Model Reference Gaussian Process Regression

Kim, Hyuntae; Chang, Hamin

Published in:
IEEE Access

DOI:
10.1109/ACCESS.2023.3336421

Publication date:
2023

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 October 2023, accepted 7 November 2023, date of publication 23 November 2023,
date of current version 4 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3336421

RESEARCH ARTICLE

Model Reference Gaussian Process Regression: Data-Driven State Feedback Controller

HYUNTAE KIM[1], (Member, IEEE), AND HAMIN CHANG[1,2], (Member, IEEE)
[1]Department of Electrical and Computer Engineering, Automation and Systems Research Institute (ASRI), Seoul National University, Seoul 08826, South Korea
[2]Bernoulli Institute, University of Groningen, 9747 AG Groningen, The Netherlands

Corresponding author: Hyuntae Kim (htkim@cdsl.kr)

This work was supported by the Artificial Intelligence Research Laboratory for Flight Control funded by the Agency for Defense Development and Defense Acquisition Program Administration under Grant UD230014SD.

ABSTRACT This paper proposes a data-driven state feedback controller that enables reference tracking for nonlinear discrete-time systems. The controller is designed based on the identified inverse model of the system and a given reference model, assuming that the identification of the inverse model is carried out using only the system's state/input measurements. When its results are provided, we present conditions that guarantee a certain level of reference tracking performance, regardless of the identification method employed for the inverse model. Specifically, when Gaussian process regression (GPR) is used as the identification method, we propose sufficient conditions for the required data by applying some lemmas related to identification errors to the aforementioned conditions, ensuring that the Model Reference-GPR (MR-GPR) controller can guarantee a certain level of reference tracking performance. Finally, an example is provided to demonstrate the effectiveness of the MR-GPR controller.

INDEX TERMS Data-driven control, nonlinear system, stability.

I. INTRODUCTION

Gaussian process regression (GPR) [1] is a nonparametric regression technique commonly used in machine learning and robotics, owing to its ability to handle large and complex datasets, integrate prior knowledge, and offer a probabilistic understanding of uncertainty, making it a powerful and versatile tool for regression analysis adopted in both academia and industry [2].

The use of GPR for identifying unknown nonlinear systems using input/output or state data and then designing a model-based controller for the identified model has been extensively studied [3], [4], [5], [6], [7], [8], [9]. Furthermore, the application of GPR for inverse model identification has emerged as a promising approach to controlling complex dynamic systems, such as robotics and mechatronic systems. The concept behind using GPR for inverse model identification is to utilize the inverse model that leads from the current state to the desired next state to learn the control input through GPR. For instance, [2] and [10] present a detailed review of controlling robotic systems using inverse model GPR. The practical efficacy of using GPR for inverse model identification in achieving high precision control has been demonstrated in several studies [11], [12], [13] conducted on robot arms. Additionally, [14] presents a case of using neural networks for inverse model identification and control.

The primary emphasis of earlier studies has revolved around attaining high precision in experimental results and computational speed in real-time. Nevertheless, despite the practical effectiveness of these approaches, there is still a lack of control theoretical understanding of their fundamental principles. This theoretical gap may result in the absence of a stability guarantee within the system, thereby creating the possibility of hazardous scenarios in real-world systems. As such, it is imperative to address this issue to enhance the reliability and safety of control systems.

In this paper, we propose the data-driven inverse model-based controller as a state feedback control by using only state/input measurements of the plant and identifying the inverse model using GPR under a given reference model. We refer to this as model reference Gaussian process regression (MR-GPR) control as in [15], where the output

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

feedback version of MR-GPR controller is firstly proposed while tracking problem was not studied.

The structure of this paper is as follows. In Section II, the stability of the closed-loop system is concerned by employing bounds on the identification performance necessary for tracking control using arbitrary data-driven methods. Section III introduces GPR and several lemmas regarding the posterior variance of the GPR, which can be utilized to measure the regression error as confidence information about the regression result. Finally, based on the proposed lemmas, a stability analysis of the closed-loop system with MR-GPR is given. In Section IV, we offer an example to demonstrate the effectiveness of the MR-GPR controller. Lastly, Section V summarizes the paper.

Notation: For integers n, m, and k, let $0_{n\times m} \in \mathbb{R}^{n\times m}$ and $I_k \in \mathbb{R}^{k\times k}$ be the zero matrix and the identity matrix, respectively. For column vectors a and b, $[a; b]$ denotes $[a^T, b^T]^T$. For a set $\mathcal{A}$, we define the number of elements in the set $\mathcal{A}$ as $|\mathcal{A}|$. For discrete-time vector sequences $y(t)$ and $z(t)$, we define a set

$$\{(y(t), z(t))\}_{t=k}^{k+T} := \{(y(k), z(k)), \cdots, (y(k+T), z(k+T))\}.$$

II. SYSTEM DESCRIPTION AND DATA-DRIVEN INVERSE MODEL-BASED CONTROL

Consider a nonlinear discrete-time system

$$x(t+1) = f(x(t), u(t)), \tag{1}$$

where $t \in \mathbb{Z}$ is the discrete-time index, $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathcal{U} \subset \mathbb{R}^m$ is the input with a compact set $\mathcal{U}$, and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is smooth. Suppose that a reference model is given by

$$x(t+1) = f_R(x(t), t), \tag{2}$$

where $f_R : \mathbb{R}^n \times \mathbb{Z} \to \mathbb{R}^n$ is smooth in its first argument and uniformly bounded in the second argument. Our goal is to construct a (data-driven) state feedback controller with which the following holds for a given error bound $\epsilon > 0$:

$$\|x(t) - x_R(t)\| < \epsilon, \qquad \forall t \geq 0, \tag{3}$$

where x_R is the solution to the reference model (2) with $x_R(0) = x(0)$. Note that some external inputs to the reference model are allowed by the time index in the function f_R (see Example 1). The reference model has our desired stability, performance, robustness and so on, and we suppose that there is a compact and convex set $\mathcal{X}_R \subset \mathbb{R}^n$, which is an operation region of (2), such that $x_R(t) \in \mathcal{X}_R$ for all $t \geq 0$. In addition, we particularly assume that the reference model is contractive in the sense of [16] as follows.

Assumption 1: For a given $\epsilon > 0$, there exist a positive definite matrix $\Theta \in \mathbb{R}^{n\times n}$ and a scalar $\gamma < 1$ such that

$$\left(\frac{\partial f_R}{\partial x}(x,t)\right)^T \Theta \left(\frac{\partial f_R}{\partial x}(x,t)\right) \leq \gamma\Theta, \qquad \forall x \in \mathcal{X}, t \geq 0,$$

where $\mathcal{X} := \{x \in \mathbb{R}^n : \exists z \in \mathcal{X}_R \text{ such that } \|x - z\| \leq \epsilon\}$.

If the reference model is a linear system, then Assumption 1 is nothing but the stability of the linear system. For example, if the reference model is given as $f_R(x, t) = Ax(t)$ for some matrix $A \in \mathbb{R}^{n\times n}$ and Assumption 1 holds, then there exists a positive definite matrix Θ and a scalar $\gamma < 1$ such that $A^T\Theta A \leq \gamma\Theta$ holds. By the relation $A^T\Theta A - \Theta \leq (\gamma - 1)\Theta$, the reference model is stable. Now, to enable model reference control, we impose the following assumption.

Assumption 2: System (1) satisfies the following.

(a) For $x(t), x(t+1) \in \mathbb{R}^n$ obtained from system (1), an input $u(t) \in \mathbb{R}^m$ that satisfies $x(t+1) = f(x(t), u(t))$ is unique.

(b) For each $x \in \mathcal{X}$ and $t \geq 0$, there exists $u \in \mathbb{R}^m$ such that

$$f_R(x, t) = f(x, u).$$

From Assumption 2-(a), there exists a function $c : \mathbb{R}^{2n} \to \mathbb{R}^m$ such that

$$x(t+1) = f(x(t), c([x(t); x(t+1)]))$$

for all $x(t)$, $x(t+1)$, and $t \geq 0$. Furthermore, if system (1) is input-affine, i.e., given by

$$x(t+1) = f(x(t), u(t)) = a(x(t)) + b(x(t))u(t),$$

where $a : \mathbb{R}^n \to \mathbb{R}^n$ and $b : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are smooth, and if $b(x)$ has full column rank for all $x \in \mathbb{R}^n$, then Assumption 2-(a) holds with the smooth function c given by

$$c([x(t); x(t+1)]) = b^\dagger(x(t))(x(t+1) - a(x(t))), \tag{4}$$

where $b^\dagger(x(t))$ is a left-inverse of $b(x(t))$. In addition, if the reference model f_R is chosen to satisfy

$$f_R(x, t) - a(x) \in \mathrm{im}(b(x)), \quad \forall x \in \mathcal{X}, t \geq 0, \tag{5}$$

where $\mathrm{im}(b(x))$ represents the image of the matrix $b(x)$, then Assumption 2-(b) also holds.

It is trivial that under Assumption 2, the closed-loop system (1) with a state feedback control input

$$u(t) = c([x(t); f_R(x(t), t)]) \tag{6}$$

becomes (2) for all $t \geq 0$. However, in order to find a data-driven construction of feedback controller (6), we need to identify the function c and design a reference model f_R without the complete knowledge of f. The following example shows that even if f is not completely known, one can design a reference model f_R such that Assumption 2-(b) holds.

Example 1: Consider a system given in the controller canonical form [17]:

$$x^+ = \begin{bmatrix} x_1^+ \\ x_2^+ \\ \vdots \\ x_{n-1}^+ \\ x_n^+ \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \\ a_n(x) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b_n(x) \end{bmatrix} u,$$

where x^+, x, and u imply $x(t+1)$, $x(t)$, and $u(t)$ respectively, $a_n : \mathbb{R}^n \to \mathbb{R}$ and $b_n : \mathbb{R}^n \to \mathbb{R}$ are unknown but smooth, and

$u \in \mathbb{R}$. Suppose that $b_n(x) \neq 0$ for all $x \in \mathbb{R}^n$. Then by (4), Assumption 2-(a) holds with the (unknown) smooth function

$$c([x; x^+]) = \frac{x_n^+ - a_n(x)}{b_n(x)}. \tag{7}$$

Furthermore, (5) (and thus, Assumption 2-(b)) is satisfied as long as the reference model has the form:

$$f_R(x,t) = \begin{bmatrix} x_2 \\ \vdots \\ x_n \\ f_n(x,t) \end{bmatrix}, \tag{8}$$

where $f_n(x,t)$ is an arbitrary smooth function. Although a_n and b_n are unknown, the smooth function $f_n(x,t)$ can be designed depending on the user's control goal. For example, suppose that we want the state $x_1(t)$ to asymptotically track a given signal $\phi(t)$ by applying the control input $u = c([x; f_R(x,t)])$ of (6). Then, one can simply take

$$f_n(x,t) = -\sum_{i=1}^{n} k_i(x_i(t) - \phi(t+i-1)) + \phi(t+n)$$

in which, the state feedback gains k_i are chosen such that the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -k_1 & -k_2 & -k_3 & \cdots & -k_n \end{bmatrix}$$

is Schur stable whose eigenvalues are located to meet the convergence performance. (One can verify this using $\tilde{x}_i(t) := x_i(t) - \phi(t+i-1)$, $i = 1, \ldots, n$, because they yield a stable linear system $\tilde{x}^+ = A\tilde{x}$.) □

Given a reference model that satisfies Assumptions 1 and 2-(b), let $\hat{c} : \mathbb{R}^{2n} \to \mathbb{R}^m$ be a data-driven identification of the function c. This means that the function $\hat{c}$ is obtained by using only available state/input measurements of system (1) with some identification method on c. Before specifying the identification method, let us inspect a sufficient condition on the function $\hat{c}$ such that the closed-loop system (1) with a controller $u(t) = \hat{c}([x(t); f_R(x(t),t)])$ satisfies (3). Let

$$B := \max_{x \in \mathcal{X}, t \geq 0} \|(\partial f_R)/(\partial x)(x,t)\|$$

which is well-defined because $\mathcal{X}$ is compact and f_R is smooth in the first argument and uniformly bounded in the second argument. Also, let $\lambda_{\min}$ and $\lambda_{\max}$ be the minimum and the maximum eigenvalues of the positive definite matrix Θ in Assumption 1, respectively. Finally, let L_f be a Lipschitz constant of f such that

$$\|f(x,u_a) - f(x,u_b)\| \leq L_f \|u_a - u_b\|, \quad \forall x \in \mathcal{X}, u_a, u_b \in \mathcal{U}$$

which is well-defined due to smoothness of f and compactness of $\mathcal{X}$ and $\mathcal{U}$. Defining the set $\mathcal{C} := \mathcal{X} \times \mathcal{X}$, we present the following theorem.

Theorem 1: Under Assumptions 1 and 2, the closed-loop system (1) with a controller $u(t) = \hat{c}([x(t); f_R(x(t),t)])$ satisfies

$$\|x(t) - x_R(t)\| < \epsilon, \qquad \forall t \geq 0$$

where $x(0) \in \mathcal{X}_R$ and x_R is the solution to the reference model (2) with $x_R(0) = x(0)$, if

$$\|\hat{c}(\xi) - c(\xi)\| \leq M, \qquad \forall \xi \in \mathcal{C},$$

where $\hat{c}(\xi), c(\xi) \in \mathcal{U}$, with $M > 0$ such that

$$\frac{2L_{fM}\lambda_{\max}B\epsilon + L_{f^2}M^2\lambda_{\max}}{\lambda_{\min}(1-\gamma)} < \epsilon^2. \tag{9}$$

Proof: First, leveraging the methodology presented in [18, Appendix A], we delineate the function $g(s)$, characterized by a variable $s \in \mathbb{R}$, as illustrated below:

$$\begin{aligned} g(s) :&= (f_R(x,t) - f_R(x_R,t))^T \Theta \\ &\quad \times (f_R(sx + (1-s)x_R, t) - f_R(x_R,t)). \end{aligned}$$

Owing to the continuous differentiability of the function $g(s)$ within the real domain, we can apply the mean-value theorem to affirm the existence of a scalar $\bar{s}$ encompassed in the interval $[0,1]$ such that the equation $g(1) - g(0) = g'(\bar{s})$ is satisfied. This can further be reformulated to:

$$\begin{aligned} &(f_R(x,t) - f_R(x_R,t))^T \Theta (f_R(x,t) - f_R(x_R,t)) \\ &\quad = (f_R(x,t) - f_R(x_R,t))^T \Theta \left(\frac{\partial f_R}{\partial x}(w,t)\right)(x - x_R), \end{aligned}$$

where $w = \bar{s}x + (1-\bar{s})x_R$. Capitalizing on this relation, we derive the subsequent inequality:

$$\begin{aligned} &(f_R(x,t) - f_R(x_R,t))^T \Theta (f_R(x,t) - f_R(x_R,t)) \\ &\quad \leq (x - x_R)^T \left(\frac{\partial f_R}{\partial x}(w,t)\right)^T \Theta \left(\frac{\partial f_R}{\partial x}(w,t)\right)(x - x_R). \end{aligned}$$

Now, let

$$V(x, x_R) = (x - x_R)^T \Theta (x - x_R).$$

Then, since

$$x^+ = f(x, \hat{c}) = f_R(x,t) + [f(x,\hat{c}) - f(x,c)],$$

it follows, as long as $x \in \mathcal{X}$ and $x_R \in \mathcal{X}_R$, that

$$\begin{aligned} V^+ &= \big(f_R(x,t) - f_R(x_R,t) + [f(x,\hat{c}) - f(x,c)]\big)^T \Theta \\ &\quad \times \big(f_R(x,t) - f_R(x_R,t) + [f(x,\hat{c}) - f(x,c)]\big) \\ &\leq (x - x_R)^T \left(\frac{\partial f_R}{\partial x}(w,t)\right)^T \Theta \left(\frac{\partial f_R}{\partial x}(w,t)\right)(x - x_R) \\ &\quad + 2L_f M \lambda_{\max} B \|x - x_R\| + L_f^2 M^2 \lambda_{\max} \\ &\leq \gamma V + 2L_f M \lambda_{\max} B \|x - x_R\| + L_f^2 M^2 \lambda_{\max}. \end{aligned} \tag{10}$$

We conclude the proof with mathematical induction. Since $x(0) = x_R(0) \in \mathcal{X}_R$ when $t = 0$, the above inequality implies that

$$\|x(1) - x_R(1)\|^2 \leq V(1)/\lambda_{\min} \leq L_f^2 M^2 \lambda_{\max}/\lambda_{\min} < \epsilon^2$$

in which, the last inequality follows from (9). Since $x_R(1) \in \mathcal{X}_R$, it also follows that $x(1) \in \mathcal{X}$ by the definition of $\mathcal{X}$. Now, suppose that $\|x(t) - x_R(t)\| < \epsilon$ for $t = 0, 1, \cdots, k-1$ (so that $x(t) \in \mathcal{X}$ for $t = 0, 1, \ldots, k-1$). With

$$\beta := 2L_f M \lambda_{\max} B\epsilon + L_f^2 M^2 \lambda_{\max}$$

for convenience, we can repeatedly apply (10) to obtain

$$V(k) \leq \gamma V(k-1) + \beta \leq \gamma(\gamma V(k-2) + \beta) + \beta$$
$$\leq \cdots \leq \gamma^k V(0) + \beta \frac{1-\gamma^k}{1-\gamma} \leq \frac{\beta}{1-\gamma}.$$

Then, we achieve with (9) that

$$\|x(k) - x_R(k)\|^2 \leq \frac{V(k)}{\lambda_{\min}} \leq \frac{\beta}{\lambda_{\min}(1-\gamma)} < \epsilon^2.$$

As a result, $\|x(k) - x_R(k)\| < \epsilon$ and $x(k) \in \mathcal{X}$, which completes the proof. ∎

Theorem 1 simply shows that if we have sufficiently accurate identification of the function c, it can directly be used for designing data-driven version of controller (6) to guarantee sufficient performance of model reference control (3). In the next section, we utilize Gaussian process regression (GPR) as an identification method to obtain $\hat{c}$ and present how to collect state/input data to achieve sufficiently accurate identification result.

III. MODEL REFERENCE-GAUSSIAN PROCESS REGRESSION

In this section, we propose a controller generated by GPR using state/input data from system (1). The resulting data-driven controller can generate control inputs that emulate those produced by (6).

We use GPR to identify the function c in (6) itself and to do so, we additionally assume that the function c is infinitely differentiable in the set $\mathcal{C}$. This approach involves treating $[x(t); x(t+1)]$ as input data and $u(t)$ as output data for the function c, based on the relation $u(t) = c([x(t); x(t+1)])$ from Assumption 2-(a). For GPR, we split the function $c(\cdot)$ as $c(\cdot) =: [c_1(\cdot); c_2(\cdot); \cdots; c_m(\cdot)]$, where $c_i : \mathbb{R}^{2n} \to \mathbb{R}$.

To perform the proposal, we first collect state/input data[1] of system (1) as

$$\{(x_{\mathsf{d}}(t), u_{\mathsf{d}}(t))\}_{t=1}^{N} \tag{11}$$

where N is the total number of state/input data. Then we rearrange the data as the training input

$$\xi_{\mathsf{d}}(t) := [x_{\mathsf{d}}(t); x_{\mathsf{d}}(t+1)] \in \mathbb{R}^{2n}$$

and the training output

$$u_{\mathsf{d}}(t) \in \mathbb{R}^m,$$

yielding the training dataset:

$$\mathcal{D}_N := \{(\xi_{\mathsf{d}}(t), u_{\mathsf{d}}(t))\}_{t=1}^{N-1}. \tag{12}$$

For the training dataset $\mathcal{D}_N$, we define the set of training input as

$$\mathcal{I}_{\mathcal{D}_N} := \{\xi_{\mathsf{d}}(t)\}_{t=1}^{N-1} \subset \mathbb{R}^{2n}.$$

Collecting sufficiently long state/input data as (12) for an unstable system may present a challenge. In such cases, combining state/input data from each short experiment with varying initial conditions may be more practical. For further details, refer to [15, Remark 1].

A Gaussian process (GP) is uniquely characterized by a mean function $m_i : \mathcal{C} \to \mathbb{R}$ and a covariance function $k_i : \mathcal{C} \times \mathcal{C} \to \mathbb{R}$ for $i = 1, \cdots, m$. To identify each function c_i for $i = 1, \cdots, m$, we utilize the GP with the zero function for the mean function and a squared exponential (SE) kernel for the covariance function:

$$k_i(\xi, \xi') = \sigma_i^2 \exp\left(-\frac{1}{2}(\xi - \xi')^T H_i^{-1}(\xi - \xi')\right), \tag{13}$$

where σ_i and $H_i = \mathrm{diag}(h_{i,1}^2, \ldots, h_{i,2n}^2)$ are hyperparameters for $i = 1, \cdots, m$. The hyperparameters are determined through marginal likelihood optimization, a technique based on Bayesian principles [1, Chapter 5].

Remark 1: In this study, regression analyses were performed utilizing the SE kernel. This choice was predicated on the SE kernel's generally robust performance, which is particularly beneficial given our lack of specific information regarding the inherent properties of the system under consideration. However, in circumstances where some characteristics of the function are known, it might be more appropriate to employ different kernels tailored to the situation to achieve more accurate and insightful regression results.

For instance, the selection of alternative kernels could be informed by the specific features of the function we wish to approximate. The periodical kernel could be a choice when modeling functions with recurring patterns. The linear kernel might be suitable for approximating functions that demonstrate a linear trend. In scenarios where the underlying function exhibits more complex behaviors, the Matern or rational quadratic kernels, known for their flexibility and capability to model functions with varying smoothness and scales, might offer a more refined modeling.

Given the training dataset $\mathcal{D}_N$ defined in (12), the GP produces posterior mean and variance functions for a test input $\xi \in \mathcal{C}$ as

$$\mu^i_{\mathcal{D}_N}(\xi) := \mathbf{k}_i^T(\xi)\mathbf{K}_i^{-1}\mathbf{u}^i, \tag{14}$$
$$\sigma^i_{\mathcal{D}_N}(\xi) := k_i(\xi, \xi) - \mathbf{k}_i^T(\xi)\mathbf{K}_i^{-1}\mathbf{k}_i(\xi), \tag{15}$$

respectively, where

$$\mathbf{u}^i := [u^i_{\mathsf{d}}(1); \cdots; u^i_{\mathsf{d}}(N-1)],$$
$$\mathbf{k}_i(\xi) := [k_i(\xi_{\mathsf{d}}(1), \xi); \cdots; k_i(\xi_{\mathsf{d}}(N-1), \xi)],$$
$$\mathbf{K}_i := \begin{bmatrix} k_i(\xi_{\mathsf{d}}(1), \xi_{\mathsf{d}}(1)) & \cdots & k_i(\xi_{\mathsf{d}}(1), \xi_{\mathsf{d}}(N-1)) \\ \vdots & \ddots & \vdots \\ k_i(\xi_{\mathsf{d}}(N-1), \xi_{\mathsf{d}}(1)) & \cdots & k_i(\xi_{\mathsf{d}}(N-1), \xi_{\mathsf{d}}(N-1)) \end{bmatrix}$$

for $i = 1, \cdots, m$, where $u_{\mathsf{d}}(t) =: [u^1_{\mathsf{d}}(t); u^2_{\mathsf{d}}(t); \cdots; u^m_{\mathsf{d}}(t)]$. It is noted that each posterior mean function $\mu^i_{\mathcal{D}_N}$ is in fact the estimation of the function c_i using only the state/input data of system (1). The confidence of this estimation is represented by the posterior variance function $\sigma^i_{\mathcal{D}_N}$.

[1]The subscript d is employed to represent the data gathered from the system during an experiment.

Finally, with the reference model f_R, we construct the model reference GPR (MR-GPR) controller by combining posterior mean functions $\mu^i_{\mathcal{D}_N}$ to $\mu_{\mathcal{D}_N}$ as

$$u(t) = \mu_{\mathcal{D}_N}([x(t); f_R(x(t), t)]) = \begin{bmatrix} \mu^1_{\mathcal{D}_N}([x(t); f_R(x(t), t)]) \\ \vdots \\ \mu^m_{\mathcal{D}_N}([x(t); f_R(x(t), t)]) \end{bmatrix} \tag{16}$$

which is a state feedback controller. We also combine the posterior variance functions $\sigma^i_{\mathcal{D}_N}$ to $\sigma_{\mathcal{D}_N}$ as

$$\sigma_{\mathcal{D}_N}([x(t); f_R(x(t), t)]) = \begin{bmatrix} \sigma^1_{\mathcal{D}_N}([x(t); f_R(x(t), t)]) \\ \vdots \\ \sigma^m_{\mathcal{D}_N}([x(t); f_R(x(t), t)]) \end{bmatrix}. \tag{17}$$

The following lemma describes the difference between the ideal controller c in (6) and the MR-GPR controller $\mu_{\mathcal{D}_N}$ in (16) using the posterior variance function $\sigma_{\mathcal{D}_N}$ in (17).

Lemma 1: [19, Proposition 2] Under Assumption 2, for the training dataset $\mathcal{D}_N$, it holds that

$$||\mu_{\mathcal{D}_N}(\xi) - c(\xi)|| \leq ||c||_k \cdot ||\sigma_{\mathcal{D}_N}(\xi)||, \quad \forall \xi \in \mathcal{C},$$

where $||c||_k := \max(||c_1||_{k_1}, \cdots, ||c_m||_{k_m})$ and $||c_i||_{k_i}$ is the RKHS norm [20] of the function c_i under the kernel k_i. □

Note that the upper bound of the identification error can be reduced arbitrarily by sufficiently reducing the posterior variance $\sigma_{\mathcal{D}_N}(\xi)$.

Before going into the next lemma, we define the set of the training inputs in the closed ball centered at ξ with radius $\rho > 0$ as $\mathbb{B}_\rho(\xi) = \{\xi' \in \mathcal{I}_{\mathcal{D}_N} : ||\xi' - \xi|| \leq \rho\}$. Also, we define a Lipschitz constant of the kernel k_i as L_{k_i}.

Lemma 2: [21, Corollary 3.2] Under Assumption 2, for the training dataset $\mathcal{D}_N$ as a function of N, if there exists a function $\rho(N)$ such that

$$0 < \rho(N) \leq \min\left(\frac{k_1(\xi,\xi)}{L_{k_1}}, \cdots, \frac{k_m(\xi,\xi)}{L_{k_m}}\right), \quad \forall N \in \mathbb{N}$$
$$\lim_{N\to\infty} \rho(N) = 0$$
$$\lim_{N\to\infty} |\mathbb{B}_{\rho(N)}(\xi)| = \infty$$

for all $\xi \in \mathcal{C}$, then $||\sigma_{\mathcal{D}_N}(\xi)|| \to 0$ as $N \to \infty$ for every $\xi \in \mathcal{C}$. □

The existence of the function ρ may appear to be limited in Lemma 2, but the method of constructing the dataset $\mathcal{D}_N$ to ensure the existence of ρ is described in [21].

By Lemma 2, we can arbitrarily reduce the gap between the ideal controller $c(\xi)$ and the MR-GPR controller $\mu_{\mathcal{D}_N}(\xi)$ by using a sufficiently large and dense training input set $\mathcal{I}_{\mathcal{D}_N}$, such that there always exist sufficiently many training inputs in a ball centered at the point ξ with a radius which approaches to 0 as $N \to \infty$. Finally, combining Lemmas 1 and 2 with Theorem 1, we state the following theorem without proof.

Theorem 2: Under Assumptions 1 and 2, suppose that the training dataset $\mathcal{D}_N$ satisfies the sufficient condition of Lemma 2. Then, given $\epsilon > 0$, there exists $\bar{N} \in \mathbb{N}$, such that for every $N > \bar{N}$, the closed-loop system (1) with the MR-GPR controller $u(t) = \mu_{\mathcal{D}_N}([x(t); f_R(x(t), t)])$ sastisfies

$$\|x(t) - x_R(t)\| < \epsilon, \quad \forall t \geq 0$$

where $x(0) \in \mathcal{X}_R$ and x_R is the solution to the reference model (2) with $x_R(0) = x(0)$. □

In practice, noise is unavoidable particularly in state measurements. Since the MR-GPR controller necessitates state data of the system as a training input to learn the function c, we need a GPR which is able to handle the noise in the training input. In fact, the noise in the training input of a linear function can exactly be treated as a noise in the training output as shown in the following example.

Example 2: Let us consider a linear system

$$x(t+1) = Ax(t) + Bu(t), \tag{18}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $B \in \mathbb{R}^{n\times m}$ is left invertible. It is clear that from (4), the function c can be found as

$$c([x(t); x(t+1)]) = B^\dagger(x(t+1) - Ax(t)), \tag{19}$$

where $B^\dagger$ is a left-inverse of B. Suppose that we have noisy measurements

$$\tilde{x}(t) = x(t) + w_x(t),$$

where w_x is a white noise such that $w_x(t) \sim \mathcal{N}(0, \sigma_0^2 I)$ for every time step t. Then from (18) and (19), we obtain

$$\begin{aligned} u(t) &= B^\dagger(x(t+1) - Ax(t)) \\ &= B^\dagger(\tilde{x}(t+1) - A\tilde{x}(t)) - B^\dagger(w_x(t+1) - Aw_x(t)) \\ &= c([\tilde{x}(t); \tilde{x}(t+1)]) - B^\dagger(w_x(t+1) - Aw_x(t)). \end{aligned}$$

Defining $\tilde{u}(t) := u(t) + w_u(t)$, where

$$w_u(t) := B^\dagger(w_x(t+1) - Aw_x(t)),$$

yields

$$c([\tilde{x}(t); \tilde{x}(t+1)]) = \tilde{u}(t).$$

This indicates that the function c can be identified by using the training input

$$\tilde{\xi}(t) := [\tilde{x}(t); \tilde{x}(t+1)] \in \mathbb{R}^{2n}$$

which can be considered as a noise-free data and the corresponding (noisy) training output

$$\tilde{u}(t) = u(t) + w_u(t) \in \mathbb{R}^m,$$

where $w_u(t) \sim \mathcal{N}(0, \sigma_0^2 B^\dagger(I - AA^\top)(B^\dagger)^\top)$. □

Motivated by Example 2, we modify the Gram matrix $\mathbf{K}_i$ of (14) and (15) as if there is an output noise as:

$$\mu^i_{\mathcal{D}_N}(\xi) := \mathbf{k}_i^T(\xi)(\mathbf{K}_i + \sigma_{i,n} I)^{-1}\mathbf{u}^i,$$
$$\sigma^i_{\mathcal{D}_N}(\xi) := k_i(\xi,\xi) - \mathbf{k}_i^T(\xi)(\mathbf{K}_i + \sigma_{i,n} I)^{-1}\mathbf{k}_i(\xi),$$

where $\sigma_{i,n}$ is the additional hyperparameter.

Remark 2: The idea in Example 2, treating the input as deterministic and adding a corrective term to the output noise, can also be applied approximately for nonlinear regression target as shown in [24] and [25]. For instance, the authors of [24] utilize a Taylor expansion approximation to obtain a corrective term in the posterior mean and variance functions. The corrective term is firstly determined by using the standard GPR and improved recursively by re-optimizing hyperparameters based on the (previously corrected) posterior mean function.

Remark 3: We discuss the performance of the MR-GPR controller under conditions where Assumption 2 does not hold.

Firstly, let us consider the scenario where Assumption 2-(a) is violated. This situation implies that the uniqueness of the control input transitioning from the current state to the next state is not guaranteed. Under such circumstances, it is hard to define the inverse model c, thereby the MR-GPR controller may become inoperative.

Secondly, let us inspect the case where Assumption 2-(a) holds, but Assumption 2-(b) is not satisfied, that is, when there is no control input that enables the system to follow the reference model. In this scenario, the MR-GPR controller might fail to compute the appropriate control inputs. However, unlike the case where Assumption 2-(a) is violated, it is feasible to conduct control operations by seeking another reference model that satisfies Assumption 2-(b) based on the confidence level of the GPR. This adaptability underlines the potential for utilizing alternative reference models to ensure effective control, even when Assumption 2-(b) is not fully met.

IV. ILLUSTRATIVE EXAMPLE

In this section, an illustrative example is presented to describe the utility of the proposed data-driven controller. Consider an inverted pendulum system discretized by Euler method [22], [23] as

$$\begin{aligned} z_1(t+1) &= z_2(t), \\ z_2(t+1) &= a(z(t)) + \frac{T^2}{ml^2}u(t), \end{aligned} \tag{20}$$

where $a(z) := z_2 + \frac{gT^2}{l}\sin(z_1) + (1 - \frac{\mu T}{ml^2})(z_2 - z_1)$, and g is gravitational constant, l is the distance from the base to the center of mass of the balanced body, μ is the coefficient of rotational friction, m is the mass to be balanced, and T is a sampling period. Following Example 1, it is obvious that

$$c([z(t); z(t+1)]) = \frac{ml^2}{T^2}(z_2(t+1) - a(z(t))) \tag{21}$$

by (7). With a given reference signal $\phi(t)$, the reference model f_R is chosen following (8), namely,

$$f_R(z(t), t) = \begin{bmatrix} z_2(t) \\ \phi(t+2) \end{bmatrix}$$

TABLE 1. RMS error of each figure with the signal $\phi(t) = 0.1\sin(0.1t)$.

	Figure 1	Figure 2	Figure 3
$z_1(t)$	1.5721e-05	2.3211e+05	0.2088
$z_2(t)$	1.5721e-05	2.3835e+05	0.2011

and it also satisfies Assumption 1. Now, we design MR-GPR controller

$$u(t) = \mu_{\mathcal{D}_N}([z(t); f_R(z(t), t)]),$$

which is expected to make $z_1(t)$ track the signal $\phi(t)$ under system parameters $g = 9.8, m = l = 0.2, \mu = 0.01$, and sampling period $T = 0.1$.

Let us collect the data for identifying function c in (21). Instead of using only one trajectory of input and state as (11) and (12), we use the following training dataset

$$\mathcal{D}_N = \left\{(\xi_{\mathsf{d}}^i(0), u_{\mathsf{d}}^i(0))\right\}_{i=1}^{N-1},$$

where

$$\xi_{\mathsf{d}}^i(0) = [{z_{\mathsf{d}}}^i(0); {z_{\mathsf{d}}}^i(1)]$$

is collected by the i-th experiment with a random initial condition ${z_{1,\mathsf{d}}}^i(0) \in [-\pi, \pi]$, ${z_{2,\mathsf{d}}}^i(0) = 0$, and a random input $u_{\mathsf{d}}^i(0) \in [-1, 1]$ for $i = 1, \ldots, N-1$. In this simulation, we adopted an alternative definition for $\mathcal{D}_N$ which differs from the definition in (12), but can be justified by the approach presented in [15, Remark 1]. The hyperparameters of the SE kernel in (13) were determined by optimizing the marginal likelihood through the GPML toolbox [26] using the training data $\mathcal{D}_N$.

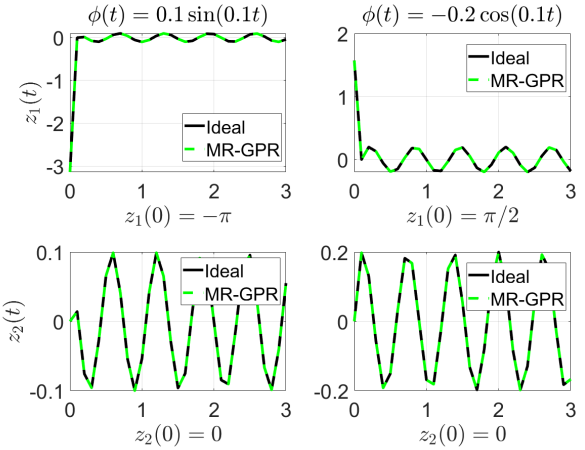

FIGURE 1. State trajectories of the system (20) with ideal controller c (black line) and MR-GPR controller $\mu_{\mathcal{D}_N}$ (green dashed line) designed by the data of $N = 300$ experiments without noise from different initial conditions and different $\phi(t)$. The hyperparameters are $\sigma_1 = 3.8057$ and $h_{1,1}^2 = h_{1,2}^2 = h_{1,3}^2 = h_{1,4}^2 = 0.5688$.

The performance of the MR-GPR controller is compared to that of the ideal controller in the closed-loop system with two different initial conditions (in all cases, we set

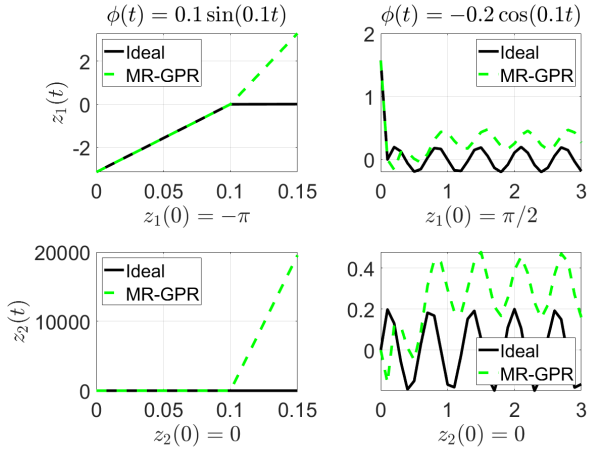

FIGURE 2. State trajectories of the system (20) with ideal controller c (black line) and MR-GPR controller $\mu_{\mathcal{D}_N}$ (green dashed line) designed by the data of $N = 3000$ experiments under the state noise in the state $z(t)$ from different initial conditions. We only draw 0.15s for the left figures because they show unstable system. The hyperparameters are $\sigma_1 = 2.8323$ and $h_{1,1}^2 = h_{1,2}^2 = h_{1,3}^2 = h_{1,4}^2 = 1.0597$.

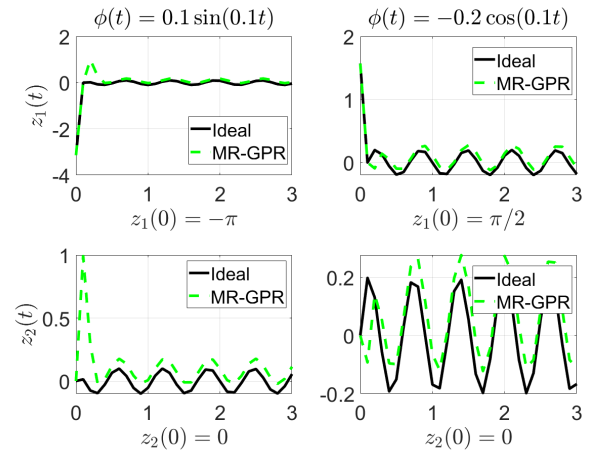

FIGURE 3. State trajectories of the system (20) with ideal controller c (black line) and MR-GPR controller $\mu_{\mathcal{D}_N}$ (green dashed line) designed by the data of $N = 3000$ experiments under the state noise treating it as input noise. The hyperparameters are $\sigma_1 = 2.8323$, $h_{1,1}^2 = h_{1,2}^2 = h_{1,3}^2 = h_{1,4}^2 = 1.0597$, and $\sigma_{1,2} = 0.3$.

the initial conditions of each system to $(z_1(0), z_2(0)) \in \{(-\pi, 0), (\pi/2, 0)\}$, where the MR-GPR controller is trained using $N = 300$ experiments. As shown in Fig. 1, the performance of the MR-GPR controller pretty well follows that of the ideal controller. Additionally, the MR-GPR controller is trained with state data in the presence of additive white Gaussian noise with a signal-to-noise ratio of 20 dB for $N = 3000$ experiments. The result is illustrated in Fig. 2. The figure reveals that the control performance is less-effective under measurement noise on the state (which plays the role of input noise for GPR). Therefore, we implement the method discussed at the end of the previous section. By setting the hyperparameter $\sigma_{1,2} = 0.3$, the result is shown in Fig. 3. It is observed that, in the presence of noise in the state, the proposed MR-GPR controller obtains better result. The root mean square (RMS) error of each experiment is detailed in Table 1 and Table 2.

TABLE 2. RMS error of each figure with the signal $\phi(t) = -0.2\cos(0.1t)$.

	Figure 1	Figure 2	Figure 3
$z_1(t)$	4.0077e-08	0.300	0.0940
$z_2(t)$	4.0773e-08	0.3094	0.0946

V. CONCLUSION

In this study, we introduced a data-driven state feedback controller, referred to as the MR-GPR controller, for nonlinear discrete-time systems under some assumptions. The proposed controller was designed based on the GPR, trained solely on the state/input data of the system, without requiring any prior knowledge of the system's underlying physics or mathematical model. Additionally, the usefulness of the proposed controller was verified through numerical experiments.

REFERENCES

[1] C. K. I. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.
[2] J. Kocijan, *Modelling and Control of Dynamic Systems Using Gaussian Process Models*. Cham, Switzerland: Springer, 2016.
[3] R. Murray-Smith, D. Sbarbaro, C. E. Rasmussen, and A. Girard, "Adaptive, cautious, predictive control with Gaussian process priors," *IFAC Proc. Volumes*, vol. 36, no. 16, pp. 1155–1160, Sep. 2003.
[4] J. Kocijan, R. Murray-Smith, C. E. Rasmussen, and A. Girard, "Gaussian process model based predictive control," in *Proc. Amer. Control Conf.*, Boston, MA, USA, 2004, pp. 2214–2219.
[5] G. Cao, E. M.-K. Lai, and F. Alam, "Gaussian process model predictive control of an unmanned quadrotor," *J. Intell. Robotic Syst.*, vol. 88, no. 1, pp. 147–162, Oct. 2017.
[6] C. J. Ostafew, A. P. Schoellig, T. D. Barfoot, and J. Collier, "Learning-based nonlinear model predictive control to improve vision-based mobile robot path tracking," *J. Field Robot.*, vol. 33, no. 1, pp. 133–152, Jan. 2016.
[7] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using Gaussian process regression," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 6, pp. 2736–2743, Nov. 2020.
[8] J. Kabzan, L. Hewing, A. Liniger, and M. N. Zeilinger, "Learning-based model predictive control for autonomous racing," *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 3363–3370, Oct. 2019.
[9] G. Torrente, E. Kaufmann, P. Fohn, and D. Scaramuzza, "Data-driven MPC for quadrotors," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 3769–3776, Apr. 2021.
[10] J. J. Craig, *Introduction to Robotics: Mechanics and Control*. Upper Saddle River, NJ, USA: Pearson Education International, 2005.
[11] D. Nguyen-Tuong, M. Seeger, and J. Peters, "Computed torque control with nonparametric regression models," in *Proc. Amer. Control Conf.*, Seattle, WA, USA, Jun. 2008, pp. 1–6.
[12] D. Nguyen-Tuong, M. Seeger, and J. Peters, "Model learning with local Gaussian process regression," *Adv. Robot.*, vol. 23, no. 15, pp. 2015–2034, Jan. 2009.
[13] D. Nguyen-Tuong, J. Peters, M. Seeger, and B. Schölkopf, "Learning inverse dynamics: A comparison," in *Proc. Eur. Symp. Artif. Neural Netw.*, Bruges, Belgium, 2008, pp. 13–18.
[14] M. Nørgaard, O. Rvn, N. K. Poulse, and L. K. Hansen, *Neural Networks for Modelling and Control of Dynamic Systems*. London, U.K.: Springer-Verlag London Limited, 2001.
[15] H. Kim, H. Chang, and H. Shim, "Model reference Gaussian process regression: Data-driven output feedback controller," in *Proc. Amer. Control Conf. (ACC)*, San Diego, CA, USA, May 2023, pp. 955–960.

[16] W. Lohmiller and J.-J.-E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, Jun. 1998.
[17] A. Isidori, *Nonlinear Control Systems: An Introduction*. Berlin, Germany: Springer, 1985.
[18] J. Woo Kim, J. Gyu Lee, D. Lee, and H. Shim, "A design method of distributed algorithms via discrete-time blended dynamics theorem," 2022, *arXiv:2210.05142*.
[19] C. Fiedler, C. W. Scherer, and S. Trimpe, "Practical and rigorous uncertainty bounds for Gaussian process regression," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 7439–7447.
[20] M. Kanagawa, P. Hennig, D. Sejdinovic, and B. K. Sriperumbudur, "Gaussian processes and kernel methods: A review on connections and equivalences," 2018, *arXiv:1807.02582*.
[21] A. Lederer, J. Umlauft, and S. Hirche, "Posterior variance analysis of Gaussian processes with application to average learning curves," 2019, *arXiv:1906.01404*.
[22] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 909–924, Mar. 2020.
[23] M. Karam and M. A. Zohdy, "Modeling a simple inverted pendulum using a model-based dynamic recurrent neural network," in *Proc. Southeastern Symp. Syst. Theory*, 2005, pp. 78–82.
[24] A. McHutchon and C. Rasmussen, "Gaussian process training with input noise," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 24, 2011, pp. 1–9.
[25] H. Bijl, T. B. Schön, J.-W. van Wingerden, and M. Verhaegen, "System identification through online sparse Gaussian process regression with input noise," *IFAC J. Syst. Control*, vol. 2, pp. 1–11, Dec. 2017.
[26] C. E. Rasmussen and H. Nickisch, "The GPML toolbox version 4.0," Tech. Document., Tech. Rep., 2016.

HYUNTAE KIM (Member, IEEE) received the B.S. degree in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2014, where he is currently pursuing the Ph.D. degree in systems and control.

His research interests include disturbance observer and data-driven control.

HAMIN CHANG (Member, IEEE) received the B.S. and M.S. degrees in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2019 and 2021, respectively. He is currently pursuing the Ph.D. degree in systems and control with the Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, The Netherlands.

His research interests include disturbance observer, data-driven control, and Gaussian process.