International Journal of Adaptive Control and Signal Processing

WILEY

RESEARCH ARTICLE

Robust Model Reference Adaptive Control Based on Reproducing Kernel Hilbert Spaces

Haoran Wang[1] | Andrew J. Kurdila[1] | Andrea L'Afflitto[2] | Derek Oesterheld[3] | Daniel J. Stilwell[3]

[1]Department of Mechanical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA | [2]Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA | [3]Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

Correspondence: Andrea L'Afflitto (a.lafflitto@vt.edu)

Received: 17 June 2024 | **Revised:** 20 January 2025 | **Accepted:** 2 March 2025

Funding: This work was supported by The National Science Foundation and DEVCOM Army Research Laboratory. NSF through Grant no. 2137159, the U.S. ARMY RESEARCH LABORATORY under Grant no. W911QX2320001, and ONR grants through grants N00014-23-1-2345 and N00014-21-1-2512.

Keywords: distributed parameter systems | model reference adaptive control | robust adaptive control

ABSTRACT

This article introduces native space embedding for robust adaptive control of ordinary differential equations that contain vector-valued functional uncertainties in a reproducing kernel Hilbert space (RKHS). The proposed approach is based on a two-phase method for analyzing and designing adaptive controllers. In the first phase, a limiting distributed parameter system (DPS), which describes the ideal closed-loop system's performance, is introduced. The limiting DPS is not realizable in practice since it evolves in a generally infinite-dimensional space. In the second phase, consistent finite-dimensional approximations of the DPS are introduced to determine realizable controllers. Uniform ultimate bounds on the trajectory tracking error dynamics are derived for the functional uncertainty classes contained in the native space. These bounds are derived in terms of the power function of the RKHS or in terms of the fill distance of centers that define the scattered basis for approximations. Two numerical examples demonstrate the applicability of the proposed results.

1 | Introduction

1.1 | Motivation and Problem Statement

This article presents a model reference adaptive control (MRAC) synthesis framework for nonlinear plants whose matched uncertainties are not represented by any real parameterization given a priori, but are only known to belong to some reproducing kernel Hilbert space (RKHS, also known as *native space*). The proposed framework uses a two-phase approach. In the first phase, an infinite-dimensional system, or distributed parameter system (DPS), is introduced. This system represents the limiting behavior of the trajectory tracking error dynamics sought in the control design problem, and, for this reason, we refer to it as the *limiting DPS system*. Due to its infinite-dimensional nature, the limiting DPS contains *nonparametric estimates* of the functional uncertainty, and, hence, the underlying control law and adaptive laws are not realizable. Thus, in the second phase, realizable controllers are determined by consistently approximating the limiting DPS in an RKHS. This overall approach is referred to as *RKHS embedding* or *native space embedding* since all of the trajectories of the ODEs determined by approximation can be understood as being embedded in the state space of the limiting DPS. To the authors' knowledge, this is the first article that describes a general approach for robust modifications of MRAC and derives the rates of convergence of the ultimate tracking error for functional uncertainty classes in an RKHS of vector-valued functions.

The arguments used in this article build on those typical of classical adaptive control theory, wherein the matched uncertainties affecting the ordinary differential equations (ODEs) are represented using a finite number of unknown real parameters and some fixed regressor vector. In this regard, the proposed framework provides a natural extension of those results in classical and robust MRAC [1–4], which constitute the standard working setting for modern adaptive control of ODEs. Because the analysis in the classical setting relies on a parameterization of the functional uncertainties for a fixed regressor vector, classical adaptive control methods as referred to as *parametric* in this article. Here we seek to formulate the adaptive control problem for ODEs in terms of *nonparametric or functional uncertainty classes*. Furthermore, we explore how well-known parametric adaptive control strategies for ODEs, such as those based on methods like those in References [1–4], can be ported to the nonparametric setting and achieve some of the advantages that arise from using an RKHS for the uncertainty. Such advantages include predicting the ultimate bounds on the tracking error in terms of properties of the functional uncertainty class such as the smoothness index and the distance among centers defining approximations of the uncertainty in the native space. Classical parametric methods, such as those in References [1–4], do not make such predictions.

The theoretical tools that are available when using RKHSs have had a deep and widespread impact on modern methods for statistical learning theory, Bayesian estimation, and Gaussian process estimation for stochastic systems [5–7]. RKHS-based adaptive control methods based on Gaussian process techniques continue to be investigated, with notable recent contributions including References [8–10] as representative examples. However, in such cases, the uncertainties appearing in the governing ODE are interpreted as stochastic uncertainties verifying some Gaussian model. The closed-loop performance guarantees obtained in these Gaussian process methods are "performance bounds with high confidence." This article continues the purely deterministic approach in [11–20]. The authors think that the derived deterministic ultimate bounds on the trajectory tracking error can be easier to state, interpret, and compute than the probabilistic bounds that arise when viewing the functional uncertainty in the setting of Gaussian processes.

In view of this claim, perhaps one of the most important attributes of native space embedding methods is how the mathematical tools of an RKHS, which are relatively straightforward to use, are used to obtain sharper error bounds than those usually obtained in real parametric adaptive control. As discussed in References [2, 21], a critical step in approximation-based, deterministic adaptive control invokes a uniform approximation assumption in $L^\infty(\Omega, \mathbb{R}^n)$ over some compact set $\Omega \subset \mathbb{R}^n$ that contains a closed-loop trajectory of interest. Of course, the particular form that the uniform approximation assumption takes can vary. The Weierstrass theorem, Cybenko's theorem, universal approximations by neural networks, and other results as described in References [2, 21] are a few examples of the theorems that are often applied. However, if the uncertainty lies in a vector-valued RKHS $\boldsymbol{\mathcal{H}}$, then the construction of the RKHS itself ensures that a type of approximation assumption holds in the norm on $\boldsymbol{\mathcal{H}}$, which in this article defines a stronger topology than that of $L^\infty(\Omega, \mathbb{R}^n)$. Even more importantly, compared with the oft-cited theorems, there are standard tools in RKHS theory to obtain closed-form expressions for bounds on the pointwise error of approximation. Instead of just assuring that a family of basis functions is dense in the continuous functions (e.g., via the Weierstrass theorem or universality of neural networks), the structure of the RKHS gives a natural way to estimate the *rates of convergence* in proofs of stability or convergence.

This article has been developed in the broader context of an approach based on learning-based control, which the authors frame in the context of native spaces. The proposed results can be used to govern the learn-then-control architecture and the switched learn-and-control strategies [22], [23, Ch. 6]. Worthy of mention is how this research trend in the area of learning theory and adaptive control has been gaining increasing interest, as shown in the recent References [24–27] to name a few.

1.2 | Advantages of Native Space Embeddings to Capture Nonlinearities

In this article, we consider the nonlinear *plant dynamics*

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

where $x(t) \in \mathbb{X} \triangleq \mathbb{R}^n$ denotes the *plant state*, $u(t) \in \mathbb{U} \triangleq \mathbb{R}^m$ denotes the *control input*, $A \in \mathbb{R}^{n\times n}$ is unknown, $B \in \mathbb{R}^{n\times m}$ is known and such that the pair (A, B) is controllable, f is unknown and denotes a nonlinear *matched uncertainty* in the vector-valued RKHS $\boldsymbol{\mathcal{H}}$, and $E_{x(t)}$ denotes the *evaluation operator* at $x(t) \in \mathbb{X}$. In this article, we let $\boldsymbol{\mathcal{H}}$ be the m-times Cartesian product of the scalar-valued RKHS $\mathcal{H}$, that is, $\boldsymbol{\mathcal{H}} \triangleq \mathcal{H}^m$. Notably, in this article, we consider f as a fixed element of the RKHS $\boldsymbol{\mathcal{H}}$. We also consider the *reference model*

$$\dot{x}_r(t) = A_r x_r(t) + B_r r(t), \quad x_r(t_0) = x_{r,0}, \quad t \geq t_0 \tag{2}$$

where $x_r(t) \in \mathbb{X}$ denotes the *reference trajectory*, $r(t) \in \mathbb{U}$ denotes the user-defined, uniformly continuous and bounded *reference command input*, $A_r \in \mathbb{R}^{n\times n}$ is Hurwitz, and $B_r \in \mathbb{R}^{n\times m}$. Given the RKHS $R > 0$ and $\epsilon > 0$, we define the uncertainty classes

$$C_R \triangleq \left\{ f \in \boldsymbol{\mathcal{H}} \;:\; \|f\|_{\boldsymbol{\mathcal{H}}} \leq R \right\} \tag{3}$$

$$C_{R,\epsilon,N} \triangleq \left\{ f \in C_R : \|(I - \boldsymbol{\Pi}_N) f\|_{\boldsymbol{\mathcal{H}}} \leq \epsilon \right\} \tag{4}$$

where the integer N indicates the number of centers that define a scattered basis in the RKHS, and, hence, define the finite-dimensional space $\boldsymbol{\mathcal{H}}_N \triangleq \mathcal{H}_N^m$ that approximates $\boldsymbol{\mathcal{H}}$, and $\boldsymbol{\Pi}_N$ denotes the $\boldsymbol{\mathcal{H}}$-orthogonal projection onto $\boldsymbol{\mathcal{H}}_N$. Assuming that $f \in C_{R,\epsilon,N}$, we are interested in obtaining uniform ultimate bounds on the *trajectory tracking error* $e_N(t) \triangleq x_N(t) - x_r(t)$, $t \geq t_0$, of the form

$$\|e_N(t)\|_{\mathbb{X}} \leq C\mathcal{F}(N), \quad \text{for all } t \geq T \tag{5}$$

where $x_N : [t_0, \infty) \to \mathbb{X}$ denotes the closed-loop plant trajectory when using $\boldsymbol{\mathcal{H}}_N \subseteq \boldsymbol{\mathcal{H}}$ to approximate $f \in \boldsymbol{\mathcal{H}}$, $C > 0$, $T \geq t_0$, and $\mathcal{F}$ is a known decreasing function of N. If $\mathcal{F}(N)$ is of the order of the error of the best offline approximation of f from the subspace $\boldsymbol{\mathcal{H}}_N$, then we say that the adaptive control strategy is *approximation theory asymptotically nearly optimal*, or AAO for short.

In the vast majority of publications on MRAC systems, it is assumed that the nonlinear matched uncertainties f in (1) are elements of the uncertainty class

$$C_{\overline{R},\Psi_N} \triangleq \left\{ f = \Psi_N^{\mathrm{T}} \Theta_N \in \boldsymbol{\mathcal{H}}_N \subseteq \boldsymbol{\mathcal{H}} : \|\Theta_N\| \leq \overline{R} \right\} \tag{6}$$

where $\overline{R} \geq 0$ is given and $\Psi_N \triangleq \{\psi_1, \ldots, \psi_{\dim(\boldsymbol{\mathcal{H}}_N)}\}$ denotes the span of some fixed set of *regressors* or *basis functions*. Assuming no unmatched uncertainties and that matched uncertainties are parameterized by the given regressors, classical MRAC systems [2, Ch. 9] ensure asymptotic convergence of the tracking error to zero, despite matched uncertainties. However, this is an idealistic result, and hence, robustifications of classical MRAC systems [2, Ch. 11] are needed. The proposed results are considerably stronger than classical robust MRAC systems. Indeed, the functional uncertainty class $C_{R,\epsilon,N}$ considered in this article, and, hence, C_R, are generally much larger than $C_{\overline{R},\Psi_N}$. For each $\overline{R} > 0$, it is easy to define $R > 0$ such that $C_{\overline{R},\Psi_N} \subseteq C_R$; on the other hand, if $\boldsymbol{\mathcal{H}}$ is infinite-dimensional, then $C_R \nsubseteq C_{\overline{R},\Psi_N}$ for any R and $\overline{R} > 0$. Also, it holds that $C_{\overline{R},\Psi_N} \subseteq C_{R,\epsilon,N}$. Therefore, the conventional uncertainty class $C_{\overline{R},\Psi}$ only contains functions in $\boldsymbol{\mathcal{H}}_N$, but $C_{R,\epsilon,N}$ contains all functions in $\boldsymbol{\mathcal{H}}$ whose approximations by elements in $\boldsymbol{\mathcal{H}}_N$ have an error of best approximation in $\boldsymbol{\mathcal{H}}_N$ less than ϵ. Thus, in general, it holds that

$$C_{\overline{R},\Psi_N} \subseteq C_{R,\epsilon,N} \subseteq C_R \tag{7}$$

Furthermore, classical robust MRAC systems do not provide estimates on the rate of convergence of the tracking error similar to (5), that is, estimates on the convergence of the tracking error based on the richness of information provided by the regressors. Finally, the ultimate bound in (5) holds the potential for functional uncertainty classes that describe a much broader collection of uncertain systems.

In this article, we choose reproducing kernels that define native spaces $\boldsymbol{\mathcal{H}}$ consisting of bounded functions on the state space $\mathbb{X}$. This makes proofs of well-posedness, stability, and convergence more closely resemble those used in classical real parametric adaptive control. On one hand, this assumption on the kernels excludes some iconic cases. For example, functions whose components contain polynomial powers x^p are not contained in any such space $\boldsymbol{\mathcal{H}}$ of bounded functions. On the other hand, however, this assumption is not particularly limiting. Indeed, we only need that f agrees with a function $g \in \boldsymbol{\mathcal{H}}$ over a bounded set $\Omega \subseteq \mathbb{X}$ that contains the closed-loop trajectory; carrying the notation for such an $\Omega \subseteq \mathbb{X}$ in all the analyses that follow unduly complicates the expressions, and we simply write $f \in \boldsymbol{\mathcal{H}}$ in the rest of the article, keeping this caveat in mind. With this interpretation, functional uncertainty that contains unbounded functions, such as polynomials, can be handled readily.

1.3 | Summary of the Original Results

The first set of key results appears in Theorems 1 and 2, which describe the limiting DPS. Theorem 1 provides sufficient conditions for a control law and set of adaptive laws, which resemble the classical control and adaptive laws for classical MRAC [2, Ch. 9], but involve elements of infinite-dimensional spaces, to assure uniform asymptotic convergence of the plant trajectory (i.e., the closed-loop limiting DPS) to the reference trajectory. Due to the infinite-dimensional elements both in the closed-loop limiting DPS and in the adaptive laws, it is impossible to leverage classical Lyapunov-like arguments, such as those presented in [28], Th. 4.1, to assure the existence of complete solutions. Thus, Theorem 2 provides sufficient conditions that ensure that the closed-loop limiting DPS is well-posed and forward complete; these key questions are not posed in the preliminary article [20], and apply to [20] as well.

A second set of key results is described in Theorems 3 and 4, which describe how the controller and the adaptive law addressed in Theorems 1 and 2 can be implemented in finite-dimensional spaces. To this goal, we use two approaches inspired by the dead-zone modification of MRAC [29] and the σ-modification of MRAC [4], respectively. Because the approximation of the functional uncertainty introduces an error, realizable controllers require a robust modification that averts the associated parameter drift. The results of Theorem 3 are based on the initial work on the dead-zone methods described in [20]. However, Theorem 3 improves the results in [20] by making the role of the uncertainty class $C_{R,\epsilon,N} \subset \mathcal{H}$ explicit. Theorem 4 is an entirely new result.

Let Ω be a bounded set that contains the closed-loop trajectory $x_N(t), t \geq t_0$, and let $\mathcal{P}_N$ denote the *power function* of the subspace $\mathcal{H}_N$ in $\mathcal{H}$, which is commonly used [30, 31] to construct upper bounds on errors of approximation from $\mathcal{H}_N$. We define

$$\epsilon_N \triangleq \|(I - \mathbf{\Pi}_N) f\|_{\boldsymbol{\mathcal{H}}} \tag{8}$$

as the error of the best (offline) approximation of $f \in \boldsymbol{\mathcal{H}} \triangleq \mathcal{H}^m$ from $(\mathcal{H}_N)^m$. Leveraging the results of Theorem 3, the ultimate bound given by (5) specializes to

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq \frac{2}{\lambda_{\min}(Q)} \|PB\| \sup_{\eta \in \Omega} \mathcal{P}_N(\eta) \epsilon_N \tag{9}$$

for all $t \geq T$ and for some $T \geq t_0$, where the pair of symmetric, positive-definite matrices $(P, Q) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ verify an algebraic Lyapunov equation, and $\lambda_{\min}(\cdot)$ denotes the eigenvalue of its argument with the smallest real part. The function $\mathcal{P}_N : \mathbb{X} \to \mathbb{R}^+$ is the power function of the subspace $\mathcal{H}_N$ in $\mathcal{H}$, which is described in detail in Section 2.3. Equation(9) improves a similar result derived in [20] in that it clarifies the role of the uncertainty class $C_{R,\epsilon,N}$ and is tighter. Applying the results of Theorem 4, the ultimate bound given by (5) specializes to

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}}^2 \leq \frac{2}{\sigma \lambda_{\min}(Q)} \|PB\|^2 \sup_{\eta \in \Omega} \mathcal{P}_N^2(\eta) \epsilon_N^2 + \mathfrak{b} \tag{10}$$

for all $t \geq T$ and for some $T \geq t_0$, where $\mathfrak{b} > 0$ captures a weighted squared norm of the uncertainties.

The trajectory tracking error bound (9) has the same form as that derived in [5–7] for offline approximations of functions based on statistical and machine learning theory using samples from discrete, stochastic systems with identically distributed observations. The error bound (10) includes additional terms that do not appear in the usual discrete stochastic framework of learning theory. To the authors' knowledge, this is the first rigorous error bound on the ultimate performance of adaptive control methods *that uses only deterministic analysis* to obtain the same type of bounds as those found in learning theory approaches for offline function estimation.

Corollaries 1 and 2 are new results that provide expressions for the ultimate bounds on the trajectory tracking error that can be attained by applying Theorems 3 and 4 and that are alternatives to (9) and (10), respectively. These results also provide a correlation between the smoothness of the unknown nonlinearity affecting (1), the kernel functions, and the ultimate bound on the controlled trajectory tracking error dynamics. Specifically, let Ξ_N denote a collection of kernel centers that define the scalar-valued vector space $\mathcal{H}_N$ used to define $\boldsymbol{\mathcal{H}}_N$. Given a sufficiently regular, bounded set Ω containing the trajectory of the closed-loop limiting DPS, let the *fill distance of the centers* Ξ_N in Ω be defined as

$$h_{\Xi_N,\Omega} \triangleq \sup_{x\in\Omega} \min_{\xi_i \in \Xi_N} \|x-\xi_i\|_{\mathbb{X}} \tag{11}$$

Corollary 1 shows that, applying the adaptive laws postulated by Theorem 3, and using exponential kernels, inverse multiquadrics kernels, or compactly supported Wendland functions of regularity r, the ultimate tracking error is captured by

$$\|x_N(t)-x_r(t)\|_{\mathbb{X}} \le \begin{cases} O\left(e^{-\left(\mu\frac{\log(h_{\Xi_N,\Omega})}{h_{\Xi_N,\Omega}}\right)}\right), & \text{Gaussian,} \\ O\left(e^{-\left(\frac{\eta}{h_{\Xi_N,\Omega}}\right)}\right), & \text{inverse multiquadric,} \\ O\left(h^{r+1/2}_{\Xi_N,\Omega}\right), & \text{Wendland, smoothness } r \end{cases} \tag{12}$$

where $\mu, \eta > 0$ are constant. Similarly, Corollary 2 shows that, applying the adaptive laws postulated by Theorem 4, and using exponential kernels, inverse multiquadrics kernels, or compactly supported Wendland functions of regularity r, the ultimate tracking error is captured by

$$\begin{aligned}&\|x_N(t)-x_r(t)\|_{\mathbb{X}} \\ &\le \begin{cases} O\left(\frac{1}{\sigma}\left(e^{-\tilde{\mu}|\log(h_{\Xi_N,\Omega}))|/h_{\Xi_N,\Omega}}+\sigma\right)\right), & \text{Gaussian,} \\ O\left(\frac{1}{\sigma}\left(e^{-\mu/h_{\Xi_N,\Omega}}+\sigma\right)\right), & \text{inverse multiquadric,} \\ O\left(\frac{1}{\sigma}\left(h^{r+1/2}_{\Xi_N,\Omega}+\sigma\right)\right), & \text{Wendland, smoothness } r \end{cases}\end{aligned} \tag{13}$$

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

These results are particularly advantageous because it is possible to use information about the smoothness of the uncertainty f to choose an appropriate kernel for the synthesis of the adaptive laws. It is known that the Gaussian kernel, the Wendland kernel of smoothness index r, and Sobolev–Matérn kernels of smoothness index $m > n/2$ over the set $\mathbb{X}$ define a scalar-valued native space $\mathcal{H}$ that is contained in $C^\infty(\mathbb{X})$, the Sobolev space $W^{r+(n+1)/2}(\mathbb{X})$, and the Sobolev space $W^m(\mathbb{X})$, respectively; recall that the Sobolev space $W^s(\mathbb{X})$ for integer $s > 0$ is the subspace of $L^2(\mathbb{X})$ that contains functions that have (weak) derivatives contained in $L^2(\mathbb{X})$ for all orders through s. The scale of Sobolev spaces $W^s(\mathbb{X})$ for $s > 0$ gives a precise way of describing the smoothness of functions. Therefore, in the proposed MRAC framework, this observation implies that, if the smoothness of the uncertainty f is known, then it is possible to tailor the kernel used in a native space embedding controller to this smoothness and produce more suitable ultimate bounds on the trajectory tracking error dynamics. Such tailoring can also have practical implications for the dimensionality of the problem. The offline approximations of smoother functions generally converge faster as a function of N than those for less smooth functions. This observation is supported by the numerical simulation results presented in Section 6.

Because (8) captures the optimal offline approximation error, the primary observations regarding Theorems 3 and 4, and their corollaries, can be summarized as follows. First, the ultimate bound on tracking performance for the dead-zone strategy in Theorem 3, for a suitably scaled dead-zone, is AAO over the functional uncertainty class $C_{R,\epsilon,N} \subset C_R \subset \boldsymbol{\mathcal{H}}$. Second, the ultimate bound on tracking performance for the σ-modification in Theorem 4 is not AAO over the functional uncertainty class $C_{R,\epsilon,N}$. The second observation does not mean, of course, that some additional hypotheses, such as the persistence of excitation, might suffice to conclude that the σ-modification is AAO. Also, the fact that the bound on the tracking error for the σ-modification in real parametric adaptive control can be large is well-known, as discussed in [4, 21]. In contrast to [4, 21], however, the ultimate bounds on the tracking error provided in this article are expressed relative to the norm on the generally infinite dimensional space $\boldsymbol{\mathcal{H}}$ for the larger uncertainty class $C_{R,\epsilon,N}$. Our numerical studies of the σ-modification suggest that the ultimate bounds on the tracking error induced by the σ-modification of MRAC can be very conservative. Refinements that give tighter bounds under additional assumptions, such as persistency of excitation, are needed and left for future work.

2 | Elements of Native Space Theory

This section describes essential results from the theory of RKHS spaces needed for the adaptive control problems in this article; for additional details, see [32, 33]. The real numbers and nonnegative real numbers are denoted by $\mathbb{R}$ and $\mathbb{R}^+$. Given the normed vector space V and the Banach space W over the field of real numbers, $\mathcal{L}(V,W)$ denotes the Banach *space of bounded linear operators* mapping V into W. For brevity, we denote $\mathcal{L}(V,V)$ by $\mathcal{L}(V)$.

2.1 | Scalar-Valued Native Spaces

Let $\Omega \subseteq \mathbb{X}$, and let $\mathfrak{K} : \Omega \times \Omega \to \mathbb{R}$ denote the *reproducing kernel function*. The *kernel section or basis function centered at* $\xi \in \Omega$ is defined as $\mathfrak{K}_\xi(\cdot) \triangleq \mathfrak{K}(\xi,\cdot)$. The native space generated by $\mathfrak{K}(\cdot,\cdot)$ is a Hilbert space $\mathcal{H}$ of real-valued functions defined as the closed linear span of the kernel sections centered at all $\xi \in \Omega$, that is,

$$\mathcal{H} \triangleq \overline{\text{span}\{\mathfrak{K}_\xi(\cdot) : \xi \in \Omega\}} \tag{14}$$

note that if there are infinitely many elements in Ω, then $\mathcal{H}$ is generally infinite-dimensional. We often refer to $\mathcal{H}$ as a *scalar RKHS* since $\mathcal{H}$ contains scalar-valued functions over Ω.

The proposed definition of native spaces allows that the set Ω over which the functions in $\mathcal{H}$ are defined is a subset $\Omega \subseteq \mathbb{X} \triangleq \mathbb{R}^n$. In this article, the kernel over Ω is typically generated as the restriction of a "standard kernel" $\mathfrak{K} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ that is defined over $\mathbb{X} \triangleq \mathbb{R}^n$. We overload the notation and use the same symbol $\mathfrak{K}$ to denote the reproducing kernel over $\mathbb{X}$ or $\Omega \subset \mathbb{X}$. We also use $\mathcal{H}$ to denote the native space of functions that are defined on all of $\mathbb{X}$ in terms of the standard kernel $\mathfrak{K} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$, as well as the RKHS over Ω that is obtained from the restricted kernel; see [33, 34] for details on the theory of RKHS defined by restrictions of kernels.

Given an RKHS $\mathcal{H}$, the *evaluation operator* $E_\xi : \mathcal{H} \to \mathbb{R}$ is defined so that $E_\xi f = f(\xi)$ for each $\xi \in \Omega$. The *evaluation operator* E_ξ is bounded and linear, and, hence, its *adjoint operator* $E_\xi^* : \mathbb{R} \to \mathcal{H}$, is a well-defined bounded linear operator. These operators are related via the *reproducing property*

$$\langle E_\xi f, \alpha \rangle_{\mathbb{R}} = \left\langle f, E_\xi^* \alpha \right\rangle_{\mathcal{H}} = \left\langle f, \mathfrak{K}_\xi \alpha \right\rangle_{\mathcal{H}} = \alpha f(\xi) \tag{15}$$

for any $\alpha \in \mathbb{R}$, $\xi \in \Omega$, and $f \in \mathcal{H}$.

In this article, we assume that the kernel $\mathfrak{K}(\cdot,\cdot)$ that defines $\mathcal{H}$ is continuous on $\Omega \times \Omega$ and is *bounded on the diagonal* in the sense that there is a constant $\overline{\mathfrak{K}} > 0$ such that

$$\mathfrak{K}(\xi, \xi) = \|\mathfrak{K}_\xi\|_{\mathcal{H}}^2 \le \overline{\mathfrak{K}}^2, \quad \text{for all } \xi \in \Omega \tag{16}$$

This assumption is satisfied for many standard choices of kernels defined on $\Omega \subseteq \mathbb{R}^n$, including the Gaussian, Sobolev–Matérn, inverse multiquadric, or Wendland kernels [30]. The boundedness on the diagonal of $\mathfrak{K}(\cdot,\cdot)$ is important in our closed-loop analysis since it ensures that the operator norm $\|E_\xi\| = \|E_\xi^*\|$ is *uniformly* bounded. Indeed, considering E_ξ as acting on $\mathcal{H}$ into $\mathbb{R}$, we have that $\|E_\xi\| \le \overline{\mathfrak{K}}$ since, by the Cauchy-Schwarz inequality,

$$\left| E_\xi f \right| = |f(\xi)| = \left| \left\langle f, \mathfrak{K}_\xi \right\rangle_{\mathcal{H}} \right| \le \overline{\mathfrak{K}} \|f\|_{\mathcal{H}} \tag{17}$$

Since f is considered a fixed element of the RKHS $\mathcal{H}$ and the functional dependency of the matched uncertainty on the state is captured by the evaluation operator E_ξ for any $\xi \in \Omega$, (17) provides an upper bound on the matched functional uncertainty that explicitly distinguishes between a contribution given by the chosen RKHS, namely $\overline{\mathfrak{K}}$, and a contribution given by the uncertainty itself, namely $\|f\|_{\mathcal{H}}$, which is constant.

This article's implementable control schemes are derived from specific approximations in the native space $\mathcal{H}$. To start with, let $\Xi_N \triangleq \{\xi_i : 1 \le i \le N\} \subset \Omega$ denote the *set of kernel centers*. A *finite-dimensional subspace of approximating functions* is defined as the linear span of the kernel sections located at the set of kernel centers $\mathfrak{K}_{\xi_i}$, that is,

$$\mathcal{H}_N \triangleq \text{span}\left\{ \mathfrak{K}_{\xi_i}(\cdot) : \xi_i \in \Xi_N \right\} \tag{18}$$

An RKHS $\mathcal{H}$ and an associated finite-dimensional subspace of approximating functions $\mathcal{H}_N$ are correlated by $\Pi_N : \mathcal{H} \to \mathcal{H}_N$, which captures the *orthogonal projection* of any function f that resides in $\mathcal{H}$ onto the finite dimensional subspace $\mathcal{H}_N$. With this definition and the reproducing properties, for any $f \in \mathcal{H}$, it holds that $f(\xi) = \left\langle \mathfrak{K}_\xi, \Pi_N f \right\rangle_{\mathcal{H}} + \left\langle \mathfrak{K}_\xi, (I - \Pi_N) f \right\rangle_{\mathcal{H}}$ for any $\xi \in \Omega$.

2.2 | Vector-Valued Native Spaces

In this article, we consider native spaces $\boldsymbol{\mathcal{H}}$ of vector-valued functions given by the Cartesian product $\boldsymbol{\mathcal{H}} \triangleq \mathcal{H}^m = \mathcal{H} \times \cdots \times \mathcal{H}$, where $\mathcal{H}$ denotes a scalar-valued native space defined by means of the scalar kernel $\mathfrak{K}(\cdot,\cdot)$ over the set Ω. Similar to the scalar-valued case, we refer to the spaces $\boldsymbol{\mathcal{H}} = \mathcal{H}^m$ as *vector-valued native spaces*. The general theory of vector-valued native spaces $\boldsymbol{\mathcal{H}}$ that contain $\mathbb{U}$-valued functions over Ω is based on operator-valued kernels [32, 35], and this theoretical framework is more abstract than it is for the scalar-valued case. Approaches by the authors in the full generality of operator-valued kernels have been presented in Reference [17] for the development of observers and References [18, 20] for the adaptive control schemes.

Assuming that vector-valued RKHSs are defined as the Cartesian product of scalar-valued RKHSs, given $\mathbb{U} \triangleq \mathbb{R}^m$, the evaluation operator $E_\xi : \boldsymbol{\mathcal{H}} \to \mathbb{U}$ is such that

$$E_\xi f = f(\xi), \quad \text{for all } \xi \in \Omega, f \in \boldsymbol{\mathcal{H}} \tag{19}$$

We also define the associated *operator kernel* and its *operator kernel section* as $\mathcal{K}(\xi,\eta) \triangleq \mathfrak{K}(\xi,\eta) I$ and $\mathcal{K}_\xi(\cdot) \triangleq \mathfrak{K}(\xi,\cdot) I$, respectively, for all $\xi, \eta \in \Omega$, where I denotes the identity on $\mathbb{R}^m$. It holds that $\mathcal{K}(\xi,\eta) \in \mathcal{L}(\mathbb{U})$ for each $\xi, \eta \in \Omega$, and the real vector-valued native space $\boldsymbol{\mathcal{H}}$ can be equivalently expressed as

$$\boldsymbol{\mathcal{H}} = \overline{\text{span}\{\mathcal{K}_\xi \alpha \mid \xi \in \Omega, \alpha \in \mathbb{U}\}} \tag{20}$$

Thus, $E_\xi \in \mathcal{L}(\boldsymbol{\mathcal{H}}, \mathbb{U})$ and $\mathcal{K}_\xi \in \mathcal{L}(\mathbb{U}, \boldsymbol{\mathcal{H}})$ for each $\xi \in \Omega$. The reproducing property for vector-valued native spaces takes the form

$$\left\langle \mathcal{K}_\xi \alpha, f \right\rangle_{\boldsymbol{\mathcal{H}}} = \left\langle E_\xi^* \alpha, f \right\rangle_{\boldsymbol{\mathcal{H}}} = \left\langle \alpha, E_\xi f \right\rangle_{\mathbb{U}} = \alpha^{\mathrm{T}} f(\xi) \in \mathbb{R} \tag{21}$$

for any $\alpha \in \mathbb{U}$, $f \in \boldsymbol{\mathcal{H}}$, and $\xi \in \Omega$. It follows from (21) that $E_\xi^* \triangleq \left(E_\xi\right)^* = \mathcal{K}_\xi$ for each $\xi \in \Omega$. In fact, if the kernel $\mathfrak{K}(\cdot,\cdot)$ that defines $\mathcal{H}$ is bounded on the diagonal, as we assume in this article, then a uniform bound on the operator norm of the evaluation operator $E_\xi : \boldsymbol{\mathcal{H}} \to \mathbb{U}$ is given by

$$\|E_\xi\| = \|E_\xi^*\| = \|\mathcal{K}_\xi\| \le \overline{\mathfrak{K}}, \quad \text{for all } \xi \in \Omega \tag{22}$$

Similar to the scalar case, if E_ξ is considered as an operator on $\boldsymbol{\mathcal{H}}$ into $\mathbb{U}$, then this bound follows directly from the fact that

$$\|\mathcal{K}_\xi \alpha\|_{\boldsymbol{\mathcal{H}}}^2 = \|\mathfrak{K}_\xi \alpha\|_{\boldsymbol{\mathcal{H}}}^2 = \|\mathfrak{K}_\xi\|_{\mathcal{H}}^2 \sum_{i=1}^m |\alpha_i|^2 \le \overline{\mathfrak{K}}^2 \|\alpha\|_{\mathbb{U}}^2 \qquad \text{for all } \alpha \in \mathbb{U} \tag{23}$$

To introduce a finite-dimensional subspace to approximate $\boldsymbol{\mathcal{H}}$, suppose that the real-valued native space $\mathcal{H}$ contains a finite-dimensional subspace $\mathcal{H}_N$. We define a finite-dimensional subspace of $\boldsymbol{\mathcal{H}}$ as $\boldsymbol{\mathcal{H}}_N \triangleq \mathcal{H}_N \times \cdots \times \mathcal{H}_N$, where the Cartesian product is repeated m times. The corresponding *projection operator* $\mathbf{\Pi}_N : \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}_N$ is defined as

$$\mathbf{\Pi}_N f \triangleq \left[\Pi_N f_1, \dots, \Pi_N f_m\right]^{\mathrm{T}} \tag{24}$$

Following the convention of the scalar case, for any $f \in \boldsymbol{\mathcal{H}}$ it holds that

$$\alpha^{\mathrm{T}} f(\xi) = \left\langle \mathcal{K}_\xi \alpha, \mathbf{\Pi}_N f \right\rangle_{\boldsymbol{\mathcal{H}}} + \left\langle \mathcal{K}_\xi \alpha, (I - \mathbf{\Pi}_N) f \right\rangle_{\boldsymbol{\mathcal{H}}} \tag{25}$$

for all $\alpha \in \mathbb{U}$, $f \in \boldsymbol{\mathcal{H}}$, and $\xi \in \Omega$, where $I : \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}$ denotes the identity operator over $\boldsymbol{\mathcal{H}}$.

2.3 | The Power Function

A relevant interpretation of (25) is that, when computing finite-dimensional approximations of $f \in \boldsymbol{\mathcal{H}}$, the projection error is given by $(I - \mathbf{\Pi}_N) f$. Although the exact value of this error cannot be calculated explicitly, RKHS theory provides us with a tool, the *power function*, to bound this error. Following [30, 31], we define the scalar power function as

$$\mathcal{P}_N(\xi) \triangleq \|\left(I - \Pi_N\right) \mathfrak{K}_\xi\|_{\mathcal{H}}, \quad \text{for each } \xi \in \Omega \tag{26}$$

which can be calculated as

$$\mathcal{P}_N(\xi) = \sqrt{\mathfrak{K}(\xi, \xi) - \mathfrak{K}_N(\xi, \xi)} \quad \text{for each } \xi \in \Omega \tag{27}$$

where $\mathfrak{K}_N(\cdot, \cdot)$ denotes the reproducing kernel of $\mathcal{H}_N$. It is well-known [30–32] that the reproducing kernel $\mathfrak{K}_N$ is given in closed form by the expression

$$\mathfrak{K}_N(x, y) \triangleq \mathfrak{K}_{\Xi_N}^{\mathrm{T}}(x) \mathbb{K}_N^{-1} \mathfrak{K}_{\Xi_N}^{\mathrm{T}}(y) \tag{28}$$

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

$$\mathbb{K}_N = \begin{bmatrix} \langle \mathfrak{K}_{\xi_1}, \mathfrak{K}_{\xi_1} \rangle_{\mathcal{H}} & \dots & \langle \mathfrak{K}_{\xi_1}, \mathfrak{K}_{\xi_N} \rangle_{\mathcal{H}} \\ \vdots & \ddots & \vdots \\ \langle \mathfrak{K}_{\xi_N}, \mathfrak{K}_{\xi_1} \rangle_{\mathcal{H}} & \dots & \langle \mathfrak{K}_{\xi_N}, \mathfrak{K}_{\xi_N} \rangle_{\mathcal{H}} \end{bmatrix} \in \mathbb{R}^{N \times N} \tag{29}$$

denotes the *Grammian matrix*.

The finite-dimensional approximation error can be bounded by the power function [36] as

$$\left| E_\xi \left(I - \Pi_N\right) f \right| \leq \mathcal{P}_N(\xi) \|(I - \Pi_N) f\|_{\mathcal{H}} \leq \mathcal{P}_N(\xi) \|f\|_{\mathcal{H}}, \quad \text{for all } \xi \in \Omega, f \in \mathcal{H} \tag{30}$$

Because, in this article, we define $\boldsymbol{\mathcal{H}}$ by means of a Cartesian product, (30) implies that

$$\left\| E_\xi (I - \mathbf{\Pi}_N) f \right\|_{\boldsymbol{\mathcal{H}}}^2 = \sum_{i=1}^{m} \left| E_\xi (I - \Pi_N) f_i \right|^2 = \mathcal{P}_N^2(\xi) \|(I - \mathbf{\Pi}_N) f\|_{\boldsymbol{\mathcal{H}}}^2 \leq \mathcal{P}_N^2(\xi) \|f\|_{\boldsymbol{\mathcal{H}}}^2, \quad \text{for all } \xi \in \Omega, f \in \boldsymbol{\mathcal{H}} \tag{31}$$

This means that the component-wise error in the $\boldsymbol{\mathcal{H}}_N$ and $\mathcal{H}_N$ can be bounded using the same *scalar* power function when $\boldsymbol{\mathcal{H}}$ is defined as the m-fold Cartesian product of $\mathcal{H}$. For vector-valued RKHSs that are not defined using the Cartesian product, expressions for the upper bound on $\left\| E_\xi \left(I - \mathbf{\Pi}_N\right) f \right\|_{\mathcal{H}}$ are presented in [31]. These more general expressions have been used in Reference [20], which does not assume that the operator-valued kernel is diagonal.

3 | MRAC for Nonparametric Matched Uncertainties

3.1 | The Limiting DPS

The next theorem is the first key result of this section, and it provides an adaptive law so that the *limiting trajectory tracking error* $e(t) \triangleq x(t) - x_r(t), t \geq t_0$, asymptotically converges to zero, despite the uncertainties in both $A \in \mathbb{R}^{n \times n}$ and $f \in \boldsymbol{\mathcal{H}}$. For the statement of this result, consider the plant model given by (1) and the reference model given by (2), and assume that the *matching conditions*

$$A_r = A + B\alpha^{\mathrm{T}}, \quad B_r = B\beta^{\mathrm{T}} \tag{32}$$

are verified by $\alpha \in \mathbb{R}^{n \times m}$ and $\beta \in \mathbb{R}^{m \times m}$. It is worthwhile to note that, in general, the structure of A is known [2, Ch. 9], and hence, although it is impossible to compute α, its existence can be verified. The matching conditions impose that there exists at least one solution that allows the plant trajectory $x(\cdot)$ to follow the reference trajectory $x_r(\cdot)$. Let $\mathbb{Z} \triangleq \mathbb{R}^n \times \mathbb{R}^{n \times n} \times \mathbb{R}^{m \times m} \times \boldsymbol{\mathcal{H}}$, and for $t \geq t_0$. For the statement of the next result, consider the *adaptive laws*

$$\dot{\hat{\alpha}}(t) = -\Gamma_\alpha x(t) e^{\mathrm{T}}(t) P B, \quad \alpha(t_0) = \alpha_0, \quad t \geq t_0 \tag{33}$$

$$\dot{\hat{\beta}}(t) = -\Gamma_\beta r(t) e^{\mathrm{T}}(t) P B, \quad \beta(t_0) = \beta_0 \tag{34}$$

$$\frac{\partial \hat{f}(t, \cdot)}{\partial t} = \Gamma_f \mathcal{K}_{x(t)} B^{\mathrm{T}} P e(t), \quad \hat{f}(t_0, \cdot) = f_0 \tag{35}$$

where $\hat{\alpha} : [t_0, \infty) \to \mathbb{R}^{n \times m}$, $\hat{\beta} : [t_0, \infty) \to \mathbb{R}^{m \times m}$, and $\hat{f} : [t_0, \infty) \times \boldsymbol{\mathcal{H}} \to \mathbb{U}$ denote the *adaptive gain matrices*, $\Gamma_\alpha \in \mathbb{R}^{n \times n}$, $\Gamma_\beta \in \mathbb{R}^{m \times m}$, and $\Gamma_f \in \mathbb{R}^{m \times m}$ denote the user-defined, symmetric, and positive-definite *adaptive rate matrices*, $P \in \mathbb{R}^{n \times n}$ denotes the symmetric, positive-definite solution of the *algebraic Lyapunov equation*

$$-Q = A_r^{\mathrm{T}} P + P A_{\mathrm{r}} \tag{36}$$

and $Q \in \mathbb{R}^{n \times n}$ is symmetric, positive-definite, and user-defined. At this point, we note that the adaptive law (35) is a partial differential equation (PDE), not an ODE.

Applying the *limiting control input*

$$u(t) = \hat{\alpha}^{\mathrm{T}}(t) x(t) + \hat{\beta}^{\mathrm{T}}(t) r(t) - E_{x(t)} \hat{f}(t, \cdot), \quad t \geq t_0 \tag{37}$$

the trajectory tracking error dynamics are given by

$$\dot{e}(t) = A_r e(t) - B\left(\tilde{\alpha}^{\mathrm{T}}(t) x(t) + \tilde{\beta}^{\mathrm{T}}(t) r(t) - E_{x(t)} \tilde{f}(t, \cdot)\right), \quad e(t_0) = x_0 - x_{r,0} \tag{38}$$

where $\tilde{\alpha}(t) \triangleq \alpha - \hat{\alpha}(t)$, $\tilde{\beta}(t) \triangleq \beta - \hat{\beta}(t)$, and $\tilde{f}(t, \cdot) \triangleq f(\cdot) - \hat{f}(t, \cdot)$. Thus, the time-varying dynamical system given by the trajectory tracking error dynamics (38) and the adaptive laws (33–35) define a *limiting DPS* with a state $(e(t), \hat{\alpha}(t), \hat{\beta}(t), \hat{f}(t,\cdot))$ that evolves in $\mathbb{Z}$. Note that $\left(e, \hat{\alpha}, \hat{\beta}, E_x\hat{f}\right) = \left(0, \alpha, \beta, E_x f\right)$ is an equilibrium point of the nonlinear dynamical system given by (38) and (33–35). The next result proves the boundedness and asymptotic convergence of solutions of this system. For clarity of exposition, the proof of the existence of complete solutions of (38) and (33–35) is deferred to Section 3.2.

Theorem 1. *Consider the nonlinear plant dynamics (1), the reference model (2), the control input (37), and the adaptive laws (33–35). Assume that the matching conditions in (32) are verified, let $\Omega = \mathbb{X}$, and assume that the kernel $\mathfrak{K}(\cdot,\cdot)$ that defines $\mathcal{H}$ is such that (16) is verified for all $\xi \in \mathbb{X}$. If the system given by (38) and (33–35) has a unique solution over the set $[t_0, \infty) \subseteq [0, \infty)$ for every initial condition, then the trajectories of (38) and (33–35) are uniformly bounded and the tracking error is convergent,*

$$\|x(t) - x_r(t)\|_{\mathbb{X}} \to 0 \quad \text{as } t \to \infty \tag{39}$$

uniformly in $t_0 \geq 0$.

Proof. Consider the Lyapunov function candidate

$$V(e, \tilde{\alpha}, \tilde{\beta}, \tilde{f}) = \langle Pe, e\rangle_{\mathbb{X}} + \left\langle \Gamma_\alpha^{-1}\tilde{\alpha}, \tilde{\alpha}\right\rangle_{\mathrm{tr}} + \left\langle \Gamma_\beta^{-1}\tilde{\beta}, \tilde{\beta}\right\rangle_{\mathrm{tr}} + \left\langle \Gamma_f^{-1}\tilde{f}(\cdot,\cdot), \tilde{f}(\cdot,\cdot)\right\rangle_{\mathcal{H}}, \quad (e, \tilde{\alpha}, \tilde{\beta}, \tilde{f}) \in \mathbb{Z} \tag{40}$$

and, assuming that $\hat{f}(\cdot) \in C^1(\mathbb{R}^+, \mathbb{X})$, along the trajectories of (38) and (33–35) it holds that

$$\begin{aligned}
&\dot{V}\left(e(t), \tilde{\alpha}(t), \tilde{\beta}(t), \tilde{f}(t,\cdot)\right) \\
&= \langle P\dot{e}(t), e(t)\rangle_{\mathbb{X}} + \langle Pe(t), \dot{e}(t)\rangle_{\mathbb{X}} \\
&\quad + 2\Big(\left\langle \Gamma_\alpha^{-1}\tilde{\alpha}^{\mathrm{T}}(t), \dot{\hat{\alpha}}(t)\right\rangle_{\mathrm{tr}} + \left\langle \Gamma_\beta^{-1}\tilde{\beta}(t), \dot{\hat{\beta}}(t)\right\rangle_{\mathrm{tr}} \\
&\quad + \left\langle \Gamma_f^{-1}\tilde{f}(t,\cdot), \dot{\hat{f}}(t,\cdot)\right\rangle_{\mathcal{H}}\Big), \\
&= -e^{\mathrm{T}}(t)Qe(t) \\
&\quad + 2\left[-\left\langle \tilde{\alpha}(t), x(t)e^{\mathrm{T}}(t)PB\right\rangle_{\mathrm{tr}} - \left\langle \tilde{\alpha}(t), \Gamma_\alpha^{-1}\dot{\hat{\alpha}}(t)\right\rangle_{\mathrm{tr}}\right] \\
&\quad + 2\left[-\left\langle \tilde{\beta}(t), r(t)e^{\mathrm{T}}(t)PB\right\rangle_{\mathrm{tr}} - \left\langle \tilde{\beta}(t), \Gamma_\beta^{-1}\dot{\hat{\beta}}(t)\right\rangle_{\mathrm{tr}}\right] \\
&\quad + 2\left[\left\langle \mathcal{K}_{x(t)}B^{\mathrm{T}}Pe(t), \tilde{f}(t,\cdot)\right\rangle_{\mathcal{H}} - \left\langle \tilde{f}(t,\cdot), \Gamma_f^{-1}\dot{\hat{f}}(t,\cdot)\right\rangle_{\mathcal{H}}\right]
\end{aligned}$$

where we used the property whereby $a^{\mathrm{T}}b = \mathrm{tr}\left(ba^{\mathrm{T}}\right)$ for any $a, b \in \mathbb{X}$. Thus, utilizing the adaptive laws (33–35), we deduce that

$$\dot{V}(e(t), \tilde{\alpha}(t), \tilde{\beta}(t), \tilde{f}(t)) = -e^{\mathrm{T}}(t)Qe(t) \leq 0, \quad t \geq t_0 \tag{41}$$

along the trajectories of (38) and (33–35).

Since $\dot{V}(e(t), \tilde{\alpha}(t), \tilde{\beta}(t), \tilde{f}(t)) \leq 0$ for all $t \geq t_0$, $V(e(t), \tilde{\alpha}(t), \tilde{\beta}(t), \tilde{f}(t))$ is non-increasing, and since $V(\cdot,\cdot,\cdot,\cdot)$ is positive-definite, it follows from the monotone convergence theorem that $\lim_{t\to\infty} V\left(e(t), \tilde{\alpha}(t), \tilde{\beta}(t), \tilde{f}(t)\right) = V_\infty$, where

$$V_\infty \leq V(e(t_0), \tilde{\alpha}(t_0), \tilde{\beta}(t_0), \tilde{f}(t_0)) \tag{42}$$

The boundedness of $V(e(t), \tilde{\alpha}(t), \tilde{\beta}(t), \tilde{f}(t))$ for all $t \geq t_0$ also implies that the states of (38) and (33–35) are bounded uniformly in $t_0 \geq 0$, that is, $e(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathbb{R}^n\right)$, $\tilde{\alpha}(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathbb{R}^{n\times n}\right)$, $\tilde{\beta}(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathbb{R}^{m\times m}\right)$, and $\tilde{f}(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathcal{H}\right)$. By integrating both sides of (42), we also conclude that $e(\cdot) \in L^2\left(\mathbb{R}^+, \mathbb{R}^n\right)$.

Next, to conclude the proof, we prove that $\dot{e}(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathbb{R}^n\right)$. It follows from (38) that

$$\begin{aligned}
\|\dot{e}(t)\|_{\mathbb{X}} \leq{}& \|A_r\|\|e\|_{L^\infty(\mathbb{R}^+,\mathbb{X})} + \|\tilde{\beta}\|_{L^\infty(\mathbb{R}^+,\mathbb{R}^{m\times m})}\|r\|_{L^\infty(\mathbb{R}^+,\mathbb{U})} \\
&+ \|B\|\|e\|_{L^\infty(\mathbb{R}^+,\mathbb{X})}\|\tilde{\alpha}\|_{L^\infty(\mathbb{R}^+,\mathbb{R}^{n\times n})} \\
&+ \|B\|\|E_{x(t)}\|_{\mathcal{L}(\mathcal{H},\mathbb{U})}\|\tilde{f}(t,\cdot)\|_{\mathcal{H}}
\end{aligned}$$

for all $t \geq t_0$. Now, $E_{x(t)} \in \mathcal{L}(\mathcal{H}, \mathbb{U})$, and under the assumption that the kernel $\mathfrak{K}$ is bounded on the diagonal, we have the uniform bound on the operator norm $\|E_{x(t)}\|_{\mathcal{L}(\mathcal{H},\mathbb{U})} \leq \overline{\mathfrak{K}}$. Furthermore, since $\|\tilde{f}(t,\cdot)\|_{\mathcal{H}} \leq \|\tilde{f}(t,\cdot)\|_{L^\infty(\mathbb{R}^+,\mathcal{H})}$ for all $t \geq t_0$, we have proven that $\dot{e}(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathbb{X}\right)$. Since $e(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathbb{X}\right) \cap L^2\left(\mathbb{R}^+, \mathbb{X}\right)$ and $\dot{e}(\cdot) \in L^\infty\left(\mathbb{R}^+, \mathbb{X}\right)$, it follows from Barbalat's lemma that (39) is verified uniformly in $t_0 \geq 0$. ◻

Theorem 1 proves that applying the limiting control input (37) and the adaptive laws (33–35), the closed-loop plant trajectory is bounded at all times and asymptotically converges to the reference model's trajectory despite parametric uncertainties in α, β and *nonparametric uncertainty in f*, uniformly in the initial time. Although both (33) and (34) closely resemble those of classical MRAC [2, Ch. 9], (35) is unique in its kind. Indeed, to the authors' knowledge, there does not exist an MRAC system that accounts for *nonparametric* matched uncertainties, such as $E_{x(\cdot)}f$ that appear in the limiting DPS, that are known to belong to some functional space, but whose representation is unknown. Thus, the framework introduced by Theorem 1 breaks from a conventional approach, whereby the matched uncertainty needs to be expressed in the form of the product of a regressor vector, which is either provided with a plant model or to be identified through neural networks or machine learning techniques [2, 21], and an unknown matrix.

The proof of the classical Lyapunov's direct method (see the proof of Theorem 4.1 in Reference [28] for example) shows the invariance of the level sets of the Lyapunov function. The same arguments apply to the proof of Theorem 1. In principle, these arguments can be used to deduce the invariance of the set Ω. In practice, however, it is difficult to correlate the level sets of the Lyapunov function (40) to the set Ω, where the closed-loop trajectory $x(\cdot)$ is expected to lie. For this reason, the set Ω should be chosen sufficiently large and according to some insights into the plant dynamics. Future research directions require leveraging Theorem 1 and the results in References [37–39] on MRAC systems and barrier Lyapunov functions to impose explicitly that the closed-loop plant trajectory lies in a user-defined set Ω.

Theorem 1 relies on the assumption whereby the nonlinear dynamical system given by the limiting DPS (38) and (33–35) is complete, that is, has a unique solution over the set $[t_0, \infty) \subseteq [0, \infty)$ for every initial condition. In Section 3.2, we provide sufficient conditions for the existence of complete solutions of (38) and (33–35).

Furthermore, the limiting control input (37) is not implementable as-is. Indeed, $\hat{f}(t,\cdot)$, $t \geq t_0$, is the solution of a PDE and resides in an infinite dimensional space. Indeed, for this reason, we refer to (37) as the *limiting* feedback control input. Finite-dimensional approximations of $E_{x(t)}\hat{f}(t,\cdot)$, and, hence, of the limiting control input are discussed in Section 4. Finally, the classical MRAC the guarantees Lyapunov stability of the equilibrium point of (38) and (33–35), which is a stronger condition than the boundedness of the solutions, uniformly in the initial time. The results discussed in Section 4 also guarantee this result.

3.2 | Forward Completeness of the Limiting DPS

Theorem 1 assumes the existence of a complete solution of (38) and (33–35), that is, of solutions over the interval $[t_0,\infty)$ for any initial condition. The next theorem provides sufficient conditions to guarantee the existence of solutions. For the statement of this result, note that the dynamical system given by (38) and (33–35) can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}z(t) = \mathcal{A}z(t) + \mathcal{G}(t,z(t)), \quad z(t_0)=z_0, \quad t\geq t_0 \tag{43}$$

where $z(t)\triangleq(e(t),\tilde{\alpha}(t),\tilde{\beta}(t),\tilde{f}(t))\in\mathbb{Z}$, the operator matrix $\mathcal{A}\triangleq\operatorname{diag}\{A_r,0,0,0\}$, and

$$\mathcal{G}(t,z)\triangleq\begin{bmatrix}-B\tilde{\alpha}^{\mathrm{T}}(x_r(t)+e)-B\tilde{\beta}+BE_{x_r(t)+e}\tilde{f}\\ \Gamma_\alpha(x_r(t)+e)e^{\mathrm{T}}PB\\ \Gamma_\beta r(t)e^{\mathrm{T}}PB\\ -\Gamma_f\mathcal{K}_{x_r(t)+e}B^{\mathrm{T}}Pe\end{bmatrix}$$

Theorem 2. *Consider the closed-loop trajectory tracking error dynamics given by (43), and assume that the reference command input $t\mapsto r(t)$ is continuous and uniformly bounded. Furthermore, assume that*

1. *the operator kernel $\mathcal{K}(\cdot,\cdot)$ is bounded on the diagonal in the sense that there is a constant $\overline{\mathfrak{K}}>0$ such that*

$$\|\mathcal{K}(x,x)\|_{\mathcal{L}(\mathbb{U})}\leq\overline{\mathfrak{K}}, \quad \text{for all } x\in\mathbb{X} \tag{44}$$

2. *the mapping $(x,y)\mapsto\mathcal{K}(x,y)$ is Lipschitz continuous on $\mathbb{X}$, that is, there is a constant $L>0$ such that $\|\mathcal{K}(x,y)-\mathcal{K}(p,q)\|_{\mathcal{L}(\mathbb{U})}\leq L(\|x-p\|+\|y-q\|)$ for all $x,y,p,q\in\mathbb{X}$;*
3. *the operator kernel $\mathcal{K}$ induces a native space $\boldsymbol{\mathcal{H}}$ that is continuously embedded in the Lipschitz continuous functions*

$$\boldsymbol{\mathcal{H}}\hookrightarrow\left(C^{0,1}(\mathbb{X})\right)^m \tag{45}$$

Then, for any initial condition $z_0\in\mathbb{Z}$, (43) has a unique solution $t\mapsto z(t)$, with $z(\cdot)\in C^1\left([t_0,\infty),\mathbb{Z}\right)$. Furthermore, the map $z_0\mapsto z$ is Lipshitz continuous from $\mathbb{Z}$ to $C([t_0,\infty),\mathbb{Z})$.

Before proceeding with the proof of Theorem 2, it is worthwhile making the following remark.

Remark 1. Assumptions 1–3 of Theorem 2 are not particularly restrictive for the class of problems discussed in this article. Indeed, since the vector-valued space $\boldsymbol{\mathcal{H}}$ is assumed to be the Cartesian product of a scalar-valued space, that is, $\boldsymbol{\mathcal{H}}\triangleq\mathcal{H}^m$, these conditions hold for many types of common kernels $\mathfrak{K}(\cdot,\cdot)$ of $\mathcal{H}$ defined over $\mathbb{X}$. For instance, if $\boldsymbol{\mathcal{H}}\triangleq\mathcal{H}^m$ and $\mathcal{H}$ is generated by the exponential, Sobolev–Matérn, inverse multiquadric, or Wendland kernels, then Assumption 1 is verified. Equation (44) follows from (23) and (16). If $\boldsymbol{\mathcal{H}}\triangleq\mathcal{H}^m$, then Assumption 2 can be verified, for example, if the scalar-valued kernel $\mathfrak{K}(\cdot,\cdot)$ is contained in the space of Lipschitz continuous functions $C^{0,1}(\mathbb{X})$, and the exponential, inverse multiquadric, sufficiently smooth Sobolev–Matérn, and sufficiently smooth Wendland kernels can be chosen to satisfy this property. Finally, Assumption 3 also holds for some common selections of the scalar-valued kernels that define $\mathcal{H}$, which in turn defines $\boldsymbol{\mathcal{H}}$. For example, the collection of Sobolev–Matérn kernels and Wendland kernels define native spaces $\mathcal{H}$ that are types of Sobolev spaces. These families can define Sobolev spaces of arbitrarily high order [30], and Sobolev spaces of sufficiently high order are continuously embedded in the spaces of Lipschitz continuous functions [40].

We now turn to the proof of Theorem 2.

Proof. This result follows from Theorem 1.4 of Chapter 6 of Reference [41]. To apply this result, we require that $e^{\mathcal{A}t}\triangleq\sum_{n=0}^{\infty}\frac{(\mathcal{A}t)^n}{n!}$ is a C_0-semigroup [41], Def. 2.1 generated by $\mathcal{A}$; for each $z\in\mathbb{Z}$, the map $t\mapsto\mathcal{G}(t,z)$ is continuous over $[t_0,\infty)$; and, for each bounded subinterval of $[t_0,\infty)$, the map $z\mapsto\mathcal{G}(t,z)$ is Lipschitz continuous in bounded neighborhoods of $z=0$, uniformly in time.

Because $\mathcal{A}$ is a bounded linear operator on $\mathbb{Z}$, it follows from Theorem 1.2 of Reference [41] that the exponential $e^{\mathcal{A}t}\triangleq\sum_{n=0}^{\infty}\frac{(\mathcal{A}t)^n}{n!}$, $t\in[t_0,\infty)$, is a C_0-semigroup generated by $\mathcal{A}$. Thus, the first condition to apply Theorem 6.1.4 of Reference [41] is verified.

Next, let $\mathcal{G}=\left[\mathcal{G}_1(t,z),\mathcal{G}_2(t,z),\mathcal{G}_3(t,z),\mathcal{G}_4(t,z)\right]^{\mathrm{T}}$, $(t,z)\in[t_0,\infty)\times\mathbb{Z}$, where $\mathcal{G}_i(t,z)$ denotes the ith component of $\mathcal{G}(t,z)$, $i=1,\ldots,4$. For each $z\in\mathbb{Z}$, the continuity in time of $\mathcal{G}_2(t,z)$ and $\mathcal{G}_3(t,z)$ is a direct consequence of the continuity of $x_r(\cdot)$ and $r(\cdot)$ on $[t_0,\infty)$. Furthermore, since $\tilde{f}\in\boldsymbol{\mathcal{H}}\in(C^{1,0}(\mathbb{X}))^m\subseteq(C_b(\mathbb{X}))^m$, the composite term $E_{x_r(t)+e}$ is continuous for all $t\in[t_0,\infty)$, which, along with the continuity of $r(\cdot)$ and $x_r(\cdot)$, implies that $\mathcal{G}_1(t,z)$ is continuous in time. Finally, it holds that

$$\begin{aligned}\left\|\mathcal{G}_4(t,z)-\mathcal{G}_4(\tau,z)\right\|_{\boldsymbol{\mathcal{H}}}\leq\|\mathcal{K}_{x_r(t)+e}-\mathcal{K}_{x_r(\tau)+e}\|_{\mathcal{L}(\mathbb{U},\boldsymbol{\mathcal{H}})}\left\|B^{\mathrm{T}}Pe\right\|_{\mathbb{U}}\\ \text{for all } (t,\tau)\in[t_0,\infty)\times[t_0,\infty)\end{aligned} \tag{46}$$

the continuity of the map $(x,z)\to\mathcal{K}(x,z)\in\mathcal{L}(\mathbb{U})$ implies the continuity of the map $x\mapsto\mathcal{K}_x\in\mathcal{L}(\mathbb{U},\boldsymbol{\mathcal{H}})$, and, since $x_r(\cdot)$ is continuous, it follows that $\mathcal{G}_4(t,z)$ is continuous in t for each fixed z. Thus, the second condition to apply Theorem 6.1.4 of Reference [41] is verified.

Finally, we prove that, for all t in bounded subintervals of $[t_0,\infty)$, the map $z\mapsto\mathcal{G}(t,z)$ is Lipschitz continuous in bounded neighborhoods of $z=0$, uniformly in time. For each bounded subinterval of $[t_0,\infty)$, the map $z\mapsto\mathcal{G}(t,z)$ is Lipschitz continuous in bounded neighborhoods of $z=0$, if, for each $t>t_0$ and for each open ball $B_R(0)\subset\mathbb{Z}$ of radius $R>0$ centered at $z=0$, there exists a constant $L_{t,R}>0$ such that

$$\|\mathcal{G}(\tau,y)-\mathcal{G}(\tau,z)\|\leq L_{t,R}\|y-z\| \tag{47}$$

for all $\tau \in [t_0, t]$ and $y, z \in B_R(0)$. To prove that (47) is verified, let $y = (\tilde{y}, \tilde{c}, \tilde{d}, \tilde{g}) \in B_R(0) \subset \mathbb{Z}$ and $z = (e, \tilde{\alpha}, \tilde{\beta}, \tilde{f}) \in B_R(0) \subset \mathbb{Z}$. Then, for each $t \in [t_0, \infty)$, $\mathcal{G}_3(t, \cdot)$ is Lipschitz continuous since

$$\begin{aligned}\|\mathcal{G}_3(\tau, y) - \mathcal{G}_3(\tau, z)\|_{\mathbb{R}^{m\times m}} &= \|\Gamma_\beta r(\tau)(e - \tilde{y})^{\mathrm{T}} PB\|_{\mathbb{R}^{m\times m}} \\ &\leq \Gamma_\beta r_{\max} \|PB\| \|(e - \tilde{y})\|_{\mathbb{X}} \end{aligned} \tag{48}$$

for all $\tau \in [t_0, t)$ and for all $(y, z) \in \mathbb{Z} \times \mathbb{Z}$, where $r_{\max} \geq 0$ denotes the uniform upper bound on $\|r(\cdot)\|$. It follows from the triangle inequality and the submultiplicative property of equi-induced norms that, for each $t \geq t_0$,

$$\begin{aligned}&\|\mathcal{G}_2(\tau, y) - \mathcal{G}_2(\tau, z)\|_{\mathbb{R}^{n\times m}} \\ &\quad = \|\Gamma_\alpha\left((x_r(\tau) + \tilde{y})\tilde{y}^{\mathrm{T}} PB - (x_r(\tau) + e)e^{\mathrm{T}} PB\right)\| \\ &\quad \leq \|\Gamma_\alpha\| \|PB\| \left(x_{r,\max} \underbrace{\|e - \tilde{y}\|_{\mathbb{X}}}_{\text{Term 1}} + \underbrace{\|(\tilde{y}\tilde{y}^{\mathrm{T}} - ee^{\mathrm{T}})\|}_{\text{Term 2}} \right)\end{aligned} \tag{49}$$

for all $\tau \in [t_0, t)$ and for all $(y, z) \in \mathbb{Z} \times \mathbb{Z}$, where $\|x_r(t)\| \leq x_{r,\max}$, $t \geq t_0$, and the uniform boundedness of $x_r(\cdot)$ is guaranteed both by the Hurwitz property of A_r and the uniform boundedness of $r(\cdot)$. Term 1 in (49) is Lipschitz continuous. Term 2 in (49) is such that

$$\begin{aligned}&\|(\tilde{y}\tilde{y}^{\mathrm{T}} - ee^{\mathrm{T}})\|_{\mathbb{R}^{n\times n}} \\ &\quad \leq \|\tilde{y}\tilde{y}^{\mathrm{T}} - e\tilde{y}^{\mathrm{T}} + e\tilde{y}^{\mathrm{T}} - ee^{\mathrm{T}}\|_{\mathbb{R}^{n\times n}} \\ &\quad \leq \|(\tilde{y} - e)\tilde{y}^{\mathrm{T}}\| + \|e(\tilde{y} - e)^{\mathrm{T}}\|_{\mathbb{R}^{n\times n}} \leq 2R\|(e - \tilde{y})\|_{\mathbb{X}}\end{aligned} \tag{50}$$

since $y, z \in B_R(0) \subseteq \mathbb{Z}$. Because both Term 1 and Term 2 in (49) are Lipschitz continuous, for each bounded subinterval of $[t_0, \infty)$, $z \mapsto \mathcal{G}_2(\tau, z)$ is Lipschitz continuous in bounded neighborhoods of $z = 0$, uniformly in time. With regards to the term $\mathcal{G}_4(t, z)$, for each $t \in [t_0, \infty)$, we have that

$$\begin{aligned}&\|\mathcal{G}_4(\tau, y) - \mathcal{G}_4(\tau, z)\|_{\mathcal{H}} \\ &\quad \leq \|\Gamma_f\| \left\| \mathcal{K}_{x_r(\tau)+\tilde{y}} B^{\mathrm{T}} P\tilde{y} - \mathcal{K}_{x_r(\tau)+e} B^{\mathrm{T}} Pe \right\|_{\mathcal{H}} \\ &\quad \leq \|\Gamma_f\| \left(\bar{\mathfrak{K}} \|B^{\mathrm{T}} P\| \|e - \tilde{y}\|_{\mathbb{X}} + \|B^{\mathrm{T}} P\| \; R \; \|\mathcal{K}_{x_r(\tau)+\tilde{y}} - \mathcal{K}_{x_r(\tau)+e}\| \right)\end{aligned} \tag{51}$$

for all $\tau \in [t_0, t)$ and for all $(y, z) \in B_R(0) \times B_R(0)$. Thus, we conclude that $z \mapsto \mathcal{G}_4(t, z)$ is Lipschitz continuous since the mapping $z \mapsto \mathcal{K}_z$ is assumed Lipschitz continuous in bounded neighborhoods of $z = 0$, uniformly in time. Finally, observe that, for each $t \in [t_0, \infty)$,

$$\begin{aligned}&\|\mathcal{G}_1(\tau, y) - \mathcal{G}_1(\tau, z)\| \\ &\quad \leq \|B\| \left(\|\tilde{c}^{\mathrm{T}}(x_r(\tau) + \tilde{y}) - \tilde{\alpha}^{\mathrm{T}}(x_r(\tau) + e)\| + r_{\max} \|\tilde{\beta} - \tilde{d}\| \right. \\ &\qquad \left. + \|\tilde{f}(x_r(\tau) + e) - \tilde{g}(x_r(\tau) + \tilde{y})\| \right) \\ &\quad \leq \|B\| \left(x_{r\max} \|\tilde{\alpha} - \tilde{c}\| + \|\tilde{c}^{\mathrm{T}}\tilde{y} - \tilde{\alpha}^{\mathrm{T}} e\| + \|\tilde{f}(x_r(\tau) + e) \right. \\ &\qquad \left. - \tilde{g}(x_r(\tau) + e)\| + \|\tilde{g}(x_r(\tau) + \tilde{y}) - \tilde{g}(x_r(\tau) - e)\| \right)\end{aligned} \tag{52}$$

for all $\tau \in [t_0, t)$ and for all $(y, z) \in B_R(0) \times B_R(0)$. Now, similarly to (50), it holds that

$$\|\tilde{c}^{\mathrm{T}}\tilde{y} - \tilde{\alpha}^{\mathrm{T}} e\| \leq 2R\|e - \tilde{y}\|_{\mathbb{X}}$$

since $y, z \in B_R(0) \subseteq \mathbb{Z}$. Furthermore, for each $t \in [t_0, \infty)$, it holds that

$$\|\tilde{f}(x_r(\tau) + e) - \tilde{g}(x_r(\tau) + e)\| \leq \|\tilde{f} - \tilde{g}\|_{C_b(\mathbb{X})} \leq \hat{C}\|\tilde{f} - \tilde{g}\|_{\mathcal{H}} \tag{53}$$

$$\begin{aligned}\|\tilde{g}(x_r(\tau) + \tilde{y}) - \tilde{g}(x_r(\tau) - e)\| &\leq \|\tilde{g}\|_{(C^{0,1}(\mathbb{X}))^m} \|e - \tilde{y}\|_{\mathbb{X}} \\ &\leq \hat{D}\|\tilde{g}\|_{\mathcal{H}} \|e - \tilde{y}\|_{\mathbb{X}} \leq \hat{D}R\|e - \tilde{y}\|_{\mathbb{X}}\end{aligned} \tag{54}$$

for all $\tau \in [t_0, t)$ and for all $(y, z) \in B_R(0) \times B_R(0)$. In both (53) and (54), we have used the fact that $\mathcal{H}$ is continuously embedded in $(C_b(\mathbb{X}))^m$ and $(C^{0,1}(\mathbb{X}))^m$, respectively, that is, $\mathcal{H} \hookrightarrow (C_b(\mathbb{X}))^m$ and $\mathcal{H} \hookrightarrow (C^{0,1}(\mathbb{X}))^m$. The former embedding holds since the operator kernel is uniformly bounded on the diagonal, and the latter holds by assumption. These embeddings mean that there are positive constants $\hat{C}$ and $\hat{D}$ such that

$$\begin{aligned}\|q\|_{(C_b(\mathbb{X}))^m} &\leq \hat{C}\|q\|_{\mathcal{H}}, \\ \|q\|_{(C^{0,1}(\mathbb{X}))^m} &\leq \hat{D}\|q\|_{\mathcal{H}}\end{aligned}$$

for all $q \in \mathcal{H}$. Thus, it follows from (52–54) that the map $z \mapsto \mathcal{G}_1(t, z)$ is Lipschitz continuous in bounded neighborhoods of $z = 0$, uniformly in time, and the third condition to apply Theorem 6.1.4 of Reference [41] is verified.

It follows from Theorem 1.4 in Chapter 6 of Reference [41] that there exists $t_{\max} \in [t_0, \infty]$ such that $t \mapsto z(t)$ exists and is unique for all $t \in [t_0, t_{\max})$, and if $t_{\max}$ is finite, then $\lim_{t\to t_{\max}} \|z(t)\| = \infty$. However, as shown in the proof of Theorem 1, $z(t)$ is uniformly bounded for all $t \geq t_0$, and, hence, $t_{\max} = \infty$. Thus, solutions of (43), or, equivalently, (38) are complete.

Using standard arguments, such as those in [28], Th. 3.4, the continuous dependence of solutions on initial conditions follows from Gronwall's inequality and the Lipschitz conditions on $\mathcal{G}(\cdot, \cdot)$, and, hence, is omitted for brevity. □

Remark 2. Assumptions 1–3 of Theorem 2 have been stated in terms of some general properties of the operator kernel $\mathcal{K}(\cdot, \cdot)$ that defines $\mathcal{H}$. In this article, the space $\mathcal{H}$ is given as the Cartesian product of scalar-valued vector spaces. However, by proceeding as in the proof of Theorem 2, the same conclusions can be attained by considering alternative (non-diagonal) admissible operator kernels $\mathcal{K}(\cdot, \cdot)$ as defined, for example, in Chapter 6 of Reference [32]. Thus, Theorem 2 holds for a larger class of operator-valued kernels $\mathcal{K}(\cdot, \cdot)$ that are not diagonal. For instance, the abovementioned theorem applies as stated to the results in Reference [20], where more general operator-valued kernels are treated, ones that are not necessarily diagonal.

4 | Finite-Dimensional Approximations of the Limiting DPS

As discussed in Section 3.1, the control input (37) is not implementable as-is because the term $\hat{f}(t, \cdot)$, $t \geq t_0$, resides in an infinite-dimensional space. In this section, we present finite-dimensional approximations of $\hat{f}(t, \cdot)$, $t \geq t_0$, and, hence, of the limiting control input. To this goal, we introduce the finite-dimensional subspace

$$\boldsymbol{\mathcal{H}}_N \triangleq (\mathcal{H}_N)^m = \mathrm{span}\left\{\mathcal{K}(\cdot,\xi_i)\alpha : \xi_i \in \Xi_N, \alpha \in \mathbb{U}\right\} \tag{55}$$

and define the *approximation of* $\hat{f}(t,\cdot)$ as

$$\hat{f}_N(t,\cdot) \triangleq \sum_{i=1}^{N} \mathcal{K}(\cdot,\xi_i)\hat{\Theta}_{N,i}(t), \qquad t \geq t_0 \tag{56}$$

where $\hat{\Theta}_{N,i}(t) \in \mathbb{U}$. In this case, the limiting control input (37) can be approximated by the finite-dimensional controller

$$u_N(t) = \hat{\alpha}_N^{\mathrm{T}}(t)x_N(t) + \hat{\beta}_N^{\mathrm{T}}(t)r(t) - E_{x_N(t)}\hat{f}_N(t,\cdot) \tag{57}$$

where $\hat{f}_N(t,\cdot) \in \boldsymbol{\mathcal{H}}_N$.

With the finite-dimensional subspace defined, we can use its corresponding projection operator $\boldsymbol{\Pi}_N$ to express the unknown vector nonlinearity as the sum of two orthogonal components, namely

$$E_{x(t)}f = E_{x(t)}\left(\boldsymbol{\Pi}_N f + \left(I - \boldsymbol{\Pi}_N\right)f\right), \qquad t \geq t_0 \tag{58}$$

where $\boldsymbol{\Pi}_N f$ denotes the *orthogonal projection* of an arbitrary function f residing in the infinite-dimensional space $\boldsymbol{\mathcal{H}}$ onto the finite-dimensional space of approximants. Furthermore, using (57), the plant dynamics (1) reduce to

$$\begin{aligned}\dot{x}_N(t) &= Ax_N(t) + B\left(u_N(t) + E_{x_N(t)}\boldsymbol{\Pi}_N f\right) + BE_{x_N(t)}\left(I - \boldsymbol{\Pi}_N\right)f \\ x_N(t_0) &= x_0, \quad t \geq t_0\end{aligned} \tag{59}$$

Thus, the trajectory tracking error dynamics are now defined as $e_N(t) \triangleq x_N(t) - x_r(t), t \geq t_0$, and, letting $\tilde{f}_N(t,\cdot) \triangleq f - \hat{f}_N(t,\cdot)$, the trajectory tracking error dynamics for the initial condition $e_N(t_0) = x_0 - x_{r,0}$ are given by

$$\begin{aligned}\dot{e}_N(t) &= A_r e_N(t) - B\Big(\tilde{\alpha}_N^{\mathrm{T}}(t)x_N(t) + \tilde{\beta}_N^{\mathrm{T}}(t)r(t) \\ &\quad - E_{x_N(t)}\boldsymbol{\Pi}_N\tilde{f}_N(t,\cdot)\Big) + BE_{x_N(t)}\underbrace{\left((I - \boldsymbol{\Pi}_N)\tilde{f}_N(t,\cdot)\right)}_{(I-\boldsymbol{\Pi}_N)f} \\ e_N(t_0) &= x_N(t_0) - x_r(t_0), \quad t \geq t_0\end{aligned} \tag{60}$$

Comparing (38) and (60), it appears that the trajectory tracking error dynamics obtained using a finite-dimensional approximation of the matched uncertainties result in an additional term, namely, $BE_{x_N(t)}\left((I - \boldsymbol{\Pi}_N)f\right)$, $t \geq t_0$, which can be interpreted as an unmatched, nonparametric, external disturbance. In the following, we present two adaptive laws that account for this term.

4.1 | The Dead-Zone Modification

To account for the projection error term in (60), in this section, we use an approach inspired by the dead-zone modification of classical MRAC presented in Reference [29]. For the statement of this result, let $\mathbb{Z}_N \triangleq \mathbb{R}^n \times \mathbb{R}^{n\times n} \times \mathbb{R}^{m\times m} \times \boldsymbol{\mathcal{H}}_N$, let $P \in \mathbb{R}^{n\times n}$ denote the symmetric, positive-definite solution of (36) for some symmetric, positive-definite user-defined matrix $Q \in \mathbb{R}^{n\times n}$, let $R > 0$ denote the radius of the uncertainty class C_R defined in (3), ϵ_N be upper bound on the projection error in the uncertainty class $C_{R,\epsilon,N}$ in (4), and let the subset $\Omega_0 \supseteq \bigcup_{\tau \geq t_0 > 0} x_N(\tau)$. We define two different choices for minimum dead-zone width. If we assume the trajectory is contained in Ω_0, then we use the *dead-zone width*

$$d_{0,N} \triangleq \frac{2\|PB\|}{\lambda_{\min}(Q)} \sup_{\eta\in\Omega_0} \mathcal{P}_N(\eta)\epsilon_N \tag{61}$$

and note that, as discussed in Section 1.3, $d_{0,N}$ is finite. Alternatively, if we do not make this hypothesis on Ω_0, but only assume that $f \in C_{R,\epsilon,N}$, then we define

$$d_\epsilon \triangleq \frac{2\|PB\|}{\lambda_{\min}(Q)}\overline{\mathfrak{K}}\epsilon \tag{62}$$

Furthermore, let d denote the user-defined size of the dead-zone such that $d \geq d_{0,N}$ or $d \geq d_\epsilon$ according to cases. We consider the *adaptive laws*

$$\dot{\hat{\alpha}}_N(t) = \begin{cases} -\Gamma_\alpha x_N(t)e_N^{\mathrm{T}}(t)PB, & \|e_N(t)\| \geq d, \\ 0, & \|e_N(t)\| < d, \end{cases}$$
$$\hat{\alpha}_N(t_0) = \hat{\alpha}_{N,0}, \quad t \geq t_0 \tag{63}$$

$$\dot{\hat{\beta}}_N(t) = \begin{cases} -\Gamma_\beta r(t)e_N^{\mathrm{T}}(t)PB, & \|e_N(t)\| \geq d, \\ 0, & \|e_N(t)\| < d, \end{cases}$$
$$\hat{\beta}_N(t_0) = \hat{\beta}_{N,0} \tag{64}$$

$$\dot{\hat{f}}_N(t) = \begin{cases} \Gamma_f \boldsymbol{\Pi}_N \mathcal{K}_{x_N(t)} B^{\mathrm{T}} P e_N(t), & \|e_N(t)\| \geq d, \\ 0, & \|e_N(t)\| < d, \end{cases}$$
$$\hat{f}_N(t_0) = \hat{f}_{N,0} \tag{65}$$

where $\hat{\alpha}_N(t_0) = \hat{\alpha}_{N,0}$, $\hat{\beta}_N(t_0) = \hat{\beta}_{N,0}$, $\hat{f}_N(t_0) = \hat{f}_{N,0}$, $\Gamma_\alpha \in \mathbb{R}^{n\times n}$, $\Gamma_\beta \in \mathbb{R}^{m\times m}$, and $\Gamma_f \in \mathbb{R}^{m\times m}$ and user-defined, symmetric, and positive-definite, and P denotes the symmetric, positive-definite solution of (36).

Theorem 3. *Consider the closed-loop trajectory tracking error dynamics (60) and the adaptive laws (63–65). Assume that the assumptions of Theorem 2 are verified. Furthermore, consider the two alternative cases*

1. $x_N(t) \in \Omega_0$ *for all* $t \geq t_0$ *and choose* $d \geq d_{0,N}$;
2. $f \in C_{R,\epsilon,N}$ *and choose* $d \geq d_\epsilon$.

Then, there exists a finite time $T \geq t_0$ *such that* $\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq d$ *for all* $t \geq T$.

Proof. Consider the Lyapunov function candidate

$$\begin{aligned} &V(e_N, \tilde{\alpha}_N, \tilde{\beta}_N, \tilde{f}_N) \\ &= \langle Pe_N(t), e_N(t)\rangle_{\mathbb{X}} + \left\langle \Gamma_\alpha^{-1}\tilde{\alpha}_N(t), \tilde{\alpha}_N(t)\right\rangle_{\mathrm{tr}} \\ &\quad + \left\langle \Gamma_\beta^{-1}\tilde{\beta}_N(t), \tilde{\beta}_N(t)\right\rangle_{\mathrm{tr}} + \left\langle \Gamma_f^{-1}\tilde{f}_N(t,\cdot), \tilde{f}_N(t,\cdot)\right\rangle_{\boldsymbol{\mathcal{H}}} \\ &\quad (e_N, \tilde{\alpha}_N, \tilde{\beta}_N, \tilde{f}_N) \in \mathbb{Z}_N \end{aligned} \tag{66}$$

First, let us assume that $\Omega_0 \supseteq \bigcup_{t\geq t_0} x_N(t)$ and $d > d_{0,N}$. In this case, if $\|e_N(t)\|_{\mathbb{X}} \geq d \geq d_{0,N}$ for some $t \geq t_0$, then, along the trajectories of (60) and (63–65), it holds that

$$\begin{aligned}
&\dot{V}(t) \\
&= -2\langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}} - 2\left\langle \tilde{\alpha}_N(t), x_N(t)e_N^T(t)PB + \Gamma_\alpha^{-1}\dot{\hat{\alpha}}_N(t)\right\rangle_{\mathrm{tr}} \\
&\quad - 2\left\langle \tilde{\beta}_N(t), r(t)e_N^T(t)PB + \Gamma_\beta^{-1}\dot{\hat{\beta}}_N(t)\right\rangle_{\mathrm{tr}} \\
&\quad + 2\left\langle \tilde{f}_N(t,\cdot), \mathbf{\Pi}_N\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t) - \Gamma_f^{-1}\dot{\hat{f}}_N(t,\cdot)\right\rangle_{\mathcal{H}} \\
&\quad + 2\left\langle \tilde{f}_N(t,\cdot), \left(I - \mathbf{\Pi}_N\right)\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t)\right\rangle_{\mathcal{H}} \\
&= -\langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}} + 2\left\langle B^{\mathrm{T}}Pe_N(t), E_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f\right\rangle_{\mathbb{U}} \\
&\leq -\langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}} + 2\|PB\|\|e_N(t)\|_{\mathbb{X}}\|E_{x_N(t)}(I - \mathbf{\Pi}_N)f\|_{\mathbb{U}}
\end{aligned} \tag{67}$$

$$\begin{aligned}
&\leq -\lambda_{\min}(Q)\|e_N(t)\|_{\mathbb{X}}\left(\|e_N(t)\|_{\mathbb{X}} - d_{0,N}\right) \\
&\leq -\lambda_{\min}(Q)\|e_N(t)\|_{\mathbb{X}}\left(\|e_N(t)\|_{\mathbb{X}} - d\right)
\end{aligned} \tag{68}$$

where $\dot{V}(\cdot)$ represents $\dot{V}(e_N(\cdot), \tilde{\alpha}_N(\cdot), \tilde{\beta}_N(\cdot), \tilde{f}_N(\cdot))$ for notational brevity. Thus, if $e_N(t)$ is outside the dead-zone $B_d \subset \mathbb{X}$ for some $t \geq t_0$, then the time derivative of (66) along the trajectories of (60) and (63–65) is negative. Therefore, following standard analysis in real parametric adaptive control [2, 21], there exists a finite-time $T > t_0$ such that $\|e_N(t)\|_{\mathbb{X}} \leq d$ for all $t \geq T$.

Next, let us assume that $f \in C_{R,\epsilon,N}$ and $d > d_\epsilon$. In this case, if $\|e_N(t)\|_{\mathbb{X}} \geq d \geq d_\epsilon$, then, an upper bound on the time derivative of the Lyapunov function candidate is given by (67). Since we know that $f \in C_{R,\epsilon,N}$ and $\mathfrak{K}$ is uniformly bounded on the diagonal, it holds that $\|E_{x_N(t)}(I - \Pi_N)f\|_{\mathbb{U}} \leq \overline{\mathfrak{K}}\epsilon$. The remainder of the proof now follows as in the previous case. □

Similar to the classical dead-zone modification of MRAC, to apply Theorem 3, it is necessary to define a sufficiently large dead-zone, that is, choose either $d > d_{0,N}$ or $d > d_\epsilon$. In problems of practical interest, it is impossible to compute a priori a bounded set Ω_0 that contains the controlled plant trajectory $t \mapsto x_N(t)$, and, hence, to compute or even estimate $d_{0,N}$. In practice, for fixed $\mathcal{H}_N$ and N, Theorem 3 should be applied iteratively, starting with a sufficiently large value of d, measuring the bound on the trajectory tracking error after a sufficiently long time, and then decreasing d until such a bound is no longer a decreasing function of d. The numerical example discussed in Section 6 clarifies this point. In contrast, note that if we know the functional uncertainty $f \in C_{R,\epsilon,N}$, then we know ϵ and d_ϵ can be computed. We emphasize that if we choose d to be equal to $d_{0,N}$ or d_ϵ, then we obtain the ultimate bounds

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq O\left(\sup_{\eta\in\Omega} \mathcal{P}_N(\eta)\epsilon_N\right) \tag{69}$$

$$\|x_N(t) - x_r(t)\|_{\mathbb{X}} \leq O(\epsilon) \tag{70}$$

respectively. Equation (69) yields a tighter performance bound, and (70) uses information that is more likely available.

4.2 | The σ-Modification

As an alternative to the dead-zone modification discussed in Section 4.1, we present an approach based on the σ-modification of MRAC that in its classical formulation, is discussed in Reference [4]. For the statement of this result, consider the trajectory tracking error dynamics (60) and the adaptive laws

$$\dot{\hat{\alpha}}_N(t) = -\Gamma_\alpha x_N(t)e_N^{\mathrm{T}}(t)PB - \sigma\hat{\alpha}_N(t), \quad \hat{\alpha}_N(t_0) = \hat{\alpha}_{N,0}, \quad t \geq t_0 \tag{71}$$

$$\dot{\hat{\beta}}_N(t) = -\Gamma_\beta r(t)e_N^{\mathrm{T}}(t)PB - \sigma\hat{\beta}_N(t), \quad \hat{\beta}_N(t_0) = \hat{\beta}_{N,0} \tag{72}$$

$$\dot{\hat{f}}_N(t) = \Gamma_f\mathbf{\Pi}_N\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t) - \sigma\hat{f}_N(t), \quad \hat{f}_N(t_0) = \hat{f}_{N,0} \tag{73}$$

where $\Gamma_\alpha \in \mathbb{R}^{n\times n}$, $\Gamma_\beta \in \mathbb{R}^{m\times m}$, and $\Gamma_f \in \mathbb{R}^{m\times m}$ are user-defined, symmetric, and positive-definite, P denotes the symmetric, positive-definite solution of the Lyapunov equation (36) for some symmetric, positive-definite user-defined matrix $Q \in \mathbb{R}^{n\times n}$, and $0 < \sigma < \frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)}$, where $\lambda_{\max}(\cdot)$ denotes the eigenvalue of its argument with the largest real part.

Theorem 4. *Consider the closed-loop trajectory tracking error dynamics (60) and the adaptive laws (71)–(73). If the assumptions of Theorem 2 are verified, then for any bounded subset $\Omega \supseteq \bigcup_{\tau\geq t_0>0} x(\tau)$ there is $T \geq t_0$ such that*

$$\begin{aligned}
\|x_N(t) - x_r(t)\|_{\mathbb{X}}^2 &\leq \frac{2}{\sigma\lambda_{\min}(Q)}\|PB\|^2 \sup_{\eta\in\Omega} \mathcal{P}_N^2(\eta)\epsilon_N^2 \\
&\quad + \left\langle\alpha, \Gamma_\alpha^{-1}\alpha\right\rangle_{\mathrm{tr}} + \left\langle\beta, \Gamma_\beta^{-1}\beta\right\rangle_{\mathrm{tr}} + \left\langle f, \Gamma_f^{-1}f\right\rangle_{\mathcal{H}}
\end{aligned} \tag{74}$$

for all $t \geq T$, where ϵ_N is given by (8) and $\mathcal{P}_N(\cdot)$ is given by (27).

Proof. Consider the Lyapunov function candidate (66). By proceeding as in (68), along the trajectories of (60) and (71–73) it holds that

$$\begin{aligned}
&\dot{V}(t) \\
&= -\langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}} + 2\left\langle\left(I - \mathbf{\Pi}_N\right)\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t), \tilde{f}_N(t,\cdot)\right\rangle_{\mathcal{H}} \\
&\quad + 2\sigma\left\langle\tilde{\alpha}_N(t), \Gamma_\alpha^{-1}\hat{\alpha}_N\right\rangle_{\mathrm{tr}} + 2\sigma\left\langle\tilde{\beta}_N(t), \Gamma_\beta^{-1}\hat{\beta}_N(t)\right\rangle_{\mathrm{tr}} \\
&\quad + 2\sigma\left\langle\tilde{f}(t,\cdot), \Gamma_f^{-1}\hat{f}_N(t,\cdot)\right\rangle_{\mathcal{H}}, \quad t \geq t_0
\end{aligned} \tag{75}$$

where $\dot{V}(\cdot)$ represents $\dot{V}(e_N(\cdot), \tilde{\alpha}_N(\cdot), \tilde{\beta}_N(\cdot), \tilde{f}_N(\cdot))$ for notational brevity.

Next, for brevity, let $\langle\cdot,\cdot\rangle_G$ denote any weighted inner product so that, for example, $\langle\hat{\alpha}_N, \hat{\alpha}_N\rangle_G = \left\langle\hat{\alpha}_N, \Gamma_\alpha^{-1}\hat{\alpha}_N\right\rangle_{\mathrm{tr}}$. Furthermore, recalling that the inner product is bilinear and symmetric over the field of real numbers, we obtain that, for all $t \geq t_0$,

$$\begin{aligned}
\langle\tilde{\alpha}_N, \tilde{\alpha}_N\rangle_G &= \langle\alpha - \hat{\alpha}_N, \alpha\rangle_G - \langle\alpha - \hat{\alpha}_N, \hat{\alpha}_N\rangle_G \\
&= \langle\alpha, \alpha\rangle_G - 2\langle\alpha, \hat{\alpha}_N\rangle_G + \langle\hat{\alpha}_N, \hat{\alpha}_N\rangle_G
\end{aligned}$$

which yields

$$2\langle\alpha, \hat{\alpha}_N\rangle_G = \langle\alpha, \alpha\rangle_G + \langle\hat{\alpha}_N, \hat{\alpha}_N\rangle_G - \langle\tilde{\alpha}_N, \tilde{\alpha}_N\rangle_G \tag{76}$$

Subtracting $2\langle\hat{\alpha}_N, \hat{\alpha}_N\rangle_G$ from both sides of (76) yields

$$2\langle\tilde{\alpha}_N(t), \hat{\alpha}_N(t)\rangle_G = \langle\alpha, \alpha\rangle_G - \langle\tilde{\alpha}_N(t), \tilde{\alpha}_N(t)\rangle_G - \langle\hat{\alpha}_N, \hat{\alpha}_N\rangle_G \tag{77}$$

and similar properties can be inferred for $\langle \tilde{\beta}_N(t), \hat{\beta}_N(t)\rangle_G$ and $\langle \tilde{f}_N(t), \hat{f}_N(t)\rangle_G$. Therefore, it follows from (75) that, for all $t \geq t_0$,

$$\begin{aligned}
&\dot{V}(t)\\
&= -\langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}} + 2\left\langle \left(I - \mathbf{\Pi}_N\right)\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t), \tilde{f}_N(t,\cdot)\right\rangle_{\mathcal{H}}\\
&\quad - \sigma\langle \tilde{\alpha}_N(t), \tilde{\alpha}_N(t)\rangle_G + \sigma\langle \alpha, \alpha\rangle_G\\
&\qquad - \sigma\langle \tilde{\beta}_N(t), \tilde{\beta}_N(t)\rangle_G + \sigma\langle \beta, \beta\rangle_G - \sigma\langle \tilde{f}_N(t), \tilde{f}_N(t)\rangle_G\\
&\quad + \sigma\langle f, f\rangle_G - \sigma\langle \hat{\alpha}_N, \hat{\alpha}_N\rangle_G - \sigma\langle \hat{\beta}_N, \hat{\beta}_N\rangle_G - \sigma\langle \hat{f}_N, \hat{f}_N\rangle_G
\end{aligned} \tag{78}$$

As the terms $-\sigma\langle \hat{\alpha}_N, \hat{\alpha}_N\rangle_G$, $-\sigma\langle \hat{\beta}_N, \hat{\beta}_N\rangle_G$, and $-\sigma\langle \hat{f}_N, \hat{f}_N\rangle_G$ are non-positive, it follows from (78) that

$$\begin{aligned}
\dot{V}(t) \leq &- \langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}}\\
&+ 2\left\langle \left(I - \mathbf{\Pi}_N\right)\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t), \tilde{f}_N(t,\cdot)\right\rangle_{\mathcal{H}}\\
&- \sigma\langle \tilde{\alpha}_N(t), \tilde{\alpha}_N(t)\rangle_G + \sigma\langle \alpha, \alpha\rangle_G - \sigma\langle \tilde{\beta}_N(t), \tilde{\beta}_N(t)\rangle_G\\
&+ \sigma\langle \beta, \beta\rangle_G - \sigma\langle \tilde{f}_N(t), \tilde{f}_N(t)\rangle_G + \sigma\langle f, f\rangle_G
\end{aligned} \tag{79}$$

for all $t \geq t_0$. Adding and subtracting $\sigma\langle Pe_N(t), e_N(t)\rangle_{\mathbb{X}}$, $t \geq t_0$, from (79) yields

$$\begin{aligned}
\dot{V}(t) \leq &- \langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}}\\
&+ 2\left\langle \left(I - \mathbf{\Pi}_N\right)\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t), \tilde{f}_N(t,\cdot)\right\rangle_{\mathcal{H}}\\
&+ \sigma\langle Pe_N(t), e_N(t)\rangle_{\mathbb{X}} + \sigma\langle \alpha, \alpha\rangle_G + \sigma\langle \beta, \beta\rangle_G v\\
&+ \sigma\langle f, f\rangle_G - \sigma V(t)
\end{aligned} \tag{80}$$

Now, note that

$$\begin{aligned}
&2\left\langle \left(I - \mathbf{\Pi}_N\right)\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t), \tilde{f}_N(t,\cdot)\right\rangle_{\mathcal{H}}\\
&= 2\left\langle B^{\mathrm{T}}Pe_N(t), E_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)\tilde{f}_N(t,\cdot)\right\rangle_{\mathbb{U}}\\
&= 2\left\langle e_N(t), PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)(f(\cdot) - \hat{f}_N(t,\cdot))\right\rangle_{\mathbb{X}}
\end{aligned} \tag{81}$$

for all $t \geq t_0$. Furthermore, since $\hat{f}_N(t,\cdot) \in \mathcal{H}_N$, it holds that

$$E_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)\hat{f}_N(t,\cdot) = 0, \quad \text{for all } t \geq t_0 \tag{82}$$

Therefore, it follows from (81) and the Cauchy-Schwarz inequality that

$$\begin{aligned}
&- \langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}}\\
&\quad + 2\left\langle \left(I - \mathbf{\Pi}_N\right)\mathcal{K}_{x_N(t)}B^{\mathrm{T}}Pe_N(t), \tilde{f}_N(t,\cdot)\right\rangle_{\mathcal{H}}\\
&= -\langle Qe_N(t), e_N(t)\rangle_{\mathbb{X}} + 2\left\langle e_N(t), PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\right\rangle_{\mathbb{X}}\\
&\leq -\lambda_{\min}(Q)\|e_N(t)\|_{\mathbb{X}}^2 + 2\|e_N(t)\|_{\mathbb{X}}\|PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\|_{\mathbb{X}}\\
&\leq -\frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|_{\mathbb{X}}^2 - 2\lambda_{\min}(Q)\\
&\left(\frac{1}{4}\|e_N(t)\|_{\mathbb{X}}^2 - \frac{\|e_N(t)\|_{\mathbb{X}}}{\lambda_{\min}(Q)}\left\|PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\right\|_{\mathbb{X}}\right)
\end{aligned} \tag{83}$$

Now, for each $t \geq t_0$, let $v(t) \triangleq \lambda_{\min}^{-1}(Q)\left\|PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f\right\|_{\mathbb{X}}$, and note that

$$\begin{aligned}
&\frac{1}{4}\|e_N(t)\|_{\mathbb{X}}^2 - \frac{\|e_N(t)\|_{\mathbb{X}}}{\lambda_{\min}(Q)}\left\|PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\right\|_{\mathbb{X}}\\
&= \frac{1}{4}\|e_N(t)\|_{\mathbb{X}}^2 - v(t)\|e_N(t)\|_{\mathbb{X}}\\
&= \left(\frac{1}{2}\|e_N(t)\|_{\mathbb{X}} - v(t)\right)^2 - v^2(t)
\end{aligned} \tag{84}$$

Thus, substituting (84) and (83) into (80) yields

$$\begin{aligned}
\dot{V}(t) \leq &-\frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|_{\mathbb{X}}^2\\
&- 2\lambda_{\min}(Q)\left[\left(\frac{1}{2}\|e_N(t)\|_{\mathbb{X}} - v(t)\right)^2 - v^2(t)\right]\\
&+ \sigma\langle \alpha, \alpha\rangle_G + \sigma\langle \beta, \beta\rangle_G + \sigma\langle f, f\rangle_G - \sigma V(t)\\
&+ \sigma\langle Pe_N(t), e_N(t)\rangle_{\mathbb{X}}
\end{aligned} \tag{85}$$

As $-\left(\frac{1}{2}\|e_N(t)\|_{\mathbb{X}} - v(t)\right)^2 \leq 0$ for all $t \geq t_0$, we obtain that

$$\begin{aligned}
\dot{V}(t) \leq &-\frac{1}{2}\lambda_{\min}(Q)\|e_N(t)\|_{\mathbb{X}}^2\\
&+ 2\lambda_{\min}^{-1}(Q)\left\|PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\right\|_{\mathbb{X}}^2\\
&+ \sigma\langle \alpha, \alpha\rangle_G + \sigma\langle \beta, \beta\rangle_G + \sigma\langle f, f\rangle_G - \sigma V(t)\\
&+ \sigma\langle Pe_N(t), e_N(t)\rangle_{\mathbb{X}}\\
\leq &-\sigma V(t) - \left(\frac{1}{2}\lambda_{\min}(Q) - \sigma\lambda_{\max}(P)\right)\|e_N(t)\|_{\mathbb{X}}^2\\
&+ \frac{2}{\lambda_{\min}(Q)}\left\|PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\right\|_{\mathbb{X}}^2\\
&+ \sigma\left(\langle \alpha, \alpha\rangle_G + \langle \beta, \beta\rangle_G + \langle f, f\rangle_G\right)
\end{aligned}$$

By assumption, σ is such that $\frac{1}{2}\lambda_{\min}(Q) - \sigma\lambda_{\max}(P) > 0$, and, hence, we obtain that

$$\begin{aligned}
\dot{V}(t) \leq &- \sigma V(t) + \frac{2}{\lambda_{\min}(Q)}\left\|PBE_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\right\|_{\mathbb{X}}^2\\
&+ \sigma\left(\langle \alpha, \alpha\rangle_G + \langle \beta, \beta\rangle_G + \langle f, f\rangle_G\right)
\end{aligned}$$

Next, initially set $\Omega = \mathbb{X}$. It follows from (31) that $\|E_{x_N(t)}\left(I - \mathbf{\Pi}_N\right)f(\cdot)\|_{\mathbb{U}} \leq \sup_{\eta\in\Omega}\mathcal{P}_N(\eta)\epsilon_N < \infty$, and, hence,

$$\begin{aligned}
\dot{V}(t) \leq &-\sigma V(t) + \frac{2}{\lambda_{\min}(Q)}\|PB\|^2\sup_{\eta\in\Omega}\mathcal{P}_N^2(\eta)\epsilon_N^2 + \sigma\left(\langle \alpha, \alpha\rangle_G\right.\\
&\left.+\langle \beta, \beta\rangle_G + \langle f, f\rangle_G\right), \quad \text{for all } t \geq t_0
\end{aligned} \tag{86}$$

note that since $\mathfrak{K}(\cdot,\cdot)$ is uniformly bounded on the diagonal, the power function $\mathcal{P}_N$ is also uniformly bounded over $\mathbb{X}$. Now, let $D \triangleq \frac{2}{\lambda_{\min}(Q)}\|PB\|^2 \sup_{\eta\in\Omega}\mathcal{P}_N^2(\eta)\epsilon_N^2 + \sigma\left(\langle \alpha, \alpha\rangle_G + \langle \beta, \beta\rangle_G + \langle f, f\rangle_G\right)$. It follows from (86) that

$$V(t) \leq c_1 e^{-\sigma t} + \frac{D}{\sigma}, \quad \text{for all } t \geq t_0 \tag{87}$$

where $c_1 \in \mathbb{R}$ is a coefficient that depends on the initial condition of the system. Using arguments similar to those used in the proof of Theorem 3, (87) implies that there exists $T \geq t_0$ such that $(e_N(T), \tilde{\alpha}_N(T), \tilde{\beta}_N(T), \tilde{f}_N(T)) \in S_0$, where

$$\begin{aligned}
S_0 \triangleq \Big\{&\left(e_N, \tilde{\alpha}_N, \tilde{\beta}_N, \tilde{f}_N\right) \in \mathbb{R}^n \times \mathbb{R}^{m\times m} \times \mathbb{R}^{m\times m}\\
&\times\mathcal{H}_N : V\left(e_N, \tilde{\alpha}_N, \tilde{\beta}_N, \tilde{f}_N\right) \leq \frac{D}{\sigma}\Big\}
\end{aligned} \tag{88}$$

which implies (74). The proposed argument can be iterated for smaller sets $\Omega \subset \mathbb{X}$ containing the closed-loop trajectory to attain tighter bounds on the closed-loop system's uniform ultimate bounded given by (74). □

Note that, as discussed in Section 1.3, the ultimate bound (74), in contrast to the dead-zone method, is not AAO as defined in Section 1.2. However, (74) improves on the bounds in References [4, 21], because, as discussed in Section 1.2, the uncertainty class C_R over which the bound holds is generally larger than those that arise in real parametric adaptive control.

As a concluding remark, it is worthwhile noting that Theorems 3 and 4 apply to dynamical models for the tracking error affected by unmatched uncertainties, that is, captured by

$$\begin{aligned}\dot{e}_N(t) &= A_r e_N(t) - B\left(\tilde{\alpha}_N^{\mathrm{T}}(t)x_N(t) + \tilde{\beta}_N^{\mathrm{T}}(t)r(t) - E_{x_N(t)}\mathbf{\Pi}_N\tilde{f}_N(t,\cdot)\right)\\ &\quad + BE_{x_N(t)}\left((I-\mathbf{\Pi}_N)\tilde{f}_N(t,\cdot)\right) + \eta(t)\\ e_N(t_0) &= x_N(t_0) - x_r(t_0), \quad t \geq t_0 \end{aligned} \tag{89}$$

where $\eta : [t_0, \infty) \to \mathbb{R}^n$ is continuous and bounded. Indeed, the unmatched uncertainty can be seamlessly included in the projection error term $\left((I-\mathbf{\Pi}_N)\tilde{f}_N(t,\cdot)\right)$. However, it should be noted that the proposed framework, as with any MRAC system, performs worse as the norm of the unmatched uncertainties increases. If the norm of the unmatched uncertainty is expected to be considerably larger than the norm of the matched uncertainty, then alternative techniques, such as adaptive error bounding should be considered. These techniques, however, may suffer from chattering due to their switched mechanisms. For details, see Reference [23], Th. 5.10.

4.3 | Controller Performance as a Function of the Fill Distance

One of the advantages of restricting the analysis in this article to vector-valued native spaces $\boldsymbol{\mathcal{H}} \triangleq \mathcal{H}^m$ is that it is possible to derive versions of Theorems 3 and 4 that are easy to interpret geometrically. These alternative versions of Theorems 3 and 4 are based on upper bounds on the power function $\mathcal{P}_N(\cdot)$ of common scalar-valued native spaces that have the form

$$\sup_{\xi\in\Omega} \mathcal{P}_N(\xi) \leq \mathcal{F}(N) \triangleq \sqrt{\mathcal{N}(h_{\Xi_N,\Omega})} \tag{90}$$

where $\mathcal{N}$ is defined in Table 1; for additional details, see also Reference [30, Ch. 11]. A similar table with slightly different bounds can be found in Reference [36].

TABLE 1 | Examples of kernel basis functions and corresponding upper bounds [30, Ch. 11].

Kernel name	$\mathfrak{K}(x,0) = \mathfrak{K}_0(r)$, $r = \Vert x\Vert_{\mathbb{X}}$	$\mathcal{N}(h)$
Gaussian	$e^{-\alpha r^2}, \quad \alpha > 0$	$e^{-c\vert\log(h)\vert/h}$
Inverse multi-quadratic	$(c^2+r^2)^\beta, \quad \beta < 0$	$e^{-\tilde{c}/h}$
Wendland	$\phi_{d,k}(r)$	h^{2k+1}

Corollary 1. *Assume the hypotheses of Theorem 3 are verified, the scalar-valued native space $\mathcal{H}$ is defined in terms of one of the admissible kernels $\mathfrak{K}(\cdot,\cdot)$ in Table 1, and let $\mathcal{N}$ be defined as in Table 1. Then, there exists a finite time $T \geq t_0$ such that*

$$\Vert x_N(t) - x_r(t)\Vert_{\mathbb{X}} \leq O\left(\sqrt{\mathcal{N}(h_{\Xi_N,S})}\right) \tag{91}$$

for all $t \geq T$. In particular, (12) is verified.

Proof. This result follows as a direct consequence of Theorem 3. □

A similar result stems from Theorem 4.

Corollary 2. *Assume the hypotheses of Theorem 4 are verified, the scalar-valued native space $\mathcal{H}$ is defined in terms of an admissible kernel $\mathfrak{K}(\cdot,\cdot)$ chosen from Table 1, and let $\mathcal{N}$ be defined as in Table 1. Then, there exists a finite time $T \geq t_0$ such that*

$$\Vert x_N(t) - x_r(t)\Vert_{\mathbb{X}} \leq O\left(\frac{1}{\sigma}\left(\sqrt{\mathcal{N}(h_{\Xi_N,S})} + \sigma\right)\right) \tag{92}$$

for all $t \geq T$. In particular, (13) is verified.

Proof. This result is a direct consequence of Theorem 4. □

5 | Coordinate Implementation

In this section, we study in detail the coordinate realizations of the adaptive laws (65) and (73) with diagonal, positive-definite, adaptive rate matrix $\Gamma_f \triangleq \mathrm{diag}(\gamma_f^1, \ldots, \gamma_f^m)$. Given that the finite-dimensional approximation of the uncertainty is in $\boldsymbol{\mathcal{H}}_N$, which is a Cartesian product of $\mathcal{H}_N$, it is possible to write $\hat{f}_N(t,\cdot)$ in the vector form

$$\begin{aligned}\hat{f}_N(t,\cdot) &= \left[\sum_{j=1}^N \hat{\theta}_{N,j}^1(t)\mathfrak{K}_{\xi_{N,j}}, \ldots, \sum_{j=1}^N \hat{\theta}_{N,j}^m(t)\mathfrak{K}_{\xi_{N,j}}\right]^{\mathrm{T}}\\ &= \sum_{j=1}^N \mathcal{K}\left(\cdot, \xi_{N,j}\right)\hat{\Theta}_{N,j}(t)\end{aligned} \tag{93}$$

where $\hat{\Theta}_{N,j}(\cdot) = \left[\hat{\theta}_{N,j}^1(\cdot), \ldots, \hat{\theta}_{N,j}^m(\cdot)\right]^{\mathrm{T}} \in RR^m$, $\mathfrak{K}_{\xi_{N,j}}(\cdot)$ denotes the kernel basis function centered at $\xi_{N,j}$, and $\hat{\theta}_{N,j}^l$ denotes the coefficient associated with the kernel basis function in the l-th subspace of $\boldsymbol{\mathcal{H}}$. Taking the l-th component of $\hat{f}_N \triangleq \left[\hat{f}_N^1, \ldots, \hat{f}_N^m\right]^{\mathrm{T}}$, $l \in \{1, \ldots, m\}$, it follows from (73)

$$\begin{aligned}\frac{\partial \hat{f}_N^l}{\partial t}(t,\cdot) &\triangleq \sum_{j=1}^N \dot{\hat{\theta}}_{N,j}^l(t)\mathfrak{K}_{\xi_{N,j}}\\ &= \gamma_f^l \Pi_N \mathfrak{K}_{x_N(t)}[B^{\mathrm{T}}Pe_N(t)]_l - \sigma\sum_{j=1}^N \hat{\theta}_{N,j}^l \mathfrak{K}_{\xi_{N,j}}\\ &\quad \text{for all } l \in \{1, \ldots, m\}, \text{for all } t \geq t_0,\\ &\quad \text{and for each } i \in \{1, \ldots, N\}\end{aligned} \tag{94}$$

Here, $[B^{\mathrm{T}}Pe_N(\cdot)]_l$ denotes the l-th component of $B^{\mathrm{T}}Pe_N(\cdot)$. It follows from (15) that $E_{x_N(t)}^* \alpha = \mathfrak{K}_{x_N(t)}\alpha$ for any scalar $\alpha \in \mathbb{R}$.

Taking the inner product of (94) with $\mathfrak{K}_{\xi_{N,i}}$ for $1 \le i \le N$, we obtain that

$$\sum_{j=1}^{N} \langle \mathfrak{K}_{\xi_{N,i}}, \mathfrak{K}_{\xi_{N,j}} \rangle_{\mathcal{H}} \dot{\hat{\theta}}^{l}_{N,j}(t) = \gamma_f^l \langle \mathfrak{K}_{\xi_{N,i}}, \Pi_N \mathfrak{K}_{x_N(t)} \rangle_{\mathcal{H}} [B^{\mathrm{T}} P e_N(t)]_l$$
$$- \sigma \sum_{j=1}^{N} \langle \mathfrak{K}_{\xi_{N,i}}, \mathfrak{K}_{\xi_{N,j}} \rangle_{\mathcal{H}} \hat{\theta}^{l}_{N,j}(t)$$
$$\text{for all } l \in \{1, \dots, m\}, \text{for all } t \ge t_0, \quad \text{and for each } i \in \{1, \dots, N\} \tag{95}$$

and it follows from (29) that

$$\sum_{j=1}^{N} \mathbb{K}_{N,ij} \dot{\hat{\theta}}^{l}_{N,j}(t) = \gamma_f^l \mathfrak{K}\left(\xi_{N,i}, x_N(t)\right) [B^{\mathrm{T}} P e_N(t)]_l$$
$$- \sigma \sum_{j=1}^{N} \mathbb{K}_{N,ij} \hat{\theta}^{l}_{N,j}(t)$$
$$\text{for all } l \in \{1, \dots, m\}, \text{for all } t \ge t_0, \quad \text{and for each } i \in \{1, \dots, N\} \tag{96}$$

where $\mathbb{K}_{N,ij}$ denotes the element of $\mathbb{K}_N$ on the i-th row and j-th column, and, hence,

$$\dot{\hat{\theta}}^{l}_{N}(t) = \gamma_f^l \mathbb{K}_N^{-1} \begin{bmatrix} \mathfrak{K}\left(\xi_{N,1}, x_N(t)\right) \\ \vdots \\ \mathfrak{K}\left(\xi_{N,N}, x_N(t)\right) \end{bmatrix} [B^{\mathrm{T}} P e_N(t)]_l - \sigma \hat{\theta}^{l}_{N}(t)$$
$$\hat{\theta}_N(t_0) = \hat{\theta}_{N,0}, \quad t \ge t_0 \tag{97}$$

which is the adaptive law for l-th entry of the vector of functions in $\hat{f}_N$.

Remark 3. The coordinate realization (97) takes the classical form of a σ-modification in $\mathbb{R}^n$ commonly used in real parametric adaptive control, as summarized in References [1, 3, 4]. One way to interpret the results of this section is that the inverse of the Grammian matrix $\mathbb{K}_N$ scales the adaptive laws, and, hence, the diagonal matrix of adaptive rates Γ_f must be tuned accounting for this effect.

By proceeding similarly, we deduce that, if the adaptive rate matrix Γ_f is diagonal, then a coordinate implementation of (65) is given by

$$\dot{\hat{\theta}}^{l}_{N}(t) = \begin{cases} \gamma_f^l \mathbb{K}_N^{-1} \begin{bmatrix} \mathfrak{K}\left(\xi_{N,1}, x_N(t)\right) \\ \vdots \\ \mathfrak{K}\left(\xi_{N,N}, x_N(t)\right) \end{bmatrix} [B^{\mathrm{T}} P e_N(t)]_l, & \|e_N(t)\| \ge d, \\ 0, & \|e_N(t)\| < d, \end{cases}$$
$$\hat{\theta}_N(t_0) = \hat{\theta}_{N,0} \tag{98}$$

It follows from (97) and (98) that the computational cost of the proposed control nonparametric control system grows with the number of centers N since $\mathbb{K}_N \in \mathbb{R}^{N \times N}$. Thus, sampling a given domain Ω with a larger number of centers N will produce smaller tracking errors at a higher computational cost. Similarly, to counter large uncertainties on where the closed-loop trajectory will lie, larger domains Ω may need to be investigated. However, fixed the fill distance, larger values for the diameter of Ω imply larger values of N and higher computational costs.

As a final remark, it is worthwhile noting that the proposed extensions of the dead-zone modification of MRAC and σ-modification of MRAC may be considered alternatives to one another. However, as discussed in Reference [2, Ch. 11], best industrial practices recommend that the dead-zone modification of MRAC be compounded with other techniques, such as the σ-modification of MRAC, to activate the adaptation mechanism only if the tracking error is sufficiently large. Thus, (98) should be compounded with (97) to give

$$\dot{\hat{\theta}}^{l}_{N}(t) = \begin{cases} \gamma_f^l \mathbb{K}_N^{-1} \begin{bmatrix} \mathfrak{K}\left(\xi_{N,1}, x_N(t)\right) \\ \vdots \\ \mathfrak{K}\left(\xi_{N,N}, x_N(t)\right) \end{bmatrix} [B^{\mathrm{T}} P e_N(t)]_l - \sigma_l \hat{\theta}^{l}_{N}(t), & \|e_N(t)\| \ge d, \\ 0, & \|e_N(t)\| < d, \end{cases}$$
$$\hat{\theta}_N(t_0) = \hat{\theta}_{N,0} \tag{99}$$

where $\sigma_l > 0$ is user-defined. If the tracking error is sufficiently small, then a nominal baseline controller, typically a PID (proportional, integral, derivative) controller, should be used. For details on the problem of augmenting baseline controllers with adaptive controllers, see Reference [2, Ch. 10].

6 | Numerical Examples

To demonstrate the applicability of the proposed results, in this section, we present two numerical examples. The first example concerns the problem of controlling an unstable plant model. The second example concerns the problem of controlling the rotational dynamics of a rigid body. The Matlab codes showing an implementation of these examples can be found in Reference [42].

6.1 | A Two-Dimensional Dimensional Problem

Consider the unstable plant model

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \left(\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \underbrace{\frac{x_1(t)}{2\|x(t)\|} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} -x_1 \|x(t)\|^2 \\ -x_2 \|x(t)\|^2 \end{bmatrix}}_{f(x)} \right)$$
$$x(t_0) = x_0, \quad t \ge t_0 \tag{100}$$

which is in the same form as (1) with $n = 2$, $m = 2$, $t_0 = 0$, the matrix $B = I$ known, and both the matrix A and function $f(\cdot)$ unknown. The reference model is given by (2) with

$A_r = \begin{bmatrix} -2 & -0.1 \\ 0.1 & -2 \end{bmatrix}$, $B_r = \begin{bmatrix} 2 & 0.1 \\ -0.1 & 2 \end{bmatrix}$, and

$$r(t) = \begin{bmatrix} 2\cos(0.1\pi t) \\ 4\sin(0.1\pi t) \end{bmatrix}, \quad t \geq 0 \tag{101}$$

it is easy to show that the matching conditions (32) are verified.

The reproducing kernels used in this example come from three families of functions. The first family of kernels is specific types of *Sobolev–Matérn kernels* archived by Abramowitz [43], which are defined as

$$\mathfrak{K}_{p+1/2}(d) = \exp\left(-\frac{\sqrt{2p+1}d}{l}\right)\frac{p!}{(2p)!} \cdot \sum_{i=0}^{p} \frac{(p+i)!}{i!(p-i)!}\left(\frac{2\sqrt{2p+1}d}{l}\right)^{p-i} \tag{102}$$

where $d = \|x-y\|_{\mathbb{X}}$ denotes the distance between x and $y \in \mathbb{X}$, and $p \in \mathbb{N}$ denotes the power of the polynomial. The power index corresponds to the Sobolev space defined by the Sobolev–Matérn kernel via half integers. For example, if $p = 1, 2, 3$, then the $3/2$, $5/2$, and $7/2$ Sobolev–Matérn kernels are induced, respectively. Note that the larger the index, the smoother the corresponding Sobolev space. The second family of reproducing kernels is the *Wendland kernels* introduced in Reference [30]. In particular, Wendland kernels that correspond to 3 different smoothness parameters are chosen. For kernels lying in $C^2(\mathbb{R}^n)$, $C^4(\mathbb{R}^n)$, and $C^6(\mathbb{R}^N)$, the corresponding kernel functions are defined as [30]

$$\mathfrak{K}_2(r) = (1-r)_+^4(4r+1) \tag{103}$$

$$\mathfrak{K}_4(r) = \frac{1}{3}(1-r)_+^6(35r^2+18r+3) \tag{104}$$

$$\mathfrak{K}_6(r) = (1-r)_+^8(32r^3+25r^2+8r+1) \tag{105}$$

respectively, where $r = \frac{\|x-y\|_{\mathbb{X}}}{3l}$ is the radius, and $l > 0$ is user-defined. In this expression, $(\cdot)_+ : \mathbb{R} \to \mathbb{R}$ denotes the *cutoff function* with $(x)_+ \triangleq x$ for $x > 0$ and $(x)_+ \triangleq 0$ otherwise. The last type of reproducing kernel studied in this section is the *Gaussian kernel*

$$\mathfrak{K}_{\exp}(x, y) = \exp\left(-\frac{\|x-y\|_{\mathbb{X}}^2}{2l^2}\right), \quad (x, y) \in \mathbb{X} \times \mathbb{X} \tag{106}$$

this is the smoothest of all the kernels used in this example. The parameter l in (103–106) is set to a value of $l = 1$. The native space generated by the Gaussian kernel is contained in $C^\infty(\Omega)$. The Sobolev–Matérn kernel having smoothness index $m > n/2$ generates a native space that is equivalent to the Sobolev space W^m for $m > n/2 = 1$.

By Corollaries 1 and 2, we expect that the tracking error depends on the fill distance of the centers $\Xi_N \subset \Omega$. The numerical conditioning of Grammian matrices depends strongly on the uniformity of the distribution of centers Ξ_N in Ω; see Reference [30]. For this reason, kernel centers are distributed "approximately uniformly" over a set Ω that contains the reference trajectory.

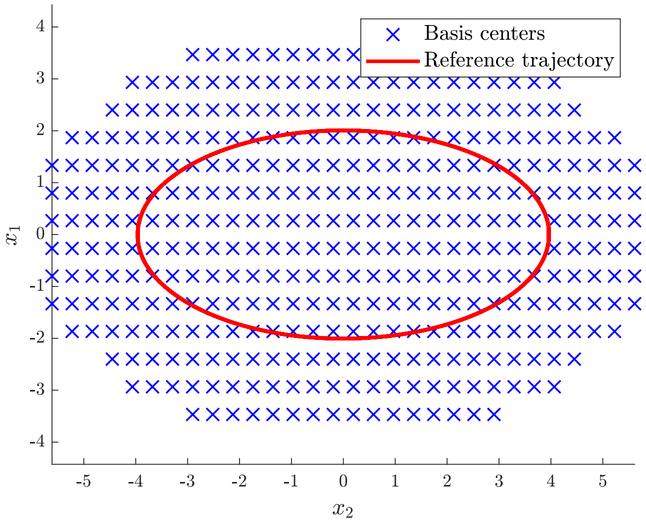

FIGURE 1 | Limit cycle of the reference trajectory $x_r(\cdot)$ and kernel centers. around the reference trajectory. Knowing that the reference trajectory eventually reaches a limit cycle allowed us to distribute centers in a sufficiently large neighborhood of this set.

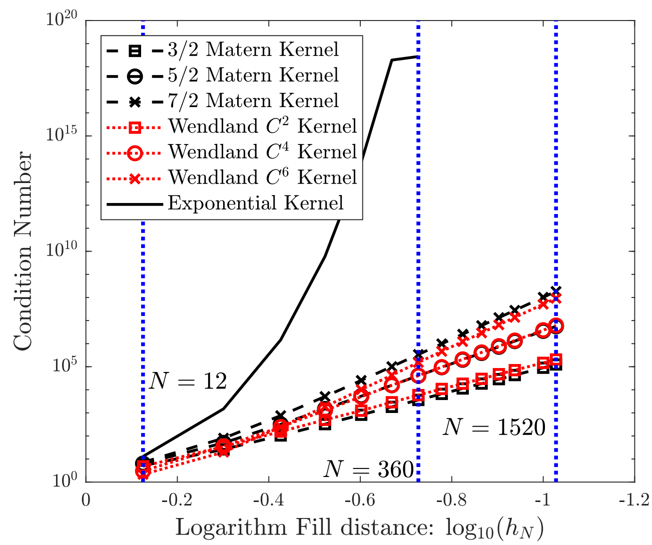

FIGURE 2 | Condition number of the matrix $\mathbb{K}$ for kernels used in the example.

The larger the number N of centers in Ξ_N, the denser the kernel centers, and, consequently, the smaller the fill distance $h_{\Xi_N,\Omega}$.

Figure 1 shows the limit cycle achieved by the reference trajectory $x_r(\cdot)$ and the set of centers Ξ_N used to capture the matched uncertainty. Knowing that the reference trajectory eventually reaches a limit cycle allowed us to distribute centers in a sufficiently large neighborhood of this set. In this example, the set Ω is defined as the convex hull of Ξ_N.

Figures 2–8 capture some key results from 182 numerical simulations. In particular, Figure 2 shows the condition number of the Grammian matrix $\mathbb{K}_N$ for different choices of reproducing kernel grows as a function of the fill distance. As is shown in Figure 2, the condition number of the Grammian matrix $\mathbb{K}_N$ associated

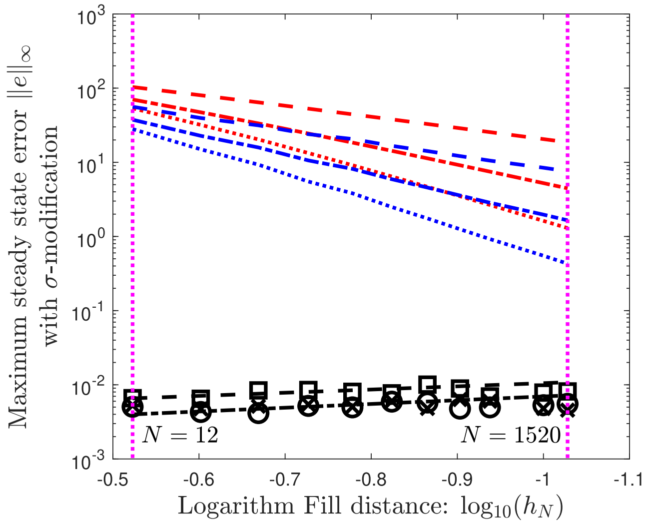

FIGURE 3 | Normalized steady-state tracking error of system using σ-modification and a family of Matern kernels. The conservative bound has the form $\sqrt{\frac{2}{\sigma\lambda_{\min}(Q)}}\|PB\| \sup_{\eta\in\Omega}\mathcal{P}_N(\eta)R$. The tighter bound has the form $\sqrt{\frac{2}{\sigma\lambda_{\min}(Q)}}\|PB\| \sup_{\eta\in\Omega}\mathcal{P}_N(\eta)\epsilon_N$. The legend is the same as Figure 4.

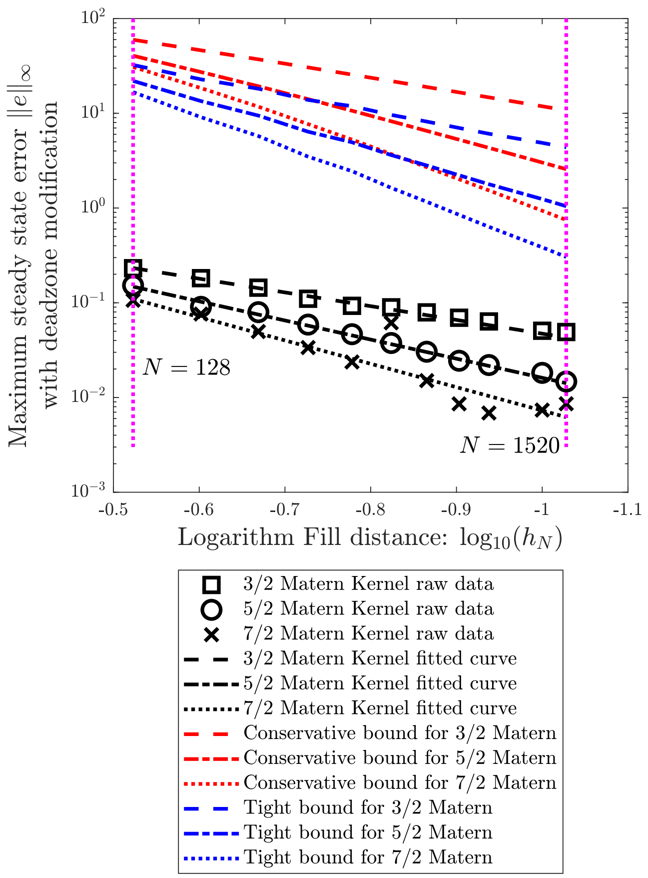

FIGURE 4 | Steady-state tracking error of system using dead-zone modification and a family of Matern kernels. The conservative bound has the form $\frac{2}{\sigma\lambda_{\min}(Q)}\|PB\| \sup_{\eta\in\Omega}\mathcal{P}_N(\eta)R$. The tighter bound has the form $\frac{2}{\sigma\lambda_{\min}(Q)}\|PB\| \sup_{\eta\in\Omega}\mathcal{P}_N(\eta)\epsilon_N$.

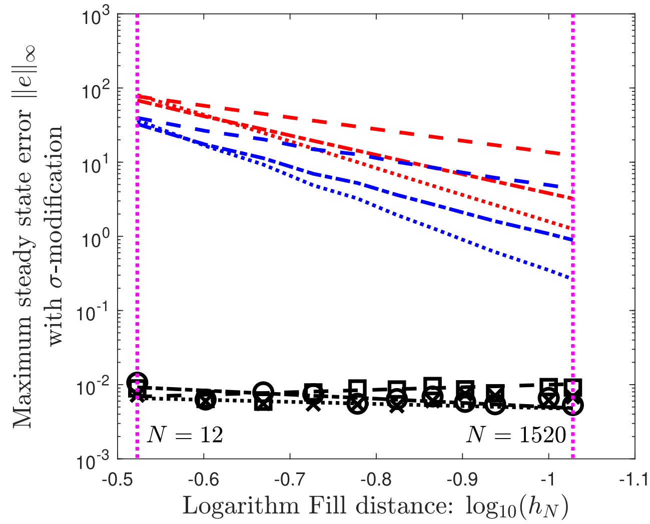

FIGURE 5 | Normalized steady-state tracking error of system using σ-modification and a family of Wendland kernels. The conservative and aggressive bound both share the same expressions as Figure 3. The legend is the same as Figure 6.

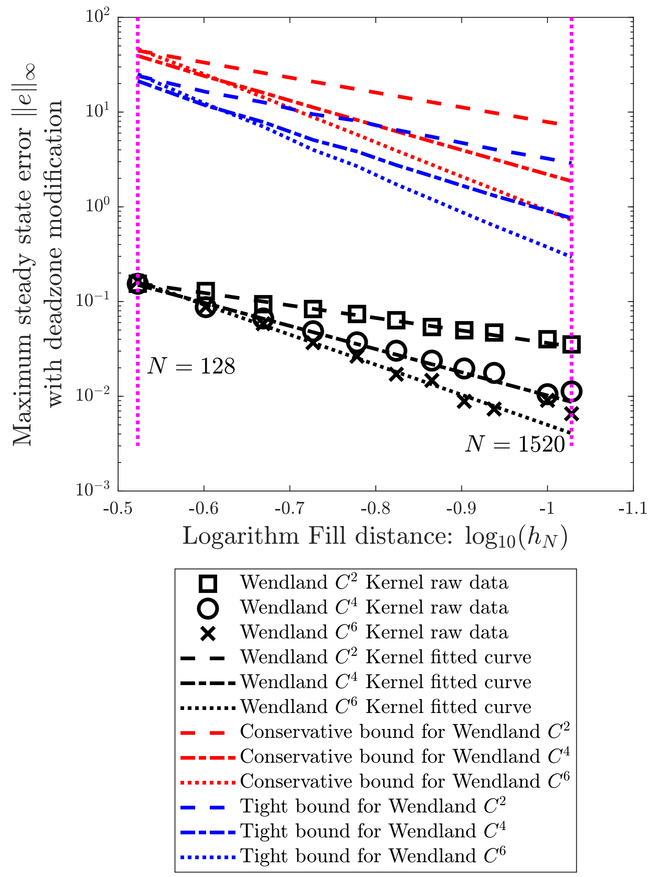

FIGURE 6 | Steady-state tracking error of system using dead-zone modification and a family of Wendland kernels. The conservative and aggressive bounds both share the same expressions as Figure 4.

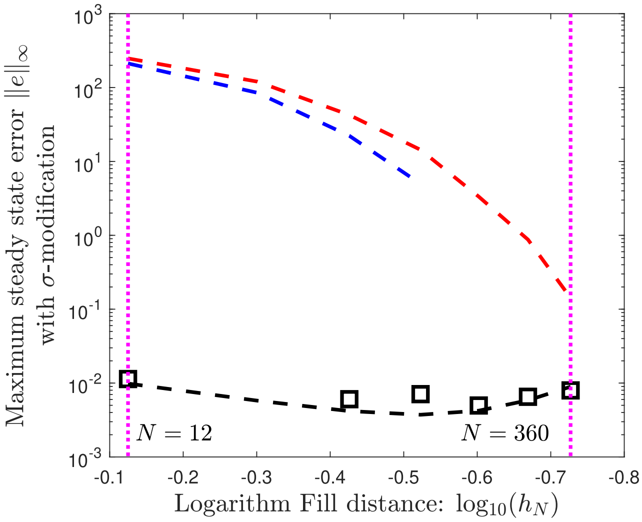

FIGURE 7 | Steady-state tracking error of system using dead-zone modification and a Gaussian kernel.

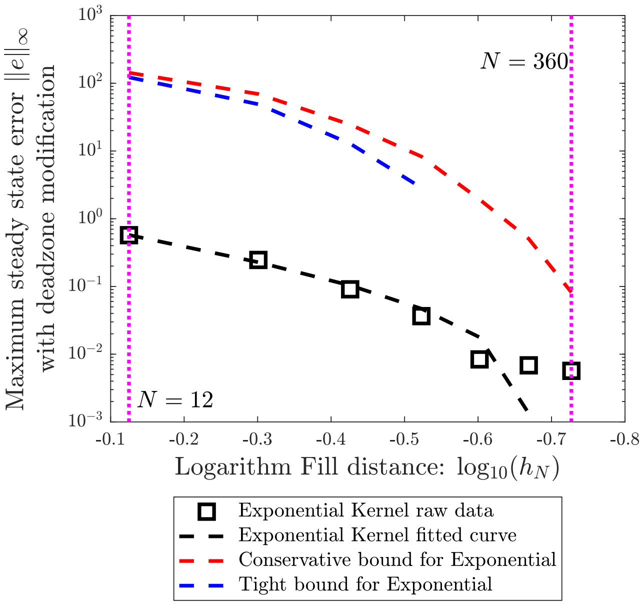

FIGURE 8 | Steady-state tracking error of system using dead-zone modification and Exponential Kernel.

with kernels of greater smoothness tends to grow faster than the condition number of the Grammian matrix associated with less smooth kernels. The Gaussian kernel produces the smoothest reproducing kernel among all kernels considered in this article, and its condition number grows much faster than for the other cases.

Figures 3–8 show that the ultimate bounds on the trajectory tracking error derived in this article are verified. Each figure explicitly depicts the boundaries of the uncertainty classes C_R and $C_{R,\epsilon,N}$; the expressions "conservative bound" captures $\sup_{\eta\in\Omega} \mathcal{P}_N(\eta)R$, and the expression "tighter bound" captures $\sup_{\eta\in\Omega} \mathcal{P}_N(\eta)\|(I - \mathbf{\Pi}_N)f\|_{\mathcal{H}}$, where $R \geq \|f\|_{\mathcal{H}}$. As expected, the larger the uncertainty class, the poorer the guarantee of performance. Worthy of attention is that the rates of decrease in Figures 4 and 6 because the achieved ultimate performance error closely matches that of the theoretical bounds.

Although smoother kernel functions may lead to poorly conditioned Grammian matrices, smoother kernels can be advantageous in terms of the rate of convergence of the approximations they generate. Smoother kernels generate more accurate approximations of smoother functions using fewer unknowns. Figures 3–8 capture numerical estimates of the maximum steady-state error in comparison with the size of the power function, as the fill distance between kernel basis centers decreases. It appears from these figures that the power functions for smoother kernels, such as the 7/2 Matern kernel or the Wendland C^6 function, have a much steeper reduction of tracking error than less smooth kernels, such as the 3/2 Sobolev–Matérn kernel or the Wendland C^2 kernel. The Gaussian kernel produces the fastest decrease in tracking error and the worst increase in the condition number, as the distribution of centers becomes denser and $N \to \infty$.

In general, since the power function determines an upper bound on tracking error, in theory, if the fill distance becomes arbitrarily small and N grows to infinity, then the tracking performance should improve. As shown in Figures 4, 6, and 8, when using dead-zone modification, the tracking performance improves as N increases and smoother kernel functions provide stronger results. Such behavior is expected since the bound d in Theorem 3 depends on $\sup_{\eta\in\Omega_0} \mathcal{P}_N(\eta)$, which is used to scale the approximation error ϵ_N. However, the computational cost of the proposed approach grows with N; this effect can be seen as a form of curse of dimensionality. Furthermore, if the fill distance becomes sufficiently small, then the tracking error oscillates. This behavior can be explained by the deterioration of the conditioning of the kernel matrix $\mathbb{K}_N$, the placement of centers, and rapid oscillations at the edge of the dead-zone, which is typical of this method. Note also that if the fill distance becomes sufficiently small, then the condition number of the Grammian for the smoother kernels increases beyond 10^5. Substantially, this means that perturbations due to the numerical integration error and numerical noise degrade the tracking error. In addition, for some basis center placements, there may be fewer basis centers that are centered directly on the trajectory, and the tracking performance may worsen.

In comparison with the dead-zone modification, which is sensitive to the power function value, the tracking performance of the σ-modification of MRAC decreases rapidly to the set tolerance of the simulation as shown in Figures 3, 5, and 7. The derived rates of convergence in Theorem 4 appear to be very conservative. This suggests an important direction of future research to improve the bounds in Theorem 3, and especially Theorem 4. A strategy may be to use arguments of the persistency of excitation, in the native space embedding context.

6.2 | Rotational Dynamics of a Fully Actuated Body With Unsteady Appendage

In this section, we discuss the problem of controlling the rotational dynamics of a fully actuated rigid body connected to an unsteady appendage. This problem is at the core of the inner loop control system design of aerial systems, such as uncrewed aerial vehicles (UAVs), which are usually fully actuated in their rotational dynamics [44]. In this case, the equations of motion are given by

$$\begin{bmatrix}\dot{\phi}(t)\\ \dot{\theta}(t)\\ \dot{\psi}(t)\end{bmatrix} = J^{-1}(\phi(t),\theta(t))\omega(t), \qquad \begin{bmatrix}\phi(0)\\ \theta(0)\\ \psi(0)\end{bmatrix} = \begin{bmatrix}\phi_0\\ \theta_0\\ \psi_0\end{bmatrix}, \quad t \geq 0 \tag{107}$$

$$\dot{\omega}(t) = \underbrace{I^{-1}}_{B}\left[\underbrace{M(t)}_{u(t)} - \underbrace{\begin{bmatrix}I_y - I_z & 0 & 0\\ 0 & I_z - I_x & 0\\ 0 & 0 & I_x - I_y\end{bmatrix}}_{\Theta_{\mathrm{param}}^{\mathrm{T}}}\underbrace{\begin{bmatrix}\omega_y(t)\omega_z(t)\\ \omega_x(t)\omega_z(t)\\ \omega_x(t)\omega_y(t)\end{bmatrix}}_{\Phi_{\mathrm{param}}(\omega(t))} + \underbrace{5\begin{bmatrix}|\cos(\omega_2)\cos(\omega_3)| - 0.5\\ |\cos(\omega_1)\cos(\omega_3)| - 0.5\\ |\cos(\omega_1)\cos(\omega_2)| - 0.5\end{bmatrix}}_{E_{\omega(t)}f}\right], \quad \omega(0) = \omega_0 \tag{108}$$

for details, see Reference [45, Ch. 1]. Here, $\phi : [t_0,\infty) \to [0,2\pi)$, $\theta : [t_0,\infty) \to (-\frac{\pi}{2},\frac{\pi}{2})$, and $\psi : [t_0,\infty) \to [0,2\pi)$ denote the *roll*, *pitch*, and *yaw* angles of a principal body reference frame relative to an inertial frame according to a 3-2-1 rotation sequence, respectively. The matrix

$$J(\phi,\theta) \triangleq \begin{bmatrix}1 & 0 & -\sin\theta\\ 0 & \cos\phi & \sin\phi\cos\theta\\ 0 & -\sin\phi & \cos\phi\cos\theta\end{bmatrix}, \quad (\phi,\theta) \in [0,2\pi) \times \left(-\frac{\pi}{2},\frac{\pi}{2}\right) \tag{109}$$

denotes the *Jacobian matrix*. The positive-definite matrix $I \triangleq \mathrm{diag}(I_x, I_y, I_z) \in \mathbb{R}^{3\times3}$ denotes the *inertia matrix*, and $\omega(t) \triangleq \left[\omega_x(t), \omega_y(t), \omega_z(t)\right]^{\mathrm{T}}$, $t \geq 0$, denotes the *angular velocity* of the body reference frame relative to the inertial reference frame, expressed in the body frame.

The rotational kinematic equations (107) are unaffected by uncertainties. Hence, to impose the twice continuously differentiable *user-defined roll*, *pitch*, and *yaw angles* $\phi_{\mathrm{user}} : [t_0,\infty) \to [0,2\pi)$, $\theta_{\mathrm{user}} : [t_0,\infty) \to (-\frac{\pi}{2},\frac{\pi}{2})$, $\psi_{\mathrm{user}} : [t_0,\infty) \to [0,2\pi)$, we would like to set $\lim_{t\to\infty} \|\omega(t) - \omega_{\mathrm{cmd}}(t)\| = 0$, where

$$\omega_{\mathrm{cmd}}(t) \triangleq J(\phi_{\mathrm{user}}(t),\theta_{\mathrm{user}}(t))\left[-K_P\left(\begin{bmatrix}\phi(t)\\ \theta(t)\\ \psi(t)\end{bmatrix} - \begin{bmatrix}\phi_{\mathrm{user}}(t)\\ \theta_{\mathrm{user}}(t)\\ \psi_{\mathrm{user}}(t)\end{bmatrix}\right) - K_I\int_0^t\left(\begin{bmatrix}\phi(\tau)\\ \theta(\tau)\\ \psi(\tau)\end{bmatrix} - \begin{bmatrix}\phi_{\mathrm{user}}(\tau)\\ \theta_{\mathrm{user}}(\tau)\\ \psi_{\mathrm{user}}(\tau)\end{bmatrix}\right)\mathrm{d}\tau + \begin{bmatrix}\dot{\phi}_{\mathrm{user}}(t)\\ \dot{\theta}_{\mathrm{user}}(t)\\ \dot{\psi}_{\mathrm{user}}(t)\end{bmatrix}\right] \tag{110}$$

The symmetric, positive-definite matrices $K_P, K_I \in \mathbb{R}^{3\times3}$ denote the proportional feedback and integration feedback coefficients.

The rotational dynamic equations (108) are in the same form as (1) with $n = 3$, $m = 3$, $A = 0$, $B = I^{-1}$, $x(t) = \omega(t)$, and $u(t) = M(t)$. In this article, we assume that the inertia matrix I is unknown. Hence, to compute the adaptive laws, we set B equal to the identity matrix. This formulation of the plant model, which includes uncertainties in the matrix B, is not explicitly admitted by the proposed theoretical results. Hence, this example shows how the proposed results can be extended to additional classes of functional uncertainties; a formal justification of this result, which is to be sought in the diagonal structure of I, is left to future research. Uncertainties in the matrix of inertia are captured not only by diagonal, right-multiplicative uncertainties to the matrix B (for details, see Reference [46]) but also by the term $\Theta_{\mathrm{param}}^{\mathrm{T}}\Phi_{\mathrm{param}}(\omega(t))$ in (108). Finally, the term $E_{\omega(t)}f$ in (108) accounts for the effect of the unsteady payload.

In this article, we use the proposed nonparametric MRAC systems to augment a classical parametric MRAC system. To this goal, we consider the control input

$$u_N(t) = \hat{\alpha}_N^{\mathrm{T}}(t)x_N(t) + \hat{\beta}_N^{\mathrm{T}}(t)r(t) - \hat{\Theta}_{\mathrm{param}}^{\mathrm{T}}\Phi_{\mathrm{param}}(x_N(t)) - E_{x_N(t)}\hat{f}_N(t,\cdot), \quad t \geq 0 \tag{111}$$

where the adaptive gain matrices $\hat{\alpha}(\cdot)$ and $\hat{\beta}(\cdot)$ verify (63) and (64) in the case of the dead-zone modification of MRAC and (71) and (72) in the case of the σ-modification of MRAC. The *adaptive gain matrix* $\hat{\Theta}_{\mathrm{param}} : [0,\infty) \to \mathbb{R}^{3\times3}$ verifies either

$$\dot{\hat{\Theta}}_{\mathrm{param}}(t) = \begin{cases}-\Gamma_\gamma\Phi_{\mathrm{param}}(x_N(t))e_N^{\mathrm{T}}(t)PB, & \|e_N(t)\| \geq d,\\ 0, & \|e_N(t)\| < d,\end{cases} \qquad \hat{\Theta}_{\mathrm{param}}(0) = \hat{\Theta}_{\mathrm{param},0}, \quad t \geq 0 \tag{112}$$

in the case of the dead-zone modification of MRAC, where the *adaptive rate matrix* $\Gamma_\gamma \in \mathbb{R}^{3\times3}$ is user-defined, symmetric, and positive-definite, or

$$\dot{\hat{\Theta}}_{\mathrm{param}}(t) = -\Gamma_\gamma\Phi_{\mathrm{param}}(x_N(t))e_N^{\mathrm{T}}(t)PB - \sigma\hat{\Theta}_{\mathrm{param}}(t), \qquad \hat{\Theta}_{\mathrm{param}}(0) = \hat{\Theta}_{\mathrm{param},0}, \quad t \geq 0 \tag{113}$$

in the case of the σ-modification of MRAC, where $\sigma > 0$. Finally, the adaptive gains $\hat{f}_N(t,\cdot)$, $t \geq 0$, verify (65) in the case of the dead-zone modification of MRAC and (73) in the case of the σ-modification of MRAC.

To enforce that $\lim_{t\to\infty}\|\omega(t) - \omega_{\mathrm{cmd}}(t)\| = 0$ using the MRAC framework, we define the reference model

$$\dot{\omega}_{\mathrm{ref}}(t) = \underbrace{-K_{P,\omega}\omega_{\mathrm{ref}}(t)}_{A_{ref}\omega_{ref}} + \underbrace{\left[K_{I,\omega}\int_0^t\left(\omega_{\mathrm{ref}}(\tau) - \omega_{\mathrm{cmd}}(\tau)\right)\mathrm{d}\tau + K_{P,\omega}\dot{\omega}_{\mathrm{cmd}}(t)\right]}_{r(t)}, \qquad \omega_{\mathrm{ref}}(0) = \omega_{\mathrm{ref},0}, \quad t \geq 0 \tag{114}$$

where $K_{P,\omega}, K_{I,\omega} \in \mathbb{R}^{3\times3}$ are user-defined, symmetric, and positive-definite. Remarkably, this dynamical model captures

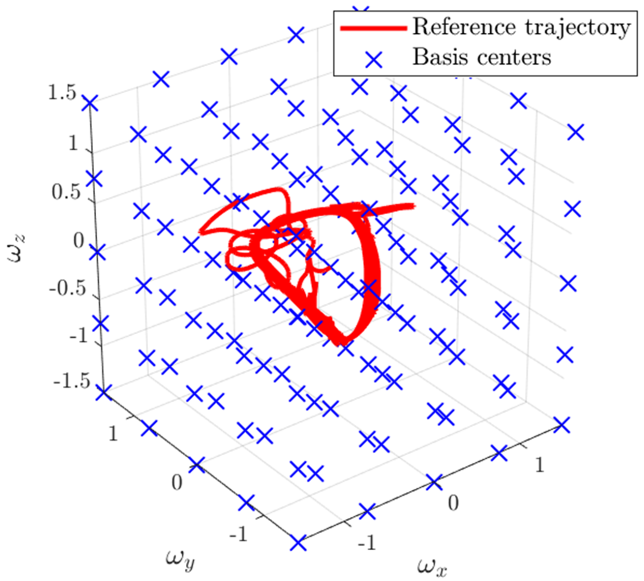

FIGURE 9 | Location of basis centers relative to reference trajectory in rotation velocity space.

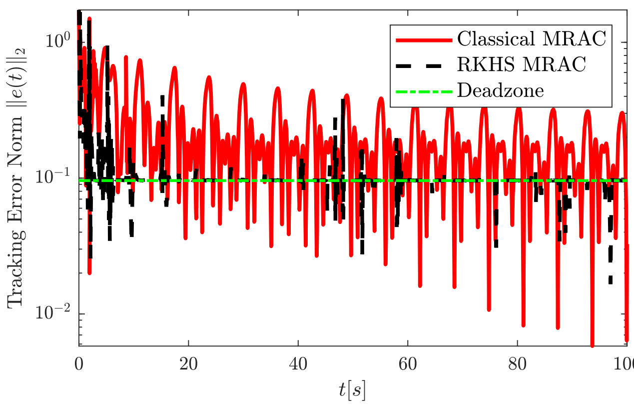

FIGURE 10 | Norm of tracking error as a function of time applying the dead-zone modification of MRAC. Nonparametric MRAC consistently reduces the tracking error at each time instant.

a PI-controlled system so that $\lim_{t\to\infty} \|\omega_{\text{ref}}(t) - \omega_{\text{cmd}}(t)\| = 0$, and, since the MRAC framework guarantees that $\lim_{t\to\infty} \|\omega(t) - \omega_{\text{ref}}(t)\| = 0$, the desired attitude is eventually attained.

Figures 10–15 show numerical simulation results to attain the user-defined attitude captured by $\begin{bmatrix} \phi_{\text{user}}(t) \\ \theta_{\text{user}}(t) \\ \psi_{\text{user}}(t) \end{bmatrix} = 0.5[\cos(t), \sin(t), \sin(t)]^{\mathrm{T}}$. To capture the rotational dynamics of an actual X8-copter, we let $I = \operatorname{diag}(2.27199, 2.2020, 1.6146) \cdot 10^{-2}$ kg $\cdot$ m^2. We consider the homogeneous distribution of 125 kernel centers depicted in Figure 9, and the tunable parameters

$$K_P = 1.5 \cdot \operatorname{diag}(18.0, 20.0, 10.0), \quad K_I = 1.5 \cdot \operatorname{diag}(0.2, 0.2, 0.5),$$
$$K_{P,\omega} = \operatorname{diag}(14.0, 12.0, 2.0), \quad K_{I,\omega} = \operatorname{diag}(0.1, 0.15, 0.4),$$
$$\Gamma_\alpha = 100 \cdot \mathbf{1}_3, \quad \Gamma_\beta = \mathbf{1}_3, \quad \Gamma_\gamma = 1500 \cdot \mathbf{1}_3,$$
$$\Gamma_f = 1000 \cdot \mathbf{1}_3, \quad Q = \mathbf{1}_3, \quad \sigma = 0.01, \quad \epsilon_N = 0.5$$

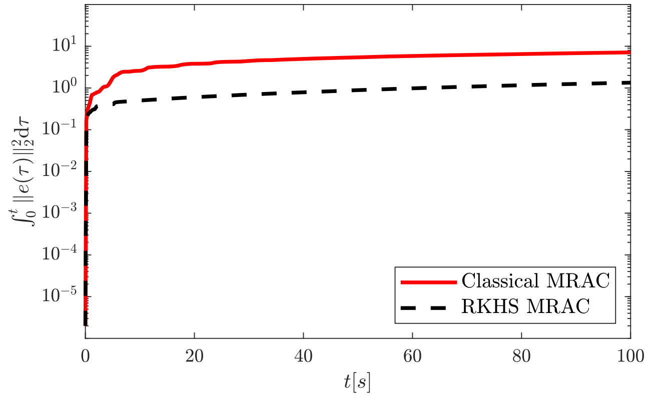

FIGURE 11 | L^2 norm of tracking error as a function of time applying the dead-zone modification of MRAC. Nonparametric MRAC consistently reduces the cumulative tracking error over time.

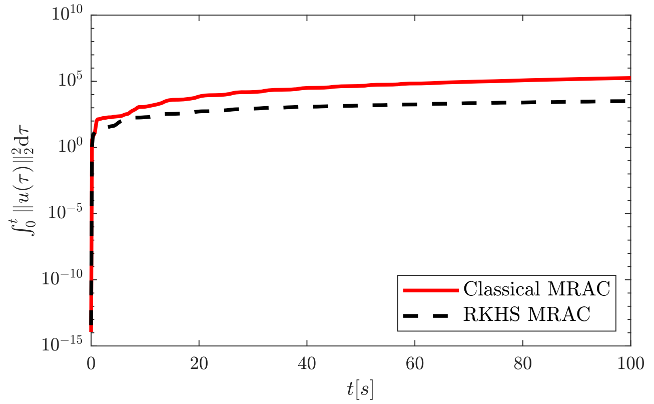

FIGURE 12 | L^2 norm of control effort as a function of time applying the dead-zone modification of MRAC. Nonparametric MRAC consistently reduces the cumulative control effort over time.

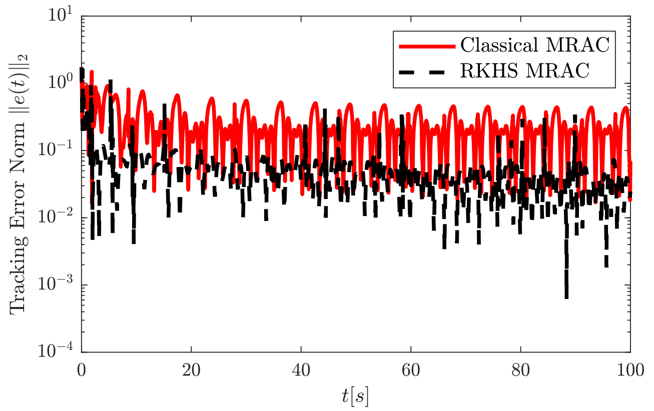

FIGURE 13 | Norm of tracking error as a function of time applying the σ-modification of MRAC. Nonparametric MRAC consistently reduces the tracking error at each time instant.

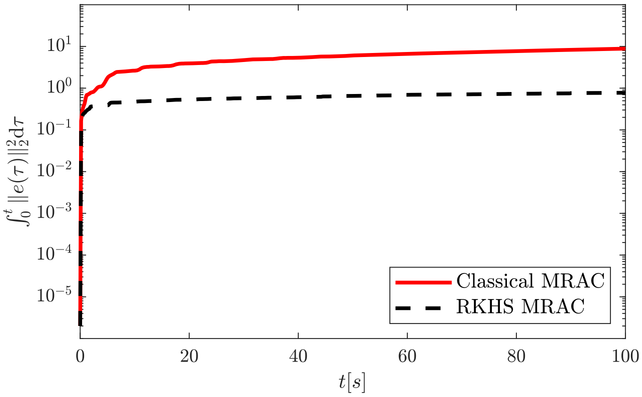

FIGURE 14 | L^2 norm of tracking error as a function of time applying the σ-modification of MRAC. Nonparametric MRAC consistently reduces the cumulative tracking error over time.

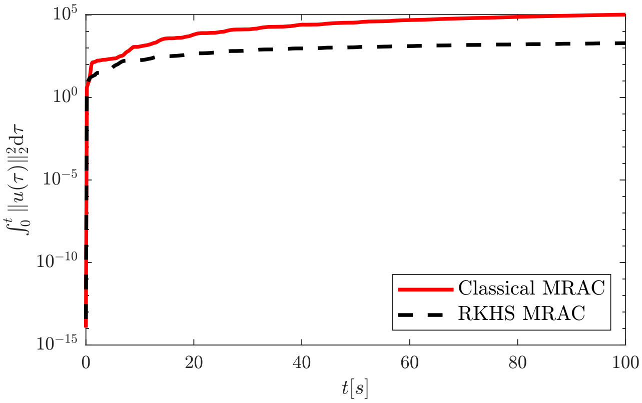

FIGURE 15 | L^2 norm of control effort as a function of time applying the σ-modification of MRAC. Nonparametric MRAC consistently reduces the cumulative control effort over time.

We chose 3/2 Matern-Sobolev kernels with $p = 1$ and $l = 1$. In this example, the set Ω is defined as the convex hull of Ξ_N.

Figures 10–12 show the norm of the trajectory tracking error, the L_2-norm of the trajectory tracking error, and the L_2-norm of the control effort obtained by applying the dead-zone modification of MRAC both with and without the proposed nonparametric MRAC augmentation. Figures 13–15 show similar results applying the σ-modification of MRAC. It is apparent how the proposed nonparametric MRAC produces smaller tracking errors for smaller cumulative control efforts.

7 | Conclusion

This article presented the first set of results toward a general theory of robust adaptive control for functional uncertainty classes in a vector-valued native space. In particular, this article showed how, under mild assumptions on the uncertainties' functional space, and without any information on the structure or shape of functional uncertainties in terms of regressors, as typically occurs in classical MRAC, it is possible to construct a control law and adaptive laws that guarantee uniform ultimate boundedness of the trajectory tracking error. Two numerical examples show the applicability of the proposed results, their advantages over classical parametric MRAC, and the effect of user-defined choices such as the density of kernel centers distribution and the smoothness of the underlying reproducing kernels.

The proposed results give insight into future research directions. Examples of open questions are related to the conservative nature of the ultimate bounds on the trajectory tracking error; the dimension, computational cost, and numerical stability of the realizable controllers; and the sometimes delicate interplay among the achievable error bounds, functional uncertainty class, approximation error, numerical conditioning, and external disturbances. Additional future works concern the application of the proposed results to complex problems engineering problems such as the design of control systems for underactuated UAVs or marine vehicles.

Author Contributions

Haoran Wang was responsible for writing this article, performing simulations, and refining proofs. Drs. Kurdila, L'Afflitto, and Stilwell were responsible for outlining proofs and conceptualizing the essential ideas underlying this work. Derek Oesterheld supported Haoran Wang in preparing numerical simulations and proofreading this article.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Dr. L'Afflitto website at https://lafflitto.com/Codes/Vector_RKHS_MRAC_Examples.zip.

References

1. S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence and Robustness* (Prentice Hall, 2011).

2. E. Lavretsky and K. Wise, *Robust and Adaptive Control: With Aerospace Applications* (Springer, 2013).

3. K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems* (Prentice Hall, 1989).

4. P. A. Ioannou and J. Sun, *Robust Adaptive Control* (Courier Corporation, 2012).

5. T. Gao, S. Z. Kovalsky, D. M. Boyer, and I. Daubechies, "Gaussian Process Landmarking for Three-Dimensional Geometric Morphometrics," *SIAM Journal on Mathematics of Data Science* 1, no. 1 (2019): 237–267.

6. T. Wenzel, G. Santin, and B. Haasdonk, "Analysis of Target Data-Dependent Greedy Kernel Algorithms: Convergence Rates for f-, $f \cdot P$- and f/P-Greedy," *Constructive Approximation* 57, no. 1 (2023): 45–74, https://doi.org/10.1007/s00365-022-09592-3.

7. T. Wenzel, G. Santin, and B. Haasdonk, "A Novel Class of Stabilized Greedy Kernel Approximation Algorithms: Convergence, Stability and Uniform Point Distribution," *Journal of Approximation Theory* 262 (2021): 105508.

8. N. M. Boffi, S. Tu, and J. J. E. Slotine, "Nonparametric Adaptive Control and Prediction: Theory and Randomized Algorithms," *Journal of Machine Learning Research* 23, no. 1 (2022): 12841–12886.

9. A. Gahlawat, P. Zhao, A. Patterson, N. Hovakimyan, and E. Theodorou, "L1-GP: L1 Adaptive Control With Bayesian Learning," in *Learning for Dynamics and Control* (PMLR, 2020), 826–837.

10. A. Lederer, J. Umlauft, and S. Hirche, "Uniform Error Bounds for Gaussian Process Regression With Application to Safe Control," in *Proceedings of the 33rd International Conference on Neural Information Processing Systems* (2019): 659–669.

11. A. Kurdila and Y. Lei, "Adaptive Control via Embedding in Reproducing Kernel Hilbert Spaces," in *American Control Conference* (IEEE, 2013), 3384–3389.

12. J. Guo, S. T. Paruchuri, and A. J. Kurdila, "Partial Persistence of Excitation in RKHS Embedded Adaptive Estimation," *IEEE Transactions on Automatic Control* 68, no. 10 (2023): 5850–5861, https://doi.org/10.1109/TAC.2022.3226710.

13. S. T. Paruchuri, J. Guo, and A. Kurdila, "Reproducing Kernel Hilbert Space Embedding for Adaptive Estimation of Nonlinearities in Piezoelectric Systems," *Nonlinear Dynamics* 101, no. 2 (2020): 1397–1415.

14. M. A. Demetriou, "Functional Estimation of Perturbed Positive Real Infinite Dimensional Systems Using Adaptive Compensators," in *American Control Conference* (IEEE, 2020), 1582–1587.

15. S. T. Paruchuri, J. Guo, and A. Kurdila, "Sufficient Conditions for Parameter Convergence Over Embedded Manifolds Using Kernel Techniques," *IEEE Transactions on Automatic Control* 68, no. 2 (2023): 753–765, https://doi.org/10.1109/TAC.2022.3148716.

16. J. Guo, S. T. Paruchuri, and A. J. Kurdila, "Approximations of the Reproducing Kernel Hilbert Space (RKHS) Embedding Method Over Manifolds," in *Conference on Decision and Control* (IEEE, 2020), 1596–1601.

17. J. A. Burns, J. Guo, A. J. Kurdila, S. T. Paruchuri, and H. Wang, "Error Bounds for Native Space Embedding Observers With Operator-Valued Kernels," in *American Control Conference* (IEEE, 2023).

18. J. Burns, A. Kurdila, D. Oesterheld, D. Stilwell, and H. Wang, "Learning Theory Convergence Rates for Observers and Controllers in Native Space Embedding," in *American Control Conference* (IEEE, 2023), 979–986.

19. H. Wang, N. Powell, A. L'Afflitto, A. J. Kurdila, and J. Burns, "The Power Function for Adaptive Control in Native Space Embedding," in *Control and Decision Conference* (IEEE, 2023).

20. D. I. Oesterheld, D. J. Stilwell, A. J. Kurdila, and J. Guo, "A Model Reference Adaptive Controller Based on Operator-Valued Kernel Functions," in *Conference on Decision and Control* (IEEE, 2023).

21. J. A. Farrell and M. M. Polycarpou, *Adaptive Approximation Based Control: Unifying Neural, Fuzzy and Traditional Adaptive Approximation Approaches*, vol. 48 (John Wiley & Sons, 2006).

22. S. Niu, A. Bouland, H. Wang, et al., "Convergence Rates of Online Critic Value Function Approximation in Native Spaces," *IEEE Control Systems Letters* 8 (2024): 2145–2150, https://doi.org/10.1109/LCSYS.2024.3417178.

23. A. J. Kurdila, A. L'Afflitto, and J. A. Burns, *Data-Driven, Nonparametric, Adaptive Control Theory* (Springer, 2024).

24. H. Tsukamoto, S. J. Chung, and J. J. Slotine, "Learning-Based Adaptive Control Using Contraction Theory," in *Conference on Decision and Control* (IEEE, 2021), 2533–2538.

25. L. K. Held, L. Vermeylen, D. Dignath, W. Notebaert, R. M. Krebs, and S. Braem, "Reinforcement Learning of Adaptive Control Strategies," *Communications Psychology* 2, no. 1 (2024): 8, https://doi.org/10.1038/s44271-024-00055-y.

26. Q. Wen, D. Liu, J. Wang, and S. Baldi, "An Adaptive Longitudinal Platooning Design Based on Concurrent Learning," *IEEE Control Systems Letters* 8 (2024): 303–308, https://doi.org/10.1109/LCSYS.2023.3333004.

27. N. Powell, J. Guo, H. Wang, A. L'Afflitto, J. Burns, and A. Kurdila, "Greedy Native Space Embedding MRAC via Newton Bases," *International Journal of Adaptive Control and Signal Processing* (2025).

28. H. K. Khalil, *Nonlinear Systems* (Prentice Hall, 2002).

29. B. Peterson and K. Narendra, "Bounded Error Adaptive Control," *IEEE Transactions on Automatic Control* 27, no. 6 (1982): 1161–1168, https://doi.org/10.1109/TAC.1982.1103112.

30. H. Wendland, *Scattered Data Approximation* (Cambridge University Press, 2004).

31. D. Wittwar, G. Santin, and B. Haasdonk, "Interpolation With Uncoupled Separable Matrix-Valued Kernels," *Seminari Padovani di Analisi Numerica* 11 (2018): 23–39.

32. V. I. Paulsen and M. Raghupathi, *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, vol. 152 (Cambridge University Press, 2016).

33. S. Saitoh and Y. Sawano, *Theory of Reproducing Kernels and Applications* (Springer, 2016).

34. A. Berlinet and C. Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics* (Springer Science & Business Media, 2011).

35. T. Evgeniou, C. A. Micchelli, and M. Pontil, "Learning Multiple Tasks With Kernel Methods," *Journal of Machine Learning Research* 6 (2005): 615–637.

36. R. Schaback, "Error Estimates and Condition Numbers for Radial Basis Function Interpolation," *Advances in Computational Mathematics* 3 (1994): 251–264.

37. R. B. Anderson, J. A. Marshall, and A. L'Afflitto, "Novel Model Reference Adaptive Control Laws for Improved Transient Dynamics and Guaranteed Saturation Constraints," *Journal of the Franklin Institute* 358, no. 12 (2021): 6281–6308, https://doi.org/10.1016/j.jfranklin.2021.06.020.

38. A. L'Afflitto, "Barrier Lyapunov Functions and Constrained Model Reference Adaptive Control," *IEEE Control Systems Letters* 2, no. 3 (2018): 441–446, https://doi.org/10.1109/LCSYS.2018.2842148.

39. A. L'Afflitto and T. A. Blackford, "Constrained Dynamical Systems, Robust Model Reference Adaptive Control, and Unreliable Reference Signals," *International Journal of Control* 93, no. 5 (2020): 1039–1052, https://doi.org/10.1080/00207179.2018.1489147.

40. R. A. Adams and J. J. Fournier, *Sobolev Spaces* (Elsevier, 2003).

41. A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Springer, 2012).

42. H. Wang, A. L'Afflitto, and A. J. Kurdila, "Robust Model Reference Adaptive Control Based on Reproducing Kernel Hilbert Spaces: Matlab Codes," (2025), https://lafflitto.com/Codes/Vector_RKHS_MRAC_Examples.zip.

43. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables* (Dover, 1972).

44. J. A. Marshall, W. Sun, and A. L'Afflitto, "A Survey of Guidance, Navigation, and Control Systems for Autonomous Multi-Rotor Small Unmanned Aerial Systems," *Annual Reviews in Control* 52 (2021): 390–427, https://doi.org/10.1016/j.arcontrol.2021.10.013.

45. A. L'Afflitto, *A Mathematical Perspective on Flight Dynamics and Control* (Springer, 2017).

46. A. L'Afflitto, R. B. Anderson, and K. Mohammadi, "An Introduction to Nonlinear Robust Control for Unmanned Quadrotor Aircraft: How to Design Control Algorithms for Quadrotors Using Sliding Mode Control and Adaptive Control Techniques," *IEEE Control Systems Magazine* 38, no. 3 (2018): 102–121, https://doi.org/10.1109/MCS.2018.2810559.