A Unified Approach of Controller Design and System Identification for Optimal Control of Single-Input Nonlinear Systems

Yuji Ito *, Kenji Fujimoto **, Takayoshi Yoshimura *, and Yukihiro Tadokoro *

Abstract: This paper presents a novel unified approach of controller design and identification for unknown input affine nonlinear systems. An issue with obtaining the best performance of optimal control is that identification errors degrade control performance. One solution to overcome it is direct controller tuning without system identification, which is expected to give high quality control. However, many experiments with various control inputs are required in the design procedure. The proposed framework simultaneously implements system identification and controller design. This method adopts a weighted least squares method to cope with various identification criteria. An unknown plant system is identified with a weight selected appropriately for control. It is selected in a simple manner based on an analytical discussion of nonlinear optimal control theory, namely the Hamilton-Jacobi-Bellman equation. An iterative calculation procedure for a preliminarily given data set is derived. This iterative algorithm gives a (sub-)optimal pair of the weight and the model parameter that improves the control performance. Numerical results demonstrate that the proposed approach achieves better optimal control performance than the standard least squares method.

Key Words: nonlinear optimal control, nonlinear systems, learning control, unknown systems.

1. Introduction

Many researchers have attempted to control various unknown nonlinear systems. One example is to control a vehicle with a human driver, whose dynamics is complex and hence unknown, causing nonlinear behavior [1], [2]. A rechargeable battery is also an example because its physical dynamics is still unclear [3]. The aim of this paper is to control an unknown input affine nonlinear system with weak nonlinearity. The control strategy on which we focus is optimal control.

A typical approach in optimal control of an unknown system consists of two steps. The first step is identification of the system, and the second is building a control law based on the obtained model. Many identification methods try to find a model minimizing the prediction error between the observed training data and the system model [4]. However, as will be described in Section 2, control laws for such a model do not give optimal control performance.

The above mentioned problem motivates to propose an alternative framework focusing on the control performance. System identification methods based on control performance criteria have been proposed in [5], [6]. Another solution is direct controller design such as model-free approaches in [7]–[10]. The method in [11] refines a controller based on iterative experiments to improve the performance. Although these methods are expected to give high quality control, they require iterative experiments in their design procedures. Such experiments are repeated with various control inputs. However, this is not practical in real world applications, including human-in-the-loop systems and electric power networks. It is desirable to obtain a controller and/or a system model from only a preliminarily given data set without iterative experiments.

Several attempts have been made to find direct controller design methods without iterative experiments. Methods that tune a controller using a single data set have been proposed [12], [13]. However, since these methods focus on feedforward control, they are not applicable to feedback optimal control. Another method based on frequency response has been applied to linear feedback controllers [14]. Unfortunately, it is not applicable to nonlinear systems. A model-free predictive control method [15] for nonlinear systems is proposed, but its control performance is not discussed.

To maximize the performance of optimal control of input affine nonlinear systems without iterative experiments, we propose a unified approach of control and identification. In system identification, a weighted least squares (WLSQ) method is introduced to cope with various types of criteria evaluating model parameters. The criterion is defined variably by the weight of a given training data set. To obtain a model parameter that maximizes control performance, the weight of the WLSQ method is selected based on knowledge of the system dynamics and the controller by a procedure as shown in the following steps. First, the maximization problem to be solved is analytically described by an optimization problem with respect to the weight of the WLSQ method in system identification. Second, the optimized weight is approximated by a function of the model parameter, which is still unknown before executing system identification. Because the model parameter and the weight are interdependent, an iterative calculation method is employed. Thus, system identification and controller design are implemented simultaneously using a WLSQ method with the designed weight. The proposed approach is applicable to optimal control of in-
put affine nonlinear systems because it is based on nonlinear optimal control theory, namely the Hamilton-Jacobi-Bellman (HJB) equation. The resulting controller is designed using a preliminarily given data set only without iterative experiments. A numerical results demonstrate that the optimal control performance of the proposed approach is better than that of the standard least squares (LSQ) method.

The outline of this paper is as follows. Section 2 introduces the problem setting, which is derived from a nonlinear optimal control problem, and the system identification methods. Section 3 proposes a unified approach of control and identification for solving the problem. The proposed method is derived from transformations based on theorems presented in this section. Section 4 evaluates the effectiveness of the proposed method in comparison with the LSQ method through numerical examples using various types of systems and training data. Finally, conclusions and future work are discussed in Section 5.

2. Problem Setting

Section 2.1 introduces an optimal control problem of input affine nonlinear systems using system models. Section 2.2 explains that a typical procedure of identification and control degrades the optimal control performance. Section 2.3 provides a novel strategy to overcome it.

2.1 Model-Based Nonlinear Optimal Control

Let us consider the following class of single-input affine nonlinear dynamical systems:

\[ \Sigma : \dot{x} = f(x) + g(x)u, \]  

where \( x \in \mathbb{R}^n \) and \( u \in \mathbb{R} \) are the state and control input at time \( t \), respectively. The symbol \( \Sigma \) denotes an unknown actual system. The system dynamics functions are \( f(x) \in \mathbb{R}^n \) and \( g(x) \in \mathbb{R}^n \). It is assumed that \( f(x) \) and \( g(x) \) are sufficiently smooth with respect to \( x \) and \( f(0) = 0 \) holds. These assumptions imply that the system is weakly nonlinear near the origin. Although system noise is assumed to be negligible in (1), observation noise will be considered in (6).

This paper focuses on a control performance maximization problem, which is regarded as minimization of a cost function \( J(x_0, u) \). An optimal control problem is introduced to minimize \( J(x_0, u) \) for a given initial state \( x_0 \):

\[
u_*(x_0) = \arg \min_u J(x_0, u),
\]

\[ J(x_0, u) := \int_0^\infty q(x) + \frac{1}{2} R(x)u^2 \, dt, \]

where a positive definite function \( q(x) \in \mathbb{R} \) and a positive function \( R(x) > 0 \in \mathbb{R} \) are design parameters. To obtain the optimal input \( u_*(x_0, u) \), the HJB equation [16] is introduced:

\[ H_{HJB}(p(x), f(x), g(x)) := p(x)^T f(x) + q(x) - \frac{1}{2} p(x)^T g(x) R(x)^{-1} g(x)^T p(x) = 0, \]

\[ p(x) := \frac{\partial V(x)}{\partial x} \in \mathbb{R}^n, \]

\[ V(x_0) := \min_u J(x_0, u) \]

where the value function \( V(x) \in \mathbb{R} \) is a positive definite function. Let us assume that such a solution \( V(x) \) exists and that it is a \( C^2 \) function of \( x \). The function \( p(x) \in \mathbb{R}^n \) corresponds to the controller because \( u_*(x) \) is a function of \( p(x) \) [16]:

\[ u_*(x) = -R(x)^{-1} g(x)^T p(x). \]

However, \( u_*(x) \) cannot be determined because the actual system \( \Sigma \) in (1) is unknown.

Alternatively, a parametric system model can be used to estimate the unknown actual system \( \Sigma \) as

\[ \hat{\Sigma}(\theta) : \dot{x} = \hat{f}(x, \theta) + \hat{g}(x, \theta)u, \]

where \( \hat{f}(x, \theta) \in \mathbb{R}^n \) and \( \hat{g}(x, \theta) \in \mathbb{R}^n \) constitute a parametric model \( \hat{\Sigma}(\theta) \) of the system dynamics. The symbol \( \theta \in \Omega \) is the model parameter of \( \hat{\Sigma}(\theta) \) in a given region \( \Omega \in \mathbb{R}^n \). The model \( \hat{\Sigma}(\theta) \) is identified from a training data set, consisting of the inputs \( u^{(i)} \) and the states \( x^{(i)} \) and their time derivatives \( \dot{x}^{(i)} \) (\( i = 1, 2, \ldots, N \)), and \( N \) is the number of training data. In general, it is desirable that the training data is evenly distributed on the state and input spaces to reduce the influence due to the bias of the data set. Let us assume that \( \dot{x}^{(i)} \) is measurable which is subject to an unknown observation noise \( v^{(i)} \in \mathbb{R}^n \) as

\[ \dot{x}^{(i)} = f(x^{(i)}) + g(x^{(i)})u^{(i)} + v^{(i)}. \]

If the system model \( \hat{\Sigma}(\theta) \) is obtained using the training data set, then the optimal input \( \hat{u}_*(x, \theta) \) with respect to \( \hat{\Sigma}(\theta) \) is calculated instead of \( u_*(x) \) in (4):

\[ \hat{u}_*(x, \theta) = -R(x)^{-1} \hat{g}(x, \theta)^T \hat{p}(x, \theta), \]

where \( \hat{p}(x, \theta) \) is a solution to the HJB equation

\[ H_{HJB}(\hat{p}(x, \theta), \hat{f}(x, \theta), \hat{g}(x, \theta)) = 0. \]

Let us assume that there exists a positive definite \( C^2 \) function \( \hat{V}(x) \) of \( x \) such that \( \hat{p}(x, \theta) := \hat{V}(x)/\hat{\theta}x \) holds. The HJB equation (8) can be solved using several techniques in [17]–[19]. In some special cases, explicit solutions to (8) are obtained [20], [21]. If the model \( \hat{\Sigma}(\theta) \) is linear, \( R(x) \) is a constant, and \( q(x) = x^T Q x \) holds for a positive definite \( Q \), \( \hat{p}(x, \theta) := \hat{P}(\theta)x \) is calculated by solving the Riccati equation

\[ \hat{P}(\theta)\hat{A}(\theta) + \hat{A}(\theta)^T \hat{P}(\theta) - \hat{P}(\theta)\hat{B}(\theta)R^{-1}\hat{B}(\theta)^T \hat{P}(\theta) + 2Q = 0, \]

where \( \hat{A}(\theta) \) and \( \hat{B}(\theta) \) are the system matrices of \( \hat{\Sigma}(\theta) \).

2.2 Control Performance Degradation due to Identification Error

This paper treats the case where \( \hat{u}_*(\theta) \neq u_*(x) \) because of identification error. The identification error of \( \hat{\Sigma}(\theta) \) necessarily exists if the unknown actual system \( \Sigma \) in (1) is not included in the set \( \{ \hat{\Sigma}(\theta) \theta \in \Omega \} \). This case often occurs because it is difficult in practice to find a model that is perfectly identical to \( \Sigma \). A suboptimal input \( \hat{u}_*(\theta) \) gives

\[ J(x_0, \hat{u}_*(\theta)) \geq J(x_0, u_*) = V(x_0). \]

It is desirable to find a model parameter \( \theta \) that minimizes \( J(x_0, \hat{u}_*(\theta)) \). However, \( J(x_0, \hat{u}_*(\theta)) \) is unknown because the use of iterative control experiments is outside the scope of this paper. We try to rewrite the cost function \( J(x_0, \hat{u}_*(\theta)) \) as a function of the preliminarily obtained training data set in (6). To this end, the following lemma in [16] is employed.
We try to find a model parameter \( \theta \) using the training data set. The prediction error identification method that obtains a model parameter \( \theta \) cannot take care of the control performance. Hence, we introduce a standard method to improve control performance. Before solving the problem (13), we introduce a standard problem in (13) which is an approximation of the original criterion \( J(\theta) \) defined variously by \( W(1:N) \) [22]. Moreover, \( J(\theta) \) is a convex function of \( \theta \) under the conditions that \( e^{\theta}(\theta) \) is a linear function of \( \theta \) and each \( W(i) \) is positive semidefinite. Thus the globally optimal parameter can be obtained.

The model parameter \( \theta_{\text{ws}} \) in (16) clearly depends on the weight \( W(1:N) \). If the two criteria in (16) and (13) coincide with each other, \( \theta_{\text{ws}} \) is equal to the optimal parameter \( \theta_{\text{data}} \) in (13) for improving the control performance

\[
J(\theta, W(1:N)) = J_{\text{data}}(\theta), \quad \forall \theta \in \Omega \Rightarrow \theta_{\text{ws}} = \theta_{\text{data}},
\]

This motivates us to design \( W(1:N) \) so that we approximate \( \theta_{\text{data}} \) in (13) by \( \theta_{\text{ws}} \) using the WLSQ method (16). There may not exist \( W(1:N) \) such that \( J(\theta, W(1:N)) = J_{\text{data}}(\theta) \) for all \( \theta \in \Omega \) due to the characteristics of the training data set. We try to find \( W(1:N) \) minimizing the difference between \( J(\theta, W(1:N)) \) and \( J_{\text{data}}(\theta) \) in terms of the mean square error with respect to \( \theta \).

**Main Problem:** Find the optimal weight \( W_{\text{ws}}(1:N) \) defined as

\[
W_{\text{ws}}(1:N) = \arg \min_{W(1:N)} \Delta J(\theta, W(1:N)),
\]

\[
\Delta J(\theta, W(1:N)) = E_{\Omega}[J(\theta, W(1:N)) - J_{\text{data}}(\theta)],
\]

where \( E_{\Omega}[\cdot] \) denotes the expectation with respect to \( \theta \), which is assumed to be uniformly distributed on \( \Omega \).

After finding \( W_{\text{ws}}(1:N) \), the system model \( \Sigma(\theta_{\text{ws}}, W_{\text{ws}}(1:N)) \) is identified using (16). Solving (8) or (9) using the identified model yields a suboptimal input \( \hat{u}(\theta_{\text{ws}}, W_{\text{ws}}(1:N)) \). The following section derives a solution \( W_{\text{ws}}(1:N) \) of the main problem (18).

### 3. Unified Approach of Control and Identification

#### 3.1 Notations

The following notations are used in this paper.

- \( I_n \): the \( n \times n \) identity matrix
- \( \|M\|_F \): the Frobenius norm of a matrix \( M \in \mathbb{R}^{n \times n} \)
- \( [v]_i \): the \( i \)-th component of a vector \( v \in \mathbb{R}^n \)
- \( [M]_{ij} \): the component in the \( i \)-th row and \( j \)-th column of a matrix \( M \in \mathbb{R}^{n \times n} \)
- \( \partial \): the partial derivative \( \partial_a(f, g, p) \)
- \( \partial \): the partial derivative \( \partial_a(u, f, g, p) \) evaluated at the training data \( x(t) \)
- \( \Delta u^{(i)}(\theta) \): the difference \( \hat{u}_t(x(t), \theta) - u(x(t)) \) between the functions related to the system model \( \Sigma(\theta) \) and the actual system \( \Sigma \)
- \( O(\epsilon) \): the set of terms which are of order \( \epsilon \) or smaller, \( \|O(\epsilon)\|_F \leq M \epsilon \)
- \( \Delta f^{(i)}(\theta), \Delta g^{(i)}(\theta), \) and \( \Delta p^{(i)}(\theta) \) are defined in a similar manner. Note that \( \Delta u^{(i)}(\theta), \Delta f^{(i)}(\theta), \Delta g^{(i)}(\theta), \) and \( \Delta p^{(i)}(\theta) \) are unknown because they depend on the unknown functions \( u(x), f(x), g(x), \) and \( p(x) \).
3.2 Overview of the Proposed Method

An optimal weight \( W_{1:N} \) in (18) is a function of \( f(x^{(i)}), g(x^{(i)}) \), and \( p(x^{(i)}) \) (\( i = 1, \ldots, N \)) because \( J_{data}(\theta) \) in (13) and (18) already contains \( u(x) \) which depends on \( f, g, \) and \( p \). This makes the problem difficult to solve. In this subsection, we propose a unified approach of the control and identification to obtain \( W_{1:N} \) overcoming this issue.

An optimal weight \( W_{1:N} \) is approximated using the following steps corresponding to Sections 3.3, 3.4, and 3.5, respectively:

\[
W_{d}(27), (28) \Rightarrow W_{d}(\tilde{d}_{1}u_{1}^{*}, \tilde{d}_{2}u_{2}^{*})
\]

\[
= W_{d}(f(x^{(i)}), g(x^{(i)}), p(x^{(i)}))
\]

\[
(39) \Rightarrow W_{d}(f(x^{(i)}), g(x^{(i)}), p(x^{(i)}))
\]

\[
= W_{d}(\theta, k).
\]

First, \( W_{1:N} \) is described by a function \( W_{1:N} \) of the partial derivatives \( \tilde{d}_{1}u_{1}^{*} \) and \( \tilde{d}_{2}u_{2}^{*} \) (Section 3.3). Second, \( \tilde{d}_{1}u_{1}^{*} \) and \( \tilde{d}_{2}u_{2}^{*} \) are derived as functions of \( f, g, \) and \( p \). This yields \( W_{1:N} \) as a function \( W_{1:N} \) of \( f, g, \) and \( p \) (Section 3.4). Third, recall that \( f(\theta), g(\theta), \) and \( p(\theta) \) in (5) and (7) are estimates of \( f, g, \) and \( p \). The weight \( W_{1:N} \) is given by a function \( W_{1:N} \) of the model parameter \( \theta \) (Section 3.5). A regularization matrix \( k \) is introduced to ensure that the system identification using (16) is numerically stable. The weight \( W_{1:N} \) depends on both the system dynamics \( f, g, \) and the controller \( p, \) meaning that the proposed method is a unified approach.

Unfortunately, \( W_{1:N} \) cannot be calculated directly because this calculation requires the model parameter \( \theta \), but \( \theta \) is identified by \( W_{1:N} \). To overcome this interdependence, we propose an iterative scheme to simultaneously obtain both parameters (Section 3.6). Through the above procedure, suboptimal \( W_{1:N} \) and \( \theta_{w} \) are calculated. Finally, suboptimal functions \( f(x, \theta_{w}), g(x, \theta_{w}), \tilde{p}(x, \theta_{w}), \) and \( \tilde{u}_{a}(x, \theta_{w}) \) are obtained.

The proposed approach is summarized in Algorithm 1. Algorithm 1 indicates that the unified approach iterates system identification, calculating the controller, and updating the criterion to improve control performance. The convergence of \( \theta_{w}^{*} \) will be discussed in Section 4.3. The following subsections describe the detail of the steps.

3.3 Weight Design Based on Criterion Transformation

To solve the main problem (18) of optimizing \( W_{1:N} \), this subsection approximates \( J_{data}(W_{1:N}) \) in (18) as an explicit form of \( W_{1:N} \). Note that \( J_{data}(\theta) \) in (12) and \( J_{W}(\theta, W_{1:N}) \) in (16) are the functions of \( \tilde{d}_{1}u_{1}^{*} \) and \( \tilde{d}_{2}u_{2}^{*} \), respectively. The function \( p \) is regarded as a function of \( f, g \) and because of (3). Let us assume that \( u, g, p(f, g) \) in (4) is a C^2 function of \( f, g \). Taylor’s theorem is applied to \( u, g, p(f, g) \) with respect to \( f, g \) as. \( \Delta f^{(0)}(\theta)^{T}, \Delta g^{(0)}(\theta)^{T} \rightarrow 0, \Delta u^{(0)}(\theta) \) is represented as

\[
\Delta u^{(0)}(\theta) = \tilde{d}_{1}u_{1}^{*}(f, g, \Delta f^{(0)}(\theta) + \tilde{d}_{2}u_{2}^{*}(f, g, \Delta g^{(0)}(\theta)), \Delta u^{(0)}(\theta)) = O(||\Delta f^{(0)}(\theta)^{T}, \Delta g^{(0)}(\theta)^{T}||^{2})
\]

\[
= \left\{ \begin{array}{l}
\Delta u(\theta) = \Delta u^{(0)}(\theta)
\end{array} \right\} + O(||\Delta f^{(0)}(\theta)^{T}, \Delta g^{(0)}(\theta)^{T}||^{2}),
\]

where \( \Delta u^{(0)}(\theta) \in R^{2n} \) is defined by

\[
\Delta u^{(0)}(\theta) = \Delta f^{(0)}(\theta) \tilde{d}_{1}(\theta) + \Delta g^{(0)}(\theta) \tilde{d}_{2}(\theta) + O(||\Delta f^{(0)}(\theta)^{T}, \Delta g^{(0)}(\theta)^{T}||^{2}).
\]
Proof The proof is given in Appendix A.

Theorem 1 shows that $\Delta J(W^{(1:N)})$ is a quadratic function of $\{M^{(0)}_w(W^{(0)}) - M^{(0)}\}$ if the residual $r_j$ is negligible. This means that the minimizer of $\Delta J(W^{(1:N)})$ is equivalent to that of the quadratic function in (23). However, it is difficult to optimize $W^{(1:N)}$ because $\Delta J(W^{(1:N)})$ is unknown. The following extension of Theorem 1 focuses on the order of $\Delta J(W^{(1:N)})$ with respect to the Frobenius norms $\|M^{(0)}_w(W^{(0)}) - M^{(0)}\|_{F}$. Theorem 2, these are still unknown. The partial derivatives $\partial J/\partial u_i$ and $\partial J/\partial y$ are represented by explicit functions of $f$, $g$, and $p$ in the next subsection to calculate $W^{(1:N)}_{\partial u_i}$. The proof is given in Appendix C.

While $W^{(1:N)}_{\partial u_i}$ is written as a function of $\partial_1 u_i$ and $\partial_2 u_i$ in Theorem 2, these are still unknown. The partial derivatives $\partial_1 u_i$ and $\partial_2 u_i$ are represented by explicit functions of $f$, $g$, and $p$ in the next subsection to calculate $W^{(1:N)}_{\partial u_i}$.

3.4 Partial Derivatives of the Optimal Input

This subsection derives the unknown partial derivatives $\partial_1 u_i$ and $\partial_2 u_i$, as functions of $f$, $g$, and $p$, as shown in (19). Recall that $u_i$ and $p$ are regarded as functions of $f$ and $g$ because of (3) and (4). However, their direct derivation is difficult because $\partial_1 u_i$ and $\partial_2 u_i$ depend on unknown partial derivatives $\partial_1 p$ and $\partial_2 p$ from the definition of $u_i$. in (4):

$$
\partial_1 u_i(f, g) = -\frac{1}{R(x)} g^T \partial_1 p,
\partial_2 u_i(f, g) = -\frac{1}{R(x)} (g^T \partial_2 p + p^T).
$$

To overcome the difficulty, we propose an procedure consisting of two steps, which obtains their approximations. First, approximations $\hat{\partial_1 p}$ and $\hat{\partial_2 p}$ of $\partial_1 p$ and $\partial_2 p$, respectively, are derived by linearizing the system dynamics $f(x)$, $g(x)$, and the controller $p(x)$. This is explained in Theorem 3. Second, Theorem 4 derives approximations $\hat{\partial_1 u}$ and $\hat{\partial_2 u}$ of $\partial_1 u$ and $\partial_2 u$, so that approximation error arising from $\hat{\partial_1 p} - \partial_1 p$ and $\hat{\partial_2 p} - \partial_2 p$ is minimized.

The vector field of the feedback system of (1) with the optimal input $u(x)$ is defined as $f_2(x)$. Suppose that $p(f, g)$ is partially differentiable with respect to $f$ and $g$, that $f(x)$, $f_2(x)$, $p(x)$, $\partial_1 p(x)$, and $\partial_2 p(x)$ are $C^2$ functions of $x$, and that $g(x)$ and $R(x)^{-1}$ are $C^1$ functions of $x$. Note that $f_2(0) = 0$ and $p(0) = 0$. As $x \rightarrow 0$, the Taylor’s theorem is applied:

$$
f_2(x) := f(x) + g(x) u(x) = A_4 x + O_s(\|x\|^2),
p(x) := P x + O_s(\|x\|^3),
g(x) = B + O_s(\|x\|),
R(x)^{-1} := R_0^{-1} + O_s(\|x\|),$$

where $A_4 := \partial_1 f_2(x)|_{x=0} \in \mathbb{R}^{n \times n}$, $P := \partial_1 p(x)|_{x=0} \in \mathbb{R}^{n \times n}$, $B := g(x)|_{x=0} \in \mathbb{R}^n$, and $R_0 := R(x)|_{x=0} > 0 \in \mathbb{R}$ are constant. Supposing that $A_4^{-1}$ exists, we derive the following theorem.

Theorem 3 (Linearized partial derivative of $p$) Let us define

$$
\hat{\partial_1 p} := -A_4^{-1} P^T \in \mathbb{R}^{n \times n},
\hat{\partial_2 p}(x) := \{\Pi_1 x, \Pi_2 x, \ldots, \Pi_N x\} \in \mathbb{R}^{n \times n},
$$

where $\Pi_k \in \mathbb{R}^{n \times n}$ is a symmetric matrix satisfying

$$
A_4 \Pi_k + \Pi_k A_4 = \frac{1}{R_0} (e_k B^T + B e_k^T) P,
$$

and $e_k \in \mathbb{R}^n$ is a basis vector such that $|e_k| = 1$ and $[e_k]_l = 0$ for $l \neq k$. The following relations hold:

$$
\hat{\partial_1 p} - \partial_1 p(x) = O_{acu}(\|x\|) \text{ as } x \rightarrow 0,
\hat{\partial_2 p}(x) - \partial_2 p(x) = O_{acu}(\|x\|^2) \text{ as } x \rightarrow 0.
$$

Proof The proof is given in Appendix D.

Based on Theorem 3, $\partial_1 p$ and $\partial_2 p$ are approximated by $\hat{\partial_1 p}$ and $\hat{\partial_2 p}$, respectively. Solving the linear matrix equation (33) obtains $\Pi_k$ in (32). Next, the following theorem introduces approximations $\hat{\partial_1 u}$ and $\hat{\partial_2 u}$ of $\partial_1 u$ and $\partial_2 u$, to minimize approximation error arising from $\hat{\partial_1 p} - \partial_1 p$ and $\hat{\partial_2 p} - \partial_2 p$.

Theorem 4 (Representation of partial derivatives of $u_i$) Suppose that $p(x)^T g(x) \neq 0$ and $f_2(x) \neq 0$. Let us define

$$
\hat{\partial_1 u}(s, f, g, p) := \frac{s}{p} g^T \hat{\partial_1 p},
\hat{\partial_2 u}(s, f, g, p) := \frac{s}{R(x)^{-1}} p^T + \tau_p(s)^T \hat{\partial_2 p},
\tau_p(s) := \frac{s}{p} g + \frac{s}{R(x)^{-1}} g,
$$

where $s \in \mathbb{R}$ is a free parameter. The following relations hold:

$$
\|\hat{\partial_1 u}(s, f, g, p) - \partial_1 u(x)\| \leq \|\tau_p(s)\| \|\hat{\partial_1 p} - \partial_1 p\|,
\|\hat{\partial_2 u}(s, f, g, p) - \partial_2 u(x)\| \leq \|\tau_p(s)\| \|\hat{\partial_2 p} - \partial_2 p\|.
$$

A minimizer $s$, and the minimum value of $\|\tau_p(s)\|$ are
\[
\begin{align*}
s_* := \arg \min_s ||\tau_r(s)|| = \frac{(p^T g)(g^T f_k)}{R(s)||f_k||^2},
\end{align*}
\]

\[
||\tau_r(s_*)|| = \sqrt{\frac{1}{R(s)} g^T f_k} \left( \frac{g^T f_k}{R(s)||f_k||^2} \right). \tag{38}
\]

**Proof** The proof is given in Appendix E.

Based on Theorem 4, \( \partial_\mu u \) and \( \partial_\mu u_* \) are approximated by \( \tilde{\partial}_\mu u(s_*) \) and \( \tilde{\partial}_\mu u_*(s_*) \), respectively. As shown in (37), minimizing the norm \( ||\tau_r(s)|| \) by the free parameter \( s \) is employed to reduce the approximation error of \( \tilde{\partial}_\mu u \) and \( \tilde{\partial}_\mu u_* \).

Recall that \( \tilde{\partial}_\mu u(s_*) \) and \( \tilde{\partial}_\mu u_* \) in (36) are the functions of \( f_r, g_r \) and \( p_r \). On the basis of Theorems 3 and 4, the weight \( W_r^0 \) in (18) is approximated by a function of \( f(x_i^0), g(x_i^0), \) and \( p(x_i^0) \) as shown in (27) and explained in (19):

\[
W_r^0 \approx \tilde{W}_r^0 \left( \tilde{f}(x_i^0, \theta), \tilde{g}(x_i^0, \theta), \tilde{p}(x_i^0, \theta) \right) =: W_r^{0}(\theta). \tag{39}
\]

It implies that the system dynamics \( f(x_i^0), g(x_i^0), \) and the controller \( p(x_i^0) \) are necessary for obtaining \( W_r^0 \). The next subsection addresses how to obtain \( f(x_i^0), g(x_i^0), \) and \( p(x_i^0) \).

### 3.5 Expression of the Weight Using the Model Parameter

Although \( f(x_i^0), g(x_i^0), \) and \( p(x_i^0) \) are needed to calculate the weight \( W_r^0 \) in (39), they are still unknown. This subsection regards \( f(x_i^0), g(x_i^0), \) and \( p(x_i^0) \) as the functions \( \tilde{f}(x_i^0, \theta) \) and \( \tilde{g}(x_i^0, \theta) \) in (5) and \( \tilde{p}(x_i^0, \theta) \) in (7), respectively, assuming the existence of \( \tilde{p}(x_i^0, \theta) \). Based on (39), \( W_r^0 \) is approximated by the following function of the model parameter \( \theta \)

\[
W_r^{(0)} \approx \tilde{W}_r^{(0)}(\tilde{f}(x_i^0, \theta), \tilde{g}(x_i^0, \theta), \tilde{p}(x_i^0, \theta)) =: W_r^{(0)}(\theta). \tag{40}
\]

The weight \( W_r^{(0)}(\theta) \) may yield unstable results because of approximation errors. In fact, \( W_r^{(0)}(\theta) \) does not have full rank because \( W_r^{(0)}_m \) in (28) is not positive definite but positive semidefinite in general. Thus, minimizing \( J_\theta(\theta, W_r^{(0)}) \) may not be robust against approximation error, so it may yield an incorrect value for \( \theta \). To realize a robust identification procedure of \( \theta \), a regularization matrix \( \kappa I \) is introduced as explained in (19):

\[
W_r^{(0)} \approx \tilde{W}_r^{(0)}(\theta) + \kappa I =: W_r^{(0)}(\theta, \kappa). \tag{41}
\]

If \( \kappa > 0 \), \( W_r^{(0)}(\theta, \kappa) \) is positive definite with full rank. How to select the regularization parameter \( \kappa \) in \( \mathbb{R} \) is described in Appendix G.

The above derivation demonstrates that \( \tilde{W}_r^{(0)}(\theta, \kappa) \) in (41) requires the unknown \( \theta \), whereas the WLSQ method in (16) requires \( \tilde{W}_r^{(0)}(\theta, \kappa) \) to identify \( \theta \). The next subsection solves such a problem on the interdependence of \( \tilde{W}_r^{(0)}(\theta, \kappa) \) and \( \theta \).

### 3.6 Iterative Learning of the Model Parameter

An iterative scheme is proposed to overcome the problem on the interdependence of \( \tilde{W}_r^{(1:N)}(\theta, \kappa) \) in (41) and \( \theta \) in (16). The initial model parameter \( \theta_r^{(0)} \) is defined as \( \theta_r^{LSQ} \) in (15). Substituting \( \theta_r^{(0)} \) into (41) yields the initial weight \( \tilde{W}_r^{(1:N)}(\kappa, \theta_r^{(0)}) \). To design the weight, we propose the following iterative calculation procedure with respect to the parameter \( \theta_r^{(1)} \):

\[
\theta_r^{(0)} := \arg \min_{\theta} J_\theta(\theta, \tilde{W}_r^{(1:N)}(\theta_r^{(0)}, \kappa)), \quad (j = 0, 1, \ldots), \tag{42}
\]

where \( \theta_r^{(0)} := \theta_r^{LSQ} \). This iteration works using only a preliminarily given data set. We derive properties to obtain \( \theta_r^{(1)} \).

**Theorem 5** (The solution \( \theta_r^{(1)} \) of the model parameter)

(i) The criterion in (42) is decomposed as

\[
J_\theta(\theta, \tilde{W}_r^{(1:N)}(\theta_r^{(0)}, \kappa)) = J_\theta(\theta, W_r^{(1:N)}(\theta_r^{(0)})) + \kappa J_\theta(\theta). \tag{43}
\]

(ii) Supposing that \( \tilde{f}(x, \theta) \) and \( \tilde{g}(x, \theta) \) are linear in \( \theta \) (allowing for nonlinearity in \( x \)), the criteria \( J_\theta(\theta) \) in (15) and \( J_\theta(\theta, W_r^{(1:N)}) \) in (16) are represented as

\[
J_\theta(\theta) := \theta^T H_\theta \theta + 2 \theta^T K_\theta + C_\theta,
\]

\[
J_\theta(\theta, W_r^{(1:N)}) := \theta^T H_\theta (W_r^{(1:N)}) \theta + 2 \theta^T K_\theta (W_r^{(1:N)}) + C_w (W_r^{(1:N)}),
\]

where the symmetric matrix \( H_\theta \in \mathbb{R}^{n_x \times n_x} \), \( C_\theta \in \mathbb{R}^{n_x} \), \( C_w \in \mathbb{R}^{n_x \times n_x} \), \( W_r^{(1:N)} \in \mathbb{R}^{n_x \times n_x} \), and \( C_w \in \mathbb{R}^{n_x} \) are constant. If \( H_\theta \) or \( H_\theta (W_r^{(1:N)}) \) is positive definite, there exists a unique solution \( \theta_r^{(1)} \) in (42) given by

\[
\theta_r^{(1)} = -(H_\theta (W_r^{(1:N)}))^{-1} (K_\theta (W_r^{(1:N)})) + \kappa H_\theta \tag{45}
\]

**Proof** The proof is given in Appendix F.

Based on Theorem 5, the solution \( \theta_r^{(1)} \) is analytically given under some conditions. Even when the item (ii) of Theorem 5 cannot be applied, \( \theta_r^{(1)} \) can be (approximately) found through nonlinear optimization. This completes the derivation of the proposed approach, which is summarized in Algorithm 1.

### 4. Numerical Example

Numerical evaluation in this section compares the performance of the proposed method with that of the LSQ method. Sections 4.1 and 4.2 introduce the plant system and the training data set, respectively. The simulation results under the settings in Section 4.3 are given in Section 4.4.

#### 4.1 Plant System

Let us consider the case where the unknown actual system \( \Sigma \) is not included in the set \( \{ \Sigma(\theta) \in \Omega \} \). The input affine nonlinear system \( \Sigma \) is identified using the linear model \( \tilde{\Sigma}(\theta) \):

\[
\Sigma : x = [Ax + f_\theta(x)] + [B + g_\theta(x)]u,
\]

\[
\tilde{\Sigma}(\theta) : \hat{x} = \hat{\dot{x}} = \hat{\dot{\theta}} + Bu,
\]

where \( \theta \) consists of all the components of \( \hat{\theta} \in \mathbb{R}^{n_\theta} \) and \( \hat{\theta} \in \mathbb{R}^{n_\theta} \). The functions \( f_\theta(x) \in \mathbb{R}^{n_x} \) and \( g_\theta(x) \in \mathbb{R}^{n_x} \) invoke the nonlinearity of \( \Sigma \). The following quadratic cost function is considered:

\[
J(x_0, u) := \int_0^\infty \left[ x^T Q x + \frac{1}{2} R u^2 \right] dr
\]

where \( Q := \text{diag}(3, 2, 4) \) and \( R := 2 \). The suboptimal input \( \tilde{u}(\theta) \) is compared to the theoretical optimal input \( u^* \). The following proposition obtains the exact theoretical expression for \( u^* \), regardless of the nonlinearity of the system.
Proposition 1 (Exactly solvable systems) For the system $\Sigma$ in (46) and the cost function $J(x_0, u)$ in (48), if $f_{\text{all}}(x)$ in (46) and a positive definite matrix $P$ satisfy
\[
 f_{\text{all}}(x) = \frac{1}{2} R_u(x) R_u^{-1}(x) + BR^{-1} g_a(x) + g_a(x) R_a^{-1} B^T P x, \\
 0 = PA + AT R_u^{-1} B^T P + 2Q.
\]
then the value function is given by $V(x) = x^T P x/2$.

Proof Substituting the gradient $p(x) = P x$, (49) and (50) into the HJB equation with respect to (46) and (48) yields $H_{\text{HJB}}(P x, [A x + f_{\text{all}}(x)], [B + g_a(x)]) = 0$. Because $V(x) = x^T P x/2$ is positive definite, it is the value function. \hfill \blacksquare

Based on Proposition 1, substituting $p(x) = P x$ obtained in (50) into (4) yields the theoretical optimal input $u_{\text{opt}}(x)$ as in (4).

In the numerical simulations, two different systems are used to evaluate the performance for various system types. They are parameterized as follows.

Parameter setting 1 (P1):
\[
 A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}, \quad g_a(x) = 0.05 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} x.
\]

Parameter setting 2 (P2):
\[
 A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 3 & 1 \\ 3 & 6 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad g_a(x) = 0.05 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} x.
\]

Here, $f_{\text{all}}(x)$ and $P$ are determined by (49) and (50).

4.2 Training Data Sets

The system is identified via a training data set $(u^{(i)}, x^{(i)}, \hat{x}^{(i)})$ for $i = 1, \ldots , N$ with $N = 5000$. Three types of training data sets are prepared to evaluate the performance for various data sets.

Data type 1 (D1): Random input and random state
\[
 u^{(i)} \sim U(-5, 5), \\
 x^{(i)} \sim U(-10, 10), \\
 \hat{x}^{(i)} = x^{(i)} + x^{(i)} \Delta t \quad \text{(otherwise)}.
\]

Data type 2 (D2): Random input and sequential state
\[
 u^{(i)} \sim U(-5, 5), \\
 x^{(i)} \sim U(-10, 10), \\
 \hat{x}^{(i)} = x^{(i)} + x^{(i)} \Delta t \quad \text{if} \ |x^{(i)}| > 10, \\
 \hat{x}^{(i)} = x^{(i)} \Delta t \quad \text{otherwise}.
\]

Data type 3 (D3): Sequential input and sequential state
\[
 u^{(i)} = 10 \sin(\pi i/180) + u_t, \quad u_t \sim U(-0.5, 0.5), \\
 x^{(i)} \sim U(-10, 10), \\
 \hat{x}^{(i)} = x^{(i)} + x^{(i)} \Delta t \quad \text{if} \ |x^{(i)}| > 10, \\
 \hat{x}^{(i)} = x^{(i)} + x^{(i)} \Delta t \quad \text{otherwise}.
\]

Here, $U(a,b)$ denotes a continuous uniform distribution on $[a,b]^T$. To collect the samples in a bounded state space in the cases of D2 and D3, the state $x^{(i)}$ is reset to a value generated from $U(-10, 10)$ if $|x^{(i)}| > 10$. The sampling time is set to $\Delta t = 0.001$. For all data types, $x^{(i)}$ is collected by substituting $x^{(i)}$ and $u^{(i)}$ into (46).

4.3 Simulation Settings

Using the setting in Sections 4.1 and 4.2, the performance of the proposed unified approach using the WLSQ method (16) is compared with that of the LSQ method (15). The proposed and LSQ methods both identify the system $\Sigma(\theta)$. After the identification step, the optimal controllers in (7) are designed based on the identified system models, and the actual nonlinear system $\Sigma$ is controlled. The above procedure of identification, controller design, and actual system control are repeated 500 times with different random seeds.

The values of the training data set are randomly changed for each trial. The initial state $x_0$ is also set to a random value uniformly distributed on the set $[-10, 10]^T$. In the simulation, the continuous system $\Sigma$ is discretized using a forward difference approximation with a sampling time of $\Delta t = 0.001$. The optimal control of the actual system is performed until $t = 20000\Delta t$. Control performance is defined as the value of the cost function $J(x_0, u)$ in (48), with the integral evaluated from $t = 0$ to $t = 20000\Delta t$ instead of $t = 0$ to $t = \infty$. When any of $||x^{(i)}||_k$ $(k = 1, \ldots , n)$ at $t = 20000\Delta t$ is larger than 0.1, the trial is regarded as non-convergent.

The proposed method is also evaluated both with and without the iteration scheme given in (42). The method without iteration has the model parameter $\theta_{\text{w}}$, set to $\theta_{\text{w}}$, which is obtained in (42). The method with iteration calculates the model parameter as $\theta_{\text{w}} := \theta_{\text{w}}^{(j)}$, if $\theta_{\text{w}}^{(j)}$ is converged, i.e., $||\theta_{\text{w}}^{(j)} - \theta_{\text{w}}^{(j+1)}|| < ||\theta_{\text{w}}^{(0)}|| \times 10^{-3}$ hold. If the iterative scheme (42) in a trial does not converge within 100 iterations, $\theta_{\text{w}}$ is set to $\theta_{\text{w}}^{(0)}$ because such a non-convergent $\theta_{\text{w}}^{(0)}$ may have an inappropriate value. The regularization parameter $\kappa$ is determined in (G. 1).

4.4 Simulation Results

This subsection evaluates the performance with regard to four measures, which are winning rate, average cost, proportion of convergent trials, and proportion of trials such that the resultant costs $J(x_0, \hat{u}_*)$ are close to their minimum values $J(x_0, u_*)$.

Table 1 shows the results for the winning rate of the proposed method compared with the LSQ method. The winning rate indicates the proportion of trials in which the value of $J(x_0, \hat{u}_*)$ for the proposed method was smaller than that for the LSQ method across all trials. We excluded the trials in which both the proposed method and the LSQ method yielded non-convergent results. For all plant and data types, the winning rates were higher than 0.718. This implies that, in most cases, the proposed unified approach performed better than the LSQ method with regard to the winning rate.

The average performance of the proposed method is also evaluated. Figure 1 shows the cost function $J(x_0, \hat{u}_*)$ averaged over all trials. We exclude trials in which the proposed method or the LSQ method yields a non-convergent result. As shown in Fig. 1, the proposed method decreases the average cost by 3.8% to 17.7% in comparison with the LSQ method. These

| Plant, data | [P1.D1] | [P1.D2] | [P1.D3] | [P2.D1] | [P2.D2] | [P2.D3] |
|------------|--------|--------|--------|--------|--------|--------|
| Proposed method without iteration | 0.841 | 0.793 | 0.876 | 0.836 | 0.731 | 0.807 |
| Proposed method with iteration | 0.843 | 0.799 | 0.868 | 0.852 | 0.718 | 0.772 |
results demonstrate that the proposed method performs better than the LSQ method with regard to cost. This means that control performance is improved by identifying the model parameter $\theta$ using the proposed method with an appropriate weight.

The convergence performance is evaluated according to the proportion of the convergence trials across all the trials, as shown in Fig. 2. The proposed method has fewer non-convergent trials than the LSQ method for most data types. This means that the proposed method yields a more stable controller than the LSQ method even when the system model $\hat{\theta}$ cannot completely express the actual system $\Sigma$.

Finally, we evaluate the proportion of trials such that the resultant costs $J(x_0, \hat{u}_*)$ are close to their minimum values $J(x_0, u_*)$, i.e., $J(x_0, \hat{u}_*) - J(x_0, u_*) \leq c_{sd} J(x_0, u_*)$, for a small positive parameter $c_{sd} > 0 \in \mathbb{R}$. In Fig. 3, the horizontal axis represents the parameter $c_{sd}$. The vertical axis represents the proportion of trials that achieve $J(x_0, \hat{u}_*) \leq (1 + c_{sd}) J(x_0, u_*)$. The proposed method yields a high achievement rate for any $c_{sd}$ compared with the LSQ method as shown in Fig. 3.

### 5. Conclusions

In this paper, a unified approach of control and identification was proposed to maximize optimal control performance of unknown input affine nonlinear systems with weak nonlinearity. The proposed approach is based on a new control performance criterion, which unfortunately cannot be calculated directly. We develop the control performance criterion by employing the WLSQ method. A parametric system model is identified based on the WLSQ method. The weight is derived under effective approximations where the approximation error is clearly analyzed. Because the derived weight is a function of the model parameter, they are interdependent. An iteration scheme is introduced to simultaneously calculate the model parameter and the weight. Finally, system identification and controller design are realized using the derived weight. A numerical simulation was performed to demonstrate the effectiveness of the proposed approach.

### References

[1] L. Saleh, P. Chevrel, F. Claveau, J.-F. Lafay, and F. Mars: Shared steering control between a driver and an automation: Stability in the presence of driver behavior uncertainty, *IEEE Trans. on Intelligent Transportation Systems*, Vol. 14, No. 2, pp. 974–983, 2013.

[2] M. Kamal, M. Mukai, J. Murata, and T. Kawabe: Ecological driving based on preceding vehicle prediction using MPC, *Proc. 18th IFAC World Congress*, pp. 3843–3848, 2011.

[3] L. Languang, H. Xuebing, L. Jianqiu, H. Jianfeng, and O. Minggao: A review on the key issues for lithium-ion battery management in electric vehicles, *Journal of Power Sources*, Vol. 226, pp. 272–288, 2013.

[4] K.J. Keensman: *System Identification: An Introduction*, Springer, 2011.

[5] M. Gevers: Identification for control: From the early achievements to the revival of experiment design, *European Journal of Control*, Vol. 11, No. 4-5, pp. 335–352, 2005.

[6] Z. Zang, R.R. Bitmead, and M. Gevers: Iterative weighted least-squares identification and weighted LQG control design, *Automatica*, Vol. 31, No. 11, pp. 1577–1594, 1995.

[7] J.J. Murray, C.J. Cox, G.G. Lendaris, and R. Saeks: Adaptive dynamic programming, *IEEE Trans. Systems, Man, and Cybernetics - Part C*, Vol. 32, No. 2, pp. 140–153, 2002.

[8] D. Vrabie and F. Lewis: Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems, *Neural Networks*, Vol. 22, No. 3, pp. 237–246, 2009.

[9] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin: Iterative feedback tuning: Theory and applications, *IEEE Control Systems Magazine*, Vol. 18, No. 4, pp. 26–41, 1998.

[10] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller: Playing atari with deep reinforcement learning, *NIPS Deep Learning Workshop*, 2013.

[11] M. Okada, H. Fukushima, and T. Sugie: LQ optimal controller design method based on experimental data, *Trans. Institute of
Therefore, calculating the expectation $E$ completes the proof. (23)

Because the components of $M_M$ in (B.1) are bounded, there exists $m_j > 0 \in \mathbb{R}$ such that $m_j \Delta w_{i,n} > M_M$. This yields that there exist $\delta_j > 0 \in \mathbb{R}$ and $m_j > 0 \in \mathbb{R}$ such that

$$\sum_{i=1}^{N} \| M_M^{ij}(W^{ij}(t)) - M_M^{ij} \|_{\Delta}^2 < \delta_j$$

This completes the proof. (B.3)

Therefore, (26) holds. This completes the proof. 

Appendix C Proof of Theorem 2

Based on definitions (24) and (25), the square of each Frobenius norm in (27) is

$$\| M_M^{ij}(W^{ij}(t)) - M_M^{ij} \|^2_{\Delta}$$

Also, (20) and (25) reduces $J_{data}(\theta)$ in (12) to

$$J_{data}(\theta) = \sum_{i=1}^{N} \Delta g_i(\theta)^T M_M^{ij}(W^{ij}(t)) \Delta g_i(\theta) + O(\text{max}\{\| \Delta f_i(\theta) \|, \| \Delta g_i(\theta) \| \}).$$

Therefore, calculating the expectation $E_{W^{ij}(t)}$ of the square difference between $J_W(\theta, W^{ij}(t))$ in (A.1) and $J_{data}(\theta)$ in (A.2) gives (23). This completes the proof.

Appendix B Proof of Corollary 1

Because of (23), $J(W^{ij}(t)) - r_j$ is expressed as a quadratic function of $M_M^{ij}(W^{ij}(t)) - M_M^{ij}$ as follows:

$$M_M \geq 0 \in \mathbb{R}^{4N \times 4N} \text{ s.t. } \Delta J(W^{ij}(t)) - r_j = \frac{1}{2} M_M \Delta w_{i,n} M_M - M_M^{ij}.$$

where $\Delta w_{i,n} \in \mathbb{R}^{4N \times 4N}$ is the vector in which $[M_M^{ij}(W^{ij}(t)) - M_M^{ij}]_{ii}$ for all $i, k$ and $l$ are arrayed. Because $\Delta M_M^{ij}(\theta)$ is bounded, $M_M$ is also bounded. On the other hand, we obtain the form

$$\sum_{i=1}^{N} \| M_M^{ij}(W^{ij}(t)) - M_M^{ij} \|^2_{\Delta} = \frac{1}{2} \sum_{i=1}^{N} \| M_M^{ij} \|^2_{\Delta}.$$
\[
x^T(A^TP_f + P^T) = O_n(||x||^2)^T, \\
x^TA^TP_{f0} = O_n(||x||^2)^T. \\
\]

(D. 5)

Because (D. 5) must be satisfied for all \(x, A^TP_f + P^T = 0\) and \(A^TP_{f0} = 0\) hold. Multiplying these by \(A^TP\) yields \(P_f = -A^TP_0 = \bar{\theta}^{-1}_f p\) and \(P_{f0} = 0\). Thus, because of (D. 4), (34) holds.

Next, transposing the higher order terms \(O_n(||x||^4)\) in the HJB equation (3) yields

\[
\frac{1}{2}x^T(A^TP + PA - P^T) + 2p = -q(x) + O_n(||x||^4), \\
\]

where \(A := \partial_x f(x)|_{x=0} \in \mathbb{R}^{n \times n}\). The partial derivative of (D. 6) with respect to \([g, l]\) with taking \(g = B + O_n(||x||)\) is

\[
x^T(A^TP + PA - P^T) + 2p = -q(x) + O_n(||x||^4), \\
\]

where \(P_{p_1}(x)\) in (D. 4) equals \([\partial_x g, \partial_t g, \partial_x p, \ldots, \partial_t p, \partial_x p]\). Because \(P\) is a symmetric matrix, its partial derivative \(\partial_x P \in \mathbb{R}^{n \times n}\) is also a symmetric matrix. Substituting \(\Pi^k = \partial_x g, P\) and \(\Pi^k = A - (BB^TP)R_0\) into (D. 7) gives the relation (33). Therefore, \(\partial_x g, p(x)\) in (32) equals \(P_{p_1}(x)\) in (D. 4). Because of \(P_{f0} = 0\) and \(\bar{\theta}_f p(x) = P_{p_1}(x)\), (35) holds. This completes the proof.

**Appendix E  Proof of Theorem 4**

We multiply (30) and (D. 2) by \((1 - s)\) and \(s/p^2 \), respectively. Calculating the sum of the results yields the following expression using the free parameter \(s\):

\[
\partial_x u_x(f, s, g) := \frac{s}{p^2}g + \tau_p(s)\partial_f p(f, g), \\
\partial_x u_x(f, s, g) := \frac{s - 1}{R}p + \tau_p(s)\partial_f p(f, g). \\
\]

The difference between (36) and (E. 1) reduces to

\[
\frac{\partial_t u_x(s) - \partial_t u_x} = \tau_p(s)^\top(\partial_f p - \partial_t p), \\
\frac{\partial_t u_x(s) - \partial_t u_x} = \tau_p(s)^\top(\partial_f p - \partial_t p). \\
\]

Taking the norm of (E. 2) yields (37).

Next, \(\|\tau_p(s, f, g, p)\|_2^2\) can be calculated as

\[
\|\tau_p(s, f, g, p)\|_2^2 = \left\| - \frac{f_1}{p^2}g + \frac{1}{R}g \right\|^2 \\
= \left\| \frac{f_1}{p^2}g \right\|^2 + \left\| \frac{1}{R}g \right\|^2 - \left( \frac{f_1}{p^2}g \right)^2. \\
\]

Thus, the minimizer \(s_\star\) and the minimum value of \(\|\tau_p(s, f, g, p)\|_2^2\) are given by (38). This completes the proof.

**Appendix F  Proof of Theorem 5**

Substituting (41) decomposes \(J_H(\theta, \hat{W}^{1,3}(\hat{\theta}^{(1)}_{W} + \kappa))\) as shown in (43). It proves the item (i). In the item (ii), if \(f(x, \theta) \) and \(g(x, \theta)\) are linear in \(\theta\), then \(\hat{e}(\theta)\) is also linear in \(\theta\). Thus, the criteria \(J_{H_{SO}}(\theta)\) in (15) and \(J_W(\theta, \hat{W}^{1,3}(\hat{\theta}^{(1)}_{W} + \kappa))\) in (16) can be represented as quadratic functions of \(\theta\) as (44). Because \(\hat{W}^{(1)}_{W}(\theta)\) is at least positive semidefinite and \(\hat{e}(\theta)\) is linear in \(\theta, \mathcal{H}_{SO}(\hat{W}^{(1,3)}_{W}(\theta))\) and \(\mathcal{H}_{SQ}(\theta)\) are at least positive semidefinite. If \(\mathcal{H}_{SO}\) or \(\mathcal{H}_{SQ}(\hat{W}^{(1,3)}_{W}(\theta))\) is positive definite, the Hessian matrix \(\left(\mathcal{H}_{SO}(\hat{W}^{(1,3)}_{W}(\theta))\right)\) is positive definite and \(J_W(\theta, \hat{W}^{1,3}(\hat{\theta}^{(1)}_{W} + \kappa))\) is strictly convex in \(\theta\). The necessary and sufficient condition for \(\hat{\theta}^{(1)}_{W} + \kappa\) is

\[
\frac{\partial_J}{\partial \hat{\theta}^{(1)}_{W}} J_W(\theta, \hat{W}^{1,3}(\hat{\theta}^{(1)}_{W} + \kappa)) = 0, \\
\]

which gives (45).

**Appendix G  Regularization Parameter**

How to select the regularization parameter \(\kappa\) is discussed. A small \(\kappa\) lets the minimum eigenvalue of \(\hat{W}^{(1)}_{W}\) in (41) small. The identification via the WLSQ method is numerically unstable due to the small eigenvalues, i.e., \(\hat{W}^{(1)}_{W}\) is nearly a low rank matrix. Conversely, the identification with a large \(\kappa\) is numerically stable because of a full rank matrix \(\hat{W}^{(1)}_{W}\). However, the control performance of the proposed method is not so improved compared with that of the LSQ method because \(J_W(\theta, \hat{W}^{1,3}(\hat{\theta}^{(1)}_{W} + \kappa))\) is close to \(J_{SO}(\theta)\) as in (43). We select \(\kappa\) to tune the trade-off between the control performance improvement and the numerical stability of the proposed method. To obtain a reasonable \(\kappa\), for example, we introduce the setting that \(\kappa\) is the maximum value of the eigenvalues of all \(\hat{W}^{(1)}_{W}(\theta)\):

\[
\kappa = \max_{i \leq k} \lambda_i(\hat{W}^{(1)}_{W}(\theta)), \\
\]

where \(\lambda_i(W)\) denotes the \(i\)-th eigenvalue of \(W\). The influence of \(W^{(1)}(\theta)\) and \(\kappa\) in (G. 1) are equivalent in the criterion \(J_W(\theta, \hat{W}^{1,3}(\hat{\theta}^{(1)}_{W} + \kappa))\), it is expected to balance the trade-off between the control performance improvement and numerical stability.

---

**Yuji Ito (Member)**

He received his B.S., M.S., and Ph.D. degrees from Nagoya University, Japan, in 2009, 2011, and 2014, respectively. He joined TOYOTA CENTRAL R&D LABS., INC. in 2011. His research interests include stochastic control. He is a member of IEEE.

**Kenji Fujimoto (Member)**

He received his B.S., M.S., and Ph.D. degrees from Kyoto University, Japan, in 1994, 1996, and 2001, respectively. He is currently a Professor of the Graduate School of Engineering, Kyoto University, Japan. He is a member of JMSE, IEEE, etc.

**Takayoshi Yoshimura**

He received his B.S., M.S., and Ph.D. degrees from Nagoya Institute of Technology, Japan, in 1997, 1999, and 2002, respectively. He is currently with TOYOTA CENTRAL R&D LABS., INC. His research interests include in-vehicle signal processing.

**Yukihiro Tadokoro**

He received his B.S., M.S., and Ph.D. degrees from Nagoya University, Japan, in 2000, 2002, and 2005, respectively. He is currently with TOYOTA CENTRAL R&D LABS., INC. His research interests include nanoscale wireless communication.