The Power Function for Adaptive Control in Native Space Embedding

Haoran Wang Nathan Powell Andrea L'Afflitto Andrew J. Kurdila John Burns

***Abstract*—This paper studies how the power function in a reproducing kernel Hilbert space (RKHS) can be used systematically to design error bounding methods of adaptive estimation and control via the native space embedding method. The approach is based on viewing the original system of ordinary differential equations (ODEs) as a type of distributed parameter system (DPS), and subsequently defining realizable controllers by approximating the DPS with scattered bases over a domain of interest. The approach provides rigorous bounds on ultimate performance guarantees for uncertainty classes defined in the native space. One result derives an upper bound on the ultimate performance of the adaptive controller in terms of the power function. Another version of this upper bound shows how the ultimate performance can bounded in terms of a fill distance of centers of approximation in subsets that contain the closed loop trajectory. In contrast to the general theory for error bounding adaptive controllers in Euclidean space, the general approach in this paper works for functional uncertainties in any RKHS.**

***Index Terms*—Native space, Distributed parameter system, Adaptive control**

I. Introduction

A. Motivation

Over the past decade, a variety of problems associated with machine and statistical learning theory, Bayesian estimation, and Gaussian process estimation have been solved using the theoretical setting of reproducing kernel Hilbert spaces (RKHS). The general study of such problems can be found in the books [1]–[3]. This paper continues the development in the recent collection of papers [4]–[12] that study how well-known techniques from real parametric adaptive estimation and control theory can be lifted to a native space setting.

In real parametric adaptive estimation and control theory [13]–[19], the unknown matched nonlinear dynamics in a governing ordinary differential equation (ODE) are usually represented as the product of an unknown matrix by a regressor vector of dimension N, which is selected *a priori* and meets sufficient regularity conditions to guarantee the existence and uniqueness of solutions of the ODE. Such methods work well when there is enough prior knowledge to construct the regressor vector, and such knowledge is usually deduced from physical modeling [20, Ch. 9] or data-driven techniques such as neural networks [21], [22]. The uniform approximation assumption is critical in these methods [23]. The stability and convergence analysis of estimation and control tasks is then performed for a fixed number N of real parameters. If the original ODE evolves in $\mathbb{R}^d$, then trajectories of the state and parameter estimates evolve in the Euclidean configuration space $\mathbb{R}^d \times \mathbb{R}^N$.

The native space or RKHS embedding method allows casting estimation and control problems in a nonparametric framework, which does not involve an *a priori* restriction to a fixed Euclidean space. Let $f \in \mathcal{H}$ denote a matched uncertainty in the governing ODEs, where $\mathcal{H}$ denotes an RKHS space. In the language of [12], the uncertainty is functional, and the estimation or control problem is considered a simple type of distributed parameter system (DPS). Practical, implementable estimation and control algorithms are obtained by consistent approximation of the DPS, which generates coordinate expressions that satisfy the governing ODEs. The overall approach is referred to as *native space embedding* since the trajectories of the ODEs evolve in the configuration space $\mathbb{R}^d \times \mathbb{R}^N$, or, equivalently, can be viewed as being embedded in the configuration space $\mathbb{R}^d \times \mathcal{H}$ of the DPS. The configuration space of the DPS, and in particular the native space $\mathcal{H}$, is used to characterize the limiting response as $N \to \infty$. A distinct advantage of this setting is that it is possible to derive explicit relations between the performance of the controller or observer and the dimension N, and these bounds hold for quite general functional uncertainty classes defined in the native space $\mathcal{H}$.

References [4]–[12] study how the real parametric gradient learning law, and its robust refinements, such as the dead-zone method [24] or the σ-modification [25] can be lifted to the setting of RKHS embedding. The conclusions of these papers are guarantees on the performance of the adaptive controller or observer. In this paper we show how the power function can be used in a systematic and general way to prove stability and convergence of error bounding methods of adaptive control in the native space embedding method; see [23, Ch. 4, 6] for a good background for error bounding methods for Euclidean or real parametric adaptive control. Again, we obtain bounds that relate ultimate controller performance to the dimension N of the approximates. These bounds hold for all functions in a functional uncertainty class defined in an RKHS, not just for some single finite dimensional approximation or its coordinate expression. Also, while standard theory as described in [23] leaves the definition of the error bounding function to be derived by the designer via a problem-specific analysis, the approach in this paper gives a general definition of the error

H. Wang, N. Powell, and A. Kurdila are with the Department of Mechanical Engineering, Virginia Tech, Blacksburg, VA , USA. A. L'Afflitto is with the Grado Department of Industrial and Systems Engineering, Virginia Tech. A. Burns is with the Department of Mathematics, Virginia Tech. (`[haoran9, nrpowell, a.lafflitto, kurdila, jaburns]@vt.edu`).

bounding function that works *for uncertainty contained in any native space $\mathcal{H}$.*

B. Problem Statement

This paper initially studies a simple, common governing ODE that has matched uncertainty, with

$$\dot{x}(t) = Ax(t) + B(u(t) + E_{x(t)}f), \tag{1}$$

where $x(t) \in \mathbb{R}^d$ is the *state vector*, and $u(t) \in \mathbb{R}$ is the *control input* with $t \geq t_0$. The *system matrix* $A \in \mathbb{R}^{d\times d}$ and the *system matrix* $B \in \mathbb{R}^d$ are known and the pair (A, B) is controllable. The unknown nonlinear scalar-valued function $f \in \mathcal{H}$ is referred to as the *functional uncertainty* in the native space $\mathcal{H}$, and $E_{x(\cdot)}$ denotes the *evaluation operator* at $x(\cdot)$. The control task is to derive an adaptive control strategy to track the *reference trajectory* $x_r(t)$ of the *reference model*

$$\dot{x}_r(t) = A_r x_r(t) + B_r r(t) \tag{2}$$

where the *reference command input* $r : [t_0, \infty) \to \mathbb{R}$ is bounded, $A_r \in \mathbb{R}^{d\times d}$ is Hurwitz, $B_r \in \mathbb{R}^d$ and such that the pair (A_r, B_r) is controllable and the *matching conditions*

$$A_r = A - BK^{\mathrm{T}}, \tag{3}$$

$$B_r = BL \tag{4}$$

are verified for some $K \in \mathbb{R}^d$ and $L \in \mathbb{R}$.

We would like to design an adaptive control strategy, whose performance is guaranteed for all functional uncertainties in the class

$$\mathcal{C}_R := \{f \in H \mid \|f\|_{\mathcal{H}} \leq R\} \subset \mathcal{H}, \tag{5}$$

where $R > 0$. When realizable controllers $u_N(t)$ are derived that depend on the dimension N of finite dimensional subspaces $\mathcal{H}_N \subset \mathcal{H}$ used in the approximation of the functional uncertainty, we seek explicit bounds that relate the performance of the feedback controller to the dimension N.

C. Summary of New Results

There are several new results in this paper. They are captured by Theorem 3.1 and Corollary 3.1, as well as in the analysis of numerical examples in Section IV. The reader is referred to them for precise statements. In this short paper, we omit detailed proofs as they will be contained in a forthcoming journal paper.

Let $\mathfrak{K}(\cdot,\cdot)$ denote the kernel that defines the RKHS space $\mathcal{H}$, and N be the number of scattered centers $\Xi_N \subset S \subseteq \mathbb{R}^d$ used to define a finite-dimensional *space $\mathcal{H}_N \subset \mathcal{H}$ of approximants* of the functional uncertainty $f \in \mathcal{H}$ over a domain of interest $S \subset \mathbb{X} := \mathbb{R}^d$. The domain S is assumed to contain the closed-loop trajectory x. We denote by $\mathfrak{K}_N(\cdot,\cdot)$ the kernel that defines the RKHS $\mathcal{H}_N$, and by $x_N(\cdot)$ the closed-loop trajectory when the adaptive controller is implemented using the N-dimensional space $\mathcal{H}_N$. In Theorem 3.1, we show how, for the approach derived in this paper, for any arbitrarily small constant $\eta > 0$, there exists a finite time $T := T(\eta) > t_0 \geq 0$ such that

$$\|x_N(t) - x_r(t)\|_{\mathbb{R}^d} \leq (1+\eta)\frac{4\|B^{\mathrm{T}}P\|_{\mathbb{R}^d}}{\lambda_{\min}(Q)} \sup_{\xi\in\bar{S}_\varepsilon} \sqrt{\mathfrak{K}(\xi,\xi) - \mathfrak{K}_N(\xi,\xi)}R, \tag{6}$$

for all $t \geq T$ and for all functions in the uncertainty class $\mathcal{C}_R$; the pair of matrices (P, Q) verify the algebraic Lyapunov equation and the set $\bar{S}_\varepsilon$ is defined precisely in Theorem 3.1 below. Corollary 3.1 shows how, for some standard choices of the kernel function $\mathfrak{K}(\cdot,\cdot)$, there exists $T > t_0$ such that

$$\|x_N(t) - x_r(t)\|_{\mathbb{R}^d} \leq O\left(\sqrt{\mathcal{G}(h_{\Xi_N,S})}\right) \text{ for all } t \geq T,$$

where $\mathcal{G} : \mathbb{R}^+ \to \mathbb{R}$ denotes a known function, $h_{\Xi_N,S}$ denotes the fill distance of the centers Ξ_N in S, and $O(\cdot)$ denotes the *big O* in the Bachmann–Landau notation. The *fill distance* is defined as

$$h_{\Xi_N,S} := \sup_{s\in S} \min_{\xi\in\Xi_N} d_S(s, \xi_i) \tag{7}$$

where d_S denotes the metric on S. For instance, if the kernel $\mathfrak{K}(\cdot,\cdot)$ is selected to be the Sobolev-Matern kernel (as used in the numerical examples) with smoothness index r, the above corollary takes the form

$$\|x_N(t) - x_r(t)\|_{\mathbb{R}^d} \leq O(h_{\Xi_N,S}^{r-1/2}) \text{ for all } t \geq T(\eta).$$

Other explicit expressions for the uniform upper bound on the trajectory tracking error are available for the exponential, inverse multiquadric, and compactly supported Wendland functions [26].

A key observation is that classical methods, such as the σ-modification of model reference adaptive control, to name one, employed to regulate a governing ODE in the presence of unknown finite-dimensional uncertainties, can not be employed directly in the proposed framework. Indeed, f in (1) is the element of an infinite-dimensional vector space, whereas classical methods require that a realization of the uncertainty is determined, although such realization is unknown to the user. An additional observation is that the results proposed in this paper do not consider $BE_{x(\cdot)}f$ as an unmatched uncertainty in its entirety. Indeed, the proposed control system perfectly counters the effect of the projection of f onto an RKHS $\mathcal{H}_N$ and bounds the effect of the component of f orthogonal to the same RKHS. In the limit, if f were contained in the chosen RKHS, then asymptotic stability, and not just uniform ultimate boundedness, of the trajectory tracking error would be attained. This is what the behavior of the limiting DPS describes. Considering $BE_{x(\cdot)}f$ as an unmatched uncertainty in its entirety would produce larger ultimate bounds on the trajectory tracking error.

II. Background

A. Collection of Uncommon Symbols

The table below is a collection of important recurring symbols associated with an RKHS that are used in the derivations below.

TABLE I
COLLECTION OF RKHS SYMBOLS

Symbol	Definition
S	Domain of interest in $\mathbb{X}$
$\mathbb{X}$	Subspace of $\mathbb{R}^d$
$\mathcal{H}$	Infinite dimensional native space
$\mathcal{H}_N$	Finite dimensional subspace of $\mathcal{H}$
$\mathfrak{K}(\cdot,\cdot)$	Kernel function that defines $\mathcal{H}$
$\mathfrak{K}_x(\cdot)$	Kernel section or basis centered at x
$\mathfrak{K}_N(\cdot,\cdot)$	Kernel function that defines $\mathcal{H}_N$
Ξ_N	Collection of kernel centers that define $\mathcal{H}_N$
$h_{\Xi_N,S}$	Fill distance of the kernel centers Ξ_N in S
E_x	Evaluation functional
E_x^*	Adjoint operator of E_x
Π_N	Orthogonal projection from $\mathcal{H}$ to $\mathcal{H}_N$
$\mathcal{P}_{\mathcal{H}_N}(x)$	Power function of $\mathcal{H}_N$ evaluated at x
$\mathbb{K}(\Xi_N,\Xi_N)$	Grammian matrix in $\mathbb{R}^{N\times N}$

B. RKHS or Native Spaces

The RKHS embedding method relies on several key properties of reproducing kernel functions $\mathfrak{K}:\mathbb{X}\times\mathbb{X}\to\mathbb{R}$ defined over a domain $\mathbb{X}$. A *kernel section*, or a *kernel function centered at* $x\in\mathbb{X}$, is written as $\mathfrak{K}_x(\cdot):=\mathfrak{K}(x,\cdot)\in\mathcal{H}$. The closed linear span of these functions forms the RKHS $\mathcal{H}=\overline{\text{span}\{\mathfrak{K}_x|x\in\mathbb{X}\}}$. If there are infinitely many $x\in\mathbb{X}$, then the resulting RKHS $\mathcal{H}$ is generally infinite dimensional. The RKHS space can be thought of as the space obtained by superposition of the "template" $\mathfrak{K}_x$ as it moves around the domain. An RKHS space is often referred to as the *native space* generated by the kernel sections $\mathfrak{K}_x$, and $\mathfrak{K}_x$ is often an example of radial basis function centered at $x\in\mathbb{X}$.

The reproducing property of an RKHS states that any $f\in\mathcal{H}$ satisfies

$$f(x)=\langle f,\mathfrak{K}_x\rangle_{\mathcal{H}} \quad \text{for all } x\in\mathbb{X}, \tag{8}$$

where $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ denotes an inner product in $\mathcal{H}$; this property is used in all the derivations associated with RKHS embedding. The reproducing property can be written in terms of the *evaluation functional* $E_x:\mathcal{H}\to\mathbb{R}$, which is defined as $E_xf:=f(x)$ for any $f\in\mathcal{H}$ and $x\in\mathbb{X}$, and its adjoint is $E_x^*:\mathbb{R}\to\mathcal{H}$. The functionals E_x and E_x^* are linear bounded operators. By the reproducing property [27, Def. 1.1] and proceeding as in the proof of Lemma 1.9 of [27], we have

$$\langle E_xf,\alpha\rangle_{\mathbb{R}}=\langle f,E_x^*\alpha\rangle_{\mathcal{H}}=\langle f,\mathfrak{K}_x\alpha\rangle_{\mathcal{H}}=\alpha\langle f,\mathfrak{K}_x\rangle_{\mathcal{H}} \tag{9}$$

for all $\alpha\in\mathbb{R}$. Therefore, it follows from (9) that $E_x^*\alpha:=\mathfrak{K}_x\alpha$ for all $\alpha\in\mathbb{R}$, and this expression in also used repeatedly in the proofs and in the derivation of algorithms.

In general, an RKHS space $\mathcal{H}$ can contain functions that are not bounded. In this paper, we always assume that $\mathfrak{K}(\cdot,\cdot)$ is bounded on the diagonal. This means that there is a constant $\bar{\mathfrak{K}}>0$ such that $\mathfrak{K}(x,x)\le\bar{\mathfrak{K}}^2$ for all $x\in\mathbb{X}$. This condition is sufficient to ensure that all the functions in the native space $\mathcal{H}$ are bounded, and, furthermore, that $\|E_x\|_{\mathcal{H}}\le\bar{\mathfrak{K}}$ for all $x\in\mathbb{X}$. There are many standard kernels that satisfy this assumption including exponential, Sobolev-Matern, inverse multiquadric, and Wendland kernels. [26]

Since practical algorithms require finite-dimensional approximations of the functions in the native space, we use projections or interpolations for this purpose. In particular, we build finite dimensional approximations in terms of kernel sections located at a finite number of centers $\Xi_N=\{\xi_i|1\le i\le N\}\subset\mathbb{X}$. The corresponding space of approximants is then $\mathcal{H}_N=\text{span}\{\mathfrak{K}_{\xi_i}|\xi_i\in\Xi_N\}$. Properties of RKHS ease the task of building approximations using the *orthogonal projection operator* $\Pi_N:\mathcal{H}\to\mathcal{H}_N$. For any $f\in\mathcal{H}$, we have

$$f=\Pi_Nf+(I-\Pi_N)f, \tag{10}$$

and, by definition, it holds that $((I-\Pi_N)f,\mathfrak{K}_{\xi_i})_H=0$ for $\xi_i\in\Xi_N$ and $1\le i\le N$. Thus, it follows from the reproducing property that the projection operator interpolates at the centers in Ξ_N, with

$$f(\xi_i)=(\Pi_Nf)(\xi_i) \quad \text{for all } \xi_i\in\Xi_N.$$

Note that $(I-\Pi_N)f$ captures the approximation error term that will appear in the stability analysis in the later sections. To quantify and establish bounds on the approximation error, the general theory of RKHS defines the power function. Suppose that $U\subseteq\mathcal{H}$ is a closed subspace of $\mathcal{H}$ and Π_U is the $\mathcal{H}$-orthogonal projection from $\mathcal{H}$ onto U. A relevant result of the theory of RKHS is that U is also a native space for the kernel $\mathfrak{K}_U(x,y):=(\Pi_U\mathfrak{K}_x,\Pi_U\mathfrak{K}_y)_H$ [27, Th. 2.5]. The *power function* $\mathcal{P}_{\mathcal{H}_N}:\mathbb{X}\to\mathbb{R}$ is defined as

$$\mathcal{P}_{\mathcal{H}_N}(x):=\sqrt{\mathfrak{K}(x,x)-\mathfrak{K}_N(x,x)} \quad \text{for all } x\in\mathbb{X}, \tag{11}$$

where $\mathfrak{K}_N(x,y):=\Pi_N\mathfrak{K}(x,y)$, $(x,y)\in\mathbb{X}\times\mathbb{X}$. It is not difficult to show that

$$\mathfrak{K}_N(x,y)=\mathfrak{K}_{\Xi_N}^{\mathrm{T}}(x)\mathbb{K}^{-1}(\Xi_N,\Xi_N)\mathfrak{K}_{\Xi_N}(y), \tag{12}$$

where $\mathfrak{K}_{\Xi_N}(x):=[\mathfrak{K}_{\xi_1}(x),\ldots,\mathfrak{K}_{\xi_N}(x)]^{\mathrm{T}}\in\mathbb{R}^N$ and $\mathbb{K}(\Xi_N,\Xi_N):=[\mathfrak{K}(\xi_i,\xi_j)]\in\mathbb{R}^{N\times N}$ denotes the *Grammian matrix*. This matrix can be easily evaluated once a set of centers Ξ_N have been selected.

The importance of the power function is that it provides a convenient pointwise upper bound on the approximation error for any function $f\in\mathcal{H}$ and any $x\in\mathbb{X}$, which is given by

$$\begin{aligned}|E_x(I-\Pi_{\mathcal{H}_N})f|&=|f(x)-(\Pi_{\mathcal{H}_N}f)(x)|\\&\le\mathcal{P}_{\mathcal{H}_N}(x)\|f\|_{\mathcal{H}}.\end{aligned} \tag{13}$$

This upper bound is crucial for many of the stability and convergence proofs derived in this paper. This bound improves on the fidelity of common universal approximation assumptions that are prevalent in adaptive control: see Chapter [13] for a few common such universal approximation theorems.

III. BOUNDING ADAPTIVE CONTROL AND POWER FUNCTIONS

The goal of this paper is to define a model reference control law and an adaptive law that steer the trajectory of the governing ODE (1) toward the reference the trajectory of the reference model (2) despite the marched functional uncertainty $f(\cdot)$. The next theorem provides the first key theoretical result to address this problem by providing both a control law and an adaptive law that assure uniform ultimate boundedness of the trajectory tracking error.

For the statement of the next theorem, let

$$e_N(x) := |E_x(I - \Pi_N)f| \tag{14}$$

denote the *pointwise approximation error* for each $x \in \mathbb{X}$, and note that

$$e_N(x) \leq \mathcal{P}_{\mathcal{H}_N}(x)\|f\|_{\mathcal{H}} \text{ for all } x \in \mathbb{X}, \tag{15}$$

where $\mathcal{P}_{\mathcal{H}_N}(x)$ is given by (11). Thus, let $\bar{e}_N(x) := \mathcal{P}_{\mathcal{H}_N}(x)R$ denote a known *pointwise approximation error bound*, since $e_N(x) \leq \bar{e}_N(x)$ for all $x \in \mathbb{X}$. Additionally, let $P \in \mathbb{R}^{d\times d}$ denote the symmetric, positive-definite solution to the *Lyapunov equation*

$$A_r^{\mathrm{T}}P + PA_r = -Q \tag{16}$$

with $Q \in \mathbb{R}^{d\times d}$ user-defined, symmetric, and positive definite. Let $\bar{S}_\varepsilon := \{\tilde{x}_N \in \mathbb{X} : |B^{\mathrm{T}}P\tilde{x}_N| < \varepsilon\}$, where $\varepsilon \in \mathbb{R}^+$ is a small constant, $\tilde{x}_N(t) := x_N(t) - x_r(t)$, $t \geq t_0$, denotes the *tracking error* of the system, $x_N(\cdot)$ verifies the governing ODE (1) with *adaptive control input*

$$\begin{aligned} u_N(t) = &- K^{\mathrm{T}}x_N(t) + Lr(t) \\ &- E_{x_N(t)}\left(\hat{f}_N(t,\cdot) + v_N(\cdot,\tilde{x}_N(t))\right), \end{aligned} \tag{17}$$

$K \in \mathbb{R}^d$ and $L \in \mathbb{R}$ verify the matching conditions (3) and (4), $\hat{f}_N(t,\cdot)$ denotes a finite-dimensional estimate of the functional uncertainty f in (1) and verifies the *adaptive law*

$$\frac{\partial \hat{f}_N(t,\cdot)}{\partial t} = \gamma \Pi_N E^*_{x_N(t)} B^{\mathrm{T}}P\tilde{x}_N(t), \tag{18}$$

$\gamma > 0$ denotes the *adaptive rate*, and

$$v_N(\cdot,\tilde{x}_N(t)) := \begin{cases} \mathrm{sign}(B^{\mathrm{T}}P\tilde{x}_N(t))\bar{e}_N(\cdot) & \text{if } |B^{\mathrm{T}}P\tilde{x}_N(t)| \geq \varepsilon, \\ \frac{1}{\varepsilon}B^{\mathrm{T}}P\tilde{x}_N(t)\bar{e}_N(\cdot) & \text{otherwise,} \end{cases} \tag{19}$$

denotes the *compensator*; it is worthwhile noting that $v_N(\cdot,\cdot)$ is continuous in both arguments.

Also note that the adaptive law in (18) is a partial differential equation (PDE). It is an approximation of the learning law

$$\frac{\partial \hat{f}(t,\cdot)}{\partial t} = \gamma E^*_{x_N(t)} B^{\mathrm{T}}P\tilde{x}(t),$$

which we refer to as the *nonparametric gradient learning law.* This nonparametric form of the learning law is also a PDE. We define the associated *nonparametric feedback control* $u(t) = -K^T x(t) + Lr(t) - E_{x(t)}\hat{f}(t,\cdot)$, $t \geq t_0$. When we use the nonparametric control law and nonparametric gradient learning law, the resulting closed loop system has a state $(x(t), \hat{f}(t,\cdot))$ that evolves in $\mathbb{X}\times\mathcal{H}$, and this system defines the *limiting DPS*. Note in general that the nonparametric feedback control law cannot be implemented in practice since $\hat{f}(t,\cdot)$ resides in an infinite dimensional native space $\mathcal{H}$. In general, the question of existence and uniqueness of solutions of the limiting DPS, and the convergence of solutions of consistent approximations to the limiting system, requires a rather lengthy proof that far exceeds the confines of this short conference paper. Throughout the rest of this paper we assume that the limiting DPS is forward complete and approximations of the limiting equations are also forward complete. We leave the associated proofs of existence and uniqueness to the full length journal article.

Theorem 3.1: Consider the governing ODE (1) and the reference model (2). Let $\mathfrak{K}(\cdot,\cdot)$ be an admissible kernel that defines the native space $\mathcal{H}(S)$ over the set $S \subset \mathbb{X} := \mathbb{R}^d$ such that $\mathfrak{K}(x,x) \leq \bar{\mathfrak{K}}^2$, where $\bar{\mathfrak{K}} > 0$ is constant. Also, assume that the matching conditions (3) and (4) are verified and that the functional uncertainty f in (1) is such that $f \in \mathcal{C}_R$ for some $R > 0$, where $\mathcal{C}_R$ is given by (5). Then, for any (arbitrarily small) constant $\eta > 0$ there exists a time $T := T(\eta) > t_0$ such that (6) is verified with kernel $\mathfrak{K}_N(x,x)$ given by (12) and induced approximating finite-dimensional subspace $\mathcal{H}_N$.

Theorem 3.1 provides an estimate of the trajectory tracking error's uniform ultimate bound, which is captured by (6). This ultimate bound is a function of the uncertainty on the uncertainties on f, which is captured by R, the difference between the kernel function $\mathfrak{K}(\cdot,\cdot)$ and its approximation $\mathfrak{K}_N(\cdot,\cdot)$, and the user-defined parameter $\lambda_{\min}(Q)$. In general, larger values of $\lambda_{\min}(Q)$ produce slower convergence of the tracking error. An approach to employ larger values of $\lambda_{\min}(Q)$ and retain fast convergence of the trajectory tracking error involves the use of two-layer model reference adaptive control [28]; this approach will be investigated in the future.

It is worthwhile to note how the proof of Theorem 3.1 follows along the lines of classical arguments that apply to finite-dimensional dynamical systems, such as, for instance, those in [29, p. 170], and extends these results to infinite-dimensional systems. Theorem 3.1 does not address a key point, namely the boundedness of $\tilde{x}(\cdot)$, $\tilde{x}_N(\cdot)$, and $\hat{f}_N(\cdot)$. This result can be obtained by applying Theorem 1.4 of Chapter 6 of [30], and it has been discussed in detail in [31].

The next result specializes Theorem 3.1 to classical kernel functions $\mathfrak{K}$. For the statement of this result, consider the kernel functions in Table II, the corresponding function $\mathcal{G} : \mathbb{R}^+ \to \mathbb{R}$, and the *fill distance* (7) with Ξ_N denoting any set of basis centers. To state the next result, define the *approximation error*

$$\begin{aligned} e_N(x) &:= |E_x(I - \Pi_N)f| \\ &\leq \mathcal{P}_{\mathcal{H}_N}(x)\|f\|_{\mathcal{H}} \\ &\leq \bar{e}_N(x) \\ &:= \mathcal{P}_N(x)R \text{ for all } x \in \mathbb{X}, \end{aligned} \tag{20}$$

where $\bar{e}_N(x)$ captures a known pointwise approximation error bound, which is defined in terms of the power function $\mathcal{P}_{\mathcal{H}_N}(x)$.

Corollary 3.1: Consider the governing ODE (1) and the reference model (2). Let $\mathfrak{K}(\cdot,\cdot)$ be an admissible kernel that defines the native space $\mathcal{H}(S)$ over the set $S \subset \mathbb{X}$ chosen from Table II. Also, assume that the matching conditions (3) and

TABLE II
EXAMPLES OF KERNEL BASIS FUNCTIONS AND CORRESPONDING UPPER BOUNDS [26, CH. 11], [32] TABLE 1.

Kernel name	$\mathfrak{K}(x,0)=\mathfrak{K}_0(r)$, $r=\|x\|_{\mathbb{R}^d}$	$\mathcal{G}(h)$
Gaussian	$e^{-\alpha r^2}, \quad \alpha>0$	$e^{-c\|\log(h)\|/h}$
Inverse MQs	$(\alpha^2+r^2)^\beta, \quad \alpha>0, \beta<0$	$e^{-c/h}$
Compactly supported functions (a.k.a. Wendland)	$\phi_{d,k}(r)$	h^{2k+1}
Sobolev-Matern	$K_{k-d/2}(r)*(r/2)^{k-d/2}$	h^{2k-d}

(4) are verified. Then, there exists a finite time $T>t_0$ such that the closed-loop trajectory tracking error satisfies

$$\|x_N(t)-x_r(t)\|_{\mathbb{R}^d} \le O\left(\sqrt{\mathcal{G}(h_{\Xi_N,S})}\right), \qquad \text{for all } t\ge T. \tag{21}$$

Proof: Consider the kernel functions listed in Table II, which are deduced from Table 11.1 of [26] and Table 1 of [32]. For the Gaussian, inverse multi-quadratic, Wendland kernel functions, and Sobolev-Matern kernels the Table II provides a function $\mathcal{G}$ such that

$$\sup_{\xi\in S}\mathcal{P}_{\mathcal{H}_N}(\xi)\le O\left(\sqrt{\mathcal{G}(h_{\Xi_N,S})}\right).$$

It follows from Theorem 3.1 that each $\eta>0$, there exists $C(\eta)>0$ and $T(\eta)>t_0$ such that

$$\begin{aligned}\|x_N(t)-x_r(t)\|_{\mathbb{R}^d} &\le C(\eta)\sup_{\xi\in\bar{S}_\varepsilon}\mathcal{P}_{\mathcal{H}_N}(\xi)\\ &\le \tilde{C}(\eta)\sqrt{\mathcal{G}(h_{\Xi_N,S})}\end{aligned}$$

for all $t\ge T$. Thus, the result follows directly from (6) and (20). □

In this paper, we assume that matrix A and vector B are known. If these system parameters are unknown, then we can apply the model reference adaptive control method of native space embedding introduced in [33].

IV. NUMERICAL EXAMPLES

In this section, we present and analyze in detail the results of a numerical example that proves the applicability of the framework outlined by Theorem 3.1. Let $\hat{f}_N(t,\cdot)\in\mathcal{H}_N$ be given by

$$\hat{f}_N(t,\cdot)=\hat{\Theta}_N^{\mathrm{T}}(t)\mathfrak{K}_{\Xi_N}(\cdot),$$

where $\hat{\Theta}_N(t)\in\mathbb{R}^N$, $t\ge t_0$, denotes the vector of weights at the basis centers. The vector $\mathfrak{K}_{\Xi_N}(x) := [\mathfrak{K}_{\xi_1}(x),\dots,\mathfrak{K}_{\xi_N}(x)]^{\mathrm{T}}\in\mathbb{R}^N$ captures the collection of kernel basis functions located at the centers in Ξ_N. We then recast the adaptive law (18) as

$$\dot{\hat{\Theta}}_N(t)=\gamma_f\mathbb{K}^{-1}(\Xi_N,\Xi_N)\mathfrak{K}_{\Xi_N}(x_N(t))B^{\mathrm{T}}P\tilde{x}_N(t). \tag{22}$$

We emphasize that, for the error bounds captured by (6) and (21) to be realizable and define a consistent approximation scheme for the governing DPS, the Grammian matrix $\mathbb{K}(\Xi_N,\Xi_N)$ in (22) must be included in the coordinate implementation. This is a significant difference between the native space embedding method and standard practices for most learning laws described in Euclidean adaptive control theory.

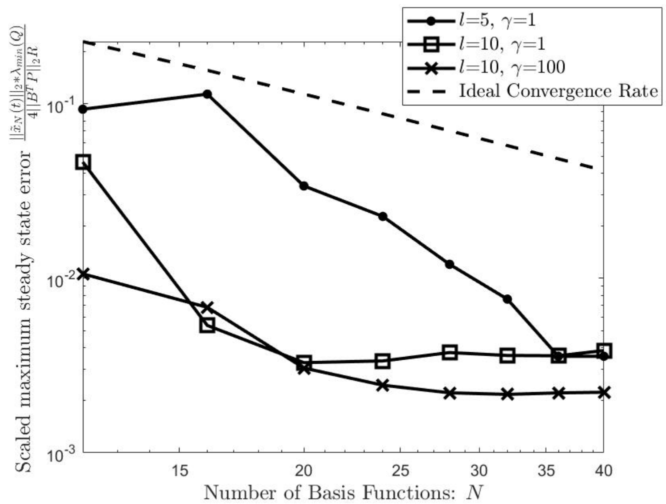

Fig. 1. Scaled norm of the maximum steady-state error $\frac{\|\tilde{x}_N(t)\|_{\mathbb{R}^d}\lambda_{\min}(Q)}{4\|B^{\mathrm{T}}P\|_{\mathbb{R}^d}R}$

The governing ODE studied in this section is given by

$$A=\begin{bmatrix}0 & 1\\ -\omega_n^2 & -2\zeta_n\omega_n\end{bmatrix} \quad B=\begin{bmatrix}0\\ \omega_n^2\end{bmatrix} \quad x=\begin{bmatrix}x_1\\ x_2\end{bmatrix}, \tag{23}$$

where $\omega_n=1$ rad/s and $\zeta_n=0.2$. The nonlinear uncertainty is chosen to be

$$f(x)=\tanh(x_1^3+0.001x_2^5), \qquad x\in\mathbb{R}^2. \tag{24}$$

Offline calculation indicates that the unknown uncertainty f satisfies $\|f\|_{\mathcal{H}}\approx\sqrt{7.23}\le R$. To track the reference command input $r(t)=\cos(5t)$, $t\ge 0$, the reference system is designed with

$$A_r=\begin{bmatrix}0 & 1\\ -\omega_r^2 & -2\zeta_r\omega_r\end{bmatrix} \quad B_r=\begin{bmatrix}0,\\ \omega_r^2\end{bmatrix}, \tag{25}$$

where $\omega_r=20$ rad/s denotes the reference model's natural frequency and $\zeta_r=\frac{\sqrt{2}}{2}$ denotes the damping ratio of the reference system. The kernel function for this problem is selected to be the "3/2 Matern kernel" [34], which is defined as

$$\mathfrak{K}_{3,2}(x,y)=\left(1+\frac{\sqrt{3}\|x-y\|_{\mathbb{R}^d}}{l}\right)e^{-\frac{\sqrt{3}\|x-y\|_{\mathbb{R}^d}}{l}}, \tag{26}$$

where $l\in\mathbb{R}$. The remaining system parameters are $Q=\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}$ and $\varepsilon=10^{-4}$.

Figure 1 captures the ultimate bound found in Theorem 3.1. There are a few important observations regarding this plot. Asymptotically, as $N \to \infty$, we expect that the scaled ultimate tracking error should *theoretically* decay at a rate that is no worse than the dashed line. This is indeed the case for the kernel with $l = 5$. Also, this rate of decrease holds for low values of N with $l = 10$. However, for $l = 10$ and large values of N, the rate of convergence plateaus. While the theoretical rates of convergence hold for steady state errors in finite time, the time to enter the dead-zone may be prohibitively large. While increasing γ_f should lead faster convergence in time of online approximation $\hat{f}_N$, larger γ_f can result in undesirable high frequency chattering that may take an intractable amount of time to decay and contribute to large tracking error. These simulations are run over a time span of 20 seconds.

V. Conclusions

In this paper, by viewing the original uncertain ODE as a DPS, we describe a general and systematic way to use the power function to define error bounding adaptive controllers in the native space setting. We use power functions to derive an upper bound on the tracking performance when the functional uncertainty f belongs to a fixed uncertainty class $\mathcal{C}_R \subset \mathcal{H}$ in the native space $\mathcal{H}$. For some kernels, we also derive a simpler form for the uniform ultimate bound for the tracking error in terms of the fill distance of the kernel basis centers.

Numerical examples demonstrate the applicability of the proposed adaptive control framework. These simulations show how, despite matched uncertainties, whose functional shape is unknown, satisfactory tracking performance can be attained. Further study of the interplay of approximation error, numerical conditioning, discrete integration error, and external noise in the performance of the native space embedding method is an important topic for future research.

Acknowledgments

This work has been partly supported by the US Department of the Navy under Grant no. N004212110004

References

[1] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[2] S. Thrun, "Probabilistic robotics," *Communications of the ACM*, vol. 45, no. 3, pp. 52–57, 2002.

[3] B. S. A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.

[4] P. Bobade, D. Panagou, and A. J. Kurdila, "Multi-agent adaptive estimation with consensus in reproducing kernel hilbert spaces," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 572–577.

[5] P. Bobade, S. Majumdar, S. Pereira, A. J. Kurdila, and J. B. Ferris, "Adaptive estimation in reproducing kernel hilbert spaces," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 5678–5683.

[6] P. Bobade, D. Panagou, and A. J. Kurdila, "Multi-agent adaptive estimation with consensus in reproducing kernel Hilbert spaces," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 572–577.

[7] P. Bobade, S. Majumdar, S. Pereira, A. J. Kurdila, and J. B. Ferris, "Adaptive estimation for nonlinear systems using reproducing kernel Hilbert spaces," *Advances in Computational Mathematics*, vol. 45, no. 2, pp. 869–896, 2019.

[8] S. T. Paruchuri, J. Guo, and A. Kurdila, "Kernel center adaptation in the reproducing kernel Hilbert space embedding method," *International Journal of Adaptive Control and Signal Processing*, vol. 36, no. 7, pp. 1562–1583, JUL 2022.

[9] J. Guo, S. T. Paruchuri, and A. J. Kurdila, "Partial Persistence of Excitation in RKHS Embedded Adaptive Estimation," *IEEE Transactions on Automatic Control*, 2022.

[10] J. Guo, M. E. Kepler, S. Tej Paruchuri, H. Wang, A. J. Kurdila, and D. J. Stilwell, "Adaptive estimation of external fields in reproducing kernel Hilbert spaces," *International Journal of Adaptive Control and Signal Processing*, vol. 36, no. 8, pp. 1931–1957, AUG 2022.

[11] J. Guo, S. T. Paruchuri, and A. J. Kurdila, "Approximations of the reproducing kernel hilbert space (rkhs) embedding method over manifolds," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 1596–1601.

[12] M. A. Demetriou, "Functional estimation of perturbed positive real infinite dimensional systems using adaptive compensators," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1582–1587.

[13] E. Lavretsky and K. Wise, *Robust and Adaptive Control: With Aerospace Applications*. Springer, 2013.

[14] P. A. Ioannou and J. Sun, *Robust adaptive control*. Courier Corporation, 2012.

[15] S. Sastry and M. Bodson, *Adaptive control: stability, convergence and robustness*. Courier Corporation, 2011.

[16] M. Krstic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design*. John Wiley & Sons, 1995.

[17] N. Hovakimyan and C. Cao, *$\mathcal{L}_1$ Adaptive Control Theory: Guaranteed Robustness with Fast Adaptation*. SIAM, 2010.

[18] G. Tao, *Adaptive control design and analysis*. John Wiley & Sons, 2003, vol. 37.

[19] J.-J. E. Slotine, W. Li *et al.*, *Applied nonlinear control*. Prentice hall Englewood Cliffs, NJ, 1991, vol. 199, no. 1.

[20] E. Lavretsky and K. Wise, *Robust and Adaptive Control: With Aerospace Applications*. London, UK: Springer, 2012.

[21] R. Sanner and J.-J. Slotine, "Gaussian networks for direct adaptive control," *IEEE Transactions on Neural Networks*, vol. 3, no. 6, pp. 837–863, 1992.

[22] S. Slama, A. Errachdi, and M. Benrejeb, "Online mrac method using neural networks based on variable learning rate for non-linear systems," vol. 60, no. 3, p. 215–224, 2019.

[23] J. A. Farrell and M. M. Polycarpou, *Adaptive approximation based control: unifying neural, fuzzy and traditional adaptive approximation approaches*. John Wiley & Sons, 2006, vol. 48.

[24] B. Peterson and K. Narendra, "Bounded error adaptive control," *IEEE Transactions on Automatic Control*, vol. 27, no. 6, pp. 1161–1168, 1982.

[25] P. Ioannou and P. Kokotović, *Adaptive Systems with Reduced Models*. Berlin: Springer, 1983.

[26] H. Wendland, *Scattered data approximation*. Cambridge university press, 2004, vol. 17.

[27] S. Saitoh and Y. Sawano, *Theory of reproducing kernels and applications*. Springer, 2016.

[28] R. B. Anderson, J. A. Marshall, and A. L'Afflitto, "Novel model reference adaptive control laws for improved transient dynamics and guaranteed saturation constraints," *Journal of the Franklin Institute*, vol. 358, no. 12, pp. 6281–6308, 2021.

[29] *Parameter Estimation Methods*. John Wiley & Sons, Ltd, 2006.

[30] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*. London, UK: Springer Science, 2012, vol. 44.

[31] H. Wang, A. J. Kurdila, A. L'Afflitto, D. Oesterheld, and D. J. Stilwell, "Robust model reference adaptive control based on reproducing kernel hilbert spaces," *IEEE Transactions on Automatic Control*, 2023.

[32] R. Schaback, "Error estimates and condition numbers for radial basis function interpolation," *Advances in Computational Mathematics*, vol. 3, pp. 251–264, 1994.

[33] J. Burns, A. Kurdila, D. Oesterheld, D. Stilwell, and H. Wang, "Learning theory convergence rates for observers and controllers in native space embedding," in *2023 American Control Conference (ACC)*. IEEE, 2023.

[34] E. Fuselier and G. B. Wright, "Scattered data interpolation on embedded submanifolds with restricted positive definite kernels: Sobolev error estimates," *SIAM Journal on Numerical Analysis*, vol. 50, no. 3, pp. 1753–1776, 2012.