Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Approximation Theory 129 (2004) 230–239

JOURNAL OF Approximation Theory

www.elsevier.com/locate/jat

Approximation by ridge function fields over compact sets

Bruno Pelletier*

Laboratoire de Mathématiques Appliquées, Université du Havre, 25 rue Philippe Lebon, 76063 Le Havre, France

Received 29 November 2003; received in revised form 26 March 2004

Abstract

We study the approximation of a continuous function field over a compact set T by a continuous field of ridge approximants over T, named ridge function fields. We first give general density results about function fields and show how they apply to ridge function fields. We next discuss the parameterization of sets of ridge function fields and give additional density results for a class of continuous ridge function fields that admits a weak parameterization. Finally, we discuss the construction of the elements in that class.

Keywords: Ridge approximation; Nonlinear approximation; Function field; Density

1. Introduction

In this work, we study the problem of approximating a continuous function field over a compact set T by a continuous field of approximants over T. By a function field over a compact set T is meant a map defined on T and valued in a function space. Let $\mathcal{C}(X, \mathbf{R})$ be the space of continuous real-valued functions on X, a subset of $\mathbf{R}^d$, and let $\mathcal{M}$ be a subset of $\mathcal{C}(X, \mathbf{R})$. We shall study the approximation of a map $T \to \mathcal{C}(X, \mathbf{R})$ by a map $T \to \mathcal{M}$, with special emphasis on the case where $\mathcal{M}$ is a set of ridge function-based approximants.

A ridge function over $\mathbf{R}^d$ is a function of the type $h(\mathbf{ax})$, where $h : \mathbf{R} \to \mathbf{R}$, $\mathbf{a}$ is a point of $\mathbf{R}^d$, and $\mathbf{ax}$ is the usual inner product in $\mathbf{R}^d$.

* Fax: +33-2-32-74-40-24.
E-mail address: bruno.pelletier@univ-lehavre.fr (B. Pelletier).

0021-9045/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jat.2004.06.007

Approximation by ridge function refers to approximation by linear combinations of n ridge functions, for some integer n. In the most general setting, the function h is allowed to vary in $\mathcal{C}(\mathbf{R})$; i.e., we consider the sets

$$\mathcal{R}_n(A) = \left\{ \sum_{i=1}^{n} c_i h_i(\mathbf{a}_i \mathbf{x}), c_i \in \mathbf{R}, \mathbf{a}_i \in A \subset \mathbf{R}^d, h_i \in \mathcal{C}(\mathbf{R}, \mathbf{R}) \right\} \tag{1}$$

and

$$\mathcal{R}(A) = \cup_n \mathcal{R}_n(A) \tag{2}$$

of approximants, where the directions $\mathbf{a}_i$ belong to some subset A of $\mathbf{R}^d$. If the function h is fixed, the above sets of approximants become

$$\mathcal{M}_n = \left\{ \sum_{i=1}^{n} c_i h(\mathbf{a}_i \mathbf{x}), c_i \in \mathbf{R}, \mathbf{a}_i \in \mathbf{R}^d \right\} \tag{3}$$

and

$$\mathcal{M} = \cup_n \mathcal{M}_n. \tag{4}$$

A slight variation on the theme consists in approximating by linear combinations of shifted ridge functions, i.e., functions of the form $h(\mathbf{a}\mathbf{x} + b)$, where $\mathbf{a} \in \mathbf{R}^d$, and where $b \in \mathbf{R}$ is the shift. Note that $\mathcal{R}$ and $\mathcal{M}$ are not linear spaces.

This kind of approximation has been studied by several authors, and density results, as well as bounds on the approximation rate, have been obtained. In Lin and Pinkus [11], necessary and sufficient conditions on the set A are given for $\mathcal{R}(A)$ to be dense in $\mathcal{C}(\mathbf{R}^d)$, in the topology of uniform convergence on compact sets (see also the paper by Vostrecov and Kreines [17]). An asymptotic expression of the approximation rate has been obtained by Maiorov [12]. Approximation by elements of the set $\mathcal{M}$ arose from the field of neural networks, where $\mathcal{M}$ has been shown to be dense in $\mathcal{C}(\mathbf{R}^d)$ if the function h is of sigmoidal form, i.e., if $\lim_{-\infty} h(t) = 0$ and $\lim_{+\infty} h(t) = 1$ [7,8], and Barron [2] obtained the dimension-independant upper bound $\mathcal{O}(n^{-1/2})$ on the approximation rate. Additional results may be found in [6,13,14,15].

The particular form of approximation studied here is motivated by a physical problem coming from the field of geosciences, i.e., ocean sciences, atmosphere sciences, and earth sciences, for which the above-approximation methods do not match all of the physical requirements. This problem is known as the ocean color problem. It consists in estimating the concentrations of several oceanic constituents, such as phytoplankton, from a vector $\mathbf{x}$ of radiometric measurements acquired by a sensor aboard a satellite. Thus, if one wishes to estimate the phytoplankton concentration from $\mathbf{x}$, a real-valued function of $\mathbf{x}$ is sought. In fact, those radiometric measurements depend continuously on a vector $\mathbf{t}$ of angular variables that are used to characterize the positions of the sun and of the satellite, relative to the observed point of the Earth' surface. Hence, the ocean color problem may be seen as a collection of similar problems continuously indexed by $\mathbf{t}$. In this context, a solution may be expressed as a function field over T, the set of allowable values for $\mathbf{t}$.

It is the purpose of this paper to give a grounding to this methodology by stating results related to its efficiency, i.e., density results. In Section 2, general density results about function fields over a compact set are given and applied to fields of ridge approximants, called ridge function fields for shorteness. Next, in Section 3, we discuss the parameterization of sets of continuous ridge function fields, which is necessary for their construction. Additional density results for a class of continuous ridge function fields are obtained, leading to the main results of Propositions 10 and 11. We conclude the paper with a brief exposition of the perspectives of this work.

2. Density results

Let us start by recalling some facts related to the compact-open topology. Let X be a locally compact Hausdorff space, and let Y be a Hausdorff space. In the following, Y^X will stand for the set of *continuous* functions from X to Y.

The compact-open topology on Y^X is generated by the sets $S(K, U) = \{f \in Y^X | f(K) \subset U\}$, where K is a compact subset of X, and where U is an open subset of Y. Furthermore, if X is a compact Hausdorff space, and if Y is metric, then the compact-open topology on Y^X is induced by the metric of uniform convergence,

$$\mathrm{dist}(f, g) = \sup\{\mathrm{dist}(f(x), g(x)) | x \in X\}.$$

Hence, if X is locally compact and Hausdorff, then the compact-open topology on Y^X is the topology of uniform convergence on compact sets.

There are also the following important two theorems, a proof of which may be found in [3], for example.

Theorem 1. *Let X be a locally compact Hausdorff space, and let Y and T be Hausdorff spaces. Let $f : X \times T \to Y$ be a function, and let f_t be the functions defined for each t by $f_t(x) = f(x, t)$. Then f being continuous, is equivalent to both of the following conditions holding*
(i) *each f_t is continuous*; *and*
(ii) *the function $T \ni t \mapsto f_t \in Y^X$ carrying t to f_t is continuous*.

Theorem 2. *Let X and T be locally compact Hausdorff spaces, and let Y be a Hausdorff space. Then there is the homeomorphism*

$$Y^{X\times T} \xrightarrow{\approx} \left(Y^X\right)^T .$$

Hence, by Theorem 1, for a function $T \to Y^X$ to be continuous, it suffices that the associated function $X \times T \to Y$ is continuous. Theorem 2 is also known as the *exponential law*.

Now let X be a locally compact Hausdorff space, and let $\mathbf{R}^X$ be the set of continuous real-valued functions on X. Let T be a compact Hausdorff space. We introduce the following notation. Given a function $f : T \to \mathbf{R}^X$, define the function $f_* : X \times T \to \mathbf{R}$ by $f_*(x, t) = f(t)(x)$.

Theorem 3. *Let X be a locally compact Hausdorff space, let T be a compact metric Hausdorff space, and let* $\mathcal{M}$ *be a dense subset of* $\mathbf{R}^X$. *Then the set* $\mathcal{M}^T$ *is dense in* $\left(\mathbf{R}^X\right)^T$.

Proof. By the exponential law theorem, it suffices to show that the set *S* of continuous functions of the form

$$X \times T \ni (x,t) \mapsto f_*(x,t) := f(t)(x) \in \mathbf{R}$$

is dense in $\mathbf{R}^{X\times T}$, where $f \in \mathcal{M}^T$.

Let $\varepsilon > 0$, and let $f \in \left(\mathbf{R}^X\right)^T$. Let

$$\mathcal{B}_\varepsilon(t) = \{g \in \mathcal{M} : \sup_{x\in K} |f_*(x,t) - g(x)| < \varepsilon; \forall K \subset X \text{ compact }\}.$$

Since *f* is continuous,

$$\forall \varepsilon' > 0, \forall t_0 \in T, \exists \eta(\varepsilon') > 0 : |t - t_0| < \eta => \sup_K |f_*(x,t) - f_*(x,t_0)| < \varepsilon',$$

for all $K \subset X$ compact.

Let $g_t \in \mathcal{B}_{\varepsilon'}(t)$. We have that

$$\begin{aligned}\sup_K |g_t(x) - f(x,t_0)| &< \sup_K |g_t(x) - f_*(x,t)| + \sup_K |f_*(x,t) - f_*(x,t_0)| \\ &< 2\varepsilon'.\end{aligned}$$

Hence $\mathcal{B}_{\varepsilon'}(t) \cap \mathcal{B}_{2\varepsilon'}(t) \neq \emptyset$, which implies that $\mathcal{B}_{2\varepsilon'}(t) \cap \mathcal{B}_{2\varepsilon'}(t_0) \neq \emptyset$, since $\mathcal{B}_{\varepsilon'}(t) \subset \mathcal{B}_{2\varepsilon'}(t)$. Hence $\forall \varepsilon' > 0, \forall t_0 \in T, \exists \eta > 0 : \text{dist}(t,t_0) < \eta => \mathcal{B}_{\varepsilon'}(t) \cap \mathcal{B}_{\varepsilon'}(t_0) \neq \emptyset$.

Consequently, for all $\varepsilon > 0$, there exists a continuous map $\hat{f} : T \to \mathcal{M}$ such that $\forall t \in T$, $\hat{f}(t) \in \mathcal{B}_\varepsilon(t)$. Finally, since *T* is compact, we have that $\sup_T \sup_K |\hat{f}_*(x,t) - f_*(x,t)| < \varepsilon$, for all compact $K \subset X$. □

Corollary 4. *Let* $\mathcal{G}$ *be a fundamental set in* $\mathbf{R}^X$; *i.e.*, span $\mathcal{G}$ *is dense in* $\mathbf{R}^X$. *Then the set of continuous maps* $T \to$ span $\mathcal{G}$ *is dense in* $\left(\mathbf{R}^X\right)^T$.

We now move on to the definition of a ridge function field. Let *T* be a compact Hausdorff metric space. We define a *ridge function field* over *T* as being a continuous map $T \to \mathcal{M}$, where $\mathcal{M}$ is the set of ridge function approximants defined by Eq. (4). It is assumed here, and in the sequel of the paper, that the generator function *h* in Eq. (4) is such that $\mathcal{M}$ is dense in $\mathcal{C}(\mathbf{R}^d)$. For instance, *h* may be of sigmoidal form. Under this assumption, we have, as a corollary of Theorem 3, the following proposition.

Proposition 5. *Let T and* $\mathcal{M}$ *be as above. Then the set* $\{T \to \mathcal{M}\}$ *of continuous ridge function fields over T is dense in* $\left(\mathbf{R}^X\right)^T$.

The set of continuous ridge function fields over *T* will be denoted by $\mathcal{M}^T$.

Remark. In the case where $\mathcal{M}$ in Theorem 3, or span $\mathcal{G}$ in Corollary 4, is a linear space, one immediately obtains a characterization of $\mathcal{M}^T$. For instance, consider the case where span $\mathcal{G}$ is the set of polynomials in several variables. Then $\mathcal{M}^T$ may be identified with the tensor product $\mathcal{C}(T) \otimes \mathcal{G}$, providing one with a practical way of constructing a continuous field of polynomials over *T*. Things go differently when $\mathcal{M}$ is not a linear space, as will be discussed in the next section.

3. Parameterization of ridge function fields

Let $\zeta \in \mathcal{M}^T$ be a ridge function field. Since *T* is compact, we may assume, without loss of generality, that ζ belongs to $\mathcal{M}_n^T$, for some integer *n*, where $\mathcal{M}_n$ is the set of linear combinations of at most *n* shifted ridge functions. We are willing to characterize the elements of $\mathcal{M}_n^T$ for the purpose of defining a simple and practical construction method of ridge function fields. Let us start with some definitions and general points.

Let $\mathcal{A}$ be a topological space. We call a set $\mathcal{P}$ a *parameter space* for $\mathcal{A}$ if $\mathcal{P}$ is homeomorphic to $\mathcal{A}$, and the homeomorphism $i_p : \mathcal{P} \to \mathcal{A}$ will be referred to as a *parameterization* for $\mathcal{A}$. We call a set $\mathcal{P}$ a *weak-parameter space* for $\mathcal{A}$ if there exists a continuous surjection $i_p : \mathcal{P} \to \mathcal{A}$, and i_p will be referred to as a *weak-parameterization* for $\mathcal{A}$. Note that if $\mathcal{A}$ admits a weak-parameterization $i_p : \mathcal{P} \to \mathcal{A}$, then it admits a parameterization if and only if the map i_p is open. In fact, if i_p is open, then $\mathcal{A}$ is homeomorphic to $\mathcal{P}/\sim$, the quotient space with the quotient topology, being the quotient of $\mathcal{P}$ given by the equivalence relation

$$p_1 \sim p_2 \text{ iff } i_p(p_1) = i_p(p_2). \tag{5}$$

The set $\mathcal{M}_n$ of ridge approximants admits a weak parameterization. More precisely, each element of $\mathcal{M}_n$ depends on parameters c_i, $\mathbf{a}_i$, b_i, for $i = 1, \ldots, n$, which we shall summarize by a vector θ_n. Let Θ_n be the set of allowable values for θ_n, i.e., $\Theta_n = \prod_{i=1}^n \mathbf{R} \times \mathbf{R}^d \times \mathbf{R}$, and let $i_n : \Theta_n \to \mathcal{M}_n$ be the continuous map sending a parameter vector θ_n to the corresponding ridge approximant of $\mathcal{M}_n$. The map i_n is a continuous surjection, i.e., a weak-parameterization for $\mathcal{M}_n$. Hence, a ridge approximant of $\mathcal{M}_n$ is constructed by specifying an element of the weak-parameter space Θ_n.

We ask if the set $\mathcal{M}_n^T$ of continuous ridge function fields over *T* admits a weak parameterization. In fact, we are to be faced with the following difficulties. For each continuous ridge function field $\zeta \in \mathcal{M}_n^T$, there exists at least one vector-valued function $\xi : T \to \Theta_n$ such that $\zeta = i_n \circ \xi$. Note that a discontinuous function $\xi : T \to \Theta_n$ may yield a continuous ridge function field $\zeta = i_n \circ \xi$, since i_n is only surjective. Therefore, there exists an appropriate subset of the set of eventually discontinuous functions $T \to \Theta_n$, which is a weak-parameter space for $\mathcal{M}_n^T$; the difficulty here resides in its characterization. An alternative approach, to ensuring the continuity of ζ is to proceed conversely, by constructing ζ via a continuous parameter map $\xi : T \to \Theta_n$; i.e., $\zeta = i_n \circ \xi$. By doing so, the field ζ is continuous, but we are not sure to get all of $\mathcal{M}_n^T$ when ξ varies in $\mathcal{C}(T)$; i.e., it is not sure that the set Θ_n^T of continuous maps $T \to \Theta_n$ is a weak-parameter space for $\mathcal{M}_n^T$.

In fact, the above difficulties come from the fact that very little is known about the quotient space $\Theta_n/\sim$, though the equivalence relation on Θ_n has been pointed out and studied by several authors, mainly in the context of neural networks. The case where *h* is the hyperbolic tangent has been studied in [1,16], and extended in [9,10] to the case where *h* is asymptotically bounded. In fact, the major concern of those works has been to reduce the set of allowable values for the parameter vector θ_n to improve the optimization procedure involved in function approximation from a finite data set. There is also the closely related work of Buhmann and Pinkus, which is to be found in [4,5].

We do not pursue this direction here. Instead, we ask if one may obtain a dense set of ridge function fields, built via continuous parameter maps $\xi : T \to \Theta_n$. More precisely, we ask if

$$\cup_n \{\zeta \in \mathcal{M}_n^T : \zeta = i_n \circ \xi, \xi \in \Theta_n^T\} \tag{6}$$

is dense in $(\mathcal{C}(X))^T$, where we recall that Θ_n^T denotes the set of continuous maps $T \to \Theta_n$. The answer is yes, and we begin by stating the following proposition.

Proposition 6. *Let T be a compact metric Hausdorff space, and let $\mathcal{G}$ be a fundamental set in $\mathcal{C}(X)$. Then the set*

$$\cup_n \left\{ f \in \mathcal{C}(X \times T) : f(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} f_i(\mathbf{t}) g_i(\mathbf{x});\ f_i \in \mathcal{C}(X);\ g_i \in \mathcal{G} \right\} \tag{7}$$

is dense in $\mathcal{C}(X \times T)$.

Lemma 7. *Let X and T be two locally compact Hausdorff spaces, and let A be the set spanned by functions of the form $f(\mathbf{x})g(\mathbf{t})$, where $f \in \mathcal{C}(X)$, and where $g \in \mathcal{C}(T)$. Then A is dense in $\mathcal{C}(X \times T)$.*

Proof. Clearly, A is a subalgebra of $\mathcal{C}(X \times T)$ which separates points and vanishes at no point of $X \times T$. Let h_0 be an element of $\mathcal{C}(X \times T)$. To show that $\overline{A} = \mathcal{C}(X \times T)$, it suffices to show that for each compact K of $X \times T$ and each $\varepsilon > 0$, the set

$$B_{K,\varepsilon} = \left\{ h \in \mathcal{C}(X \times T) : \sup_{(x,t) \in K} |h(x,t) - h_0(x,t)| < \varepsilon \right\}$$

has a nonempty intersection with A. Let K_1 and K_2 be compact subsets of, respectively, X and T, such that $K \subset K_1 \times K_2$. The set $\{f_{K1 \times K_2} : f \in A\}$ of restrictions of elements of A to $K_1 \times K_2$ is still an algebra containing constants and vanishing at no point, and is therefore dense in $\mathcal{C}(K_1 \times K_2)$ by the Stone–Weierstrass theorem. Consequently A intersects $B_{K_1 \times K_2, \varepsilon}$. Noting that $B_{K_1 \times K_2, \varepsilon} \subset B_{K,\varepsilon}$, we have also that A intersects $B_{K,\varepsilon}$. □

Proof of Proposition 6. The set defined by Eq. (7) contains the set of functions of the form $f(\mathbf{t}) \sum_{i=1}^n g_i(\mathbf{x})$, which is easily seen to be dense in $\mathcal{C}(X \times T)$ by Lemma 7. □

Corollary 8. *Let $\mathcal{G}$ be fundamental in $\mathcal{C}(X)$, and let T be a compact metric Hausdorff space. Then the set of function fields $\zeta : T \to \mathcal{C}(X)$ such that*

$$\zeta_*(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} c_i(\mathbf{t}) g_i(\mathbf{x}), \tag{8}$$

for some integer n, $c_i \in \mathcal{C}(T)$, and $g_i \in \mathcal{G}$, is dense in $(\mathcal{C}(X))^T$.

Proof. By Theorem 2, there is the homeomorphism $\mathcal{C}(X \times T) \stackrel{\approx}{\longrightarrow} (\mathcal{C}(X))^T$. $\square$

Finally, we arrive at the following two propositions. Recall that the generator function h is such that $\mathcal{M}$ is dense in $\mathcal{C}(\mathbf{R}^d)$.

Proposition 9. *The set of ridge function fields* $\zeta : T \to \mathcal{M}$ *such that*

$$\zeta_*(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} c_i(\mathbf{t}) h(\mathbf{a}_i \mathbf{x} + b_i), \tag{9}$$

for some integer n, $c_i \in \mathcal{C}(T)$, $\mathbf{a}_i \in \mathbf{R}^d$, *and* $b_i \in \mathbf{R}$, *is dense in* $(\mathcal{C}(X))^T$.

Proof. Take $\mathcal{G} = \{\tilde{h} : \mathbf{R}^d \ni \mathbf{x} \mapsto h(\mathbf{a}\mathbf{x} + b) \in \mathbf{R};\, \mathbf{a} \in \mathbf{R}^d;\, b \in \mathbf{R}\}$, i.e., $\mathcal{M} = \operatorname{span} \mathcal{G}$, and apply Corollary 8. $\square$

Proposition 10. *The set of ridge function fields* $\zeta : T \to \mathcal{M}$ *such that*

$$\zeta_*(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} c_i(\mathbf{t}) h\left(\mathbf{a}_i(\mathbf{t})\mathbf{x} + b_i(\mathbf{t})\right), \tag{10}$$

for some integer n, $c_i \in \mathcal{C}(T)$, $\mathbf{a}_i \in \mathcal{C}(T, \mathbf{R}^d)$, *and* $b_i \in \mathcal{C}(T)$, *is dense in* $(\mathcal{C}(X))^T$.

Proof. The set of ridge function fields satisfying (9) is included in the set of ridge function fields satisfying (10). $\square$

In the above proposition, b_i and c_i vary in all of $\mathcal{C}(T)$, and $\mathbf{a}_i$ varies in all of $\mathcal{C}(T, \mathbf{R}^d)$. Now given subsets $\mathcal{F}_{\mathbf{a}} \subset \mathcal{C}(T, \mathbf{R}^d)$, $\mathcal{F}_b \subset \mathcal{C}(T)$, and $\mathcal{F}_c \subset \mathcal{C}(T)$, we consider the set of ridge function fields satisfying (10), where $c_i \in \mathcal{F}_c$, $\mathbf{a}_i \in \mathcal{F}_{\mathbf{a}}$, and $b_i \in \mathcal{F}_b$. In the following proposition, sufficient conditions on $\mathcal{F}_{\mathbf{a}}$, $\mathcal{F}_b$, and $\mathcal{F}_c$ are given for such a set of ridge function fields to be dense in $(\mathcal{C}(X))^T$.

Proposition 11. *Let* $\mathcal{F}_c$ *and* $\mathcal{F}_b$ *be subsets of* $\mathcal{C}(T)$, *and let* $\mathcal{F}_{\mathbf{a}}$ *be a subset of* $\mathcal{C}(T, \mathbf{R}^d)$. *Let* $\mathcal{R}(\mathcal{F}_c, \mathcal{F}_{\mathbf{a}}, \mathcal{F}_b)$ *be the set of ridge function fields* $\zeta : T \to \mathcal{M}$ *such that*

$$\zeta_*(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} c_i(\mathbf{t}) h\left(\mathbf{a}_i(\mathbf{t})\mathbf{x} + b_i(\mathbf{t})\right), \tag{11}$$

for some integer n, $c_i \in \mathcal{F}_c$, $\mathbf{a}_i \in \mathcal{F}_{\mathbf{a}}$, *and* $b_i \in \mathcal{F}_b$. *For* $\mathcal{R}(\mathcal{F}_c, \mathcal{F}_{\mathbf{a}}, \mathcal{F}_b)$ *to be dense in* $(\mathcal{C}(X))^T$, *it is sufficient that* $\mathcal{F}_c$ *and* $\mathcal{F}_{\mathbf{a}}$ *contain the constant functions, and that* $\mathcal{F}_b$ *contains the affine functions.*

Proof. Let $\mathcal{A}$ be the set of approximants of the ridge form over $X \times T$; i.e., $\mathcal{A}$ is spanned by functions of the form

$$f(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} c_i h\left(\mathbf{a}_i \mathbf{x} + \tilde{\mathbf{a}}_i \mathbf{t} + \tilde{b}_i\right), \tag{12}$$

where $c_i, \tilde{b}_i \in \mathbf{R}$, $\mathbf{a}_i \in \mathbf{R}^d$, and $\tilde{\mathbf{a}}_i \in \mathbf{R}^{\dim(T)}$. Let

$$\mathcal{R}_*(\mathcal{F}_c, \mathcal{F}_{\mathbf{a}}, \mathcal{F}_b) = \{\zeta_* : \zeta \in \mathcal{R}(\mathcal{F}_c, \mathcal{F}_{\mathbf{a}}, \mathcal{F}_b)\}. \tag{13}$$

The density of $\mathcal{R}(\mathcal{F}_c, \mathcal{F}_{\mathbf{a}}, \mathcal{F}_b)$ in $(\mathcal{C}(X))^T$ comes from the inclusion $\mathcal{A} \subset \mathcal{R}_*(\mathcal{F}_c, \mathcal{F}_{\mathbf{a}}, \mathcal{F}_b)$, from the density of $\mathcal{A}$ in $\mathcal{C}(X \times T)$, and from the homeomorphism $\mathcal{C}(X \times T) \overset{\approx}{\to} (\mathcal{C}(X))^T$. □

Of particular interest is the ridge function field of the special kind described below. Assume T is a compact subset of $\mathbf{R}^p$. Let $\{\mathbf{t}_1, \ldots \mathbf{t}_{k^p}\}$ be k^p points of $\mathbf{R}^p$, being the k^p vertices of a regular grid of $\mathbf{R}^p$, such that T is included in the smallest p-dimensional cube Ξ containing all of the $\mathbf{t}_i$. Hence $T \subset \Xi$, and $\mathbf{t}_i \in \Xi$ for all $i = 1, \ldots, k^p$.

Let $\gamma_1, \ldots, \gamma_{k^p}$ be k^p real numbers. Let f be a continuous and piecewise differentiable function on Ξ such that $f(\mathbf{t}_i) = \gamma_i$ for all $i = 1, \ldots, k^p$, and defined for all $\mathbf{t} \in \Xi$ such that $\mathbf{t} \neq \mathbf{t}_i \forall i$ by

$$f(\mathbf{t}) = \sum_{j=1}^{2^p} \alpha_j(\mathbf{t}) f(\mathbf{t}_{i_j}). \tag{14}$$

In this equation, the $\mathbf{t}_{i_j}$ are the 2^p immediate neighbours of $\mathbf{t}$ on the grid—i.e., they are the vertices of the p-cube such that $\mathbf{t}$ belongs to its interior and the coefficients $\alpha_j(\mathbf{t})$ are the coefficients of the standard p-dimensional interpolation procedure on the interior of a p-cube. We shall denote by $\mathcal{F}_k$ the set of all such maps. Note that $\mathcal{F}_k \subset \mathcal{F}_{k+1}$. The construction of these maps is illustrated in Fig. 1, in the case where $p = 2$.

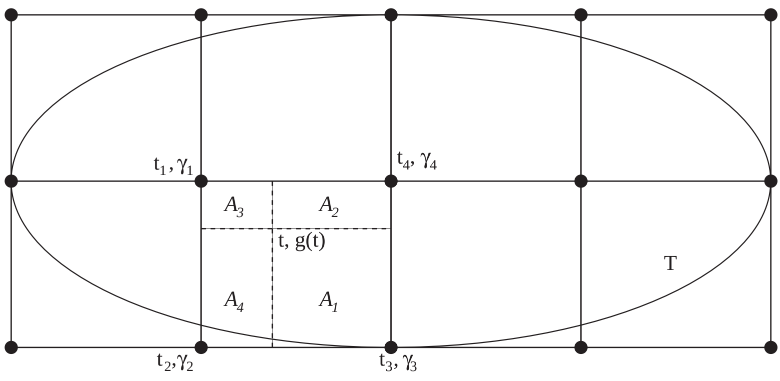

Fig. 1. Construction of a real-valued piecewise differentiable map f by multilinear interpolation. In this picture, T is a two-dimensional ellipsoidal domain covered by a regular 5×3 grid. The value $f(\mathbf{t})$ of f at some point $\mathbf{t}$ in the square with corners located at $\mathbf{t}_1, \ldots, \mathbf{t}_4$ is defined by

$$f(\mathbf{t}) = \sum_{i=1}^{4} \frac{A_i}{A_1 + A_2 + A_3 + A_4} \gamma_i,$$

where the A_i's are the areas of the rectangles defined on the picture. Note that the A_i depend on t.

Let $\mathcal{R}(\mathcal{F}_k)$ be the set of ridge function fields $\zeta : T \to \mathcal{M}$ such that ζ_* is of the form

$$\zeta_*(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} \tilde{c}_i(\mathbf{t}) h\left(\tilde{\mathbf{a}}_i(\mathbf{t})\mathbf{x} + \tilde{b}_i(\mathbf{t})\right), \tag{15}$$

where $\tilde{\mathbf{a}}_i$, $\tilde{b}_i$, $\tilde{c}_i$ are restrictions to T of functions $\mathbf{a}_i$, b_i, c_i defined on Ξ, and such that b_i, c_i, and the components a_i^j, $j = 1, ..., d$ of $\mathbf{a}_i$ belong to $\mathcal{F}_k$. As above, $\mathbf{x} \in X \subset \mathbf{R}^d$, $\mathbf{a}_i \in \mathcal{C}(T, \mathbf{R}^d)$, $b_i \in \mathcal{C}(T)$, and $c_i \in \mathcal{C}(T)$.

From a practical point of view, the sets $\mathcal{R}(\mathcal{F}_k)$ are especially interesting, since their elements may be constructed in a rather simple way. Furthermore, by Proposition 11, the sets $\mathcal{R}(\mathcal{F}_k)$ are dense in $(\mathcal{C}(X))^T$, for all $k \geq 1$, which illustrates the significance of the above results.

4. Concluding remarks

So far we have given density results on sets of ridge function fields. It would be interesting to pursue this work by investigating the rate of approximation of some class of function fields, by ridge function fields. One might expect an interplay between the number n of ridge functions and the complexity of the parameter map. For instance, in the above field, constructed on a regular grid, it would be interesting to examine the dependence of the approximation rate on n, and on the number of points in the grid. Another research direction that would be worth exploring is the geometry and topology of sets of ridge function-based approximants. We have seen that this is the major source of problems in getting a parameterization, or weak parameterization, of sets of continuous ridge function fields.

References

[1] F. Albertini, E. Sontag, V. Maillot, Uniqueness of weights for neural networks, in: R. Mammone (Ed.), Artificial Neural Networks for Speech and Vision, Chapman & Hall, London, 1993, pp. 113–125.
[2] A. Barron, Universal approximation bounds for superpositions of a sigmoidal function, IEEE Trans. Inform. Theory 39 (3) (1993) 930–945.
[3] G. Bredon, Topology and Geometry, Graduate Texts in Mathematics, vol. 139, Springer, Berlin, New York, 1993.
[4] M. Buhmann, A. Pinkus, On a recovery problem, Ann. Numer. Math. 4 (1997) 129–142.
[5] M. Buhmann, A. Pinkus, Identifying linear combinations of ridge functions, Adv. Appl. Math. 22 (1999) 103–118.
[6] M. Burger, A. Neubauer, Error bounds for approximation with neural networks, J. Approx. Theory 112 (2001) 235–250.
[7] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Systems 2 (1989) 303–314.
[8] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989) 359–366.
[9] V. Kurkova, P. Kainen, Functionally equivalent feedforward neural networks, Neural Comput. 6 (1994) 543–558.
[10] V. Kurkova, P. Kainen, Singularities of finite scaling functions, Appl. Math. Lett. 9 (2) (1996) 33–37.
[11] V.Y. Lin, A. Pinkus, Fundamentality of ridge functions, J. Approx. Theory 75 (1993) 295–311.

[12] V. Maiorov, On best approximation by ridge functions, J. Approx. Theory 99 (1999) 68–94.
[13] Y. Makovoz, Random approximants and neural networks, J. Approx. Theory 85 (1996) 98–109.
[14] Y. Makovoz, Uniform approximation by neural networks, J. Approx. Theory 95 (1998) 215–228.
[15] P. Niyogi, F. Girosi, Generalization bounds for function approximation from scattered noise data, Adv. Comput. Math. 10 (1999) 51–80.
[16] H. Sussman, Uniqueness of the weights for minimal feedforward nets with a given input–output map, Neural Networks 5 (4) (1992) 589–593.
[17] B. Vostrecov, M. Kreines, Approximation of continuous functions by superpositions of plane waves, Dokl. Akad. Nauk SSSR 2 (1961) 1237–1240.