

## A GRADIENT ALGORITHM FOR OPTIMAL CONTROL PROBLEMS WITH MODEL-REALITY DIFFERENCES

SIE LONG KEK

Department of Mathematics and Statistics  
Universiti Tun Hussein Onn Malaysia  
86400 Parit Raja, Malaysia

MOHD ISMAIL ABD AZIZ

Department of Mathematical Sciences  
Universiti Teknologi Malaysia  
81310 UTM, Skudai, Malaysia

KOK LAY TEO

Department of Mathematics and Statistics  
Curtin University of Technology  
Perth, W.A. 6845, Australia

**ABSTRACT.** In this paper, we propose a computational approach to solve a model-based optimal control problem. Our aim is to obtain the optimal solution of the nonlinear optimal control problem. Since the structures of both problems are different, only solving the model-based optimal control problem will not give the optimal solution of the nonlinear optimal control problem. In our approach, the adjusted parameters are added into the model used so as the differences between the real plant and the model can be measured. On this basis, an expanded optimal control problem is introduced, where system optimization and parameter estimation are integrated interactively. The Hamiltonian function, which adjoins the cost function, the state equation and the additional constraints, is defined. By applying the calculus of variation, a set of the necessary optimality conditions, which defines modified model-based optimal control problem, parameter estimation problem and computation of modifiers, is then derived. To obtain the optimal solution, the modified model-based optimal control problem is converted in a nonlinear programming problem through the canonical formulation, where the gradient formulation can be made. During the iterative procedure, the control sequences are generated as the admissible control law of the model used, together with the corresponding state sequences. Consequently, the optimal solution is updated repeatedly by the adjusted parameters. At the end of iteration, the converged solution approaches to the correct optimal solution of the original optimal control problem in spite of model-reality differences. For illustration, two examples are studied and the results show the efficiency of the approach proposed.

---

2010 *Mathematics Subject Classification.* Primary: 93C05, 93C10; Secondary: 93B40.  
*Key words and phrases.* Nonlinear optimal control, model-reality differences, gradient algorithm, adjusted parameters, iterative solution.

The reviewing process of the paper was handled by Cedric Yiu as Guest Editor.

1. **Introduction.** Linear quadratic regulator (LQR) problem is a standard optimal control problem, where the cost functional is in quadratic criterion and the state dynamics is in a linear form. Solving this optimal control problem is simple and the corresponding optimal solution is guaranteed [2], [5], [15]. Further from this, the applications of the LQR problem have been widely explored; see for examples, [13], [14], [10], [18], [9]. However, the nonlinear state dynamics is always linearized before a decision control policy is determined to minimize the cost function. In this point of view, the adjustable parameters are introduced in the LQR model such that the differences between the real plant and the model used can be measured repeatedly. At the end of iteration, the iterative solution could converge to the correct optimal solution of the original optimal control problem, in spite of model-reality differences [11], [12], [1].

Usually, the sweep method is applied to construct the feedback control law in solving the LQR model-based optimal control problem [2], [5]. It is the same work as those done by [11], [12], [1]. In this paper, we propose an efficient computation approach to construct the control sequences for the optimal control problem with model-reality differences. On this basis, the model-based optimal control problem, which is added with the adjusted parameters, is solved iteratively. Our aim is to obtain the true optimal solution of the original optimal control problem via solving the model-based optimal control problem repeatedly. For doing so, the initial control sequences are defined from the LQR optimal control model. Then, the modified model-based optimal control problem is formulated as a nonlinear programming problem [15], [6], [3]. During each iteration step, the differences between the real plant and the model used are measured by the adjusted parameters. It follows that the value of the control sequences is updated through the gradient algorithm, where the mathematical optimization technique is applicable. Within a given tolerance, the iterative algorithm gives the correct optimal solution of the original optimal control problem despite model-reality differences. It is highly recommended that the gradient algorithm can make the way of solving optimal control problems with model-reality differences more flexible.

The rest of the paper is organized as follows. In Section 2, a general class of optimal control problem is described. In Section 3, a simplified model-based optimal control problem is discussed, where the adjusted parameters are added into the model used. It points out that the interactive between system optimization and parameter estimation gives a modified optimal control problem, which can be solved by the gradient algorithm. Consequently, an efficient iterative algorithm is resulted. In Section 4, two illustrative examples are demonstrated and the efficiency of the approach proposed is shown. Finally, some concluding remarks are made.

2. **Problem Description.** Consider a general class of optimal control problem given below:

$$\begin{aligned} \min_{u(k)} J_0(u) &= \varphi(x(N), N) + \sum_{k=0}^{N-1} L(x(k), u(k), k) \\ \text{subject to} & \\ x(k+1) &= f(x(k), u(k), k) \\ x(0) &= x_0 \end{aligned} \tag{1}$$

where  $u(k) \in \mathfrak{R}^m$ ,  $k = 0, 1, \dots, N-1$ , and  $x(k) \in \mathfrak{R}^n$ ,  $k = 0, 1, \dots, N$ , are, respectively, control and state sequences, whereas  $f : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \rightarrow \mathfrak{R}^n$  represents the real plant,  $\varphi : \mathfrak{R}^n \times \mathfrak{R} \rightarrow \mathfrak{R}$  is the terminal cost and  $L : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \rightarrow \mathfrak{R}$  is the cost under summation. Here,  $J_0$  is the scalar cost function and the initial state  $x_0$  is a known vector. It is assumed that all functions in (1) are continuously differentiable with respect to their respective arguments.

This problem is regarded as the real optimal control problem, and is referred to as Problem (P). Notice that this problem is a complex problem, where the structure of the problem is in nonlinear manner. Solving this kind of the problem is computationally demanding. In view of this, we propose to solve a simplified model-based optimal control problem iteratively in order to obtain the correct optimal solution of Problem (P). Let this simplified model-based optimal control problem, which is referred to as Problem (M), be given below.

$$\begin{aligned} \min_{u(k)} J_1(u) &= \frac{1}{2}x(N)^\top S(N)x(N) + \gamma(N) \\ &+ \sum_{k=0}^{N-1} \frac{1}{2}(x(k)^\top Qx(k) + u(k)^\top Ru(k)) + \gamma(k) \\ \text{subject to} & \\ x(k+1) &= Ax(k) + Bu(k) + \alpha(k) \\ x(0) &= x_0 \end{aligned} \tag{2}$$

where  $\alpha(k) \in \mathfrak{R}^n$ ,  $k = 0, 1, \dots, N-1$ , and  $\gamma(k) \in \mathfrak{R}$ ,  $k = 0, 1, \dots, N$ , are the adjustable parameters, while  $A$  is an  $n \times n$  state transition matrix and  $B$  is an  $n \times m$  control coefficient matrix.  $J_1$  is the model cost function,  $S(N)$  and  $Q$  are  $n \times n$  positive semi-definite matrices, and  $R$  is a  $m \times m$  positive definite matrix.

Notice that solving Problem (M) iteratively would give the true optimal solution of Problem (P). This could be done because of the adjustable parameters that introduced into the model are able to measure the differences between the real plant and the model used repeatedly. In such way, we aim at approximating the correct optimal solution of Problem (P) by solving Problem (M), in spite of model-reality differences.

**3. Gradient Algorithm with Model-Reality Differences.** Now, let us introduce an expanded optimal control problem, which is referred to as Problem (E), given below.

$$\begin{aligned} \min_{u(k)} J_2(u) &= \frac{1}{2}x(N)^\top S(N)x(N) + \gamma(N) \\ &+ \sum_{k=0}^{N-1} \frac{1}{2}(x(k)^\top Qx(k) + u(k)^\top Ru(k)) + \gamma(k) \\ &+ \frac{1}{2}r_1 \|u(k) - v(k)\|^2 + \frac{1}{2}r_2 \|x(k) - z(k)\|^2 \\ \text{subject to} & \\ x(k+1) &= Ax(k) + Bu(k) + \alpha(k) \\ x(0) &= x_0 \\ \frac{1}{2}z(N)^\top S(N)z(N) + \gamma(N) &= \varphi(z(N), N) \end{aligned} \tag{3}$$

$$\begin{aligned}
\frac{1}{2}(z(k)^\top Qz(k) + v(k)^\top Rv(k)) + \gamma(k) &= L(z(k), v(k), k) \\
Az(k) + Bv(k) + \alpha(k) &= f(z(k), v(k), k) \\
v(k) &= u(k) \\
z(k) &= x(k)
\end{aligned}$$

where  $v(k) \in \mathfrak{R}^m$ ,  $k = 0, 1, \dots, N-1$ , and  $z(k) \in \mathfrak{R}^n$ ,  $k = 0, 1, \dots, N$ , are introduced to separate the control sequence and the state sequence in the optimization problem 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  $\frac{1}{2}r_2 \|x(k) - z(k)\|^2$  with  $r_1 \in \mathfrak{R}$  and  $r_2 \in \mathfrak{R}$  are introduced to improve convexity and to facilitate convergence of the resulting iterative algorithm. It is important to note that the algorithm is designed such that the constraints  $v(k) = u(k)$  and  $z(k) = x(k)$  are satisfied upon termination of the iterations, assuming that convergence is achieved. The state constraint  $z(k)$  and the control constraint  $v(k)$  are used for the computation of the parameter estimation and matching schemes, while the corresponding state constraint  $x(k)$  and control constraint  $u(k)$  are reserved for optimizing the model-based optimal control problem. In this way, system optimization and the parameter estimation are mutually interactive.

**3.1. Necessary optimality conditions.** Consider that the Hamiltonian function for Problem (E) is defined by

$$\begin{aligned}
H_2(k) &= \frac{1}{2}(x(k)^\top Qx(k) + u(k)^\top Ru(k)) + \gamma(k) \\
&\quad + \frac{1}{2}r_1 \|u(k) - v(k)\|^2 + \frac{1}{2}r_2 \|x(k) - z(k)\|^2 \\
&\quad + p(k+1)^\top (Ax(k) + Bu(k) + \alpha(k)) \\
&\quad - \lambda(k)^\top u(k) - \beta(k)^\top x(k)
\end{aligned} \tag{4}$$

where  $\lambda(k) \in \mathfrak{R}^m$ ,  $k = 0, 1, \dots, N-1$ , and  $\beta(k) \in \mathfrak{R}^n$ ,  $k = 0, 1, \dots, N-1$ , are modifiers. Then, the augmented cost function becomes

$$\begin{aligned}
J'_2(k) &= \frac{1}{2}x(N)^\top S(N)x(N) + \gamma(N) + p(0)^\top x(0) - p(N)^\top x(N) \\
&\quad + \xi(N)(\varphi(z(N), N) - \frac{1}{2}x(N)^\top S(N)x(N) - \gamma(N)) \\
&\quad + \Gamma^\top (x(N) - z(N)) \\
&\quad + \sum_{k=0}^{N-1} H_2(k) - p(k)^\top x(k) + \lambda(k)^\top v(k) + \beta(k)^\top z(k) \\
&\quad + \xi(k)(L(z(k), v(k), k) - \frac{1}{2}(z(k)^\top Qz(k) + v(k)^\top Rv(k)) - \gamma(k)) \\
&\quad + \mu(k)^\top (f(z(k), v(k), k) - Az(k) - Bv(k) - \alpha(k))
\end{aligned} \tag{5}$$

where  $p(k)$ ,  $\gamma(k)$ ,  $\xi(k)$ ,  $\mu(k)$ ,  $\Gamma$ ,  $\lambda(k)$ , and  $\beta(k)$ , are the appropriate multipliers to be determined later.

Applying the calculus of variation [2], [5], the following necessary optimality conditions are obtained.

(a) Stationary condition:

$$Ru(k) + B^\top p(k+1) - \lambda(k) + r_1(u(k) - v(k)) = 0 \tag{6a}$$

(b) Co-state equation:

$$p(k) = Qx(k) + A^\top p(k+1) - \beta(k) + r_2(x(k) - z(k)) \quad (6b)$$

(c) State equation:

$$x(k+1) = Ax(k) + Bu(k) + \alpha(k) \quad (6c)$$

(d) Boundary conditions:

$$p(N) = S(N)x(N) + \Gamma \text{ and } x(0) = x_0 \quad (6d)$$

(e) Adjustable parameter equations:

$$\varphi(z(N), N) = \frac{1}{2}z(N)^\top S(N)z(N) + \gamma(N) \quad (7a)$$

$$L(z(k), v(k), k) = \frac{1}{2}(z(k)^\top Qz(k) + v(k)^\top Rv(k)) + \gamma(k) \quad (7b)$$

$$f(z(k), v(k), k) = Az(k) + Bv(k) + \alpha(k) \quad (7c)$$

(f) Multiplier equations:

$$\Gamma = \nabla_{z(N)}\varphi - S(N)z(N) \quad (8a)$$

$$\lambda(k) = -(\nabla_{v(k)}L - Rv(k)) - \left( \frac{\partial f}{\partial v(k)} - B \right)^\top \hat{p}(k+1) \quad (8b)$$

$$\beta(k) = -(\nabla_{z(k)}L - Qz(k)) - \left( \frac{\partial f}{\partial z(k)} - A \right)^\top \hat{p}(k+1) \quad (8c)$$

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

(g) Separable variables:

$$z(k) = x(k), v(k) = u(k), \hat{p}(k) = p(k) \quad (9)$$

Notice that the parameter estimation problem is defined by (7) and the multipliers can be calculated from (8). Equations (6a) C (6d) are the necessary conditions for the modified model-based optimal control problem.

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

$$\begin{aligned} \min_{u(k)} J_3(u) &= \frac{1}{2}x(N)^\top S(N)x(N) + \Gamma^\top x(N) + \gamma(N) \\ &+ \sum_{k=0}^{N-1} \frac{1}{2}(x(k)^\top Qx(k) + u(k)^\top Ru(k)) + \gamma(k) \\ &+ \frac{1}{2}r_1 \|u(k) - v(k)\|^2 + \frac{1}{2}r_2 \|x(k) - z(k)\|^2 \\ &- \lambda(k)^\top u(k) - \beta(k)^\top x(k) \end{aligned}$$

subject to (10)

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + \alpha(k) \\ x(0) &= x_0 \end{aligned}$$

with the specified  $\alpha(k), \gamma(k), \Gamma, \lambda(k), \beta(k), v(k)$  and  $z(k)$ , where the boundary conditions  $x(0)$  and  $p(N)$  are given with the specified modifier  $\Gamma$ .

Once the state sequences are determined corresponding to the control sequences which can be defined through the gradient formulation, Problem (MM) could be converted in a nonlinear programming problem as given below [15], [6], [7], [8], [17]:

$$\min_{u(k)} J_3(u) \text{ subject to } u(k) \in \mathfrak{R}^m, \quad k = 0, 1, \dots, N-1.$$

Let this problem be Problem (MM').

**3.3. Admissible control law.** Define

$$V = \{v = [v_1, \dots, v_m]^\top \in \mathfrak{R}^m : a_i \leq v_i \leq b_i, i = 1, \dots, m\} \quad (11)$$

where  $a_i$ ,  $i = 1, \dots, m$ , and  $b_i$ ,  $i = 1, \dots, m$ , are given real numbers. Notice that  $V$  is compact and convex subset of  $\mathfrak{R}^m$ . Let  $u$  denote a control sequence  $\{u(k) : k = 0, 1, \dots, N-1\}$  in  $V$ . Then,  $u$  is called an admissible control. Let  $U$  be the class of all such admissible controls.

For each  $u \in U$ , let  $x(k|u)$ ,  $k = 0, 1, \dots, N$ , be a sequence in  $\mathfrak{R}^n$  such that the difference equations with the initial condition as mentioned in Problem (MM) are satisfied. This discrete function is called the solution of the system in Problem (MM) corresponding to  $u \in U$ .

Let the control vector  $u$  be perturbed by  $\varepsilon \hat{u}$ , where  $\varepsilon > 0$  is a small real number and  $\hat{u}$  is an arbitrary but fixed perturbation of  $u$  given by

$$\hat{u} = [(\hat{u}(0))^\top, (\hat{u}(1))^\top, \dots, (\hat{u}(N-1))^\top]^\top. \quad (12)$$

This gives

$$u_\varepsilon = u + \varepsilon \hat{u} = [(u(0, \varepsilon))^\top, (u(1, \varepsilon))^\top, \dots, (u(N-1, \varepsilon))^\top]^\top \quad (13)$$

where

$$u(k, \varepsilon) = u(k) + \varepsilon \hat{u}(k), \quad k = 0, 1, \dots, N-1. \quad (14)$$

Following from (14), the state of the model used in Problem (MM) will be perturbed, and so is the cost functional  $J_3$ .

**3.4. Gradient formula.** Define the state sequences

$$x(k, \varepsilon) = x(k|\varepsilon), \quad k = 1, 2, \dots, N. \quad (15)$$

Then, the system of difference equations in Problem (MM) becomes

$$x(k+1, \varepsilon) = Ax(k, \varepsilon) + Bu(k, \varepsilon) + \alpha(k). \quad (16)$$

The variation of the state (16) for  $k = 0, 1, \dots, N-1$ , is

$$\Delta x(k+1) = \left. \frac{dx(k+1, \varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = A\Delta x(k, \varepsilon) + B\hat{u}(k) \quad (17a)$$

with

$$\Delta x(0) = 0 \quad (17b)$$

For the cost functional, it is considered that

$$\begin{aligned} \frac{\partial J_3(u)}{\partial u} \hat{u} &= \lim_{\varepsilon \rightarrow 0} \frac{J_3(u_\varepsilon) - J_3(u)}{\varepsilon} \equiv \left. \frac{dJ_3(u_\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} \\ &= (S(N)x(N) + \Gamma)^\top \Delta x(N) \\ &\quad + \sum_{k=0}^{N-1} (Ru(k) + r_1(u(k) - v(k)) - \lambda(k))^\top \hat{u} \\ &\quad + (Qx(k) + r_2(x(k) - z(k)) - \beta(k))^\top \Delta x(k) \end{aligned} \quad (18)$$

Consider the Hamiltonian function defined by (4) and the corresponding necessary conditions (6a) -C (6b), we obtain

$$\begin{aligned} \frac{\partial J_3(u)}{\partial u} \hat{u} &= (S(N)x(N) + \Gamma)^\top \Delta x(N) \\ &+ \sum_{k=0}^{N-1} \left( \frac{\partial H_2(k)}{\partial u} - p(k+1)^\top B \right) \hat{u}(k) + (p(k) - p(k+1)^\top A) \Delta x(k) \end{aligned}$$

Then, it follows from (17a) that

$$\begin{aligned} \frac{\partial J_3}{\partial u} \hat{u} &= (S(N)x(N) + \Gamma)^\top \Delta x(N) \\ &+ \sum_{k=0}^{N-1} p(k)^\top \Delta x(k) - p(k+1)^\top \Delta x(k+1) + \frac{\partial H_2(k)}{\partial u} \hat{u} \end{aligned} \quad (19)$$

Hence, from the boundary conditions (6d) and (17b), it yields that

$$\frac{\partial J_3(u)}{\partial u} \hat{u} = \left[ \frac{\partial H_2(0)}{\partial u(0)}, \dots, \frac{\partial H_2(N-1)}{\partial u(N-1)} \right] \hat{u}.$$

Because of  $\hat{u}$  is arbitrary, we obtain the following gradient formula

$$\frac{\partial J_3(u)}{\partial u} = \left[ \frac{\partial H_2(0)}{\partial u(0)}, \dots, \frac{\partial H_2(N-1)}{\partial u(N-1)} \right]. \quad (20)$$

We present this result in the following as a theorem [15], [6], [7], [8], [17].

**Theorem 3.1.** *Consider Problem (MM'). The gradient of  $J_3(u)$ , where*

$$u = [(u(0))^\top, (u(1))^\top, \dots, (u(N-1))^\top]^\top,$$

*is given by (20).*

**3.5. Gradient algorithm.** The computation of the gradient of the cost functional  $J_3(u)$  is stated in the following algorithm.

*Gradient algorithm*

Data For a given  $u = [(u(0))^\top, (u(1))^\top, \dots, (u(N-1))^\top]^\top$  such that (11) is satisfied.

Step 1 Solve the system of difference equations (6c) forward in time from  $k = 0$  to  $k = N$  with initial condition (6d) to obtain  $x(k|u)$ ,  $k = 0, 1, \dots, N$ .

Step 2 Solve the system of the co-state difference equations (6b) backward in time from  $k = N$  to  $k = 1$ . Let  $p(k|u)$  be the solution obtained.

Step 3 Calculate the value of the cost functional  $J_3(u)$  from (10).

Step 4 Compute the gradient of  $J_3(u)$  according to (20).

**Remark:** The gradient algorithm is used for updating the control sequence, solving the system of difference equations, to calculate the value of  $J_3(u)$  and the corresponding gradient of  $J_3(u)$  in Problem (MM').

**3.6. Iterative algorithm.** From the discussion above, we shall summarize the result as an iterative algorithm, and the computation procedure is given below.

#### The iterative computation procedure

Data  $A, B, Q, R, S(N), x_0, N, r_1, r_2, k_v, k_z, k_p, f, L, \varphi$ . Note that  $A$  and  $B$  may be chosen based on the linearization of  $f$  at  $x_0$  or the linear terms of  $f$ .

- Step 0 Compute a nominal solution. Assuming that  $\alpha(k) = 0$ ,  $k = 0, 1, \dots, N-1$ , and  $r_1 = r_2 = 0$ . Solve Problem (M) defined by (2) to obtain  $u(k)^0$ ,  $k = 0, 1, \dots, N-1$ , and  $x(k)^0, p(k)^0$ ,  $k = 0, 1, \dots, N$ . Then, with  $\alpha(k) = 0$ ,  $k = 0, 1, \dots, N-1$ , and using  $r_1, r_2$  from the data. Set  $i = 0, v(k)^0 = u(k)^0, z(k)^0 = x(k)^0$ , and  $\hat{p}(k)^0 = p(k)^0$ .
- Step 1 Compute the parameters  $\gamma(k)^i, k = 0, 1, \dots, N$ , and  $\alpha(k)^i, k = 0, 1, \dots, N-1$ , from (7). This is called the *parameter estimation step*.
- Step 2 Compute the modifiers  $\Gamma^i, \lambda(k)^i$  and  $\beta(k)^i, k = 0, 1, \dots, N-1$ , from (8). Note that this step requires taking the derivatives of  $f$  and  $L$  with respect to  $v(k)^i$  and  $z(k)^i$ .
- Step 3 Using  $\alpha(k)^i, \gamma(k)^i, \Gamma^i, \lambda(k)^i, \beta(k)^i, v(k)^i$  and  $z(k)^i$ , solve Problem (MM') using the result that is presented in Theorem 3.1 and the gradient algorithm. This is called the *system optimization step*.
- 3.1 Use (20) to obtain the new control  $u(k)^i, k = 0, 1, \dots, N-1$ .
- 3.2 Use (6c) to obtain the new state  $\bar{x}(k)^i, k = 0, 1, \dots, N$ .
- 3.3 Use (6b) to obtain the new costate  $p(k)^i, k = 0, 1, \dots, N$ .
- Step 4 Test the convergence and update the optimal solution of Problem (P). In order to provide a mechanism for regulating convergence, a simple relaxation method is employed:

$$v(k)^{i+1} = v(k)^i + k_v(u(k)^i - v(k)^i) \quad (21a)$$

$$z(k)^{i+1} = z(k)^i + k_z(x(k)^i - z(k)^i) \quad (21b)$$

$$\hat{p}(k)^{i+1} = \hat{p}(k)^i + k_p(p(k)^i - \hat{p}(k)^i) \quad (21c)$$

where  $k_v, k_z, k_p \in (0, 1]$  are scalar gains. If  $v(k)^{i+1} = v(k)^i$ ,  $k = 0, 1, \dots, N-1$ , and  $z(k)^{i+1} = z(k)^i$ ,  $k = 0, 1, \dots, N$ , within a given tolerance, stop; else set  $i = i + 1$ , and repeat the procedure starting with Step 1.

**Remarks:**

- (a) A set of control sequences, which is for solving Problem (M) in Step 0, and for solving Problem (MM') in Step 3, respectively, is determined from (20) by using the gradient algorithm.
- (b) The parameters  $\alpha(k)^i, \gamma(k)^i, \Gamma^i, \lambda(k)^i$ , and  $\beta(k)^i$  are zero in Step 0. Their calculated values, where  $\alpha(k)^i$  and  $\gamma(k)^i$  in Step 1, and  $\Gamma^i, \lambda(k)^i, \beta(k)^i$  in Step 2, change from iteration to iteration.
- (c) Problem (P) is not necessary to be linear or to have a quadratic cost function.
- (d) The conditions  $v(k)^{i+1} = v(k)^i$  and  $z(k)^{i+1} = z(k)^i$  are required to be satisfied for the converged optimal control sequence and the converged state sequence, respectively. The following averaged 2-norms are computed, and then they are compared with a given tolerance to verify the convergence of  $v(k)$  and  $z(k)$ :

$$\|v^{i+1} - v^i\|_2 = \left( \frac{1}{N-1} \sum_{k=0}^{N-1} \|v(k)^{i+1} - v(k)^i\| \right)^{1/2} \quad (22a)$$

$$\|z^{i+1} - z^i\|_2 = \left( \frac{1}{N} \sum_{k=0}^N \|z(k)^{i+1} - z(k)^i\| \right)^{1/2} \quad (22b)$$

- (e) The relaxation scalars ( $k_v, k_z, k_p$ ) are step-sizes that regulate the convergence mechanism. They are normally chosen from the interval  $(0, 1]$ , but this choice



may not result in an optimal number of iterations. It is important to note that the optimal choice of  $k_v, k_z, k_p \in (0, 1]$  is problem dependent, requiring that the proposed algorithm is run several times from Step 1 to Step 4. These values are initially set as  $k_v = k_z = k_p = 1$  for the first run of the algorithm from Step 1 to Step 4, and then the algorithm is run with different values ranging from 0.1 to 0.9. The value that provides the optimal number of iterations can then be determined. The parameters  $r_1$  and  $r_2$  are to enhance convexity, leading to the improvement of the convergence of the algorithm.

**4. Illustrative Examples.** Two illustrative examples are demonstrated here. They are continuous stirred-tank reactor problem [4] and inverted pendulum balancing problem [16].

**Example 1:** Consider a continuous stirred-tank reactor problem. The real plant is given by

$$\begin{aligned} x_1(k+1) &= x_1(k) - 0.02(x_1(k) + 0.25) + 0.01(x_2(k) + 0.5) \exp\left(\frac{25x_1(k)}{x_1(k) + 2}\right) \\ &\quad - 0.01(x_1(k) + 0.25)u(k) \\ x_2(k+1) &= 0.99x_2(k) - 0.005 - 0.01(x_2(k) + 0.5) \exp\left(\frac{25x_1(k)}{x_1(k) + 2}\right) \end{aligned}$$

for  $k = 0, 1, \dots, 77$ , with initial condition

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

Our aim is to determine the control sequences  $u(k)$ ,  $k = 0, 1, \dots, N-1$ , so that the cost function

$$\min_{u(k)} J_0(u) = 0.01 \sum_{k=0}^{N-1} (x_1(k))^2 + (x_2(k))^2 + 0.1(u(k))^2$$

is minimized over the state dynamics.

This problem is referred to as Problem (P).

The simplified model-based optimal control problem, which is referred to as Problem (M), is given below.

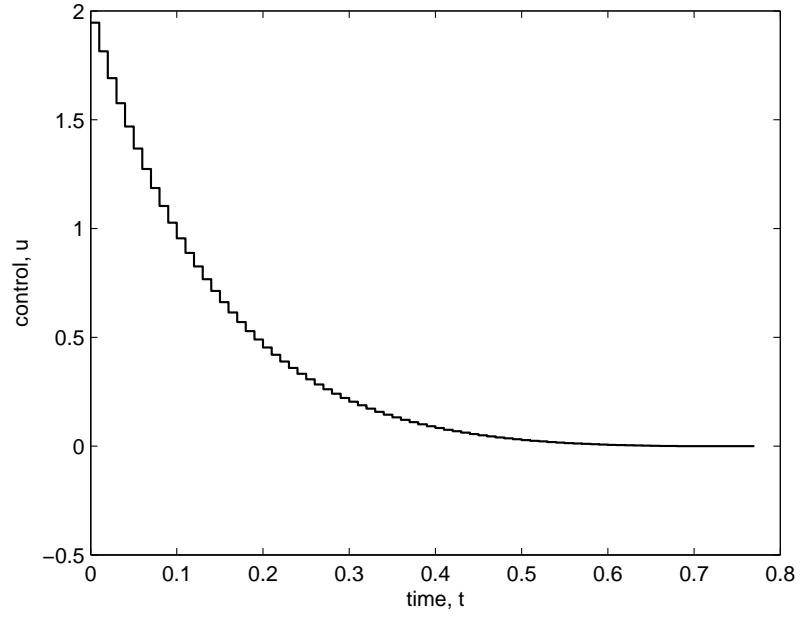
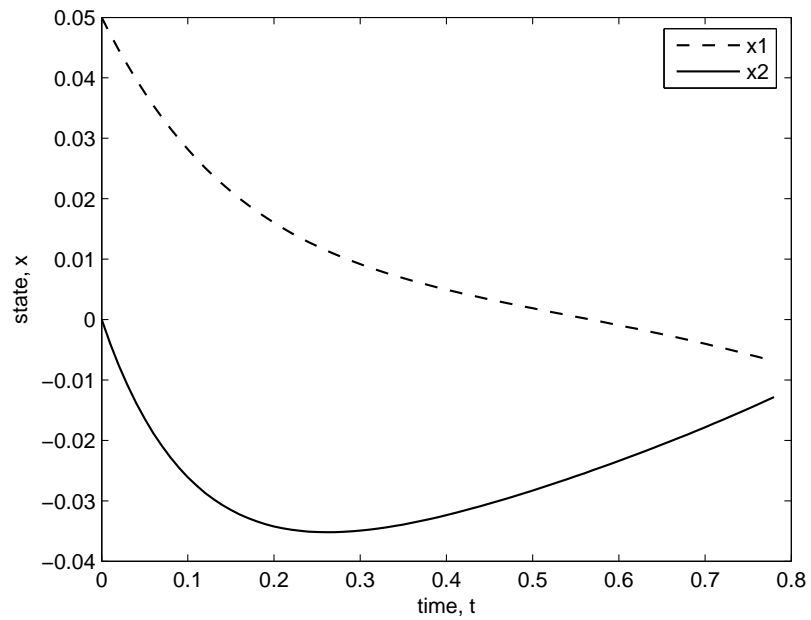
$$\min_{u(k)} J_1(u) = \frac{1}{2} \sum_{k=0}^{N-1} [(x_1(k))^2 + (x_2(k))^2 + 0.1(u(k))^2 + 2\gamma(k)]$$

subject to

$$\begin{bmatrix} \bar{x}_1(k+1) \\ \bar{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 1.0895 & 0.0184 \\ -0.1095 & 0.9716 \end{bmatrix} \begin{bmatrix} \bar{x}_1(k) \\ \bar{x}_2(k) \end{bmatrix} + \begin{bmatrix} -0.003 \\ 0.000 \end{bmatrix} u(k) + \begin{bmatrix} \alpha_1(k) \\ \alpha_2(k) \end{bmatrix}$$

with the initial condition  $x(0) = [0.05 \ 0]^\top$ , and the adjusted parameters  $\gamma(k)$ , and  $\alpha(k) = [\alpha_1(k) \ \alpha_2(k)]^\top$ . The tolerance is set to  $10^{-5}$ .

The simulation result that is shown in Table 1 presents the efficiency of the algorithm proposed, where 99.3% of the initial cost has been reduced. The trajectories of control and state are, respectively, shown in Figures 1 and 2, while the adjusted parameters  $\alpha(k)$  and  $\gamma(k)$  are, respectively, shown in Figures 3 and 4. From the value of the adjusted parameters, **which their values are approximated to zero within the given tolerance**, it seems that the correct optimal solution of the original optimal control problem is obtained in spite of model-reality differences.

FIGURE 1. Final control trajectory  $u(k)$ , Example 1FIGURE 2. Final state trajectory  $x(k)$ , Example 1

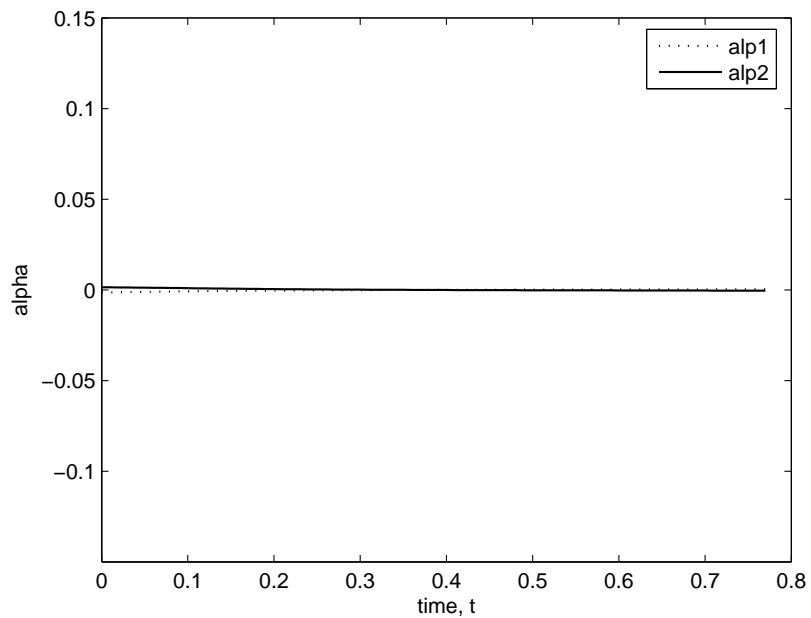


FIGURE 3. Adjusted parameter  $\alpha(k)$ , Example 1

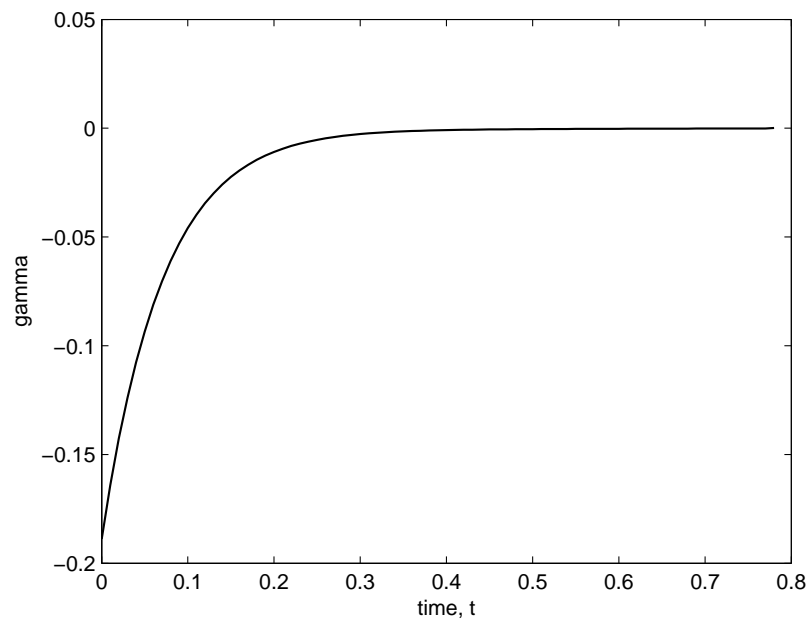


FIGURE 4. Adjusted parameter  $\gamma(k)$ , Example 1

TABLE 1. Simulation result, Example 1

Iteration number	Elapsed time (s)	Initial cost $J_1^*$	Final cost $J_9^*$
9	35.1300	2.0859	0.014634

**Example 2:** Consider an inverted pendulum balancing problem. Problem (P) is described as follows. The state dynamic equations are discretized and given by

$$\begin{aligned} x_1(k+1) &= x_1(k) + 0.1x_2(k) \\ x_2(k+1) &= x_2(k) + \frac{3.234 \sin x_1(k) - 0.015(x_2(k))^2 \cos x_1(k) \sin x_1(k)}{2.2 - 0.15(\cos x_1(k))^2} \\ &\quad + \frac{0.3u(k) \cos x_1(k)}{2.2 - 0.15(\cos x_1(k))^2} \end{aligned} \quad (23)$$

for  $k = 0, 1, \dots, 29$ , with initial condition

$$x_1(0) = 1.0, \quad x_2(0) = 0.5.$$

We aim to determine a set of the control sequences  $u(k)$ ,  $k = 0, 1, \dots, N-1$ , such that the cost function

$$\min_{u(k)} J_0(u) = 0.05 \sum_{k=0}^{N-1} (x_1(k))^2 + (x_2(k))^2 + (u(k))^2$$

is to be minimized over the state dynamics.

The corresponding Problem (M) is given below.

$$\min_{u(k)} J_1(u) = 0.05 \sum_{k=0}^{N-1} [(x_1(k))^2 + (x_2(k))^2 + 0.1(u(k))^2 + 2\gamma(k)]$$

subject to

$$\begin{bmatrix} \bar{x}_1(k+1) \\ \bar{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0.1 \\ 1.5776 & 1 \end{bmatrix} \begin{bmatrix} \bar{x}_1(k) \\ \bar{x}_2(k) \end{bmatrix} + \begin{bmatrix} 0.0000 \\ 0.1463 \end{bmatrix} u(k) + \begin{bmatrix} \alpha_1(k) \\ \alpha_2(k) \end{bmatrix}$$

with the initial condition  $x(0) = [1.0 \ 0.5]^\top$ , and the adjusted parameters  $\gamma(k)$  and  $\alpha(k) = [\alpha_1(k) \ \alpha_2(k)]^\top$ . The given tolerance is  $10^{-3}$ .

The simulation result is shown in Table 2 with the 79% efficiency of the algorithm proposed. Figures 5 and 6 show the trajectories of control and state, respectively, while Figures 7 and 8 show the adjusted parameters  $\alpha(k)$  and  $\gamma(k)$ , respectively. Since the values of the adjusted parameters tend to zero at the end of iteration step, it shows that the true optimal solution of the original optimal control problem is obtained despite model-reality differences.

TABLE 2. Simulation result, Example 2

Iteration number	Elapsed time (s)	Initial cost $J_1^*$	Final cost $J_7^*$
7	2.063	452.5640	94.3415

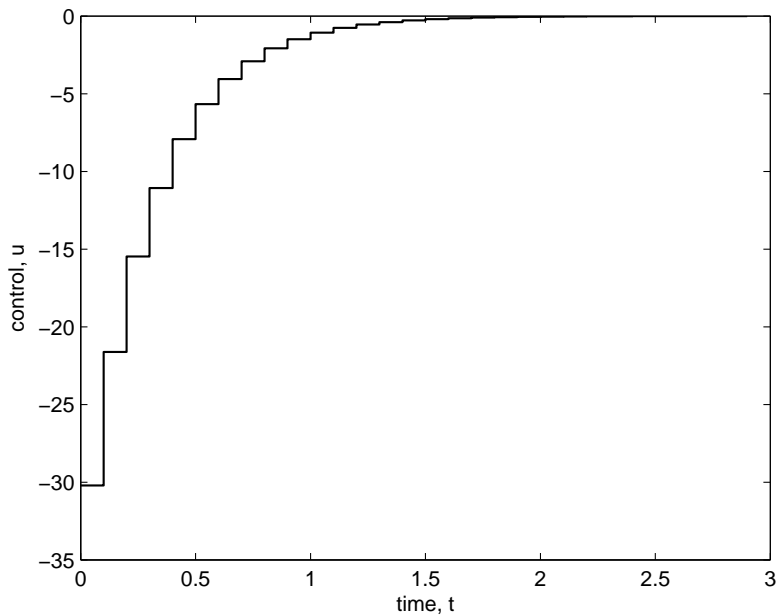


FIGURE 5. Final control trajectory  $u(k)$ , Example 2

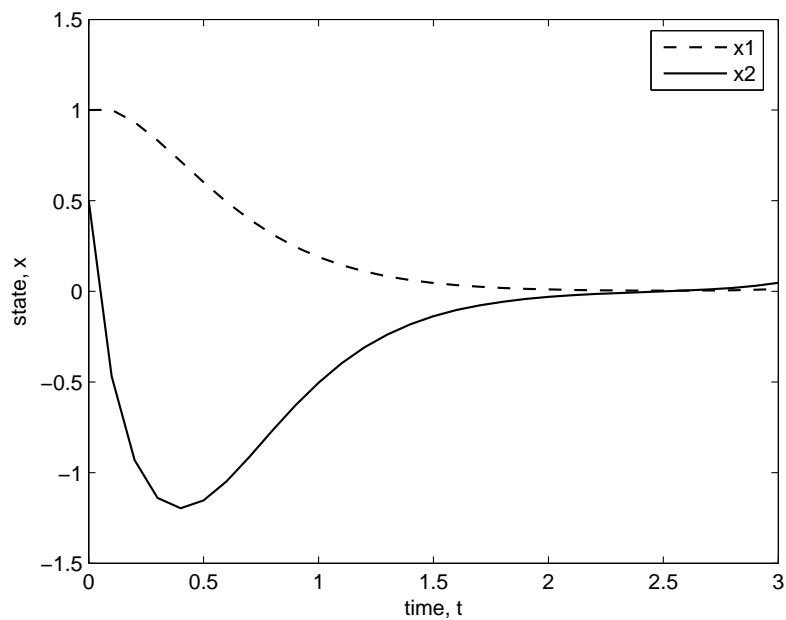
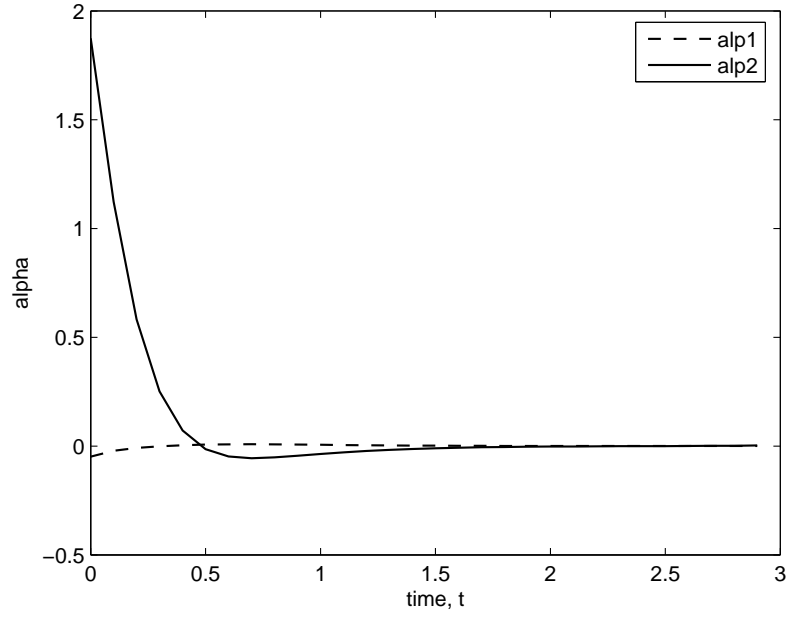
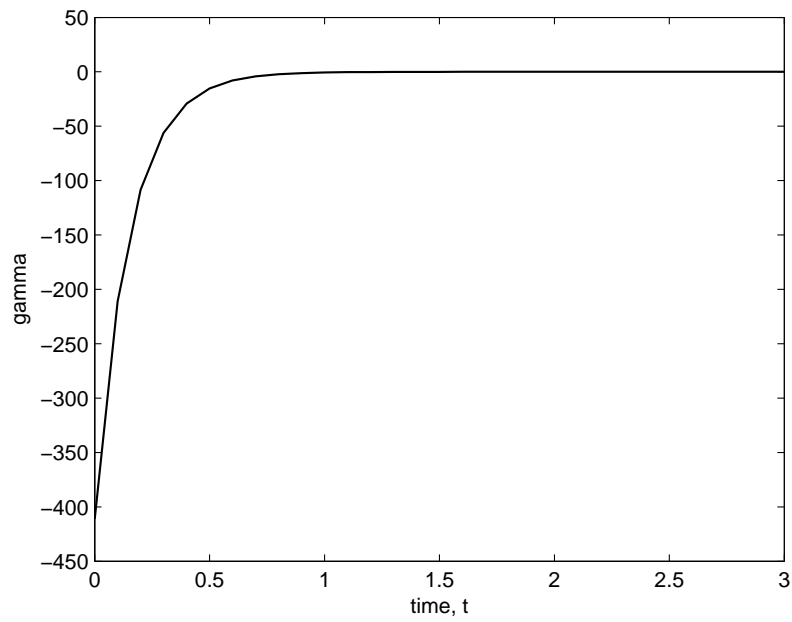


FIGURE 6. Final state trajectory  $x(k)$ , Example 2

FIGURE 7. Adjusted parameter  $\alpha(k)$ , Example 2FIGURE 8. Adjusted parameter  $\gamma(k)$ , Example 2

**5. Concluding Remarks.** A general class of discrete-time optimal control problems, where model-reality differences is taken in account, was discussed in this paper. Because of the complexity of the original optimal control problem, a simplified model-based optimal control problem was proposed to be solved iteratively such that the true optimal solution of the original optimal control problem could be obtained. With introducing an expanded optimal control problem, we integrated system optimization and parameter estimation interactively. In addition to this, a modified optimal control problem was formulated as a nonlinear programming problem and it was solved by using the gradient approach. The resulting iterative algorithm, which integrates the gradient algorithm and the model-reality differences, shows the efficiency through the illustrative examples discussed. **On the other hand, the convergence of the adjusted parameters is guaranteed as Lipschitz condition is satisfied.** In conclusion, the applicability of the algorithm proposed is highly recommended for solving nonlinear optimal control problems.

#### REFERENCES

- [1] V. M. Becerra and P. D. Roberts, [Dynamic integrated system optimization and parameter estimation for discrete time optimal control of nonlinear systems](#), *Int. J. Control*, **63** (1996), 257–281.
- [2] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*, Hemisphere Publishing Company, New York, 1975.
- [3] S. L. Kek, Nonlinear programming approach for optimal control problems, Proceeding of the 2nd International Conference on Global Optimization and Its Applications, (2013), 20–25.
- [4] D. E. Kirk, *Optimal Control Theory: An Introduction*, Mineola, NY: Dover Publications, 2004.
- [5] F. L. Lewis and V. L. Syrmos, *Optimal Control*, 2nd ed, John Wiley & Sons, 1995.
- [6] Q. Lin, R. Loxton and K. L. Teo, [The control parameterization method for nonlinear optimal control: a survey](#), *Journal of Industrial and Management Optimization*, **10** (2014), 275–309.
- [7] R. Loxton, K. L. Teo and V. Rehbock, [Computational method for a class of switched system optimal control problems](#), *IEEE Transactions on Automatic Control*, **54** (2009), 2455–2460.
- [8] R. Loxton, K. L. Teo, V. Rehbock and K. F. C. Yiu, [Optimal control problems with a continuous inequality constraint on the state and the control](#), *Automatica*, **45** (2009), 2250–2257.
- [9] L. F. Lupián and J. R. Rabadán-Martin, LQR control methods for trajectory execution in omnidirectional mobile robots, *Recent Advances in Mobile Robotics*, (2011), 385–400.
- [10] L. H. Nguyen, S. Park, A. Turnip and K. S. Hong, Application of LQR control theory to the design of modified skyhook control gains for semi-active suspension systems, Proceeding of ICROS-SICE International Joint Conference, (2009), 4698–4703.
- [11] P. D. Roberts and T. W. C. Williams, [On an algorithm for combined system optimization and parameter estimation](#), *Automatica*, **17** (1981), 199–209.
- [12] P. D. Roberts, Optimal control of nonlinear systems with model-reality differences, Proceedings of the 31st IEEE Conference on Decision and Control, **1** (1992), 257–258.
- [13] R. C. H. del Rosario and R. C. Smith, LQR control of shell vibrations via piezoceramic actuators, NASA Contractor Report 201673, ICASE Report No. 97-19, 1997.
- [14] J. Saak and P. Benner, Application of LQR techniques to the adaptive control of quasilinear parabolic PDEs, Proceedings in Applied Mathematics and Mechanics, 2007.
- [15] K. L. Teo, C. J. Goh and K. H. Wong, *A Unified Computational Approach to Optimal Control Problem*, Longman Scientific and Technical, Essex, 1991.
- [16] L. X. Wang, *A Course in Fuzzy Systems and Control*, Upper Saddle River, NJ, 1997.

- [17] C. Z. Wu, K. L. Teo and V. Rehbock, [Optimal control of piecewise affine systems with piecewise affine state feedback](#), *Journal of Industrial and Management Optimization*, **5** (2009), 737–747.
- [18] B. Yang and B. Xiong, Application of LQR techniques to the anti-sway controller of overhead crane, *Advanced Material Research*, **139-141** (2010), 1933–1936.

Received May 2014; 1<sup>st</sup> revision June 2014; final revision March 2015.

*E-mail address:* [slkek@uthm.edu.my](mailto:slkek@uthm.edu.my)

*E-mail address:* [mismail@utm.my](mailto:mismail@utm.my)

*E-mail address:* [K.L.Teo@curtin.edu.au](mailto:K.L.Teo@curtin.edu.au)