} \alpha_{11}(k) \\ \alpha_{12}(k) \end{bmatrix}$$

$$\overline{y}(k) = \overline{x}_1(k) + \alpha_2(k)$$

with the initial condition $\overline{x}(0) = \overline{x}_0$ and the boundary value $\hat{x}_s(N) = \overline{x}(N)$. Here, $\gamma(k)$, $\alpha_2(k)$ and $\alpha_1(k) = [\alpha_{11}(k) \ \alpha_{12}(k)]^T$ are the adjusted parameters.

Model	Iteration number	Elapsed time	Initial cost	Final cost	Output residual
Filtering	6	0.782772	3.7910	0.021271	0.034731
Smoothing	8	1.026919	3.5095	0.000734	0.018294

Table 1. Iteration result.

The iteration results, both for filtering and smoothing models, are shown in **Table 1**. The final cost of the smoothing model is the least compared to the final cost of the filtering model. When the trace matrix terms are considered in the cost function, the total final cost of the smoothing model is 0.019188 unit, while the total final cost of the filtering model is 0.039725 unit. The value of the trace matrix terms is 0.0185 unit. It is noticed that the output residual could be dropped to almost 52% from the filtering output residual by using the approach proposed in

this chapter. This statement is valid since the output residual of smoothing model is least than the output residual of filtering model.
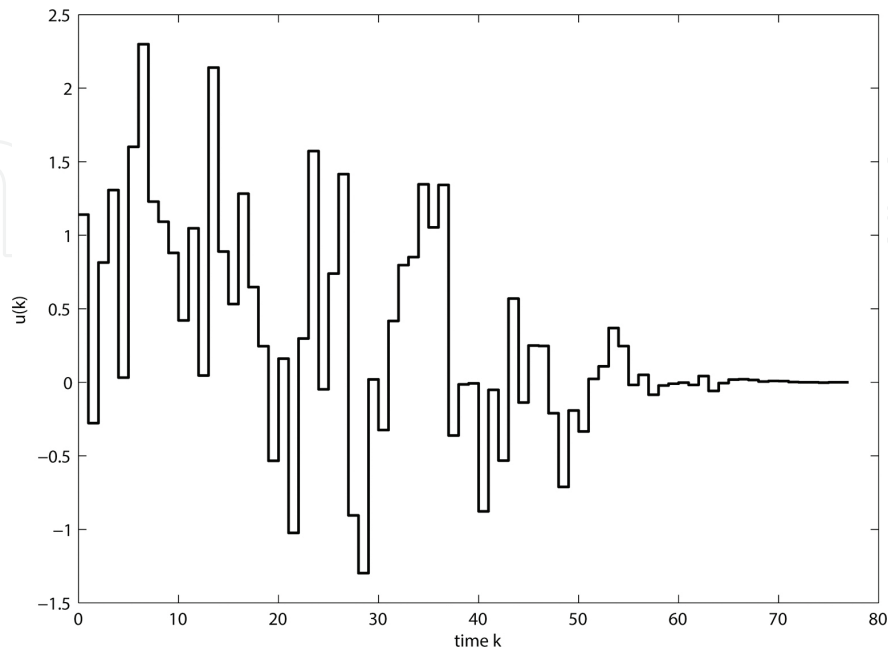

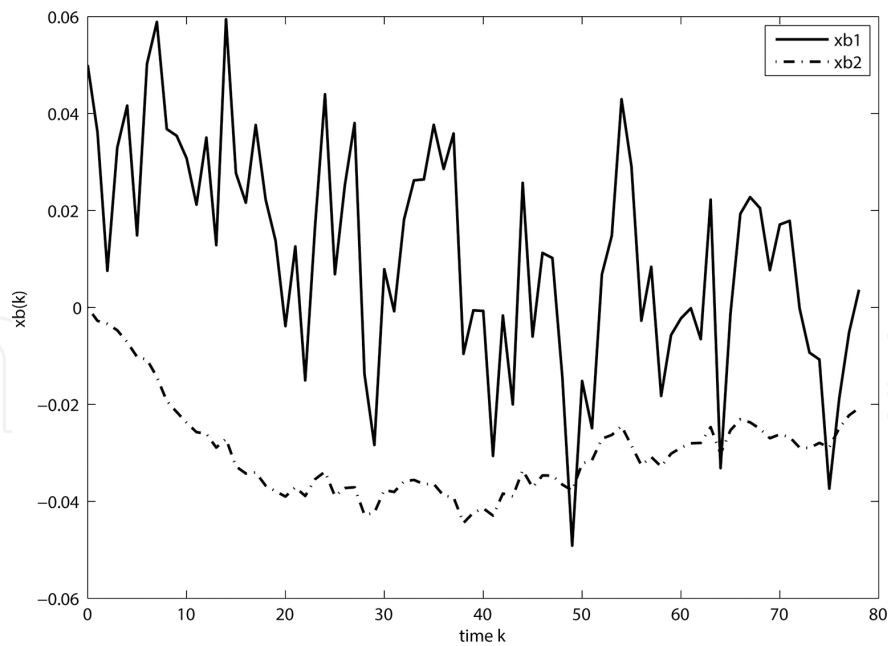**Figure 2.** Filtering trajectory for final control.

Figure 3. Filtering trajectory for final state.

To identify the accuracy of the resulting algorithm, the norms of the differences between the real plant and the model used at the end of iteration, which are 0.0128 unit for filtering model

and 0.0099 unit for smoothing model, are calculated. These values show that the smoothing model can approximate closely to the correct optimal solution of the original optimal control problem rather than the filtering model. Hence, the accuracy of the smoothing model is proven.
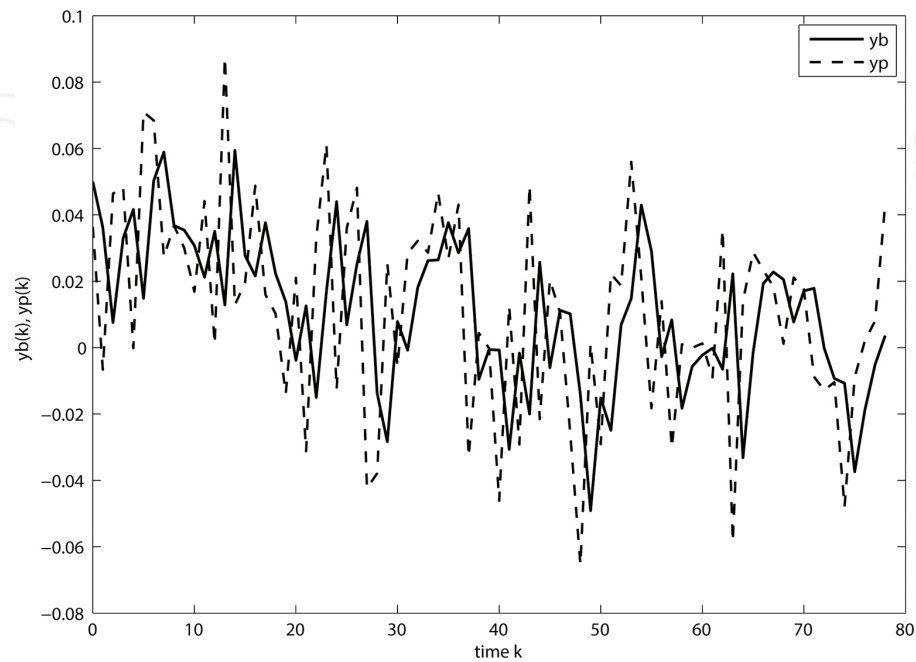
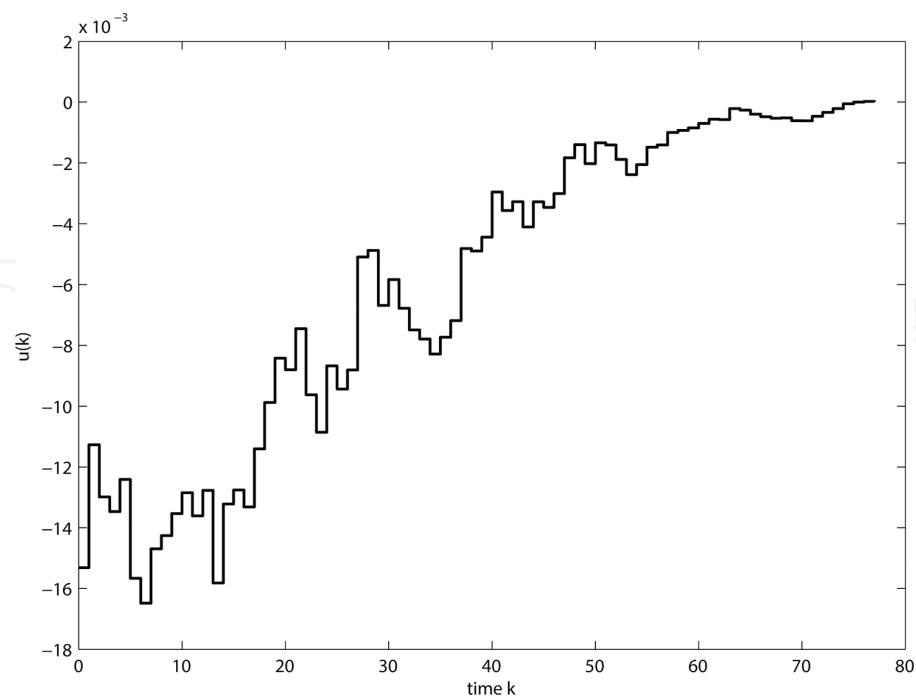**Figure 4.** Filtering trajectory for final output and real output.

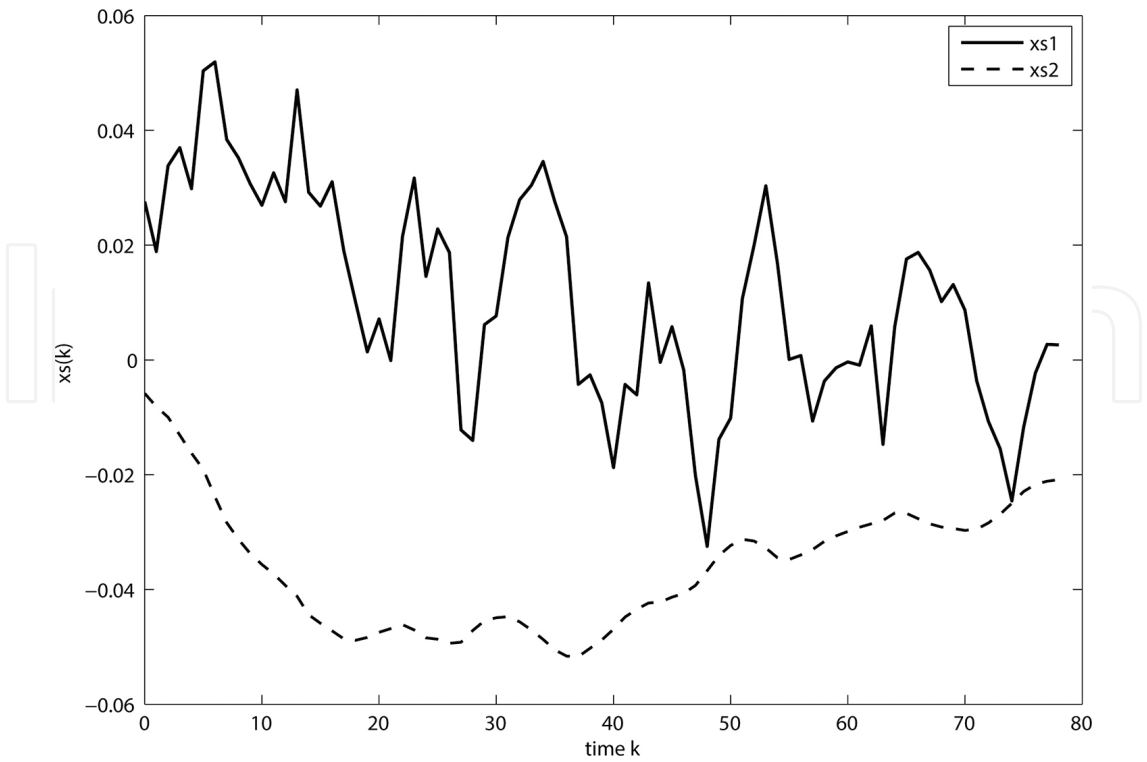**Figure 5.** Smoothing trajectory for final control.

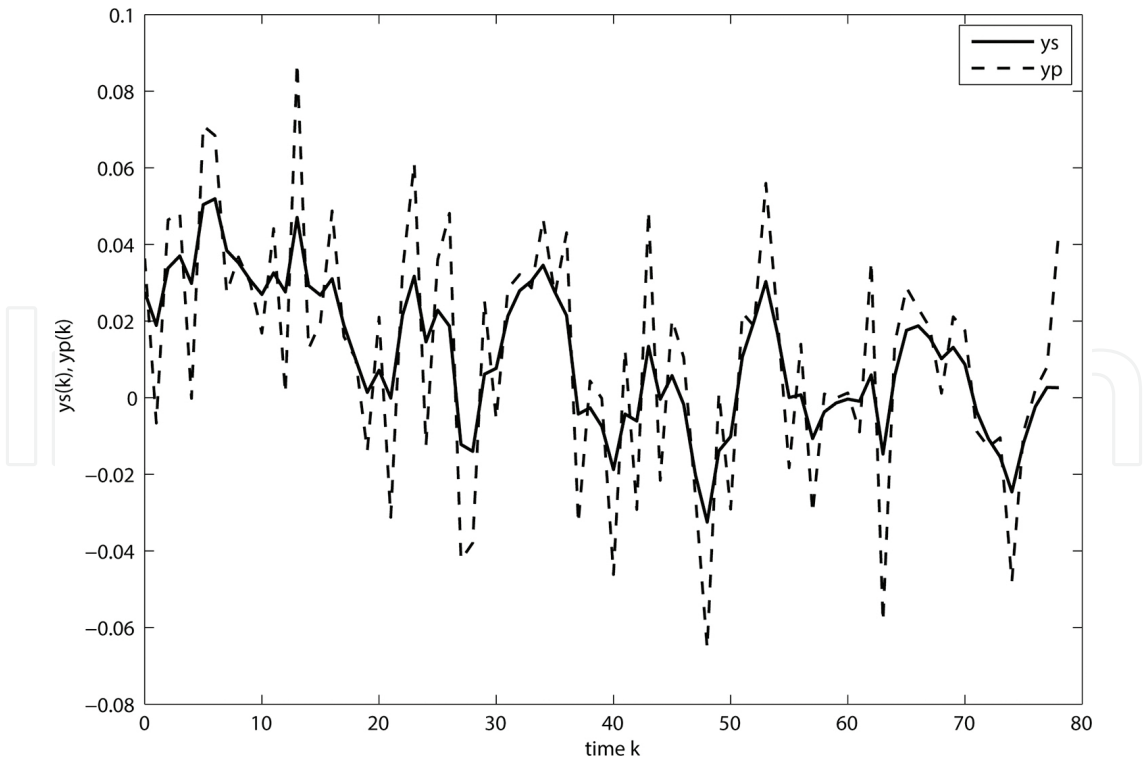**Figure 6.** Smoothing trajectory for final state.

Figure 7. Smoothing trajectory for final output and real output.

The trajectories of final control, final state and final output for filtering, and smoothing models are shown in **Figures 2–7**. With the smallest output residual, the output, which is associated with the smoothed state estimate, is definitely applicable to measure the real output trajectory.

6. Concluding remarks

A fixed-interval smoothing scheme was modified in this chapter for solving the discrete-time nonlinear stochastic optimal control problem. The state estimation procedure, which is using the Kalman filtering theory and is followed by the fixed-interval smoothing, is applied to estimate the system dynamics. Then, the smoothed state estimate is used in designing the feedback optimal control law. By employing this smoothed state estimate, system optimization and parameter estimation are integrated. During the computation procedure, the differences between the real plant and the model used are calculated iteratively. On the other hand, the output measured from the real plant is fed back into the model used, in turn, updates the iterative solution. Once the convergence is achieved, the iterative solution approaches to the correct optimal solution of the original optimal control problem, in spite of model-reality differences. The illustrative example on the optimal control of the continuous stirred-tank reactor problem was studied. The results obtained demonstrated the applicable of the approach proposed, and the efficiency of the approach proposed is highly presented.

Author details

Sie Long Kek[1*], Kok Lay Teo[2] and Mohd Ismail Abd Aziz[3]

*Address all correspondence to: slkek@uthm.edu.my

1 Center for Research in Computational Mathematics, Universiti Tun Hussein Onn Malaysia, Parit Raja, Malaysia

2 Department of Mathematics and Statistics, Curtin University of Technology, Perth, WA, Australia

3 Department of Mathematical Sciences, Universiti Teknologi Malaysia, UTM, Skudai, Malaysia

References

[1] Kalman R. E. A new approach to linear filtering and prediction problems. Journal of Basic Engineering. 1960; 82(1):35–45.

[2] Bryson A. E. and Ho Y. C. Applied Optimal Control. Washington: Hemisphere; 1975.

[3] Bertsekas D. P. Dynamic Programming and Optimal Control (Vol. 1, No. 2). Belmont: Athena Scientific; 1995.

[4] Lewis F. L. and Syrmos V. L. Optimal Control. 2nd ed. USA: John Wiley & Sons; 1995.

[5] Feng Z.G. and Teo K. L. Optimal feedback control for stochastic impulsive linear systems subject to Poisson processes. In: Optimization and Optimal Control. New York: Springer; 2010. p. 241–258.

[6] Misiran M., Wu C., Lu Z. and Teo K.L. Optimal filtering of linear system driven by fractional Brownian motion. Dynamic Systems and Applications. 2010; 19(3):495–514.

[7] Loxton R., Lin Q. and Teo K. L. A stochastic fleet composition problem. Computers & Operations Research. 2012; 39(12):3177–3183. DOI: 10.1016/j.cor.2012.04.004.

[8] Liu C. M., Feng Z. G. and Teo K. L. On a class of stochastic impulsive optimal parameter selection problems. International Journal of Innovation, Computer and Information Control. 2009; 5:1043–1054.

[9] Yin Y., Shi P., Liu F. and Teo K. L. Robust $L_2 - L_\infty$ filtering for a class of dynamical systems with nonhomogeneous Markov jump process. International Journal of Systems Science. 2015; 46(4):599–608. DOI: 10.1080/00207721.2013.792976.

[10] Moura S. J., Fathy H. K., Callaway D. S. and Stein J. L. A stochastic optimal control approach for power management in plug-in hybrid electric vehicles. IEEE Transactions on Control Systems Technology. 2011; 19(3):545–555. DOI: 10.1109/TCST.2010.2043736.

[11] Wiegerinck W. Broek B. V. D. and Kappen H. Stochastic optimal control in continuous space-time multi-agent systems. Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence (UAI'06), Arlington, Virginia. 2006; 528–535.

[12] Zhu Y. Uncertain optimal control with application to portfolio selection model. International Journal of Cybernetics and Systems. 2010; 41(7):535–547. DOI: 10.1080/01969722.2010.511552.

[13] Hać A. Suspension optimization of a 2-DOF vehicle model using a stochastic optimal control technique. Journal of Sound and Vibration. 1985; 100(3):343–357. DOI: 10.1016/0022-460X(85)90291-3.

[14] Todorov E. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. Neural Computation. 2005; 17(5):1084–1108.

[15] Sethi S. P. Deterministic and stochastic optimization of a dynamic advertising model. Optimal Control Applications and Methods. 1983; 4(2):179–184. DOI: 10.1002/oca. 4660040207.

[16] Kek S. L., Teo K. L. and Mohd Ismail A. A. An integrated optimal control algorithm for discrete-time nonlinear stochastic system. International Journal of Control. 2010; 83:2536–2545. DOI: 10.1080/00207179.2010.531766.

[17] Kek S. L., Teo K. L. and Mohd Ismail A. A. Filtering solution of nonlinear stochastic optimal control problem in discrete-time with model-reality differences. Numerical Algebra, Control and Optimization. 2012; 2(1):207–222. DOI: 10.3934/naco.2012.2.207.

[18] Kek S. L., Mohd Ismail A. A., Teo K. L. and Ahmad R. An iterative algorithm based on model-reality differences for discrete-time nonlinear stochastic optimal control problems. Numerical Algebra, Control and Optimization. 2013; 3(1):109–125. DOI: 10.3934/naco.2013.3.109.

[19] Kirk D. E. Optimal Control Theory: An Introduction. Mineola, New York: Dover Publications; 2004.