Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Fuzzy Sets and Systems 138 (2003) 441 – 453

FUZZY sets and systems

www.elsevier.com/locate/fss

Notes on the approximation rate of fuzzy KH interpolators☆

Domonkos Tikk[a,b]

[a] Department of Telecommunications and Telematics, Budapest University of Technology and Economics, H-1117 Budapest, Magyar Tudósok Körútja 2, Hungary
[b] Integrated Intelligent Systems Japanese–Hungarian Laboratory, 1111 Budapest, Műegyetem rakpart 3, Hungary

Received 1 March 2001; received in revised form 25 July 2002; accepted 2 August 2002

Abstract

This paper investigates the approximation behaviour of the Kóczy–Hirota (KH) interpolative fuzzy controllers. First, in accordance with the remarks in (Fuzzy Sets and Systems 125(1) (2002) 105), it is pointed out that it is a fuzzy generalization of the Shepard operator. Shepard operator has thoroughly studied by approximation theorist since the mid-1970s. Exploiting the aforementioned relationship, we establish analog results on the approximation rate of KH controllers. The optimal order and class of approximation (saturation problem) are determined for certain values of the exponent λ. Corresponding results on the modified alpha-cut based interpolation method, being an improvement of the KH interpolator, are also provided. The results offer trade-off facilities between approximation accuracy and the number of rules. As a consequence, the necessary and sufficient number of rules can be determined for a prescribed accuracy.

Keywords: Fuzzy inference systems; Analysis; KH interpolator; Shepard operator; Universal approximation; Approximation error estimates; Saturation problem; Trade-off between accuracy and number of rules

1. Introduction

Recently, it was pointed out [28] that a special version of the interpolative fuzzy Kóczy–Hirota (KH) interpolator (see details in Section 2.1) is universal approximator in the sense that it is able to approximate any continuous function on a compact domain with arbitrary accuracy with respect

☆ Research funded by the Hungarian Scientific Research Fund (OTKA) Grants No. D034614, T34212 and T34233.
E-mail address: tikk@ttt.bme.hu (D. Tikk).

PII: S0165-0114(02)00387-1

to the L_p norm $p \in [1, \infty]$. In the same paper it was remarked without any proof that the KH interpolator can be considered as a fuzzy generalization of the Shepard interpolator (Section 2.2). Now we show their exact connection in Section 3.

Universal approximation theorems on soft computing techniques (see e.g. [4,17,30] for fuzzy systems, and [6,9] for neural nets) have been usually criticized due to their solely existential nature [11,29]. In the recent past, there arose an effort to give constructive proofs, or to determine the number of "building blocks" (antecedents or rules in fuzzy, hidden neurons in neural terminology) as a function of the accuracy (see e.g. [3,16,20,32,33]). See [29] for a survey on this topic. Naturally, to obtain such results one has to restrict somehow the set of continuous functions, usually requiring some smoothness conditions on the approximated function. In approximation theory, the optimal order of convergence is called the saturation order, and the subset of continuous functions which can be approximated with the specified order is termed the saturation class. Hitherto, saturation classes and orders have not been determined for soft computing techniques.

The approximation result for KH controllers published in [28] is purely existential, no trade-off between the number of the rules and the accuracy was determined. In this paper, we make up this deficiency, and derive theorems for the approximation rate of KH controllers based on their relationship with Shepard operator (Section 4).

2. Preliminaries

2.1. The KH interpolator

The fuzzy rule based interpolation technique (or briefly: fuzzy interpolation) was proposed to provide an inference mechanism suitable for rule bases containing gaps. The first method was introduced in [12], and is termed KH interpolation. It creates the conclusion using its α-cuts based on the extension principle and the resolution principle. For every α value of the important level set (e.g., for triangular and trapezoidal shaped membership function that is 0 and 1), it determines the conclusion as

$$B^*_{\alpha C} = \frac{\sum_{i=1}^{n} B_{i\alpha C} 1/d_C(A^*_{\alpha C}, A_{i\alpha C})}{\sum_{k=1}^{n} 1/d_C(A^*_{\alpha C}, A_{k\alpha C})}, \tag{1}$$

where A_i and B_i $(i = 1, \ldots, n)$ denote the antecedents neighbouring the observation A^*, and the corresponding consequents fuzzy sets, respectively. $C \in \{L, U\}$, where L and U refer to the lower and upper extreme of the α-cut. Finally, the function $d_C : \mathbb{R} \times \mathbb{R} \to [0, +\infty)$ is an appropriate lower/upper distance function (cf. [13]). If $n = 2$, the approach is termed linear interpolation.

For the applicability of the method the involved fuzzy sets should fulfil the following ordering:

$$A_i \prec_X A^* \prec_X A_{i+1}, \quad B_i \prec_Y B_{i+1}, \tag{2}$$

where $\prec_X$ and $\prec_Y$ are proper partial orderings on the multidimensional input and the one-dimensional output space, respectively. These relations mean that the observation should be flanked by antecedents from both sides (no extrapolation is allowed), and a similar ordering should exist for the corresponding consequents.

In [28], the stabilized version of the KH interpolation is introduced, when we take λth power of the distance function requiring that the exponent λ cannot be smaller than N, the dimension of the antecedents

$$B^*_{\alpha C} = \frac{\sum_{i=1}^{n} B_{i\alpha C} 1/d_C^{\lambda}(A^*_{\alpha C}, A_{i\alpha C})}{\sum_{k=1}^{n} 1/d_C^{\lambda}(A^*_{\alpha C}, A_{k\alpha C})} \quad (\lambda \geqslant N) \tag{3}$$

Now, we recall the universal approximation theorem of the stabilized KH approach [28].

Definition 1. Let $\mathbb{R}^N \supset \Omega = [a_1, b_1] \times \cdots \times [a_N, b_N]$, further let $\{\Gamma_n\}_{n=1}^{\infty}$ be a sequence of finite subsets of Ω with $\#\Gamma_n = n$. If

$$\forall \varepsilon > 0 \exists n_0 \ \forall \omega \in \Omega \ \forall n \geqslant n_0 : \left| \frac{\#(\Gamma_n \cap \omega)}{\#\Gamma_n} - \frac{|\omega|}{|\Omega|} \right| < \varepsilon \tag{4}$$

then the set Γ_n are uniformly distributed on the domain Ω. Here $\#(\Gamma_n \cap \omega)$ denotes the cardinality of the finite set $(\Gamma_n \cap \omega)$ and $|\omega|$ is the Lebesque measure of ω.

Theorem 2. *Consider the L_p norm $\|\cdot\|_p$ with $p \in [1, \infty]$, the domain $\mathbb{R}^N \supset \Omega = [a_1, b_1] \times \cdots \times [a_N, b_N]$ and a continuous function $f : \Omega \to \mathbb{R}$, then for all $x \in \Omega$ the expression*

$$\lim_{n\to\infty} K_n^{\lambda}(f, x) := \lim_{n\to\infty} \sum_{k=1}^{n} f(x_k^{(n)}) \frac{1/\|x - x_k^{(n)}\|_p^{\lambda}}{\sum_{j=1}^{n} 1/\|x - x_j^{(n)}\|_p^{\lambda}} \tag{5}$$

is equal to $f(x)$, where measurement points $x_k^{(n)}$ are uniformly distributed on Ω in the sense of (4), and $\lambda \geqslant N$.

Observe that when $N = 1$ (5) yields (3) with some apparent substitutions: x_k for $A_{i\alpha C}$, $f(x_k)$ for $B_{i\alpha C}$, x for $A^*_{\alpha C}$ and $|x - x_k|_p$ for $d_C(A^*_{\alpha C}, A_{i\alpha C})$.

Without the loss of generality, in the next we suppose that $\bar{\Omega} = [0, 1]$, because with a proper linear transformation every compact domain can be mapped into each other. From now, we mean the stabilized KH interpolator (3) on the term "KH interpolator".

Let we briefly summarize here the state-of-the-art of fuzzy rule based interpolation. The original KH method was criticized by several authors because it does not always give directly interpretable conclusion. To alleviate partly or completely this shortcoming various approaches were proposed: among others the so-called solid cutting method [1], the use of approximate reasoning in vague environment [18,19], the modified alpha-cut based interpolation (MACI) method [26], and a method that aims at preserving the relative fuzziness of the conclusion [8,14]. The approach described in [31] shows advantageous properties if the input is multi-dimensional. Recently, an axiomatic characterization of fuzzy rule based interpolation was worked out [10]. For an overview see e.g. [15] or the comparative work [21].

2.2. The Shepard operator

The Shepard interpolation method was first introduced in [22] for arbitrarily placed bivariate data as

$$S_0(f,x,y) = \begin{cases} f(x_i,y_i) & \text{if } (x,y)=(x_i,y_i) \text{ for some } i, \\ \dfrac{\sum_{i=0}^n f(x_i,y_i)/d_i^\lambda}{\sum_{i=0}^n 1/d_i^\lambda} & \text{otherwise,} \end{cases} \tag{6}$$

where measurement points $x_i,\ y_i$ $(i=0,\dots,n)$ are irregularly spaced on the domain of $f\in\mathbb{R}^2\to\mathbb{R}$, $\lambda>0$, and $d_i=[(x-x_i)^2+(y-y_i)^2]^{1/2}$ is the Euclidean distance from the ith knot point (x_i,y_i). This function can be used typically when a surface model is required to interpolate scattered spatial measurements (e.g. in pattern recognition, geology, cartography, earth sciences, fluid dynamics and many others).

Beside the application oriented investigation of Shepard's method such as [2], an increasing interest has arisen from mathematical researchers to examine the approximation property of formula (6). For a more general analysis of these properties, operator (6) was reformulated in the more suitable and more concise form as

$$S_n^\lambda(f,x) = \frac{\sum_{i=0}^n f(x_i)(x-x_i)^{-\lambda}}{\sum_{i=0}^n (x-x_i)^{-\lambda}}, \quad \lambda>0,\ n=1,2\dots \tag{7}$$

for an arbitrary $f\in C[0,1]$, where x_i $(i=0,\dots,n)$, in general, denotes the nodes of the equidistant distribution of the domain $[0,1]$. We recall that fixing the domain to the interval $[0,1]$ does not mean any restriction.

The possible use of rational functions of type (7) as approximating means was first discovered by J. Balázs. (After his name this operator is often termed Balázs–Shepard operator in the literature of approximation theory.) The main advantages of rational function interpolator (7) compared to traditional polynomial or trigonometrical approximators is that it always converges to the approximated function independently from the selection of knot points x_i. The properties of operator (7) were widely investigated by mostly Hungarian and Italian mathematicians; see e.g. [5,7,23–25].

3. The fuzziness of the approximation

As it was shown in [28], for a fixed $\alpha\in[0,1]$ and $C\in\{L,U\}$ the input–output function of the (stabilized) KH interpolator, $K_n^\lambda(f,x)$ coincides with the Shepard operator. Therefore, we can consider the family of $K_n^\lambda(f,x)$ functions as a generalization of the $S_n^\lambda(f,x)$. Here, first we aim at clarifying what is nature of this generalization, how the approximated functions for various α and/or C values differ. It is obvious, that the family of KH functions tailor the same approximated function, if all the involved fuzzy sets are crisp. In the next, we investigate how the conclusion depends on the fuzziness of the antecedents and consequents.

First, let us only assume that the modulus of continuity of the approximated function $f:[0,1]\to\mathbb{R}$ is known

$$\omega(f,n^{-1}) = \max_{\substack{x,y\in[0,1]\\ |x-y|\leqslant n^{-1}}} |f(x)-f(y)|. \tag{8}$$

Due to the uniform distribution (4) of the knot points we can estimate the difference of the adjacent knot points by n^{-1}.

Therefore, we can estimate the support of a consequent fuzzy set, or in other words, its fuzziness by quantity defined in (8):

$$\max_{i=1}^{n}(\mathrm{supp}(B_i)) \leqslant \omega(f,n^{-1}). \tag{9}$$

Note, that this is a very rough estimation. We can sharpen it easily under certain circumstances to be discussed later.

Let us estimate now the difference of two $K_n^\lambda(f,x)$ operators. The largest difference appears when the minimum and the maximum of the support are calculated. That is when $\alpha=0$ is fixed, and $C=L$ and $C=U$, respectively.

Let

$$K_1^{(n)} = \frac{\sum_{i=1}^n (\inf B_{i0})^{(n)} 1/d^\lambda((\inf A_0^*)^{(n)},(\inf A_{i0})^{(n)})}{\sum_{j=1}^n 1/d^\lambda((\inf A_0^*)^{(n)},(\inf A_{j0})^{(n)})}$$

$$K_2^{(n)} = \frac{\sum_{i=1}^n (\sup B_{i0})^{(n)} 1/d^\lambda((\sup A_0^*)^{(n)},(\sup A_{i0})^{(n)})}{\sum_{j=1}^n 1/d^\lambda((\sup A_0^*)^{(n)},(\sup A_{j0})^{(n)})}$$

be the two farthest points of the interpolated conclusion. The superscript $^{(n)}$ refers to the knot point system consisting of n points. Then

$$\begin{aligned}\lim_{n\to\infty} |K_1^{(n)}-K_2^{(n)}| &\leqslant \Bigg| \frac{\sum_{i=1}^n (\inf B_{i0})^{(n)} 1/d^\lambda((\inf A_0^*)^{(n)},(\inf A_{i0})^{(n)})}{\sum_{j=1}^n 1/d^\lambda((\inf A_0^*)^{(n)},(\inf A_{j0})^{(n)})} \\ &\quad - \frac{\sum_{i=1}^n ((\inf B_{i0})^{(n)}+\omega(f,n^{-1})) 1/d^\lambda((\sup A_0^*)^{(n)},(\sup A_{i0})^{(n)})}{\sum_{j=1}^n 1/d^\lambda((\sup A_0^*)^{(n)},(\sup A_{j0})^{(n)})}\Bigg| \\ &\leqslant \Bigg| \frac{\sum_{i=1}^n (\inf B_{i0})^{(n)} 1/d^\lambda((\inf A_0^*)^{(n)},(\inf A_{i0})^{(n)})}{\sum_{j=1}^n 1/d^\lambda((\inf A_0^*)^{(n)},(\inf A_{j0})^{(n)})} \\ &\quad - \frac{\sum_{i=1}^n (\inf B_{i0})^{(n)} 1/d^\lambda((\sup A_0^*)^{(n)},(\sup A_{i0})^{(n)})}{\sum_{j=1}^n 1/d^\lambda((\sup A_0^*)^{(n)},(\sup A_{j0})^{(n))}}\Bigg| + \omega(f,n^{-1}) \\ &\leqslant \omega(f,n^{-1}). \end{aligned} \tag{10}$$

As n^{-1} converges to zero, and the knot points are uniformly distributed, the support of antecedents and observation fuzzy sets should also vanish due to the condition of their ordering (cf. (2)).

Therefore, the two expressions of distance become identical, hence the absolute value of the difference of the fractions vanishes. Note, that the support of the consequents does not necessarily vanishes as n^{-1} tends to zero.

The obtained result means that the maximal difference between two operators in a family of Shepard type approximations is bounded by the modulus of continuity of the approximated functions, and this value can be theoretically arbitrary large. Observe that the result only depends on the estimated support size (9). However, it is not very reasonable to model a significant change in the approximated function with only one consequent fuzzy set. Therefore, the rough estimation of (9) and (10) can be improved, if we suppose that antecedents and consequents are fuzzy numbers modelling the measured input–output samples of the approximated function. In this case we can set the consequents' maximum support length to δ, being e.g. the margin of error of the measuring tool

$$\max_{i=1}^{n}(\mathrm{supp}(B_i)) \leqslant \delta. \tag{11}$$

With analog reasoning as in (10), we can get

$$\lim_{n\to\infty} |K_1^{(n)} - K_2^{(n)}| \leqslant \delta.$$

Based on the previous thoughts we can state the following

Proposition 3. *Under the conditions of Theorem* 2, *the fuzziness of the conclusion of the stabilized* KH *interpolator is bounded by the maximum fuzziness of the consequent fuzzy sets*, *if* n, *the number of knot points tends to* ∞.

It means that the fuzziness of the conclusion is estimated from above by the maximum fuzziness of consequent fuzzy sets, i.e. we can select the number of knot points in such a way that the estimate is valid with an error of order at last $\mathrm{O}(n^{-1})$. This error margin is estimated based on the uniform distribution of knot points on the unit interval, and this is usually incomparably smaller than $\omega(f, n^{-1})$.

As it was already remarked, the KH operator does not always give directly interpretable fuzzy conclusion. This problem was investigated by several researchers, and alternative solutions and methods were proposed [1,18,19,8,14]. Among those, the MACI method [26] is the most advantageous, because it eliminates the abnormality problem and maintains the low computational requirements of the KH method. Moreover, its generalized version also possesses the universal approximation property [26,27].

The general MACI method tailors the conclusion as a finite sum of KH interpolators. Formally, using the denotation of (5)

$$\lim_{n\to\infty}(\mathrm{MACI}^{(k)})_n^{\lambda}\left(\sum_{j=1}^{k} f^j, x\right) = \sum_{j=1}^{k}\sum_{i=1}^{n} f^j(x_{ij}^{(n)})\frac{1/\|x - x_{ij}^{(n)}\|_p^{\lambda}}{\sum_{\ell=1}^{n} 1/\|x - x_{\ell j}^{(n)}\|_p^{\lambda}} \quad \lambda \geqslant N, \tag{12}$$

where f^j $(j = 1, \ldots, k)$ are the functions approximated by the KH interpolators.

It was shown in [2], the fuzziness of the conclusion determined by the MACI method is usually larger than the one created by the KH interpolator, due to the construction of the conclusion. In consequence, it is not surprising that the conclusion is more fuzzy even if the number of rules tends to infinity.

Proposition 4. *Let k be the number of* KH *interpolators needed to create* MACIs *conclusion, approximating the functions f_i with δ_i accuracy $(i=1,\dots,k)$. Further, let conditions of Theorem* 2 *hold. Then the conclusion's fuzziness is less than $\sum_{i=1}^{k}|\delta_i|$, if the number of knot points tends to ∞.*

The proof is straightforward, it employs the result of Proposition 3 and combines it with the triangle inequality.

Because of Propositions 3 and 4, theorems for Shepard operators are convertible for the KH (or MACI) interpolators, bearing in mind that all derived results have an uncertainty factor (or fuzziness) due to the fuzziness of the consequents.

4. Main results

The approximation property and the saturation problem of the Shepard operator were investigated for various λ values in [5,23–25]. Here we recall the proof of Szabados [25], which gives a complete analysis for all $\lambda \geqslant 1$. When $\lambda < 1$, the operator does not converge for all $f(x) \in C[0,1]$, so it is of no interest for our investigations.

Theorem 5 (Szabados [25]). *The approximation order of the operator $S_n^{\lambda}(f,x)$ is*

$$\|f(x) - S_n^{\lambda}(f,x)\| = \begin{cases} \mathrm{O}(\omega(f,n^{-1})) & \text{if } \lambda > 2, \\ \mathrm{O}(n^{1-\lambda})\displaystyle\int_{1/n}^{1} t^{-\lambda}\omega(f,t)\,\mathrm{d}t & \text{if } 1 < \lambda \leqslant 2, \\ \mathrm{O}(\log^{-1} n)\displaystyle\int_{1/n}^{1} t^{-\lambda}\omega(f,t)\,\mathrm{d}t & \text{if } \lambda = 1, \end{cases} \tag{13}$$

for any $f \in C[0,1]$.

Proof. Let us suppose equispaced knot points systems on the unit interval, i.e.

$$x_i = \frac{i}{n}, \quad \{x_i\}^{(n)} = \{x_i | i = 0,\dots,n\}. \tag{14}$$

Further, let $x_j \in \{x_i\}^{(n)}$, $j=0,\dots,n$ be the closest knot point to an arbitrary $x \in [0,1]$, $x \notin \{x_i\}^{(n)}$

$$\left|x - \frac{j}{n}\right| = \min_{0 \leqslant i \leqslant n}\left|x - \frac{i}{n}\right|$$

(if this definition is not unique take any of the two possibilities). Then

$$|f(x)-S_n^{\lambda}(f,x)| \leqslant \frac{\sum_{i=0}^{n}|f(x)-f(i/n)|\cdot|x-i/n|^{-\lambda}}{\sum_{i=0}^{n}|x-i/n|^{-\lambda}}.$$

The order of the denominator is evidently

$$\left(\sum_{i=0}^{n}\left|x-\frac{i}{n}\right|^{-\lambda}\right)^{-1} = \mathrm{O}(n^{-\lambda}\log^{-\lceil 1/\lambda\rceil} n) \quad (\lambda \geqslant 1), \tag{15}$$

so

$$\begin{aligned}|f(x)-S_n^{\lambda}(f,x)| &\leqslant \left|f(x)-f\left(\frac{j}{n}\right)\right| + \mathrm{O}(n^{-\lambda}\log^{-\lceil 1/\lambda\rceil} n)\\ &\quad \sum_{i\neq j}\left|f(x)-f\left(\frac{i}{n}\right)\right|\cdot\left|x-\frac{i}{n}\right|^{-\lambda}\\ &\leqslant \omega\left(f,\frac{1}{2n}\right) + \mathrm{O}(n^{-\lambda}\log^{-\lceil 1/\lambda\rceil} n)\\ &\quad \sum_{i\neq j}\left(\frac{|j-i|-1/2}{n}\right)^{-\lambda}\omega\left(f,\frac{|j-i|-1/2}{n}\right)\\ &\leqslant \omega(f,n^{-1}) + \mathrm{O}(n^{-\lambda}\log^{-\lceil 1/\lambda\rceil} n)\sum_{i=1}^{n}\left(\frac{i}{n}\right)^{-\lambda}\omega\left(\frac{i}{n}\right).\end{aligned} \tag{16}$$

Because $\omega(f,k/n)\leqslant k\omega(f,1/n)$, for $\lambda>2$ we have

$$\begin{aligned}|f(x)-S_n^{\lambda}(f,x)| &\leqslant \omega(f,n^{-1})\left\{1+\mathrm{O}(n^{-\lambda})\sum_{i=1}^{n}\frac{i^{1-\lambda}}{n^{-\lambda}}\right\}\\ &= \mathrm{O}(\omega(f,n^{-1})).\end{aligned}$$

If $1\leqslant\lambda\leqslant 2$, then

$$|f(x)-S_n^{\lambda}(f,x)| \leqslant \omega(f,n^{-1}) + \mathrm{O}(n^{1-\lambda}\log^{-\lceil 1/\lambda\rceil} n)\int_{1/n}^{1} t^{-\lambda}\omega(f,t)\,\mathrm{d}t,$$

whence the last two statements of the theorem follows. □

The theorem gives some immediate results on the saturation problem. For its precise characterization we need the following definition.

Definition 6. A function $f:[0,1]\to\mathbb{R}$ is called Lipschitz continuous with Lipschitz coefficient α (notation: $f\in\mathrm{Lip}\,\alpha$) if

$$|f(x)-f(y)|\leqslant\alpha|x-y| \quad \text{for all } x,y\in[0,1]. \tag{17}$$

From Theorem 5 it is obvious that if $f(x) \in \mathrm{Lip}\,1$ and $\lambda > 2$, then the saturation order is

$$\|f(x) - S_n^{\lambda}(f,x)\| = \mathrm{O}(n^{-1}). \tag{18}$$

In fact, as it is proved in [23], (18) holds if and only if $f(x) \in \mathrm{Lip}\,1$. Furthermore,

$$\|f(x) - S_n^{\lambda}(f,x)\| = \mathrm{o}(n^{-1}) \tag{19}$$

if and only if $f(x) = \text{const}$. Thus the saturation problem is completed for $\lambda > 2$.

If $\lambda = 2$, even with $f(x) \in \mathrm{Lip}\,1$ Theorem 5 gives only

$$\|f(x) - S_n^{\lambda}(f,x)\| = \mathrm{O}\left(\frac{\log n}{n}\right). \tag{20}$$

This result can be improved to $\mathrm{O}(n^{-1})$ under stronger restriction on $f(x)$:

Theorem 7 (Szabados [25]). *If* $f'(x) \in [0,1]$ *and*

$$\int_0^1 t^{-1}\omega(f',t)\,\mathrm{d}t < \infty \tag{21}$$

further

$$f'(0) = f'(1) = 0 \tag{22}$$

then

$$\|f(x) - S_n^{2}(f,x)\| = \mathrm{O}(n^{-1}). \tag{23}$$

It is also shown that we need both conditions, so expressions (21)–(22) on $f(x)$ cannot be weakened. This is because, e.g. $f(x) = x$ satisfies (21) but not (22), and on the other hand

$$f(x) = \frac{x(1-x)}{\log(x(1-x))}$$

satisfies (22) but not (21), and (23) does not hold for either function. However, the saturation problem is not solved for $\lambda = 2$, because the converse result of Theorem 7, that is (23) implies conditions (21) and (22), has not been proved yet.

Even less is known for the $1 \leqslant \lambda \leqslant 2$ case, when only best error estimates are obtained. From Theorem 5,

$$\mathrm{O}(n^{1-\lambda}), \quad 1 < \lambda < 2 \quad \text{and} \quad \mathrm{O}(\log^{-1} n), \quad \lambda = 1 \tag{24}$$

provided that

$$\int_0^1 t^{-\lambda}\omega(f,t)\,\mathrm{d}t < \infty \tag{25}$$

holds.

The results of Theorems 5 and 7 can be carried over directly for KH and MACI interpolators under the following conditions:

(1) The approximated function f is univariate, i.e. the input universe of the interpolator is one dimensional.
(2) The knot points are equispaced on the unit interval (14).

The second condition can be weakened, and changed to uniform distribution, since (4) ensures asymptotically the same behaviour, which yield that expression (15) and estimate (16) remain valid in the proof of Theorem 5, if we substitute the equispaced system (14) by (4).

Corollary 8. *The* KH *interpolator* (5) *saturates with order* $\mathrm{O}(n^{-1})$ *on the class of the functions* $f(x)\in \mathrm{Lip}\,1$, *if* $\lambda>2$, *the knots point system satisfies* (4), *and the input is one dimensional.*

Proof. Because of $f(x)\in \mathrm{Lip}\,1$

$$\omega(f,n^{-1}) = \max_{\substack{x,y\in[0,1]\\ |x-y|\leqslant n^{-1}}} |f(x)-f(y)| \leqslant \frac{1}{n}$$

thus (10) can be estimated as

$$|K_1^{(n)} - K_2^{(n)}| = \mathrm{O}(n^{-1}).$$

Hence, as each member of the family of K_n^{λ} ($\lambda>2$) saturate on the class of Lip 1 with order $\mathrm{O}(n^{-1})$ and the difference between the family members are also in order $\mathrm{O}(n^{-1})$, then the saturation order of the whole family is $\mathrm{O}(n^{-1})$ with the class Lip 1. □

Analog result holds for MACI interpolators:

Corollary 9. *The* MACI *interpolator being a finite sum of* KH *interpolators* (5) *saturates with order* $\mathrm{O}(n^{-1})$ *on the class of the functions* $f(x)\in \mathrm{Lip}\,1$, *if* $\lambda>2$, *the knots point system satisfies* (4), *and the input is one dimensional.*

The proof is analog with the previous one, thus omitted.
The following statements are derived from Theorem 7.

Corollary 10. *The best approximation order of the* KH *interpolator* $K_n^{\lambda}(f,x)$ $(1\leqslant\lambda\leqslant 2)$ *is*

$$\|f(x)-K_n^{\lambda}(f,x)\| = \begin{cases} \mathrm{O}(n^{-1}) & \text{if } \lambda = 2,\\ \mathrm{O}(n^{1-\lambda}) & \text{if } 1\leqslant\lambda<2,\\ \mathrm{O}(\log^{-1} n) & \text{if } \lambda = 1 \end{cases}$$

on the class of the functions $f(x)\in \mathrm{Lip}\,1$, *if* $f(x)$ *further satisfies conditions* (21) *and* (22) *for* $\lambda=2$, *and condition* (25) *for* $1\leqslant\lambda<2$, *and furthermore if the knots point system is in accordance with* (4), *and the input is one dimensional.*

Corollary 11. *The best approximation order of the* MACI *interpolator*, *being a finite sum of* KH *interpolators* $K_n^{\lambda}(f,x)$ $(1 \leqslant \lambda \leqslant 2)$,

$$\|f(x) - \mathrm{MACI}_n^{\lambda}(f,x)\| = \begin{cases} \mathrm{O}(n^{-1}) & \text{if } \lambda = 2, \\ \mathrm{O}(n^{1-\lambda}) & \text{if } 1 \leqslant \lambda < 2, \\ \mathrm{O}(\log^{-1} n) & \text{if } \lambda = 1 \end{cases}$$

on the class of the functions $f(x) \in \mathrm{Lip}\,1$, *if* $f(x)$ *further satisfies conditions* (21) *and* (22) *for* $\lambda = 2$, *and condition* (25) *for* $1 \leqslant \lambda < 2$, *and furthermore*, *under the uniform distribution of the knot points* (4), *and with one dimensional input*.

The proofs are similar as in Corollary 8.

The results obtained so far in this paper contribute to the investigation of the approximation behaviour of fuzzy inference methods. The saturation order and class of an operator determine the optimal order of convergence, or in other words, the best achievable approximation speed by the given operator. In our results the approximation error is determined in terms of the modulus of continuity. This construction offers a natural way to determine the *necessary and sufficient* number of rules for the approximation of a given function. If the required number of rules is known in terms of accuracy we can simply do trade-off between them: if we prefer higher accuracy in approximation the number of rules should be set based on Corollary 10 or 11; if we search for the best achievable approximation for a given rule set its order is also determined by the above expressions.

Let us assume that modulus of continuity w.r.t. to the approximated function is known or estimated in terms of the number of knot points. Therefore, given a uniformly distributed knot point set, its density uniquely determines the value of $\omega(f, n^{-1})$, because it also sets the number of rules, n, as the location of knot points determines the core of fuzzy sets (see also the remark after Theorem 2). Thus, the density of the distribution defines the order of approximation error through the modulus of continuity by Corollary 10 or 11 for KH and MACI method, respectively. Vice versa, given a prescribed accuracy, we can calculate the order of the sufficient number of rules based on the modulus of continuity. This value coincides with the necessary number of rules, because Corollaries 8 and 9 circumscribes the conditions of the best achievable approximation. Naturally, these considerations hold if the conditions of Corollaries 8 and 9 are satisfied.

In comparison with related works [3,20,32,33] this paper sets out a new direction in universal approximation field of soft computing techniques. While hitherto only upper bounds for the sufficient number of rules were determined or estimated, we give estimations also for the necessary number of rules and determine the optimal order of convergence and the corresponding class of functions.

5. Conclusion

In this paper we derived approximation rates for the KH interpolator and for its modification the MACI method. The achieved results exploited the connection of Shepard and KH interpolators. We determined for $\lambda > 2$ the saturation class and order for KH interpolator, and we got best approximation error estimates for $1 \leqslant \lambda \leqslant 2$ case. Analog results for MACI method were also outlined. The results offer trade-off facilities between the number of rules and best possible accuracy.

References

[1] P. Baranyi, T.D. Gedeon, Rule interpolation by spatial geometric representation, in: Proc. IPMU'96, Granada, 1996, pp. 483–488.

[2] R.E. Barnhill, R.P. Dube, F.F. Little, Properties of Shepard's surfaces, Rocky Mountain J. Math 13 (1991) 365–382.

[3] E.K. Blum, L.K. Li, Approximation theory and feedforward networks, Neural Networks 4 (4) (1991) 511–515.

[4] J.L. Castro, Fuzzy logic controllers are universal approximators, IEEE Trans. Systems Man Cybernet. 25 (1995) 629–635.

[5] G. Criscuolo, G. Mastroianni, Estimates of Shepard interpolatory procedure, Acta Math. Hungar. 61 (1993) 79–91.

[6] G. Cybenko, Approximation by superposition of sigmoidal functions, Math. Control Signals Systems 2 (1989) 303–314.

[7] B. Della Vecchia, G. Mastroianni, V. Totik, Saturation of the Shepard operators, Approx. Theory Appl. 6 (4) (1990) 76–84.

[8] T.D. Gedeon, L.T. Kóczy, Conservation of fuzziness in rule interpolation, in: Proc. Symp. on New Trends in Control of Large Scale Systems, vol. 1, Herl'any, 1996, pp. 13–19.

[9] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989) 359–366.

[10] S. Jenei, A new approach for interpolation and extrapolation of compact fuzzy quantities, in: Proc. 21th Linz Seminar on Fuzzy Set Theory, 2000, pp. 13–18.

[11] E.P. Klement, L.T. Kóczy, B. Moser, Are fuzzy systems universal approximators? Internat. J. Gen. Systems 28 (2–3) (1999) 259–282.

[12] L.T. Kóczy, K. Hirota, Approximate reasoning by linear rule interpolation and general approximation, Internat. J. Approx. Reason. 9 (1993) 197–225.

[13] L.T. Kóczy, K. Hirota, Ordering, distance and closeness of fuzzy sets, Fuzzy Sets and Systems 60 (1993) 281–293.

[14] L.T. Kóczy, K. Hirota, T.D. Gedeon, Fuzzy rule interpolation by the conservation of relative fuzziness, Technical Report 97/2, Hirota Lab, Department of Intelligent Comp. and Sys. Sci., Tokyo Institute of Technology, Yokohama, 1997.

[15] L.T. Kóczy, D. Tikk, L. Muresan, Fuzzy systems with interpolation, An overview, in: Proc. Joint Ninth IFSA World Congress and 20th NAFIPS Internat. Conf. (IFSA/NAFIPS'2001), Vancouver, BC, Canada, 2001, pp. 2494–2498.

[16] L.T. Kóczy, A. Zorat, Fuzzy systems and approximation, Fuzzy Sets and Systems 85 (1997) 203–222.

[17] B. Kosko, Fuzzy systems as universal approximators, IEEE Trans. Comput. 43 (11) (1994) 1329–1333.

[18] Sz. Kovács, New aspect of interpolative reasoning, in: Proc. IPMU'96, Granada, 1996, pp. 477–482.

[19] Sz. Kovács, L.T. Kóczy, The use of the concept of vague environment in approximate fuzzy reasoning, Tatra Mountains Math. Publ. 12 (1997) 169–181.

[20] V. Kůrková, Approximation of functions by perceptron networks with bounded number of hidden units, Neural Networks 8 (5) (1995) 745–750.

[21] S. Mizik, D. Szabó, P. Korondi, Survey on fuzzy interpolation techniques, in: Proc. IEEE Internat. Conf. on Intelligent Engineering Systems (INES'99), Poprad, Slovakia, 1999, pp. 587–592.

[22] D. Shepard, A two dimensional interpolation function for irregularly spaced data, in: Proc. 23rd ACM Internat. Conf., 1968, pp. 517–524.

[23] G. Somorjai, On a saturation problem, Acta Math. Acad. Sci. Hungar. 32 (1978) 377–381.

[24] J. Szabados, On a problem of R. DeVore, Acta Math. Acad. Sci. Hungar. 27 (1976) 219–223.

[25] J. Szabados, Direct and converse approximation theorems for Shepard operator, J. Approx. Theory Appl. 7 (1991) 63–76.

[26] D. Tikk, P. Baranyi, Comprehensive analysis of a new fuzzy rule interpolation method, IEEE Trans. Fuzzy Systems 8 (3) (2000) 281–296.

[27] D. Tikk, P. Baranyi, Y. Yam, L.T. Kóczy, Stability of a new interpolation method, in: Proc. IEEE Internat. Conf. on System, Man, and Cybernetics (IEEE SMC'99), vol. III, Tokyo, Japan, October, 1999, pp. 7–9.

[28] D. Tikk, I. Joó, L.T. Kóczy, P. Várlaki, B. Moser, T.D. Gedeon, Stability of interpolative fuzzy KH-controllers, Fuzzy Sets and Systems 125 (1) (2002) 105–119.

[29] D. Tikk, L.T. Kóczy, T.D. Gedeon, A survey on the universal approximation and its limits in soft computing techniques, Research Working Paper RWP-IT-01-2001, School of Information Technology, Murdoch University, Perth, WA, 2001, p. 20.
[30] L.X. Wang, Fuzzy systems are universal approximators, in: Proc. IEEE Internat. Conf. on Fuzzy Systems, San Diego, 1992, pp. 1163–1169.
[31] K.W. Wong, T.D. Gedeon, D. Tikk, An improved multidimensional α-cut based fuzzy interpolation technique, in: V. Karri, M. Negnevitsky (Eds.), Proc. Internat. Conf. on Artificial Intelligence in Science and Technology (AISAT 2000), Hobart, Tasmania, Australia, December 2000, pp. 33–38.
[32] H. Ying, Sufficient conditions on uniform approximation of multivariate functions by general Takagi–Sugeno fuzzy systems with linear rule consequents, IEEE Trans. Systems Man Cybernat. Part A 28 (4) (1998) 515–520.
[33] K. Zeng, N.-Y. Zhang, W.-L. Xu, A comparative study on sufficient conditions for Takagi–Sugeno fuzzy systems as universal approximators, IEEE Trans. Fuzzy Systems 8 (6) (2000) 773–780.