IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 16, 2020, accepted July 5, 2020, date of publication July 15, 2020, date of current version July 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3009357

Kernel-Based Hamilton–Jacobi Equations for Data-Driven Optimal and H-Infinity Control

YUJI ITO[1], (Member, IEEE), KENJI FUJIMOTO[2], (Member, IEEE), AND YUKIHIRO TADOKORO[1], (Senior Member, IEEE)

[1]Toyota Central R&D Labs., Inc., Nagakute 480-1192, Japan
[2]Department of Aeronautics and Astronautics, Graduate School of Engineering, Kyoto University, Kyoto 615-8540, Japan

Corresponding author: Yuji Ito (ito-yuji@mosk.tytlabs.co.jp)

This work was partly supported by JSPS KAKENHI Grant Number JP18K04222.

ABSTRACT This paper presents a data-driven method for designing optimal controllers and robust controllers for unknown nonlinear systems. Mathematical models for the realization of the control are difficult to develop owing to a lack of knowledge regarding such systems. The proposed multidisciplinary method, based on optimal control theory and machine learning with kernel functions, facilitates designing appropriate controllers using a data set. Kernel-based system models are useful for representing nonlinear systems. An optimal and an H-infinity controller can be designed by solving Hamilton–Jacobi (HJ) equations, which unfortunately, are difficult to solve owing to the nonlinearity and complexity of the kernel-based models. The objective of this study consists of overcoming two challenges. The first challenge is to derive exact solutions to the HJ equations for a class of kernel-based system models. A key technique in overcoming this challenge is to reduce the HJ equations to easily solvable algebraic matrix equations, from which optimal and H-infinity controllers are designed. The second challenge is to control an unknown system using the obtained controllers, wherein the system is identified as a kernel-based model. Additionally, this study analyzes probabilistic stability of the feedback system with the proposed controllers. Numerical simulations demonstrate control performances of both the derived optimal and H-infinity controllers and stability of the feedback system.

INDEX TERMS Gaussian processes, H-infinity control, optimal control.

I. INTRODUCTION

There exist various unknown nonlinear systems that it is useful to control. Examples of such systems are (semi-)autonomous vehicles that comprise a human driver with unknown nonlinear dynamics [1], [2] and batteries in electric vehicles that should be managed by taking into account their unknown dynamics [3]. It is desirable that such systems are controlled optimally and safely, albeit their dynamics remain partially unknown. This study focuses on controlling such unknown nonlinear systems.

When model-based approaches are used to design controllers for unknown systems, system identification is needed to deduce mathematical models. Precise system modeling is crucial in controller design to realize high control performance without the need for iterative experiments. Various model types have been developed to identify nonlinear systems in fields of system identification and machine learning. Data-driven models using kernel functions are promising because they describe nonlinear dynamics while requiring limited knowledge regarding the dynamics. Such kernel-based models include kernel ridge regression models [4], Gaussian processes (GPs) [5]–[7], and GP-based state-dependent coefficient models [8]. Successful utilization of the GPs can be seen in control systems [9]–[13]. This study focuses on data-driven methods using kernel-based models to design controllers for unknown systems. The data-driven methods indicate that the controllers are designed using data sets of the systems. Although estimate models of the systems are developed from the data sets, true equations describing the systems are not used in the controller design. Such methods are efficient for controlling unknown systems, regarding which only limited information is available.

The associate editor coordinating the review of this manuscript and approving it for publication was Moussa Boukhnifer.

Various methods have been developed to control kernel-based system models. Model predictive control [14]–[19] can design (sub-) optimal controllers to minimize cost functions. Local (sub-) optimal controllers near reference trajectories have been calculated using differential dynamic programming [20] and an iterative linear quadratic regulator [21]. Several promising techniques, such as the Dijkstra algorithm [22] and a gradient-based method [11], have been employed for controller design. Stabilizing controllers have been proposed using sampling-based approaches [23], [24]. Unfortunately, several drawbacks are associated with the use of these methods. First, the controllers are designed via approximations that are sub-optimal for system models. Next, designing the controllers involves high computational costs. Lastly, stability of feedback system models with some controllers is not guaranteed. An underlying cause of these drawbacks is the difficulty in solving nonlinear optimal control problems for the kernel-based models. This difficulty also exists when employing other data-driven approaches based on neural networks [25]–[28].

To overcome these drawbacks, this paper presents a method to design optimal and H_∞ controllers for a class of kernel-based models. We design not only optimal controllers but also H_∞ controllers to ensure robustness against disturbances. The proposed method derives exact solutions to optimal control and H_∞ control problems via an analytical approach. The control problems can be automatically solved without the need for huge computation once the system model is obtained. Moreover, stability of the feedback system model during the application of the designed controllers is automatically guaranteed under certain assumptions. The proposed controllers are data-driven; that is, a true system is identified as a data-driven model, using which the controllers can be designed. The resultant controllers consist of kernel functions with a training data set.

The proposed method focuses on solving Hamilton–Jacobi (HJ) equations [29], which are powerful tools for analyzing optimal and H_∞ control problems. Solving the HJ equations yields optimal and H_∞ controllers. Unfortunately, it is difficult to acquire exact solutions to the HJ equations for kernel-based system models, though approximate solutions can be obtained, e.g., in [23]. To address this difficulty, we find a class of kernel-based models for which exact solutions to the HJ equations can be obtained. A key technique to find such a model class involves reducing the HJ equations to easily solvable algebraic matrix equations. The main originality and novelty of this study involve the implementation of this technique. The algebraic equations consist of free parameters to determine a kernel-based model and controllers. The parameters are selected such that the algebraic equations hold. Satisfying the algebraic equations implies that the original HJ equations are solved. The exact corresponding controllers are subsequently obtained for the kernel-based model. If the kernel-based model is equivalent to a true plant system, the designed controllers stabilize it under certain assumptions. However, there almost always exists a modeling error between the true plant system and its kernel-based model. This study analyzes probabilistic stability of the true system when the proposed controllers are employed.

The remainder of this paper is organized as follows. Mathematical notations are described in Section II. Section III states two main problems associated with control problems that are addressed in this study. The details of the proposed method are described in Sections IV and V. Section IV addresses the first problem, which refers to finding kernel-based models for which exact solutions to optimal and H_∞ control problems can be obtained. Based on the solutions to the first problem, Section V solves the second problem, which concerns the development of a kernel-based system model and (stabilizing) controllers through the use of a data-driven approach. In Section V-B, Algorithm 1 summarizes the method. Section VI demonstrates the effectiveness of the proposed method via numerical simulations. Section VII summarizes the advantages of the proposed method compared to existing methods. Section VIII concludes this paper and describes future work. Contents of this paper have been presented in part at the 2018 American Control Conference [30].

II. NOTATION

The following notations are used in this paper.

- Scalars are denoted by symbols in regular-weight font, e.g., $t \in \mathbb{R}$ and $J \in \mathbb{R}$.
- Vectors and matrices are denoted by bold letters, e.g., $\boldsymbol{v} \in \mathbb{R}^n$ and $\boldsymbol{A} \in \mathbb{R}^{n\times m}$.
- $\boldsymbol{I}_n$: the $n \times n$ identity matrix
- $[\boldsymbol{v}]_i$: the i-th component of a vector $\boldsymbol{v} \in \mathbb{R}^n$
- $[\boldsymbol{A}]_{i,j}$: the component in the i-th row and j-th column of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$
- $[\boldsymbol{A}]_{i,\cdot} \in \mathbb{R}^m$: the i-th row vector of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$
- $[\boldsymbol{A}]_{\cdot,j} \in \mathbb{R}^n$: the j-th column vector of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$
- $\mathrm{vec}(\boldsymbol{A}) := [[\boldsymbol{A}]_{\cdot,1}^{\mathrm{T}}, \ldots, [\boldsymbol{A}]_{\cdot,m}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{nm}$: the vector form of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$
- $\mathrm{vech}(\boldsymbol{A}) := [[\boldsymbol{A}]_{1,1}, [\boldsymbol{A}]_{1,2}, [\boldsymbol{A}]_{2,2}, \ldots, [\boldsymbol{A}]_{1,j}, \ldots, [\boldsymbol{A}]_{j,j}, \ldots, [\boldsymbol{A}]_{1,n}, \ldots, [\boldsymbol{A}]_{n,n}]^{\mathrm{T}} \in \mathbb{R}^{n(n+1)/2}$: the half-vectorization of the upper triangular components of a symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$
- $\mathrm{diag}(\boldsymbol{v}) \in \mathbb{R}^{n\times n}$: the diagonal matrix whose diagonal components are the components of a vector $\boldsymbol{v} \in \mathbb{R}^n$
- $\boldsymbol{A}_{\mathrm{a}} \otimes \boldsymbol{A}_{\mathrm{b}} \in \mathbb{R}^{n_{\mathrm{a}}n_{\mathrm{b}}\times m_{\mathrm{a}}m_{\mathrm{b}}}$: the Kronecker product of matrices $\boldsymbol{A}_{\mathrm{a}} \in \mathbb{R}^{n_{\mathrm{a}}\times m_{\mathrm{a}}}$ and $\boldsymbol{A}_{\mathrm{b}} \in \mathbb{R}^{n_{\mathrm{b}}\times m_{\mathrm{b}}}$, given by

$$\boldsymbol{A}_{\mathrm{a}} \otimes \boldsymbol{A}_{\mathrm{b}} = \begin{bmatrix} [\boldsymbol{A}_{\mathrm{a}}]_{1,1}\boldsymbol{A}_{\mathrm{b}} & \cdots & [\boldsymbol{A}_{\mathrm{a}}]_{1,m_{\mathrm{a}}}\boldsymbol{A}_{\mathrm{b}} \\ \vdots & \ddots & \vdots \\ [\boldsymbol{A}_{\mathrm{a}}]_{n_{\mathrm{a}},1}\boldsymbol{A}_{\mathrm{b}} & \cdots & [\boldsymbol{A}_{\mathrm{a}}]_{n_{\mathrm{a}},m_{\mathrm{a}}}\boldsymbol{A}_{\mathrm{b}} \end{bmatrix} \quad (1)$$

- $\partial_{\boldsymbol{v}}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{v}) \in \mathbb{R}^{n\times m}$: the partial derivative $\partial\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{v})/\partial\boldsymbol{v}$ of a function $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$ with respect to $\boldsymbol{v} \in \mathbb{R}^n$

III. PROBLEM SETTING

This study focuses on H_∞ control and optimal control problems, both of which are introduced in Section III-A.

Section III-B defines kernel-based system models to identify unknown nonlinear systems. On the basis of these preliminaries, the two main problems to be solved in this study are described in Section III-C.

A. NONLINEAR H_∞ CONTROL AND OPTIMAL CONTROL

Consider a nonlinear system affine with respect to its input and disturbance:

$$\mathrm{d}\boldsymbol{x}(t)/\mathrm{d}t = \boldsymbol{f}(\boldsymbol{x}(t)) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{B}_\mathrm{d}\boldsymbol{w}(t), \tag{2}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$, $\boldsymbol{u}(t) \in \mathbb{R}^{n_u}$, and $\boldsymbol{w}(t) \in \mathbb{R}^{n_w}$ are the state, control input, and disturbance, respectively, at the time $t \in \mathbb{R}$. The symbols $\boldsymbol{f} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$, $\boldsymbol{B} \in \mathbb{R}^{n_x \times n_u}$, and $\boldsymbol{B}_\mathrm{d} \in \mathbb{R}^{n_x \times n_w}$ are the drift term, input matrix, and disturbance matrix, respectively. Let us assume the following: $\boldsymbol{f}(\boldsymbol{x})$ is locally Lipschitz; $\boldsymbol{f}(0) = 0$ holds; $\boldsymbol{u}(t)$ is continuous; $\sup_t \|\boldsymbol{w}(t)\| < \infty$ holds; $\boldsymbol{w}(t)$ is continuous and square-integrable; i.e., $\int_0^\infty \|\boldsymbol{w}(t)\|^2 \mathrm{d}t < \infty$. The drift term $\boldsymbol{f}(\boldsymbol{x})$ indicates autonomous dynamics without any control input. The state $\boldsymbol{x}$ drifts according to $\boldsymbol{f}(\boldsymbol{x})$ when the control input and disturbance are not applied to the system.

This study considers a nonlinear H_∞ control problem. Let $\boldsymbol{z}(t)$ be the performance output at t:

$$\boldsymbol{z}(t) := [\boldsymbol{h}(\boldsymbol{x}(t))^\mathrm{T}, (\boldsymbol{R}_\mathrm{sqr}\boldsymbol{u}(t))^\mathrm{T}]^\mathrm{T} \in \mathbb{R}^{n_h+n_u}. \tag{3}$$

Here, $\boldsymbol{h} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_h}$ and $\boldsymbol{R}_\mathrm{sqr} \in \mathbb{R}^{n_u \times n_u}$ are a continuous function and matrix, respectively, both of which are arbitrarily designed under the condition that the state cost function $q(\boldsymbol{x}) := \boldsymbol{h}(\boldsymbol{x})^\mathrm{T}\boldsymbol{h}(\boldsymbol{x})/2$ and input cost matrix $\boldsymbol{R} := \boldsymbol{R}_\mathrm{sqr}^\mathrm{T}\boldsymbol{R}_\mathrm{sqr}$ are positive definite. For an L_2 gain parameter $\gamma \in (0, \infty)$, let us define the cost function:

$$\begin{aligned} &J(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{x}(0)) \\ &:= \int_0^\infty \frac{1}{2}\Big(\|\boldsymbol{z}(t)\|^2 - \gamma^2\|\boldsymbol{w}(t)\|^2\Big)\mathrm{d}t \\ &= \int_0^\infty \Big(q(\boldsymbol{x}(t)) + \frac{1}{2}\boldsymbol{u}(t)^\mathrm{T}\boldsymbol{R}\boldsymbol{u}(t) - \frac{1}{2}\gamma^2\|\boldsymbol{w}(t)\|^2\Big)\mathrm{d}t. \end{aligned} \tag{4}$$

The nonlinear H_∞ control problem considered in this study is to determine an H_∞ state feedback controller that makes the L_2 gain from $\boldsymbol{w}$ to $\boldsymbol{z}$ less than or equal to γ. That is,

$$\Big(\int_0^\infty \|\boldsymbol{z}(t)\|^2\mathrm{d}t\Big)^{\frac{1}{2}} \Big/ \Big(\int_0^\infty \|\boldsymbol{w}(t)\|^2\mathrm{d}t\Big)^{\frac{1}{2}} \le \gamma, \tag{5}$$

for $\boldsymbol{x}(0) = 0$ and any $\boldsymbol{w}$ such that $0 < \int_0^\infty \|\boldsymbol{w}(t)\|^2\mathrm{d}t < \infty$. This inequality (5) indicates that $J(\boldsymbol{u}, \boldsymbol{w}, 0) \le 0$.

Designing an H_∞ controller reduces to solving the Hamilton–Jacobi–Isaacs (HJI) equation [31, Theorem 10.3-1]

$$\begin{aligned} H_\mathrm{HJI}(\boldsymbol{x}) &:= \partial_{\boldsymbol{x}}V(\boldsymbol{x})^\mathrm{T}\boldsymbol{f}(\boldsymbol{x}) \\ &\quad - \frac{1}{2}\partial_{\boldsymbol{x}}V(\boldsymbol{x})^\mathrm{T}\boldsymbol{S}(\gamma)\partial_{\boldsymbol{x}}V(\boldsymbol{x}) + q(\boldsymbol{x}) \\ &= 0, \quad \forall \boldsymbol{x} \in \mathbb{R}^{n_x}, \end{aligned} \tag{6}$$

where

$$\boldsymbol{S}(\gamma) := \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\mathrm{T} - \frac{1}{\gamma^2}\boldsymbol{B}_\mathrm{d}\boldsymbol{B}_\mathrm{d}^\mathrm{T}. \tag{7}$$

A solution $V : \mathbb{R}^{n_x} \to \mathbb{R}$ to the HJI equation (6) is termed the value function that is assumed to be C^1 continuous and positive definite. If such a positive definite solution $V(\boldsymbol{x})$ exists, the corresponding H_∞ controller $\boldsymbol{u}_*(\boldsymbol{x})$ is expressed as

$$\boldsymbol{u}_*(\boldsymbol{x}) := -\boldsymbol{R}^{-1}\boldsymbol{B}^\mathrm{T}\partial_{\boldsymbol{x}}V(\boldsymbol{x}). \tag{8}$$

Unfortunately, the HJI equation is difficult to solve if $\boldsymbol{f}(\boldsymbol{x})$ is nonlinear. This difficulty is tackled in Section III-C.

Next, a nonlinear optimal control problem is introduced. The H_∞ control described above covers nonlinear optimal control. Under the condition that $\boldsymbol{B}_\mathrm{d} = 0$ and $\boldsymbol{w}(t) = 0$ hold for any t in (2), the optimal control problem corresponds to designing a feedback controller that minimizes the cost function:

$$\min_{\boldsymbol{u}} J(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{x}(0))|_{\boldsymbol{w}=0}, \tag{9}$$

over all admissible controllers [31, Theorem 10.1-2]. The optimal feedback controller for the problem (9) is expressed as $\boldsymbol{u}_*(\boldsymbol{x})$ with substitution of $V(\boldsymbol{x}) = V(\boldsymbol{x})|_{\boldsymbol{B}_\mathrm{d}=0}$ in (8). The function $V(\boldsymbol{x})|_{\boldsymbol{B}_\mathrm{d}=0}$ represents a positive definite solution to the Hamilton–Jacobi–Bellman (HJB) equation given by

$$H_\mathrm{HJB}(\boldsymbol{x}) := H_\mathrm{HJI}(\boldsymbol{x})|_{\boldsymbol{B}_\mathrm{d}=0} = 0, \quad \forall \boldsymbol{x} \in \mathbb{R}^{n_x}. \tag{10}$$

Therefore, the HJB equation (10) is a special case of the HJI equation (6). Subsequent sections exclusively consider the HJI equation (6) for handling both optimal and H_∞ controllers.

B. IDENTIFICATION OF DRIFT TERMS USING KERNEL-BASED FUNCTIONS

Let us focus on a partially unknown nonlinear system for which the input matrix $\boldsymbol{B}$ is known but a true drift term $\boldsymbol{f}_\mathrm{tr}(\boldsymbol{x})$ is unknown. For example, such a partially unknown system can arise owing to the interaction between autonomous vehicles and manually operated vehicles [1]. A partially unknown system can be expressed as

$$\mathrm{d}\boldsymbol{x}(t)/\mathrm{d}t = \boldsymbol{f}_\mathrm{tr}(\boldsymbol{x}(t)) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{\omega}(t), \tag{11}$$

where $\boldsymbol{\omega}(t) \in \mathbb{R}^{n_x}$ denotes system noise or disturbance. For controlling the system (11), $\boldsymbol{f}(\boldsymbol{x})$ in (2) corresponds to a mathematical model of the true drift term $\boldsymbol{f}_\mathrm{tr}(\boldsymbol{x})$ in (11). This indicates that (2) is an approximation of (11). The disturbance in (2) could be considered as the difference between the true drift term and its model with noise; i.e., $\boldsymbol{B}_\mathrm{d}\boldsymbol{w}(t) = \boldsymbol{f}_\mathrm{tr}(\boldsymbol{x}(t)) - \boldsymbol{f}(\boldsymbol{x}(t)) + \boldsymbol{\omega}(t)$. If there exists no noise/disturbance ($\boldsymbol{\omega}(t) = 0$), the optimal control problem for (2) with the setting of $\boldsymbol{B}_\mathrm{d} = 0$ can be considered.

The true drift term $\boldsymbol{f}_\mathrm{tr}(\boldsymbol{x})$ can be identified as the drift term model $\boldsymbol{f}(\boldsymbol{x})$ using a given training data set. The training data set consists of D pairs of the states $\boldsymbol{x}_d$ and true drift terms $\boldsymbol{f}_{\mathrm{tr},d}$ that obey

$$\boldsymbol{f}_{\mathrm{tr},d} = \boldsymbol{f}_\mathrm{tr}(\boldsymbol{x}_d) + \boldsymbol{\omega}_d, \quad (d = 1, 2, \ldots, D). \tag{12}$$

Suppose that $\boldsymbol{\omega}_d \in \mathbb{R}^{n_x}$ for each d is independently and identically distributed as a normal distribution with mean zero.

Measuring the values of $\boldsymbol{x}$ and $\mathrm{d}\boldsymbol{x}/\mathrm{d}t - \boldsymbol{Bu}$ in (11) can yield such a data set.

Recall that kernel-based functions have the potential to describe various nonlinear dynamics, as described in Section I. This study focuses on kernel-based drift term models $\boldsymbol{f}(\boldsymbol{x})$ expressed as

$$\boldsymbol{f}(\boldsymbol{x}) := \boldsymbol{C}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}), \tag{13}$$

where

$$\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) := [k(\boldsymbol{x}, \boldsymbol{x}_1), \ldots, k(\boldsymbol{x}, \boldsymbol{x}_D)]^{\mathrm{T}} \in \mathbb{R}^D. \tag{14}$$

Here, the coefficient matrix $\boldsymbol{C}(\boldsymbol{x}) \in \mathbb{R}^{n_x \times D}$ is a function of $\boldsymbol{x}$, and $k(\boldsymbol{x}, \boldsymbol{x}_d) \in \mathbb{R}$ denotes a positive definite kernel function, such as a squared exponential kernel or a polynomial kernel. Suppose that $k(\boldsymbol{x}, \boldsymbol{x}_d)$ is C^2 continuous in $\boldsymbol{x}$.

While practical systems may not be completely described by (13), kernel functions have recently become popular and promising for representing unknown nonlinear systems in fields of system identification and machine learning [32]. Indeed, the kernel-based drift term model $\boldsymbol{f}(\boldsymbol{x})$ in (13) includes several functions, such as kernel ridge regression models [4], GP models [5], and GP-based state-dependent coefficient models [8].

C. MAIN PROBLEMS

This study aims to solve two main problems.

Kernel-based drift terms $\boldsymbol{f}(\boldsymbol{x})$ can potentially represent various nonlinear behaviors in a data-driven manner. Unfortunately, it is difficult to solve the HJI equation (6) for $\boldsymbol{f}(\boldsymbol{x})$ owing to nonlinearities. For addressing this difficulty, the first problem is described as follows.

Problem 1: Find a set $\mathcal{F}$ of kernel-based drift terms $\boldsymbol{f}(\boldsymbol{x})$ in the form of (13). For this set, an exact solution $V(\boldsymbol{x})$ to the HJI equation (6) is to be obtained.

The set $\mathcal{F}$ is a set of functions and is clarified at the end of Section IV. It is challenging to find a new set for which the HJI equation can be solved. In Section IV, Problem 1 is solved under certain assumptions by reducing the HJI equation to an algebraic matrix equation.

After Problem 1 is solved, the true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ is identified as a kernel-based model $\boldsymbol{f}(\boldsymbol{x})$ in the set $\mathcal{F}$ using a training data set. A data-driven H_∞ or optimal controller $\boldsymbol{u}_*(\boldsymbol{x})$ for the model $\boldsymbol{f}(\boldsymbol{x})$ is automatically obtained from (8) because $V(\boldsymbol{x})$ can be obtained. Note that the controller $\boldsymbol{u}_*(\boldsymbol{x})$ may not be optimal for a true system owing to the existence of modeling errors. To avoid unexpected control failures, stability of the true feedback system involving $\boldsymbol{u}_*(\boldsymbol{x})$ should be ensured. Thus, the second problem is stated as follows.

Problem 2: For a given training data set $(\boldsymbol{x}_d, \boldsymbol{f}_{\mathrm{tr},d})_{d=1}^D$, find a best estimate $\boldsymbol{f}(\boldsymbol{x})$ in the set $\mathcal{F}$ to represent the true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$, and design an H_∞ or optimal controller $\boldsymbol{u}_*(\boldsymbol{x})$ that stabilizes the true plant system in a probabilistic sense.

Section V addresses Problem 2 under certain assumptions pertaining to the true system.

IV. SOLUTION TO PROBLEM 1: KERNEL-BASED HJI EQUATIONS

In this section, we propose a method to solve Problem 1 stated in Section III-C. It is difficult to solve the HJI equation directly because the kernel functions $\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ included in $\boldsymbol{f}(\boldsymbol{x})$ increase the equation's complexity. A key technique to overcome this difficulty is to reduce the HJI equation to a solvable algebraic matrix equation. Supposing that $\boldsymbol{f}(\boldsymbol{x})$ and $V(\boldsymbol{x})$ are parametric functions, we attempt to find a symmetric matrix $\boldsymbol{M}_{\mathrm{HJI}}$ and nonlinear function $\boldsymbol{\Phi}(\boldsymbol{x})$ that decompose $H_{\mathrm{HJI}}(\boldsymbol{x})$ in (6):

$$H_{\mathrm{HJI}}(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{M}_{\mathrm{HJI}}\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}), \tag{15}$$

where $\boldsymbol{M}_{\mathrm{HJI}}$ contains free parameters that determine $\boldsymbol{f}(\boldsymbol{x})$ and $V(\boldsymbol{x})$. If such $\boldsymbol{M}_{\mathrm{HJI}}$ and $\boldsymbol{\Phi}(\boldsymbol{x})$ exist, the algebraic matrix equation $\boldsymbol{M}_{\mathrm{HJI}} = 0$ is a sufficient condition for the HJI equation (6) to be satisfied as follows:

$$\boldsymbol{M}_{\mathrm{HJI}} = 0 \Rightarrow \forall \boldsymbol{x} \in \mathbb{R}^{n_x}, \quad H_{\mathrm{HJI}}(\boldsymbol{x}) = 0. \tag{16}$$

The algebraic matrix equation $\boldsymbol{M}_{\mathrm{HJI}} = 0$ is more tractable compared to the original HJI equation (6) because the matrix equation is independent of the kernels and state. Therefore, we find a set of $\boldsymbol{f}(\boldsymbol{x})$ for which the matrix equation $\boldsymbol{M}_{\mathrm{HJI}} = 0$ can be solved.

On the basis of this framework, Section IV-A derives $\boldsymbol{M}_{\mathrm{HJI}}$ and $\boldsymbol{\Phi}(\boldsymbol{x})$ for a set of $\boldsymbol{f}(\boldsymbol{x})$. Section IV-B extends the derived matrix equation $\boldsymbol{M}_{\mathrm{HJI}} = 0$ such that it can be solved in an analytical manner. A set $\mathcal{F}$ of kernel-based drift terms $\boldsymbol{f}(\boldsymbol{x})$ can then be determined as a solution to Problem 1.

A. REDUCING THE HJI EQUATION TO AN ALGEBRAIC MATRIX EQUATION

This subsection describes how the HJI equation (6) is reduced to an algebraic matrix equation, to find $\boldsymbol{M}_{\mathrm{HJI}}$ and $\boldsymbol{\Phi}(\boldsymbol{x})$ in (15). The value function $V(\boldsymbol{x})$ should be a flexible parametric function so that the HJI equation is decomposed, as described in (15). It is important to parameterize $V(\boldsymbol{x})$ by considering the form of the drift term $\boldsymbol{f}(\boldsymbol{x})$ because $V(\boldsymbol{x})$ is related to $\boldsymbol{f}(\boldsymbol{x})$ through the HJI equation. In a manner similar to $\boldsymbol{f}(\boldsymbol{x})$, $V(\boldsymbol{x})$ is parameterized using kernel functions $\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$. Such a parametrization is expressive and used to approximate or design several functions, e.g., value functions [33], [34], Lyapunov functions [35], [36], and controllability and/or observability energy functions [37]. Additionally, sums of basis functions approximate value functions [38], [39]. We try to derive an exact $V(\boldsymbol{x})$ for a set of $\boldsymbol{f}(\boldsymbol{x})$ apart from these approximation methods.

Here, we represent the value function $V(\boldsymbol{x})$ as a kernel-based parametric function. Recall that the HJI equation (6) is a partial differential equation of $V(\boldsymbol{x})$. It is suitable that a parameter included in the partial derivative $\partial_{\boldsymbol{x}} V(\boldsymbol{x})$ is separated from functions of $\boldsymbol{x}$ so that the HJI equation is decomposed, as shown in (15). The value function $V(\boldsymbol{x})$ is assumed to be expressed as the following kernel-based parametric

function:

$$V(\boldsymbol{x}) = [\boldsymbol{\psi}_1(\boldsymbol{x})\tilde{\boldsymbol{p}}_1, \dots, \boldsymbol{\psi}_D(\boldsymbol{x})\tilde{\boldsymbol{p}}_D]\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) \in \mathbb{R}, \tag{17}$$

where $\boldsymbol{\psi}_d(\boldsymbol{x}) \in \mathbb{R}^{1\times n_p}$ and $\tilde{\boldsymbol{p}}_d \in \mathbb{R}^{n_p}$ are a C^2 function of $\boldsymbol{x}$ and a parameter, respectively. The function $\boldsymbol{\psi}_d(\boldsymbol{x})$ is arbitrarily defined such that the following conditions hold:

$$\boldsymbol{\psi}_d(0) = 0, \tag{18}$$
$$k(0, \boldsymbol{x}_d)\partial_{\boldsymbol{x}}\boldsymbol{\psi}_d(0) = 0, \tag{19}$$
$$k(\boldsymbol{x}, \boldsymbol{x}_d) = 0 \Rightarrow \boldsymbol{\psi}_d(\boldsymbol{x}) = 0. \tag{20}$$

The conditions (18) and (19) were applied to ensure $V(0) = 0$ and $\partial_{\boldsymbol{x}}V(0) = 0$. The condition (20) was introduced for the following lemma, and it is not restrictive, because nonzero kernel functions $k(\boldsymbol{x}, \boldsymbol{x}_d) \neq 0$ for any $\boldsymbol{x}$ can be employed; examples include squared-exponential kernels.

Kernel functions are regarded as basis functions to express various types of nonlinear functions. It is reasonable to employ same basis functions for determining the system model $\boldsymbol{f}(\boldsymbol{x})$ and value function $V(\boldsymbol{x})$. Furthermore, exact solutions to the HJ equations can be obtained using the expression in (17). By virtue of this expression, the partial derivative $\partial_{\boldsymbol{x}}V(\boldsymbol{x})$ is linear in $\tilde{\boldsymbol{p}}_d$, as shown in the following lemma.

Lemma 1 (Partial Derivative of the Value Function): The partial derivative $\partial_{\boldsymbol{x}}V(\boldsymbol{x})$ is given by

$$\partial_{\boldsymbol{x}}V(\boldsymbol{x}) = (\tilde{\boldsymbol{p}}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}), \tag{21}$$

where

$$\tilde{\boldsymbol{p}} := [\tilde{\boldsymbol{p}}_1^{\mathrm{T}}, \dots, \tilde{\boldsymbol{p}}_D^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n_pD}, \tag{22}$$

$$\boldsymbol{\Phi}(\boldsymbol{x}) := \begin{bmatrix} \mathrm{vec}(\boldsymbol{\Phi}'_1(\boldsymbol{x})) & & \\ & \ddots & \\ & & \mathrm{vec}(\boldsymbol{\Phi}'_D(\boldsymbol{x})) \end{bmatrix} \in \mathbb{R}^{n_pn_xD\times D}, \tag{23}$$

$$\boldsymbol{\Phi}'_d(\boldsymbol{x}) := \boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d)\boldsymbol{\psi}_d(\boldsymbol{x}) + \partial_{\boldsymbol{x}}\boldsymbol{\psi}_d(\boldsymbol{x}) \in \mathbb{R}^{n_x\times n_p}. \tag{24}$$

The function $\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d) \in \mathbb{R}^{n_x}$ is given by

$$\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d) := \begin{cases} \partial_{\boldsymbol{x}}k(\boldsymbol{x}, \boldsymbol{x}_d)/k(\boldsymbol{x}, \boldsymbol{x}_d) & (k(\boldsymbol{x}, \boldsymbol{x}_d) \neq 0) \\ 0 & (k(\boldsymbol{x}, \boldsymbol{x}_d) = 0). \end{cases} \tag{25}$$

Proof: The proof is given in Section A of Appendix. □

Remark 1: The function $\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d)$ in (25) is expressed in a simple form for specific kernels. For squared-exponential kernels, $k(\boldsymbol{x}, \boldsymbol{x}_d)$ and $\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d)$ are given by

$$k(\boldsymbol{x}, \boldsymbol{x}_d) = \alpha_f \exp\left(\frac{-1}{2}(\boldsymbol{x} - \boldsymbol{x}_d)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\boldsymbol{x} - \boldsymbol{x}_d)\right), \tag{26}$$
$$\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d) = -\boldsymbol{\Gamma}^{-1}(\boldsymbol{x} - \boldsymbol{x}_d). \tag{27}$$

For rational quadratic kernels, $k(\boldsymbol{x}, \boldsymbol{x}_d)$ and $\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d)$ are

$$k(\boldsymbol{x}, \boldsymbol{x}_d) = \alpha_f\left(1 + \frac{1}{2\alpha_b}(\boldsymbol{x} - \boldsymbol{x}_d)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\boldsymbol{x} - \boldsymbol{x}_d)\right)^{-\alpha_b}, \tag{28}$$
$$\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d) = \frac{-2\alpha_b\boldsymbol{\Gamma}^{-1}(\boldsymbol{x} - \boldsymbol{x}_d)}{2\alpha_b + (\boldsymbol{x} - \boldsymbol{x}_d)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\boldsymbol{x} - \boldsymbol{x}_d)}. \tag{29}$$

Here, $\alpha_f > 0 \in \mathbb{R}$, $\boldsymbol{\Gamma} \succ 0 \in \mathbb{R}^{n_x\times n_x}$, and $\alpha_b > 0 \in \mathbb{R}$ are the hyperparameters of the kernels.

Lemma 1 separates the constant parameter $\tilde{\boldsymbol{p}}_d$ from the state-dependent function $\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ in $\partial_{\boldsymbol{x}}V(\boldsymbol{x})$. Such a separation reduces the HJI equation (6) to an algebraic matrix equation, as described in (16).

Theorem 1 (Kernel-Based HJI Equation): For given parameters $\tilde{\boldsymbol{A}} \in \mathbb{R}^{n_x\times n_pn_xD}$, $\tilde{\boldsymbol{p}} \in \mathbb{R}^{n_pD}$ in (22), and $\tilde{\boldsymbol{Q}} = \tilde{\boldsymbol{Q}}^{\mathrm{T}} \succeq 0 \in \mathbb{R}^{n_pn_xD\times n_pn_xD}$, suppose that $V(\boldsymbol{x})$ obeys (17) and that $\boldsymbol{f}(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are given by

$$\boldsymbol{f}(\boldsymbol{x}) = \tilde{\boldsymbol{A}}\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}), \tag{30}$$
$$q(\boldsymbol{x}) = \boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x})^{\mathrm{T}}\tilde{\boldsymbol{Q}}\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}). \tag{31}$$

If the algebraic matrix equation

$$\begin{aligned}\boldsymbol{M}_{\mathrm{HJI}} &= (\tilde{\boldsymbol{p}} \otimes \boldsymbol{I}_{n_x})\tilde{\boldsymbol{A}} + \tilde{\boldsymbol{A}}^{\mathrm{T}}(\tilde{\boldsymbol{p}}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x}) + 2\tilde{\boldsymbol{Q}} \\ &\quad - (\tilde{\boldsymbol{p}} \otimes \boldsymbol{I}_{n_x})\boldsymbol{S}(\gamma)(\tilde{\boldsymbol{p}}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x}) \\ &= 0 \in \mathbb{R}^{n_pn_xD\times n_pn_xD}\end{aligned} \tag{32}$$

holds, then the HJI equation (6) holds as well. Furthermore, $\boldsymbol{f}(0) = 0$ and the local Lipschitz continuity of $\boldsymbol{f}(\boldsymbol{x})$ are satisfied.

Proof: The proof is given in Section B of Appendix. □

Remark 2: Theorem 1 is useful for solving not only the H_∞ control problem but also the optimal control problem introduced in Section III-A. Recall that the optimal controller is obtained by solving the HJB equation (10). If the algebraic matrix equation (32) is satisfied under the condition $\boldsymbol{S}(\gamma) = \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\mathrm{T}}$ with $\boldsymbol{B}_{\mathrm{d}} = 0$, the HJB equation holds.

Remark 3: The HJI equation (6) is reduced to the algebraic matrix equation (32) independent of the kernels $\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ and state $\boldsymbol{x}$. It is easy to deal with such an algebraic equation compared to the HJI equation that is dependent on $\boldsymbol{x}$. If $\boldsymbol{M}_{\mathrm{HJI}} = 0$ holds, a set of $\boldsymbol{f}(\boldsymbol{x})$ obeying (30) corresponds to a solution to Problem 1 under certain assumptions (see Remark 4).

Remark 4: We must satisfy several assumptions to obey $\boldsymbol{f}(\boldsymbol{x})$ in (30), $V(\boldsymbol{x})$ in (17), and $q(\boldsymbol{x})$ in (31). Theorem 1 satisfies all the assumptions for $\boldsymbol{f}(\boldsymbol{x})$, which are $\boldsymbol{f}(0) = 0$ and the local Lipschitz continuity. The positive definiteness of $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$ is also assumed; this is discussed in Remark 8 in the next subsection.

Remark 5: The functions $\boldsymbol{f}(\boldsymbol{x})$, $V(\boldsymbol{x})$, and $q(\boldsymbol{x})$ are associated with the parameters $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{p}}$, and $\tilde{\boldsymbol{Q}}$, respectively. The parameter $\tilde{\boldsymbol{A}}$ can be determined according to the true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ (via identification methods; e.g., the least-squares method). The parameters $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{Q}}$ can be designed such that the matrix equation (32) holds. Some numerical methods can be used to obtain parameters that satisfy the matrix equation approximately owing to it being independent of $\boldsymbol{x}$. For example, the following non-convex program is considered:

$$\min_{\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{Q}}} \|\boldsymbol{M}_{\mathrm{HJI}}\|_{\mathrm{Fro}}^2, \quad \text{s.t.} \quad \tilde{\boldsymbol{p}} \neq 0, \ \tilde{\boldsymbol{Q}} \succeq 0, \tag{33}$$

where $\|\cdot\|_{\mathrm{Fro}}$ denotes the Frobenius norm. The constrained optimization problem (33) can be relaxed by an unconstrained optimization problem using the barrier method [40,

Sections 9.1.1 and 11.2.1] provided that the initial condition satisfies $\tilde{\boldsymbol{Q}} \succ 0$; for example,

$$\min_{\tilde{p},\tilde{\boldsymbol{Q}}} \left(\|\boldsymbol{M}_{\mathrm{HJI}}\|_{\mathrm{Fro}}^2 - \eta_1 \ln|\tilde{p}^2| - \eta_2 \ln\left(\det(\tilde{\boldsymbol{Q}})\right) \right), \quad (34)$$

where η_1 and η_2 are coefficients. As an alternative to such numerical approaches, an analytical approach for solving the matrix equation (32) is presented in the next subsection.

B. ANALYTICAL APPROACH FOR FINDING A SET OF KERNEL-BASED DRIFT TERMS

This subsection extends Theorem 1 such that the derived matrix equation $\boldsymbol{M}_{\mathrm{HJI}} = 0$ can be solved in an analytical manner. Such an extension results in the following theorem yielding one solution to Problem 1.

Theorem 2 (Analytical Kernel-Based HJI Equation): For given parameters $\boldsymbol{A} \in \mathbb{R}^{n_x \times n_x}$, $\boldsymbol{v} := [\boldsymbol{v}_1{}^{\mathrm{T}}, \ldots, \boldsymbol{v}_D{}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n_p D}$ with $\boldsymbol{v}_d \in \mathbb{R}^{n_p}$, $p \in \mathbb{R}$, and $\boldsymbol{Q} = \boldsymbol{Q}^{\mathrm{T}} \succ 0 \in \mathbb{R}^{n_x \times n_x}$, suppose that $\boldsymbol{f}(\boldsymbol{x})$, $V(\boldsymbol{x})$, and $q(\boldsymbol{x})$ are given by

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{v}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}), \quad (35)$$

$$V(\boldsymbol{x}) = p[\boldsymbol{\psi}_1(\boldsymbol{x})\boldsymbol{v}_1, \ldots, \boldsymbol{\psi}_D(\boldsymbol{x})\boldsymbol{v}_D]\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}), \quad (36)$$

$$q(\boldsymbol{x}) = \boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{v} \otimes \boldsymbol{I}_{n_x}) \times \boldsymbol{Q}(\boldsymbol{v}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}). \quad (37)$$

If the algebraic matrix equation

$$\boldsymbol{M}_{\mathrm{HJI}} = (\boldsymbol{v} \otimes \boldsymbol{I}_{n_x})\big(p\boldsymbol{A} + p\boldsymbol{A}^{\mathrm{T}} - p^2\boldsymbol{S}(\gamma) + 2\boldsymbol{Q}\big)(\boldsymbol{v}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x}) = 0 \in \mathbb{R}^{n_p n_x D \times n_p n_x D} \quad (38)$$

holds, then the HJI equation (6) holds as well. Furthermore, $\boldsymbol{f}(0) = 0$ and the local Lipschitz continuity of $\boldsymbol{f}(\boldsymbol{x})$ are satisfied.

Proof: The proof is given in Section C of Appendix. □

Remark 6: If $\boldsymbol{Q} \succ 0$ is set as

$$\boldsymbol{Q} = \frac{-p(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})+p^2\boldsymbol{S}(\gamma)}{2}, \quad (39)$$

both the algebraic matrix equation (38) and HJI equation (6) hold automatically. In this sense, Problem 1 can be solved using a analytical approach. A set of $\boldsymbol{f}(\boldsymbol{x})$ obeying (35) is thus a solution to Problem 1 under certain assumptions (see Remark 8). The drift term $\boldsymbol{f}(\boldsymbol{x})$, state cost $q(\boldsymbol{x})$, and value function $V(\boldsymbol{x})$ are determined from the parameters $\boldsymbol{v}$, $\boldsymbol{A}$, and p. Section V proposes a method for determining the values of the parameters such that $\boldsymbol{f}(\boldsymbol{x})$ corresponds to a best estimate of a true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$, considered in Problem 2.

Remark 7: As discussed in Remark 2 for Theorem 1, Theorem 2 is likewise efficient for solving not only the H_∞ control problem but also the optimal control problem by setting $\boldsymbol{S}(\gamma) = \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\mathrm{T}}$ with $\boldsymbol{B}_{\mathrm{d}} = 0$.

Remark 8: Recall Remark 4 discussing the assumptions for $\boldsymbol{f}(\boldsymbol{x})$, $q(\boldsymbol{x})$, and $V(\boldsymbol{x})$. Theorem 2 satisfies all the assumptions for $\boldsymbol{f}(\boldsymbol{x})$, which are $\boldsymbol{f}(0) = 0$ and the local Lipschitz continuity. The positive definiteness of $q(\boldsymbol{x})$ and $V(\boldsymbol{x})$ is assumed. The state cost function $q(\boldsymbol{x})$ in (37) is positive definite if $\boldsymbol{Q}$ in (39) is positive definite, $(\boldsymbol{v}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) \neq 0$ holds for all $\boldsymbol{x} \in \mathbb{R}^{n_x} \setminus \{0\}$, and $(\boldsymbol{v}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(0)\boldsymbol{k}_{\mathrm{vec}}(0) = 0$ holds. Section V-B discusses how to select the values of the parameters $\boldsymbol{v}$, $\boldsymbol{A}$, and p such that $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are positive definite. Numerical evaluations described in Section VI demonstrate that $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are positive definite at least approximately.

Furthermore, we show that Theorem 2 is consistent with both linear true systems and linear kernels.

Theorem 3: (Kernel-Based HJI Equation for Linear Systems): Suppose that (35)–(37) in Theorem 2 hold. For matrices $\boldsymbol{A}_{\mathrm{tr}} \in \mathbb{R}^{n_x \times n_x}$ and $\boldsymbol{Q}_{\mathrm{tr}} = \boldsymbol{Q}_{\mathrm{tr}}^{\mathrm{T}} \succ 0 \in \mathbb{R}^{n_x \times n_x}$, suppose that the pair $(\boldsymbol{A}_{\mathrm{tr}}, \boldsymbol{B})$ is stabilizable, that there exists a positive definite symmetric matrix $\boldsymbol{P}_{\mathrm{tr}} \succ 0$ satisfying

$$\boldsymbol{P}_{\mathrm{tr}}^{\mathrm{T}}\boldsymbol{A}_{\mathrm{tr}} + \boldsymbol{A}_{\mathrm{tr}}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{tr}} - \boldsymbol{P}_{\mathrm{tr}}^{\mathrm{T}}\boldsymbol{S}(\gamma)\boldsymbol{P}_{\mathrm{tr}} + 2\boldsymbol{Q}_{\mathrm{tr}} = 0, \quad (40)$$

and that the condition

$$\mathrm{rank}([\mathrm{vech}(\boldsymbol{x}_1\boldsymbol{x}_1^{\mathrm{T}}), \ldots, \mathrm{vech}(\boldsymbol{x}_D\boldsymbol{x}_D^{\mathrm{T}})]) = \frac{n_x(n_x+1)}{2} \quad (41)$$

holds, i.e., the matrix $[\mathrm{vech}(\boldsymbol{x}_1\boldsymbol{x}_1^{\mathrm{T}}), \ldots, \mathrm{vech}(\boldsymbol{x}_D\boldsymbol{x}_D^{\mathrm{T}})]$ has full rank. If

$$\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x}) = \boldsymbol{A}_{\mathrm{tr}}\boldsymbol{x}, \quad (42)$$

$$q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_{\mathrm{tr}}\boldsymbol{x}, \quad (43)$$

$$k(\boldsymbol{x}, \boldsymbol{x}_d) = \alpha_f \boldsymbol{x}_d^{\mathrm{T}}\boldsymbol{x}, \quad (44)$$

$$\psi_d(\boldsymbol{x}) = \boldsymbol{x}_d^{\mathrm{T}}\boldsymbol{x}, \quad (45)$$

then there exist parameters p, $\boldsymbol{v}$, $\boldsymbol{A}$, and $\boldsymbol{Q}$ such that the algebraic matrix equation (38) holds, $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are positive definite functions, and $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ holds.

Proof: The proof is given in Section D of Appendix. □

Remark 9: Theorem 3 validates the algebraic matrix equation (38) and positive definiteness of $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$. Therefore, Theorem 2 is consistent with linear systems and linear kernels.

Consequently, $\mathcal{F}$ obtained as a solution to Problem 1 can be defined as a set, all members of which are $\boldsymbol{f}(\boldsymbol{x})$ obeying (35) under the assumption of $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$ being positive definite (refer Remark 8 for discussion pertaining to this assumption). The next section addresses Problem 2 based on the set $\mathcal{F}$.

V. SOLUTION TO PROBLEM 2: IDENTIFICATION AND CONTROLLER DESIGN

For a given training data set $(\boldsymbol{x}_d, \boldsymbol{f}_{\mathrm{tr},d})_{d=1}^D$, we find a best estimate $\boldsymbol{f}(\boldsymbol{x})$ in the set $\mathcal{F}$ to represent the true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ and design a stabilizing H_∞ or optimal controller $\boldsymbol{u}_*(\boldsymbol{x})$. The set $\mathcal{F}$ includes $\boldsymbol{f}(\boldsymbol{x})$ obeying (35), which constitutes the solution to Problem 1. In the following analyses, $\boldsymbol{f}(\boldsymbol{x})$, $V(\boldsymbol{x})$, and $q(\boldsymbol{x})$ obey (35), (36), (37), and (39) that depend on the values of the parameters $\boldsymbol{v}$, $\boldsymbol{A}$, and p. The corresponding controller $\boldsymbol{u}_*(\boldsymbol{x})$ in (8) also depends on these parameters. The corresponding functions are denoted by $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A})$, $V(\boldsymbol{x}; \boldsymbol{v}, p)$, $q(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A}, p)$, and $\boldsymbol{u}_*(\boldsymbol{x}; \boldsymbol{v}, p)$, respectively.

In this section, the true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ in (11) is identified as the kernel-based model $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A})$. The corresponding functions $V(\boldsymbol{x}; \boldsymbol{v}, p)$ and $q(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A}, p)$ are simultaneously determined. Applying the resulting controller $\boldsymbol{u}_*(\boldsymbol{x}; p, \boldsymbol{v})$ to

the true system (11) yields the feedback system. Stability of the feedback system is analyzed because the controller $\boldsymbol{u}_*(\boldsymbol{x};\boldsymbol{\nu},p)$ may not be optimal for the true system.

The idea behind developing the model $\boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A})$ and analyzing the stability is to employ GP regression, which is a successful means to represent system uncertainty. The GP regression is reviewed in Section V-A. Section V-B describes determination of the model parameters in $\boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A})$ using the GP model. In Section V-C, probabilistic stability of the true feedback system is analyzed.

A. DRIFT TERM IDENTIFICATION USING GPs

This subsection reviews GP regression [5] and describes application of GPs to n_x-dimensional $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$. The true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ in (11) is represented by a GP model using the training data set $(\boldsymbol{x}_d,\boldsymbol{f}_{\mathrm{tr},d})_{d=1}^D$ in Section III-B, where

$$\boldsymbol{X} := [\boldsymbol{x}_1,\boldsymbol{x}_2,\dots,\boldsymbol{x}_D]^{\mathrm{T}} \in \mathbb{R}^{D\times n_x}, \tag{46}$$

$$\boldsymbol{F}_{\mathrm{tr}} := [\boldsymbol{f}_{\mathrm{tr},1},\boldsymbol{f}_{\mathrm{tr},2},\dots,\boldsymbol{f}_{\mathrm{tr},D}]^{\mathrm{T}} \in \mathbb{R}^{D\times n_x}. \tag{47}$$

Let us assume that each component of $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ obeys an identical GP independently, and that the covariance of the noise $\boldsymbol{\omega}_d$ in (12) is equal to $\alpha_n \boldsymbol{I}_{n_x}$. Then, the following relation holds:

$$\begin{bmatrix}[\boldsymbol{F}_{\mathrm{tr}}]_{\cdot,s}\\ [\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})]_s\end{bmatrix} \sim \mathcal{N}\left(0,\begin{bmatrix}\boldsymbol{K}_{\mathrm{mat}} & \boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})\\ \boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^{\mathrm{T}} & k(\boldsymbol{x},\boldsymbol{x})\end{bmatrix}\right), \tag{48}$$

$$[\boldsymbol{K}_{\mathrm{mat}}]_{d,d'} := k(\boldsymbol{x}_d,\boldsymbol{x}_{d'}) + \delta_{d,d'}\alpha_n, \tag{49}$$

where $\alpha_n > 0 \in \mathbb{R}$ is the hyperparameter and $\delta_{d,d'}$ is the Kronecker delta. The symbol $\boldsymbol{K}_{\mathrm{mat}} \in \mathbb{R}^{D\times D}$ denotes an array of the kernel functions as defined in (49), and $\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) \in \mathbb{R}^D$ was defined in (14). Let the hyperparameter vector $\boldsymbol{\theta} \in \mathbb{S}_{\boldsymbol{\theta}} \subset \mathbb{R}^{n_{\boldsymbol{\theta}}}$ denote a $n_{\boldsymbol{\theta}}$-dimensional vector in a set $\mathbb{S}_{\boldsymbol{\theta}}$. The components of $\boldsymbol{\theta}$ consist of α_n and all hyperparameters included within the kernel functions. The hyperparameter vector $\boldsymbol{\theta}$ is trained to maximize the following log-likelihood function:

$$\begin{aligned}\boldsymbol{\theta}_* &\in \operatorname*{arg\,max}_{\boldsymbol{\theta}\in\mathbb{S}_{\boldsymbol{\theta}}} \ln \Pr(\boldsymbol{F}_{\mathrm{tr}}|\boldsymbol{X},\boldsymbol{\theta})\\ &= \operatorname*{arg\,max}_{\boldsymbol{\theta}\in\mathbb{S}_{\boldsymbol{\theta}}} \ln \prod_{s=1}^{n_x} \Pr([\boldsymbol{F}_{\mathrm{tr}}]_{\cdot,s}|\boldsymbol{X},\boldsymbol{\theta}),\end{aligned} \tag{50}$$

where $\boldsymbol{\theta}_*$ is an optimal hyperparameter vector that maximizes the log-likelihood function. The problem (50) is solved to obtain $\boldsymbol{\theta}_*$ (in an local optimal sense) using optimization methods, such as the conjugate gradient method [41]. Finally, the GP model obeys a normal distribution, the mean $\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x}) \in \mathbb{R}^{n_x}$ and covariance $\mathrm{diag}(\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x}))^2$ with $\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x}) \in \mathbb{R}^{n_x}$ of which are defined by

$$\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x}) := \boldsymbol{F}_{\mathrm{tr}}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{mat}}^{-1}\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}), \tag{51}$$

$$[\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x})]_s := \sqrt{k(\boldsymbol{x},\boldsymbol{x}) - \boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{K}_{\mathrm{mat}}^{-1}\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})}. \tag{52}$$

B. DETERMINING THE PARAMETERS OF THE KERNEL-BASED MODEL

The true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ is identified as the kernel-based model $\boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A})$. Here, a difficulty arises in the facts that the parameters $\boldsymbol{\nu} \in \mathbb{R}^{n_p D}$ and $\boldsymbol{A} \in \mathbb{R}^{n_x\times n_x}$ have many dimensions, and that the kernel functions $\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ include hyperparameters. Such a multitude of parameters may lead to model overfitting.

To address this concern, we employ a two-stage identification method to determine the parameters $\boldsymbol{\nu}$, $\boldsymbol{A}$, and p along with the hyperparameters of the kernel functions $\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ similarly to that in [8]. Recall that the kernel-based model $\boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A})$ has a form similar to the GP mean model $\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x})$ in (51), namely, linear combinations of kernels. We first derive the GP mean model $\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x})$ using the training data set, as explained in Section V-A. The hyperparameters of the kernel functions $\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ are then obtained. In the second stage, the kernel-based model $\boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A})$ is determined such that the size of the following difference $\Delta\boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A})$ between the GP mean and kernel-based models is reduced:

$$\Delta\boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A}) := \boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A}). \tag{53}$$

It is not straightforward to determine the values of the parameters $\boldsymbol{\nu}$, $\boldsymbol{A}$, and p because the state cost function $q(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A},p)$ and value function $V(\boldsymbol{x};\boldsymbol{\nu},p)$ also depend on the parameters in the algebraic matrix equation (38). We propose a method to obtain all of the parameters simultaneously, which yields the kernel-based model $\boldsymbol{f}(\boldsymbol{x})$, state cost function $q(\boldsymbol{x};\boldsymbol{\nu},\boldsymbol{A},p)$, and value function $V(\boldsymbol{x};\boldsymbol{\nu},p)$.

To obtain optimal parameters $(\boldsymbol{\nu}_*,\boldsymbol{A}_*,p_*)$ that minimize the objective function $g(\boldsymbol{\nu},\boldsymbol{A},p)$, the proposed method is formulated as follows:

$$(\boldsymbol{\nu}_*,\boldsymbol{A}_*,p_*) \in \operatorname*{arg\,min}_{\boldsymbol{\nu},\boldsymbol{A},p} g(\boldsymbol{\nu},\boldsymbol{A},p), \tag{54}$$

$$\begin{aligned}g(\boldsymbol{\nu},\boldsymbol{A},p) := \sum_{d=1}^{\acute{D}} \Big(&\|\Delta\boldsymbol{f}(\acute{\boldsymbol{x}}_d;\boldsymbol{\nu},\boldsymbol{A})\|^2\\ &+ \eta_V\|V(\acute{\boldsymbol{x}}_d;\boldsymbol{\nu},p) - V_{\mathrm{ref}}(\acute{\boldsymbol{x}}_d)\|^2\Big)\\ &- \eta_Q \ln(\det(-p\boldsymbol{A} - p\boldsymbol{A}^{\mathrm{T}} + p^2\boldsymbol{S}(\gamma)))\\ &+ \eta_{\mathrm{reg}}\big(\|\boldsymbol{\nu}\|^2 + \|\boldsymbol{A}\|_{\mathrm{Fro}}^2 + p^2\big),\end{aligned} \tag{55}$$

where $\acute{\boldsymbol{x}}_d \in \mathbb{R}^{n_x}$ $(d = 1,\dots,\acute{D})$ are predefined states. The symbols $\eta_V \geq 0 \in \mathbb{R}$, $\eta_Q \geq 0 \in \mathbb{R}$, and $\eta_{\mathrm{reg}} > 0 \in \mathbb{R}$ are coefficients. The operator $\sum_{d=1}^{\acute{D}}(\cdots)$ in (55) is employed instead of the integration $\int(\cdots)\mathrm{d}\boldsymbol{x}$ over a state space. Details concerning the objective function $g(\boldsymbol{\nu},\boldsymbol{A},p)$ are explained as follows. The term $\|\Delta\boldsymbol{f}(\acute{\boldsymbol{x}}_d;\boldsymbol{\nu},\boldsymbol{A})\|^2$ is introduced to reduce the difference between the GP mean and kernel-based models, as discussed above. The term $\eta_V\|V(\acute{\boldsymbol{x}}_d;\boldsymbol{\nu},p) - V_{\mathrm{ref}}(\acute{\boldsymbol{x}}_d)\|^2$ helps satisfy the positive definiteness of $V(\boldsymbol{x};\boldsymbol{\nu},p)$ by defining a positive definite function $V_{\mathrm{ref}}(\boldsymbol{x})$, whereas this term is not necessary for the positive definiteness, as demonstrated in Section VI. The term $\eta_Q\ln(\det(\cdots))$ is the barrier function employed based on the barrier method [40, Sections 9.1.1 and 11.2.1]. This function constrains the parameters so that the positive definiteness of $\boldsymbol{Q}$ in (39) is satisfied. The state cost function $q(\boldsymbol{x})$ in (37) becomes positive definite if $\boldsymbol{Q}$ is positive definite and $(\boldsymbol{\nu}^{\mathrm{T}} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) \neq 0$ holds for

all $\boldsymbol{x} \in \mathbb{R}^{n_x} \setminus \{0\}$. The last term $\eta_{\mathrm{reg}}(\cdots)$ regularizes the size of the parameters.

The above problem (54) must optimize the many parameters $\boldsymbol{v} \in \mathbb{R}^{n_p D}$, $\boldsymbol{A} \in \mathbb{R}^{n_x \times n_x}$, and $p \in \mathbb{R}$. This problem is difficult to solve owing to its high computational complexity. We derive the following property to reduce the complexity.

Proposition 1 (Explicit Optimal Parameter): Suppose that (35), (36), (37), and (39) hold so that $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A})$ is included in the set $\mathcal{F}$. The optimal $\boldsymbol{v}$ in (54) for each $\boldsymbol{A}$ and p is explicitly given by the following function $\overline{\boldsymbol{v}}$ of $\boldsymbol{A}$ and p:

$$\overline{\boldsymbol{v}}(\boldsymbol{A}, p) = (\boldsymbol{Y}(\boldsymbol{A}, p)^{\mathrm{T}} \boldsymbol{Y}(\boldsymbol{A}, p) + \eta_{\mathrm{reg}} \boldsymbol{I}_{n_p D})^{-1} \boldsymbol{Y}(\boldsymbol{A}, p)^{\mathrm{T}} \boldsymbol{y}, \quad (56)$$

where $\boldsymbol{Y}(\boldsymbol{A}, p) \in \mathbb{R}^{\acute{D}(n_x+1) \times n_p D}$ and $\boldsymbol{y} \in \mathbb{R}^{\acute{D}(n_x+1)}$ are defined as

$$\boldsymbol{Y}(\boldsymbol{A}, p) := \begin{bmatrix} \boldsymbol{A}k(\acute{\boldsymbol{x}}_1, \boldsymbol{x}_1)\boldsymbol{\Phi}_1'(\acute{\boldsymbol{x}}_1) & \cdots & \boldsymbol{A}k(\acute{\boldsymbol{x}}_1, \boldsymbol{x}_D)\boldsymbol{\Phi}_D'(\acute{\boldsymbol{x}}_1) \\ \vdots & & \vdots \\ \boldsymbol{A}k(\acute{\boldsymbol{x}}_{\acute{D}}, \boldsymbol{x}_1)\boldsymbol{\Phi}_1'(\acute{\boldsymbol{x}}_{\acute{D}}) & \cdots & \boldsymbol{A}k(\acute{\boldsymbol{x}}_{\acute{D}}, \boldsymbol{x}_D)\boldsymbol{\Phi}_D'(\acute{\boldsymbol{x}}_{\acute{D}}) \\ \eta_V pk(\acute{\boldsymbol{x}}_1, \boldsymbol{x}_1)\boldsymbol{\psi}_1(\acute{\boldsymbol{x}}_1) & \cdots & \eta_V pk(\acute{\boldsymbol{x}}_1, \boldsymbol{x}_D)\boldsymbol{\psi}_D(\acute{\boldsymbol{x}}_1) \\ \vdots & & \vdots \\ \eta_V pk(\acute{\boldsymbol{x}}_{\acute{D}}, \boldsymbol{x}_1)\boldsymbol{\psi}_1(\acute{\boldsymbol{x}}_{\acute{D}}) & \cdots & \eta_V pk(\acute{\boldsymbol{x}}_{\acute{D}}, \boldsymbol{x}_D)\boldsymbol{\psi}_D(\acute{\boldsymbol{x}}_{\acute{D}}) \end{bmatrix}, \quad (57)$$

$$\boldsymbol{y} := \begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{f}}(\acute{\boldsymbol{x}}_1) \\ \vdots \\ \boldsymbol{\mu}_{\boldsymbol{f}}(\acute{\boldsymbol{x}}_{\acute{D}}) \\ \eta_V V_{\mathrm{ref}}(\acute{\boldsymbol{x}}_1) \\ \vdots \\ \eta_V V_{\mathrm{ref}}(\acute{\boldsymbol{x}}_{\acute{D}}) \end{bmatrix}. \quad (58)$$

Proof: The proof is given in Section E of Appendix. □

By virtue of the above proposition, optimal solutions $(\boldsymbol{A}_*, p_*)$ can be given as follows:

$$(\boldsymbol{A}_*, p_*) \in \arg\min_{\boldsymbol{A}, p} g(\overline{\boldsymbol{v}}(\boldsymbol{A}, p), \boldsymbol{A}, p), \quad (59)$$

where $\boldsymbol{A}$ and p become the only decision variables. The corresponding optimal $\boldsymbol{v}_*$ is then given by the following explicit function:

$$\boldsymbol{v}_* = \overline{\boldsymbol{v}}(\boldsymbol{A}_*, p_*). \quad (60)$$

Gradient methods, such as the quasi-Newton and conjugate gradient methods [41], can be used to obtain $\boldsymbol{A}_*$ and p_* in a local optimal sense because the problem (59) concerns nonconvex optimization.

The proposed method is summarized in Algorithm 1. The best drift term model $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}_*, \boldsymbol{A}_*)$ in the set $\mathcal{F}$ and the corresponding H_∞ or optimal controller $\boldsymbol{u}_*(\boldsymbol{x}; \boldsymbol{v}_*, p_*)$ are obtained simultaneously using the training data set.

Algorithm 1 Simultaneous Derivation of the Drift Term Model and H_∞ or Optimal Controller

Input: the L_2 gain parameter γ and training data set $(\boldsymbol{x}_d, \boldsymbol{f}_{\mathrm{tr},d})_{d=1}^{D}$

Output: the drift term model $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}_*, \boldsymbol{A}_*)$ and H_∞ or optimal controller $\boldsymbol{u}_*(\boldsymbol{x}; \boldsymbol{v}_*, p_*)$

1: Define the function $\boldsymbol{\psi}_d(\boldsymbol{x})$ under the conditions (18)–(20)
2: Develop a GP model $\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x})$ using the training data set as shown in Section V-A
3: Determine the parameters $\boldsymbol{A}_*$ and p_* by solving (59)
4: Determine the parameter $\boldsymbol{v}_*$ in (60)
5: Calculate $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}_*, \boldsymbol{A}_*)$, $q(\boldsymbol{x}; \boldsymbol{v}_*, \boldsymbol{A}_*, p_*)$, and $V(\boldsymbol{x}; \boldsymbol{v}_*, p_*)$ in (35), (36), and (37) under the definition (39)
6: Calculate $\boldsymbol{u}_*(\boldsymbol{x}; \boldsymbol{v}_*, p_*)$ by substituting $V(\boldsymbol{x}; \boldsymbol{v}_*, p_*)$ into (8)

C. PROBABILISTIC STABILITY OF THE TRUE FEEDBACK SYSTEMS

This subsection analyzes stability of true feedback systems when the proposed controller $\boldsymbol{u}_*(\boldsymbol{x}; \boldsymbol{v}, p)$ is employed. The GP model in Section V-A represents the uncertainty of a true system as the standard deviation $\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x})$. The following assumption is introduced to analyze the stability on a given bounded set $\mathbb{X} \subset \mathbb{R}^{n_x}$. Let $\beta > 0 \in \mathbb{R}$ be a given constant and $\mathbb{F}_{\mathrm{tr}}(\boldsymbol{x}; \beta) \subseteq \mathbb{R}^{n_x}$ be the following set:

$$\mathbb{F}_{\mathrm{tr}}(\boldsymbol{x}; \beta) := \{\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x}) + \mathrm{diag}(\beta_1, \ldots, \beta_{n_x})\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x}) \mid \forall s,\ |\beta_s| \le \beta\}. \quad (61)$$

Assumption 1 (Inclusion of the True Drift Term in GPs): For the given $\beta > 0 \in \mathbb{R}$, there exists a constant $\delta \in [0, 1)$ such that the following relation holds with probability $1 - \delta$:

$$\forall \boldsymbol{x} \in \mathbb{X}, \quad \boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x}) \in \mathbb{F}_{\mathrm{tr}}(\boldsymbol{x}; \beta). \quad (62)$$

Assumption 1 implies that the true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ is included in the bounded model set $\mathbb{F}_{\mathrm{tr}}(\boldsymbol{x}; \beta)$ with a probability, albeit it may be difficult to check this property of Assumption 1 for a system in practice. Note that a large value of β yields a high probability $1 - \delta$ because the size of $\mathbb{F}_{\mathrm{tr}}(\boldsymbol{x}; \beta)$ with nonzero $\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x})$ is monotonically increasing in β; i.e., $\mathbb{F}_{\mathrm{tr}}(\boldsymbol{x}; \beta) \subset \mathbb{F}_{\mathrm{tr}}(\boldsymbol{x}; \beta')$ for any $0 < \beta < \beta'$. The value of $1 - \delta$ depends on the smoothness (which means a reproducing kernel Hilbert space norm) of the true system (refer [42] for details).

Assumption 1 describes the relation between the GP mean model $\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x})$ and true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$. Thus, combining Assumption 1 with the difference $\Delta\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A})$, defined in (53), provides a characterization of the difference between the true system $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ and kernel-based model $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A})$. This characterization derives a property for probabilistic stability of the true feedback system. This study focuses on the Lyapunov stability theory. Because the Lyapunov inequality for the true feedback system cannot be evaluated, we employ another inequality.

Theorem 4 (Probabilistic Stability of the True Feedback System): Suppose that (35), (36), (37), and (39) hold so that $\boldsymbol{f}(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{A})$ is included in the set $\mathcal{F}$. Suppose that Assumption 1 holds, that the true drift term $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$ in (11) is C^1 continuous, and that the state $\boldsymbol{x}(t)$ obeys the true feedback system

in (11) with the controller $\boldsymbol{u}_*(\boldsymbol{x};\boldsymbol{v},p)$ *applied and zero noise* $\boldsymbol{\omega}(t)=0$. *For any set* $\tilde{\mathbb{X}}\subseteq\mathbb{X}$, *the following relation holds with a probability of at least* $1-\delta$:

$$W(\boldsymbol{x})<0,\quad \forall \boldsymbol{x}\in\tilde{\mathbb{X}}\ \Rightarrow\ \frac{\mathrm{d}V}{\mathrm{d}t}(\boldsymbol{x}(t);\boldsymbol{v},p)<0,\ \forall \boldsymbol{x}(t)\in\tilde{\mathbb{X}}, \tag{63}$$

where the function $W:\mathbb{R}^{n_x}\to\mathbb{R}$ *is given as follows*:

$$W(\boldsymbol{x}):=p\boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v})^{\mathrm{T}}\Delta\boldsymbol{f}(\boldsymbol{x};\boldsymbol{v},\boldsymbol{A})+\beta p\sum_{s=1}^{n_x}\left|[\boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v})]_s[\boldsymbol{\sigma}_f(\boldsymbol{x})]_s\right| - \boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v})^{\mathrm{T}}\left(\frac{p^2}{2}\left(\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\mathrm{T}}+\frac{1}{\gamma^2}\boldsymbol{B}_{\mathrm{d}}\boldsymbol{B}_{\mathrm{d}}^{\mathrm{T}}\right)+\boldsymbol{Q}\right)\boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v}), \tag{64}$$

and

$$\boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v}):=(\boldsymbol{v}^{\mathrm{T}}\otimes\boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}). \tag{65}$$

Proof: The proof is given in Section F of Appendix. □

Remark 10: Theorem 4 indicates that the Lyapunov inequality for the true feedback system can be estimated using the proposed function $W(\boldsymbol{x})$ *in a probabilistic sense. Such a probabilistic approach is based on [42]. If there exists a set* $\tilde{\mathbb{X}}$ *wherein* $W(\boldsymbol{x})<0$ *holds and* $V(\boldsymbol{x};\boldsymbol{v},p)$ *is positive definite, a region of attraction and small invariant set near the origin can be estimated [43]. For the stability notion in this study, any state in the region of attraction arrives in the invariant set asymptotically. Section VI-D demonstrates evaluation of the proposed inequality* $W(\boldsymbol{x})<0$ *along with estimation of a region of attraction and small invariant set.*

Remark 11: The function $W(\boldsymbol{x})$ *consists of first- and second-order terms with respect to* $\|\boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v})\|$. *If* $\boldsymbol{Q}$ *is positive definite, the second-order terms are negative for* $\|\boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v})\|\neq 0$. *Thus, it is expected that* $W(\boldsymbol{x})<0$ *holds at least for states* $\boldsymbol{x}$ *for which* $\|\boldsymbol{\zeta}(\boldsymbol{x};\boldsymbol{v})\|$ *is sufficiently large.*

Remark 12: Because of Assumption 1, a large value of β *ensures a high probability* $1-\delta$ *for the stability. Reducing the value of* $\|\Delta\boldsymbol{f}(\boldsymbol{x};\boldsymbol{v},\boldsymbol{A})\|$ *increases an upper bound of admissible* β *to satisfy the inequality* $W(\boldsymbol{x})<0$. *In this sense, determining the values of the parameters in accordance with (54) is appropriate to enhance the probability* $1-\delta$.

In Section V, we have found the best drift term model $\boldsymbol{f}(\boldsymbol{x};\boldsymbol{v}_*,\boldsymbol{A}_*)$ in the set $\mathcal{F}$ and the corresponding H_∞ or optimal controller $\boldsymbol{u}_*(\boldsymbol{x};\boldsymbol{v}_*,p_*)$ along with the analysis of the probabilistic stability, as summarized in Algorithm 1. Recall that when solving the optimal control problem, optimality of the proposed controller can be guaranteed under technical assumptions, because an exact solution $V(\boldsymbol{x})$ to the HJB equation (10) can be obtained. While this optimality refers to control of the model $\boldsymbol{f}(\boldsymbol{x};\boldsymbol{v}_*,\boldsymbol{A}_*)$ instead of the true system $\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x})$, it also corresponds to a general limitation due to treating unknown systems. The proposed H_∞ controller focuses on robustness to the error between the model and true system.

VI. NUMERICAL EXAMPLES

In this section, the utility of the proposed method is demonstrated by performing numerical simulations. Section VI-A introduces the settings of the simulation. At the beginning of subsequent Sections VI-B and VI-C, we describe how to evaluate the effectiveness of the proposed method associated with performance measures. Section VI-B describes results obtained by using the proposed optimal controller. The robustness of the proposed H_∞ controller is evaluated in Section VI-C. Stability analysis of the true feedback system is described in Section VI-D. Section VI-E describes a practical implementation of the proposed method on an inverted pendulum subjected to a nonlinear torque. In addition, in Sections VI-B, VI-C, and VI-E, the proposed method is compared with other standard methods and the effect of the randomness of the simulation is evaluated.

A. PLANT SYSTEM AND SIMULATION SETTINGS

Let us consider the true system (11) with the drift term:

$$\boldsymbol{f}_{\mathrm{tr}}(\boldsymbol{x}):=\begin{bmatrix}-[\boldsymbol{x}]_1+0.5[\boldsymbol{x}]_2\\ 0.5[\boldsymbol{x}]_2+0.2[\boldsymbol{x}]_1[\boldsymbol{x}]_2+0.3([\boldsymbol{x}]_2)^3\end{bmatrix}. \tag{66}$$

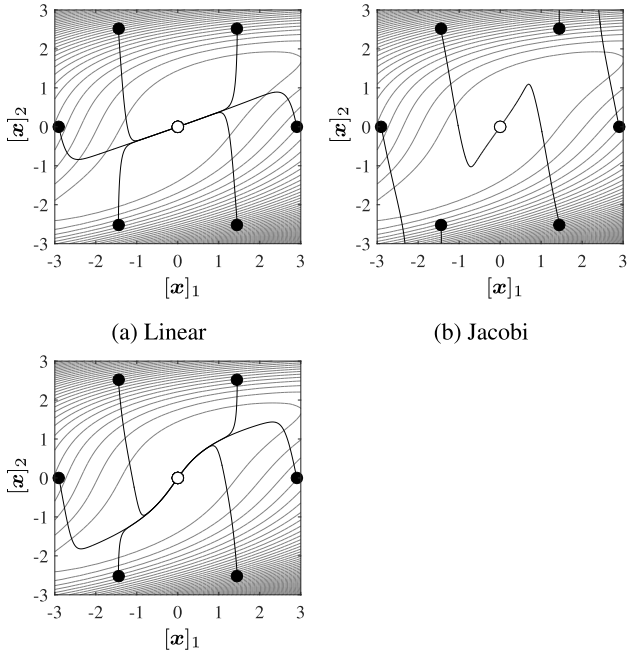
The input matrix is given by $\boldsymbol{B}:=[0,1]^{\mathrm{T}}$. The equation (66) is not used for controller design but only serves as the ground truth to evaluate the proposed method. The parameter in the performance output $\boldsymbol{z}(t)$ is set to $\boldsymbol{R}_{\mathrm{sqr}}=1$, and thus, $\boldsymbol{R}=\boldsymbol{R}_{\mathrm{sqr}}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{sqr}}=1$ holds. The training data set $(\boldsymbol{x}_d,\boldsymbol{f}_{\mathrm{tr},d})_{d=1}^{D}$ with $D=49$ is given as follows. The states $\boldsymbol{x}_d$ within the data set are sampled at regular intervals on $[-3,3]\times[-3,3]$. The drift terms $\boldsymbol{f}_{\mathrm{tr},d}$ within the data set are given by (12), wherein each noise $\boldsymbol{\omega}_d$ independently obeys the normal distribution with mean zero and covariance $0.04\boldsymbol{I}_2$. The L_2 gain parameter is set to $\gamma=3$.

Settings of the proposed method with Algorithm 1 is described as follows. In Line 1 of Algorithm 1, we define $\boldsymbol{\psi}_d(\boldsymbol{x}):=[([\boldsymbol{x}]_1)^2,([\boldsymbol{x}]_2)^2,[\boldsymbol{x}]_1[\boldsymbol{x}]_2]$. In Line 2, the GP modeling is implemented using the GPML package [44]. The squared-exponential kernel defined in (26) is applied, where $\boldsymbol{\Gamma}$ is a diagonal matrix. To optimize the hyperparameter vector $\boldsymbol{\theta}:=(\boldsymbol{\Gamma},\alpha_f,\alpha_n)$, we define $\tilde{\boldsymbol{\theta}}:=(1/2)[\ln[\boldsymbol{\Gamma}]_{1,1},\ln[\boldsymbol{\Gamma}]_{2,2},\ln\alpha_f,\ln\alpha_n]^{\mathrm{T}}\in\mathbb{R}^4$. The conjugate gradient method [41] solves (50) with respect to $\tilde{\boldsymbol{\theta}}$, where the initial value of $\tilde{\boldsymbol{\theta}}$ is $[0,0,0,\ln(0.1)]^{\mathrm{T}}$. In Line 3, the coefficients in the objective function $g(\boldsymbol{v},\boldsymbol{A},p)$ in (55) are set to $\eta_V=0$, $\eta_Q=5$, and $\eta_{\mathrm{reg}}=0.5$. The predefined states $\acute{\boldsymbol{x}}_d$ with $\acute{D}=121$ are sampled at regular intervals on $[-3,3]\times[-3,3]$. The optimization problem described in (59) is solved via the quasi-Newton (BFGS) method [41]. The values of the initial parameters are set as $p=\gamma$ and $\boldsymbol{A}=(p\boldsymbol{S}(\gamma)-\boldsymbol{I}_2)/2$, such that the positive definiteness of $\boldsymbol{Q}$ in (39) holds.

Performance of the proposed optimal and H_∞ controllers are compared against those of existing optimal and H_∞ controllers developed for a linear model $\boldsymbol{f}(\boldsymbol{x}):=\boldsymbol{A}_{\mathrm{Linear}}\boldsymbol{x}$ identified from the training data set (termed Linear) and a linearized model $\boldsymbol{f}(\boldsymbol{x}):=(\partial\boldsymbol{\mu}_f(\boldsymbol{x})/\partial\boldsymbol{x}^{\mathrm{T}})|_{\boldsymbol{x}=0}\boldsymbol{x}$ around the origin

TABLE 1. Average values of the cost function (67) for optimal control.

Linear	Jacobi	**Proposed optimal controller**
66.4	∞ (diverged)	**45.8**



(a) Linear (b) Jacobi

(c) Proposed controller

FIGURE 1. Control results for optimal control. The markers • and ∘ indicate the initial and terminal states, respectively, with thin contour lines representing the designed state cost function $q(x)$.

(termed Jacobi). The optimal and H_∞ controllers for Linear and Jacobi models can be readily calculated, because the state cost function $q(\boldsymbol{x})$ in (37) reduces to a quadratic function of $\boldsymbol{x}$; i.e., $q(\boldsymbol{x}) = \boldsymbol{\zeta}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\zeta}(\boldsymbol{x})$, where $\boldsymbol{\zeta}(\boldsymbol{x}) = \boldsymbol{A}^{-1}\boldsymbol{f}(\boldsymbol{x})$ is linear in $\boldsymbol{x}$. The control simulation is performed using the Dormand–Prince method [45] with a sampling time of 0.005.

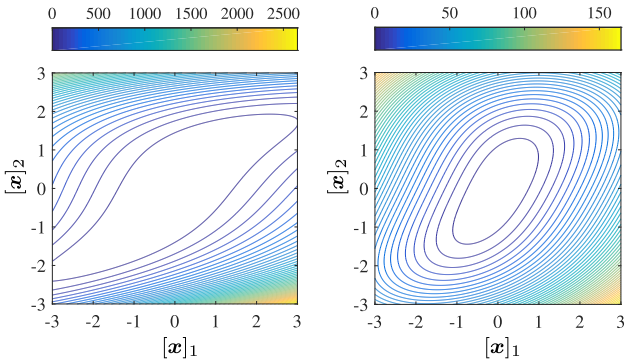
B. RESULTS FOR OPTIMAL CONTROL

This subsection evaluates control performance of the optimal controller designed using the proposed method, where the disturbance matrix is set to $\boldsymbol{B}_{\mathrm{d}} = 0$. There is no system noise ($\boldsymbol{\omega}(t) = 0$) in the true system considered in this evaluation. The control performance is evaluated in terms of the cost function over the time horizon [0, 20]:

$$\int_0^{20} \Big(q(\boldsymbol{x}(t)) + \frac{1}{2}\boldsymbol{u}(t)^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}(t) \Big) \mathrm{d}t. \tag{67}$$

The true system in (11) is controlled from the initial states $\boldsymbol{x}(0) := [2.9\cos(2\pi l/6), 2.9\sin(2\pi l/6)]^{\mathrm{T}}$ for $l = 1, \ldots, 6$.

Table 1 lists the average values of the cost function (67) with respect to the six initial states. Using the proposed method, the average cost was successfully reduced from that of the existing methods. Figure 1 depicts control results and the state cost function $q(\boldsymbol{x})$ that was obtained using the proposed method. The proposed controller confirmed that for all the initial states, the state trajectories terminated at the origin.



(a) State cost function $q(\boldsymbol{x})$. (b) Value function $V(\boldsymbol{x})$.

FIGURE 2. State cost function $q(x)$ and value function $V(x)$ designed using the proposed method.

In Fig. 1 (b), we consider that the divergence (while using Jacobi controller) was caused by the initial states far from the origin. Although Jacobi controller may stabilize the system in a small neighborhood of the origin, the stability in a large region is not guaranteed.

The effect of the randomness of the numerical simulation is evaluated. The variation in the system model is caused by the random noise included in the training data set. The control performance depends on this variation because the proposed controller is based on the system model. Let us define the performance ratio as the ratio of the average cost of the proposed controller divided by that of Linear controller. The value of this ratio was calculated as 0.690 in the case of Table 1. We evaluated the ratios for 20 different random seeds to investigate statistical results of the control performance. The values of the mean and sample standard deviation of the ratio were 0.682 and 0.035, respectively.

We also evaluated the positive definiteness of the state cost function $q(\boldsymbol{x})$ and value function $V(\boldsymbol{x})$, which is assumed in Theorems 1, 2, and 4. The values of $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$ were evaluated for states $\boldsymbol{x}$ sampled at regular intervals of 0.01 on $[-3, 3] \times [-3, 3]$. The positive definiteness was numerically (approximately) confirmed because, the values of $V(\boldsymbol{x})$ and $q(\boldsymbol{x})$ were observed to be positive for all the sampled states except the origin, as described in Fig. 2.

C. RESULTS FOR H_∞ CONTROL

Control performance associated with the robustness of the proposed H_∞ controller is evaluated in this subsection. The disturbance matrix is set to $\boldsymbol{B}_{\mathrm{d}} = \boldsymbol{I}_2$. The control performance is evaluated in terms of the L_2 gain from the disturbance to the performance output over the time horizon [0, 20]:

$$\Big(\int_0^{20} \|z(t)\|^2 \mathrm{d}t \Big)^{\frac{1}{2}} \Big/ \Big(\int_0^{20} \|\boldsymbol{\omega}(t)\|^2 \mathrm{d}t \Big)^{\frac{1}{2}}. \tag{68}$$

The system noise $\boldsymbol{\omega}(t)$ in the true system (11) is regarded as the disturbance. The true system is controlled from the initial state at the origin $\boldsymbol{x}(0) = [0, 0]^{\mathrm{T}}$, where six evaluations under deterministic disturbances are performed along with one hundred evaluations under stochastic disturbances.

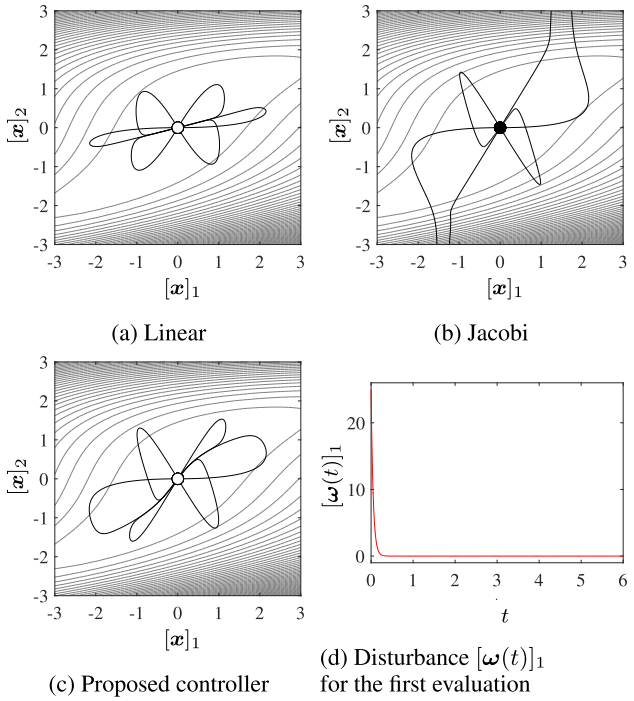


FIGURE 3. Control results for H_∞ control under the deterministic disturbance for the six evaluations with the initial state $x(0) = [0, 0]^{\mathsf{T}}$. The markers ∘ indicate the terminal states. The thin contour lines represent the designed state cost function $q(x)$.

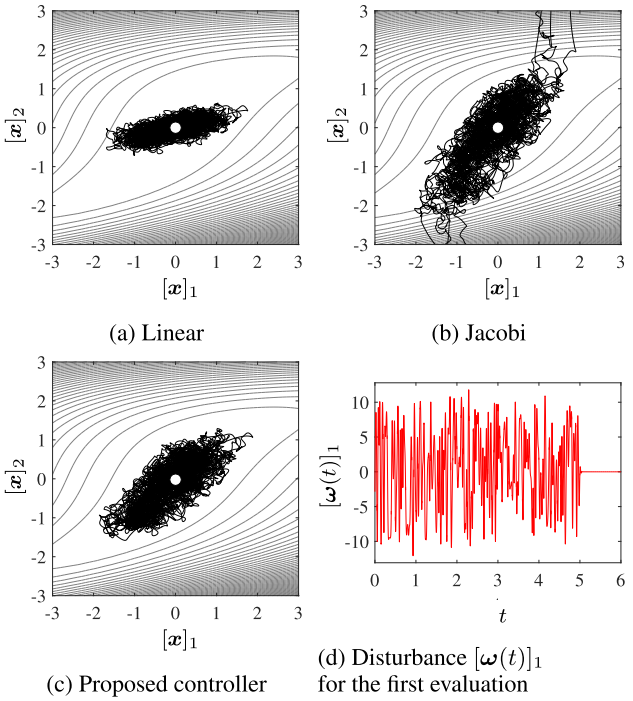


FIGURE 4. Control results for H_∞ control under the stochastic disturbance for the first twenty evaluations with the initial state $x(0) = [0, 0]^{\mathsf{T}}$. The markers ∘ indicate the terminal states. The thin contour lines represent the designed state cost function $q(x)$.

TABLE 2. Average values of the L_2 gain (68) for H_∞ control under the deterministic disturbance with respect to the six evaluations.

Linear	Jacobi	**Proposed H_∞ controller**
0.747	∞ (diverged)	**0.638**

TABLE 3. Average values of the L_2 gain (68) for H_∞ control under the stochastic disturbance with respect to the one hundred evaluations.

Linear	Jacobi	**Proposed H_∞ controller**
0.397	∞ (diverged)	**0.324**

The deterministic disturbance for the l-th evaluation ($l = 1, \ldots, 6$) is defined as the exponentially decaying function $\boldsymbol{\omega}(t) = [\cos(2\pi l/6), \sin(2\pi l/6)]^{\mathsf{T}} \times 50\exp(-20t)$. The stochastic disturbance $\boldsymbol{\omega}(t)$ for each $t \in \{0, 0.02, 0.04, \ldots, 4.98\}$ and $l \in \{1, \ldots, 100\}$ independently obeys the uniform distribution on $[-10, 10] \times [-10, 10]$, and $\boldsymbol{\omega}(t)$ is zero for $t \in \{5, 5.02, \ldots, 20\}$, where $\boldsymbol{\omega}(t)$ for $t \notin \{0, 0.02, 0.04, \ldots\}$ is interpolated using spline functions. Examples of $\boldsymbol{\omega}(t)$ are depicted in Fig. 3 (d) and Fig. 4 (d).

The average values of the L_2 gain (68) with respect to all the evaluations under the deterministic and stochastic disturbances are listed in Tables 2 and 3, respectively. For both disturbances, using the proposed method reduced the average L_2 gains from those obtained using the existing methods. Figures 3 and 4 depict control results and the state cost functions $q(\boldsymbol{x})$ under the deterministic and stochastic disturbances, respectively. In Fig. 3, the state trajectories converged to the origin when the proposed controller was applied. Figure 4 indicates that the state was maintained when using the proposed controller without divergence even though the true system was subjected to the stochastic disturbance. The performance ratios were evaluated for 20 different random seeds to investigate statistical results of the control performance. Let us redefine the performance ratio as the ratio of the average L_2 gain of the proposed controller divided by that of Linear controller. The mean and sample standard deviation of the ratio under the deterministic disturbances were 0.849 and 0.020, respectively. Under the stochastic disturbances, the corresponding values were 0.813 and 0.012, respectively.

The positive definiteness of the state cost function $q(\boldsymbol{x})$ and value function $V(\boldsymbol{x})$ was evaluated in a manner similar to that described for optimal control in Section VI-B. The positive definiteness was numerically (approximately) confirmed, because the proposed H_∞ control method obtained results similar to those depicted in Fig. 2.

D. RESULTS PERTAINING TO STABILITY OF THE TRUE FEEDBACK SYSTEM

Based on Theorem 4, the probabilistic stability of the true feedback system is evaluated under the condition of no disturbance ($\boldsymbol{\omega}(t) = 0$) when the proposed optimal controller is employed. The inequality $W(\boldsymbol{x}) < 0$ in (63) is evaluated for states $\boldsymbol{x}$ sampled at regular intervals of 0.01 on $\mathbb{X} := [-3, 3] \times [-3, 3]$. Such grid evaluations provided an estimate of the region $\tilde{\mathbb{X}}$ wherein $W(\boldsymbol{x}) < 0$ holds in Fig. 5. A region of attraction and small invariant set near the origin were estimated using the obtained $\tilde{\mathbb{X}}$, as depicted in Fig. 5.

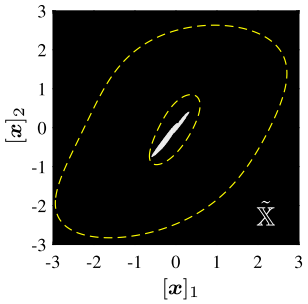


FIGURE 5. Stability evaluation with $\beta = 3$. The black region depicts an estimate of the region $\tilde{\mathbb{X}}$ wherein $W(x) < 0$. The dashed outer line and inner line indicate estimates of a region of attraction and small invariant set near the origin, respectively.

Recall that any state in the region of attraction arrives in the invariant set (see Remark 10). The drift term $\boldsymbol{f}_{\rm tr}(\boldsymbol{x})$ in (66) was included in $\mathbb{F}_{\rm tr}(\boldsymbol{x};\beta)$ for $\beta = 3$ at the sampled states. This result numerically (approximately) confirmed that the true feedback system with the proposed controller is practically stable. Note that the stability is ensured for drift terms $\boldsymbol{f}_{\rm tr}(\boldsymbol{x})$ included in $\mathbb{F}_{\rm tr}(\boldsymbol{x};\beta)$; however, the drift term $\boldsymbol{f}_{\rm tr}(\boldsymbol{x})$ in (66) is not always included in $\mathbb{F}_{\rm tr}(\boldsymbol{x};\beta)$. The training data set involving random noise affects whether a target $\boldsymbol{f}_{\rm tr}(\boldsymbol{x})$ is included or not. The proposed H_∞ controller obtained results similar to those obtained for the optimal controller.

E. PRACTICAL EXAMPLE

This subsection describes the application of the proposed method to a practical example. Consider an inverted pendulum subjected to a nonlinear torque, as illustrated in Fig. 6. The corresponding equation of motion is given by

$$ML^2\ddot{\rho} = MGL\sin\rho + T(\rho,\dot{\rho}) + \kappa_{\rm u}u, \tag{69}$$

where τ, M, L, G, ρ [rad], $\kappa_{\rm u}$, and u denote the time, mass, length, gravity acceleration, angle, torque coefficient, and control input of the pendulum, respectively. Let $\dot{\rho}$ and $\ddot{\rho}$ denote $\mathrm{d}\rho/\mathrm{d}\tau$ and $\mathrm{d}^2\rho/\mathrm{d}\tau^2$, respectively. The nonlinear torque $T(\rho,\dot{\rho})$ depends on ρ and $\dot{\rho}$ as follows:

$$T(\rho,\dot{\rho}) = \kappa_1\dot{\rho} + \kappa_2\rho^2\dot{\rho} + \kappa_3\rho\dot{\rho}^2 + \kappa_4\dot{\rho}^3. \tag{70}$$

By defining $t := \tau/\kappa_\tau$ and $\boldsymbol{x} := [\rho/\kappa_\rho, \dot{\rho}\kappa_\tau/\kappa_\rho]^{\rm T}$ with constants κ_τ and κ_ρ, the true drift term is expressed by

$$\boldsymbol{f}_{\rm tr}(\boldsymbol{x}) = \begin{bmatrix} [\boldsymbol{x}]_2 \\ \frac{\kappa_\tau^2 G}{\kappa_\rho L}\sin(\kappa_\rho[\boldsymbol{x}]_1) + \frac{\kappa_\tau^2}{\kappa_\rho ML^2}T(\kappa_\rho[\boldsymbol{x}]_1, \frac{\kappa_\rho}{\kappa_\tau}[\boldsymbol{x}]_2) \end{bmatrix}. \tag{71}$$

During evaluation, the parameters are set as follows: $M = 1.0$ [kg], $L = 20.0$ [m], $G = 9.80665$ [m/s^2], $\kappa_1 = 400$ [N·m·s], $\kappa_2 = 400$ [N·m·s], $\kappa_3 = 400$ [N·m·s^2], $\kappa_4 = 200$ [N·m·s^3], $\kappa_\tau = 1.0$ [s], $\kappa_\rho = 0.3$, and $\kappa_{\rm u} = \kappa_\rho ML^2/\kappa_\tau^2$. The other parameters, functions such as $\boldsymbol{\psi}_d(\boldsymbol{x})$, and settings are equivalent to those described in Sections VI-A and VI-B.

The proposed controller was evaluated in terms of control performance of the optimal controller with $\boldsymbol{B}_{\rm d} = 0$ and $\boldsymbol{\omega}(t) = 0$. Table 4 lists the average values of the cost function (67) with respect to the six initial states defined in Section VI-B. With the proposed method, the average cost was substantially reduced from that of the existing methods. We also evaluated the performance ratios for 50 different random seeds to investigate the control performance statistically. The performance ratio is redefined as the ratio of the average cost of the proposed controller divided by that of Linear controller. The values of the mean and sample standard deviation of the ratio were 0.791 and 0.018, respectively. For the ratio of the average cost of the proposed controller divided by that of Jacobi controller, the corresponding values were 0.702 and 0.095, respectively. We confirmed the effectiveness of the proposed method via the practical example.

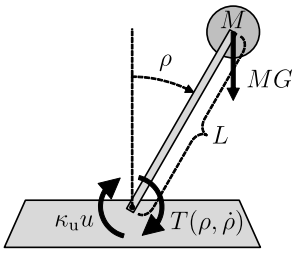


FIGURE 6. Practical example: an inverted pendulum subjected to a nonlinear torque $T(\rho,\dot{\rho})$.

TABLE 4. Average values of the cost function (67) for optimal control with the practical system (71).

Linear	Jacobi	Proposed optimal controller
49.5	56.7	**39.8**

VII. COMPARISON AGAINST EXISTING METHODS

The proposed method offers several advantages compared to the existing methods [11], [14]–[24]. The following advantages constitute the major contributions of this study.

(i) Exact solutions to optimal control and H_∞ control problems can be obtained for kernel-based system models, whereas the existing methods only yield approximate solutions.

(ii) Once a system model is obtained, the control problems can be automatically solved without the need for huge computation, whereas the existing methods involve huge computations. This advantage makes it easy to (re)tune the values of parameters (e.g., $\boldsymbol{Q}, \boldsymbol{R}, \gamma$) that affect the control policies.

(iii) Stability of a feedback system model with the proposed controllers is automatically guaranteed under certain assumptions. In contrast, the existing methods, except those in [16], [22]–[24], do not guarantee stability.

(iv) The numerical example described in Section VI demonstrates that the proposed controllers are superior to the other existing controllers (termed Linear and Jacobi) in terms of control performance.

VIII. CONCLUSION

This paper has presented a method to design nonlinear optimal and H_∞ controllers for partially unknown nonlinear systems described by kernel-based functions. Major contributions of this study towards solving Problems 1 and 2 are summarized as follows. First, Theorems 1 and 2 reduce HJI equations to algebraic matrix equations for a class of kernel-based system models. The algebraic matrix equations can be solved via analytical and/or numerical approaches, because complex nonlinear functions of the state are not included in the equations. Solving the matrix equations gives exact solutions to the HJI equations, and thus, optimal and H_∞ controllers are obtained. Secondly, the true drift term is identified as a kernel-based model. Data-driven optimal and H_∞ controllers are designed based on the model. For a true feedback system with the designed controllers, probabilistic stability is analyzed using GPs and Theorem 4. Compared to existing methods, the proposed method offers several advantages, as described in Section VII.

In future work, the proposed method should be extended for applications involving the control of unknown nonlinear systems, e.g., autonomous vehicles involving human interactions [1]. The realization of effective control requires the development of mathematical models of such unknown systems, regarding which only limited information is available. Data-driven approaches are important for overcoming this difficulty owing to their ability to design appropriate controllers using a data set. Additionally, we focus on a theoretical extension of the proposed method to find other classes of kernel-based models for which exact solutions to HJI equations can be obtained. Other extensions involve output feedback control problems associated with state estimations.

APPENDIX
PROOFS

A. PROOF OF LEMMA 1

From the definition of $V(\boldsymbol{x})$ in (17), the partial derivative $\partial_{\boldsymbol{x}} V(\boldsymbol{x})$ is calculated as follows under the condition given in (20):

$$\begin{aligned}\partial_{\boldsymbol{x}} V(\boldsymbol{x}) &= \sum_{d=1}^{D} \Big(\boldsymbol{\psi}_d(\boldsymbol{x}) \tilde{\boldsymbol{p}}_d \partial_{\boldsymbol{x}} k(\boldsymbol{x}, \boldsymbol{x}_d) + \partial_{\boldsymbol{x}} \boldsymbol{\psi}_d(\boldsymbol{x}) \tilde{\boldsymbol{p}}_d k(\boldsymbol{x}, \boldsymbol{x}_d) \Big) \\ &= \sum_{d=1}^{D} \Big(\boldsymbol{c}_V(\boldsymbol{x}, \boldsymbol{x}_d) \boldsymbol{\psi}_d(\boldsymbol{x}) \tilde{\boldsymbol{p}}_d + \partial_{\boldsymbol{x}} \boldsymbol{\psi}_d(\boldsymbol{x}) \tilde{\boldsymbol{p}}_d \Big) k(\boldsymbol{x}, \boldsymbol{x}_d) \\ &= \sum_{d=1}^{D} \boldsymbol{\Phi}'_d(\boldsymbol{x}) \tilde{\boldsymbol{p}}_d k(\boldsymbol{x}, \boldsymbol{x}_d). \end{aligned} \tag{72}$$

Using the property of $\mathrm{vec}(\boldsymbol{A}_\mathrm{a}\boldsymbol{A}_\mathrm{b}\boldsymbol{A}_\mathrm{c}) = (\boldsymbol{A}_\mathrm{c}^\mathrm{T} \otimes \boldsymbol{A}_\mathrm{a})\mathrm{vec}(\boldsymbol{A}_\mathrm{b})$ for given matrices $\boldsymbol{A}_\mathrm{a}$, $\boldsymbol{A}_\mathrm{b}$, and $\boldsymbol{A}_\mathrm{c}$ [46, (3.76)], the term $\boldsymbol{\Phi}'_d(\boldsymbol{x})\tilde{\boldsymbol{p}}_d$ in (72) becomes

$$\begin{aligned}\boldsymbol{\Phi}'_d(\boldsymbol{x})\tilde{\boldsymbol{p}}_d &= \mathrm{vec}(\boldsymbol{\Phi}'_d(\boldsymbol{x})\tilde{\boldsymbol{p}}_d) = \mathrm{vec}(\boldsymbol{I}_{n_x}\boldsymbol{\Phi}'_d(\boldsymbol{x})\tilde{\boldsymbol{p}}_d) \\ &= (\tilde{\boldsymbol{p}}_d^\mathrm{T} \otimes \boldsymbol{I}_{n_x})\mathrm{vec}(\boldsymbol{\Phi}'_d(\boldsymbol{x})). \end{aligned} \tag{73}$$

Because $\tilde{\boldsymbol{p}}^\mathrm{T} = [\tilde{\boldsymbol{p}}_1^\mathrm{T}, \ldots, \tilde{\boldsymbol{p}}_D^\mathrm{T}]$ in (22) is the row vector,

$$[\tilde{\boldsymbol{p}}_1^\mathrm{T} \otimes \boldsymbol{I}_{n_x}, \ldots, \tilde{\boldsymbol{p}}_D^\mathrm{T} \otimes \boldsymbol{I}_{n_x}] = \tilde{\boldsymbol{p}}^\mathrm{T} \otimes \boldsymbol{I}_{n_x} \tag{74}$$

holds. Substituting (24), (73), and (74) into (21) yields (72). This completes the proof. □

B. PROOF OF THEOREM 1

Substituting $\partial_{\boldsymbol{x}} V(\boldsymbol{x})$ in (21), $\boldsymbol{f}(\boldsymbol{x})$ in (30), and $q(\boldsymbol{x})$ in (31) into the HJI equation (6) yields the relation described in (15) as follows:

$$\begin{aligned} H_{\mathrm{HJI}}(\boldsymbol{x}) &= \boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^\mathrm{T}\boldsymbol{\Phi}(\boldsymbol{x})^\mathrm{T}(\tilde{\boldsymbol{p}}^\mathrm{T} \otimes \boldsymbol{I}_{n_x})^\mathrm{T}\tilde{\boldsymbol{A}}\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) \\ &\quad - \frac{1}{2}\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^\mathrm{T}\boldsymbol{\Phi}(\boldsymbol{x})^\mathrm{T}(\tilde{\boldsymbol{p}}^\mathrm{T} \otimes \boldsymbol{I}_{n_x})^\mathrm{T} \\ &\quad \times \boldsymbol{S}(\gamma)(\tilde{\boldsymbol{p}}^\mathrm{T} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) \\ &\quad + \boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^\mathrm{T}\boldsymbol{\Phi}(\boldsymbol{x})^\mathrm{T}\tilde{\boldsymbol{Q}}\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}) \\ &= \frac{1}{2}\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})^\mathrm{T}\boldsymbol{\Phi}(\boldsymbol{x})^\mathrm{T}\boldsymbol{M}_{\mathrm{HJI}}\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x}). \end{aligned} \tag{75}$$

Because of (75), $\boldsymbol{M}_{\mathrm{HJI}} = 0$ is a sufficient condition for $H_{\mathrm{HJI}}(\boldsymbol{x}) = 0$ to be satisfied for all $\boldsymbol{x} \in \mathbb{R}^{n_x}$.

Next, substituting (18) and (19) into $\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ yields $\boldsymbol{\Phi}(0)\boldsymbol{k}_{\mathrm{vec}}(0) = 0$. Note that $\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$ is a linear combination of $\partial_{\boldsymbol{x}} k(\boldsymbol{x}, \boldsymbol{x}_d)\boldsymbol{\psi}_d(\boldsymbol{x})$ and $k(\boldsymbol{x}, \boldsymbol{x}_d)\partial_{\boldsymbol{x}}\boldsymbol{\psi}_d(\boldsymbol{x})$. This relation leads to the local Lipschitz continuity of $\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\mathrm{vec}}(\boldsymbol{x})$, because $k(\boldsymbol{x}, \boldsymbol{x}_d)$ and $\boldsymbol{\psi}_d(\boldsymbol{x})$ are C^2 continuous. Therefore, $\boldsymbol{f}(0) = 0$ holds, and $\boldsymbol{f}(\boldsymbol{x})$ is locally Lipschitz continuous. This completes the proof. □

C. PROOF OF THEOREM 2

The statement can be proved in a manner similar to Theorem 1. We replace $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{p}}$, and $\tilde{\boldsymbol{Q}} \succeq 0$ in Theorem 1 with $\boldsymbol{A}(\boldsymbol{v}^\mathrm{T} \otimes \boldsymbol{I}_{n_x})$, $p\boldsymbol{v}$, and $(\boldsymbol{v}^\mathrm{T} \otimes \boldsymbol{I}_{n_x})^\mathrm{T}\boldsymbol{Q}(\boldsymbol{v}^\mathrm{T} \otimes \boldsymbol{I}_{n_x}) \succeq 0$, respectively. Performing these replacements in (30), (17), (31), and (32) satisfies (35)–(38). Therefore, the algebraic matrix equation (38) corresponds to a special case of (32), and thus, a sufficient condition that the HJI equation (6) holds because of Theorem 1. The proof of $\boldsymbol{f}(0) = 0$ and the local Lipschitz continuity of $\boldsymbol{f}(\boldsymbol{x})$ are equivalent to those described for Theorem 1. This completes the proof. □

D. PROOF OF THEOREM 3

Let $c_d \in \mathbb{R}$ be a parameter. We consider the setting of $\nu_d = c_d\alpha_f \in \mathbb{R}$ and $p = 1$. Using the relations (44) and (45), we represent $\psi_d(\boldsymbol{x})\nu_d$ in (36) as follows:

$$\psi_d(\boldsymbol{x})\nu_d = c_d\alpha_f\boldsymbol{x}_d^\mathrm{T}\boldsymbol{x} = c_d k(\boldsymbol{x}, \boldsymbol{x}_d). \tag{76}$$

Let us define a symmetric matrix $\hat{\boldsymbol{P}}_{\mathrm{tr}}$ as

$$\hat{\boldsymbol{P}}_{\mathrm{tr}} := 2\sum_{d=1}^{D} c_d\alpha_f^2\boldsymbol{x}_d\boldsymbol{x}_d^\mathrm{T}. \tag{77}$$

From (36) and (77), $V(\boldsymbol{x})$ is given as

$$\begin{aligned} V(\boldsymbol{x}) &= \sum_{d=1}^{D} c_d k(\boldsymbol{x}, \boldsymbol{x}_d)^2 = \boldsymbol{x}^\mathrm{T}\Big(\sum_{d=1}^{D} c_d\alpha_f^2\boldsymbol{x}_d\boldsymbol{x}_d^\mathrm{T}\Big)\boldsymbol{x} \\ &= \frac{1}{2}\boldsymbol{x}^\mathrm{T}\hat{\boldsymbol{P}}_{\mathrm{tr}}\boldsymbol{x}. \end{aligned} \tag{78}$$

For any symmetric matrix $\hat{\boldsymbol{P}}_{\rm tr}$, the definition (77) is equivalent to the following vectorized form:

$$\mathrm{vech}(\hat{\boldsymbol{P}}_{\rm tr}) = 2\sum_{d=1}^{D} c_d \alpha_f^2 \mathrm{vech}(\boldsymbol{x}_d \boldsymbol{x}_d^{\rm T}) \in \mathbb{R}^{n_x(n_x+1)/2}. \quad (79)$$

From the condition (41), for any symmetric matrix $\hat{\boldsymbol{P}}_{\rm tr}$, there exist $[c_1, c_2, \ldots, c_D]$ such that (79) and (77) hold. Therefore, we choose $\hat{\boldsymbol{P}}_{\rm tr} = \boldsymbol{P}_{\rm tr}$ that satisfies the positive definiteness of $V(\boldsymbol{x})$. In addition, the partial derivative of $V(\boldsymbol{x})$ calculated in (78) is equivalent to the relation (21) with $p = 1$ as follows:

$$\partial_{\boldsymbol{x}} V(\boldsymbol{x}) = \boldsymbol{P}_{\rm tr}\boldsymbol{x} = (\boldsymbol{v}^{\rm T} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\rm vec}(\boldsymbol{x}). \quad (80)$$

Substituting this relation into (35) and (37) yields

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{P}_{\rm tr}\boldsymbol{x}, \quad (81)$$
$$q(\boldsymbol{x}) = \boldsymbol{x}^{\rm T}\boldsymbol{P}_{\rm tr}^{\rm T}\boldsymbol{Q}\boldsymbol{P}_{\rm tr}\boldsymbol{x}. \quad (82)$$

We obtain the positive definite $\boldsymbol{Q} = \boldsymbol{P}_{\rm tr}^{-{\rm T}}\boldsymbol{Q}_{\rm tr}\boldsymbol{P}_{\rm tr}^{-1} \succ 0$ to satisfy (43) and (82), where $q(\boldsymbol{x})$ is positive definite. Choosing $\boldsymbol{A} = \boldsymbol{A}_{\rm tr}\boldsymbol{P}_{\rm tr}^{-1}$ satisfies $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}_{\rm tr}(\boldsymbol{x})$. Using these parameters with $p = 1$ along with the condition (40), the algebraic matrix equation (38) holds owing to the following relation:

$$\begin{aligned}
&(p\boldsymbol{A} + p\boldsymbol{A}^{\rm T} - p^2\boldsymbol{S}(\gamma) + 2\boldsymbol{Q}) \\
&\quad = \boldsymbol{A}_{\rm tr}\boldsymbol{P}_{\rm tr}^{-1} + \boldsymbol{P}_{\rm tr}^{-{\rm T}}\boldsymbol{A}_{\rm tr}^{\rm T} - \boldsymbol{S}(\gamma) + 2\boldsymbol{P}_{\rm tr}^{-{\rm T}}\boldsymbol{Q}_{\rm tr}\boldsymbol{P}_{\rm tr}^{-1} \\
&\quad = \boldsymbol{P}_{\rm tr}^{-{\rm T}}(\boldsymbol{P}_{\rm tr}^{\rm T}\boldsymbol{A}_{\rm tr} + \boldsymbol{A}_{\rm tr}^{\rm T}\boldsymbol{P}_{\rm tr} - \boldsymbol{P}_{\rm tr}^{\rm T}\boldsymbol{S}(\gamma)\boldsymbol{P}_{\rm tr} + 2\boldsymbol{Q}_{\rm tr})\boldsymbol{P}_{\rm tr}^{-1} \\
&\quad = 0. \qquad (83)
\end{aligned}$$

This completes the proof. □

E. PROOF OF PROPOSITION 1

Using the relation between (21) and (73), $\boldsymbol{f}(\boldsymbol{x})$ in (35) can be represented as follows:

$$\begin{aligned}
\boldsymbol{f}(\boldsymbol{x};\boldsymbol{v},p) &= \boldsymbol{A}(\boldsymbol{v}^{\rm T} \otimes \boldsymbol{I}_{n_x})\boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{k}_{\rm vec}(\boldsymbol{x}) \\
&= \boldsymbol{A}\sum_{d=1}^{D} \boldsymbol{\Phi}_d'(\boldsymbol{x})\boldsymbol{v}_d k(\boldsymbol{x},\boldsymbol{x}_d) \\
&= [\boldsymbol{A}k(\boldsymbol{x},\boldsymbol{x}_1)\boldsymbol{\Phi}_1'(\boldsymbol{x}), \ldots, \boldsymbol{A}k(\boldsymbol{x},\boldsymbol{x}_D)\boldsymbol{\Phi}_D'(\boldsymbol{x})]\boldsymbol{v}. \quad (84)
\end{aligned}$$

The value function $V(\boldsymbol{x})$ defined in (36) is expressed as

$$\begin{aligned}
V(\boldsymbol{x};\boldsymbol{v},p) &= \sum_{d=1}^{D} \boldsymbol{\psi}_d(\boldsymbol{x})p\boldsymbol{v}_d k(\boldsymbol{x},\boldsymbol{x}_d) \\
&= [pk(\boldsymbol{x},\boldsymbol{x}_1)\boldsymbol{\psi}_1(\boldsymbol{x}), \ldots, pk(\boldsymbol{x},\boldsymbol{x}_D)\boldsymbol{\psi}_D(\boldsymbol{x})]\boldsymbol{v}. \quad (85)
\end{aligned}$$

By substituting (84) and (85) into the expression for $g(\boldsymbol{v},\boldsymbol{A},p)$ described in (55), $g(\boldsymbol{v},\boldsymbol{A},p)$ is given as the following quadratic function of $\boldsymbol{v}$:

$$g(\boldsymbol{v},\boldsymbol{A},p) = \|\boldsymbol{Y}(\boldsymbol{A},p)\boldsymbol{v} - \boldsymbol{y}\|^2 + \eta_{\rm reg}\|\boldsymbol{v}\|^2 + g_{\rm c}(\boldsymbol{A},p), \quad (86)$$

where $g_{\rm c}(\boldsymbol{A},p)$ is a function of $\boldsymbol{A}$ and p and is independent of $\boldsymbol{v}$. Therefore, for each fixed $\boldsymbol{A}$ and p, (54) represents a regularized least-squares minimization problem with respect to $\boldsymbol{v}$. The solution to the problem is explicitly obtained as described in (56), where the matrix $(\boldsymbol{Y}(\boldsymbol{A},p)^{\rm T}\boldsymbol{Y}(\boldsymbol{A},p) + \eta_{\rm reg}\boldsymbol{I}_{n_pD})$ is nonsingular because $\eta_{\rm reg} > 0$ holds. This completes the proof. □

F. PROOF OF THEOREM 4

In this proof, the parameters $(\boldsymbol{v},\boldsymbol{A},p)$ included in functions are omitted for the sake of brevity. For example, $V(\boldsymbol{x};\boldsymbol{v},\boldsymbol{A})$ is simply denoted by $V(\boldsymbol{x})$. For the value function $V(\boldsymbol{x})$ in (36) and true system (11) with $\boldsymbol{u}_*(\boldsymbol{x})$ applied and $\boldsymbol{\omega}(t) = 0$, Assumption 1 leads to the following relation with a probability of at least $1-\delta$:

$$\begin{aligned}
\frac{{\rm d}V}{{\rm d}t}(\boldsymbol{x}(t)) &= \partial_{\boldsymbol{x}}V(\boldsymbol{x}(t))^{\rm T}(\boldsymbol{f}_{\rm tr}(\boldsymbol{x}(t)) + \boldsymbol{B}\boldsymbol{u}_*(\boldsymbol{x}(t))) \\
&\le \partial_{\boldsymbol{x}}V(\boldsymbol{x}(t))^{\rm T}(\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x}(t)) + \boldsymbol{B}\boldsymbol{u}_*(\boldsymbol{x}(t))) \\
&\quad + \beta\sum_{s=1}^{n_x}\left|[\partial_{\boldsymbol{x}}V(\boldsymbol{x}(t))]_s[\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x}(t))]_s\right|, \quad \forall \boldsymbol{x}(t) \in \mathbb{X}. \qquad (87)
\end{aligned}$$

Using the definition of $\Delta\boldsymbol{f}(\boldsymbol{x};\boldsymbol{v},\boldsymbol{A})$ in (53) yields

$$\partial_{\boldsymbol{x}}V(\boldsymbol{x})^{\rm T}\boldsymbol{\mu}_{\boldsymbol{f}}(\boldsymbol{x}) = \partial_{\boldsymbol{x}}V(\boldsymbol{x})^{\rm T}(\boldsymbol{f}(\boldsymbol{x}) + \Delta\boldsymbol{f}(\boldsymbol{x})). \quad (88)$$

Because the HJI equation (6) holds, using the definition of $\boldsymbol{u}_*(\boldsymbol{x})$ in (8), the following relation can be obtained:

$$\begin{aligned}
&\partial_{\boldsymbol{x}}V(\boldsymbol{x})^{\rm T}(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{u}_*(\boldsymbol{x})) \\
&\quad = \left(\frac{1}{2}\partial_{\boldsymbol{x}}V(\boldsymbol{x})^{\rm T}\boldsymbol{S}(\gamma)\partial_{\boldsymbol{x}}V(\boldsymbol{x}) - q(\boldsymbol{x})\right) \\
&\qquad + \partial_{\boldsymbol{x}}V(\boldsymbol{x})^{\rm T}\boldsymbol{B}\boldsymbol{u}_*(\boldsymbol{x}) \\
&\quad = \frac{-1}{2}\partial_{\boldsymbol{x}}V(\boldsymbol{x})^{\rm T}\left(\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\rm T} + \frac{1}{\gamma^2}\boldsymbol{B}_{\rm d}\boldsymbol{B}_{\rm d}^{\rm T}\right)\partial_{\boldsymbol{x}}V(\boldsymbol{x}) \\
&\qquad - q(\boldsymbol{x}). \qquad (89)
\end{aligned}$$

Substituting (88) and (89) into (87) yields

$$\begin{aligned}
&\frac{{\rm d}V}{{\rm d}t}(\boldsymbol{x}(t)) \\
&\le \partial_{\boldsymbol{x}}V(\boldsymbol{x}(t))^{\rm T}\Delta\boldsymbol{f}(\boldsymbol{x}(t)) \\
&\quad + \beta\sum_{s=1}^{n_x}\left|[\partial_{\boldsymbol{x}}V(\boldsymbol{x}(t))]_s[\boldsymbol{\sigma}_{\boldsymbol{f}}(\boldsymbol{x}(t))]_s\right| \\
&\quad - \frac{1}{2}\partial_{\boldsymbol{x}}V(\boldsymbol{x}(t))^{\rm T}\left(\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\rm T} + \frac{1}{\gamma^2}\boldsymbol{B}_{\rm d}\boldsymbol{B}_{\rm d}^{\rm T}\right)\partial_{\boldsymbol{x}}V(\boldsymbol{x}(t)) \\
&\quad - q(\boldsymbol{x}(t)), \quad \forall \boldsymbol{x}(t) \in \mathbb{X}, \qquad (90)
\end{aligned}$$

with a probability of at least $1-\delta$. Here, the relation $q(\boldsymbol{x}) = \boldsymbol{\zeta}(\boldsymbol{x})^{\rm T}\boldsymbol{Q}\boldsymbol{\zeta}(\boldsymbol{x})$ holds because of (37) and (65). Additionally, the relation $\partial_{\boldsymbol{x}}V(\boldsymbol{x}) = p\boldsymbol{\zeta}(\boldsymbol{x})$ is obtained from Lemma 1. Substituting these relations replaces (90) with the following inequality:

$$\frac{{\rm d}V}{{\rm d}t}(\boldsymbol{x}(t)) \le W(\boldsymbol{x}(t)), \quad \forall \boldsymbol{x}(t) \in \mathbb{X}. \quad (91)$$

Therefore, the statement (63) holds with a probability of at least $1-\delta$ owing to (91). This completes the proof. □

ACKNOWLEDGMENT

The authors would like to thank Editage (www.editage.jp) for English language editing.

REFERENCES

[1] T. Nishi, P. Doshi, and D. Prokhorov, "Merging in congested freeway traffic using multipolicy decision making and passive actor-critic learning," *IEEE Trans. Intell. Veh.*, vol. 4, no. 2, pp. 287–297, Jun. 2019.

[2] 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. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 974–983, Jun. 2013.

[3] M. Torchio, N. A. Wolff, D. M. Raimondo, L. Magni, U. Krewer, R. B. Gopaluni, J. A. Paulson, and R. D. Braatz, "Real-time model predictive control for the optimal charging of a lithium-ion battery," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2015, pp. 4536–4541.

[4] V. Vovk, "Kernel ridge regression," in *Empirical Inference. Festschrift in Honor of Vladimir N. Vapnik*. Berlin, Germany: Springer-Verlag, 2013, pp. 105–116.

[5] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[6] B. Hu, G. Su, J. Jiang, J. Sheng, and J. Li, "Uncertain prediction for slope displacement time-series using Gaussian process machine learning," *IEEE Access*, vol. 7, pp. 27535–27546, 2019.

[7] W. Guo, T. Pan, Z. Li, and S. Chen, "Model calibration method for soft sensors using adaptive Gaussian process regression," *IEEE Access*, vol. 7, pp. 168436–168443, 2019.

[8] Y. Ito, K. Fujimoto, Y. Tadokoro, and T. Yoshimura, "On stabilizing control of Gaussian processes for unknown nonlinear systems," in *Proc. 20th IFAC World Congr.*, 2017, pp. 15955–15960.

[9] M. P. Deisenroth, C. E. Rasmussen, and J. Peters, "Gaussian process dynamic programming," *Neurocomputing*, vol. 72, nos. 7–9, pp. 1508–1524, Mar. 2009.

[10] P. Hemakumara and S. Sukkarieh, "Learning UAV stability and control derivatives using Gaussian processes," *IEEE Trans. Robot.*, vol. 29, no. 4, pp. 813–824, Aug. 2013.

[11] M. P. Deisenroth, D. Fox, and C. E. Rasmussen, "Gaussian processes for data-efficient learning in robotics and control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 2, pp. 408–423, Feb. 2015.

[12] F. Xie, W. Hong, W. Wu, K. Liang, and C. Qiu, "Current distribution method of induction motor for electric vehicle in whole speed range based on Gaussian process," *IEEE Access*, vol. 7, pp. 165974–165984, 2019.

[13] K. Chen, J. Yi, and D. Song, "Gaussian processes model-based control of underactuated balance robots," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 4458–4464.

[14] J. Kocijan, R. Murray-Smith, C. E. Rasmussen, and B. Likar, "Predictive control with Gaussian process models," in *Proc. IEEE Region 8 EUROCON Comput. Tool*, Sep. 2003, pp. 352–356.

[15] A. Grancharova, J. Kocijan, and T. A. Johansen, "Explicit stochastic nonlinear predictive control based on Gaussian process models," in *Proc. Eur. Control Conf. (ECC)*, Jul. 2007, pp. 2340–2347.

[16] G. Cao, "Gaussian process based model predictive control," Ph.D. dissertation, School Eng. Adv. Technol., Massey Univ., Palmerston North, New Zealand, 2017. [Online]. Available: https://mro.massey.ac.nz/handle/10179/11443

[17] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using Gaussian process regression," *IEEE Trans. Control Syst. Technol.*, early access, Nov. 21, 2019, doi: 10.1109/TCST.2019.2949757.

[18] T. X. Nghiem, "Linearized Gaussian processes for fast data-driven model predictive control," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2019, pp. 1629–1634.

[19] E. Bradford, L. Imsland, and E. A. del Rio-Chanona, "Nonlinear model predictive control with explicit back-offs for Gaussian process state space models," in *Proc. IEEE 58th Conf. Decis. Control (CDC)*, Dec. 2019, pp. 4747–4754.

[20] Y. Pan and E. A. Theodorou, "Data-driven differential dynamic programming using Gaussian processes," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2015, pp. 4467–4472.

[21] J. Boedecker, J. T. Springenberg, J. Wulfing, and M. Riedmiller, "Approximate real-time optimal control based on sparse Gaussian process models," in *Proc. IEEE Symp. Adapt. Dyn. Program. Reinforcement Learn. (ADPRL)*, Dec. 2014, pp. 1–8.

[22] J. Umlauft, L. Pohler, and S. Hirche, "An uncertainty-based control Lyapunov approach for control-affine systems modeled by Gaussian process," *IEEE Control Syst. Lett.*, vol. 2, no. 3, pp. 483–488, Jul. 2018.

[23] Y. Ito, K. Fujimoto, and Y. Tadokoro, "Second-order bounds of Gaussian kernel-based functions and its application to nonlinear optimal control with stability," 2017, *arXiv:1707.06240*. [Online]. Available: http://arxiv.org/abs/1707.06240

[24] F. Berkenkamp, M. Turchetta, A. P. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 908–918.

[25] Y.-C. Chang, N. Roohi, and S. Gao, "Neural Lyapunov control," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 3245–3254.

[26] Y. Li, S. Guo, L. Zhu, T. Mukai, and Z. Gan, "Enhanced probabilistic inference algorithm using probabilistic neural networks for learning control," *IEEE Access*, vol. 7, pp. 184457–184467, 2019.

[27] C. Mu, D. Wang, and H. He, "Novel iterative neural dynamic programming for data-based approximate optimal control design," *Automatica*, vol. 81, pp. 240–252, Jul. 2017.

[28] C. Mu, C. Sun, D. Wang, and A. Song, "Adaptive tracking control for a class of continuous-time uncertain nonlinear systems using the approximate solution of HJB equation," *Neurocomputing*, vol. 260, pp. 432–442, Oct. 2017.

[29] A. van der Schaft, "L_2-gain analysis of nonlinear systems and nonlinear state feedback H_∞ control," *IEEE Trans. Autom. Control*, vol. 37, no. 6, pp. 770–784, Jun. 1992.

[30] Y. Ito, K. Fujimoto, T. Yoshimura, and Y. Tadokoro, "On Gaussian kernel-based Hamilton-Jacobi-Bellman equations for nonlinear optimal control," in *Proc. Annu. Amer. Control Conf. (ACC)*, Jun. 2018, pp. 1835–1840.

[31] F. L. Lewis, D. L. Vrabie, and V. L. Syrmos, *Optimal Control*, 3rd ed. Hoboken, NJ, USA: Wiley, 2012.

[32] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag, 2006.

[33] E. Todorov, "Eigenfunction approximation methods for linearly-solvable optimal control problems," in *Proc. IEEE Symp. Adapt. Dyn. Program. Reinforcement Learn.*, Mar. 2009, pp. 161–168.

[34] R. Kamalapurkar, J. A. Rosenfeld, and W. E. Dixon, "Efficient model-based reinforcement learning for approximate online optimal control," *Automatica*, vol. 74, pp. 247–258, Dec. 2016.

[35] P. Giesl, "Construction of a local and global Lyapunov function for discrete dynamical systems using radial basis functions," *J. Approximation Theory*, vol. 153, no. 2, pp. 184–211, Aug. 2008.

[36] P. Giesl and S. Hafstein, "Review on computational methods for Lyapunov functions," *Discrete Continuous Dyn. Syst.-Ser. B*, vol. 20, no. 8, pp. 2291–2331, Aug. 2015.

[37] J. Bouvrie and B. Hamzi, "Kernel methods for the approximation of some key quantities of nonlinear systems," *J. Comput. Dyn.*, vol. 4, pp. 1–19, Jun. 2017.

[38] R. W. Beard, G. Saridis, and J. Wen, "Approximate solutions to the time-invariant Hamilton–Jacobi–Bellman equation," *J. Optim. Theory Appl.*, vol. 96, no. 3, pp. 589–626, 1998.

[39] Y. Zhu, D. Zhao, X. Yang, and Q. Zhang, "Policy iteration for H_∞ optimal control of polynomial nonlinear systems via sum of squares programming," *IEEE Trans. Cybern.*, vol. 48, no. 2, pp. 500–509, Feb. 2018.

[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[41] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.

[42] F. Berkenkamp, R. Moriconi, A. P. Schoellig, and A. Krause, "Safe learning of regions of attraction for uncertain, nonlinear systems with Gaussian processes," in *Proc. IEEE 55th Conf. Decis. Control (CDC)*, Dec. 2016, pp. 4661–4666.

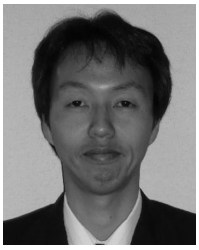
[43] Y. Ito, K. Fujimoto, and Y. Tadokoro, "Sampling-based stability evaluation with second-order margins for unknown systems with Gaussian processes," in *Proc. IEEE 58th Conf. Decis. Control (CDC)*, Dec. 2019, pp. 3471–3477.

[44] C. E. Rasmussen and H. Nickisch. (2018). *Gaussian Process Regression Classification Toolbox Version 4.2*. [Online]. Available: http://www.gaussianprocess.org/gpml/code/matlab/doc/

[45] J. R. Dormand and P. J. Prince, "A family of embedded Runge-Kutta formulae," *J. Comput. Appl. Math.*, vol. 6, no. 1, pp. 19–26, Mar. 1980.

[46] J. E. Gentle, *Matrix Algebra: Theory, Computations, and Applications in Statistics*. New York, NY, USA: Springer, 2007.

YUJI ITO (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Nagoya University, Aichi, Japan, in 2009, 2011, and 2014, respectively. Since 2011, he has been with Toyota Central R&D Labs., Inc., Japan. His research interests include stochastic systems theory and nonlinear optimal control. He is a member of the Society of Instrument and Control Engineers.



KENJI FUJIMOTO (Member, IEEE) received the B.S. and M.S. degrees in engineering and the Ph.D. degree in informatics from Kyoto University, Kyoto, Japan, in 1994, 1996, and 2001, respectively. From 1997 to 2004, he was a Research Associate with the Graduate School of Engineering and the Graduate School of Informatics, Kyoto University. From 1999 to 2000, he was a Research Fellow with the Department of Electrical Engineering, Delft University of Technology, The Netherlands. He held visiting research positions with The Australian National University, Australia, in 1999, and the Delft University of Technology, in 2002. From 2004 to 2012, he was an Associate Professor with the Graduate School of Engineering, Nagoya University, Japan. He is currently a Professor with the Graduate School of Engineering, Kyoto University. His research interests include nonlinear control and stochastic systems theory. He received the IFAC Congress Young Author Prize from the IFAC World Congress, in 2005, and the SICE Control Division Pioneer Award, in 2007.

YUKIHIRO TADOKORO (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees in information electronics engineering from Nagoya University, Aichi, Japan, in 2000, 2002, and 2005, respectively. Since 2006, he has been with Toyota Central R&D Labs., Inc., Japan. He was a Research Scholar with the Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA, from 2011 to 2012, to study nonlinear phenomena for future applications in signal and information processing fields. Since 2019, he has been a Visiting Professor with the Graduate School of Informatics, Nagoya University. His current research interests include nanoscale wireless communication and noise-related phenomena in nonlinear systems and their applications in vehicles. He is a Senior Member of the Institute of Electronic, Information and Communication Engineers, Japan. He received the IEICE Best Paper Award and the JSAP Outstanding Paper Award in 2020.