Electronic Delivery Cover Sheet

Annual Reviews in Control 58 (2024) 100968

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Tutorial article

Nonparametric adaptive control in native spaces: Finite-dimensional implementations, Part II

Andrew J. Kurdila [a], Andrea L'Afflitto [b,*], John A. Burns [c], Haoran Wang [a]

[a] Department of Mechanical Engineering, Virginia Tech, 1145 Perry St., Blacksburg, 24060, VA, USA
[b] Grado Department of Industrial and Systems, Virginia Tech, 1145 Perry St., Blacksburg, 24060, VA, USA
[c] Interdisciplinary Center for Applied Mathematics, Virginia Tech, 765 West Campus Drive, Blacksburg, 24061, VA, USA

A R T I C L E I N F O

MSC:
93C40
93D21
93C10
47B32

Keywords:
Adaptive control systems
Adaptive stabilization
Nonlinear systems in control theory
Linear operators in reproducing kernel Hilbert spaces

A B S T R A C T

This two-part work presents a novel theory for model reference adaptive control (MRAC) of deterministic nonlinear ordinary differential equations (ODEs) that contain functional, nonparametric uncertainties that reside in a native space, also called a reproducing kernel Hilbert space (RKHS). As discussed in the first paper of this two-part work, the proposed framework relies on a limiting distributed parameter system (DPS). To allow implementations of this framework in finite dimensions, this paper shows how several techniques developed in parametric MRAC, such as the σ-modification method, the deadzone modification, adaptive error bounding methods, and projection methods, can be generalized to the proposed nonparametric setting. Some of these techniques assure uniform ultimate boundedness of the trajectory tracking error, while others guarantee its asymptotic convergence to zero. This paper introduces nonparametric metrics of performance that are cast in terms of the functional uncertainty classes in the native space. These performance metrics are relative to the best offline approximation error of the functional uncertainty. All the provided performance bounds are explicit in the dimension of the approximations of the functional uncertainty. Numerical examples show the applicability of the proposed theoretical results.

Contents

* Corresponding author.
E-mail address: a.lafflitto@vt.edu (A. L'Afflitto).

https://doi.org/10.1016/j.arcontrol.2024.100968
Received 6 December 2023; Received in revised form 14 September 2024; Accepted 17 September 2024
Available online 27 September 2024

1. Introduction

This paper is the second of a two-part work that presents in a systematic and tutorial manner both the theoretical foundation and several specific algorithms for a general theory of *nonparametric* model reference adaptive control (MRAC). In the first part of this work, we illustrated in detail the essential idea underlying the proposed approach to adaptive control systems design. Initially, in the overall approach, the nonlinear matched uncertainty is neither parameterized by a regressor vector nor reconstructed through estimators or other techniques but is assumed to be an element of a generally infinite dimensional reproducing kernel Hilbert space (RKHS), also known as native space. Substantial effort was devoted to explaining not only the differences but also the advantages of the proposed approach over classical approaches to MRAC design in particular and similar adaptive control design techniques in general. These techniques, indeed, which we refer to as *parametric methods*, require the user to either provide a parameterization of the matched uncertainties or devise some tools to parameterize such uncertainties.

This paper discussed in detail how the general theory of nonparametric MRAC can be applied in practice. Practical implementations of the proposed theory involve using finite-dimensional approximations of the space of uncertainties. Nevertheless, when applied to finite-dimensional problems of practical interest, using some classical tools of RKHSs, the proposed nonparametric framework allows quantifying explicitly the controller's performance as a function of the dimensions of the finite-dimensional functional space that approximates the infinite-dimensional space in which the uncertainty resides. In the classical parametric framework, on the contrary, such results are generally unavailable. Furthermore, by their nature, parametric MRAC systems can address considerably smaller classes of functional uncertainties than those that can be addressed by applying the proposed framework.

The core of the proposed approach lies in defining a so-called limiting distributed parameter system (DPS). This DPS captures both the trajectory tracking error dynamics and the adaptive laws. Since the adaptive law corresponding to the functional uncertainty evolves in an infinite-dimensional space, the first paper of this two-part work concluded that the adaptive laws comprised in the limiting DPS were not directly applicable to problems of practical interest. This paper shows how the limiting DPS can be approximated in finite dimensions, and, hence, applied to problems of practical interest. This result is attained by extending to an RKHS setting several classical adaptive control techniques such as the deadzone modification of MRAC (Peterson & Narendra, 1982), the σ-modification of MRAC (Ioannou & Fidan, 2006), the error bounding method, the adaptive error bounding method, and the use of continuous convex projection operators (Pomet & Praly, 1992). The proposed framework introduces metrics to measure the performance of all adaptive controllers proposed herein. In this paper, this metric captures the ultimate bound on the trajectory tracking error, and, as already discussed, classical tools of RKHS theory allow expressing explicitly this metric as a function of the dimension of the approximating RKHS. While retaining the same framework, alternative metrics, which capture other key aspects of the trajectory tracking error dynamics, such as its rate of convergence, can be considered.

This paper is structured as follows. Section 2 draws from the analogous section in the first paper of this two-part work, and succinctly provides a formal statement of the problem. Section 3 summarizes the main results of the first paper of this two-part work. These results are the MRAC laws for a plant model, whose nonlinearities are captured by infinite-dimensional RKHS, and sufficient conditions for the existence of solutions of both the adaptive and the closed-loop trajectory tracking error dynamics at all times. Before discussing how these results can be applied to finite-dimensional problems of practical interest,

Sections 4 and 5 present two sets of results instrumental to the scope of this paper. Specifically, Section 4 discusses a key tool in RKHS theory, that is, power functions, and how those can be used to characterize the error while approximating infinite-dimensional RKHSs using finite-dimensional Hilbert spaces. Section 5 introduces the novel notion of approximation theory suboptimality. This metric of performance will be used to compare the performance of all the controllers considered in this paper for multiple uncertainty classes using a standard measure.

Since the limiting controller presented in Section 3 yields in infinite dimensional spaces, and, hence, cannot be implemented in practice, in Sections 6–11, we present multiple generalizations of classical robust adaptive systems to the nonparametric setting that allow a finite-dimensional implementation of these results. The robust adaptive control techniques considered include the σ-modification of MRAC, the deadzone modification, error bounding and adaptive error bounding methods, and the continuous convex projection operator. All the control systems presented in this paper can be derived or interpreted as resulting from a two-stage design process. In the first stage, the analyst defines the limiting DPS discussed in Section 3. In the second stage, this ideal control system is specialized to a finite-dimensional setting to provide robustness to errors due to the attempt to approximate an infinite-dimensional RKHS employing a finite-dimensional one.

Section 13 provides a summary of the results attained in this paper. A central role in this section is played by Tables 2 and 3, where the performance of the proposed control techniques is compared for multiple uncertainty classes. Finally, in Section 14, we present the outcome of numerical simulations, and, in Section 15, we draw conclusions about this work and recommend future research directions.

2. Problem statement

In this paper, we consider *plants* in the form

$$\dot{x}(t) = Ax(t) + B\left(u(t) + f(x(t))\right), \quad x(t_0) = x_0, \quad t \geq t_0, \tag{1}$$

where $x : [t_0, \infty) \to \mathbb{X}$ denotes the *plant trajectory*, $\mathbb{X} = \mathbb{R}^n$ is the state space, $u : [t_0, \infty) \to \mathbb{R}$ denotes the *control input*, the *system matrix* $A \in \mathbb{R}^{n\times n}$ is unknown, the *control influence operator* $B \in \mathbb{R}^n$ is known and such that the pair (A, B) is controllable, and $f : \mathbb{R}^n \to \mathbb{R}$ is the nonlinear matched uncertainty. In the first paper of this two-part work, we provided a detailed discussion of the usefulness of this model and how, despite some limitations, such as the scalar input and the lack of unmatched uncertainties, it is sufficiently generic to explain the novel framework presented here.

Our goal is to find a *state-feedback control law* $\mu : \mathbb{X} \to \mathbb{R}$ such that the input $u(t) = \mu(x(t))$ steers the plant trajectory toward the reference trajectory $x_r : [t_0, \infty) \to \mathbb{X}$ defined as the solution of the *reference model*

$$\dot{x}_r(t) = A_{\text{ref}}x_r(t) + B_{\text{ref}}r(t), \quad x_r(t_0) = x_{r,0}, \quad t \geq t_0, \tag{2}$$

where the *reference command input* $r : [t_0, \infty) \to \mathbb{R}$ is continuous and bounded, $A_{\rm ref} \in \mathbb{R}^{n\times n}$ is Hurwitz, $B_{\rm ref} \in \mathbb{R}^n$, the pair $(A_{\rm ref}, B_{\rm ref})$ is controllable, and the *matching conditions*

$$A_{\rm ref} = A + B\alpha^{\rm T}, \tag{3}$$

$$B_{\rm ref} = B\beta, \tag{4}$$

are verified by for some $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$; note that α, in general, is unknown, whereas β must be known.

Defining the *trajectory tracking error* $e(t) \triangleq x(t) - x_r(t)$, $t \geq t_0$, our goal is to design a feedback control law $\mu(\cdot)$ so that *asymptotic tracking*, that is,

$$\lim_{t\to\infty} \|x(t;\mu) - x_r(t)\|_{\mathbb{R}^n} = 0, \tag{5}$$

where $x(t;\mu)$ denotes the solution of (1) with control input $u(t) = \mu(x(t))$. In some cases, such a result is not attainable. Thus, a weaker, but sufficiently useful objective is pursued, namely *practical tracking* (Aulisa, Burns, & Gilliam, 2023; Aulisa & Gilliam, 2015). This objective consists in finding a feedback control law $\mu(\cdot)$ so that, given a user-defined tolerance $\delta > 0$,

$$\limsup_{t\to\infty} \|x(t;\mu) - x_r(t)\|_{\mathbb{R}^n} \leq \delta, \tag{6}$$

where the limit supremum of a bounded function $\varphi : [t_0, \infty) \to \mathbb{R}^n$ is defined as

$$\limsup_{t\to\infty} \varphi(t) \stackrel{\Delta}{=} \lim_{T\to\infty} \left\{ \sup_{t\geq T} \varphi(t) \right\}. \tag{7}$$

Observe that if (5) is attained, then (6) is satisfied for all $\delta > 0$.

As discussed in Section 4 of the first paper of this two-part work, we assume that the functional matched uncertainty f resides in a native space

$$\mathcal{H} \triangleq \overline{\operatorname{span}\{\mathfrak{K}_x \mid x \in \mathbb{X}\}} \tag{8}$$

of real-valued functions over $\mathbb{X}$, where $\mathfrak{K}_x(\cdot) \triangleq \mathfrak{K}(\cdot, x)$, $\mathfrak{K} : \mathbb{X}\times\mathbb{X} \to \mathbb{R}$ denotes an admissible, bounded on the diagonal kernel function, and the closure is taken with respect to the inner product $\langle \mathfrak{K}_x, \mathfrak{K}_y\rangle_{\mathcal{H}} \triangleq \mathfrak{K}(x,y)$ for all $x, y \in \mathbb{X}$. In particular, we consider the following three classes of functional uncertainties.

Definition 2.1 (*Functional Uncertainty Classes*). Let $N \in \mathbb{N}$, $R > 0$, and $\epsilon > 0$. The *real parametric uncertainty class of radius* R is defined as

$$C_{\Phi_N,R} \triangleq \left\{ f = \Phi_N^{\rm T}\Theta_N \mid \Theta_N \in \mathbb{R}^N, \|f\| \leq R \right\} \subset \operatorname{span}\{\Phi_N\}. \tag{9}$$

The *nonparametric uncertainty class of radius* R *and projection error less than* ϵ is defined as

$$C_{N,\epsilon,R} \triangleq \left\{ f \in \mathcal{H} \mid \|(I - \Pi_N)f\|_{\mathcal{H}} \leq \epsilon, \|f\|_{\mathcal{H}} \leq R \right\} \subset \mathcal{H}. \tag{10}$$

The *nonparametric uncertainty class of radius* R is defined as

$$C_R \triangleq \left\{ f \in \mathcal{H} \mid \|f\|_{\mathcal{H}} \leq R \right\} \subset \mathcal{H}. \tag{11}$$

As extensively discussed in the first paper of this two-part work, these three classes are such that

$$C_{\Phi_N,R} \subset C_{N,\epsilon,R} \subset C_R. \tag{12}$$

Classical parametric adaptive control systems only consider the class of functional uncertainties $C_{\Phi_N,R}$. The proposed framework, instead, allows addressing the broader classes $C_{N,\epsilon,R}$ and C_R.

3. The nonparametric control law and the limiting DPS

In this section, we briefly summarize the key result of the first paper of this two-part work. Assuming that f in (1) is contained in the native space $\mathcal{H}$ of real-valued functions over $\mathbb{X}$ given by (8), it holds that

$$f(x) = E_x f = \langle \mathfrak{K}_x, f\rangle_{\mathcal{H}}, \quad \text{for all } x \in \mathbb{X}, \tag{13}$$

and (1) can be written as

$$\dot{x}(t) = Ax(t) + B\left(u(t) + E_{x(t)}f\right), \quad x(t_0) = x_0, \quad t \geq t_0.$$

To design MRAC laws that account for uncertainties in both A and f, we consider the nonparametric adaptive control input

$$u(t) = \mu(t, x(t), \hat{\alpha}(t), \mathfrak{K}_{x(t)}, \hat{f}(t,\cdot)), \quad t \geq t_0, \tag{14}$$

where the *control law* is defined as

$$\mu(t, x, \alpha, \mathfrak{K}_x, f) \triangleq \alpha^{\rm T}x + \beta r(t) - \langle \mathfrak{K}_x, f\rangle_{\mathcal{H}}, \quad (t, x, \alpha, \mathfrak{K}_x, f) \in [t_0, \infty)\times\mathbb{X}\times\mathbb{R}^n\times\mathcal{H}\times\mathcal{H}, \tag{15}$$

$\hat{\alpha}(t) : [t_0, \infty) \to \mathbb{R}^n$ and $\hat{f}(t,\cdot) : [t_0, \infty) \to \mathcal{H}$ denote *adaptive gains* and verify the *adaptive laws*

$$\dot{\hat{\alpha}}(t) = -\Gamma_\alpha x(t)e^{\rm T}(t)PB, \quad \hat{\alpha}(t_0) = \hat{\alpha}_0, \tag{16}$$

$$\begin{aligned}\frac{\partial \hat{f}(t,\cdot)}{\partial t} &= \gamma_f \mathfrak{K}(\cdot, x(t))e^{\rm T}(t)PB \\ &= \gamma_f E^*_{x(t)}B^{\rm T}Pe(t), \quad \hat{f}(t_0,\cdot) = \hat{f}_0(\cdot),\end{aligned} \tag{17}$$

It is worthwhile noting that (17) is a partial differential equation (PDE). With the control input (14), the trajectory tracking error dynamics are given by

$$\dot{e}(t) = A_{\rm ref}e(t) - B\left(\tilde{\alpha}^{\rm T}(t)x(t) - \langle \mathfrak{K}_{x(t)}, \tilde{f}(t,\cdot)\rangle\right), \quad e(t_0) = x_0 - x_{r,0}, \quad t \geq t_0, \tag{18}$$

where $\tilde{\alpha}(t) \triangleq \alpha - \hat{\alpha}(t)$ and $\tilde{f}(t,\cdot) \triangleq f - \hat{f}(t,\cdot)$. Furthermore, (16)–(18) can be collected to write the *limiting DPS error equations* in matrix operator form

$$\frac{\partial}{\partial t}\begin{Bmatrix} e(t) \\ \tilde{\alpha}(t) \\ \tilde{f}(t,\cdot)\end{Bmatrix} = \begin{bmatrix} A_{\rm ref} & -Bx^{\rm T}(t) & BE_{x(t)} \\ \Gamma_\alpha x(t)B^{\rm T}P & 0 & 0 \\ -\gamma_f E^*_{x(t)}B^{\rm T}P & 0 & 0 \end{bmatrix}\begin{Bmatrix} e(t) \\ \tilde{\alpha}(t) \\ \tilde{f}(t,\cdot)\end{Bmatrix} \tag{19}$$

with the same initial conditions as in (16)–(18). In (19), we do not have a matrix of real numbers but an operator since it contains the bounded linear operators $E_x : \mathcal{H} \to \mathbb{R}$ and $E^*_x : \mathbb{R} \to \mathcal{H}$.

The next theorem shows the boundedness of solutions of (19), that is, of

$$t \mapsto z(t) \triangleq \left(e(t), \tilde{\alpha}(t), \tilde{f}(t,\cdot)\right) \in \mathbb{R}^n\times\mathbb{R}^n\times\mathcal{H} \triangleq \mathbb{Z} \tag{20}$$

and the convergence of the trajectory tracking error $e(\cdot)$ to zero despite the uncertainties in A and f.

Theorem 3.1. *Consider the limiting DPS (19), suppose that the kernel $\mathfrak{K}(\cdot,\cdot)$ that defines the native space $\mathcal{H}$ is bounded on the diagonal, and (19) is forward complete. Then, the trajectory of the limiting error DPS is uniformly bounded, and*

$$\lim_{t\to\infty} e(t) = 0$$

uniformly in $t_0 \in [0, \infty)$.

Theorem 3.1 assumes the forward completeness of the map (20). The next result provides sufficient conditions for this result to hold.

Theorem 3.2. *Consider the limiting error DPS given by (19) and the reference model given by (2). Suppose that the kernel $\mathfrak{K}(\cdot,\cdot)$ that defines the native space $\mathcal{H}$ of real-valued functions is bounded on the diagonal, and that the native space $\mathcal{H}$ is continuously embedded in the space of Lipschitz continuous functions $C^{0,1}(\mathbb{X})$, that is, $\mathcal{H} \hookrightarrow C^{0,1}(\mathbb{X})$. Then, for any $z(t_0) \in \mathbb{Z}$, (19) is forward complete with $z \in C^1([t_0,\infty), \mathbb{Z})$, that is, the mapping $t \mapsto z(t) = \left(e(t), \tilde{\alpha}(t), \tilde{f}(t,\cdot)\right)$ is defined on $[t_0, \infty)$.*

4. Power functions and approximation bounds

In native space theory, a powerful tool to characterize the error in approximating an infinite-dimensional space using a finite-dimensional space is given by the power function. This tool allows quantifying the error due to projecting the DPS (17) into a finite-dimensional RKHS and will be used extensively in the remainder of this work. For a brief discussion on relevant properties of orthogonal projections, see Section 4 of the first paper of this two-part work.

Consider the admissible kernel function $\mathfrak{K} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$, and the associated RKHS (8). The power function can be defined as follows.

Definition 4.1 (*Power Function*)**.** Let $U \subseteq \mathcal{H}$ denote a closed subspace of $\mathcal{H}$, let Π_U denote the $\mathcal{H}$-orthogonal projection onto U, and let $\mathcal{H}^*$ denote the topological dual of $\mathcal{H}$. The mapping $\mathcal{P}_U : \mathcal{H}^* \to \mathbb{R}$ such that

$$\mathcal{P}_U(h^*) \triangleq \sup_{f\in\mathcal{H}\setminus\{0\}} \frac{\left|h^*(f) - h^*(\Pi_U f)\right|}{\|f\|_{\mathcal{H}}}, \quad \text{for all } h^* \in \mathcal{H}^*, \tag{21}$$

is the *power function associated with* U.

Let $x \in \mathbb{X}$, $\alpha \in \mathbb{R}$, and $\delta_x^\alpha(f) \triangleq \left(E_x f\right)\alpha$. If $h = \delta_x^\alpha(f)$, then we use the short-hand notation $\mathcal{P}_U(x)$ for $\mathcal{P}_U(h^*)$. Consider also the set of centers

$$\Xi_N \triangleq \left\{\xi_1, \ldots, \xi_N\right\} \subset \Omega \subset \mathbb{X}. \tag{22}$$

It is shown in Wittwar, Santin, and Haasdonk (2018) that we always have the pointwise error bound

$$|E_x\left(I - \Pi_N\right) f| \le \mathcal{P}_N(x) \|\left(I - \Pi_N\right) f\|_{\mathcal{H}} \le \mathcal{P}_N(x)\|f\|_{\mathcal{H}}, \tag{23}$$

for all $x \in \mathbb{X}$ and $f \in \mathcal{H}$, where $I : \mathcal{H} \to \mathcal{H}$ denotes the *identity operator*, $\Pi_N : \mathcal{H} \to \mathcal{H}_N$ denotes the *orthogonal projection* onto

$$\mathcal{H}_N \triangleq \text{span}\left\{\mathfrak{K}_{\xi_i} \mid \xi_i \in \Xi_N,\ 1 \le i \le N\right\} \subset \mathcal{H}, \tag{24}$$

and $\mathcal{P}_N(\cdot)$ denotes the *power function* associated with the finite-dimensional space $\mathcal{H}_N$; with a minor abuse of notation, $\mathcal{P}_N(\cdot)$ stands for $\mathcal{P}_{\mathcal{H}_N}(\cdot)$. Given $f \in \mathcal{H}$, $\Pi_N f$ denotes the orthogonal projection of f into the finite-dimensional space $\mathcal{H}_N$, and $\left(I - \Pi_N\right) f$ denotes its orthogonal complement. Thus, through (23), the power function $\mathcal{P}_N(\cdot)$ provides an upper bound on the magnitude of $\left(I - \Pi_N\right) f$. Now, as discussed in Section 6 below, it possible to deduce adaptive control laws that assure asymptotic convergence of the trajectory tracking error to zero irrespectively of $\Pi_N f$, that is, irrespectively of the effect on (1) of the component of f in a user-defined subspace $\mathcal{H}_N$. Thus, only the effect of the component of f orthogonal of $\mathcal{H}_N$ on (1), namely $\left(I - \Pi_N\right) f$, must be assessed, and (23) provides an essential tool to this goal.

Definition 4.1 is not particularly useful for algorithms, at least in most problems of practical interest. To overcome this limitation, we recall the following result.

Theorem 4.1. *Let $\mathfrak{K} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ denote the kernel that induces the native space $\mathcal{H}$ and let $U \subseteq \mathcal{H}$ be a closed subspace induced by the kernel $\mathfrak{K}_U(x, y) = (\Pi_U \mathfrak{K}_x)(y)$. Then, it holds that*

$$\mathcal{P}_U^2(x) = \mathfrak{K}(x, x) - \mathfrak{K}_U(x, x), \quad \textit{for all } x \in \mathbb{X}. \tag{25}$$

Let $\mathcal{H}$ denote a native space defined by the kernel function $\mathfrak{K}(\cdot, \cdot)$, let $U \subseteq \mathcal{H}$ be a closed subspace induced by the kernel $\mathfrak{K}_U(x, y) = (\Pi_U \mathfrak{K}_x)(y)$. It holds that $\mathfrak{K}(x, x) \ge \mathfrak{K}_U(x, x)$ for all $x \in \mathbb{X}$. Thus, given the set of centers (22) and finite-dimensional subspace

$$\mathcal{H}_N \triangleq \text{span}\{\mathfrak{K}_{\xi_i} \mid \xi_i \in \Xi_N\} \subset \mathcal{H}, \tag{26}$$

it follows from Theorem 4.1 that

$$\mathcal{P}_N(x) = \sqrt{\mathfrak{K}(x, x) - \mathfrak{K}_N(x, x)}, \tag{27}$$

Box 1: To apply the proposed nonparametric MRAC framework to problems of practical interest, the underlying DPS needs to be projected onto finite-dimensional RKHSs. Some of the proposed adaptive control systems are affected by the orthogonal complement of this projection. The power function, a classical tool from RKHS theory, allows estimating the L^∞-norm of the projection error and, ultimately, its effect on the trajectory tracking error performance.

where the reproducing kernel $\mathfrak{K}_N(\cdot, \cdot)$ of $\mathcal{H}_N$ is given by

$$\mathfrak{K}_N(x, y) \triangleq \mathfrak{K}_{\Xi_N}^{\mathrm{T}}(x)\mathbb{K}_N^{-1}\mathfrak{K}_{\Xi_N}(y), \quad \text{for all } x, y \in \mathbb{X}, \tag{28}$$

$$\mathfrak{K}_{\Xi_N}(\cdot) \triangleq \left[\mathfrak{K}_{\xi_1}(\cdot), \ldots, \mathfrak{K}_{\xi_N}(\cdot)\right]^{\mathrm{T}}, \text{ and}$$

$$\mathbb{K}_N \triangleq [\mathfrak{K}(\xi_i, \xi_j)]_{(i,j)} = [(\mathfrak{K}_{\xi_i}, \mathfrak{K}_{\xi_j})]_{(i,j)} \in \mathbb{R}^{N\times N} \tag{29}$$

denotes the *Grammian matrix of the set of kernel basis functions* $\{\mathfrak{K}_{\xi_i}\}_{i=1}^N$ located at the set of N centers given by (22), and $[\cdot]_{(i,j)}$ denotes a matrix by specifying the element on the ith row and jth column. Thus, having chosen the set of centers $\Xi_N \subset \Omega$, we can compute (28) and (27), and hence, quantify the effect of the unknown matched uncertainty employing (23).

It is worthwhile reflecting on the novelty of (27) and (23). These equations give a geometric condition, defined in closed form in terms of the centers Ξ_N, that can be used to bound the pointwise approximation error at any point x in a domain of interest $\Omega \subseteq \mathbb{X}$. The authors are unaware of such a simple and computable result when the space of approximants is defined in terms of finite element, spline, Fourier, ridge function, or common single layer or multilayer neural network bases, for example.

5. Approximation theory suboptimality

In this work, we use a convenient overarching strategy to assess the performance of various types of adaptive control methods when the uncertainty belongs to one of the classes (9)–(11). To present this strategy, consider the following definition. For the statement of this definition, recall that *identity operator* $I : \mathbb{R}^N \to \mathbb{R}^N$ is defined such that $Ix = x$.

Definition 5.1 (*Approximation Theory Suboptimality*)**.** Consider the uncertainty class $C_{N,\epsilon,R}$ given by (10), and let the closed-loop trajectory of (1) with $f \in C_{N,\epsilon,R}$ be denoted by $x_N : [t_0, \infty) \to \mathbb{X}$. Let the functional $J : C([t_0, \infty), \mathbb{X}) \to \overline{\mathbb{R}}_+$ measure the performance of an adaptive control method, where if $J(x_N) = 0$, then the best performance is attained. An adaptive control system is *approximation theory $\mathcal{F}(\epsilon)$-suboptimal over* $C_{N,\epsilon,R}$ if, controlling the plant model (1) by means of this adaptive control system, there exist a constant $C > 0$ and a known function $\mathcal{F} : \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$ such that $\mathcal{F}(\epsilon) \ge \epsilon$ for all $\epsilon \ge 0$ and

$$J(x_N) \le C\mathcal{F}(\epsilon), \quad \text{for all } f \in C_{N,\epsilon,R}. \tag{30}$$

If $\mathcal{F}(\epsilon) = I$ for all $\epsilon \ge 0$, then (30) takes the form

$$J(x_N) \le C\epsilon, \quad \text{for all } f \in C_{N,\epsilon,R}, \tag{31}$$

and we say that the controller is *nearly approximation theory optimal* over $C_{N,\epsilon,R}$.

Intuitively, the notion of approximation theory $\mathcal{F}(\epsilon)$-suboptimal over $C_{N,\epsilon,R}$ reflects the fact that we expect the approximation of the uncertainty to influence the performance of the adaptive control system. In general, we expect that the poorer the approximation, the lower the performance. Inequality (30) says that if the uncertainty $f \in C_R$ is approximated from $\mathcal{H}_N$ with error ϵ, then the performance of the closed-loop controller is of the order $O(\mathcal{F}(\epsilon))$. Inequality (31) is to be interpreted as a special, easily quantifiable case of (30). The positive

Box 2: The notion of approximation theory suboptimality, which is inspired by the notion of near optimality commonly used in approximation theory and statistical learning theory, explicitly correlates the controller's performance to the ability to approximate the infinite-dimensional space of functional uncertainties by means of finite-dimensional spaces. Approximation theory suboptimality requires establishing an upper-bound on a user-defined performance measure. Such an upper bound must be a function of the largest component of a functional uncertainty that is orthogonal to the user-defined, finite-dimensional approximating space.

constant C in (30) and (31) is to be interpreted as an offline rate of convergence of best approximations. Sometimes, it is possible to derive alternative forms that are explicit in the dimension N. Thus, we introduce

$$\epsilon_N(x) \triangleq ((I - \Pi_N)f)(x), \quad \text{for all } x \in \mathbb{X}, \tag{32}$$

which denotes the pointwise error of the best offline approximation of f, and the following notion.

Definition 5.2 (*Approximation Theory Suboptimality, cont.*). Consider the uncertainty class $C_{N,\epsilon,R}$ given by (10), and let the closed-loop trajectory of (1) with $f \in C_{N,\epsilon,R}$ be denoted by $x_N : [t_0, \infty) \to \mathbb{X}$. Let the functional $J : C([t_0, \infty), \mathbb{X}) \to \overline{\mathbb{R}}_+$ measure the performance of an adaptive control system, where if $J(x_N) = 0$, then the best performance is attained. An adaptive control scheme is *$\mathcal{F}(N)$-suboptimal over $C_{N,\epsilon,R}$* if, applying this system to the plant model (1), there exist a constant $C > 0$ and a known function $\mathcal{F} : \mathbb{N} \to \overline{\mathbb{R}}_+$ such that $\mathcal{F}(N) \geq \|\epsilon_N\|$, where

$$J(x_N) \leq C\mathcal{F}(N). \tag{33}$$

If $\mathcal{F}(N) = \|\epsilon_N\| = \|(I - \Pi_N)f\|$, then (33) becomes

$$J(x_N) \leq C\|\epsilon_N\| = C\|(I - \Pi_N)f\|. \tag{34}$$

Definition 5.2 can be considered an extension of Definition 5.1 because any method that satisfies (34) satisfies (31). Indeed, $\|\epsilon_N\|_{\mathcal{H}} = \|(I - \Pi_N)f\|_{\mathcal{H}} \leq \epsilon$ for all $f \in C_{N,\epsilon,R}$. However, it is more difficult to establish (33) and (34).

The description of estimates as "nearly optimal" is common in approximation theory (DeVore, 1998; DeVore & Lorentz, 1993), and statistical learning theory (Temlyakov, 2011). Some early connections between adaptive control schemes and cost functions within a differential game framework are presented in L'Afflitto (2017). However, to the authors' knowledge, this notion has not been systematically used in adaptive control theory to date. For brevity, we just call such methods *approximation theory optimal* for short. The relevance of inequalities (30)–(34) lies in that they measure the performance of the closed-loop controller against a well-known, universal standard, namely the error of best offline approximation of f from $\mathcal{H}_N$.

In this paper, we capture the performance of adaptive systems by means of the performance measure

$$J(x_N(\cdot)) \triangleq \limsup_{t\to\infty} \|x_N(t) - x_r(t)\|_{\mathbb{R}^n}. \tag{35}$$

This index captures the ultimate bound on the tracking error of the trajectory in a model reference adaptive control strategy that uses $\mathcal{H}_N$ to approximate the functional uncertainty f that is known to lie in $\mathcal{H}$. If, for some method, we derive an ultimate bound on the tracking error of the form $\bar{e}_{N,\infty}$, then it must hold that $J(x_N(\cdot)) \leq \bar{e}_{N,\infty}$. However, it must be noted that, in general, $\bar{e}_{N,\infty}$ is only a conservative estimate of the performance of the adaptive control system. For this reason, we introduce the following definition.

Box 3: If the user-defined performance measure captures the ultimate bound on the trajectory tracking error, then the concept approximation theory suboptimality specializes to asymptotic approximation optimality.

Definition 5.3 (*AAO Adaptive Systems*). Let $x_N : [t_0, \infty) \to \mathbb{X}$ denote the closed-loop trajectory of (1) with $f \in C_{N,\epsilon,R}$. If (30) or (34) are verified with performance measure (35), then an adaptive control system is *asymptotically approximation optimal (AAO)*. If $\bar{e}_{N,\infty} = 0$ for any N, then we say that the adaptive control system is AAO^*.

For some control systems, such as the deadzone method with a suitably defined deadzone size (see Section 8 below), we establish that

$$J(x_N(\cdot)) \leq \bar{e}_{N,\infty} \leq C\epsilon. \tag{36}$$

Using well-known tools from the theory of native spaces, in this paper, we derive further upper bounds on $e_{N,\infty}$ that are explicit functions of the number of bases N as in (34). For some other methods, such as the σ-modification (see Section 7 below), the function $\mathcal{F}$ is more complicated, and the bounds derived in this paper are only suboptimal.

It is instructive to compare the bound given by (36) to the situation when a classical real parametric MRAC system is used and the plant uncertainty f is contained in the smaller uncertainty class $f \in C_{\Phi_N,R}$. In this case the performance metric $J(\cdot)$ for the closed-loop trajectory $t \mapsto x_N(t)$ reduces to

$$J(x_N) \triangleq \lim_{t\to\infty} \|x_N(t) - x_r(t)\| \equiv \|(I - \Pi_N)f\|_{\mathcal{H}} = 0, \quad \text{for all } f \in C_{\Phi_N,R}.$$

Given our terminology, the performance of the classical real parametric gradient learning law, for the performance functional that captures the ultimate bound on the tracking error, is *approximation theory optimal* over $C_{\Phi_N,R}$ for every N for which $f \in \mathcal{H}_N$. We are interested in generalizing this statement of classical parametric adaptive control to obtain similar guarantees, but in a setting that includes nonparametric uncertainty classes.

6. Finite-dimensional approximations of DPS

As discussed in Section 3, for each $t \in [t_0, \infty)$, the adaptive gain $\hat{f}(t, \cdot) \in \mathcal{H}$ is nonparametric and, hence, not implementable. Thus, to overcome this limitation, throughout the remainder of this paper, we devise finite-dimensional adaptive gains $\hat{f}_N(t, \cdot) \in \mathcal{H}_N$ that approximate $\hat{f}(t, \cdot)$ and reside in the finite-dimensional space $\mathcal{H}_N \subseteq \mathcal{H}$ given by (26). Several finite-dimensional control adaptive systems are considered, and all these systems stem from the very same set of considerations presented in this section.

Before we proceed, a comment about notation is in order. When building approximations, we eventually restrict our analysis to some compact set $\Omega \subset \mathbb{X}$ wherein the uniform approximation assumption is verified (see Definition 3.1 of the first paper in this two-part work). Restricting the kernel $\mathfrak{K}(\cdot, \cdot)$ initially defined on $\mathbb{X} \times \mathbb{X}$ to the subset $\Omega \times \Omega$ defines a native space of real-valued functions over the set Ω. This is a standard construction of RKHS over subsets; see the discussion in Berlinet and Thomas-Agnan (2011), Paulsen and Raghupathi (2016), Saitoh and Sawano (2016) for details. This restriction process is so common in many analyses and applications that it is made without comment; see Wendland (2004). In this paper, we abuse notation and use the same symbols $\mathcal{H}$ and $\mathfrak{K}(\cdot, \cdot)$ both for $\mathcal{H}(\Omega)$ and $\mathfrak{K} : \Omega \times \Omega \to \mathbb{R}$, which concern the analysis over the subset Ω, and $\mathcal{H}(\mathbb{X})$ and $\mathfrak{K} : \mathbb{X}\times\mathbb{X} \to \mathbb{R}$. Whether we refer to functions defined over the specific domain Ω or $\mathbb{X}$ in any particular situation should be clear from the context.

In light of Theorem 4.1 of the first part of this work, for each $t \in [t_0, \infty)$, we let

$$E_{x(t)}f = E_{x(t)}\Pi_N f + E_{x(t)}\left(I - \Pi_N\right)f. \tag{37}$$

Box 4: Employing finite-dimensional RKHS to approximate infinite-dimensional spaces of functional uncertainties, the approximation error perturbs the closed-loop plant dynamics in a manner that is comparable to an unmatched uncertainty. Thus, when employing finite-dimensional approximations of uncertainties, robust adaptive control systems need to be employed. A key advantage of the proposed nonparametric framework over parametric results lies in the ability to correlate explicitly the controller's performance with the approximation error.

With this representation of the functional uncertainty, (1) becomes

$$\dot{x}(t) = Ax(t) + B\left(u(t) + E_{x(t)}\Pi_N f\right) + \underbrace{BE_{x(t)}(I - \Pi_N)f}_{\text{approximation error}},$$
$$x(t_0) = x_0, \quad t \geq t_0. \tag{38}$$

This equation should be compared to common forms that appear in approximation-based MRAC in real parametric adaptive control theory; see Section 4.6 of Farrell and Polycarpou (2006) or Chapter 11 of Lavretsky and Wise (2012). An important difference between parametric MRAC and the proposed nonparametric framework is that, in the nonparametric setting, we exploit the Hilbert space structure of the functional uncertainty.

Remark 6.1. In this paper, for brevity, we do not address problems wherein the plant dynamics are affected by unmatched uncertainties, that is, plant in the same form as

$$\dot{x}(t) = Ax(t) + B(u(t) + f(x(t))) + d(t), \quad x(t_0) = x_0, \quad t \geq t_0,$$

where $d : [t_0, \infty) \to \mathbb{R}^n$ is unknown, continuous, and bounded. This case can be addressed by applying any of the results presented in this paper by replacing $BE_{x(t)}(I - \Pi_N)f$ with $BE_{x(t)}(I - \Pi_N)f + d(t)$ for all $t \geq t_0$.

We denote by $x_N(\cdot)$ the closed-loop trajectory of (38) generated by the control input

$$u_N(t) = \hat{\alpha}_N^{\mathrm{T}}(t)x_N(t) + \beta r(t) - E_{x_N(t)}\hat{f}_N(t, \cdot), \quad t \geq t_0, \tag{39}$$

with $\hat{f}_N(t, \cdot) \in \mathcal{H}_N$ to be defined. In this case, the tracking error dynamics are

$$\begin{aligned}
\dot{e}_N(t) &= \dot{x}_N(t) - \dot{x}_{\mathrm{ref}}(t)\\
&= Ax_N(t) + B\left(\hat{\alpha}_N^{\mathrm{T}}(t)x_N(t) - E_{x_N(t)}\hat{f}_N(t, \cdot) + E_{x_N(t)}f\right)\\
&\quad - A_{\mathrm{ref}}x_r(t) - B_{\mathrm{ref}}r(t)\\
&= \left(A + B\hat{\alpha}_N^{\mathrm{T}}(t)\right)x_N(t) + BE_{x_N(t)}\tilde{f}_N(t, \cdot) - A_{\mathrm{ref}}x_r(t) - B_{\mathrm{ref}}r(t)\\
&= \left(A + B\hat{\alpha}_N^{\mathrm{T}}(t) + B\alpha_N^{\mathrm{T}} - B\alpha_N^{\mathrm{T}}\right)x_N(t) + BE_{x_N(t)}\tilde{f}_N(t, \cdot)\\
&\quad - A_{\mathrm{ref}}x_r(t) - B\beta r(t)\\
&= A_{\mathrm{ref}}x_N(t) - B\tilde{\alpha}_N^{\mathrm{T}}(t)x_N(t) + E_{x_N(t)}\tilde{f}_N(t, \cdot) - A_{\mathrm{ref}}x_r(t)\\
&= A_{\mathrm{ref}}e_N(t) - B\left(\tilde{\alpha}_N^{\mathrm{T}}(t)x_N(t) - E_{x_N(t)}\tilde{f}_N(t, \cdot)\right),
\end{aligned}$$
$$e_N(t_0) = x_0 - x_{r,0}, \quad t \geq t_0, \tag{40}$$

where

$$\tilde{f}_N(t, \cdot) \triangleq f - \hat{f}_N(t, \cdot). \tag{41}$$

To define the adaptive laws, assess the boundedness of the trajectory tracking error and of the adaptive gains, and prove the asymptotic convergence properties of the tracking error, we consider the Lyapunov function candidate

Box 5: All the adaptive control systems presented in this paper stem from the adaptive laws (45) and (47), and are certified by the Lyapunov function candidate (42). These adaptive laws are modified according to the technique considered. In all cases, uniform boundedness of the trajectory tracking error and adaptive gains is guaranteed. In some cases, uniform ultimate boundedness of the trajectory tracking error is ensured. In other cases, uniform asymptotic convergence of the trajectory tracking error is proven.

$$V\left(e_N, \tilde{\alpha}_N, \tilde{f}_N\right) = \langle Pe_N, e_N\rangle_{\mathbb{R}^n} + \left\langle \Gamma_\alpha^{-1}\tilde{\alpha}_N, \tilde{\alpha}_N\right\rangle_{\mathbb{R}^n} + \gamma_f^{-1}\left\langle \tilde{f}_N, \tilde{f}_N\right\rangle_{\mathcal{H}},$$
$$\left(e_N, \tilde{\alpha}_N, \tilde{f}_N\right) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{H}_N. \tag{42}$$

The time derivative of (42) along trajectories of (40) is given by

$$\begin{aligned}
&\dot{V}\left(e_N(t), \tilde{\alpha}_N(t), \tilde{f}_N(t, \cdot)\right)\\
&= \langle P\dot{e}_N(t), e_N(t)\rangle_{\mathbb{R}^n} + \langle Pe_N(t), \dot{e}_N(t)\rangle_{\mathbb{R}^n}\\
&\quad + 2\left(\left\langle \Gamma_\alpha^{-1}\tilde{\alpha}_N^{\mathrm{T}}(t), \dot{\tilde{\alpha}}_N(t)\right\rangle_{\mathbb{R}^n} + \gamma_f^{-1}\left\langle \tilde{f}_N(t, \cdot), \dot{\tilde{f}}_N(t, \cdot)\right\rangle_{\mathcal{H}}\right)\\
&= \left\langle \left(A_{\mathrm{ref}}^{\mathrm{T}}P + PA_{\mathrm{ref}}\right)e_N(t), e_N(t)\right\rangle_{\mathbb{R}^n}\\
&\quad - 2\left\langle PB\left(\tilde{\alpha}_N^{\mathrm{T}}(t)x(t) - E_{x_N(t)}\tilde{f}_N(t, \cdot)\right), e_N(t)\right\rangle_{\mathbb{R}^n} \qquad (43)\\
&\quad + 2\left(-\left\langle \tilde{\alpha}_N(t), \Gamma_\alpha^{-1}\dot{\hat{\alpha}}_N(t)\right\rangle_{\mathbb{R}^n} - \gamma_f^{-1}\left\langle \tilde{f}_N(t, \cdot), \dot{\hat{f}}_N(t, \cdot)\right\rangle_{\mathcal{H}}\right)\\
&= -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\left[-\left\langle PB\tilde{\alpha}_N^{\mathrm{T}}(t)x(t), e_N(t)\right\rangle_{\mathbb{R}^n} - \left\langle \tilde{\alpha}_N(t), \Gamma_\alpha^{-1}\dot{\hat{\alpha}}_N(t)\right\rangle_{\mathbb{R}^N}\right]\\
&\quad + 2\left[\left\langle PBE_{x_N(t)}\tilde{f}_N(t, \cdot), e_N(t)\right\rangle_{\mathbb{R}^n} - \gamma_f^{-1}\left\langle \tilde{f}_N(t, \cdot), \dot{\hat{f}}_N(t, \cdot)\right\rangle_{\mathcal{H}}\right]\\
&= -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} - 2\left\langle \tilde{\alpha}_N(t), x(t)e_N^{\mathrm{T}}(t)PB + \Gamma_\alpha^{-1}\dot{\hat{\alpha}}_N(t)\right\rangle_{\mathbb{R}^n}\\
&\quad + 2\left\langle \tilde{f}_N(t, \cdot), E_{x_N(t)}^*B^{\mathrm{T}}Pe_N(t) - \gamma_f^{-1}\dot{\hat{f}}_N(t, \cdot)\right\rangle_{\mathcal{H}}, \quad t \geq t_0. \qquad (44)
\end{aligned}$$

Based on the well-known strategy in the classical MRAC approach, we would like to cancel all terms other than $-\langle Qe(t), e(t)\rangle_{\mathbb{R}^n}$, $t \geq t_0$. Thus, we consider the adaptive law

$$\dot{\hat{\alpha}}_N(t) = -\Gamma_\alpha x_N(t)e_N^{\mathrm{T}}(t)PB, \quad \hat{\alpha}_N(t_0) = \hat{\alpha}_{N,0}, \quad t \geq t_0, \tag{45}$$

which has the same structure as (15). Ideally, we would like to set

$$\frac{\partial \hat{f}_N(t, \cdot)}{\partial t} = \gamma_f E_{x_N(t)}^* B^{\mathrm{T}}Pe_N(t), \quad \hat{f}_N(t_0, \cdot) = \hat{f}_0(\cdot), \quad t \geq t_0, \tag{46}$$

as is done for the limiting DPS in (17). However, this is only possible if $\mathcal{H}_N \equiv \mathcal{H}$. Indeed, we know that $\hat{f}_N(t, \cdot) \in \mathcal{H}_N$ for each $t \in [t_0, \infty)$. However, in general, it holds that $E_{x_N(t)}^* B^{\mathrm{T}}Pe_N(t) = \mathfrak{K}\left(x_N(t), \cdot\right)B^{\mathrm{T}}Pe_N(t) \notin \mathcal{H}_N$, $t \geq t_0$. Thus, we propose the adaptive law

$$\frac{\partial \hat{f}_N(t, \cdot)}{\partial t} = \gamma_f \Pi_N E_{x_N(t)}^* B^{\mathrm{T}}Pe_N(t), \quad \hat{f}_N(t_0, \cdot) = \hat{f}_0(\cdot), \quad t \geq t_0; \tag{47}$$

it is worthwhile noting that (47) is a PDE, not an ODE. With this choice, it follows from (44) that

$$\begin{aligned}
&\dot{V}\left(e_N(t), \tilde{\alpha}(t), \tilde{f}_N(t, \cdot)\right)\\
&\quad = -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle\left(I - \Pi_N\right)E_{x_N(t)}^* B^{\mathrm{T}}Pe_N(t), \tilde{f}_N\rangle_{\mathcal{H}}\\
&\quad \leq -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle B^{\mathrm{T}}Pe_N(t), E_{x(t)}(I - \Pi_N)f\rangle_{\mathcal{H}}, \quad t \geq t_0,
\end{aligned} \tag{48}$$

which is not sign-definite, and, hence, any Lyapunov argument would be inconclusive. To overcome this situation, in the following, we apply robust adaptive laws to guarantee the stability of both the closed-loop trajectory tracking error dynamics and the adaptive laws.

7. The σ-modification

7.1. A relaxed functional uncertainty approach

The first method we study to account for the fact that the right-hand side of (48) is not sign definite is captured by the following theorem. This result, which is based on the classical σ-modification of MRAC (Farrell & Polycarpou, 2006; Ioannou & Sun, 2012; Sastry & Bodson, 2011), relies on the adaptive laws

$$\dot{\hat{\alpha}}_N(t) = -\Gamma_\alpha x_N(t) e_N^{\mathrm{T}}(t) PB - \sigma\hat{\alpha}_N, \quad \hat{\alpha}_N(t_0) = \hat{\alpha}_{N,0}, \quad t \geq t_0, \tag{49}$$

$$\frac{\partial \hat{f}_N(t,\cdot)}{\partial t} = \gamma_f \Pi_N E^*_{x_N(t)} B^{\mathrm{T}} P e_N(t) - \sigma \hat{f}_N(t,\cdot), \quad \hat{f}_N(t_0,\cdot) = \hat{f}_{N,0}. \tag{50}$$

where $0 < \sigma < \frac{\lambda_{\min}(Q)}{2\lambda_{\max}(Q)}$, the pair (P, Q) verifies the algebraic Lyapunov equation

$$-Q = A_{\mathrm{ref}}^{\mathrm{T}} P + A_{\mathrm{ref}} \tag{51}$$

with $Q \in \mathbb{R}^{n\times n}$ user-defined, symmetric, and positive-definite, and $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the eigenvalues of their argument with the smallest and largest real part, respectively. For the statement of this result, it is worthwhile recalling the definition of bounded on the diagonal kernel provided in Definition 5.1 of the first part of this work.

Theorem 7.1. *Suppose that the plant dynamics (38) with control input (39), and the adaptive laws (49) and (50) generate solutions $x_N(t)$, $\hat{\alpha}_N(t)$, and $\hat{f}_N(t,\cdot)$ defined for all $t \in [t_0,\infty)$. Assume further that $f \in C_R$, where the functional uncertainty class C_R is given by (11), and the kernel $\mathfrak{K}(\cdot,\cdot)$ is bounded on the diagonal. Then, for any subset $\Omega \supseteq \bigcup_{t\geq t_0>0} x_N(t)$, there is a constant $T \in (t_0,\infty)$ such that*

$$\|x_N(t) - x_r(t)\|^2_{\mathbb{X}} \leq \frac{D}{\sigma}, \tag{52}$$

for all $t \geq T$, where

$$D \triangleq \frac{2\|PB\|^2 \left(\sup_{\xi\in\Omega} \mathcal{P}_N(\xi)\right)^2 R^2}{\lambda_{\min}(Q)} + \sigma\left(\gamma_\alpha^{-1}\alpha^2 + \gamma_f^{-1} R^2\right). \tag{53}$$

Proof. Consider the Lyapunov function candidate (42). Employing the adaptive laws (49) and (50), it holds that

$$\begin{aligned}\dot{V}(t) = &-\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle\left(I-\Pi_N\right) E^*_{x_N(t)} e_N^{\mathrm{T}}(t) PB, \tilde{f}_N\rangle_{\mathcal{H}_N} \\ &+ 2\sigma\left(\tilde{\alpha}_N^{\mathrm{T}} \Gamma_\alpha^{-1}\hat{\alpha}_N\right) + 2\gamma_f^{-1}\sigma\langle\tilde{f}_N, \hat{f}_N\rangle_{\mathcal{H}_N}, \\ = &-\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle\left(I-\Pi_N\right) E^*_{x_N(t)} e_N^{\mathrm{T}}(t) PB, \tilde{f}_N\rangle_{\mathcal{H}_N} \\ &+ 2\sigma\langle\tilde{\alpha}_N, \hat{\alpha}_N\rangle_\alpha + 2\gamma_f^{-1}\sigma\langle\tilde{f}_N, \hat{f}_N\rangle_{\mathcal{H}_N}, \quad t \geq t_0,\end{aligned}$$

where $V(t)$ denotes $V\left(e_N(t), \tilde{\alpha}_N(t), \tilde{f}_N(t)\right)$, and $\langle\tilde{\alpha}_N, \hat{\alpha}_N\rangle_\alpha \equiv \left(\tilde{\alpha}_N^{\mathrm{T}}\Gamma_\alpha^{-1}\hat{\alpha}_N\right)$ denotes a weighted inner product. We then can apply the identity

$$\langle a, b-a\rangle = -\langle a, a-b\rangle = -\left(\frac{1}{2}\langle a-b, a-b\rangle + \frac{1}{2}\langle a,a\rangle - \frac{1}{2}\langle b,b\rangle\right),$$

to $\langle\hat{\alpha},\tilde{\alpha}\rangle_\alpha = \langle\hat{\alpha}, \alpha-\hat{\alpha}\rangle_\alpha$ and $\langle\hat{f}_N,\tilde{f}_N\rangle_{\mathcal{H}_N} = \langle\hat{f}_N, f-\hat{f}_N\rangle_{\mathcal{H}_N}$. Thus,

$$\begin{aligned}\dot{V}(t) = &-\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle\left(I-\Pi_N\right) E^*_{x_N(t)} e_N^{\mathrm{T}}(t) PB, \tilde{f}\rangle \\ &- \sigma\left(\langle\tilde{\alpha}(t),\tilde{\alpha}(t)\rangle_\alpha + \langle\hat{\alpha}(t),\hat{\alpha}(t)\rangle_\alpha - \langle\alpha,\alpha\rangle_\alpha\right) \\ &- \gamma_f^{-1}\sigma\left(\|\tilde{f}_N(t,\cdot)\|^2_{\mathcal{H}_N} + \|\hat{f}_N(t,\cdot)\|^2_{\mathcal{H}_N} - \|f\|^2_{\mathcal{H}}\right), \quad t\geq t_0.\end{aligned}$$

Removing the guaranteed negative terms that contain $\hat{\alpha}^2(\cdot)$ and $\|\hat{f}_N(\cdot,\cdot)\|^2_{\mathcal{H}_N}$ gives us

$$\begin{aligned}\dot{V}(t) \leq &-\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle\left(I-\Pi_N\right) E^*_{x_N(t)} e_N^{\mathrm{T}}(t) PB, \tilde{f}_N(t,\cdot)\rangle_{\mathcal{H}_N} \\ &- \sigma\langle\tilde{\alpha}(t),\tilde{\alpha}(t)\rangle_\alpha + \sigma\langle\alpha,\alpha\rangle_\alpha - \gamma_f^{-1}\sigma\|\tilde{f}_N(t,\cdot)\|^2_{\mathcal{H}_N} + \gamma_f^{-1}\sigma\|f\|^2_{\mathcal{H}},\end{aligned} \tag{54}$$

for all $t \geq t_0$. From the first line of (54), we have

$$\begin{aligned}&-e_N^{\mathrm{T}}(t) Q e_N(t) + 2\langle\left(I-\Pi_N\right) E^*_{x_N(t)} e_N^{\mathrm{T}}(t) PB, \tilde{f}_N(t,\cdot)\rangle_{\mathcal{H}_N} \\ &= -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle B^{\mathrm{T}} P e_N^{\mathrm{T}}(t), E_{x_N(t)}\left(I-\Pi_N\right)\tilde{f}_N(t,\cdot)\rangle_{\mathbb{R}^n}, \\ &= -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle e_N(t), PBE_{x_N(t)}\left(I-\Pi_N\right) f\rangle_{\mathbb{R}^n}, \\ &\leq -\lambda_{\min}(Q)\|e_N(t)\|^2 + 2\|e_N(t)\|\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|, \\ &\leq -\frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|^2 \\ &\quad - 2\lambda_{\min}(Q)\left(\frac{1}{4}\|e_N(t)\|^2 - \frac{\|e_N(t)\|\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|}{\lambda_{\min}(Q)}\right), \\ &\quad t\geq t_0.\end{aligned}$$

Within the bracket, we get

$$\begin{aligned}&\frac{1}{4}\|e_N(t)\|^2 - \frac{\|e_N(t)\|\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|}{\lambda_{\min}(Q)} \\ &= \frac{1}{4}\|e_N(t)\|^2 - \nu(t)\|e_N(t)\| \\ &= \left(\frac{1}{2}\|e_N(t)\| - \nu(t)\right)^2 - \nu(t)^2, \quad t\geq t_0,\end{aligned} \tag{55}$$

where $\nu(t) \triangleq -\frac{\|PBE_{x_N(t)}(I-\Pi_N) f\|}{\lambda_{\min}(Q)}$. Thus, it follows from (55) that (54) is equivalent to

$$\begin{aligned}\dot{V}(t) \leq &-\frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|^2 - 2\lambda_{\min}(Q)\left[\left(\frac{1}{2}\|e_N(t)\| - \nu(t)\right)^2 - \nu^2(t)\right] \\ &- \sigma\langle\tilde{\alpha}(t),\tilde{\alpha}(t)\rangle_\alpha + \sigma\langle\alpha,\alpha\rangle_\alpha - \gamma_f^{-1}\sigma\|\tilde{f}_N(t,\cdot)\|^2_{\mathcal{H}_N} + \gamma_f^{-1}\sigma\|f\|^2_{\mathcal{H}} \\ \leq &-\frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|^2 + 2\lambda_{\min}(Q)\left(\frac{\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|}{\lambda_{\min}(Q)}\right)^2 \\ &- \sigma\langle\tilde{\alpha}(t),\tilde{\alpha}(t)\rangle_\alpha + \sigma\langle\alpha,\alpha\rangle_\alpha - \gamma_f^{-1}\sigma\|\tilde{f}_N(t,\cdot)\|^2_{\mathcal{H}_N} + \gamma_f^{-1}\sigma\|f\|^2_{\mathcal{H}}, \\ \leq &-\left[\sigma e_N^{\mathrm{T}}(t) Q e_N(t) + \sigma\left(\langle\tilde{\alpha}(t),\tilde{\alpha}(t)\rangle_\alpha + \gamma_f^{-1}\|\tilde{f}_N(t,\cdot)\|^2_{\mathcal{H}_N}\right)\right] \\ &- \frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|^2 + \sigma e_N^{\mathrm{T}}(t) Q e_N(t) + \frac{2\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|^2}{\lambda_{\min}(Q)} \\ &+ \sigma\left(\langle\alpha,\alpha\rangle_\alpha + \gamma_f^{-1}\|f\|^2_{\mathcal{H}}\right) \\ \leq &-\sigma V(t) - \frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|^2 + \sigma\lambda_{\max}(Q)\|e_N(t)\|^2 \\ &+ \frac{2\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|^2}{\lambda_{\min}(Q)} + \sigma\left(\langle\alpha,\alpha\rangle_\alpha + \gamma_f^{-1}\|f\|^2_{\mathcal{H}}\right) \\ \leq &-\sigma V(t) - \left(\frac{1}{2}\lambda_{\min}(Q) - \sigma\lambda_{\max}(Q)\right)\|e_N(t)\|^2 \\ &+ \frac{2\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|^2}{\lambda_{\min}(Q)} + \sigma\left(\langle\alpha,\alpha\rangle_\alpha + \gamma_f^{-1}\|f\|^2_{\mathcal{H}}\right), \quad t\geq t_0.\end{aligned}$$

By hypothesis, $\frac{1}{2}\lambda_{\min}(Q) - \sigma\lambda_{\max}(Q) > 0$. Therefore,

$$\begin{aligned}\dot{V}(t) \leq &-\sigma V(t) + \frac{2\|PBE_{x_N(t)}\left(I-\Pi_N\right) f\|^2}{\lambda_{\min}(Q)} \\ &+ \sigma\left(\gamma_\alpha^{-1}\alpha^2 + \gamma_f^{-1}\|f\|^2_{\mathcal{H}}\right), \quad t\geq t_0.\end{aligned}$$

Initially, let $\Omega = \mathbb{R}^n$, and recall that, by (23),

$$\left\|E_{x_N(t)}(I-\Pi_N) f\right\| \leq \sup_{\xi\in\Omega} \mathcal{P}_N(\xi)\,\|f\|_{\mathcal{H}},$$

and, by the assumption of boundedness of the kernel on the diagonal, we obtain that

$$\mathcal{P}_N(x) = \|(I-\Pi_N)\mathfrak{K}_x\|_{\mathcal{H}} \leq \|\mathfrak{K}_x\|_{\mathcal{H}} \leq \bar{\mathfrak{K}},$$

which implies that $\left|E_{x_N(t)}(I-\Pi_N) f\right| < \infty$. Thus, it follows from (53) that

$$\dot{V}(t) \leq -\sigma V(t) + D, \quad t\geq t_0, \tag{56}$$

and hence,

$$V(t) \leq C_1 e^{-\sigma t} + \frac{D}{\sigma}, \quad t\geq t_0, \tag{57}$$

where C_1 denotes a coefficient that depends on the initial condition of the system. Inequality (57) shows that all the trajectories are bounded and $V(\cdot)$ converges exponentially fast to the level set

$$S_0 \triangleq \left\{\left(e_N, \tilde{\alpha}_N, \tilde{f}_N\right) \in \mathbb{Z} : V\left(e_N, \tilde{\alpha}_N, \tilde{f}_N\right) \leq \frac{D}{\sigma}\right\} \tag{58}$$

and enters it in a finite time $T > t_0$. During this first pass, we set $\Omega = \mathbb{X} = \mathbb{R}^n$. Now that we know that trajectories are bounded, we can choose Ω to be any subset that contains the tail of the trajectory, that is, for all $t \geq t_0 \geq 0$. The smaller the set Ω, the tighter the bounds on performance. □

Theorem 7.1 proves the effectiveness of the control system comprising the control input (39), and the adaptive laws (49) and (50), and quantified the trajectory tracking error assuming that $f \in \mathcal{C}_R$. In the following, we assess the practical tracking of this control architecture for multiple classes of functional uncertainties.

7.2. Practical tracking of the σ-modification

In the following, we assess the performance of the control architecture proposed in Theorem 7.1 for the functional uncertainty classes $\mathcal{C}_{\Phi_N,R}$, $\mathcal{C}_{N,\epsilon,R}$, and $\mathcal{C}_R$ given by (9)–(11). To this goal, consider the cost function (35).

It follows from (23) that

$$\left|E_{x_N(t)}(I-\Pi_N)f\right| \leq \|(I-\Pi_N)f\|_{\mathcal{H}}\mathcal{P}_N(x_N(t)), \quad t \geq t_0.$$

Now, if $f \in \mathcal{C}_{\Phi_N,R}$, then $(I-\Pi_N)f = 0$. A review of the proof of Theorem 7.1 shows that for any uncertainty f in the smallest uncertainty class $\mathcal{C}_{\Phi_N,R}$, we have

$$J^2(x_N) \leq \frac{\gamma_\alpha^{-1}\|\alpha\|^2_{\mathbb{R}^n} + \gamma_f^{-1}R^2}{\lambda_{\min}(P)}. \tag{59}$$

Alternatively, if $f \in \mathcal{C}_{N,\epsilon,R}$, then proceeding as in the proof of Theorem 7.1, we have that

$$J^2(x_N) \leq \frac{2\|PB\|^2\epsilon^2R^2}{\sigma\lambda_{\min}(P)\lambda_{\min}(Q)} + \frac{\gamma_\alpha^{-1}\|\alpha\|^2_{\mathbb{R}^n} + \gamma_f^{-1}R^2}{\lambda_{\min}(P)}. \tag{60}$$

Finally, if $f \in \mathcal{C}_R$, then, as shown by Theorem 7.1, it holds that

$$J^2(x_N) \leq \frac{2\|PB\|^2}{\sigma\lambda_{\min}(P)\lambda_{\min}(Q)}\left(\sup_{\xi\in\Omega}\mathcal{P}_N(\xi)\right)^2 R^2 + \frac{\gamma_\alpha^{-1}\|\alpha\|^2_{\mathbb{R}^n} + \gamma_f^{-1}R^2}{\lambda_{\min}(P)}. \tag{61}$$

It is worthwhile to note how the upper bounds on $J(\cdot)$ increase as the size of the uncertainty class increases.

7.3. Coordinate implementations for realizable controllers

We can express the adaptive gain $\hat{f}_N(\cdot,\cdot)$ as

$$\hat{f}_N(t,\cdot) = \sum_{j=1}^{N}\hat{\theta}_{N,j}(t)\mathfrak{K}_{\xi_{N,j}}(\cdot), \quad t \geq t_0, \tag{62}$$

where $\mathfrak{K}_{\xi_{N,j}}$ denotes the kernel basis function centered around $\xi_{N,j}$ and $\hat{\theta}_{N,j}(\cdot)$ denotes the corresponding coefficient. Substituting (62) into (50) yields the functional equation

$$\sum_{j=1}^{N}\dot{\hat{\theta}}_{N,j}(t)\mathfrak{K}_{\xi_{N,j}}(\cdot) = \gamma_f\Pi_N\mathfrak{K}_{x_N(t)}e_N^{\mathrm{T}}(t)PB - \sigma\sum_{j=1}^{N}\hat{\theta}_{N,j}(t)\mathfrak{K}_{\xi_{N,j}}(\cdot), \tag{63}$$

for all $t \geq t_0$. By the reproducing property, for any $\alpha \in \mathbb{R}$ and for each $t \in [t_0,\infty)$, it holds that $E^*_{x_N(t)}\alpha = \mathfrak{K}_{x_N(t)}\alpha$. Thus, taking the inner product of (63) with $\mathfrak{K}_{\xi_{N,i}}$, for any $i = 1,\dots,N$, gives us

$$\sum_{j=1}^{N}\langle\mathfrak{K}_{\xi_{N,i}},\mathfrak{K}_{\xi_{N,j}}\rangle_{\mathcal{H}}\dot{\hat{\theta}}_{N,j}(t) = \gamma_f\langle\mathfrak{K}_{\xi_{N,i}},\Pi_N\mathfrak{K}_{x_N(t)}\rangle_{\mathcal{H}}e_N^{\mathrm{T}}(t)PB - \sigma\sum_{j=1}^{N}\langle\mathfrak{K}_{\xi_{N,i}},\mathfrak{K}_{\xi_{N,j}}\rangle_{\mathcal{H}}\hat{\theta}_{N,j}(t), \quad t \geq t_0. \tag{64}$$

Thus, it follows from the definition of Grammian matrix (29) that

$$\sum_{j=1}^{N}\mathbb{K}_{ij}\dot{\hat{\theta}}_{N,j}(t) = \gamma_f\mathfrak{K}\left(\xi_{N,i},x_N(t)\right)e_N^{\mathrm{T}}(t)PB - \sigma\sum_{j=1}^{N}\mathbb{K}_{ij}\hat{\theta}_{N,j}(t), \quad t \geq t_0,$$

for any $i = 1,\dots,N$. Introducing the vectors $\hat{\Theta}_N = \left[\hat{\theta}_1,\dots,\hat{\theta}_N\right]^{\mathrm{T}}$ and $\mathfrak{K}_{\Xi_N}(\cdot) = \left[\mathfrak{K}_{\xi_{N,1}}(\cdot),\dots,\mathfrak{K}_{\xi_{N,N}}(\cdot)\right]^{\mathrm{T}}$, (64) is equivalent to

$$\mathbb{K}_N\dot{\hat{\Theta}}_N = \gamma_f\mathfrak{K}_{\Xi_N}x_N(t)e_N^{\mathrm{T}}(t)PB - \sigma\mathbb{K}_N\hat{\Theta}_N. \tag{65}$$

Box 6: The adaptive law (66) is particularly similar to the adaptive law employed in the σ-modification of MRAC. However, the Grammian matrix $\mathcal{K}$ in (66) allows correlating explicitly the adaptive law to the error in approximating the infinite-dimensional space $\mathcal{H}$ with the finite-dimensional space $\mathcal{H}_N$. The Grammian matrix is required to make explicit guarantees in terms of N about all the coordinate realizations as N varies.

Premultiplying (65) by the inverse of the Grammian matrix then yields

$$\dot{\hat{\Theta}}_N(t) = \gamma_f\mathbb{K}_N^{-1}\mathfrak{K}_{\Xi_N}\left(x_N(t)\right)e_N^{\mathrm{T}}(t)PB - \sigma\hat{\Theta}_N(t), \qquad \hat{\Theta}_N(t_0) = \hat{\Theta}_{N,0}, \quad t \geq t_0, \tag{66}$$

which is the adaptive law for the coordinate representation.

Remarkably, (66) for the approximation in the nonparametric method has a form that resembles that of the σ-modification of MRAC (Lavretsky & Wise, 2012, Ch. 11). However, we emphasize that the nonparametric method involves the Grammian matrix. The Grammian matrix defines a consistent approximation scheme in the native space and allows defining the error estimates for the nonparametric methods in the native space $\mathcal{H}$ defined by the kernel $\mathfrak{K}(\cdot,\cdot)$.

Remark 7.1. The coordinate implementation of the adaptive law (50) provided by (66) allows solving a matrix differential equation instead of a partial differential equation. In the following, we discuss several other adaptive control techniques, whose adaptive gains associated with the functional uncertainty can be computed as solutions of PDEs in some form that is similar to (50). By proceeding as in this section, these adaptive gains can be alternatively computed as solutions of matrix ODEs.

8. The deadzone modification

8.1. A tight functional uncertainty approach

The deadzone method is another robust modification of the classical MRAC laws that has been a staple of robust real parametric adaptive control theory (Farrell & Polycarpou, 2006; Ioannou & Sun, 2012; Lavretsky & Wise, 2012; Sastry & Bodson, 2011). The deadzone modification can be applied to problems wherein the functional uncertainty is such that $f \in \mathcal{C}_{N,\epsilon,R}$, and is defined in terms of a user-defined constant d such that

$$d > d_\epsilon \triangleq \frac{2\|PB\|\lambda_{\max}(P)\bar{\mathfrak{K}}\epsilon}{\lambda_{\min}(Q)\sqrt{\lambda_{\min}(P)}}. \tag{67}$$

In this case, the adaptive laws are given by

$$\dot{\hat{\alpha}}(t) = \begin{cases} -\Gamma_\alpha x_N(t)e_N^{\mathrm{T}}(t)PB, & \|P^{1/2}e_N(t)\| > d, \\ 0 & \text{otherwise}, \end{cases} \tag{68}$$

$$\frac{\partial\hat{f}_N(t,\cdot)}{\partial t} = \begin{cases} \gamma_f\Pi_N E^*_{x_N(t)}B^{\mathrm{T}}Pe_N(t), & \|P^{1/2}e_N(t)\| > d, \\ 0 & \text{otherwise}, \end{cases} \tag{69}$$

with initial conditions chosen to guarantee the continuity of the adaptive gains at all times; remarkably, if $\|P^{1/2}e_N(t)\| > d$ for some $t \geq t_0$, then (68) and (69) reduce to (45) and (47), respectively.

Theorem 8.1. *Suppose that the plant dynamics (38) with control input (39) and the adaptive laws (68) and (69) generate solutions $x_N(t)$, $\hat{\alpha}_N(t)$, and $\hat{f}_N(t,\cdot)$ defined for all $t \in [t_0,\infty)$. Finally, assume that $f \in \mathcal{C}_{N,\epsilon,R}$, where the functional uncertainty class $\mathcal{C}_{N,\epsilon,R}$ is given by (9), and assume that the kernel $\mathfrak{K}(\cdot,\cdot)$ is bounded on the diagonal. Then, there exists $T \triangleq T(d) \in (t_0,\infty)$ such that*

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} < \frac{d}{\sqrt{\lambda_{\min}(P)}}, \tag{70}$$

for all $t \geq T$.

Box 7: Similarly to the classical deadzone modification method (Lavretsky & Wise, 2012, Ch. 11), the proposed deadzone modification method requires the adaptive laws to operate only when the trajectory tracking error is sufficiently large. Furthermore, it can be proven that the trajectory tracking error lies outside the deadzone for only a finite time interval. Despite the classical deadzone modification, the adaptive law corresponding to the functional uncertainty depends on both the projection operator and the adjunct of the evaluation operator.

Proof. If $\|P^{1/2}e_N(t)\| > d$ for some $t \geq t_0$, then it follows from (48) that

$$\dot{V}(t) = -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle \left(I - \Pi_N\right) E^*_{x_N(t)} e^{\mathrm{T}}_N(t) PB, \tilde{f}_N(t,\cdot)\rangle_{\mathcal{H}}$$
$$\leq -\lambda_{\min}(Q)\|e_N(t)\|^2 + 2\|e_N(t)\|\ \|PB\|\ \underbrace{\|E_{x_N(t)}\left(I - \Pi_N\right)\tilde{f}_N\|_{\mathcal{H}}}_{\|E_{x_N(t)}(I-\Pi_N)f\|_{\mathcal{H}}},$$

where $V(t) = V\left(e_N(t), \tilde{\alpha}(t), \tilde{f}_N(t)\right)$ for brevity. By the definition of the functional uncertainty class $C_{R,\epsilon,N}$, and using the fact that the kernel $\mathfrak{K}(\cdot,\cdot)$ is uniformly bounded on the diagonal by a constant $\bar{\mathfrak{K}} > 0$, we have

$$\|E_{x_N(t)}\left(I - \Pi_N\right) f\|_{\mathcal{H}} \leq \|\mathfrak{K}_{x_N(t)}\|_{\mathcal{H}}\ \|(I - \Pi_N)f\|_{\mathcal{H}} \leq \bar{\mathfrak{K}}\epsilon,$$

for all $t \geq t_0$. This implies that

$$\dot{V}(t) \leq -\lambda_{\min}(Q)\|e_N(t)\|^2 + 2\bar{\mathfrak{K}}\epsilon\|PB\|\|e_N(t)\|,$$
$$\leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\|P^{1/2}e_N(t)\|^2 + \frac{2\|PB\|\bar{\mathfrak{K}}\epsilon}{\sqrt{\lambda_{\min}(P)}}\|P^{1/2}e_N(t)\|$$
$$= -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\|P^{1/2}e_N(t)\|\left(\|P^{1/2}e_N(t)\| - \frac{2\|PB\|\lambda_{\max}(P)\bar{\mathfrak{K}}\epsilon}{\lambda_{\min}(Q)\sqrt{\lambda_{\min}(P)}}\right),$$
$$\leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\ \|P^{1/2}e_N(t)\|\left(\|P^{1/2}e_N(t)\| - d\right), \quad t \geq t_0. \tag{71}$$

Next, we show that there is $T \in (t_0, \infty)$ such that $\|P^{1/2}e_N(T)\| = d$. Assume that $\|P^{1/2}e_N(t_0)\| > d$; the case whereby $\|P^{1/2}e_N(t_0)\| \leq d$ can be easily addressed. If the tracking error leaves the deadzone at even indexed times t_{2i}, $i \in \mathbb{N}$, and subsequently re-enters the deadzone at odd indexed time t_{2i+1}, then $\dot{V}(t) < -C$ for all $t \in (t_{2i}, t_{2i+1}]$, where $C > 0$. This means that $V\left(t_{2i+1}\right) \leq V\left(t_{2i}\right) - C\left(t_{2i+1} - t_{2i}\right)$. Additionally, we know that $V\left(t_{2i}\right) = V\left(t_{2i-1}\right)$ for all $i = 1, 2, \ldots, M$, where $M \leq \infty$. Therefore, $\sum_{i=0}^{M}\left(t_{2i+1} - t_{2i}\right) \leq \frac{1}{C}\sum_{i=1}^{M}\left(V_{2i} - V_{2i-1}\right) \leq \frac{1}{C}V(t_0) < \infty$, which implies that the time spent outside the deadzone is finite. The tracking error $\|P^{1/2}e_N(t)\|$ is consequently bounded by d after some $T > t_0$. □

The proof of Theorem 8.1 provides an estimate of the performance of the deadzone modification when the uncertainty f is contained in the uncertainty class $C_{R,\epsilon,N}$. The upper bound on the trajectory tracking error given by (70) can be tightened as follows. If $\bigcup_{t\geq t_0>0} x_N(t) \subseteq \Omega$, where $\Omega \subset \mathbb{X}$ is bounded, then the fact that

$$\|E_{x_N(t)}(I - \Pi_N)f\| \leq \left(\sup_{\xi\in\Omega}\mathcal{P}_N(\xi)\right)\|(I - \Pi_N)f\| \leq \left(\sup_{\xi\in\Omega}\mathcal{P}_N(\xi)\right)\|\epsilon_N\| \tag{72}$$

for any $f \in \mathcal{H}$ and all $t \geq t_0$, which follows from (23) and (32), allows establishing a tighter bound than (70). However, the deadzone method described in Theorem 8.1 does not allow to choose Ω *a priori*. Barrier Lyapunov functions can be employed to attain this goal (Anderson, Marshall, & L'Afflitto, 2021a; Arabi, Gruenwald, Yucelen, & Nguyen, 2018; L'Afflitto, 2018). Alternatively, it is possible to include control feedback terms, in addition to the deadzone control term, which ensures the closed-loop trajectory remains in the set Ω that is chosen at the outset. This is the philosophy of approximation-based adaptive control described in Chapter 12 of Lavretsky and Wise (2012), Sections 6.2–6.4 and Chapter 7 of Farrell and Polycarpou (2006), and Theorems 7.2.1–7.2.3 and Section 7.2.4 of Farrell and Polycarpou (2006), which assume suitable addition controls to guarantee that the approximation set is entered in finite time and is positive invariant.

The next results characterize the performance of deadzone modification assuming that there is a control law that allows constraining both the trajectory tracking error and the adaptive gains within a bounded subset Ω. The goal is to obtain rather simple performance bounds on the performance of the control scheme without all the attendant details required to carry out precise, complicated Lyapunov arguments for the specific, complete control scheme. This result supposes that the deadzone is characterized by a constant d such that

$$d > d_N \triangleq \frac{2\|PB\|\lambda_{\max}(P)}{\lambda_{\min}(Q)\sqrt{\lambda_{\min}(P)}}\left(\sup_{\xi\in\Omega}\mathcal{P}_N(\xi)\right)\|\epsilon_N\|_{\mathcal{H}}. \tag{73}$$

Furthermore, the next result considers the control input

$$u_N(t) = \begin{cases} \hat{\alpha}^{\mathrm{T}}_N(t)x_N(t) - E_{x_N(t)}\hat{f}_N(t,\cdot), & \text{for all } x_N(t) \in \Omega, \\ v_N(x_N(t), e_N(t)) & \text{otherwise,} \end{cases} \tag{74}$$

where $\hat{\alpha}_N(t)$ and $\hat{f}_N(t,\cdot)$ verify the adaptive laws (45) and (47) for all $t \in [t_0, \infty)$ such that $x_N(t) \in \Omega$ and the control law $v_N : \mathbb{X} \times \mathbb{R}^n \to \mathbb{R}$ renders the approximation set Ω attractive and positive invariant. Such solutions are often useful whenever the control law $v_N(\cdot,\cdot)$, whose performance is poorly characterized or not as good as the adaptive controller's performance, which is characterized by the next theorem. Indeed, for practical purposes, the term $v_N(x_N(t), e_N(t))$ is expected to work for all t over an interval that is smaller than the interval over which the term $\hat{\alpha}^{\mathrm{T}}_N(t)x_N(t) - E_{x_N(t)}\hat{f}_N(t,\cdot)$ is expected to work.

Theorem 8.2. *Suppose that (38) with control input (74) and the adaptive laws (45) and (47) generate solutions $x_N(t)$, $\hat{\alpha}_N(t)$, and $\hat{f}_N(t,\cdot)$ defined for all $t \in [t_0, \infty)$. Assume further that the functional uncertainty $f \in C_{R,\epsilon,N}$ given by (9), and that the kernel $\mathfrak{K}(\cdot,\cdot)$ is bounded on the diagonal. Then, there exists $T \triangleq T(d) \in (t_0, \infty)$ such that*

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} < \frac{d}{\sqrt{\lambda_{\min}(P)}}, \tag{75}$$

for all $t \geq T$.

Proof. With the hypotheses on the compensator $v_N(\cdot,\cdot)$ and the approximation set Ω, the theorem statement should be verified only for $x_N(t) \in \Omega$. If $\|P^{1/2}e_N(t)\| > d$ and $x_N(t) \in \Omega$ for some $t \in [t_0, \infty)$, then the time derivative of the Lyapunov function (42) along the trajectories of (38) with control input (74) and (45) and (47) is given by

$$\dot{V}(t) = -\langle Qe(t), e(t)\rangle_{\mathbb{R}^n} + 2\langle \left(I - \Pi_N\right) E^*_{x_N(t)} e^{\mathrm{T}}_N(t) PB, \tilde{f}_N\rangle$$
$$\leq -\lambda_{\min}(Q)\|e_N(t)\|^2 + 2\|e_N(t)\|\ \|PB\|\ \underbrace{\|E_{x_N(t)}\left(I - \Pi_N\right)\tilde{f}_N\|}_{\|E_{x_N(t)}(I-\Pi_N)f\|}, \quad t \geq t_0,$$

where $V(t) = V\left(e_N(t), \tilde{\alpha}(t), \tilde{f}_N(t)\right)$ for brevity. Now, it follows from (72) that

$$\dot{V}(t) \leq -\lambda_{\min}(Q)\|e_N(t)\|^2 + 2\|PB\|\|\epsilon_N\|\|e_N(t)\|\sup_{\xi\in\Omega}\mathcal{P}_N(\xi),$$
$$\leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\|P^{1/2}e_N(t)\|^2 + \frac{2\|PB\|}{\sqrt{\lambda_{\min}(P)}}\|\epsilon_N\|\|P^{1/2}e_N(t)\|\sup_{\xi\in\Omega}\mathcal{P}_N(\xi)$$
$$= -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\|P^{1/2}e_N(t)\|\left(\|P^{1/2}e_N(t)\| - \frac{2\|PB\|\lambda_{\max}(P)}{\lambda_{\min}(Q)\sqrt{\lambda_{\min}(P)}}\|\epsilon_N\|\sup_{\xi\in\Omega}\mathcal{P}_N(\xi)\right)$$
$$\leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\ \|P^{1/2}e_N(t)\|\left(\|P^{1/2}e_N(t)\| - d\right), \quad t \geq t_0. \tag{76}$$

The tracking error $\|P^{1/2}e_N(t)\|$ is consequently bounded by d for all $t \geq T$ for some $T \in (t_0, \infty)$. The remainder of the proof follows that of Theorem 8.1. □

Following the same approach as in Section 7.3, the adaptive law for the deadzone modification can be implemented as follows If $\|e_N(t)\| > \epsilon$, for some $t \geq t_0$, then

$$\dot{\hat{\Theta}}_N(t) = \gamma_f \mathbb{K}^{-1}\left(\Xi_N, \Xi_N\right) \mathfrak{K}_{\Xi_N}\left(x_N(t)\right) e_N^{\mathrm{T}}(t) P B, \quad \hat{\Theta}_N(t_0) = \hat{\Theta}_{N,0}, \tag{77}$$

and if $\|e_N(t)\| \leq \epsilon$, then $\dot{\hat{\Theta}}_N(t) = 0$.

As it appears from (12), the assumption whereby $f \in C_{N,\epsilon,R}$ may be tight. In the following, we discuss how this assumption can be lifted and the results proposed thus far can be extended to broader classes of functional uncertainties.

8.2. Relaxed functional uncertainty approaches

In principle, the deadzone defined by d in (73) yields the tightest ultimate performance bounds if we choose $d = (1+\eta)d_N$ for an arbitrarily small constant η. In this case, when the controller trajectory remains in Ω, we conclude that we have the ultimate bound

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq (1+\eta)\frac{2\|PB\|\lambda_{\max}(P)}{\lambda_{\min}(Q)\lambda_{\min}(P)}\left(\sup_{\xi\in\Omega} \mathcal{P}_N(\xi)\right)\|(I-\Pi_N)f\|_{\mathcal{H}}, \tag{78}$$

for all $t \geq T$ and some $T \in (t_0, \infty)$. In practice, however, we cannot choose such a deadzone since $\|(I-\Pi_N)f\|_{\mathcal{H}}$ is unknown, and, hence, ϵ_N is unknown. It follows from (73) that, if $f \in C_{R,\epsilon,N}$, then an alternative to (78) is given by

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq (1+\eta)\frac{2\|PB\|\lambda_{\max}(P)}{\lambda_{\min}(Q)\lambda_{\min}(P)}\left(\sup_{\xi\in\Omega} \mathcal{P}_N(\xi)\right)\epsilon, \tag{79}$$

for all $t \geq T$ and some $T \in (t_0, \infty)$, The right-hand side of (79) can be evaluated in practice and used for the design of the deadzone d.

If we assume that $f \in C_R$, then, by proceeding as in Section 8.1, it is possible to prove that

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq (1+\eta)\frac{2\|PB\|\lambda_{\max}(P)}{\lambda_{\min}(Q)\lambda_{\min}(P)}\left(\sup_{\xi\in\Omega} \mathcal{P}_N(\xi)\right)R, \tag{80}$$

for all $t \geq T$ and some $T \in (t_0, \infty)$. The right-hand side of (80) can be evaluated in practice and used for the design of d.

9. An error bounding method

Having addressed the problem of extending classical modifications of MRAC to problems involving infinite-dimensional uncertainties, in this section, we discuss an approach involving a variable structure framework; it is worthwhile recalling that variable structure controllers comprise, among many others, sliding mode control, super-twisting control, and other higher-order control systems inspired by the sliding model technology. In this section, in particular, we leverage the native space embedding modifications of the error bounding control methods described in Chapters 4–7 of Farrell and Polycarpou (2006). While the section presents a particular error bounding adaptive control method, the overall approach can be viewed as a template for deriving other general methods based on other existing approaches of real parametric adaptive control theory. In the following, for brevity, we will assume that A in (1), or, equivalently, (38), is known, and, hence, the matching conditions (3) and (4) can be verified. Thus, the only uncertainty in (38) lies in f. By proceeding as in Sections 7 and 8, the proposed results can be extended to the case whereby A is unknown.

For the statement of the results in this section, it follows from (32) and the definition of evaluation functional that the pointwise error of best approximation from $\mathcal{H}_N$ is given by

$$\epsilon_N(x) = \|E_x(I-\Pi_N)f\|_{\mathcal{H}}, \quad \text{for all } x \in \mathbb{X}. \tag{81}$$

Thus, it follows from (23) that

$$\epsilon_N(x) \leq \mathcal{P}_N(x)\|(I-\Pi_N)f\|_{\mathcal{H}} \leq \mathcal{P}_N(x)\|f\|_{\mathcal{H}}, \tag{82}$$

> **Box 8**: The proposed error bounding method can be considered a variable structure control system. Indeed, the control input comprises both a compensator, which changes its structure according to whether the trajectory tracking error lies outside or inside a boundary layer, and an adaptive control term. Similar to any variable structure method, the smaller is the diameter of the boundary layer, the smaller is the ultimate bound on the trajectory tracking error, and the higher are the frequencies of oscillation in the control input.

where $\mathcal{P}_N(x)$ is given by (27). Finally, we define the *pointwise approximation error bound over the uncertainty class* C_R as

$$\bar{\epsilon}_N(x) \triangleq \mathcal{P}_N(x)R, \quad \text{for all } x \in \mathbb{X}, \tag{83}$$

and note that

$$\epsilon_N(x) \leq \bar{\epsilon}_N(x). \tag{84}$$

Now, it follows from (27) that $\mathcal{P}_N(\cdot)$ depends on the location of centers Ξ_N. If the location of the centers Ξ_N is user-defined, then the upper bound $\bar{\epsilon}_N(x)$ can be computed for any $x \in \mathbb{X}$.

Given the plant dynamics (1) and the reference model (2), the controller considered in this section is given by

$$u_N(t) = \alpha^{\mathrm{T}} x_N(t) + \beta r(t) - E_{x_N(t)}\left(\hat{f}_N(t,\cdot) + v_N(\cdot, e_N(t))\right), \quad t \geq t_0 \tag{85}$$

where $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ verify the matching conditions (3) and (4), respectively,

$$v_N(\cdot, e_N(t)) \triangleq \begin{cases} \operatorname{sign}(B^{\mathrm{T}} P e_N(t))\bar{\epsilon}_N(\cdot) & \text{if } e_N(t) \notin \mathcal{B}_\varepsilon(\mathcal{S}), \\ \frac{1}{\varepsilon} B^{\mathrm{T}} P e_N(t)\bar{\epsilon}_N(\cdot) & \text{otherwise,} \end{cases} \tag{86}$$

denotes the *compensator*,

$$\mathcal{S} \triangleq \left\{\eta \in \mathbb{X} : B^{\mathrm{T}} P \eta = 0\right\}, \tag{87}$$

$$\mathcal{B}_\varepsilon(\mathcal{S}) \triangleq \{\eta \in \mathbb{X} : |B^{\mathrm{T}} P \eta| \leq \varepsilon\}, \tag{88}$$

$\operatorname{sign} : \mathbb{R} \to \{-1, 0, 1\}$ denotes the *signum* function, $\varepsilon \geq 0$ is a user-defined constant, not to be confused with ϵ in (10), $\hat{f}_N : [t_0, \infty) \to \mathbb{R}$ verifies the adaptive law

$$\frac{\partial \hat{f}_N(t,\cdot)}{\partial t} = \begin{cases} \gamma_f \Pi_N E^*_{x_N(t)} B^{\mathrm{T}} P e_N(t), \\ \quad \text{if } \varepsilon = 0, \text{ or } \varepsilon > 0 \text{ and } \|e_N(t)\|_{\mathbb{X}} \leq \frac{4\|B^{\mathrm{T}} P\|\bar{\epsilon}_{N,\varepsilon}}{\lambda_{\min}(Q)}, \\ 0, \quad \text{otherwise} \end{cases} \tag{89}$$

with the same initial conditions as (47), and

$$\Omega_\varepsilon \triangleq \bigcup_{t\geq t_0}\left\{x \in \mathbb{X} : \|x - x_r(t)\|_{\mathbb{X}} \leq \varepsilon\right\}, \tag{90}$$

$$\bar{\epsilon}_{N,\varepsilon} \triangleq \sup_{\xi\in\Omega_\varepsilon} \bar{\epsilon}_N(\xi) \tag{91}$$

defining the *approximation region* and the *maximum user error bound* $\epsilon_{N,\varepsilon}$ *over* Ω_ε, respectively. Interpreting $\mathcal{S}$ as a *sliding surface* and $\mathcal{B}_\varepsilon(\mathcal{S})$ as and the *boundary layer of width* ϵ *around the sliding surface* $\mathcal{S}$, (85) provides an adaptive variable structure controller. It is worthwhile noting that if $\varepsilon > 0$, then $v_N(\cdot,\cdot)$ is continuous in both arguments. Alternatively, if $\varepsilon = 0$, then (85) is discontinuous over the sliding surface $\mathcal{S}$.

Theorem 9.1. *Suppose that the trajectory tracking error dynamics (38) with control input (85) and compensator (86) and the adaptive law (89) generate solutions $x_N(t)$ and $\hat{f}_N(t,\cdot)$ defined on $[t_0,\infty)$. Finally, assume that $f \in C_R$, where C_R is given by (11), and assume that the kernel $\mathfrak{K}(\cdot,\cdot)$ that induces the native space $\mathcal{H}$ of real-valued functions over $\mathbb{X}$ is bounded on the diagonal. If $\varepsilon = 0$, then*

$$\lim_{t\to\infty} \|x_N(t) - x_r(t)\|_{\mathbb{X}} = 0. \tag{92}$$

Alternatively, if $\varepsilon > 0$, then for any (arbitrarily small) constant $\eta > 0$ there exists a constant $T \triangleq T(\eta) \in (t_0, \infty)$ such that

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \le (1+\eta)\frac{4\|PB\|}{\lambda_{\min}(Q)}\left(\sup_{\xi\in\Omega_\varepsilon}\mathcal{P}_N(\xi)\right)R, \quad t \ge T. \tag{93}$$

Proof. Employing the control law (85), the trajectory tracking error dynamics are such that

$$\dot{e}_N(t) = A_{\mathrm{ref}}e_N(t) + BE_{x_N(t)}\left(\tilde{f}_N(t,\cdot) - v_N(\cdot, e_N(t))\right), \quad e_N(t_0) = x_0 - x_{r,0}, \quad t \ge t_0. \tag{94}$$

Now, consider the Lyapunov function candidate

$$V(e,\tilde{f}) = \frac{1}{2}\langle Pe, e\rangle_{\mathbb{R}^n} + \frac{1}{2}\gamma^{-1}\langle \tilde{f},\tilde{f}\rangle_{\mathcal{H}}, \quad (e,\tilde{f}) \in \mathbb{R}^n\times\mathcal{H}. \tag{95}$$

The time derivative of $V(\cdot,\cdot)$ along the trajectories of (1) with control input (85) and the compensator (86), and the trajectories of the adaptive law (89) is given by

$$\dot{V}\left(e_N(t),\tilde{f}_N(t,\cdot)\right) = -\frac{1}{2}\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n} + \left\langle B^{\mathrm{T}}Pe_N(t), E_{x_N(t)}\underbrace{(I-\Pi_N)\tilde{f}_N(t,\cdot)}_{(I-\Pi_N)f} - E_{x_N(t)}v_N\right\rangle_{\mathbb{R}},$$

for all $t \ge t_0$.

Next, we discuss the case whereby $\varepsilon = 0$. Successively, we address the case whereby $\varepsilon > 0$. If $\epsilon = 0$, then $v_N(\cdot, e_N(\cdot))$ is discontinuous in t whenever $e_N(t) \in \mathcal{S}$. In this case,

$$\dot{V}\left(e_N(t),\tilde{f}_N(t,\cdot)\right) = -\frac{1}{2}\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n} + |B^{\mathrm{T}}Pe_N(t)|\underbrace{\left(\mathrm{sign}(B^{\mathrm{T}}Pe_N(t))\epsilon_N(x_N(t)) - \bar{\epsilon}_N(x_N(t))\right)}_{\le 0},$$

for all $t \ge t_0$. Now, since we have

$$\dot{V}(e_N(t),\tilde{f}_N(t,\cdot)) \le -\frac{1}{2}\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n}, \quad t \ge t_0,$$

the conclusion follows from an application of Barbalat's lemma as follows. Since $V\left(e_N(t),\tilde{f}_N(t,\cdot)\right)$ is nonincreasing, we infer that $e_N \in L^\infty([t_0,\infty),\mathbb{R}^n)$ and $\tilde{f}_N \in L^\infty([t_0,\infty),\mathcal{H})$. Furthermore, integrating $\dot{V}\left(e_N(t),\tilde{f}_N(t,\cdot)\right)$, we conclude that $e_N \in L^2([t_0,\infty),\mathbb{R}^n)$. Furthermore, we can obtain that $\dot{e}_N \in L^\infty([t_0,\infty),\mathbb{R}^n)$ since

$$\|\dot{e}_N(t)\|_{\mathbb{R}^n} \le \|A\|\|e_N\|_{L^\infty([t_0,\infty),\mathbb{R}^n)} + \underbrace{\|E_{x(t)}\|}_{\le\bar{\mathfrak{K}}}\underbrace{\|\tilde{f}_N(t,\cdot)\|_{\mathcal{H}}}_{\le\|\tilde{f}_N\|_{L^\infty([t_0,\infty),\mathcal{H})}}, \quad t \ge t_0. \tag{96}$$

Since $e_N \in L^\infty([t_0,\infty),\mathbb{R}^n)\bigcap L^2([t_0,\infty),\mathbb{R}^n)$ and $\dot{e}_N \in L^\infty([t_0,\infty),\mathbb{R}^n)$, Barbalat's lemma implies that (92) is verified.

If $\varepsilon > 0$, then $v_N(\cdot, e_N(\cdot))$ is continuous in t. If $|B^{\mathrm{T}}Pe_N(t)| \ge \epsilon$ for some $t \in [t_0,\infty)$, then

$$\begin{aligned}\dot{V}\left(e_N(t),\tilde{f}_N(t,\cdot)\right) &= -\frac{1}{2}\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n} + |B^{\mathrm{T}}Pe_N(t)|\underbrace{\left(\mathrm{sign}(B^{\mathrm{T}}Pe_N(t))\epsilon_N(x_N(t)) - \bar{\epsilon}_N(x_N(t))\right)}_{\le 0}\\ &\le -\frac{1}{2}\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n}, \quad t \ge t_0.\end{aligned}$$

Alternatively, if $|B^{\mathrm{T}}Pe_N(t)| < \epsilon$ for some $t \in [t_0,\infty)$, then it holds that

$$\begin{aligned}\dot{V}\left(e_N(t),\tilde{f}_N(t,\cdot)\right) &= -\frac{1}{2}\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n} + |B^{\mathrm{T}}Pe_N(t)|\Big(\mathrm{sign}(B^{\mathrm{T}}Pe_N(t))\epsilon_N(x_N(t)) - \frac{1}{\varepsilon}\bar{\epsilon}_N(x_N(t))|B^{\mathrm{T}}Pe_N(t)|\Big)\\ &\le -\frac{1}{2}\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n} + |B^{\mathrm{T}}Pe_N(t)|\left(2\sup_{\xi\in\Omega_\varepsilon}\bar{\epsilon}_N(\xi)\right)\\ &\le -\frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|\left(\|e_N(t)\| - \frac{4|B^{\mathrm{T}}P|}{\lambda_{\min}(Q)}\bar{\epsilon}_{N,\varepsilon}\right).\end{aligned}$$

We conclude that $\dot{V}\left(e_N(t),\tilde{f}_N(t,\cdot)\right) < 0$ for all $t \in [t_0,\infty)$ such that

$$e_N(t) \notin \left\{\eta \in \mathbb{X} : \|\eta\|_{\mathbb{X}} \le \frac{4\|B^{\mathrm{T}}P\|}{\lambda_{\min}(Q)}\bar{\epsilon}_{N,\varepsilon}\right\}.$$

Finally, proceeding as in Farrell and Polycarpou (2006, pp. 251, 270) or the discussion in Khalil (2002, Sec. 4.8), we deduce that the amount of time that the tracking error is outside the residual set $\left\{x \in \mathbb{X} : \|x\|_{\mathbb{X}} \le \frac{4\|B^{\mathrm{T}}P\|}{\lambda_{\min}(Q)}\bar{\epsilon}_{N,\varepsilon}\right\}$ is finite, and (93) is proven. □

Remark 9.1. Theorem 9.1 provides an estimate of the trajectory tracking error's uniform ultimate bound, which is described by (93). This ultimate bound is a function of the upper bound R on $\|f\|_{\mathcal{H}}$, the difference between the kernel function $\mathfrak{K}(\cdot,\cdot)$ and its approximation $\mathfrak{K}_N(\cdot,\cdot)$ through the power function $\mathcal{P}_N(\cdot)$, and the user-defined parameter $\lambda_{\min}(Q)$. Smaller values of R, that is, smaller uncertainties on f, produce smaller uniform ultimate bounds. If f were known to reside in $\mathcal{H}_N$, then the uniform ultimate bound would be zero. Such is the case when f is contained in the smaller real parametric uncertainty class $f \in \mathcal{C}_{\Phi_N,R} \subset \mathcal{H}_N$. Indeed, if $f \in \mathcal{C}_{\Phi_N,R}$, then, as discussed in the first paper of this two-part work, f can be written in the form of a regressor vector of basis functions in $\mathcal{H}_N$, $\epsilon_N(x) \equiv 0$, $\bar{\epsilon}_N(x) \equiv 0$, and the adaptive law (89) reduces to the classical adaptive law of MRAC; see (15) of the first paper of this two-part work. In general, larger values of N lead to better approximations of $\mathfrak{K}(\cdot,\cdot)$ by means of $\mathfrak{K}_N(\cdot,\cdot)$ and, hence, smaller uniform ultimate bounds. However, larger values of N lead to larger computational costs. Finally, in general, larger values of $\lambda_{\min}(Q)$ produce slower convergence of the tracking error.

Remark 9.2. If $\varepsilon = 0$, then the compensator $v_N(\cdot, e_N(\cdot))$ is discontinuous across the sliding surface $\mathcal{S}$. In this case, it follows from Definition 5.3 that the control system captured by Theorem 9.1 is AAO* for all $f \in \mathcal{C}_R$ and $N \in \mathbb{N}$. However, since the compensator is discontinuous, it is arguable whether we can say that such a controller is implementable. Implementation would require an actuator that has an infinite bandwidth, and no such hardware exists. Thus, chattering may be experienced.

Remark 9.3. If $\varepsilon > 0$, then the compensator $v_N(\cdot, e_N(\cdot))$ is continuous in time, and it can be implemented with actuators having a finite bandwidth. However, as shown by (93), the control system's performance degrades. Specifically, there is constant $C > 0$ such that the practical tracking attained by the control input (85) and the adaptive law (89) is given by

$$\limsup_{t\to\infty}\|x_N(t) - x_r(t)\|_{\mathbb{X}} \le C\frac{4\|PB\|}{\lambda_{\min}(Q)}\bar{\epsilon}_{N,\varepsilon} \le \underbrace{\sup_{\xi\in\Omega_\varepsilon}\mathcal{P}_N(\xi)\|f\|}_{\mathcal{F}(N)}. \tag{97}$$

Thus, the power function $\mathcal{P}_N(\cdot)$ characterizes "how suboptimal" the method is in terms of $\mathcal{F}$.

10. An adaptive error bounding method

Section 9 described how the power function $\mathcal{P}_N(\cdot)$ of the finite-dimensional subspace $\mathcal{H}_N \subseteq \mathcal{H}$ can be used systematically to derive robust adaptive control methods for functional uncertainty in native spaces. One crucial piece of information needed to implement the control strategy in Section 9 is the constant $R > 0$ in the upper bound $\|f\|_{\mathcal{H}} \le R$. Since this constant explicitly appears in the performance upper bound in (93), it is essential in practice that this upper bound be selected as tightly as possible.

In this section, we revisit the discussion outlined in Section 6 and, leveraging the general guidelines provided in Farrell and Polycarpou (2006), devise an approach whereby the radius R that characterizes

Box 9: By adding an adaptive term, which captures the norm of the functional uncertainty, the adaptive error bounding method improves on the classical error bounding method since this method does not require the user to know an upper bound on the largest admissible functional uncertainty. The adaptive bounding technique can be formulated without or with a boundary layer. In the former case, this system ensures asymptotic convergence to zero of the trajectory tracking error but may suffer from chattering. In the latter case, this system can only ensure uniform ultimate boundedness on the trajectory tracking error.

the uncertainty class (11) is fixed *a priori*, but designed as an adaptive gain.

Once again, we study the uncertain system in (1), or, equivalently, (38), that contains the scalar-valued uncertainty $f \in \mathcal{H}$, the reference system is given by (2), and the matching conditions (3) and (4) hold. Since, Section 9 only discusses for brevity the case when the function $f \in \mathcal{H}$ is uncertain, in this section, we return to the case whereby A is also unknown as we already did in Sections 7 and 8. We employ the adaptive law (45) for the adaptive gain $\hat{\alpha}(t)$, $t \geq t_0$, while the approximation of the nonparametric adaptive law (47) is used for the function estimate $\hat{f}_N(t,\cdot)$.

In the problem at hand, we introduce the time-varying adaptive gain $\hat{\lambda} : [t_0,\infty) \to \mathbb{R}$ corresponding to $\|f\|_{\mathcal{H}}$, and define the feedback control input

$$u_N(t) = \hat{\alpha}^{\mathrm{T}}(t)x_N(t) - E_{x_N(t)}\left(\hat{f}_N(t,\cdot) + w_N(\cdot, e_N(t), \hat{\lambda}(t))\right), \quad t \geq t_0, \tag{98}$$

where

$$w_N(x,e,\lambda) \triangleq -\mathrm{sign}\left(B^{\mathrm{T}}Pe\right)\mathcal{P}_N(x)\lambda, \quad \text{for all } (x,e,\lambda) \in \mathbb{X}\times\mathbb{R}^n\times\mathbb{R}, \tag{99}$$

denotes the *compensator*. It is worthwhile noting the relationship between the compensator given by (86) for $e_N \notin \mathcal{B}_\varepsilon(\mathcal{S})$ and the compensator given by (99). It follows from (11) and (83) that if $e_N \notin \mathcal{B}_\varepsilon(\mathcal{S})$, then

$$w_N(x,e,R) = -v_N(x,e), \quad \text{for all } (x,e,R) \in \mathbb{X}\times\mathbb{R}^n\times\mathbb{R}_+. \tag{100}$$

Thus, we define $\tilde{\lambda}_N(t) \triangleq \|f\|_{\mathcal{H}} - \hat{\lambda}_N(t)$, $t \geq t_0$, which captures the error in estimating $\|f\|_{\mathcal{H}}$. The adaptive law for $\hat{\lambda}_N(\cdot)$ is given by

$$\dot{\hat{\lambda}}_N(t) = -\gamma_\lambda\left|B^{\mathrm{T}}Pe_N(t)\right|\mathcal{P}_N(x_N(t)), \quad \hat{\lambda}_N(t_0) = \hat{\lambda}_{N,0}, \quad t \geq t_0. \tag{101}$$

where $\gamma_\lambda > 0$. The next result proves the effectiveness of the proposed approach.

Theorem 10.1. *Suppose that the plant dynamics (38) with control input (98) and the adaptive laws (45), (47), and (101) generate solutions $x_N(t)$, $\hat{\alpha}_N(t)$, $\hat{f}_N(t,\cdot)$, and $\hat{\lambda}_N(t)$ defined for all $t \in [t_0,\infty)$. Finally, assume that $f \in C_R$, where C_R is given by (11), and assume that the kernel $\mathfrak{K}(\cdot,\cdot)$ that induces the native space $\mathcal{H}$ is bounded on the diagonal. Then,*

$$\lim_{t\to\infty}\|x_N(t) - x_r(t)\|_{\mathbb{X}} = 0. \tag{102}$$

Proof. Consider the Lyapunov function candidate

$$\mathcal{V}(e,\tilde{\alpha}_N,\tilde{f},\tilde{\lambda}) \triangleq V(e,\tilde{\alpha}_N,\tilde{f}) + \gamma_\lambda^{-1}\tilde{\lambda}_N^2, \quad (e,\tilde{\alpha}_N,\tilde{f},\tilde{\lambda}) \in \mathbb{R}^n\times\mathbb{R}^n\times\mathcal{H}_N\times\mathbb{R}, \tag{103}$$

where $V(e,\tilde{\alpha}_N,\tilde{f})$ is given by (42). Along the trajectories of (38) with control input (98), (45), (47), and (69), it holds that

$$\begin{aligned}\dot{\mathcal{V}}(t) \leq& -\langle Qe(t),e(t)\rangle_{\mathbb{R}^n} - 2\gamma_\lambda^{-1}\left\langle\dot{\hat{\lambda}}(t),\tilde{\lambda}(t)\right\rangle_{\mathbb{R}}\\ &+ 2\langle B^{\mathrm{T}}Pe_N(t), E_{x(t)}(I-\Pi_N)f + w_N(x_N(t),e_N(t),\hat{\lambda}_N(t))\rangle_{\mathbb{R}}\\ \leq& -\langle Qe(t),e(t)\rangle_{\mathbb{R}^n} - 2\gamma_\lambda^{-1}\left\langle\dot{\hat{\lambda}}(t),\tilde{\lambda}(t)\right\rangle_{\mathbb{R}}\\ &+ 2\langle B^{\mathrm{T}}Pe_N(t), E_{x(t)}(I-\Pi_N)f + w_N(x_N(t),e_N(t),\|f\|_{\mathcal{H}})\rangle_{\mathbb{R}^n}\\ &+ 2\langle B^{\mathrm{T}}Pe_N(t), w_N(x_N(t),e_N(t),\tilde{\lambda}_N(t))\rangle_{\mathbb{R}^n}\\ \leq& -\langle Qe(t),e(t)\rangle_{\mathbb{R}^n} - 2\left\langle\gamma_\lambda^{-1}\dot{\hat{\lambda}}(t) + |B^{\mathrm{T}}Pe_N(t)|\mathcal{P}_N(x_N(t)),\tilde{\lambda}(t)\right\rangle_{\mathbb{R}}\\ &+ 2\left|B^{\mathrm{T}}Pe_N(t)\right|\left(\underbrace{\mathrm{sign}(B^{\mathrm{T}}Pe_N(t))E_{x(t)}(I-\Pi_N)f - \mathcal{P}_N(x_N(t))\|f\|_{\mathcal{H}}}_{\leq 0}\right)\\ \leq& -\langle Qe(t),e(t)\rangle_{\mathbb{R}^n}, \quad t \geq t_0,\end{aligned}$$

where $\mathcal{V}(t)$ denotes $\mathcal{V}(e(t),\tilde{\alpha}_N(t),\tilde{f}(t),\tilde{\lambda}(t))$ for brevity. A routine application of Barbalat's lemma, as has been done in Theorem 3.1 or in Theorem 9.1, completes the proof. □

Theorem 10.1 provided an adaptive control scheme, that, despite all the schemes considered Sections 6–9, does not require the user to set *a priori* the radius of the class of admissible uncertainties. In light of (100), we observe that, if $\varepsilon = 0$, then $w_N(\cdot,\cdot,\cdot)$ is discontinuous in its second argument across the sliding surface $\mathcal{S}$ given by (87). Thus, Theorem 10.1 extends Theorem 9.1 for the case whereby $\varepsilon = 0$ to allow adaptive bounds on the radius of the class of admissible functional uncertainties. The case whereby $\varepsilon > 0$ is relatively straightforward and left to the reader. Remarkably, the discontinuous nature of the compensator (99) may induce chattering and may require actuators of infinite bandwidth to act across the sliding surface $\mathcal{S}$. To avert this problem, Theorem 10.1 can be extended by introducing a boundary layer $\mathcal{B}_\varepsilon(\mathcal{S})$ with $\varepsilon > 0$ and proceeding as in Theorem 9.1. Alternatively, a higher-order variable structure architecture can be employed. Higher-order variable structure control systems replace discontinuous terms in the control input, such as $w_N(\cdot,\cdot,\cdot)$ in (98) with terms defined as the Lebesgue integral, or multiple Lebesgue integrals, over time of discontinuous terms. In these cases, the mechanisms of the proofs are substantially identical, and chattering in the controlled plant is averted due to the smoothing effect of the integrals.

The adaptive control system outlined by Theorem 10.1 may suffer from parameter drift due to the problem of approximating infinite-dimensional uncertainties by means of finite-dimensional RKHSs, which have been discussed at length in Section 6. This problem can be overcome by constructing robust variations of Theorem 10.1 along the same lines as the results in Sections 7 or 8.

Finally, similarly to the discontinuous case of error bounding control treated in Theorem 9.1, since the tracking error converges to zero *for any* $N > 0$, the control system presented by Theorem 10.1 is AAO* since

$$J(x_N) \triangleq \limsup_{t\to\infty}\|x_N(t) - x_r(t)\|_{\mathbb{X}} = 0 \leq \|\epsilon_N\| = \|(I-\Pi_N)f\|_{\mathcal{H}}, \tag{104}$$

for each $N > 0$ and every initial condition.

11. Projection methods

In this section, we present an additional, alternative method to account for the approximation disturbance term in (38). Specifically, in the following, first, we recall the notion of continuously differentiable vector projection operators in finite-dimensional spaces. Then, we extend for the first time this result to infinite-dimensional spaces. Finally, we employ this result to address the problem of controlling the DPS, whose functional uncertainty is approximated by a finite-dimensional RKHS.

11.1. The vector projection operator

In classical, real parametric adaptive control theory, the projection operator allows modifying the adaptive laws so that the adaptive gains are guaranteed to lie within some user-defined convex constraint set at all times. The projection operator acts as follows. The user defines a second convex set that is similar to the original constraint set and contained within the constraint set. Thus, the projection operator modifies the adaptive law only if the adaptive gains lie outside the inner convex

Box 10: The convex projection operator allows to constrain the solution of an ODE within some user-defined convex constraint set by progressively modifying the component of the ODE orthogonal to the boundary of the constraint set. If the trajectory reaches the boundary of the constraint set, then the ODE has no component orthogonal to the boundary of the constraint set, and its solution is bounded to lie within the set.

set and evolve toward the boundary of the original constraint set. This modification is increasingly strong as some weighted distance between the adaptive gain and the boundary of the constraint set increases. If the adaptive gain reaches the boundary of the constraint set, then the projection operator modifies the adaptive law to allow the adaptive gain to evolve only in a direction that is locally tangential to the constraint set; for details, see Farrell and Polycarpou (2006), Ioannou and Sun (2012), Krstic, Kanellakopoulos, and Kokotovic (1995), Lavretsky and Wise (2012), to name a few classical references. This notion is explained precisely in the following definition.

Definition 11.1 (*Vector Projection Operator in* $\mathbb{R}^p$)**.** Let $\mathcal{D} \subseteq \mathbb{R}^p$ denote a convex set, and let $h : \mathcal{D} \to \mathbb{R}$ be a continuously differentiable, convex function such that $\inf_{\theta \in \mathcal{D}} h(\theta) < 0$. Further, consider the convex set

$$\overline{\Omega}_\delta \triangleq \{\theta \in \mathcal{D} : h(\theta) \le \delta\}, \quad \text{for all } \delta \in (-\infty, 1]. \tag{105}$$

The *vector projection operator* is given by

$$\operatorname{proj}(\theta, \theta_d) \triangleq \begin{cases} \theta_d - h(\theta) \dfrac{\left(\frac{\partial h(\theta)}{\partial \theta}\right)^{\mathrm{T}} \left(\frac{\partial h(\theta)}{\partial \theta}\right)}{\left(\frac{\partial h(\theta)}{\partial \theta}\right)\left(\frac{\partial h(\theta)}{\partial \theta}\right)^{\mathrm{T}}} \theta_d & \text{for} \quad (\theta, \theta_d) \in \mathcal{S}, \\ \theta_d & \text{otherwise}, \end{cases} \tag{106}$$

where $\mathcal{S} \triangleq \left\{(\theta, \theta_d) \in \mathcal{D} \times \mathbb{R}^p : \theta \in \overline{\Omega}_1 \setminus \Omega_0, \frac{\partial h(\theta)}{\partial \theta}\theta_d > 0\right\}$.

Let $\Omega_\delta \triangleq \{\theta \in \mathcal{D} : h(\theta) < \delta\}$ denote the *interior of* $\overline{\Omega}_\delta$ for some $\delta \in (-\infty, 1]$, and $\partial\Omega_\delta \triangleq \{\theta \in \mathcal{D} : h(\theta) = \delta, \delta \in (-\infty, 1]\}$ denotes the *boundary of* Ω_δ. A key property relating $h : \mathcal{D} \to \mathbb{R}$ and $\overline{\Omega}_\delta$, $\delta \in (-\infty, 1]$, is that

$$\left\langle \left.\frac{\partial h(\theta)}{\partial \theta}\right|_{\theta=\theta_b}, \theta_i - \theta_b \right\rangle_{\mathbb{R}^p} < 0, \quad \text{for all } (\theta_b, \theta_i) \in \partial\Omega_\delta \times \Omega_\delta. \tag{107}$$

Furthermore, a key property of the vector projection operator is that, given $(\theta_b, \theta_i) \in \partial\Omega_\delta \times \Omega_\delta$ for some $\delta \in (0, 1]$,

$$\left(\theta_b - \theta_i\right)^{\mathrm{T}} \left(\operatorname{proj}(\theta_b, \theta_d) - \theta_d\right) \le 0, \quad \text{for all } \theta_d \in \mathbb{R}^p. \tag{108}$$

Employing (106) and (108), it can be proven that if $\theta(t) \in \Omega_1 \setminus \Omega_0$ for some $t \in [t_0, \infty)$, and $\theta(t)$ evolves toward the boundary of $\overline{\Omega}_1$, that is, $\frac{\partial h(\theta(t))}{\partial \theta}\dot{\theta}(t) > 0$, then $\operatorname{proj}(\theta(t), \dot{\theta}(t))$ modifies $\dot{\theta}(t)$ by reducing the magnitude of the component of $\dot{\theta}(t)$ orthogonal to $\partial\Omega_1$. Ultimately, if $\theta(t) \in \partial\Omega_1$ for some $t \in [t_0, \infty)$, then $\operatorname{proj}(\theta(t), \dot{\theta}(t))$ is tangential to $\partial\Omega_1$. The term projection operator comes from the fact that $\operatorname{proj}(\theta, \theta_d)$ projects θ_d onto a hyperplane tangential to $\partial\Omega_1$ whenever $\theta \in \partial\Omega_1$.

11.2. Projection operators in native spaces

Inspired by (106) and (108), in this section, we establish a projection operator for the function space of scalar-valued uncertainties $\mathcal{H}$. Let $\mathcal{H}$ denote a scalar-valued RKHS, let $\mathcal{D} \subseteq \mathcal{H}$ denote a closed convex set, and let $h : \mathcal{D} \to \mathbb{R}$ be Fréchet differentiable, convex, and such that $\inf_{f \in \mathcal{H}} h(f) < 0$. Per definition, if $h(\cdot)$ is Fréchet differentiable over $\mathcal{D}$, then, for each $f \in \mathcal{H}$, there exists a (necessarily unique (Abraham, Marsden, & Ratiu, 2012)) operator $Dh(f) \in \mathcal{H}^*$, with $\mathcal{H}^*$ being the topological dual of $\mathcal{H}$, such that

$$h(f+g) = h(f) + \langle Dh(f), g\rangle_{\mathcal{H}^* \times \mathcal{H}} + r(f,g), \quad \text{for all } g \in \mathcal{H}, \tag{109}$$

where $\langle \cdot, \cdot \rangle_{\mathcal{H}^* \times \mathcal{H}}$ denotes the *duality pairing on* $\mathcal{H}^* \times \mathcal{H}$ and the remainder term $r(f,g)$ is of the order $O(\|g\|^2_{\mathcal{H}})$. We follow a common convention and identify the bounded linear functional $Dh(f) \in \mathcal{H}^*$ with its Riesz representer, and we use the same symbol for both so that for each $f \in \mathcal{D}$

$$\langle Dh(f), g\rangle_{\mathcal{H}^* \times \mathcal{H}} = \langle Dh(f), g\rangle_{\mathcal{H}}, \quad \text{for all } g \in \mathcal{H}. \tag{110}$$

Furthermore, per definition, since $h(\cdot)$ is convex over $\mathcal{D}$, then, for any $\alpha \in [0, 1]$, it holds that

$$h(\alpha f + (1-\alpha)g) \le \alpha h(f) + (1-\alpha)h(g), \quad \text{for all } f, g \in \mathcal{D}. \tag{111}$$

Similarly to (105), we define the closed set

$$\overline{\Omega}_\delta \triangleq \{f \in \mathcal{D} : h(f) \le \delta\}, \quad \text{for all } \delta \in (-\infty, 1]. \tag{112}$$

Since $\overline{\Omega}_\delta = \mathcal{D} \cap \{f \mid h(f) \le \delta\}$, $\mathcal{D}$ is closed by assumption, and $\{f \mid h(f) \le \delta\}$ is the inverse image under a continuous mapping of the closed set $(-\infty, \delta]$, and, hence $\{f \mid h(f) \le \delta\}$ is closed, $\overline{\Omega}_\delta$ is closed. We denote the *interior* and *boundary of* $\overline{\Omega}_\delta$, $\delta \in (-\infty, 1]$, by

$$\Omega_\delta \triangleq \{f \in D \mid h(f) < \delta\}, \tag{113}$$

$$\partial\Omega_\delta \triangleq \{f \in D \mid h(f) = \delta\}, \tag{114}$$

respectively. For each $\delta \in (\infty, 1]$, the set $\overline{\Omega}_\delta$ is convex. Thus, in the following, we prove a relationship that is analogous to (107) in Hilbert spaces.

Lemma 11.1. *Let $\mathcal{D} \subseteq \mathcal{H}$ be closed and convex, $h : \mathcal{D} \to \mathbb{R}$ be Fréchet differentiable and convex, and $\delta \in (-\infty, 1]$. Then,*

$$\langle Dh(f), f_i - f_b\rangle_{\mathcal{H}} \le 0, \quad \textit{for all } (f_b, f_i) \in \partial\Omega_\delta \times \Omega_\delta. \tag{115}$$

Proof. It follows from (111) that, for any $\beta \in [0, 1]$,

$$h(f_b + \beta(f_i - f_b)) \le h(f_b) + \beta(h(f_i) - h(f_b)),$$

and, hence,

$$\frac{h(f_b + \beta(f_i - f_b)) - h(f_b)}{\beta} \le h(f_i) - h(f_b) \le 0. \tag{116}$$

Thus, taking the limit of (116) for $\beta \to 0^+$, we deduce that the Gâteaux derivative of h at f_b in the direction given by $f_i - f_b$ (Oden, 1986), which we denote by $Dh(\cdot)$, is such that

$$\langle Dh(f_b), f_i - f_b\rangle_{\mathcal{H}^* \times \mathcal{H}} \le 0. \tag{117}$$

Fréchet differentiability implies Gâteaux differentiability, and, if this occurs, then both derivatives are equal. Thus, (115) follows from (110). □

Next, we introduce the definition of convex projection operator on Hilbert spaces and, hence, generalize both (106) and (108). To this goal, let us recall the definition of convex projection onto a closed and convex subset of a Hilbert space.

Definition 11.2 (*Convex Projection on* $\mathcal{H}$)**.** Let $\overline{C} \subset \mathcal{H}$ be closed and convex, and let $f \in \overline{C}$. The *convex projection* $P_{\overline{C}} : \mathcal{H} \to \overline{C}$ is defined as

$$P_{\overline{C}} f \triangleq \operatorname*{arg\,min}_{c \in \overline{C}} \|f - c\|_{\mathcal{H}}. \tag{118}$$

Employing standard arguments, it can be shown that, since $\overline{C}$ is closed and convex, $c^* \triangleq \min_{c \in \overline{C}} \|f - c\|_{\mathcal{H}}$ is well-defined and unique (Kreyszig, 1989). We note that, ordinarily, in our problems of RKHS embedding, the closed convex set $\overline{C}$ is bounded, too. However, although it is closed and bounded, it is generally *not* compact in $\mathcal{H}$.

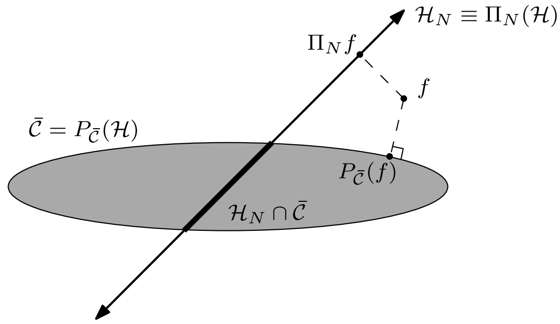

Fig. 1. The orthogonal projection $\Pi_N : \mathcal{H} \to \mathcal{H}_N$ and the convex projection $P_{\overline{C}} : \mathcal{H} \to \overline{C}$.

Only in a finite-dimensional space is a set compact if and only if it is closed and bounded. This is an important difference between the finite-dimensional setting in $\mathbb{R}^p$ and the infinite-dimensional setting in $\mathcal{H}$.

At this point, it is important to contrast the qualitative difference between orthogonal projections onto a subspace $\mathcal{H}_N \subseteq \mathcal{H}$ introduced in Definition 4.3 of the first paper of this two-paper work and convex projections onto a closed convex subset $\overline{C} \subset \mathcal{H}$ introduced in Definition 11.3. In our most common situation, wherein $\overline{C}$ is norm-bounded, $P_{\overline{C}}\mathcal{H}$ is a bounded, convex, closed subset of $\mathcal{H}$. On the other hand, the range of an orthogonal projection Π_N is a closed, finite-dimensional subspace. The range of Π_N is closed since it is finite-dimensional. However, the range of Π_N is unbounded. Indeed, if $f \in \mathcal{H}_N$, then $\alpha f \in \mathcal{H}_N$ for any $\alpha \in \mathbb{R}$. This situation is depicted graphically in Fig. 1. As shown in the figure, we have

$$\underbrace{\mathcal{H}_N \bigcap \overline{C}}_{\substack{\text{closed, convex} \\ \text{subset of vector space} \\ \text{(usually compact)}}} \triangleq \underbrace{\mathcal{H}_N}_{\substack{\text{closed vector space} \\ \text{unbounded}}} \bigcap \underbrace{\overline{C}}_{\substack{\text{closed, convex set} \\ \text{(usually bounded)}}}.$$

Thus, the following remark is essential to appreciate the intent of this section. For the statement of this remark, recall that the expression "parameter drift" means that one of the adaptive gains diverges due to the presence of unmatched disturbances or other effects not accounted for in the Lyapunov analysis that lead to the definition of the adaptive laws.

Remark 11.1. Since the range of Π_N is unbounded, the orthogonal projection Π_N by itself in an adaptive law, such as, for instance, (47), does not prevent parameter drift.

In light of this observation, we construct a projection operator for adaptive laws that embodies a notion of convex projection on $\mathcal{H}$.

Definition 11.3 (*Convex Projection Operator in $\mathcal{H}$*)**.** Let $\mathcal{D} \subseteq \mathcal{H}$ be closed and convex, and let $h : \mathcal{D} \to \mathbb{R}$ be a continuously Fréchet differentiable, convex function such that $\inf_{\theta\in\mathcal{D}} h(\theta) < 0$. Further, consider the closed, convex set (112). The *convex projection operator in $\mathcal{H}$* is given by

$$\mathrm{Proj}(f, f_d) \triangleq \begin{cases} P_h(f) f_d & (f, f_d) \in \mathcal{S}, \\ f_d & \text{otherwise}, \end{cases} \tag{119}$$

where

$$P_h(f) f_d \triangleq \left[I - h(f) \left(\cdot, \frac{Dh(f)}{\|Dh(f)\|_{\mathcal{H}}} \right)_{\mathcal{H}} \frac{Dh(f)}{\|Dh(f)\|_{\mathcal{H}}} \right] f_d, \quad (f, f_d) \in \mathcal{D} \times \mathcal{H}, \tag{120}$$

and

$$\mathcal{S} \triangleq \left\{ (f, f_d) \in \mathcal{D} \times \mathcal{H} : f \in \overline{\Omega}_1 \setminus \Omega_0, \langle Dh(f), f_d \rangle_{\mathcal{H}} > 0 \right\}. \tag{121}$$

Box 11: This paper presents the first convex projection operator over RKHSs. Similarly to the classical convex projection operator, this tool allows constraining the solution of differential equations evolving over Hilbert spaces. It is worthwhile stressing that this operator is not related to the classical projection operator, which is unable to constrain solutions of differential equations within user-defined closed and bounded sets.

Observe that the convex projection operator in $\mathcal{H}$ is nonlinear in the first argument and linear in the second argument. The next result shows how (108) can be extended to the convex projection operator in $\mathcal{H}$. In the following, for brevity, we refer to the convex projection operator in $\mathcal{H}$ as the projection operator.

Theorem 11.2. *Let $\delta \in (0, 1]$, and suppose that $f_i \in \Omega_\delta$ and $f_b \in \partial\Omega_\delta$, where Ω_δ and $\partial\Omega_\delta$ are given by (113) and (114), respectively. Then,*

$$\langle f_b - f_i, \mathrm{Proj}(f_b, f_d) - f_d \rangle_{\mathcal{H}} \leq 0 \quad \text{for all } f_d \in \mathcal{H}. \tag{122}$$

Proof. If $(f_b, f_d) \notin \mathcal{S}$, then it follows from (119) that $\mathrm{Proj}(f_b, f_d) = f_d$, and the result follows immediately. Alternatively, if $(f_b, f_d) \in \mathcal{S}$, then it follows from (119) that

$$\Big\langle f_b - f_i, \mathrm{Proj}(f_b, f_d) - f_d \Big\rangle_{\mathcal{H}} = -\underbrace{\frac{\left(f_b - f_i, Dh(f_b)\right)_{\mathcal{H}}}{\|Dh(f_b)\|}}_{>0} \underbrace{\frac{\left(Dh(f_b), f_d\right)_{\mathcal{H}}}{\|Dh(f_b)\|}}_{>0} h(f_b) \leq 0,$$

and the result is proven. □

Having outlined the general principles, in the following, we propose some specific applications of the proposed approach. The goal is to employ the convex projection operator to modify adaptive laws on $\hat{f}_N(t, \cdot)$, $t \geq t_0$, such as, for instance, (47) so that parameter drift due to the approximation error $E_{x(t)}(I - \Pi_N)f$ is averted. To this goal, it is essential to design carefully $h(\cdot)$ and, hence, both $\overline{\Omega}_0$, that is, the set wherein the projection operator does not alter the adaptive law, irrespectively of the sign of $\langle Dh(f), f_d \rangle_{\mathcal{H}}$ for any $(f, f_d) \in \overline{\Omega}_0 \times \mathcal{H}$, and $\overline{\Omega}_1$, that is, the set wherein $\hat{f}_N(t, \cdot)$ must be contained for all $t \geq t_0$. A viable option is to set $\mathcal{D} = \mathcal{H}$ and

$$h(f) = \frac{\|\Pi_N f\|_{\mathcal{H}}^2 - R^2}{\Delta}, \quad \text{for all } f \in \mathcal{D}, \tag{123}$$

where $R > 0$ captures the diameter of the uncertainty class C_R given by (11), and $\Delta > 0$ is a user-defined parameter. This way,

$$\overline{\Omega}_0 = \left\{ f \in \mathcal{H} : \|\Pi_N f\|_{\mathcal{H}} \leq R \right\}, \tag{124}$$

and, if $f \in C_R$, then $\Pi_N f \in C_R$ and

$$\overline{\Omega}_1 = \left\{ f \in \mathcal{H} : \|\Pi_N f\|_{\mathcal{H}} \leq \sqrt{R^2 + \Delta} \right\}. \tag{125}$$

An alternative option is to set $\mathcal{D} = \mathcal{H}$ and

$$h(f) = \frac{(1 + \Delta) R^{-2} \|\Pi_N f\|_{\mathcal{H}}^2 - 1}{\Delta}, \quad \text{for all } f \in \mathcal{D}, \tag{126}$$

where $R > 0$ and $\Delta > 0$ so that

$$\overline{\Omega}_0 = \left\{ f \in \mathcal{H} : \|\Pi_N f\|_{\mathcal{H}} \leq \frac{R^2}{1 + \Delta} \right\}, \tag{127}$$

$$\overline{\Omega}_1 = \left\{ f \in \mathcal{H} : \|\Pi_N f\|_{\mathcal{H}} \leq R \right\}. \tag{128}$$

The next result shows that there is a close relationship between the Fréchet derivative of h at $f \in \mathcal{H}$ and $\Pi_N f$.

Theorem 11.3. *Consider the convex function (123) with $\mathcal{D} = \mathcal{H}$. Then,*

$$Dh(f) = \Pi_N f, \quad \text{for all } f \in \mathcal{H}. \tag{129}$$

Proof. Let $g \in \mathcal{H}$ and $\Delta > 0$. The Gâteaux derivative of h evaluated at f in the direction of g, is given by

$$\langle Dh(f), g\rangle_{\mathcal{H}^*\times\mathcal{H}} = \lim_{\epsilon\to 0}\frac{h(f+\epsilon g)-h(f)}{\epsilon} \tag{130}$$

$$\begin{aligned} &= \lim_{\epsilon\to 0}\frac{\langle \Pi_N f, \Pi_N f\rangle_{\mathcal{H}} + 2\epsilon\langle \Pi_N f, \Pi_N g\rangle_{\mathcal{H}}}{\epsilon} \\ &\quad + \lim_{\epsilon\to 0}\frac{\epsilon^2\langle \Pi_N g, \Pi_N g\rangle_{\mathcal{H}} - \langle \Pi_N f, \Pi_N f\rangle_{\mathcal{H}}}{\epsilon} \\ &= \langle \Pi_N f, \Pi_N g\rangle_{\mathcal{H}} = \langle \Pi_N^2 f, g\rangle_{\mathcal{H}} = \langle \Pi_N f, g\rangle_{\mathcal{H}}, \\ &\quad \text{for all } g\in\mathcal{H}. \end{aligned} \tag{131}$$

Here, we have used the property whereby if Π_N denotes an orthogonal projection, then $\Pi_N = \Pi_N^* = \Pi_N^2$. Since Fréchet differentiability of h implies its differentiability in the sense of Gâteaux, (131) yields $Dh(f) = \Pi_N$, and the result is proven. □

Theorem 11.3 proves that, employing $h(\cdot)$ given by (123) to define the convex projection operator, the Fréchet derivative of $h(\cdot)$ at $f \in \mathcal{H}$ is equivalent to the projection of f onto $\mathcal{H}_N$. Thus, the notion of convex projection operator given by Definition 11.3 and orthogonal projection given by Definition 4.3 of the first paper of this two-part work can be tightly related. A similar result can be proven by employing (126) instead of (123).

11.3. Applying the projection operator in a native space

In this section, we discuss how the convex projection operator can be employed in the context of DPSs. For the statement of this result, consider the control input

$$u_N(t) = \hat{\alpha}_N^{\mathrm{T}}(t)x_N(t) - E_{x_N(t)}\left(\hat{f}_N(t,\cdot) + v_N(\cdot, e_N(t))\right), \quad t\geq t_0, \tag{132}$$

where the *adaptive gain* $\hat{\alpha} : [t_0,\infty) \to \mathbb{R}^n$ verifies the *adaptive law*

$$\dot{\hat{\alpha}}_N(t) = \mathrm{proj}(\hat{\alpha}_N(t), -\gamma_\alpha x_N(t)e_N^{\mathrm{T}}(t)PB), \quad \hat{\alpha}(t_0) = \hat{\alpha}_0, \tag{133}$$

the vector projection operator $\mathrm{proj}(\cdot,\cdot)$ is given by (106), $\gamma_\alpha > 0$, $P \in \mathbb{R}^{n\times n}$ denotes the symmetric, positive-definite solution of (51), $\hat{f} : [t_0,\infty) \to \mathcal{H}_N$ verifies the *adaptive law*

$$\frac{\partial \hat{f}_N(t,\cdot)}{\partial t} = -\mathrm{Proj}(\hat{f}_N(t,\cdot), \Pi_N\gamma_f E^*_{x_N(t)}B^{\mathrm{T}}Pe_N(t)), \quad \hat{f}(t_0,\cdot) = \hat{f}_0, \tag{134}$$

$\gamma_f > 0$, the projection operator $\mathrm{Proj}(\cdot,\cdot)$ is given by (119), $x_N : [t_0,\infty) \to \mathbb{X}$ denotes the solution of (38) with control input (132), $e_N(t) \triangleq x_N(t) - x_r(t)$, and the *compensator* $v_N : \mathbb{X}\times\mathbb{R}^n \to \mathbb{R}$ is user-defined.

Theorem 11.4. *Suppose that the plant dynamics (38) with control input (98) and adaptive laws (133) and (134) generate solutions $x_N(t)$, $\hat{\alpha}_N(t)$, and $\hat{f}_N(t,\cdot)$ defined for all $t \in [t_0,\infty)$. If $f \in C_R$, where C_R is given by (11), and the compensator $v_N(\cdot,\cdot)$ is such that*

$$E_{x_N(t)}v_N(\cdot, e_N(t))B^{\mathrm{T}}Pe_N(t) \leq 0, \quad \text{for all } t\geq t_0, \tag{135}$$

then $t \mapsto x_N(t)$, $t \mapsto \hat{\alpha}_N(t)$, and $t \mapsto \hat{f}_N(t,\cdot)$ are uniformly bounded on $\mathbb{X}$, $\mathbb{R}^n$, and $\mathcal{H}_N$, respectively, for all $t \in [t_0,\infty)$. Additionally, for any arbitrarily small constant $\eta > 0$, there exists $T \triangleq T(\eta) \in (t_0,\infty)$ such that

$$\left\|x_N(t) - x_r(t)\right\|_{\mathbb{X}} \leq (1+\eta)\frac{2\|B^{\mathrm{T}}P\|}{\lambda_{\min}(Q)}\bar{\epsilon}_N \sup_{\xi\in\Omega}\mathcal{P}_N(\xi), \quad \text{for all } t\geq T, \tag{136}$$

where $\bar{\epsilon}_N \triangleq \sup_{x\in\Omega}\|E_x(I-\Pi_N f)\|_{\mathcal{H}}$, and $\Omega \supseteq \bigcup_{t\geq t_0>0} x_N(t)$.

Proof. Consider the Lyapunov function candidate given by (42). Taking the derivative of (42) along the trajectories of (38) with control input (98), (133) and (134), we obtain that, for all $t \geq t_0$,

$$\begin{aligned}\dot{V}(t) = &\underbrace{-\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n}}_{\text{Term 1}} \underbrace{-2\left\langle \tilde{\alpha}_N(t), x_N(t)e_N^{\mathrm{T}}(t)PB + \gamma_\alpha^{-1}\dot{\hat{\alpha}}_N(t)\right\rangle_{\mathrm{tr}}}_{\text{Term 2}} \\ &\underbrace{+2\left\langle \tilde{f}_N(t,\cdot), \Pi_N E^*_{x_N(t)}B^{\mathrm{T}}Pe_N(t) + \gamma_f^{-1}\dot{\hat{f}}_N(t,\cdot)\right\rangle_{\mathcal{H}}}_{\text{Term 3}} \\ &+\underbrace{2\left\langle E_{x_N(t)}(I-\Pi_N)f, B^{\mathrm{T}}Pe_N(t)\right\rangle_{\mathbb{R}}}_{\text{Term 4}} + \underbrace{2\left\langle E_{x_N(t)}v_N(\cdot, e_N(t)), B^{\mathrm{T}}Pe_N(t)\right\rangle_{\mathbb{R}}}_{\text{Term 5}},\end{aligned} \tag{137}$$

with $V(t) = V(e_N(t), \tilde{\alpha}_N(t)\tilde{f}_N(t,\cdot))$ for brevity.

Applying Lemma 11.3 of Lavretsky and Wise (2012), we deduce that Term 2 is nonpositive for all $t \geq t_0$. Furthermore, for all $t \geq t_0$, it holds that

$$\text{Term 3} = 2\left\langle \Pi_N f - \hat{f}_N(t,\cdot), \Pi_N E^*_{x_N(t)}B^{\mathrm{T}}Pe_N(t) - \mathrm{Proj}(\hat{f}_N(t), \Pi_N\gamma_f E^*_{x_N(t)}B^{\mathrm{T}}Pe(t))\right\rangle_{\mathcal{H}_N}.$$

Now, let $\delta \in (-\infty, 1]$ and $\Omega_\delta = \{f \in \mathcal{D} : \|\boldsymbol{\Pi}_N f\|_{\mathcal{H}} < \delta\}$. If $f \in \Omega_\delta$, then $\Pi_N f \in \Omega_\delta$ from the definition of Ω_δ. Thus, it follows from Theorem 11.2 that Term 3 in (137) is nonpositive.

Finally, by assumption, Term 5 is nonpositive. Thus,

$$\begin{aligned}\text{Terms 1 \& 4} &= -\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n} + 2\left\langle \tilde{f}_N(t,\cdot), (I-\Pi_N)E^*_{x_N(t)}B^{\mathrm{T}}Pe_N(t)\right\rangle_{\mathcal{H}} \\ &= -\langle Qe_N(t), e_N(t)\rangle_{\mathbb{R}^n} + 2\left\langle E_{x_N(t)}(I-\Pi_N)f, B^{\mathrm{T}}Pe_N(t)\right\rangle_{\mathcal{H}} \\ &\leq -\lambda_{\min}(Q)\|e_N(t)\| \\ &\quad \cdot\left(\|e_N(t)\| - \frac{2\|B^{\mathrm{T}}P\|}{\lambda_{\min}(Q)}\|\epsilon_N\|_{\mathcal{H}}\sup_{x\in\Omega}\mathcal{P}_N(x)\right),\end{aligned} \tag{138}$$

for all $t \geq t_0$, where $\Omega \subset \mathbb{X}$ is a compact set such that $x(t) \in \Omega$ at all times. Thus, the result is attained. □

Theorem 11.4 requires that (135) is verified at all times. A compensator that verifies such a requirement is given, for instance, by $v_N(x_N(t), e_N(t)) = -\mathrm{sign}\left(B^{\mathrm{T}}Pe_N(t)\right)$, $t \geq t_0$. This solution may lead to chattering, and, to overcome this limitation, classical higher-order variable structure control systems can be designed; this extension is left for future work directions. This result also requires to construct $\Omega \supseteq \bigcup_{t\geq t_0>0} x_N(t)$ to estimate the ultimate bound given by (136). For a discussion on this point, the reader is referred to the proof of Theorem 7.1.

12. Ultimate bounds and fill distances

Sections 7–10 discussed the ultimate performance for the cost functional

$$J(x_N) \triangleq \limsup_{t\to\infty}\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq \bar{e}_{N,\infty},$$

where $\bar{e}_{N,\infty}$ is an ultimate bound on the tracking error for a specific approach. The analysis of the ultimate bounds $\bar{e}_{N,\infty}$ yields inequalities such as $\bar{e}_{N,\infty} \leq \mathcal{F}(\epsilon)$ or $\bar{e}_{N,\infty} \leq \mathcal{F}(N)$, for all $f \in C_{R,\epsilon,N}$, as the case may be.

The functions $\mathcal{F}(\cdot)$ for all of the cases of performance bounds are expressed in terms of the power function $\mathcal{P}_N(\cdot)$ of $\mathcal{H}_N \subseteq \mathcal{H}$ defined by (27). Since $\mathcal{P}_N(\cdot)$ is known in closed form, these bounds can be very powerful. They can be used offline to assess the performance of algorithms, or they can be used online as parts of more advanced adaptive control approaches. Carefully observe that the bounds on the tracking error that use the power function $\mathcal{P}_N(\cdot)$ make no restrictions on the regularity of the set Ω. In this section, we review a commonly employed technique, as summarized in Schaback (1994), Wendland (2004), which can be used to generate intuitive, alternative forms of such error bounds. These alternative forms are attractive owing to their inherent geometric structure.

The approach that follows can be applied to any bound that contains the power function $\mathcal{P}_N(\cdot)$ of the subspace $\mathcal{H}_N \in \mathcal{H}$. Here, we consider two general classes of error bounds that arise in the theorems above. In the first class, we study expressions in the form

$$\bar{e}_{N,\infty} \leq O\left(\epsilon\sup_{\xi\in\Omega}\mathcal{P}_N(\xi)\right), \tag{139}$$

that arise when the uncertainty f is contained in $C_{R,\epsilon,N}$. Note that an expression of this form arises in the study of variants of the deadzone,

Table 1
Radial functions η used to define kernels $\mathfrak{K}(x,y) \triangleq \eta(\|x-y\|_2)$, for $x, y \in \mathbb{X}$. The symbol $K_{k-d/2}$ in the definition of Sobolev–Matern kernels denotes the modified Bessel function of the second kind, and $\lceil x \rceil$ denotes the *ceiling function*. It holds that $\sup_{\xi \in \Omega_\epsilon} \mathcal{P}_N(\xi) \leq \sqrt{\mathcal{N}(h_{\Xi_N,\Omega})}$ for centers $\Xi_N \subset \Omega$. In this table, $h \triangleq h_{\Xi_N,\Omega}$ defined in (141) and $c > 0$ is a generic constant that differs for each kernel. This table has been extrapolated from Schaback (1994), Wendland (2004).

Kernel name	$\mathfrak{K}(x,0) = \mathfrak{K}_0(r)$	Parameters	$\mathcal{N}(h)$
Gaussian	$e^{-\alpha r^2}$	$\alpha > 0$	$e^{-c\lvert\log(h)\rvert/h}$
Inverse multiquadric	$(\alpha^2+\beta^2)^\beta$	$\beta < 0$	$e^{-c/h}$
Wendland	$\phi_{d,k}(r)$	$d, k \in \mathbb{N}$	h^{2k+1}
Sobolev–Matern	$\frac{2\pi^{d/2}}{\Gamma(k)} K_{k-d/2}(r/2)^{k-d/2}$	$d, k \in \mathbb{N}$	h^{2k-d}

Box 12: The fill distance is a measure of how scattered bases are distributed over some domain. In the first part of this paper, we described the performance of the proposed controllers as functions of the number of elements in the center set in terms of the power function. For some well-known kernels, the controllers' performances can be captured using fill distances as well.

σ-modification, continuous error bounding, continuous adaptive error bounding, and projection methods for the functional uncertainty class $C_{R,\epsilon,N}$. We also consider the larger class of functional uncertainties $f \in C_R$, and for this class, we consider expressions in the form

$$\bar{e}_{N,\infty} \leq O\left(\sup_{\xi \in \Omega} \mathcal{P}_N(\xi)\right). \tag{140}$$

The alternative bounds are obtained by employing Table 1.

Consider the bounded set $\Omega \subset \mathbb{X}$, the set of centers given by (22), a reproducing kernel $\mathfrak{K} : \mathbb{X}\times\mathbb{X} \to \mathbb{R}$ in the Table 1 and the corresponding function $\mathcal{N} : \mathbb{R}^+ \to \mathbb{R}$, and the *fill distance* $h_{\Xi_N,\Omega}$ *of the centers* Ξ_N *in* Ω defined as

$$h_{\Xi_N,\Omega} \triangleq \sup_{\xi \in \Omega} \min_{\xi_i \in \Xi_N} \|\xi - \xi_i\|_{\mathbb{X}}. \tag{141}$$

It follows from (82) that the norm on the projection error can be estimated by means of the power function $\mathcal{P}_N(\cdot)$, and

$$\mathcal{P}_N(x) \leq O\left(\sqrt{\mathcal{N}(h_{\Xi_N,\Omega})}\right). \tag{142}$$

Note that the expressions in (82) hold for any set Ω without regard to its regularity, but the bound in (142) requires some regularity of the set Ω, such as the verification of the interior cone condition, to apply the upper bound $\mathcal{N}(\cdot)$ for the power function (Wendland, 2004).

It is immediate, then, that we obtain for the inequality in (139) the alternate expression

$$\bar{e}_{N,\infty} \leq O\left(\epsilon\sqrt{\mathcal{N}(h_{\Xi_N,\Omega})}\right). \tag{143}$$

for all $f \in C_{R,\epsilon,N}$, and from the inequality in (140) we get

$$\bar{e}_{N,\infty} \leq O\left(\sqrt{\mathcal{N}(h_{\Xi_N,\Omega})}\right) \tag{144}$$

for all $f \in C_R$. We apply this technique specifically to Theorem 9.1, which captures the error bounding method, and Theorem 7.1, which captures the σ-modification, in the corollaries below.

Corollary 12.1. *Let the hypotheses of Theorem 9.1 hold and let $\mathcal{N}(\cdot)$ be given by Table 1 for the user-defined choice of kernel function. For the choice $\epsilon > 0$ in the compensator $v_N(\cdot,\cdot)$ given by (86), there exists a constant $T \in (t_0, \infty)$ such that the closed-loop trajectory tracking error satisfies*

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq O\left(\sqrt{\mathcal{N}(h_{\Xi_N,\Omega_\epsilon})}\right) \tag{145}$$

for all $t \geq T$. In particular, for the Gaussian, inverse multiquadric, and Wendland kernels of smoothness index $r > 0$, it holds that

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq \begin{cases} O\left(e^{-\left(\mu \frac{\log(h_{\Xi_N,\Omega_\epsilon})}{h_{\Xi_N,\Omega_\epsilon}}\right)}\right) & \textit{Gaussian}, \\ O\left(e^{-\left(\frac{\tilde{\mu}}{h_{\Xi_N,\Omega_\epsilon}}\right)}\right) & \textit{inverse multiquadric}, \\ O\left(h_{\Xi_N,\Omega_\epsilon}^{r+1/2}\right) & \textit{Wendland, smoothness } r, \end{cases} \tag{146}$$

where $\mu, \tilde{\mu} > 0$ are constant.

Proof. Consider the kernel functions listed in Table 1. For the Gaussian, inverse multi-quadratic, and Wendland kernel functions, Table 1 provides a function $\mathcal{N}(\cdot)$ such that

$$\sup_{\xi \in \Omega_\epsilon} \mathcal{P}_N(\xi) \leq O\left(\sqrt{\mathcal{N}(h_{\Xi_N,\Omega_\epsilon})}\right).$$

It follows from Theorem 9.1 that there is a constant $C > 0$ such that

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq C \sup_{\xi \in \Omega_\epsilon} \mathcal{P}_N(\xi) \leq \tilde{C}\sqrt{\mathcal{N}(h_{\Xi_N,\Omega_\epsilon})}$$

for all $t \geq T$ and some constant $T \in (t_0, \infty)$. Thus, the result follows directly from (82). □

The following corollary for the σ-modification in Theorem 7.1 is analogous.

Corollary 12.2. *Let the hypotheses of Theorem 7.1 hold and let $\mathcal{N}(\cdot)$ be given by Table 1 for the user-defined choice of kernel function. There exists a constant $T \in (t_0, \infty)$ such that the closed-loop trajectory tracking error satisfies*

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq O\left(\frac{1}{\sigma}\left(\sqrt{\mathcal{N}(h_{\Xi_N,\Omega_\epsilon})} + \sigma\right)\right) \tag{147}$$

for all $t \geq T$. In particular, for the Gaussian, inverse multiquadric, and Wendland kernels of smoothness index $r > 0$, it holds that

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq \begin{cases} O\left(\frac{1}{\sigma}\left(e^{-\tilde{\mu}\lvert\log(h_{\Xi_N,\Omega})\rvert/h_{\Xi_N,\Omega}}\right) + \sigma\right), & \textit{Gaussian}, \\ O\left(\frac{1}{\sigma}\left(e^{-\mu/h_{\Xi_N,\Omega}} + \sigma\right)\right), & \textit{inv. multiquadric}, \\ O\left(\frac{1}{\sigma}\left(h_{\Xi_N,\Omega}^{r+1/2} + \sigma\right)\right), & \textit{Wendland, smoothness } r, \end{cases} \tag{148}$$

where $\mu, \tilde{\mu} > 0$ are constant.

Proof. The proof follows by proceeding as in the proof of Corollary 12.1. □

13. Concluding remarks on theoretical results

13.1. A comparison of theoretical results

Some of the key theoretical results collected in this paper are summarized in Table 2. This table provides a concise inventory of the performance of some of the nonparametric adaptive control methods for a collection of progressively larger functional uncertainty classes given by (12) in the hypothesis space $\mathcal{H}$ that contains the functional matched uncertainty. In particular, Table 2 summarizes the controller performance, as measured by an ultimate bound $\bar{e}_{N,\infty}$ on the tracking error for multiple control schemes.

The first row of Table 2 shows the performance of the classical parametric MRAC law, which only applies to the uncertainty class $C_{\Phi_N,R}$, and assures asymptotic convergence of the trajectory tracking error to zero. The parametric MRAC law does not apply to the broader

Table 2
Summary of the performance of the nonparametric adaptive control methods for the functional uncertainty classes $\mathcal{C}_{\Phi_N,R} \subset \mathcal{C}_{N,\epsilon,R} \subset \mathcal{C}_R \subset \mathcal{H}$. Control techniques are sorted by performance level. The first row shows the performance of the classical parametric gradient adaptive law, which only applies to the uncertainty class $\mathcal{C}_{\Phi_N,R}$, and assures asymptotic convergence of the trajectory tracking error to zero. The parametric gradient adaptive law does not apply to the broader uncertainty classes $\mathcal{C}_{N,\epsilon,R}$ and $\mathcal{C}_R$. The second row shows the performance of the limiting DPS system. This ideal controller assures asymptotic convergence of the trajectory tracking error $e(\cdot)$ to zero for any uncertainty class considered. The remainder of this table shows the uniform ultimate bounds on the trajectory tracking error $e_N(\cdot)$ obtained by approximating functional uncertainties on $\mathcal{H}_N$, that is, $\bar{e}_{N,\infty}$, which we distinguish from $\bar{e}_\infty$ attained by the ideal, limiting DPS. In this table, $C_0(R) \triangleq C_1 + C_2 R^2$, the function $\mathcal{F}(\epsilon) \triangleq \sqrt{C\epsilon^2 \|\mathcal{P}_N\|_{L^\infty} + C_0(R)}$ captures the suboptimality of the σ-modification in the class $\mathcal{C}_{N,\epsilon,R}$, and $\mathcal{F}_1(N) \triangleq \sqrt{C\|\mathcal{P}_N\|_{L^\infty} R^2 + C_0(R)}$ bounds the suboptimality in $\mathcal{C}_R$. The term $\mathcal{P}_N$ denotes the power function of the subspace of approximations $\mathcal{H}_N$ in the hypothesis space $\mathcal{H}$.

Method	Type	Uncertainty Class		
		$\mathcal{C}_{\Phi_N,R}$	$\mathcal{C}_{N,\epsilon,R}$	$\mathcal{C}_R$
Parametricgradient	Classical	0	–	–
Nonparametricgradient	LimitingDPS	$\bar{e}_\infty = 0$	$\bar{e}_\infty = 0$	$\bar{e}_\infty = 0$
Boundingmethod	Ideal Approx.	0 0	0 $\epsilon\|\mathcal{P}_N\|_{L^\infty}$	0 $\|\mathcal{P}_N\|_{L^\infty} R$
Adaptiveerror bounding	Ideal Approx.	0 0	0 $\epsilon\|\mathcal{P}_N\|_{L^\infty}$	0 $\|\mathcal{P}_N\|_{L^\infty} R$
Deadzone		–	$\epsilon\|\mathcal{P}_N\|_{L^\infty}$	$\|\mathcal{P}_N\|_{L^\infty} R$
Projection		–	$\epsilon\|\mathcal{P}_N\|_{L^\infty}$	$\|\mathcal{P}_N\|_{L^\infty} R$
σ-modification		$\sqrt{C_0(R)}$	$\leq \mathcal{F}(\epsilon)$	$\mathcal{F}_1(N)$

uncertainty classes $\mathcal{C}_{N,\epsilon,R}$ and $\mathcal{C}_R$. The second row of Table 2 shows the performance of the limiting DPS system. This ideal controller assures asymptotic convergence of the trajectory tracking error $e(\cdot)$ to zero for any uncertainty class considered. The remainder of Table 2 shows the uniform ultimate bounds on the trajectory tracking error $e_N(\cdot)$ obtained by approximating functional uncertainties on $\mathcal{H}_N$, that is, $\bar{e}_{N,\infty}$, which we distinguish from $\bar{e}_\infty$ attained by the ideal, limiting DPS. The power function $\mathcal{P}_N$ defined in (27) plays a pivotal role in predicting the controller's performance *a priori* without relying on conservative assumptions on the functional uncertainty, but on clearly quantifiable characteristics of the kernel function defining the chosen RKHS $\mathcal{H}$ and its approximation $\mathcal{H}_N$. Remarkably, these performance bounds are also explicit functions of the number of centers N, whereas in classical adaptive control systems, such a dependence cannot assessed as an explicit function of the dimension of the regressor vector.

Using standard kernels, Table 2 summarizes the performance for a particularly large inventory of function spaces (Wendland, 2004). Indeed, this table does not specify $\mathfrak{K}(\cdot,\cdot)$. For example, using either the Wendland or Sobolev–Matern kernels of appropriate order, Table 2 holds when the uncertainty classes are contained in any of the Sobolev spaces $H^r(\mathbb{R}^n)$ of smoothness index $r \in \mathbb{N}$, for r large enough such that the Sobolev space is a native space. The Sobolev space of integer index $r \in \mathbb{N}$ consists of all functions in $L^2(\mathbb{R}^n)$ that have weak derivatives through order r contained in $L^2(\mathbb{R}^n)$, and it is a widely employed collection of spaces used to measure the regularity of functions. Alternatively, if we use the Gaussian kernel, Table 2 applies to all functions in the uncertainty class contained in the associated native space, which is a space of holomorphic functions (Steinwart, Hush, & Scovel, 2006). See Wendland (2004) for an inventory of other popular reproducing kernels for which the theory in this paper holds. The variety of native spaces enables the design of control methods based on knowledge about the smoothness of the uncertainty. The Sobolev scale of smoothness spaces $H^r(\mathbb{R}^n)$ is a well-known standard for categorizing the regularity of functions (Adams & Fournier, 2003), and it plays a critical role in studying approximations of PDEs. Here, we can infer a similar role for nonparametric adaptive feedback control.

Table 2 also clarifies that *the regularity of the functional uncertainty can have a predictable effect on the bounds on the performance of a nonparametric feedback control strategy*. As shown empirically in Section 14 below, smoother kernels generally yield approximations that converge faster whenever the uncertainty is correspondingly smooth. Table 2 helps appreciate that the general nonparametric control theory presented in this paper allows for improved tracking performance guarantees that exploit this general observation.

It is also worth noting that using the techniques described in Section 12, the bounds in Table 2 can be replaced in terms of fill distances (141) of centers of approximation given by (22) in subsets $\Omega \subset \mathbb{X}$ that support the closed loop trajectory. For families of well-known kernels, such as the Sobolev–Matern, inverse multiquadric, exponential, or Wendland kernels, we obtain controller performance bounds captured by (143) and (144). These geometric descriptions help make the definition of approximations in these nonparametric methods more intuitive. This geometric understanding of the role of the approximation centers can lead to the development of new approaches in data-driven adaptive control methods.

Viewed solely from an academic standpoint, results such as those in Table 2 are interesting and compelling as they summarize a reasonable next step that begins with the host of documented parametric MRAC results in approximation-based adaptive control for deterministic, generally nonlinear ODEs. Table 2 should be compared to Table 11.1 in Lavretsky and Wise (2012), Table 4.2 in Farrell and Polycarpou (2006), or Tables 8.1 through 8.5 in Ioannou and Sun (2012). Table 2 makes additional assumptions in comparison to the tables in these standard references. However, the tables in these classical references rely crucially on the fact that the uncertainty satisfies the uniform approximation assumption in $L^\infty(\Omega)$. For all the kernels discussed in this paper, we have $\mathcal{H} \subset L^\infty(\Omega)$, so Table 2 holds in a more restrictive setting that enables sharper error estimates. Table 2 still holds for a vast collection of types of functional uncertainty, and the conclusions are considerably stronger.

It is worthwhile recalling that Table 2 was developed assuming that the functional uncertainty lies in the selected uncertainty class, and that the controlled plant state $x_N(\cdot)$ remains in a set Ω. In other words, Table 2 assumed that the closed-loop plant trajectory remains in a compact set Ω that is known *a priori*. This is a strong assumption, but it is also common in the initial study of the performance of approximation-based adaptive control methods; see Farrell and Polycarpou (2006, Ch. 4, 7). There can be substantial, and nontrivial, differences among the hypotheses or operating assumptions of the methods in Table 2. This should not be overlooked in assessing the table. For instance, we expect that a significant improvement in the "uniformity of the hypotheses" would be a major improvement to the table. Here, as a significant example, we are thinking about the generalizations to classes of methods that guarantee positive invariance of the set Ω, like methods based on barrier functions. See Anderson, Marshall, and L'Afflitto (2020, 2021b) for recent work by the authors.

Also, the table suffers in that it overlooks some common technical issues related to discontinuous learning or feedback laws. Since the deadzone modification and variable structure methods include types of discontinuous feedback controls, then the Lyapunov analysis should be carried out in the context of Filippov solutions or differential inclusions. As it is standard in the classical texts on parametric adaptive control theory, we just note these potential issues and assume that high-frequency oscillations or chattering are not of substantial concern for the problems at hand.

Table 3 summarizes additional key results of this paper. In particular, this table summarizes the performance of finite-dimensional control systems presented in this paper in terms of the proposed notion of approximation theory suboptimality.

13.2. The role of the Grammian matrix

As discussed in Remark 7.1, the adaptive gains associated with the functional uncertainty can be found as solutions of PDEs, such as (47),

Table 3
Summary of the performance of native space embedding methods in terms of approximation theory suboptimality. Results on $C_{N,\epsilon,R}$ hold for any $N \in \mathbb{N}$, including $N \to \infty$. In classical parametric MRAC, N is fixed *a priori*.

Method	Class	Performance
Real Parametric Gradient	$C_{\Phi_N,R}$	AAO, $\bar{e}_{N,\infty} = 0$
Bounding Control	$C_{N,\epsilon,R}$	AAO*, $\bar{e}_{N,\infty} = 0$
Adaptive Bounding Control	$C_{N,\epsilon,R}$	AAO*, $\bar{e}_{N,\infty} = 0$
Deadzone	$C_{N,\epsilon,R}$	AAO, $\bar{e}_{N,\infty} \le O(h^r_{\Xi_N,\Omega})$
σ-modification	$C_{N,\epsilon,R}$	$\mathcal{F}(\epsilon)$-suboptimal

(50), (69), (89), and (134). However, by proceeding as in Section 7.3, these adaptive gains can be computed as solutions of matrix ODEs, such as (66). In such equations, the Grammian matrix plays a key role, which does not appear in the matrix ODEs underlying classical adaptive control systems. This difference with matrix adaptive laws of classical adaptive control systems is critical to define consistent approximations in the native space. It is the Grammian matrix that relates how implementations for each N are all related to the same space $\mathcal{H}$. These matrix ODEs and the associated Grammian matrix are needed to conclude how performance can be related to approximation error in the native space and guaranteed over the functional uncertainty class *for all* N. In classical adaptive control systems, such as MRAC, the controller's performance cannot be explicitly related to the dimension of the regressor vector.

The appearance of the Grammian matrix has other important practical implications. As we discuss more carefully in the numerical examples in the next section, the appearance of the Grammian makes clear that the bounds derived in the RKHS framework are achievable only under certain conditions that depend on the combination of the approximation error for a dimension N, the condition number of the Grammian, and size of the injected disturbance in the original equations.

14. Numerical example

In this section, we present the results of a numerical example that illustrates both the applicability of the proposed results and the qualitative character of some of the performance bounds described in this paper. For brevity, we focus on the results obtained by applying the error bounding method discussed in Section 9. Since the analysis of multiple control techniques over a single example may erroneously lead to the conclusion that a technique is universally better than another, and in consideration of the length of such a discussion, we encourage the reader to employ the suite of computer codes produced as part of this work (Wang, L'Afflitto, & Kurdila, 2024) and verify the quality of results for multiple techniques and multiple examples. The considerations made in the following regarding the quality of the error bounding control method will serve as guidelines.

Consider the plant model given by (1) with $\mathbb{X} = \mathbb{R}^2$, $\mathbb{U} = \mathbb{R}$,

$$A = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\zeta\omega \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \omega^2 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{149}$$

$\omega = 1$ rad/s, and $\zeta = 0.2$. The functional uncertainty is given by

$$f(x) = \tanh(x_1^3 + 0.001x_2^5), \qquad x \in \mathbb{R}^2. \tag{150}$$

The reference model is given by (2) with $r(t) = \cos(5t)$, $t \ge 0$,

$$A_{\rm ref} = \begin{bmatrix} 0 & 1 \\ -\omega_{\rm ref}^2 & -2\zeta_{\rm ref}\omega_{\rm ref} \end{bmatrix}, \quad B_{\rm ref} = \begin{bmatrix} 0, \\ \omega_{\rm ref}^2 \end{bmatrix}. \tag{151}$$

$\omega_{\rm ref} = 20$ rad/s, and $\zeta_{\rm ref} = \frac{\sqrt{2}}{2}$. The kernel function employed for all simulations is the 3/2 Sobolev–Matern kernel (Schaback, 1994) given by

$$\mathfrak{K}_{3,2}(x,y) = \left(1 + \frac{\sqrt{3}\,\|x-y\|_{\mathbb{X}}}{l}\right) e^{-\frac{\sqrt{3}\|x-y\|_{\mathbb{X}}}{l}}, \tag{152}$$

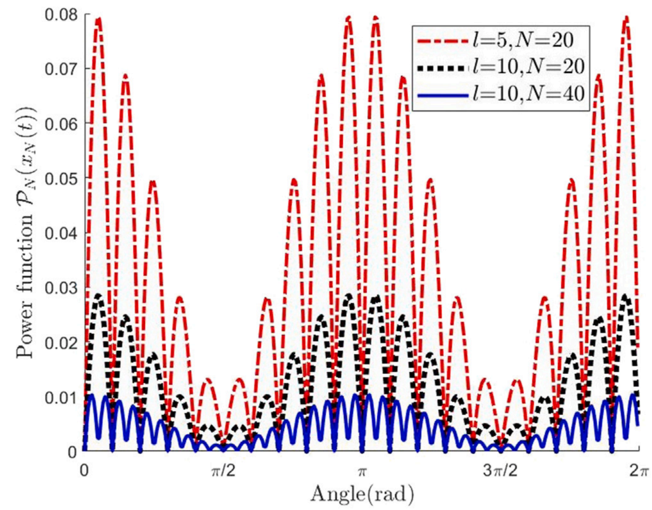

Fig. 2. The power function "unwrapped" and plotted over the circle that supports the ultimate dynamics.

with $l > 0$. The larger the hyperparameter l, the larger the "variance" of the radial basis function defined by the kernel. Finally, we set

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \epsilon = 10^{-4}. \tag{153}$$

Applying the parameterization (62), the adaptive laws for the adaptive law on the functional uncertainty (47) takes the form

$$\dot{\hat{\Theta}}_N(t) = \gamma_f \mathbb{K}_N^{-1} \mathfrak{K}_{\Xi_N}(x_N(t)) B^{\rm T} P e_N(t), \quad \hat{\Theta}_N(0) = \hat{\Theta}_{N,0}, \quad t \ge 0. \tag{154}$$

Similar expressions can be deduced for the adaptive laws in (50), (69), (89), and (134).

The basis centers Ξ_N are distributed evenly around the circular path eventually approached by the reference signal $x_r(\cdot)$. It appears from this distribution of centers that the RKHS embedding method is well-suited to choosing bases that are scattered over the domain and need not be defined on some regular grid in $\mathbb{R}^n$. The choice of centers can be understood as the locations in the domain where approximations can interpolate the uncertainty.

Applying the error bounding method, the power function $\mathcal{P}_N(\cdot)$ is plotted in Fig. 2 over the circular path that supports the asymptotic dynamics for multiple hyperparameters and numbers of basis functions. The power function is zero at the kernel centers and grows between the kernel centers. The magnitude of the growth is determined by the smoothness of the kernel function and the number of basis functions. For $l = 5$, the kernel function exhibits much greater growth away from the centers compared to the kernel with $l = 10$, for a fixed number of basis functions. For the same hyperparameter, doubling the number of basis functions effectively suppresses the growth between the basis centers, which leads to better approximations. The plots in Fig. 2 can be constructed once the centers are selected and do not reflect the performance of the control strategy. This plot suggests the usefulness of the power function as either a tool to assess the pointwise error in an *a priori* manner for controller design or in an *a posteriori* fashion for adaptive basis selection. Fig. 2 merely describes the offline error that can result from the choice of some set of centers. Similar results can be attained by applying any other adaptive control methods presented herein.

Applying the error bounding method, Fig. 3 shows the power function at $x_N(t)$ with $t \approx 19.5$ s versus the number of basis centers. The power function at the time $t \approx 19.5$ s is at a local maximum value and the transient behavior has mostly died out. The power function can be used to build an upper bound for the pointwise approximation error $(I - \Pi_N)f$ at the state $x_N(t)$. From Fig. 3, we deduce that the

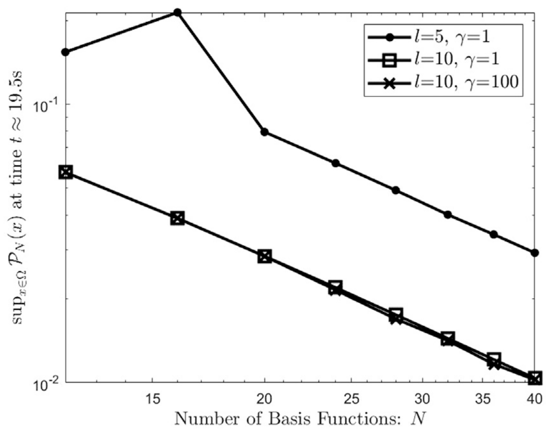

Fig. 3. Maximum steady state value of the power function $\mathcal{P}_N(x)$ at approximately 19.5 s.

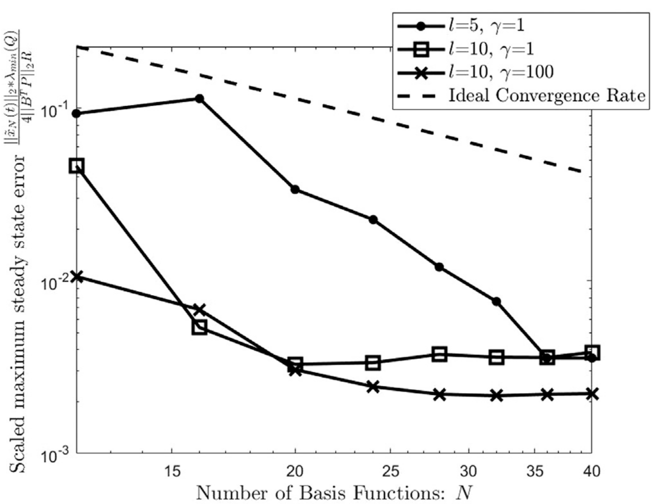

Fig. 4. Scaled norm of the maximum steady-state error $\frac{4\|B^T P\|}{\lambda_{\min}(Q)} R$ employing the error bounding method.

value of the power function depends on the smoothness of the unknown function and the kernel function that defines $\mathcal{H}$. The power function value produced by the kernel with $l = 5$ is larger than that produced by the flatter kernel with $l = 10$. In some sense, Fig. 3 illustrates an ideal behavior that we expect the closed-loop response to follow as N increases and approaches ∞.

Fig. 4 captures the ultimate bound in Theorem 9.1. There are a few important observations regarding this plot. First, the dashed line in Fig. 4 captures the slope of the lines depicted in Fig. 3. 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 reaches a plateau. This is a generic feature of the realizable controllers and is due to other unmodeled noise in the simulation such as numerical integration error and numerical conditioning of the factorization of the Grammian matrix, which is discussed further below.

Fig. 4 can also be used to introduce a new notion that helps distinguish the performance of these methods for different kernels. As already discussed, the combination of any injected external disturbance and numerical conditioning error, as well as integration error in computational studies, introduces a lower plateau. This tracking error cannot be reduced below the plateau irrespective of the number of basis functions

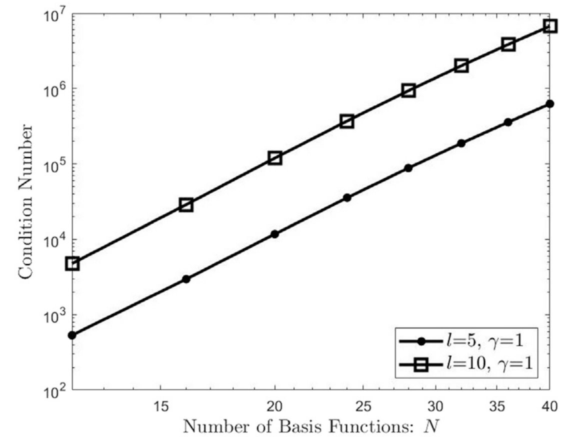

Fig. 5. Condition number of the Grammian matrix $\mathbb{K}(\cdot,\cdot)$ given by (29).

used. For each type of kernel, it is possible to define an *effective dimension* for the control problem that is equal to the dimension N_∞ at which the tracking controller reaches the noise floor. For the kernels shown in Fig. 4, $N_\infty = 35$ for $l = 5$ and $\gamma_f = 1$; $N_\infty = 20$ for $l = 10$ and $\gamma_f = 1$; and $N_\infty = 20$ for $l = 10$ and $\gamma_f = 100$. As expected, the learning rate γ_f does not influence the effective dimension, but the hyperparameter does. There is a large inventory of methods for the selection of optimal hyperparameters in offline statistical and machine-learning problems in a stochastic setting. Fig. 4 shows that understanding the relationship between the hyperparameter and the effective dimension in problems of nonparametric control theory is useful, and suggests an important topic of future research.

While increasing γ_f should lead to faster convergence in time of online approximation $\hat{f}_N$, larger γ_f can result in undesirable high-frequency oscillations that may take an intractable amount of time to decay and contribute to large tracking error.

As noted in Müller and Schaback (2009), Pazouki and Schaback (2011), Powell, Liu, and Kurdila (2022), Wendland (2004), building estimates from an RKHS can result in poor conditioning. Fig. 5 shows that the condition number of the Grammian matrix grows as the number of basis functions increases. In addition, the kernel with $l = 10$ has worse conditioning than the kernel with $l = 5$. A poorly conditioned Grammian matrix can inject large numerical noise in native RKHS embedding and eventually dominate the approximation error bounds. As already noted, ill-conditioning of the Grammian matrix results in the tracking error hitting a plateau. The scaled steady-state maximum error produced by the kernel with $l = 10$ improves as expected up to around $N = 20$ for the learning rate of $\gamma_f = 1$. It should be noted that the relative numerical integration error in this example is $O(10^{-3})$. We would, therefore, expect an estimate using an inverse of the Grammian matrix with a condition number greater than the order $O(10^3)$, as seen in Fig. 5, for $N > 15$ to be potentially sensitive to numerical integration error, which can also contribute to a plateau or even increase in the tracking error as seen in Fig. 4. Future studies will aim to address these large condition numbers.

Fig. 6 shows the evolution of the power function and the scaled norm of the tracking error. In these simulations, the power function rarely reaches the scaled norm of tracking error. However, the power function still bounds the norm of the tracking error from the above, except at a few points where numerical round-off errors appear.

15. Conclusion

This work, divided into two papers, presented a novel control framework that we called nonparametric to distinguish it from existing ones, which we refer to as parametric. The key idea underlying the

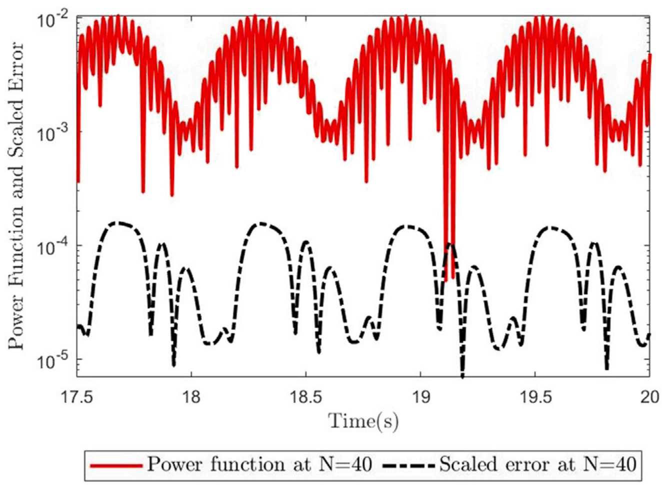

Fig. 6. This figure shows the error bound as a function of time and the steady state error throughout the simulation. Note that the pointwise error bound is minimal near the kernel centers. Consequently, we observed smaller tracking error when the state values approached the kernel centers.

proposed framework lies in assuming that the matched functional uncertainties affecting the plant dynamics, which are captured by a set of nonlinear ODEs, are elements of a native space. Thus, in the first paper, we developed an adaptive law able to steer the plant's trajectory toward the trajectory of a reference model. This result employed a DPS and, hence, could not be applied to problems of practical interest. In this paper, we presented several approaches to approximate this DPS in finite dimensions and, hence, assure both satisfactory trajectory tracking and uniform boundedness of the trajectory tracking error as well as of the adaptive gains.

The key elements of the novelty of the proposed nonparametric adaptive control framework are multiple. Firstly, despite existing parametric methods, the user is not required to provide a parameterization of the functional uncertainties using a regressor vector or to create some mechanism producing such a parameterization. Secondly, the proposed framework allows capturing classes of functional uncertainties that are considerably more vast than those addressed by parametric methods. Thirdly, and more importantly, the proposed framework allows us to explicitly correlate *a priori* the ultimate bounds on the trajectory tracking error with the dimension of the approximating native space. In classical parametric methods, there is no explicit way to correlate *a priori* the controller's performance with the dimension of the regressor vector. Additionally, the adaptive laws corresponding to the functional uncertainties can be reduced to matrix ODEs, and one of the terms of such ODEs is the Grammian matrix, which allows the user to gauge the effect of the number and distribution bases used to define the approximating native space. Finally, the proposed framework enables the introduction of standard unified metrics of the performance of the adaptive controller.

Future work directions are multiple and discussed throughout this paper. The first lies in enforcing that the closed-loop plant trajectory lies in some domain defined *a priori*. Furthermore, it is to be studied how to distribute approximation centers, whether and how those centers can be moved throughout the control process, and their effect on both the trajectory tracking error and the computational cost. Additionally, as already discussed in the introduction to the first paper of this two-part work, explicit connections to data-driven methods and deep learning should be explored. Finally, this work primarily focused on the effect of matched functional uncertainties. Neither the proposed results nor any of the existing results, to our knowledge, provide ultimate bounds on the tracking error as functions of any characterization of the unmatched functional uncertainty class other than the bound on the unmatched uncertainty. The problem of tighter characterizations of the unmatched uncertainties using a native space setting is left for future investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was in part performed with the support of National Science Foundation through Grant no. 2137159 and the US Army Research Lab under Grant no. W911QX2320001.

References

Abraham, R., Marsden, J. E., & Ratiu, T. (2012). *Manifolds, tensor analysis, and applications*: *vol. 75*, Springer Science & Business Media.

Adams, R. A., & Fournier, J. J. (2003). *Sobolev spaces*. Academic Press.

Anderson, R. B., Marshall, J. A., & L'Afflitto, A. (2020). Constrained robust model reference adaptive control of a tilt-rotor quadcopter pulling an unmodeled cart. *IEEE Transactions on Aerospace and Electronic Systems*, *57*(1), 39–54.

Anderson, R. B., Marshall, J. A., & L'Afflitto, A. (2021a). Novel model reference adaptive control laws for improved transient dynamics and guaranteed saturation constraints. *Journal of the Franklin Institute*, http://dx.doi.org/10.1016/j.jfranklin.2021.06.020.

Anderson, R. B., Marshall, J. A., & L'Afflitto, A. (2021b). Novel model reference adaptive control laws for improved transient dynamics and guaranteed saturation constraints. *Journal of the Franklin Institute*, *358*(12), 6281–6308.

Arabi, E., Gruenwald, B. C., Yucelen, T., & Nguyen, N. T. (2018). A set-theoretic model reference adaptive control architecture for disturbance rejection and uncertainty suppression with strict performance guarantees. *International Journal of Control*, *91*(5), 1195–1208. http://dx.doi.org/10.1080/00207179.2017.1312019.

Aulisa, E., Burns, J. A., & Gilliam, D. S. (2023). Approximate tracking for distributed parameter systems using only sensed data. *Systems & Control Letters*, *173*, Article 105477.

Aulisa, E., & Gilliam, D. (2015). *A practical guide to geometric regulation for distributed parameter systems*. CRC Press.

Berlinet, A., & Thomas-Agnan, C. (2011). *Reproducing kernel Hilbert spaces in probability and statistics*. Springer.

DeVore, R. A. (1998). Nonlinear approximation. *Acta Numerica*, *7*, 51–150. http://dx.doi.org/10.1017/S0962492900002816.

DeVore, R. A., & Lorentz, G. G. (1993). *Constructive approximation*. Springer-Verlag.

Farrell, J. A., & Polycarpou, M. M. (2006). *Adaptive approximation based control: Unifying neural, fuzzy and traditional adaptive approximation approaches*: *vol. 48*, John Wiley & Sons.

Ioannou, P., & Fidan, B. (2006). *Adaptive control tutorial*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Ioannou, P. A., & Sun, J. (2012). *Robust adaptive control*. Courier Corporation.

Khalil, H. K. (2002). *Nonlinear systems*. Upper Saddle River, NJ: Prentice Hall.

Kreyszig, E. (1989). *Introductory functional analysis with applications*. New York, NY: Wiley.

Krstic, M., Kanellakopoulos, I., & Kokotovic, P. (1995). *Nonlinear and adaptive control design*. John Wiley & Sons.

L'Afflitto, A. (2017). Differential games, partial-state stabilization, and model reference adaptive control. *Journal of the Franklin Institute*, *354*(1), 456–478. http://dx.doi.org/10.1016/j.jfranklin.2016.10.012.

L'Afflitto, A. (2018). Barrier Lyapunov functions and constrained model reference adaptive control. *IEEE Control Systems Letters*, *2*(3), 441–446. http://dx.doi.org/10.1109/LCSYS.2018.2842148.

Lavretsky, E., & Wise, K. A. (2012). Robust adaptive control. In *Robust and adaptive control: with aerospace applications*. Springer.

Müller, S., & Schaback, R. (2009). A Newton basis for kernel spaces. *Journal of Approximation Theory*, *161*(2), 645–655. http://dx.doi.org/10.1016/j.jat.2008.10.014.

Oden, J. (1986). *Qualitative methods in nonlinear mechanics*. Upper Saddle Hill, NJ: Prentic-Hall.

Paulsen, V. I., & Raghupathi, M. (2016). *An introduction to the theory of reproducing kernel Hilbert spaces*: *vol. 152*, Cambridge University Press.

Pazouki, M., & Schaback, R. (2011). Bases for kernel-based spaces. *Journal of Computational and Applied Mathematics*, *236*(4), 575–588.

Peterson, B., & Narendra, K. (1982). Bounded error adaptive control. *Transactions on Automatic Control*, *27*(6), 1161–1168. http://dx.doi.org/10.1109/TAC.1982.1103112.

Pomet, J. B., & Praly, L. (1992). Adaptive nonlinear regulation: Estimation from the Lyapunov equation. *Transactions on Automatic Control*, *37*(6), 729–740. http://dx.doi.org/10.1109/9.256328.

Powell, N., Liu, B., & Kurdila, A. J. (2022). Koopman methods for estimation of motion over unknown, regularly embedded submanifolds. In *American control conference* (pp. 2584–2591). IEEE, http://dx.doi.org/10.23919/ACC53348.2022.9867268.

Saitoh, S., & Sawano, Y. (2016). *Theory of reproducing kernels and applications*. Springer.

Sastry, S., & Bodson, M. (2011). *Adaptive control: stability, convergence and robustness*. Courier Corporation.

Schaback, R. (1994). Error estimates and condition Numbers for radial basis function interpolation. *Advances in Computational Mathematics*, *3*, 251–264.

Steinwart, I., Hush, D., & Scovel, C. (2006). An explicit description of the reproducing kernel Hilbert spaces of Gaussian RBF kernels. *Institute of Electrical and Electronics Engineers. Transactions on Information Theory*, *52*(10), 4635–4643.

Temlyakov, V. (2011). *Greedy approximation*: *vol. 20*, Cambridge, UK: Cambridge University Press.

Wang, H., L'Afflitto, A., & Kurdila, A. J. (2024). Computer codes for the implementation of adaptive control laws for nonlinear plants affected by scalar-valued uncertainties. URL https://lafflitto.com/Codes/ARC_Scalar_RKHS_Matlab_Codes.zip.

Wendland, H. (2004). *Scattered data approximation*: *vol. 17*, Cambridge University Press.

Wittwar, D., Santin, G., & Haasdonk, B. (2018). Interpolation with uncoupled separable matrix-valued kernels. arXiv preprint arXiv:1807.09111.