State Following (StaF) Kernel Functions for Function Approximation[☆]

Joel A. Rosenfeld[a,*], Rushikesh Kamalapurkar[a], Warren E. Dixon[a]

[a] *Department of Mechanical and Aerospace Engineering, MAE-B, University of Florida, Gainesville, FL*

Abstract

A function approximation method is developed that aims to approximate a function in a small neighborhood of a state that travels within a compact set. The development is based on the theory of universal reproducing kernel Hilbert spaces over the n-dimensional Euclidean space. Several theorems are introduced that support the development of this State Following (StaF) method. In particular, it is shown that there is a bound on the number of kernel functions required for the maintenance of an accurate function approximation as a state moves through a compact set. Additionally, a weight update law, based on gradient descent, is introduced where arbitrarily close accuracy can be achieved provided the weight update law is iterated at a sufficient frequency, as detailed in Theorem 6.1.

To illustrate the advantage, the impact of the StaF method is that for some applications the number of basis functions can be reduced. The StaF method is applied to an adaptive dynamic programming (ADP) application to demonstrate that stability is maintained with a reduced number of basis functions.

Simulation results demonstrate the utility of the StaF methodology for the maintenance of accurate function approximation as well as solving an infinite horizon optimal regulation problem through ADP. The results of the simulation indicate that fewer basis functions are required to guarantee stability and approximate optimality than are required when a global approximation approach is used.

[☆] A preliminary version of this work appeared in the proceedings of the 2015 American Control Conference (ACC) [1]

[*] Corresponding author

Email address: joelar@ufl.edu (Joel A. Rosenfeld)

1. Introduction

Often in the theory of approximation, an accurate estimation of a function over a large compact set is sought [2, 3, 4]. It is well known that the larger the compact set, a correspondingly larger number of basis functions are required to achieve an accurate function approximation. There is a large body of literature concerned with methods for the reduction of the number of basis functions required to achieve such an approximation (c.f. [5, 6, 7]).

In many control applications, function approximation is used to generate a stabilizing controller of a state in a dynamical system. For instance, in adaptive dynamic programming (ADP), an approximation of the optimal value function is leveraged to produce an approximate optimal controller [8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. Traditionally, the approximation is sought over a large compact set, and requires many basis functions. The computational resources required to tune the weights of the basis functions renders real-time implementation of controllers based on ADP methods infeasible.

Motivated by problems in control theory, this paper introduces an approximation methodology that aims to establish and maintain an accurate approximation of a function in a neighborhood of a moving state in a dynamical system. The method, deemed the state following (StaF) method, reduces the number of basis functions required to achieve an accurate approximation by focusing on the approximation of a function over a small neighborhood by linear combinations of time and state varying basis functions. Therefore, even in cases where processing power of on-board CPUs is limited, an accurate approximation of a function can be maintained.

The particular basis functions that will be employed throughout this paper are derived from kernel functions corresponding to RKHSs. In particular, the centers are selected to be continuous functions of the state variable bounded by a predetermined value. That is, given a compact set $D \subset \mathbb{R}^n$, $\epsilon > 0$, $r > 0$ and $M \in \mathbb{N}$, $c_i(x) = x + d_i(x)$ where $d_i : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable and $\sup_{x \in D} \|d_i(x)\| < r$ for $i = 1, ..., M$. The parameterization of a function $V : D \to \mathbb{R}$ in terms of StaF kernel functions is given by

$$\hat{V}(y; x(t), t) = \sum_{i=1}^{M} w_i(t) K(y, c_i(x(t)))$$

where $w_i(t)$ is a weight signal chosen to satisfy

$$\limsup_{t \to \infty} E_r(x(t), t) < \epsilon$$

where E_r is a measure of the accuracy of an approximation in a neighborhood of $x(t)$, such as that of the supremum norm:

$$E_r(x(t),t) = \sup_{y \in \overline{N_r(x(t))}} \left| V(y) - \hat{V}(y; x(t), t) \right|.$$

The goal of the StaF method is to establish and maintain an approximation of a function in a neighborhood of the state. The justification for this approach stems from the observation that an optimal controller only requires the value of the estimation of the optimal value function to be accurate at the current system state. Thus, when computational resources are limited, computational efforts should be focused on improving the accuracy of approximations near the system state.

The advantage of using RKHSs for the purpose of local approximations is twofold. RKHSs have been found to be effective for nonlinear function approximation [18], and the use of RKHS can enable accurate estimations of a wide array of nonlinear functions. Also, the ideal weights corresponding to the Hilbert space norm provided by RKHSs change smoothly with respect to smooth changes in the centers, as demonstrated in Theorem 5.1, which allows the execution of weight update laws to achieve and maintain an accurate approximation. The ideal weights in the context of the StaF approximation method become a continuous function of the state and are investigated in Section 5.

Previous efforts in the literature have performed nonlinear approximation through the adjustment of the centers of radial basis functions (c.f. [19, 20, 21]) as a means to determine the optimal centers for global approximation. These efforts are more applicable when off-line techniques can be used due to computational demands. For other applications where computational resources are limited, global approximations may not be feasible (especially as the dimension of the problem grows), nor is the optimal selection of parameters.

This paper lays the foundation for the establishment and maintenance of a real-time moving local approximation of a continuous function. Section 2 of this paper frames the particular approximation problem of the StaF method. Section 3 demonstrates accurate approximation with a fixed number of moving basis functions. Section 4 demonstrates an explicit bound on the number of required StaF basis functions for the case of the exponential kernel function. The ideal weight function arising from the StaF method is introduced and discussed in Section 5, where the existence and smoothness of the ideal weight function is established. Section 6 provides a proof of concept demonstrating the existence of weight update laws to

maintain an accurate approximation of a function in a local neighborhood, ultimately establishing a uniform ultimate bounded result. The remaining sections demonstrate the developed method through numerical experiments and discussions of applications. Specifically, Section 7 gives the results of a "gradient chase" algorithm. In Section 8, the utility of StaF methods are demonstrated in an ADP application.

2. The StaF Problem Statement

Given a continuous function $V : \mathbb{R}^n \to \mathbb{R}$, $r > 0$, an arbitrarily small $\epsilon > 0$, and a dynamical system $\dot{x} = f(x, u)$ (where f is sufficiently regular for the system to be well defined), the goal of the StaF approximation method is to select state and time varying basis functions $\sigma_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ for $i = 1, 2, ..., M$ and weight signals $w_i : \mathbb{R}_+ \to \mathbb{R}$ for $i = 1, 2, ..., M$ such that

$$\limsup_{t\to\infty} \sup_{y\in\overline{N_r(x(t))}} \left| V(y) - \sum_{i=1}^{M} w_i(t)\sigma_i(y; x(t), t) \right| < \epsilon. \tag{1}$$

In other words, the StaF approximation method aims to achieve an arbitrarily small steady state error of order ϵ in a closed neighborhood of the state, $\overline{N_r(x(t))} = \{y \in \mathbb{R}^n : \|x(t) - y\|_2 \leq r\}$.

Central problems to the StaF method include determining the basis functions and the weight signals. When reproducing kernel Hilbert spaces are used for basis functions, (1) can be relaxed to where the supremum norm is replaced with the Hilbert space norm. Since the Hilbert space norm of a RKHS dominates the supremum norm, (1) with the supremum norm is simultaneously satisfied. Moreover, when using a RKHS, the basis functions can be selected to correspond to centers placed in a moving neighborhood of the state. In particular, given a kernel function $K : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ corresponding to a (universal) RKHS, H, and continuous center functions $c_i : \mathbb{R}^n \to \mathbb{R}^n$ for which $d_i(x) := c_i(x) - x$ is bounded by r, then the StaF problem becomes the determination of weight signals $w_i : \mathbb{R}_+ \to \mathbb{R}$ for $i = 1, ..., M$ such that:

$$\limsup_{t\to\infty} \left\| V(\cdot) - \sum_{i=1}^{M} w_i(t)K(\cdot, c_i(x(t))) \right\|_{r,x(t)} < \epsilon \tag{2}$$

where $\|\cdot\|_{r,x(t)}$ is the norm of the RKHS obtained by restricting functions in H to $N_r(x(t))$.

Since (2) implies (1), the focus of this paper is to demonstrate the feasibility of satisfying (2). Theorem 3.1 demonstrates that under a certain

continuity assumption a bound on the number of kernel functions necessary for the maintenance of an approximation throughout a compact set can be determined, and Theorem 5.1 shows that a collection of continuous ideal weight functions can be determined to satisfy (2). Theorem 5.1 justifies the use of weight update laws for the maintenance of an accurate function approximation, and this is demonstrated by Theorem 6.1 as well as the numerical results contained in Section 7 and Appendix A.

The choice of RKHS for Section 7 is that which corresponds to the exponential kernel $K(x, y) = \exp(x^T y)$ where $x, y \in \mathbb{R}^n$ and will be denoted by $F^2(\mathbb{R}^n)$ since it is closely connected to the Bargmann-Fock space [23]. The RKHS corresponding to the exponential kernel is a universal RKHS [24, 25], which means that given any compact set $D \subset \mathbb{R}^n$, $\epsilon > 0$ and continuous function $f : D \to \mathbb{R}$, there exists a function $\hat{f} \in F^2(\mathbb{R}^n)$ for which $\sup_{x \in D} |f(x) - \hat{f}(x)| < \epsilon$.

3. Feasibility of the StaF Approximation and the Ideal Weight Functions

The first theorem concerning the StaF method demonstrates that if the state variable is constrained to a compact subset of $\mathbb{R}^n$, then there is a finite number of StaF basis functions required to establish the accuracy of an approximation.

Theorem 3.1. *Suppose that $K : X \times X \to \mathbb{C}$ is a continuous kernel function corresponding to a RKHS, H, over a set X equipped with a metric topology. If $V \in H$, D is a compact set of X, $r > 0$, and $\|V\|_{x,r}$ is continuous with respect to x, then for all $\epsilon > 0$ there is a $M \in \mathbb{N}$ such that for each $x \in D$ there are centers $c_1, c_2, ..., c_M \in N_r(x)$ and weights $w_i \in \mathbb{C}$ such that*

$$\left\| V(\cdot) - \sum_{i=1}^{M} w_i K(\cdot, c_i) \right\|_{r,x} < \epsilon.$$

Proof. Given $\epsilon > 0$, for each neighborhood $N_r(x)$ with $x \in D$, there exists a finite number of centers $c_1, ..., c_M \in N_r(x)$, and weights $w_1, ..., w_M \in \mathbb{C}$, such that

$$\left\| V(\cdot) - \sum_{i=1}^{M} w_i K(\cdot, c_i) \right\|_{r,x} < \epsilon.$$

Let $M_{x,\epsilon}$ be the minimum such number. The claim of the proposition is that the set $Q_\epsilon := \{M_{x,\epsilon} : x \in D\}$ is bounded. Assume by way of contradiction

that Q_ϵ is unbounded, and take a sequence $\{x_n\} \subset D$ such that $M_{x_n,\epsilon}$ is a strictly increasing sequence and $x_n \to x$ in D. It is always possible to find such a convergent sequence, since every compact subset of metric space is sequentially compact. Let $c_1, ..., c_{M_{x,\epsilon/2}} \in N_r(x)$ and $w_1, ..., w_{M_{x,\epsilon/2}} \in \mathbb{C}$ be centers and weights for which

$$E(x) := \left\| V(\cdot) - \sum_{i=1}^{M_{x,\epsilon/2}} w_i K(\cdot, c_i) \right\|_{r,x} < \epsilon/2. \tag{3}$$

For convenience, let each $c_i \in N_r(x)$ be expressed as $x + d_i$ for $d_i \in N_r(0)$. The function $E(x)$ in (3) can be written as

$$\left(\|V\|_{r,x} - 2Re \left(\sum_{i=1}^{M_{x,\epsilon/2}} w_i V(x + d_i) \right) + \sum_{i,j=1}^{M_{x,\epsilon/2}} w_i \overline{w_j} K(x + d_i, x + d_j) \right)^{1/2}.$$

By the hypothesis, K is continuous with respect to x, which implies that V is continuous [2], and $\|V\|_{r,x}$ is continuous with respect to x. Hence, there exists $\eta > 0$ for which $|E(x) - E(x_n)| < \epsilon/2$ for all $x_n \in N_\eta(x)$. Thus $E(x_n) < E(x) + \epsilon/2 < \epsilon$ for sufficiently large n. By minimality $M_{x_n,\epsilon} < M_{x,\epsilon/2}$ for sufficiently large n. This is a contradiction. □

The assumption of the continuity of $\|V\|_{r,x}$ in Theorem 3.1 is well founded. There are several examples where the assumption is known to hold. For instance, if the RKHS is a space of real entire functions, as it is for the exponential kernel, then $\|V\|_{r,x}$ is not only continuous, but it is constant.

Using a similar argument as that in Theorem 3.1, the theorem can be shown to hold when the restricted Hilbert space norm is replace by the supremum norm over $\overline{N_r(x)}$. The proof of the following theorem can be found in the preliminary work for this article in [1].

Proposition 3.2. *Let D be a compact subset of $\mathbb{R}^n$, $V : \mathbb{R}^n \to \mathbb{R}$ be a continuous function, and $K : \mathbb{R}^n \to \mathbb{R}^n \to \mathbb{R}$ be a continuous and universal kernel function. For all $\epsilon, r > 0$, there exists $M \in \mathbb{N}$ such that for each $x \in D$, there is a collection of centers $c_1, ..., c_M \in N_r(x)$ and weights $w_1, ..., w_M \in \mathbb{R}$ such that $\sup_{y \in \overline{N_r(x)}} \left| V(y) - \sum_{i=1}^{M} K(y, c_i) \right| < \epsilon$.*

4. Explicit Bound for the Exponential Kernel

Theorem 3.1 establishes a bound on the number of kernel functions required for the maintenance of the accuracy of a moving local approximation.

However, the proof does not provide an algorithm to computationally determine the upper bound. Even when the approximation with kernel functions is performed over a fixed compact set, a general bound for the number of collocation nodes required for accurate function approximation under the Hilbert space norm is unknown. Thus, it is desirable to have a computationally determinable upper bound to the number of StaF basis functions required to yield an arbitrarily close approximation. Theorem 4.1 provides a calculable bound on the number of exponential functions required to yield an arbitrarily close approximation with respect to the supremum norm. That is, Theorem 4.1 provides a computable analogue of Theorem 3.1 and Proposition 3.2 for a StaF approximation problem of the form

$$\limsup_{t\to\infty} \sup_{y\in\overline{N_r(x(t))}} \left| V(y) - \sum_{i=1}^{\infty} w_i(t) K(y, c_i(x(t))) \right| < \epsilon.$$

While error bounds have been computed for the exponential function with respect to the supremum norm (c.f. [26]), current literature allows the "frequencies" or centers of the exponential kernel functions to be unconstrained. The lack of constraints on the centers of the exponential kernel functions means that the existing results cannot be leveraged for the StaF approximation problem. The contribution of Theorem 4.1 is the development of an error bound while constraining the size of the centers.

Theorem 4.1. *Let $K : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by $K(x, y) = \exp\left(x^T y\right)$ be the exponential kernel function. Let $D \subset \mathbb{R}^n$ be a compact set, $V : D \to \mathbb{R}$ continuous, and $\epsilon, r > 0$. For each $x \in D$, there exists a finite number of centers $c_1, ..., c_{M_{x,\epsilon}} \in N_r(x)$ and weights $w_1, w_2, ..., w_{M_{x,\epsilon}} \in \mathbb{R}$, such that*

$$\sup_{y\in\overline{N_r(x)}} \left| V(y) - \sum_{i=1}^{M_{x,\epsilon}} w_i K(y, c_i) \right| < \epsilon.$$

If p is an approximating polynomial that achieves the same accuracy over $\overline{N_r(x)}$ with degree $N_{x,\epsilon}$, then an asymptotically similar bound can be found with $M_{x,\epsilon}$ kernel functions, where $M_{x,\epsilon} < \binom{n+N_{x,\epsilon}+S_{x,\epsilon}}{N_{x,\epsilon}+S_{x,\epsilon}}$ for some constant $S_{x,\epsilon}$ that is the degree of an approximating polynomial. Moreover, $N_{x,\epsilon}$ and $S_{x,\epsilon}$ can be bounded uniformly over D, and thus, so can $M_{x,\epsilon}$.

Proof. For notational simplicity, the quantity $\|f\|_{D,\infty}$ denotes the supremum norm of a function $f : D \to \mathbb{R}$ over the compact set D throughout the proof of Theorem 4.1.

First, consider the ball of radius r centered at the origin. The statement of the theorem can be proved by finding an approximation of monomials by a linear combination of exponential kernel functions.

Let $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$ be a multi-index, and define $|\alpha| = \sum \alpha_i$. Note that[2]

$$m^{|\alpha|} \prod_{i=1}^{n} \left(\exp\left(y_i/m\right) - 1\right)^{\alpha_i} = y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n} + O\left(\frac{1}{m}\right)$$

which by the binomial theorem leads to the sum

$$m^{|\alpha|} \sum_{l_i \leq \alpha_i, i=1,2,...,n} \binom{\alpha_1}{l_1}\binom{\alpha_2}{l_2} \cdots \binom{\alpha_n}{l_n} (-1)^{|\alpha| - \sum_i l_i} \exp\left(\sum_{i=1}^{n} y_i \left(\frac{l_i}{m}\right)\right)$$
$$= y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n} + O\left(\frac{1}{m}\right). \quad (4)$$

The big-oh constant indicated by $O(1/m)$ can be computed in terms of the derivatives of the exponential function via Taylor's Theorem. The centers corresponding to this approximation are of the form l_i/m where l_i is a non-negative integer satisfying $l_i < \alpha_i$. Hence, for m sufficiently large, the centers reside in $N_r(0)$.

To shift the centers so that they reside in $N_r(y)$, let $x = (x_1, x_2, ..., x_n)^T \in \mathbb{R}^n$, and multiply both sides of (4) by $\exp\left(y^T x\right)$ to get

$$m^{|\alpha|} \sum_{l_i \leq \alpha_i, i=1,2,...,n} \binom{\alpha_1}{l_1}\binom{\alpha_2}{l_2} \cdots \binom{\alpha_n}{l_n} (-1)^{|\alpha| - \sum_i l_i} \exp\left(\sum_{i=1}^{n} y_i \left(\frac{l_i}{m} + x_i\right)\right)$$
$$= e^{y^T x} \left(y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n}\right) + O\left(\frac{1}{m}\right).$$

For each multi-index, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$, the centers for the approximation of the corresponding monomial are of the form $x_i + l_i/m$ for $0 \leq l_i \leq \alpha_i$. Thus, by linear combinations of these kernel functions, a function of the form $e^{y^T x} g(y)$, where g is a multivariate polynomial, can be uniformly approximated by exponential functions over $N_r(x)$. Moreover if g is a polynomial of degree β, then this approximation can be a linear combination of $\binom{n+\beta}{\beta}$ kernel functions.

[2]The notation $g_m(x) = O(f(m))$ means that for sufficiently large m, there is a constant C for which $g_m(x) < Cf(m)$ for all $y \in \overline{N_r(0)}$.

Two polynomials, p_x and q_x are selected to approximate V and $e^{-x^T y}$, respectively, over $N_r(x)$. Since V is a continuous function, it can be approximated with arbitrary accuracy by polynomials. Subsequently, the previous development will be utilized to approximate the polynomials by linear combinations of exponential functions.

Let $\epsilon' > 0$ and suppose that p_x is polynomial with degree $N_{x,\epsilon'}$ such that

$$p_x(y) = V(y) + \epsilon_1(y)$$

where $|\epsilon_1(y)| < \|e^{y^T x}\|_{D,\infty}^{-1}\epsilon'/2$ for all $y \in N_r(x)$. Let $q_x(y)$ be a polynomial in $\mathbb{R}^n$ variables of degree $S_{x,\epsilon}$ such that

$$q_x(y) = e^{-y^T x} + \epsilon_2(y)$$

where $\epsilon_2(y) < \|V\|_{D,\infty}^{-1}\|e^{y^T x}\|_{D,\infty}^{-1}\epsilon'/2$ for all $y \in N_r(x)$.

The above construction indicates that there is a sequence of linear combinations of exponential kernel functions, $F_m(y)$, (with a fixed number of centers inside $N_r(x)$) for which

$$\begin{aligned} F_m(y) &= e^{y^T x} q_x(y) p_x(y) + O\left(\frac{1}{m}\right) \\ &= e^{y^T x}\left(e^{-y^T x} + \epsilon_2(y)\right)\left(V(y) + \epsilon_1(y)\right) + O\left(\frac{1}{m}\right). \end{aligned}$$

After multiplication and an application of the triangle inequality, the following is established:

$$|F_m(y) - V(y)| < \epsilon' + \left(\frac{\|V\|_{D,\infty}^{-1}\|e^{y^T x}\|_{D,\infty}^{-1}}{4}\right)\epsilon'^2 + O\left(\frac{1}{m}\right)$$

for all $y \in N_r(x)$. The degree of the polynomial q_x, $S_{x,\epsilon}$, can be uniformly bounded in terms of the modulus of continuity of $e^{y^T x}$ over D. Similarly, the uniform bound on the degree of p_x, $N_{x,\epsilon'}$, can be described in terms of the modulus of continuity of V over D. The number of centers required for $F_m(y)$ is determined by the degree of the polynomial $q \cdot p$ (treating the x terms of q as constant), which is sum of the two polynomial degrees. Finally for m large enough and ϵ' small enough, $|F_m(y) - V(y)| < \epsilon$, and the proof is complete. □

Theorem 4.1 demonstrates an upper bound required for the accurate approximation of a function through the estimation of approximating polynomials. Moreover, the upper bound is a function of the polynomial degrees.

For example, for a neighborhood of the origin in $\mathbb{R}$, if p is an approximating polynomial of degree N, then the same order of approximation can be achieved by a linear combination of $N+1$ exponential functions. The exponential kernel will be used for simulations in Section 7 and Appendix A.

5. Existence and Smoothness of the Ideal Weight Function

Theorem 3.1 and Proposition 3.2 establish that given a kernel function, a finite number of centers can be used to yield an arbitrarily accurate estimation of a function, for a set of ideal weights. Theorem 4.1 further establishes that for the exponential kernel function, a calculable number of centers can be determined. However, further investigation is required to understand the characteristics of the ideal weights that correspond to the moving centers. For example, in control applications involving function approximation or system identification, it is assumed that there is a collection of constant ideal weights, and much of the theory is in the demonstration of the convergence of approximate weights to the ideal weights. The subsequent Theorem 5.1 establishes that ideal weights, which are functions of the state dependent centers, are m-times continuously differentiable. This property can then be used to develop weight update laws (e.g., see Section 6).

Since the ideal weights corresponding to a Hilbert space norm are unique, Theorem 5.1 is framed in the Hilbert space setting of (2). Thus, Theorem 5.1 together with Theorem 3.1 provides the StaF framework for RKHSs.

Theorem 5.1. *Let H be a RKHS over a set $X \subset \mathbb{R}^n$ with a strictly positive kernel $K : X \times X \to \mathbb{C}$ such that $K(\cdot, c) \in C^{m_0}(\mathbb{R}^n)$ for all $c \in X$. Suppose that $V \in H$. Let C be an ordered collection of M distinct centers, $C = (c_1, c_2, ..., c_M) \in X^M$, with the associated ideal weights*

$$W(C) = \underset{(a_i)_{i=1}^M \in \mathbb{C}^M}{\arg\min} \left\| \sum_{i=1}^{M} a_i K(\cdot, c_i) - V(\cdot) \right\|_H . \tag{5}$$

The function W is m_0-times continuously differentiable with respect to each component of C.

Proof. The determination of $W(C)$ is equivalent to computing the projection of V onto the space $Y = \text{span}\{K(\cdot, c_i) : i = 1, ..., M\}$. To compute the projection, a Gram-Schmidt algorithm can be employed. The Gram-Schmidt algorithm is most easily expressed in its determinant form. Let

$D_0 = 1$ and $D_m = \det\left(K(c_j, c_i)\right)_{i,j=1}^{m}$, then for $m = 1, ..., M$ the functions

$$u_m(x) := \frac{1}{\sqrt{D_{m-1} D_m}} \det \begin{pmatrix} K(c_1, c_1) & K(c_1, c_2) & \cdots & K(c_1, c_m) \\ K(c_2, c_1) & K(c_2, c_2) & \cdots & K(c_2, c_m) \\ \vdots & \vdots & \ddots & \vdots \\ K(c_{m-1}, c_1) & K(c_{m-1}, c_2) & \cdots & K(c_{m-1}, c_m) \\ K(x, c_1) & K(x, c_2) & \cdots & K(x, c_m) \end{pmatrix}$$

constitute an orthonormal basis for Y. Since K is strictly positive definite, D_m is positive for each m and every C. The coefficient for each $K(x, c_l)$ with $l = 1, ..., m$ in u_m is a sum of products of the terms $K(c_i, c_j)$ for $i, j = 1, ...m$. Each such coefficient is m_0-times differentiable with respect to each c_i, $i = 1, ..., M$. When $\langle V, u_m \rangle$ is computed for the projection, the result is a linear combination of evaluations of V at each of the centers. The function V is m_0-times continuously differentiable, since K is m_0-times differentiable [25], therefore $\langle V, u_m \rangle$ is continuous with respect to the centers. Finally, each term in $W(C)$ is a linear combination of the coefficients determined by u_m for $m = 1, ..., M$, and thus is m_0-times continuously differentiable with respect to each c_i for $i = 1, ..., M$. □

6. The Gradient Chase Theorem

As mentioned before, control theory problems involving function approximation and system identification are centered around the concept of weight update laws. Weight update laws are a collection of rules that the approximating weights must obey which lead to convergence to the ideal weights. In the case of the StaF approximation framework, the ideal weights are replaced with ideal weight functions. Theorem 5.1 showed that if the moving centers of the StaF kernel functions are selected in such a way that the centers adjust smoothly with respect to the state x, then the ideal weight functions will also change smoothly with respect to x. Thus, in this context, weight update laws of the StaF approximation framework aim to achieve an estimation of the ideal weight function at the current state.

Theorem 6.1 provides an example of such weight update laws that achieve a UUB result. The theorem takes advantage of perfect samples of a function in the RKHS H corresponding to a real valued kernel function.

The proof of the theorem is similar to the standard proof for the convergence of the gradient descent algorithm for a quadratic programming problem [27]. The contribution of the proof is in a modification, where the mean value theorem is used to produce an extra term which results in a UUB

result, and the continuity of the largest and smallest eigenvalues of a Gram matrix are used to get a uniform bound in tandem with the Kantorovich inequality.

Theorem 6.1 (Gradient Chase Theorem). *Let H be a real valued RKHS over $\mathbb{R}^n$ with a continuously differentiable strictly positive definite kernel function $K : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Let $V \in H$, $D \subset \mathbb{R}^n$ be a compact set, and $x : \mathbb{R} \to \mathbb{R}^n$ a state variable subject to the dynamical system $\dot{x} = q(x,t)$, where $q : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$ is a bounded locally Lipschitz continuous function. Further suppose that $x(t) \in D$ for all $t > 0$. Let $c : \mathbb{R}^n \to \mathbb{R}^M$, where for each $i = 1, ..., M$, $c_i(x) = x + d_i(x)$ where $d_i \in C^1(\mathbb{R}^n)$, and let $a \in \mathbb{R}^M$. Consider the function*

$$F(a,c) = \left\| V - \sum_{i=1}^{M} a_i K(\cdot, c_i(x)) \right\|_H^2 .$$

At each time instance $t > 0$, there is a unique $W(t)$ for which $W(t) = \arg\min_{a \in \mathbb{R}^M} F(a, c(x(t)))$. Given any $\epsilon > 0$ and initial value a^0, there is a frequency $\tau > 0$, where if the gradient descent algorithm (with respect to a) is iterated at time steps $\Delta t < \tau^{-1}$, then $F(a^k, c^k) - F(w^k, c^k)$ will approach a neighborhood of radius ϵ as $k \to \infty$.

Proof. Let $\bar{\epsilon} > 0$. By the Hilbert space structure of H:

$$F(a,c) = \|V\|_H^2 - 2V(c)^T a + a^T K(c) a$$

where $V(c) = (V(c_1), ..., V(c_M))^T$ and $K(c) = (K(c_i, c_j))_{i,j=1}^M$ is the symmetric strictly positive kernel matrix corresponding to c. At each time iteration t^k, $k = 0, 1, 2, ...$, the corresponding centers and weights can be written as $c^k \in \mathbb{R}^{nM}$ and $a^k \in \mathbb{R}^M$, respectively. The ideal weights corresponding to c^k will be denoted by w^k. It can be shown that $w^k = K(c^k)^{-1} V(c^k)$ and $F(w^k, c^k) = \|V\|_H^2 - V(c^k)^T K(c^k) V(c^k)$. Theorem 5.1 ensures that the ideal weights change continuously with respect to the centers which remain in a compact set $\tilde{D}^M$, where $\tilde{D} = \{x \in \mathbb{R}^M : \|x - D\| \leq \max_{i=1,...,M} (\sup_{x \in D} |d_i(x)|)\}$, so the collection of ideal weights is bounded. Let $R > \bar{\epsilon}$ be large enough so that $N_R(0)$ contains both the initial value a^0 and the set of ideal weights. To facilitate the subsequent analysis, consider

the constants:

$$R_0 = \max_{x\in D, t>0} |q(x,t)| \qquad R_1 = \max_{a\in\overline{N_r(0)}, c\in\tilde{D}} |\nabla_a F(a,c)|$$
$$R_2 = \max_{c\in\tilde{D}} |\nabla_c F(w(c),c)| \qquad R_3 = \max_{c\in\tilde{D}} |\dot{d}_i(x(t)|$$
$$R_4 = \max_{c\in\tilde{D}} \left\| \frac{d}{dc} w(c) \right\|$$

where ∇_a is the gradient with respect to a, and let $\Delta t < \tau^{-1} := \bar{\epsilon} \cdot (2(R_0 + R_3) \cdot (R_1 \cdot R_4 \cdot (R_0 + R_3) + R_2 + 1))^{-1}$. The proof aims to show that by using the gradient descent law for choosing a^k, the following inequality can be achieved:

$$\frac{F(a^{k+1}, c^{k+1}) - F(w^{k+1}, c^{k+1})}{F(a^k, c^k) - F(w^k, c^k)} < \delta + \frac{\bar{\epsilon}}{F(a^k, c^k) - F(w^k, c^k)}$$

for some $0 < \delta < 1$. Set

$$a^{k+1} = a^k + \lambda g \tag{6}$$

where $g = -\nabla_a F(a^k, c^k) = 2V(c^k) - 2K(c^k)a^k$ and λ is selected so that the quantity $F(a^k + \lambda g, c^k)$ is minimized. The λ that minimizes this quantity is $\lambda = \left(\frac{g^T g}{2g^T K(c^k) g}\right)$ which yields $F(a^{k+1}, c^k) = F(a^k, c^k) - \frac{(g^T g)^2}{4g^T K(c^k) g}$. Since $F(a^{k+1}, c^{k+1})$ is continuously differentiable in the second variable, we have $F(a^{k+1}, c^{k+1}) = F(a^{k+1}, c^k) + \nabla_c F(a^{k+1}, \eta) \cdot (c^{k+1} - c^k)$. Since $|\dot{c}(x(t))| < R_0 + R_3$, an application of the mean value theorem demonstrates that $\|c^{k+1} - c^k\| < (R_0 + R_3)\Delta t$. Thus

$$F(a^{k+1}, c^{k+1}) = F(a^{k+1}, c^k) + \epsilon_1(t^k),$$

where $|\epsilon_1(t^k)| < \bar{\epsilon}/2$ for all k. The quantity $F(w^{k+1}, c^{k+1})$ is continuously differentiable in both variables. Thus, by the multi-variable chain rule and another application of the mean value theorem:

$$F(w^{k+1}, c^{k+1}) = F(w^k, c^k) + \epsilon_2(t^k),$$

for $|\epsilon_2(t^k)| < \bar{\epsilon}/2$ for all k. Therefore, the following is established:

$$\frac{F(a^{k+1}, c^{k+1}) - F(w^{k+1}, c^{k+1})}{F(a^k, c^k) - F(w^k, c^k)} = \frac{F(a^{k+1}, c^k) - F(w^k, c^k) + (\epsilon_1(t^k) - \epsilon_2(t^k))}{F(a^k, c^k) - F(w^k, c^k)}$$
$$= 1 - \frac{(g^T g)^2}{(g^T K(c^k) g)(g^T K(c^k)^{-1} g)} + \frac{\epsilon_1(t^k) - \epsilon_2(t^k)}{F(a^k, c^k) - F(w^k, c^k)}.$$

The Kantorovich inequality [27] yields

$$1 - \frac{(g^T g)^2}{(g^T K(c^k) g)(g^T K(c^k)^{-1} g)} \leq \left(\frac{A_{c^k}/a_{c^k} - 1}{A_{c^k}/a_{c^k} + 1} \right)^2 \tag{7}$$

where A_{c^k} is the largest eigenvalue of $K(c^k)$ and a_{c^k} is the smallest eigenvalue of $K(c^k)$. The quantity on the right of (7) is continuous with respect to A_{c^k} and a_{c^k}. In turn, A_{c^k} and a_{c^k} are continuous with respect to $K(c^k)$ (c.f. Exercise 4.1.6 [28]) which is continuous with respect to c^k. Therefore there is a largest value, δ, that the right hand side of (7) obtains on the compact set $\tilde{D}$ and this value is less than 1. Moreover, δ is independent of $\bar{\epsilon}$, so it may be declared that $\bar{\epsilon} = \epsilon(1 - \delta)$. Finally,

$$\frac{F(a^{k+1}, c^{k+1}) - F(w^{k+1}, c^{k+1})}{F(a^k, c^k) - F(w^k, c^k)} \leq \delta + \frac{(\epsilon_1(t^k) - \epsilon_2(t^k))}{F(a^k, c^k) - F(w^k, c^k)}.$$

Therefore, setting $e(k) = F(a^k, c^k) - F(w^k, c^k)$, it can be shown that $e(k+1) \leq \delta e(k) + \epsilon(1 - \delta)$ and the conclusion of the theorem follows. □

7. Simulation for the Gradient Chase Theorem

To demonstrate the effectiveness of the Gradient Chase theorem, a simulation performed on a two-dimensional linear system is presented below. The system dynamics are given by

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

which is the dynamical system corresponding to a circular trajectory. The function to be approximated is

$$V(x_1, x_2) = x_1^2 + 5x_2^2 + \tanh(x_1 \cdot x_2),$$

and the kernel function to be used for function approximation are the exponential kernels, $K(x, y) = \exp\left(x^T y\right)$. The centers are arranged in an equilateral triangle centered about the state. In particular, each center resides on a circle of radius 0.1 centered at the state:

$$c_i(x) = x + 0.1 \begin{pmatrix} \sin((i-1)2\pi/3) \\ \cos((i-1)2\pi/3) \end{pmatrix}$$

for $i = 1, 2, 3$.

The initial values selected for the weights are $a^0 = [0\ 0\ 0]^T$. The gradient descent weight update law, given by (6), are applied 10 iterations per time-step and the time-steps incremented every 0.01 seconds. Figure 1 presents the results of the simulation.

Figure 1d demonstrates that the function approximation error is regulated to a small neighborhood of zero as the Gradient Chase Theorem is implemented and validates the claim of the UUB result of Theorem 6.1. In Figure 1c, approximations of the ideal weight function can be seen to be periodic as well as smooth. The smoothness of the ideal weight function itself is given in Theorem 5.1, and the periodicity of the approximation follows from the periodicity of the selected dynamical system, as illustrated in Figure 1a. Figure 1b presents a comparison of V evaluated at the current state to that of the approximation evaluated at the current state. Approximation of the function is maintained as the system state moves through its domain as anticipated.

8. Application to Adaptive Dynamic Programming

The application of approximation theory to the theory of optimal control arises through the approximation of the optimal value function, which is the solution to the Hamilton-Jacobi-Bellman (HJB) equation. Efficient methods for the approximation of the optimal value function are essential, since an increase in dimension can lead to a exponential increase in the number of required basis functions necessary to achieve an accurate approximation, the so called "curse of dimensionality".

The optimal value function corresponds to the infinite horizon optimal regulator problem, where the cost function

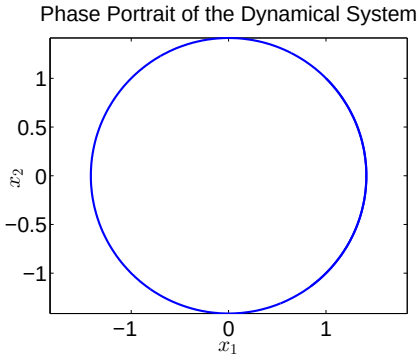
$$J(x, u) = \int_0^\infty x^T Q x + u^T R u \, dt$$

is to be minimized subject to the dynamics

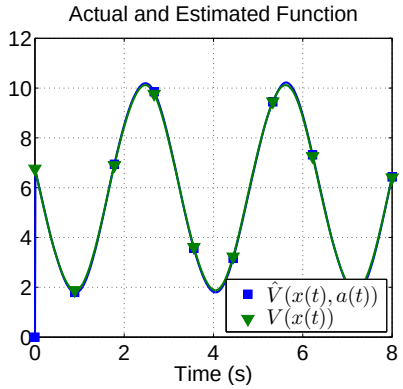
$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \tag{8}$$

where $x : \mathbb{R}_+ \to \mathbb{R}^n$, $u : \mathbb{R}_+ \to \mathbb{R}^m$, $Q \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$, with Q and R positive definite, $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$. Moreover, f and g are assumed to be locally Lipschitz. The optimal value function is given by

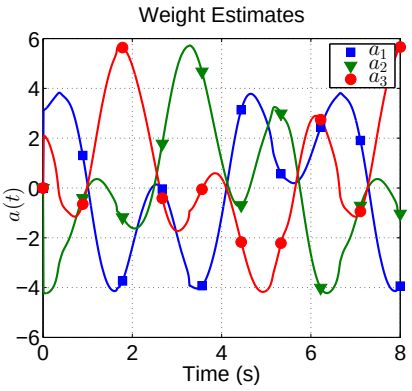
$$V(x) = \inf_{u\in\mathcal{U}} \int_0^\infty x^T Q x + u^T R u \, dt$$



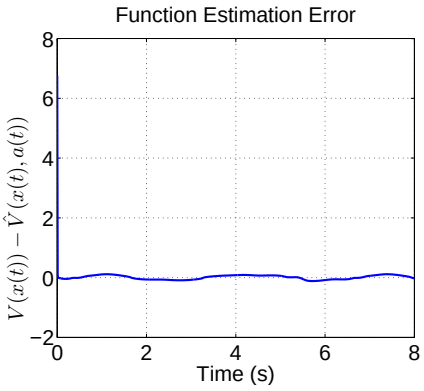
(a) Trajectory of the state vector.



(b) Comparison of V and the approximation $\hat{V}$.



(c) The values of the weight function estimates.



(d) Error committed by the approximation at the current state.

Figure 1: Results of the numerical experiment demonstrating the Gradient Chase algorithm.

where $\mathcal{U}$ is the collection of admissible controllers. When the optimal value function is continuously differentiable and an optimal controller, $u^* \in \mathcal{U}$ exists, the optimal value function is the unique solution to the HJB equation

$$0 = x^T Q x + u^{*T} R u^* + \nabla V(x)(f(x) + g(x)u^*). \tag{9}$$

Once the optimal value function is determined, the optimal controller takes the form

$$u^*(x(t)) = -\frac{1}{2} R^{-1} g(x)^T \nabla V(x(t))^T. \tag{10}$$

In many applications, an approximation of the optimal controller is used real-time to yield autonomous behavior in a dynamic environment.

For some problems, such as the linear quadratic regulator (LQR) problem, the optimal value function takes a particular form which simplifies the choice of basis functions. In the case of LQR, the optimal value function is of the form $\sum_{i,j=1}^{n} w_{i,j} x_j x_i$ (c.f. [29, 30]), so basis functions of the form $\sigma_{i,j} = x_j x_i$ will provide an accurate estimation of the optimal value function provided the weights, $w_{i,j} \in \mathbb{R}$, are tuned properly. However, in most cases, the form of the optimal value function is unknown, and generic basis functions have been proposed to parameterize the problem.

Adaptive dynamic programming (ADP) replaces V with a parametrization, $\hat{V}(x, W_c) = \sum_{i=1}^{M} w_{i,c} \sigma_i(x)$, with $W_c = (w_{1,c}, ..., w_{M,c}) \in \mathbb{R}^M$, and u^* with a parametrization $\hat{u}(x, W_a) = -\frac{1}{2} R^{-1} g(x)^T \nabla_x V(x, W_a)^T$ where $W_a \in \mathbb{R}^M$. The actor and critic weights, W_a and W_c respectively, are tuned to minimize the residual Bellman error (BE),

$$\delta(x, W_a, W_c) = x^T Q x + \hat{u}(x, W_a)^T R u(x, W_a) + \nabla_x \hat{V}(x, W_c) \left(f(x) + g(x)\hat{u}(x, W_a)\right),$$

over all x in some compact set D in real-time. The BE is used to motivate weight update laws for W_a and W_c to achieve a real-time minimization.

Throughout the ADP literature, many basis functions have been proposed for real-time (approximate) optimal control. However, in practice, it is difficult to select weight update laws that guarantee stability by achieving a good approximation of the ideal weights, especially for a system with a modest embedded processor. In the majority of cases, actual implementation of ADP is executed using only polynomial basis functions, and the StaF method enables a broader class of functions to be used for approximate optimal control of a dynamical system.

In this setting, the StaF problem becomes

$$\limsup_{t \to \infty} \sup_{x \in \overline{N_r(x)}} |\delta(x, W_a(t), W_c(t))| < \epsilon.$$

Appendix A provides more information concerning the application of the StaF method to ADP by presenting the results of a companion paper [22].

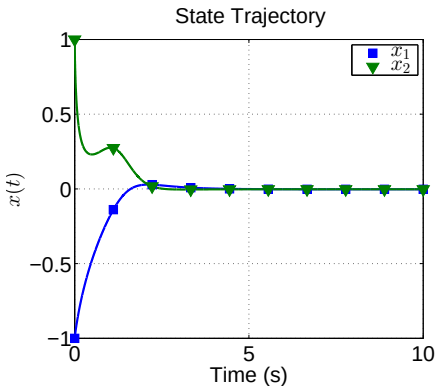
9. Conclusion

A new StaF kernel method is introduced in this paper for the purpose of function approximation. The development in this paper establishes that by using the StaF method a local approximation of a function can be maintained in real-time as a state moves through a compact domain. Heuristically, much fewer kernel functions are required in comparison to more traditional function approximation schemes, since the approximation is maintained in a smaller region. For the exponential kernels, a new theorem in this paper establishes that an explicit bound on the number of kernel functions required can calculated. Two applications of this methodology were presented. In Section 6, a "gradient chase" algorithm was developed. There it was seen that a function may be well approximated provided that the algorithm was applied with a high enough frequency. Simulations results provided in Section 7 demonstrated the performance of the gradient chase algorithm, and an application to ADP is provided in Section 8 and the Appendix for an infinite horizon optimal regulation problem.

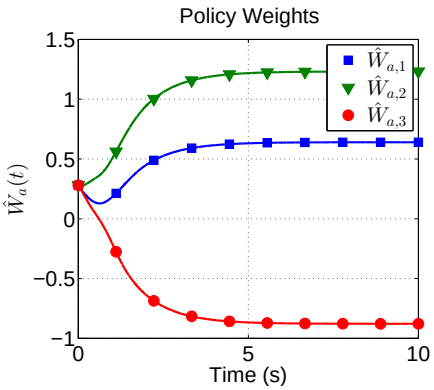
The strength of the StaF methodology is the reduction of the computational requirements for real-time implementation of a function approximation, through the reduction in the number of basis functions. As demonstrated in Appendix A, where only three basis functions were required to achieve a stabilizing approximate optimal controller for a 2-dimensional system. However, since the StaF method aims at maintaining an accurate local approximation of the value function only in a local neighborhood of the current system state, the StaF kernel method lacks memory, in the sense that the information about the ideal weights over a region of interest is lost when the state leaves the region of interest. Thus, unlike existing techniques, the StaF method generates an approximation that is valid only in a local region. A memory-based modification to the StaF kernel method that retains and reuses past information for creating a global approximation is the subject of future research.

Appendix A. Applications to Adaptive Dynamic Programming

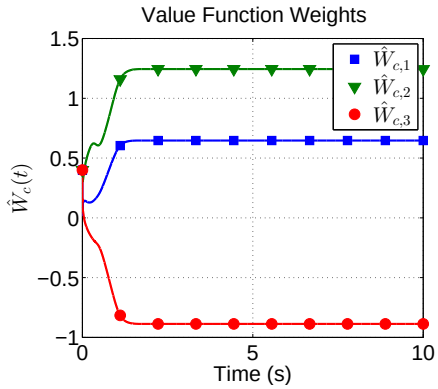
To demonstrate the effectiveness of the StaF technique in the context of optimal control, the simulation results of a companion paper are presented



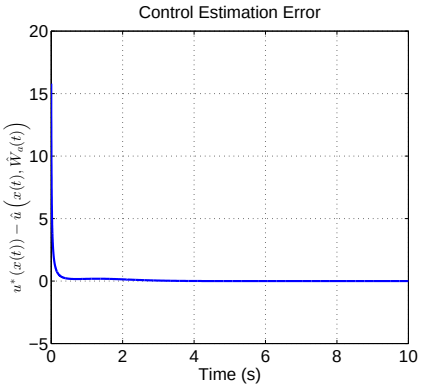
(a) Trajectory of the state vector.



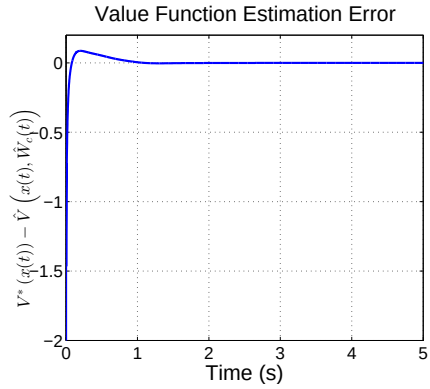
(b) Trajectory of the actor weights.



(c) The values of the critic weights.



(d) Error committed by approximate policy.



(e) Error of the estimation of the value function at the current state.

Figure A.2: Results of the numerical experiment demonstrating the convergence for the StaF ADP method. [22]

here. The details of the analysis are contained in [22]. The dynamical system in question is of the form $\dot{x} = f(x) + g(x)u$ where $x = (x_1, x_2)^T \in \mathbb{R}^2$,

$$f(x) = \begin{pmatrix} x_2 - x_1 \\ -\frac{1}{2}x_1 - \frac{1}{2}x_2(\cos(2x_1) + 2)^2 \end{pmatrix}, \text{ and } g(x) = \begin{pmatrix} 0 \\ \cos(2x_1) + 2 \end{pmatrix}. \tag{A.1}$$

Associated with this dynamical system is the cost functional

$$J(x, u) = \int_0^\infty (x^T(\tau)x(\tau) + u(\tau)^2)\, d\tau \tag{A.2}$$

In the infinite horizon regulation problem, the goal is to determine an optimal control law $u^* : \mathbb{R}^2 \to \mathbb{R}$ (assuming an optimal control law exists) that satisfies

$$u^*(x_0) = \arg\min_{u \in \mathcal{U}} \int_0^\infty (x^T(\tau)x(\tau) + u(\tau)^2)\, d\tau$$

where $\mathcal{U}$ is the collection of admissible controllers and $x(0) = x_0$ inside the integrand. The optimal value function is given by

$$V(x_0) = \min_{u \in \mathcal{U}} \int_0^\infty (x^T(\tau)x(\tau) + u(\tau)^2)\, d\tau$$

when such a minimum exists, and the optimal value function satisfies the HJB equation (9). If V^* satisfies the HJB equation and is also continuously differentiable, then it is the unique solution to (9). Furthermore, u^* can be determined from V^* by $u^*(x) = \frac{1}{2}g^T(x)\nabla V^*(x)$.

In most cases, the optimal value function cannot be determined analytically, and approximate solutions are used instead. However, for the system presented in the section, the optimal value function is known. In particular, for the infinite horizon optimal regulator problem with dynamics given by (A.1) with cost functional (A.2), the optimal value function is given by $V^*(x) = \frac{1}{2}x_1^2 + x_2^2$ and the associated optimal control law is given by $u^*(x) = -(\cos(2x_1) + 2)x_2$. More details can be found in [14].

In this example, the infinite horizon optimal regulator problem is solved in real-time. The function V^* is approximated by a function of the form

$$\hat{V}(x, \hat{W}_c) = \sum_{i=1}^{3} \hat{W}_{c,i} \left(\exp(x^T c_i(x)) - 1\right)$$

where $\hat{W}_c \in \mathbb{R}^3$ are weights to be adjusted in real-time, and $c_i(x) = x + d_i(x)$ where

$$d_i(x) = 0.7 \left(\frac{x^T x + 0.01}{1 + x^T x}\right) \left[\cos\left(\frac{2\pi}{3}i + \frac{\pi}{2}\right), \sin\left(\frac{2\pi}{3}i + \frac{\pi}{2}\right)\right]^T$$

for $i = 1, 2, 3$. The approximation of the optimal control law is given by

$$\hat{u}(x, \hat{W}_a) = -\frac{1}{2} g^T(x) \nabla_x \hat{V}(x, \hat{W}_a)$$

where $\hat{W}_a \in \mathbb{R}^3$ are weights to be adjusted in real-time. In the framework of ADP, the functions V^* and u^* are replaced by their approximations $\hat{V}$ and $\hat{u}$, respectively, in the HJB equation, yielding a residual nonzero error, called the Bellman error (BE). The goal is to minimize the BE by adjustments of the weights, $\hat{W}_a$ and $\hat{W}_c$. If the BE is identically zero after the adjustment of the weights, then the optimal value function and the approximation of the optimal value function coincide. For nonzero BE, the BE is used as a heuristic measure of the distance between $\hat{V}$ and V^*, as well as the distance between $\hat{u}$ and u^*. The weight update laws and subsequent convergence analysis can be found in [22].

The results of the numerical experiment are presented in Figure A.2. Figure A.2a indicates that the state is regulated to the origin when using the ADP algorithm combined with the StaF methodology. Figure A.2b shows that the weight vector $\hat{W}_a$ converged as well. In typical StaF implementations, the weights are not expected to converge. However, since the optimal control problem is a regulator problem, the state and the centers ultimately occupy a fixed neighborhood of the origin, and the weights converge to the ideal weights corresponding to a small neighborhood of the origin.

When the weights converge, it is expected that $\hat{W}_a$ and $\hat{W}_c$ converge to the same values. The convergence is demonstrated by comparing Figure A.2b and Figure A.2c. The approximate controller and the optimal controller converge as well, as shown in Figure A.2d, and the value function estimation error, given in Figure A.2e, vanishes rapidly.

Acknowledgments

This research is supported in part by NSF award number 1509516 and Office of Naval Research Grant N00014-13-1-0151. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsoring agencies.

References

[1] J. A. Rosenfeld, R. Kamalapurkar, W. E. Dixon, State following (StaF) kernel functions for function approximation Part I: Theory and motivation, in: Proc. Am. Control Conf., 2015, pp. 1217–1222.

[2] A. Christmann, I. Steinwart, Univeral kernels on non-standard input spaces, in: Advances in Nueral Information Processing, 2010, pp. 406–414.

[3] C. A. Micchelli, Y. Xu, H. Zhang, Universal kernels, J. Mach. Learn. Res. 7 (2006) 2651–2667.

[4] J. Park, I. W. Sandberg, Universal approximation using radial-basis-function networks, Neural Comput. 3 (2) (1991) 246–257. doi:10.1162/neco.1991.3.2.246.
URL http://dx.doi.org/10.1162/neco.1991.3.2.246

[5] C. Prud'homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, G. Turinici, Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods, AMSE Journal of Fluids Engineering 124(1) (2001) 70–80.

[6] E. Balmes, Parametric families of reduced finite element models. theory and applications, Mechanical Systems and Signal Processing 10 (4) (1996) 381 – 394. doi:http://dx.doi.org/10.1006/mssp.1996.0027.
URL http://www.sciencedirect.com/science/article/pii/S0888327096900278

[7] J. M. P. Ahmed K. Noor, Reduced basis technique for nonlinear analysis of structures, AIAA Journal 18 (2) (1980) 455–462.

[8] A. Al-Tamimi, F. L. Lewis, M. Abu-Khalaf, Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof, IEEE Trans. Syst. Man Cybern. Part B Cybern. 38 (2008) 943–949.

[9] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, W. E. Dixon, A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems, Automatica 49 (1) (2013) 89–92.

[10] T. Dierks, B. Thumati, S. Jagannathan, Optimal control of unknown affine nonlinear discrete-time systems using offline-trained neural networks with proof of convergence, Neural Netw. 22 (5-6) (2009) 851–860.

[11] F. L. Lewis, D. Vrabie, Reinforcement learning and adaptive dynamic programming for feedback control, IEEE Circuits Syst. Mag. 9 (3) (2009) 32–50.

[12] P. Mehta, S. Meyn, Q-learning and pontryagin's minimum principle, in: Proc. IEEE Conf. Decis. Control, 2009, pp. 3598 –3605.

[13] R. Padhi, N. Unnikrishnan, X. Wang, S. Balakrishnan, A single network adaptive critic (SNAC) architecture for optimal control synthesis for a class of nonlinear systems, Neural Netw. 19 (10) (2006) 1648–1660.

[14] K. Vamvoudakis, F. Lewis, Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem, Automatica 46 (5) (2010) 878–888.

[15] H. Zhang, L. Cui, Y. Luo, Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network adp, IEEE Trans. Cybern. 43 (1) (2013) 206–216.

[16] H. Zhang, L. Cui, X. Zhang, Y. Luo, Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method, IEEE Trans. Neural Netw. 22 (12) (2011) 2226–2236.

[17] H. Zhang, D. Liu, Y. Luo, D. Wang, Adaptive Dynamic Programming for Control Algorithms and Stability, Communications and Control Engineering, Springer-Verlag, London, 2013.

[18] R. A. DeVore, Nonlinear approximation, Acta Numerica 7 (1998) 51–150.

[19] M. Gaggero, G. Gnecco, M. Sanguineti, Dynamic programming and value-function approximation in sequential decision problems: Error analysis and numerical results, Journal of Optimization Theory and Applications 156.

[20] M. Gaggero, G. Gnecco, M. Sanguineti, Approximate dynamic programming for stochastic n-stage optimization with application to optimal consumption under uncertainty, Computational Optimization and Applications 58 (1) (2014) 31–85.

[21] R. Zoppoli, M. Sanguineti, T. Parisini, Approximating networks and extended ritz method for the solution of functional optimization problems, Journal of Optimization Theory and Applications 112 (2) (2002) 403–440. doi:10.1023/A:1013662124879.

[22] R. Kamalapurkar, J. A. Rosenfeld, W. E. Dixon, State following (StaF) kernel functions for function approximation Part II: Adaptive dynamic programming, in: Proc. Am. Control Conf., 2015, pp. 521–526.

[23] K. Zhu, Analysis on Fock spaces, Vol. 263 of Graduate Texts in Mathematics, Springer, New York, 2012. doi:10.1007/978-1-4419-8801-0.
URL http://dx.doi.org/10.1007/978-1-4419-8801-0

[24] A. Pinkus, Strictly positive definite functions on a real inner product space, Adv. in Comput. Math. 20 (2004) 263–271.

[25] I. Steinwart, A. Christmann, Support vector machines, Information Science and Statistics, Springer, New York, 2008.

[26] G. Beylkin, L. Monzon, On approximation of functions by exponential sums, Applied and Computational Harmonic Analysis 19 (1) (2005) 17 – 48. doi:http://dx.doi.org/10.1016/j.acha.2005.01.003.
URL http://www.sciencedirect.com/science/article/pii/S106352030500014X

[27] D. P. Bertsekas, Nonlinear Programming, Athena Scientific, Belmont, MA, 1999.

[28] G. K. Pedersen, Analysis now, Vol. 118 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1989. doi:10.1007/978-1-4612-1007-8.
URL http://dx.doi.org/10.1007/978-1-4612-1007-8

[29] D. Kirk, Optimal Control Theory: An Introduction, Dover, 2004.

[30] D. Liberzon, Calculus of variations and optimal control theory: a concise introduction, Princeton University Press, 2012.