2020 59th IEEE Conference on Decision and Control (CDC)
Jeju Island, Republic of Korea, December 14-18, 2020

Approximations of the Reproducing Kernel Hilbert Space (RKHS) Embedding Method over Manifolds

Jia Guo[1], Sai Tej Paruchuri[1] and Andrew J. Kurdila[1]

***Abstract*— The reproducing kernel Hilbert space (RKHS) embedding method is a recently introduced estimation approach that seeks to identify the unknown or uncertain function in the governing equations of a nonlinear set of ordinary differential equations (ODEs). While the original state estimate evolves in Euclidean space, the function estimate is constructed in an infinite dimensional RKHS and must be approximated in practice. When a finite dimensional approximation is constructed using a basis defined in terms of shifted kernel functions centered at the observations along a trajectory, the RKHS embedding method can be understood as a data-driven approach. This paper derives sufficient conditions that ensure that approximations of the unknown function converge in a Sobolev norm over a submanifold that supports the dynamics. Moreover, the rate of convergence for the finite dimensional approximations is derived in terms of the fill distance of the samples in the embedded manifold. A numerical simulation of an example problem is carried out to illustrate the qualitative nature of convergence results derived in the paper.**

I. INTRODUCTION

Data-driven modeling of uncertain or unknown nonlinear dynamic systems has been a topic of great interest over the past few years [1], [2]. The collection of algorithms that can, in some sense, be viewed as data-dependent methods is vast. Specific examples include the following: the collection of studies on the extended dynamic mode decomposition (EDMD) algorithm and its variants that are based on Koopman theory [3], [4]; adaptive basis methods for online adaptive estimation [5], [6]; fuzzy control methods based on neural networks [7]; and strategies from distribution-free learning theory and nonlinear regression [8].

Recently the authors have introduced a novel approach, the RKHS embedding method in [9]–[12], for the estimation of uncertain systems. This method likewise can be viewed as a type of data-dependent algorithm when bases of approximation are selected along the trajectory of an unknown system. The RKHS embedding method generalizes estimators used in conventional adaptive estimation over finite dimensional state spaces. The approach essentially lifts the learning law of the estimation scheme to an infinite dimensional RKHS H of real-valued functions defined over the state space. The unknown function $f(\cdot)$ that characterizes the uncertainty in the ordinary differential equations (ODEs) of dynamical system of is assumed to be an element of the RKHS H. The resulting overall estimator is thereby defined for both the states and the unknown function, and it defines an evolution in $\mathbb{R}^d \times H$. Since the evaluation functional $\mathcal{E}_x$ is linear and bounded in the RKHS H, the unknown nonlinearity defined by $x \mapsto f(x)$ in the original ODE can be expressed as $\mathcal{E}_x f$ in the RKHS embedding formulation, that is, a bounded linear operator acting on the function $f \in H$. In this way, the nonlinearity in the original ODEs is avoided in the RKHS error equations. The trade-off is that one has to conduct analysis in the infinite dimensional spaces, which is usually (much) more complicated.

[1]Jia Guo, Sai Tej Paruchuri and Andrew J. Kurdila are with the Department of Mechanical Engineering, Virginia Tech, Blacksburg VA 24060, USA jguo18@vt.edu, saitejp@vt.edu, kurdila@vt.edu

In a way that is analogous to the conventional adaptive estimation in finite dimensional spaces, the convergence of the RKHS embedded estimator can be guaranteed with the satisfaction of a condition of persistence of excitation (PE). The notion of the PE condition for the RKHS embedding method has been introduced and studied in the authors' previous work [13], [14]. Given a subset $\Omega \subseteq \mathbb{R}^d$, a necessary condition for Ω to be PE by a positive orbit $\Gamma^+(x_0)$ starting at x_0 is that the neighborhoods of points in Ω are "visited infinitely often" by the trajectory. This means that Ω must be a subset of the ω-limit set, in which every point is the limit of a subsequence of points extracted from the trajectory [14]. A convergence result states that over an indexing set Ω that is persistently excited, the estimate of the unknown function converges to the actual unknown function [13]. However, the function estimate generated by the RKHS embedded estimator lies in the abstract infinite dimensional RKHS. In order to compute the practical estimates, a finite dimensional approximation has to be implemented. The error generated from the approximation is the major topic of this paper.

Most of time, a discussion about the convergence of approximations in an RKHS can be transformed into a discussion about the operator $I - \mathbf{P}_{\Omega_n}$, where $\mathbf{P}_{\Omega_n} : H_X \to H_{\Omega_n}$ denotes the orthogonal projection onto the finite dimensional approximant subspace H_{Ω_n}. Additional insight, or sometimes a finer analysis, can be obtained by interpreting this projection error in other well-known spaces. It is known that many commonly used RKHS are either embedded in or equivalent to some Sobolev space $W^{\tau,2}(\mathbb{R}^d)$. The fact that the family of Sobolev spaces provides a refined characterization of the smoothness of functions is useful for estimating the approximation error. In this paper, we first review carefully the relationship between some types of RKHS and Sobolev spaces. Then the error equations for some type of RKHS embedded estimator are recast in Sobolev spaces to facilitate the error analysis. Using the recently introduced results on the Sobolev error bounds for the interpolation operator [15]–[17], the rate of convergence for the finite dimensional approximation of the RKHS em-

978-1-7281-7447-1/20/$31.00 ©2020 IEEE

bedding equations is derived in terms of the fill distance of the samples Ω_n with respect to the subset (or manifold) Ω. The most succinct form of this new error bound states that the error decays like $\mathcal{O}(h^{s-\mu}_{\Omega_n,\Omega})$ where s and μ are the smoothness indices that depend on the smoothness of the unknown function, the smoothness of approximations, and the choice of the reproducing kernel and Sobolev spaces.

II. Problem Setup

A. Reproducing Kernel Hilbert Spaces

A real RKHS is a Hilbert space H_X of real-valued functions defined over X that admits a reproducing kernel $\mathfrak{K} : X \times X \to \mathbb{R}$. The kernel $\mathfrak{K}(\cdot,\cdot)$ has reproducing property provided for all $x \in X$ and $f \in H_X$, $f(x) = (f, \mathfrak{K}_x)_{H_X}$. By the Moore-Aronszajn theorem [18], the RKHS H_X is the completion of the space spanned by the kernel basis functions $\mathfrak{K}_x := \mathfrak{K}(\cdot, x)$ centered at $x \in X$,

$$H_X = \overline{\text{span}\{\mathfrak{K}(\cdot, x) : x \in X\}}.$$

If $\Omega \subseteq X$ is a subset, then the closed subspace $H_\Omega = \overline{\text{span}\{\mathfrak{K}(\cdot, x) : x \in \Omega\}}$ is also an RKHS. Here Ω is referred to as the indexing set of the space H_Ω, to distinguish it from the support X of functions in $H_\Omega \subseteq H_X$. The whole space H_X can be expressed as the direct sum $H_X = H_\Omega \oplus V_\Omega$. A function $\phi \in H_X$ belongs to V_Ω if and only if $\phi(x) = 0$ for all $x \in \Omega$ [18]. This characterization of V_Ω is particularly useful for computing the projection $\mathbf{P}_\Omega f$ for $f \in H_X$. Since $(I - \mathbf{P}_\Omega)f \in V_\Omega$, we must have for all $x \in \Omega$,

$$((I - \mathbf{P}_\Omega)f, \mathfrak{K}_x)_{H_X} = f(x) - (\mathbf{P}_\Omega f)(x) = 0.$$

Thus for any *finite discrete* set Ω_n, the projection operation is equivalent to the interpolation, $\mathbf{P}_{\Omega_n} \equiv \mathbf{I}_{\Omega_n} : H_X \to H_{\Omega_n}$.

In this paper, we only consider an RKHS that is continuously embedded in the space of continuous functions $C(X)$, denoted $H_X \hookrightarrow C(X)$. Denote the evaluation operator on H_X by $\mathcal{E}_x : f \mapsto f(x)$ denote. It can be proven that $\mathcal{E}_x$ is a bounded linear operator. One sufficient condition for $H_X \hookrightarrow C(X)$ is that there exists a constant $\bar{k}$ such that $\sqrt{\mathfrak{K}(x,x)} \leq \bar{k} < \infty$ for all $x \in X$. In fact, by the Cauchy-Schwartz inequality we have

$$|\mathcal{E}_x f| = |(f, \mathfrak{K}_x)_{H_X}| \leq \|f\|_{H_X}\|\mathfrak{K}_x\|_{H_X} = \|f\|_{H_X}\sqrt{\mathfrak{K}(x,x)}.$$

If the aforementioned constant $\bar{k}$ exists, then we have

$$\|f\|_{C(X)} = \sup_{x\in X} |\mathcal{E}_x f| \leq \|f\|_{H_X} \sup_{x\in X} \sqrt{\mathfrak{K}(x,x)} \leq \bar{k}\|f\|_{H_X},$$

which implies $H_X \hookrightarrow C(X)$. In all the following discussions, we assume that such a constant $\bar{k}$ always exists.

In some cases it is useful to consider the space of restrictions $\mathbf{R}_\Omega H_X$. Clearly, the restriction operator $\mathbf{R}_\Omega : H_X \to \mathbf{R}_\Omega H_X$ is linear and onto. It follows from the fact that $H_X := H_\Omega \oplus V_\Omega$ that $\mathbf{R}_\Omega V_\Omega = \{0\}$. In fact, one can show that $\mathbf{R}_\Omega : H_\Omega \to \mathbf{R}_\Omega H_X$ is a bijection, and the inverse operator $(\mathbf{R}_\Omega|_{H_\Omega})^{-1} : \mathbf{R}_\Omega H_X \to H_\Omega$ defines an extension operator, denoted by $\mathbf{E}_\Omega$. It follows that

$$\mathbf{P}_\Omega = \mathbf{E}_\Omega \mathbf{R}_\Omega : H_X \to H_\Omega. \tag{1}$$

By defining the inner product in $\mathbf{R}_\Omega H_X$ as $(f, g)_{\mathbf{R}_\Omega H_X} := (\mathbf{E}_\Omega f, \mathbf{E}_\Omega g)_{H_X}$, we can show that $\mathbf{R}_\Omega H_X$ is an RKHS. The associated reproducing kernel is shown to be the restriction of the reproducing kernel $\mathfrak{K}(\cdot,\cdot)$ over $\Omega \times \Omega$, that is, $\mathbf{R}_\Omega \mathfrak{K} := \mathfrak{K}|_{\Omega\times\Omega}$ [18].

B. RKHS embedded estimator

The governing equation of the partially unknown dynamic system in this paper has the form

$$\dot{x}(t) = Ax(t) + Bf(x(t)), \tag{2}$$

where $A \in \mathbb{R}^{d\times d}$ is a Hurwitz matrix, $B \in \mathbb{R}^{d\times 1}$, and $f : \mathbb{R}^d \to \mathbb{R}$ is the unknown nonlinear function. The generality of this formulation is discussed in [13]. The governing equations of the RKHS embedded estimator have a structure that is similar to that of a classical adaptive observer in Euclidean space, but modified to include the evolution of functions in the RKHS. We set

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + B\mathcal{E}_{x(t)}\hat{f}(t), \\ \dot{\hat{f}}(t) &= \gamma(B\mathcal{E}_{x(t)})^* P(x(t) - \hat{x}(t)), \end{aligned} \tag{3}$$

where $\hat{f}(t) \in H_X$ is the time-varying function estimate in the RKHS, $\mathcal{E}_{x(t)}$ denotes the evaluation operator at the state $x(t)$, and P is the solution to the Lyapunov equation $A^T P + PA = -Q$ for a user-defined positive definite $Q \in \mathbb{R}^{d\times d}$. When we assume $f \in H_X$, then the nonlinear term $f(x(t))$ is replaced with $\mathcal{E}_{x(t)} f$. Denote the estimation errors by $\tilde{x}(t) = x(t) - \hat{x}(t)$ and $\tilde{f}(t) = f - \hat{f}(t)$, then the error equations are as follows,

$$\begin{aligned} \dot{\tilde{x}}(t) &= A\tilde{x}(t) + B\mathcal{E}_{x(t)}\tilde{f}(t), \\ \dot{\tilde{f}}(t) &= -\gamma(B\mathcal{E}_{x(t)})^* P\tilde{x}(t), \end{aligned} \tag{4}$$

which evolves in the infinite dimensional space $\mathbb{R}^d \times H_X$. It has been proven that the equilibrium of the error equations at the origin is uniformly asymptotically stable under the PE condition [13], [14].

Finite dimensional approximation is needed for implementing the RKHS embedded estimator. Denote the approximation states by $(\hat{x}_n, \hat{f}_n) \in \mathbb{R}^d \times H_{\Omega_n}$. The approximant space H_{Ω_n} is a subspace having the form

$$H_{\Omega_n} = \text{span}\{\mathfrak{K}(\cdot, x) : x \in \Omega_n\} \subseteq H_X,$$

where $\Omega_n = \{\xi_i\}_{i=1}^n$ is a finite collection of kernel centers taken from the positive orbit $\Gamma^+(x_0) = \bigcup_{t\geq 0} x(t)$ of the uncertain system. As in [9], the governing equations of the finite dimensional estimator are

$$\begin{aligned} \dot{\hat{x}}_n(t) &= A\hat{x}_n(t) + B\mathcal{E}_{x(t)}\mathbf{P}^*_{\Omega_n}\hat{f}_n(t), \\ \dot{\hat{f}}_n(t) &= \gamma\mathbf{P}_{\Omega_n}(B\mathcal{E}_{x(t)})^* P(x(t) - \hat{x}_n(t)). \end{aligned} \tag{5}$$

Now compare the Eq. 3-5. We call $\tilde{f}_n(t) = f - \hat{f}_n(t)$ the *total error* of the function estimate. It is the summation of the *infinite dimensional error* $\tilde{f}(t) = f - \hat{f}(t)$ and the *approximation error* $\bar{f}_n(t) = \hat{f}(t) - \hat{f}_n(t)$. The corresponding notations $\tilde{x}(t)$ and $\bar{x}_n(t)$ are defined accordingly for the state errors. The infinite dimensional error $(\tilde{x}, \tilde{f})$ has been proven to converge in [13], and in this paper we focus on the approximation error $(\bar{x}_n, \bar{f}_n)$, which is characterized by the following evolution,

$$\dot{\bar{x}}_n(t) = A\bar{x}_n(t) + B\mathcal{E}_{x(t)}\bar{f}_n(t), \tag{6}$$

$$\begin{aligned}\dot{\bar{f}}_n(t) = &-\gamma \mathbf{P}_{\Omega_n}(B\mathcal{E}_{x(t)})^* P\bar{x}_n(t) \\ &+ \gamma(I - \mathbf{P}_{\Omega_n})(B\mathcal{E}_{x(t)})^* P\tilde{x}(t). \end{aligned}\tag{7}$$

C. Sobolev Spaces

Of particular importance to this paper is when the RKHS space is either equivalent to, or embedded in, a Sobolev space. The Sobolev space $W^{\tau,2}(\Omega)$ is the collection of functions in $L^2(\Omega, \mu)$ that have weak derivatives of all orders less than or equal to τ that are also in $L^2(\Omega, \mu)$, where $\Omega \subseteq X = \mathbb{R}^d$ is a (sufficiently regular) subset and τ is an integer. This definitions also makes sense when Ω is a smooth and compact manifold M, with the weak derivatives replaced by the covariant intrinsic derivatives on M and the Lebesgue measure μ replaced by the volume measure of the manifold M [19]. In general, boundary values or restrictions of functions in a Sobolev space are described by trace theorems. A specific trace theorem ([15], Proposition 2) is applied in this paper to study the restriction of $W^{\tau,2}(X)$ to subdomain Ω. When Ω is a k-dimensional smooth compact embedded manifold and $\tau > (d-k)/2$, we have the following identity

$$\mathbf{R}_\Omega W^{\tau,2}(X) \approx W^{\tau-(d-k)/2,2}(\Omega) \tag{8}$$

with equivalent norms. As described in [15], the intuition here is that an amount of "smoothness" is lost due to the restriction operation onto the low-dimensional submanifold.

Our interest in Sobolev spaces arises when we can show that a certain RKHS space is isometric with a Sobolev space. Let $\hat{\mathfrak{K}}$ be the Fourier transform of $\mathfrak{K}$. If $\hat{\mathfrak{K}}$ has the algebraic decay as

$$\hat{\mathfrak{K}}(\xi) \sim (1 + \|\xi\|_2)^{-\tau}, \quad \tau > d/2,$$

then the RKHS H_X induced by $\mathfrak{K}$ is a Sobolev space of the order τ [15], [18], that is, $H_X \approx W^{\tau,2}(X)$ with equivalent norms. Then by Theorem 5 in [15], the isometry also exists between the space of restrictions. When $\Omega \subseteq X$ is the submanifold described above, we have

$$\mathbf{R}_\Omega H_X \approx \mathbf{R}_\Omega W^{\tau,2}(X) \approx W^{\tau-(d-k)/2,2}(\Omega). \tag{9}$$

III. Main Results

In the remainder of this paper, we assume that Ω is a compact, connected, k-dimensional, regularly embedded Riemannian submanifold of $\mathbb{R}^d$, and that the submanifold Ω is invariant under the trajectory $t \mapsto x(t)$ of the uncertain system.

We now derive the explicit equations governing the evolution of error over Ω. Using the identity $\mathbf{E}_\Omega \mathbf{R}_\Omega = \mathbf{P}_\Omega$, it can be shown that the weak differential operator d/dt commutes with the restriction operator $\mathbf{R}_\Omega$, that is, for $h \in C^1([0,T], H_\Omega)$,

$$\mathbf{R}_\Omega \dot{h}(t) = \frac{d}{dt}(\mathbf{R}_\Omega h(t))$$

in the weak sense. Since $\mathbf{R}_\Omega$ is a bounded linear operator, and the evolution $t \mapsto \bar{f}(t)$ in Eq. 7 is strongly differentiable, we know that $t \mapsto (\mathbf{R}_\Omega f)(t)$ is also strongly differentiable. This means the derivatives on both sides are interpreted in the strong sense in the identity above. Now with this conclusion, we can apply $\mathbf{R}_\Omega$ to both sides of Eq. 7. The resulting error equation with respect to the restriction $\mathbf{R}_\Omega \bar{f}_n(t)$ is

$$\begin{aligned}\frac{d}{dt}\left(\mathbf{R}_\Omega \bar{f}_n(t)\right) &= \mathbf{R}_\Omega \dot{\bar{f}}_n(t) \\ &= -\gamma \mathbf{R}_\Omega \mathbf{P}_{\Omega_n}(B\mathcal{E}_{x(t)})^* P\bar{x}_n(t) \\ &\quad + \gamma \mathbf{R}_\Omega (I - \mathbf{P}_{\Omega_n})(B\mathcal{E}_{x(t)})^* P\tilde{x}(t). \end{aligned}\tag{10}$$

In order to analyze the approximation error in Eq. 10, we first review several results from [20], [21] about the Sobolev error bounds for scattered data interpolation. Let r and μ be the orders of two Sobolev spaces. Given $r > \mu$, it follows that $W^{r,2}(\Omega) \subseteq W^{\mu,2}(\Omega)$. Suppose the function $u \in W^{r,2}(\Omega)$ has a set of zero points Ω_n (i.e. $u|_{\Omega_n} = 0$) distributed densely enough in Ω. Theorem 1 characterizes the relationship between $\|u\|_{W^{r,2}(\Omega)}$ and $\|u\|_{W^{\mu,2}(\Omega)}$ in the term of the *fill distance* of Ω_n in Ω.

Definition 1 (Fill Distance [15]): For a finite set of discrete points $\Omega_n = \{\xi_i\}_{i=1}^n$ in a metric space Ω, the fill distance $h_{\Omega_n,\Omega}$ of Ω_n with respect to Ω is defined as

$$h_{\Omega_n,\Omega} := \sup_{x \in \Omega} \min_{\xi_i \in \Omega_n} d(x, \xi_i),$$

where $d(\cdot,\cdot)$ is the metric on Ω.
In the case which is of the most interest to this paper, the set Ω is a compact smooth Riemannian submanifold in $\mathbb{R}^d$, and the discrete set Ω_n is the set of interpolation points in the manifold. With this definition in mind, the following theorem states the relationship between $\|u\|_{W^{r,2}(\Omega)}$ and $\|u\|_{W^{\mu,2}(\Omega)}$.

Theorem 1 ([15]): Let $\Omega \subseteq \mathbb{R}^d$ be a smooth k-dimensional manifold, $r \in \mathbb{R}$ with $r > k/2$, $\mu \in \mathbb{N}_0$ with $0 \le \mu \le \lceil r \rceil - 1$.Then there is a constant h_Ω such that if the fill distance $h_{\Omega_n,\Omega} \le h_\Omega$ and $u \in W^{r,2}(\Omega)$ satisfies $u|_{\Omega_n} = 0$, then

$$\|u\|_{W^{\mu,2}(\Omega)} \lesssim h_{\Omega_n,\Omega}^{r-\mu} \|u\|_{W^{r,2}(\Omega)}.$$

We denote the interpolation operator over Ω_n by $\mathbf{I}_{\Omega_n}$. For a function $f \in W^{r,2}(\Omega)$, the interpolation error $(I - \mathbf{I}_{\Omega_n})f$ by definition has zeros over Ω_n. A corollary is introduced in [15] to characterize the decaying rate of interpolation error in the native space (i.e. RKHS).

Corollary 1: Let $\Omega \subseteq X := \mathbb{R}^d$ be a k-dimensional smooth manifold, and let the native space H_X be continuously embedded in a Sobolev space $W^{\tau,2}(X)$ with $\tau > d/2$, so that $\|f\|_{W^{\tau,2}(\mathbb{R}^d)} \lesssim \|f\|_{H_X}$. Define $s = \tau - (d-k)/2$ and let $0 \le \mu \le \lceil s \rceil - 1$. Then there is a constant h_Ω such that if $h_{\Omega_n,\Omega} \le h_\Omega$, then for all $f \in \mathbf{R}_\Omega(H_X)$ we have

$$\|(I - \mathbf{I}_{\Omega_n})f\|_{W^{\mu,2}(\Omega)} \lesssim h_{\Omega_n,\Omega}^{s-\mu} \|f\|_{\mathbf{R}_\Omega(H_X)}.$$

With this error bound in mind for interpolation and projection, we now turn to the main result of this paper.

Theorem 2: Suppose that Ω is a compact, connected, k-dimensional, regularly embedded Riemannian submanifold of $X := \mathbb{R}^d$, $\mathbf{R}_\Omega(H_X) \hookrightarrow W^{s,2}(\Omega)$ for some $s > k/2$, and the orbit of the unknown system $\Gamma^+(x_0) \subseteq \Omega$. Then there exist two constants $a, b > 0$ such that for all $t \in [0, T]$,

$$\begin{aligned}
&\|\bar{x}_n(t)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{s,2}(\Omega)} \\
&\quad \leq e^{bt}\bigg(\|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \hat{f}_0\|^2_{W^{s,2}(\Omega)} \\
&\qquad + a\int_0^t \|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \mathfrak{K}_{x(\zeta)}\|^2_{W^{s,2}(\Omega)} d\zeta\bigg).
\end{aligned}$$

Here $\Pi_{\Omega_n} : \mathbf{R}_\Omega H_X \to \mathbf{R}_\Omega H_{\Omega_n}$ denotes the projection operator defined over the space of restrictions $\mathbf{R}_\Omega H_X$ onto the space $\mathbf{R}_\Omega H_{\Omega_n}$.

Proof: By Eq. 6 and 10 we have

$$\begin{aligned}
&\frac{d}{dt}\left(\|\bar{x}_n(t)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{s,2}(\Omega)s}\right) = (\dot{\bar{x}}_n(t), \bar{x}_n(t)) \\
&\quad + (\bar{x}_n(t), \dot{\bar{x}}_n(t)) + 2\left(\mathbf{R}_\Omega \bar{f}_n(t), \frac{d}{dt}(\mathbf{R}_\Omega \bar{f}_n(t))\right)_{W^{s,2}(\Omega)} \\
&= ((A + A^T)\bar{x}_n(t), \bar{x}_n(t)) + \underbrace{2(B\mathcal{E}_{x(t)}\bar{f}_n(t), \bar{x}_n(t))}_{\text{term 1}} \\
&\quad + \underbrace{2\gamma\left(\mathbf{R}_\Omega(I - \mathbf{P}_{\Omega_n})\mathcal{E}^*_{x(t)}B^T P\tilde{x}(t), \mathbf{R}_\Omega \bar{f}_n(t)\right)_{W^{s,2}(\Omega)}}_{\text{term 2}} \\
&\quad + \underbrace{2\gamma\left(-\mathbf{R}_\Omega \mathbf{P}_{\Omega_n}\mathcal{E}^*_{x(t)}B^T P\bar{x}_n(t), \mathbf{R}_\Omega \bar{f}_n(t)\right)_{W^{s,2}(\Omega)}}_{\text{term 3}}.
\end{aligned}$$

In several of the steps that follow, we use conclusions that result from the assumptions that certain spaces are continuously embedded in others. By choosing certain types of reproducing kernels, it can be guaranteed that the injection $j : \mathbf{R}_\Omega(H_X) \hookrightarrow W^{s,2}(\Omega)$ is a continuous embedding for some $s > k/2$. By the Sobolev embedding theorem the injection $i : W^{s,2}(\Omega) \hookrightarrow C(\Omega)$ is also continuous, which implies

$$|\mathcal{E}_{s,x}(f)| := |f(x)| \leq \|if\|_{C(\Omega)} \leq \|i\|\|f\|_{W^{s,2}(\Omega)},$$

and it follows that each evaluation functional $\mathcal{E}_{s,x} : W^{s,2}(\Omega) \to \mathbb{R}$ is uniformly bounded by $\|i\|$. By assumption we have that the forward orbit $\Gamma^+(x_0) \subseteq \Omega$, and we conclude that term 1 can be bounded by the expression

$$\begin{aligned}
&|(B\mathcal{E}_{x(t)}\mathbf{R}_\Omega \bar{f}_n(t), \bar{x}_n(t))| \\
&\quad \leq \|B\|\|i\|\|\bar{x}_n(t)\|\|\mathbf{R}_\Omega \bar{f}_n(t)\|_{W^{s,2}(\Omega)}. \qquad (11)
\end{aligned}$$

We bound term 2 by

$$\begin{aligned}
&\left|\left(\mathbf{R}_\Omega(I - \mathbf{P}_{\Omega_n})\mathcal{E}^*_{x(t)}B^T P\tilde{x}(t), \mathbf{R}_\Omega \bar{f}_n(t)\right)_{W^{s,2}(\Omega)}\right| \\
&\quad \leq \|B\|\|P\|\|\tilde{x}(t)\|\|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \mathfrak{K}_{x(t)}\|_{W^{s,2}(\Omega)} \\
&\qquad \|\mathbf{R}_\Omega \bar{f}_n(t)\|_{W^{s,2}(\Omega)}. \qquad (12)
\end{aligned}$$

We next consider term 3, which satisfies the inequality

$$\begin{aligned}
&\left|\left(\mathbf{R}_\Omega \mathbf{P}_{\Omega_n}\mathcal{E}^*_{x(t)}B^T P\bar{x}_n(t), \mathbf{R}_\Omega \bar{f}_n(t)\right)_{W^{s,2}(\Omega)}\right| \\
&\quad \leq \bar{k}\|j\|\|\mathbf{R}_\Omega\|\|B\|\|P\|\|\bar{x}_n(t)\| \\
&\qquad \times \|\mathbf{R}_\Omega \bar{f}_n(t)\|_{W^{s,2}(\Omega)}. \qquad (13)
\end{aligned}$$

Combining all the terms above, we obtain

$$\begin{aligned}
&\frac{d}{dt}\left(\|\bar{x}_n(t)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{s,2}(\Omega)}\right) \\
&\quad \leq \gamma\|B\|^2\|P\|^2\|\tilde{x}(t)\|^2\|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \mathfrak{K}_{x(t)}\|^2_{W^{s,2}(\Omega)} \\
&\qquad + \big(2\|A\| + \|i\|\|B\|\big)\|\bar{x}_n(t)\|^2 \\
&\qquad + \big(\gamma\bar{k}^2\|j\|^2\|\mathbf{R}_\Omega\|^2\|B\|^2\|P\|^2\big)\|\bar{x}_n(t)\|^2 \\
&\qquad + \big(2\gamma + \|i\|\|B\|\big)\|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{s,2}(\Omega)}.
\end{aligned}$$

Let the constants a, b be defined as follows.

$$\begin{aligned}
a &:= \gamma\|B\|^2\|P\|^2 \sup_{\zeta\in[0,T]}\|\tilde{x}(\zeta)\|^2, \\
b &:= \max\{2\gamma + \|i\|\|B\|, \\
&\qquad 2\|A\| + \|i\|\|B\| + \gamma\bar{k}^2\|j\|^2\|\mathbf{R}_\Omega\|^2\|B\|^2\|P\|^2\}.
\end{aligned}$$

When we integrate the inequality above, it follows that

$$\begin{aligned}
&\|\bar{x}_n\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{s,2}(\Omega)} \\
&\quad \leq \|\bar{x}_n(0)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(0)\|^2_{W^{s,2}(\Omega)} \\
&\qquad + \int_0^t a\|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \mathfrak{K}_{x(\zeta)}\|_{W^{s,2}(\Omega)} d\zeta \\
&\qquad + \int_0^t \big(\|\bar{x}(t)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(\zeta)\|^2_{W^{s,2}(\Omega)}\big) d\zeta.
\end{aligned}$$

But since $\hat{f}_0 \in H_\Omega$, we know that $\mathbf{P}_\Omega \hat{f}_0 = \mathbf{E}_\Omega \mathbf{R}_\Omega \hat{f}_0 = \hat{f}_0$, and $\mathbf{R}_\Omega \Pi_{\Omega_n}\hat{f}_0 = \Pi_{\Omega_n}\mathbf{R}_\Omega \hat{f}_0$. Thus

$$\|\mathbf{R}_\Omega \bar{f}_n(0)\|_{W^{s,2}(\Omega)} = \|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \hat{f}_0\|_{W^{s,2}(\Omega)}.$$

Combining the above inequalities yields

$$\begin{aligned}
&\|\bar{x}_n(t)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{s,2}(\Omega)} \\
&\quad \leq e^{bt}\|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \hat{f}_0\|^2_{W^{s,2}(\Omega)} \\
&\qquad + ae^{bt}\int_0^t \|(I - \Pi_{\Omega_n})\mathbf{R}_\Omega \mathfrak{K}_{x(\zeta)}\|^2_{W^{s,2}(\Omega)} d\zeta.
\end{aligned}$$

■

The next corollary combines the results of Theorem 2 and Corollary 1 to obtain the error rates in terms of the fill distance of samples in the manifold.

Corollary 2: Suppose that Ω is a compact, connected, regularly embedded k-dimensional submanifold of $X := \mathbb{R}^d$, the kernel $\mathfrak{K}$ is selected so that $H_X \hookrightarrow W^{\tau,2}(X)$ for $\tau > d/2$, define $s := \tau - (d-k)/2$, and let $\mu \in [k/2, \lceil s\rceil - 1]$. Then we have

$$\begin{aligned}
&\|\bar{x}_n(t)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{\mu,2}(\Omega)} \\
&\qquad \leq \left(\|\mathbf{R}_\Omega \hat{f}_0\|^2_{\mathbf{R}_\Omega(H_X)} + a\bar{k}^2 t\right) e^{\tilde{b}t} h^{2(s-\mu)}_{\Omega,\Omega_n}.
\end{aligned}$$

Proof: We first observe that under the stated hypotheses the native space $\mathbf{R}_\Omega H_X \hookrightarrow W^{\mu,2}(\Omega)$. For any two positive $r_1 \geq r_2 > 0$ the associated Sobolev spaces are a continuous scale of spaces with $W^{r_1,2}(\Omega) \hookrightarrow W^{r_2,2}(\Omega)$. This implies that $W^{s,2}(\Omega) \hookrightarrow W^{\mu,2}(\Omega)$ since $s \geq \mu$. Also, the trace theorem yields

$$\begin{aligned}
\|f\|_{W^{\mu,2}(\Omega)} &\lesssim \|f\|_{W^{s,2}(\Omega)} = \|\mathbf{E}_\Omega \mathbf{R}_\Omega f\|_{W^{s,2}(\Omega)} \\
&\lesssim \|\mathbf{E}_\Omega f\|_{W^{\tau,2}(X)} \lesssim \|f\|_{\mathbf{R}_\Omega H_X},
\end{aligned}$$

where the constants in the above string of inequalities depend on $\|\mathbf{E}_\Omega\|$, $\|\mathbf{R}_\Omega\|$, and the norm of the embedding of H_X into $W^{\tau,2}(X)$. Combining these results yields $\mathbf{R}_\Omega H_X \hookrightarrow W^{\mu,2}(\Omega)$ with $\mu \geq k/2$. We can now apply the results of Theorem 2 for the choice $s=\mu$ and write

$$\begin{aligned}
&\|\bar{x}_n(t)\|^2 + \|\mathbf{R}_\Omega \bar{f}_n(t)\|^2_{W^{\mu,2}(\Omega)} \\
&\quad\leq e^{bt}\|(I-\Pi_{\Omega_n})\mathbf{R}_\Omega \hat{f}_0\|^2_{W^{\mu,2}(\Omega)} \\
&\qquad + ae^{bt}\|(I-\Pi_{\Omega_n})\mathbf{R}_\Omega \mathfrak{K}_{x(t)}\|^2_{W^{\mu,2}(\Omega)}, \\
&\quad\leq e^{bt} h^{2(s-\mu)}_{\Omega,\Omega_n}\|\mathbf{R}_\Omega \hat{f}_0\|^2_{\mathbf{R}_\Omega H_X} \\
&\qquad + ate^{bt} h^{2(s-\mu)}_{\Omega,\Omega_n}\left(\sup_{t\in[0,T]}\|\mathbf{R}_\Omega \mathfrak{K}_{x(t)}\|_{\mathbf{R}_\Omega H_X}\right)^2
\end{aligned}$$

for each $t \in [0,T]$ by Theorem 11 of [15]. The bound now follows since

$$\sup_{t\in[0,T]}\|\mathbf{R}_\Omega \mathfrak{K}_{x(t)}\|_{\mathbf{R}_\Omega H_X} = \sup_{t\in[0,T]}\|\mathfrak{K}_{x(t)}\|_{H_X} \leq \bar{k}.$$

■

IV. Numerical Simulation

Corollary 2 gives a rate of convergence for finite dimensional approximations of the RKHS embedded estimator. The rate of convergence depends on the density of the sample set Ω_n in the manifold Ω, which is characterized by the fill distance $h_{\Omega_n,\Omega}$. In this section, this rate of convergence is illustrated qualitatively using numerical simulations. Following the formulation of Eq. 2, the governing equations of the unknown system are selected to be

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + \begin{bmatrix}x_1^2(t)\\ 0\end{bmatrix}, \tag{14}$$

where $B=[1,0]^T$ and $f(x_1,x_2)=x_1^2$. Here we assume the linear coefficient matrix A_0 is known. By adding and subtracting a selected Hurwitz term $Ax(t)$, the governing equations of the original system can be written as

$$\dot{x}(t) = Ax(t) + (A_0 - A)x(t) + B\mathcal{E}_{x(t)}f.$$

Since A_0 and A are known, the term $(A_0-A)x(t)$ can be canceled in the error equation. The governing equations of the finite dimensional RKHS embedded estimator are chosen as

$$\begin{aligned}
\dot{\hat{x}}_n(t) &= A\hat{x}_n(t) + (A_0-A)x(t) + B\mathcal{E}_{x(t)}\mathbf{P}^*_\Omega \hat{f}_n(t), \\
\dot{\hat{f}}_n(t) &= \gamma \mathbf{P}_\Omega (B\mathcal{E}_{x(t)})^* P(x(t)-\hat{x}_n(t)).
\end{aligned}$$

This choice yields the error equations that have the form studied in this paper.

The phase portraits of the original system in Eq. 14 are shown in Fig. 1. A first integral of the unknown system is

$$\Phi(x) := (x_2 + x_1^2 - 0.5)e^{2x_2} = c. \tag{15}$$

The stability of the system depends on the initial condition $(x_1(0), x_2(0))$. When the initial condition $x(0)$ is such that the constant $c<0$, the system is stable and the positive orbit $\Gamma^+(x(0))$ itself is an invariant manifold $\Omega := \{x \in \mathbb{R}^2 : \Phi(x) = c\}$. The manifold Ω is a smooth, one dimensional, regularly embedded submanifold in the phase space $\mathbb{R}^2$. In this example, we choose the trajectory for of which the constant $c=-0.1$ in Eq. 15. The samples $\Omega_n = \{\xi_i\}_{i=1}^N$ are taken uniformly along the manifold with respect to the intrinsic metric of the manifold Ω. Although in practice, such sampling procedure cannot be accomplished without knowing the manifold *a priori*, it is not difficult to picture that as $t\to\infty$, the set $\{x(t_i)\}_{i=1}^N$ gradually fills the manifold Ω. The samples are used to construct the approximant RKHS $H_{\Omega_n} := \text{span}\{\mathfrak{K}(\cdot,x) : x \in \Omega_n\}$. The Sobolev-Matérn kernel $\mathfrak{K}_\nu$ is used to induce the RKHS. The subscript ν denotes the order of the kernel. If $\nu > d/2$, then all the functions in the RKHS H_X induced by $\mathfrak{K}_\nu$ over $X=\mathbb{R}^d$ also belong to every Sobolev space $W^{\tau,2}(\mathbb{R}^d)$ where $\tau > 2\nu - d/2$ [15]. The general expression of $\mathfrak{K}_\nu$ is defined using a Bessel function, but when $\nu = p + 1/2$, $p\in\mathbb{N}$ the kernel has the following closed-form expressions

$$\begin{aligned}
\mathfrak{K}_{3/2}(x,y) &= \left(1+\frac{\sqrt{3}r}{l}\right)\exp\left(-\frac{\sqrt{3}r}{l}\right), \\
\mathfrak{K}_{5/2}(x,y) &= \left(1+\frac{\sqrt{5}r}{l}+\frac{5r^2}{3l^2}\right)\exp\left(-\frac{\sqrt{5}r}{l}\right),
\end{aligned}$$

where $r=\|x-y\|$, and l is the scaling factor of length [22].

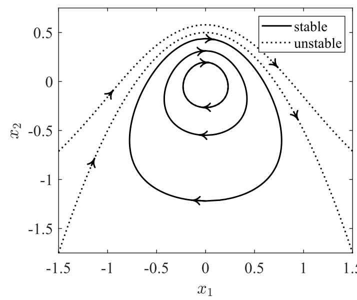

Fig. 1: Phase trajectories of the actual system.

Fig. 2 shows the contour of the estimation error $|f(x) - \hat{f}_n(x)|$ in $\mathbb{R}^d$ using when $N=100$ and $\nu=5/2$. The result is as expected. The estimate of error in the unknown function is close to zero along the manifold $\Omega = \{x\in\mathbb{R}^2 : \Phi(x) = -0.1\}$. The rate of convergence with respect to the number of samples N is shown in Fig. 3-4. Note that the manifold Ω is a closed curve, and the samples are taken uniformly in metric. As a result, the fill distance $h_{\Omega_n,\Omega} \sim N^{-1}$. With this in mind, by Corollary 2 we have the following relationship

$$\|\mathbf{R}_\Omega(\hat{f}(t) - \hat{f}_n(t))\|_{W^{\mu,2}(\Omega)} \sim N^{-(s-\mu)}.$$

In this example, the set Ω is PE, so $\hat{f}(t)\to f$ over Ω. The RKHS $H_X \hookrightarrow W^{\tau,2}(\mathbb{R}^2)$ where $\tau < 2\nu - 1$. Thus the space of restrictions $\mathbf{R}_\Omega H_X \hookrightarrow W^{\tau-0.5,2}(\Omega)$, and $s \leq \tau - 0.5 < 2\nu - 1.5$. On the other hand, we must have $\mu \in [k/2, \lceil s\rceil - 1]$ so that $W^{s,2}(\Omega) \hookrightarrow W^{\mu,2}(\Omega) \hookrightarrow C(\Omega)$. In this way, we have

$$\|f - \hat{f}_n(t)\|_{C(\Omega)} = \sup_{x\in\Omega}|f(x) - (\hat{f}_n(t))(x)| \sim N^{-(s-\mu)}.$$

From the analysis above, we obtain the rates of convergence for the $C(\Omega)$-norm. When $\nu = 3/2$, the order $s-\mu \geq 1$. When $\nu = 5/2$, the order $s-\mu\geq 2$. Taking the logarithms for both sides of the equation above, the values calculated above

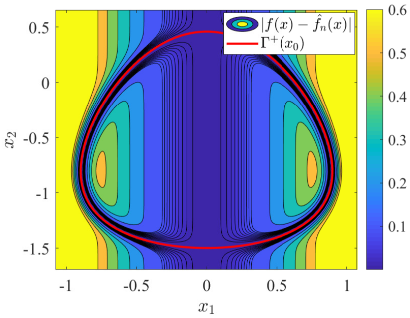

Fig. 2: Contour of function error

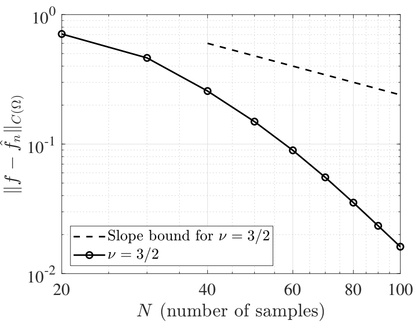

Fig. 3: Conv. rate for $\nu = 3/2$.

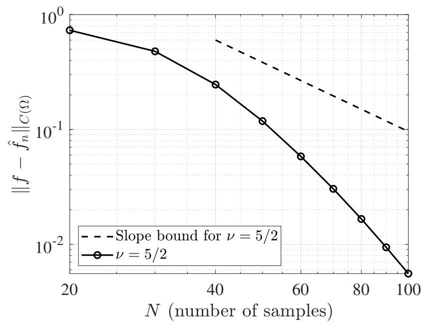

Fig. 4: Conv. rate for $\nu = 5/2$.

are the worst case of slope bounds in Fig. 3-4. In both figures, the actual error curves are below the slope bounds, which validates the conclusions in Corollary 2. One assumption for the Theorem 1 to hold is that $h_{\Omega_n,\Omega}$ must be smaller than a threshold. This assumption may explain the flat error curve when $N \leq 30$.

V. Conclusions

This paper considers the practical problem of formulating finite dimensional approximations for the RKHS embedded apative estimator. The convergence of approximations is proven, and the rates of convergence are derived. By selecting the reproducing kernel that has algebraic decaying Fourier transform, the induced RKHS is embedded in or equivalent to a Sobolev space. The error equation of approximation is recast in the Sobolev space, and bounds on the error of interpolation in Sobolev spaces are applied to analyse the error of approximation. When the trajectory of the unknown system concentrates in a compact, regularly embedded submanifold of the state space, the rate of convergence for finite dimensional approximation is derived in terms of the fill distance of the samples. It is shown that as the samples becomes increasingly dense in the submanifold, the approximation error decays accordingly.

References

[1] S. Klus, F. Nüske, P. Koltai, H. Wu, I. Kevrekidis, C. Schütte, and F. Noé, “Data-driven model reduction and transfer operator approximation,” *Journal of Nonlinear Science*, vol. 28, no. 3, pp. 985–1010, 2018.

[2] S. Pan and K. Duraisamy, “Data-driven discovery of closure models,” *SIAM Journal on Applied Dynamical Systems*, vol. 17, no. 4, pp. 2381–2413, 2018.

[3] D. Giannakis, “Data-driven spectral decomposition and forecasting of ergodic dynamical systems,” *Applied and Computational Harmonic Analysis*, vol. 47, no. 2, pp. 338–396, 2019.

[4] M. Korda and I. Mezić, “On convergence of extended dynamic mode decomposition to the Koopman operator,” *Journal of Nonlinear Science*, vol. 28, no. 2, pp. 687–710, 2018.

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

[6] C. L. P. Chen, G. X. Wen, Y. J. Liu, and Z. Liu, “Observer-based adaptive backstepping consensus tracking control for high-order nonlinear semi-strict-feedback multiagent systems,” *IEEE Transactions on Cybernetics*, vol. 46, no. 7, pp. 1591–1601, 2016.

[7] O. San, R. Maulik, and M. Ahmed, “An artificial neural network framework for reduced order modeling of transient flows,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 77, pp. 271–287, 2019.

[8] L. Györfi, M. Kohler, A. Krzyzak, and H. Walk, *A distribution-free theory of nonparametric regression*. Springer Science & Business Media, 2006.

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

[10] ——, “Adaptive estimation for nonlinear systems using reproducing kernel Hilbert spaces,” *Advances in Computational Mathematics*, vol. 45, no. 2, pp. 869–896, 2019.

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

[12] A. Kurdila and Y. Lei, “Adaptive control via embedding in reproducing kernel Hilbert spaces,” in *2013 American Control Conference*. IEEE, 2013, pp. 3384–3389.

[13] J. Guo, S. T. Paruchuri, and A. J. Kurdila, “Persistence of excitation in uniformly embedded reproducing kernel Hilbert (RKH) spaces,” *arXiv preprint arXiv:2002.07963*, 2020.

[14] A. J. Kurdila, J. Guo, S. T. Paruchuri, and P. Bobade, “Persistence of excitation in reproducing kernel Hilbert spaces, positive limit sets, and smooth manifolds,” *arXiv preprint arXiv:1909.12274*, 2019.

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

[16] T. Hangelbroek, F. J. Narcowich, and J. D. Ward, “Kernel approximation on manifolds i: bounding the Lebesgue constant,” *SIAM Journal on Mathematical Analysis*, vol. 42, no. 4, pp. 1732–1760, 2010.

[17] T. Hangelbroek, F. J. Narcowich, X. Sun, and J. D. Ward, “Kernel approximation on manifolds ii: The L_∞ norm of the L_2 projector,” *SIAM journal on mathematical analysis*, vol. 43, no. 2, pp. 662–684, 2011.

[18] A. Berlinet and C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in probability and statistics*. Springer Science & Business Media, 2011.

[19] R. A. Adams and J. J. Fournier, *Sobolev spaces*. Elsevier, 2003.

[20] F. Narcowich, J. Ward, and H. Wendland, “Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting,” *Mathematics of Computation*, vol. 74, no. 250, pp. 743–763, 2005.

[21] R. Arcangéli, M. C. L. de Silanes, and J. J. Torrens, “An extension of a bound for functions in Sobolev spaces, with applications to (m, s)-spline interpolation and smoothing,” *Numerische Mathematik*, vol. 107, no. 2, pp. 181–211, 2007.

[22] C. E. Rasmussen, “Gaussian processes in machine learning,” in *Summer School on Machine Learning*. Springer, 2003, pp. 63–71.