Dynamic Systems and Applications, **28**, No. 3 (2019), 533-543 ISSN: 1056-2176

COMMENTS ON THE YUN'S ALGEBRAIC ACTIVATION FUNCTION. SOME EXTENSIONS IN THE TRIGONOMETRIC CASE

NIKOLAY KYURKCHIEV

Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24 Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: In this paper we study the asymptotic behavior of the Hausdorff distance between Heaviside function and the Yun's activation function of algebraic type.

We also consider a new activation function of trigonometric type.

Numerical examples using *CAS Mathematica*, illustrating our results are given.

AMS Subject Classification: 41A46

Key Words: Yun's activation function of algebraic type, Heaviside function, Hausdorff distance, Upper and lower bounds

Received: December 22, 2018 ; **Revised:** April 1, 2019 ;
Published (online): May 13, 2019 **doi:** 10.12732/dsa.v28i3.1
Dynamic Publishers, Inc., Acad. Publishers, Ltd. https://acadsol.eu/dsa

1. INTRODUCTION

Sigmoidal functions (also known as "activation functions") find multiple applications to population dynamics, biostatistics, analysis of nutrient supply for cell growth in bioreactors, controllability of tumor growth, classical predator–pray models, artificial neural networks, nucleation theory, machine learning, antenna–feeder technique, debugging theory, computer viruses propagation theory and others [1]–[12], [16]–[32].

In this paper we study the asymptotic behavior of the Hausdorff distance between Heaviside function $\sigma_H(t)$ and the Yun's parametric activation function of algebraic type [12].

We also consider a new activation function of trigonometric type.

The basic approaches for approximation of functions and point sets of the plane by algebraic and trigonometric polynomials in respect of Hausdorff distance (H–distance) are connected to the work and achievements of Bl. Sendov who established a Bulgarian school in Approximation theory, particularly developing the theory of Hausdorff approximations.

2. PRELIMINARIES

Definition 1. The Heaviside function is defined by

$$\sigma_H(t) = \begin{cases} 0, & \text{if} \quad t < 0, \\ 1, & \text{if} \quad t > 0. \end{cases} \tag{1}$$

Definition 2. [13], [14] The Hausdorff distance (the H–distance) $\rho(f, g)$ between two interval functions f, g on $\Omega \subseteq \mathbb{R}$, is the distance between their completed graphs $F(f)$ and $F(g)$ considered as closed subsets of $\Omega \times \mathbb{R}$.

More precisely,

$$\rho(f, g) = \max\{ \sup_{A \in F(f)} \inf_{B \in F(g)} ||A - B||, \sup_{B \in F(g)} \inf_{A \in F(f)} ||A - B||\}, \tag{2}$$

wherein $||.||$ is any norm in $\mathbb{R}^2$, e. g. the maximum norm $||(t, x)|| = \max\{|t|, |x|\}$; hence the distance between the points $A = (t_A, x_A)$, $B = (t_B, x_B)$ in $\mathbb{R}^2$ is $||A - B|| = \max(|t_A - t_B|, |x_A - x_B|)$.

Definition 3. The Yun's parametric activation function of algebraic type is defined by [12]

$$\sigma^{[m]}(t) = \begin{cases} 0, & \text{if} \quad t < -L, \\ \dfrac{(L+t)^m}{(L+t)^m + (L-t)^m}, & \text{if} \quad |t| \leq L, \\ 1, & \text{if} \quad t > L, \end{cases} \tag{3}$$

for a fixed $L > 0$.

In [12], the author proves the following properties of the new activation function: (A1) $\sigma^{[m]}$ is strictly increasing over $[-L, L]$ and $\sigma^{[m]} \in C^\infty(-L, L) \cap C^{m-1}\mathbf{R})$ for an integer $m \leq 1$.

The Hausdorff distance d between the $\sigma_H(t)$ and the function $\sigma^{[m]}(t)$ satisfies

$$\left(\frac{L+d}{L-d}\right)^m = \frac{1}{d} - 1, \quad 0 \leq d \leq \min\{\frac{1}{2}, L\}.$$

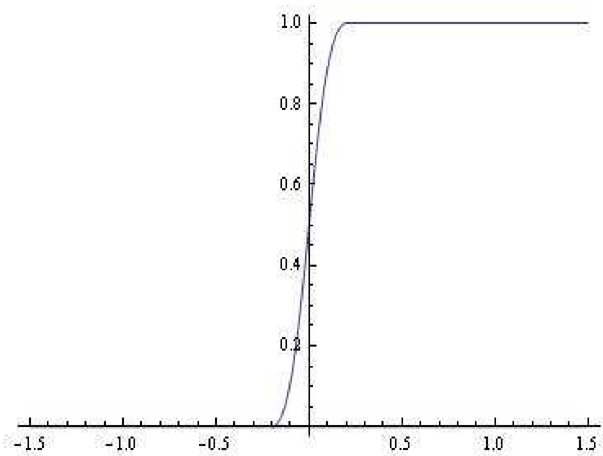

Figure 1: The Yun's activation function for $m = 1.6$ and $L = 0.2$.

That is, $m = O\left(\dfrac{\ln(1/d)}{\ln(1+d)}\right)$ for d small enough.

(A2) For m large enough $\sigma^{[m]}$ has the asymptotic behavior

$$\sigma^{[m]}(t) = \begin{cases} O\left(\theta(t)^m\right), & \text{if} \quad -L \le t < 0, \\ 1 + O\left(\theta(t)^m\right), & \text{if} \quad 0 < t \le L, \end{cases}$$

where

$$\theta(t) = \left(\frac{L-t}{L+t}\right)^{sgn(t)},$$

satisfying $0 \le \theta(t) < 1$ for all $t \in [-L, L]\backslash\{0\}$.

Based on the new activation function, the author consider a constructive feed–forward neural network approximation on a closed interval.

3. MAIN RESULTS

The Yun's considerations (see (A1)–(A2)) can be precised. Without loss of generality, we consider the case $L = 1$. Let $m = [\frac{n}{2}]$.

It is interesting to study the asymptotic behavior of $d = d(n)$ when n tends to infinity.

Theorem 1. For $d = d(n)$ the following is valid

$$d = \frac{\ln n}{n} + O\left(\frac{1}{n}\right). \tag{4}$$

Proof. Let us examine the relation:

$$\left(\frac{1+d}{1-d}\right)^m + 1 = \frac{1}{d}. \tag{5}$$

Following the ideas given in [14], [15] we have

$$\left(\frac{1+t}{1-t}\right)^m - e^{2mt} = \frac{2}{3}mt^3 + \frac{4}{3}m^2t^4 + ct^5 + \cdots$$

and

$$\left(\frac{1+t}{1-t}\right)^m = e^{2mt} + O(mt^3) \tag{6}$$

for small t.

Using (6), from (5) we obtain $e^{2md} + 1 + O(md^3) = \frac{1}{d}$, that is for $m = [\frac{n}{2}]$

$$e^{nd} + 1 + O(nd^3) = \frac{1}{d}. \tag{7}$$

From (7) we are lead to write

$$d = \frac{\ln n}{n} + \left(\frac{\theta_1(n)}{n}\right), \tag{8}$$

where $\theta_1(n)$ is to be estimated.

Insertion of d into (7) gives

$$e^{n\left(\frac{\ln n}{n} + \frac{\theta_1(n)}{n}\right)} + 1 + O\left(n\left(\frac{\ln n}{n} + \frac{\theta_1(n)}{n}\right)^3\right) = \frac{n}{\ln n + \theta_1(n)}$$

or

$$ne^{\theta_1(n)} + 1 + O\left(\frac{1}{n^2}\ln^3(n)\right) = \frac{n}{\ln n + \theta_1(n)}$$

and hence the function $\theta_1(n)$ is bounded.

In this way we obtain from (8) that

$$d = \frac{\ln n}{n} + O\left(\frac{1}{n}\right).$$

This completes the proof of the theorem.

Remark. In a number of cases, researchers working in the following areas: population dynamics, neural networks, biostatistics and others, for the values of the best Hausdorff approximations reliable interval estimates are required.

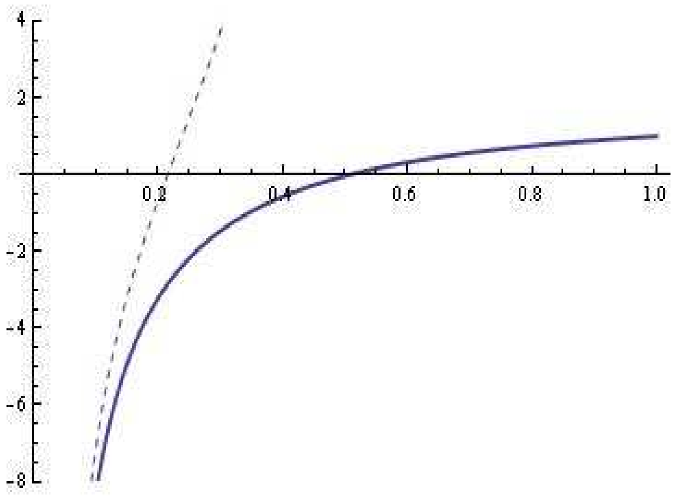

Figure 2: The functions $F(d)$ and $G(d)$ for $m = 3$; $L = 1$.

In this connection, let us examine the functions:

$$F(d) = \left(\frac{1+d}{1-d}\right)^m + 1 - \frac{1}{d},$$

$$G(d) = e^{2md} + 1 - \frac{1}{d}.$$

The following upper and lower bounds for d are valid

$$d_l = \frac{1}{2m} < d < \frac{\ln 2m}{2m} = d_r.$$

Evidently $G(d)$ approximates $F(d)$ with $d \to 0$ as $O(d^2)$ (see Fig. 2) and for $m > 2$ we have

$$G(d_l) < 0; \quad G(d_r) > 0.$$

Approximations of the $\sigma_H(t)$ by $\sigma^{[m]}(t)$ for various m are visualized on Fig. 3–Fig. 5.

4. A NEW ACTIVATION FUNCTION OF TRIGONOMETRIC TYPE

Definition 4. We consider the following activation function of trigonometric type

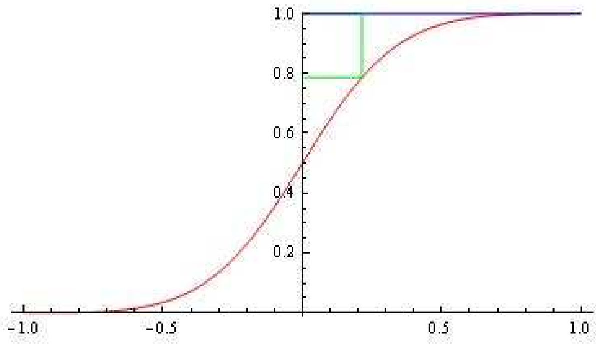

Figure 3: The case $m = 3$; $L = 1$; Hausdorff distance $d = 0.21374$; $d_l = 0.166667$; $d_r = 0.298027$.

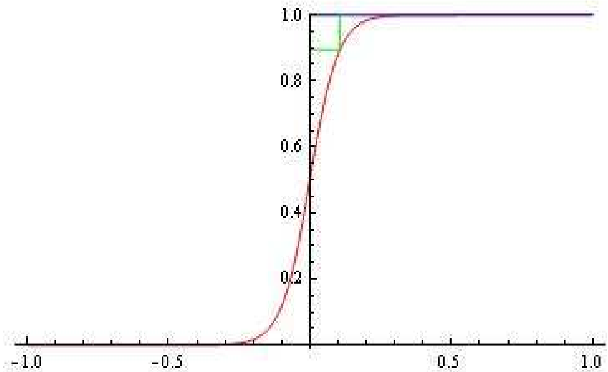

Figure 4: The case $m = 10$; $L = 1$; Hausdorff distance $d = 0.106139$; $d_l = 0.05$; $d_r = 0.149787$.

$$\sigma_T^{[m]}(t) = \begin{cases} 0, & \text{if} \quad t < -\frac{\pi}{m}, \\ \dfrac{(1+\tan(t))^m}{(1+\tan(t))^m + (1-\tan(t))^m}, & \text{if} \quad |t| \leq \frac{\pi}{m}, \\ 1, & \text{if} \quad t > \frac{\pi}{m}. \end{cases} \tag{9}$$

The Hausdorff distance d between the $\sigma_H(t)$ and $\sigma_T^{[m]}(t)$ satisfies

$$\left(\frac{1+\tan(d)}{1-\tan(d)}\right)^m + 1 = \frac{1}{d}.$$

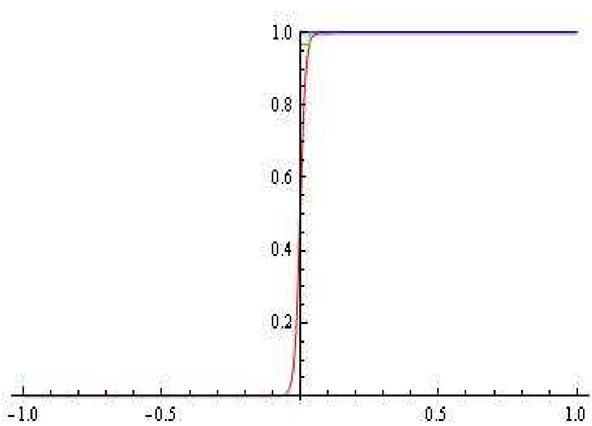

Figure 5: The case $m = 50$; $L = 1$; Hausdorff distance $d = 0.0335831$; $d_l = 0.01$; $d_r = 0.0460517$.

Then

$$\left(\frac{1+\tan(d)}{1-\tan(d)}\right)^m - e^{2mt} = \frac{4}{3}mt^3 + c_1 t^4 + \cdots$$

and

$$\left(\frac{1+\tan(d)}{1-\tan(d)}\right)^m = e^{2mt} + O(mt^3)$$

for small t.

The asymptotic behavior of the function $\sigma_T^{[m]}(t)$ can be studied in the manner outlined in Theorem 1.

Approximation of the $\sigma_H(t)$ by $\sigma_T^{[m]}(t)$ for $m = 15$ is visualized on Fig. 6.

In this case the following bounds for d are valid:

$$0.0333333 \le d = 0.0807295 \le 0.11373.$$

Remark. The reader may also consider the following function

$$E(t) = \frac{e^{(1+t)^m}}{e^{(1+t)^m} + e^{(1-t)^m}}$$

for $|t| \le \frac{2}{m}$.

Evidently, the Hausdorff distance d between the $\sigma_H(t)$ and $E(T)$ satisfies

$$\frac{e^{(1+d)^m}}{e^{(1+d)^m} + e^{(1-d)^m}} = 1 - d$$

or

$$e^{(1+d)^m - (1-d)^m} = \frac{1}{d} - 1.$$

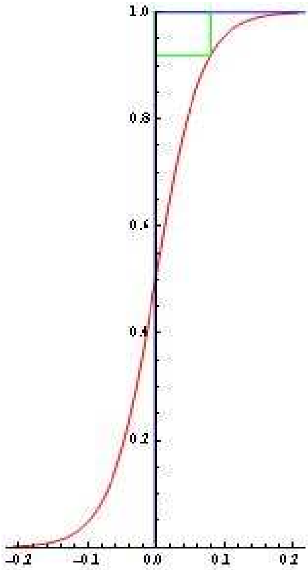

Figure 6: The case $m = 15$; H-distance between $\sigma_H(t)$ and $\sigma_T^{[m]}(t)$: $d = 0.0807295$; $d_l = 0.0333333$; $d_r = 0.11373$.

Then

$$e^{(1+d)^m-(1-d)^m} - e^{2md} = \frac{1}{3}m(2-3m+m^2)d^3 + O(md^4)$$

for small d.

For example, let $m = 15$. Then for the Hausdorff distance d between the $\sigma_H(t)$ and $E(T)$ we have $d = 0.0727489$ and the model has good "saturation to horizontal asymptote".

Constructive results of approximation by superposition of sigmoidal functions can be obtained using the methodology given, for example, in the articles [6]–[12].

Of course, we will explicitly note that the estimates are in line with the basic approaches for the approximation of functions by algebraic and trigonometric polynomials with respect to H-distance [14].

5. CONCLUSION

In this paper we study the asymptotic behavior of the Hausdorff distance between Heaviside function $\sigma_H(t)$ and the Yun's parametric activation function of algebraic type [12].

A new activation function of trigonometric type is also analyzed.

We hope that the results will be useful for specialists working in this scientific area.

6. ACKNOWLEDGMENTS

This paper is supported by the National Scientific Program "Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICTinSES)", financed by the Ministry of Education and Science.

REFERENCES

[1] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Math. Control Signal.*, **2** (1989), 303–314.

[2] L. N. Trefethen, *Approximation Theory and Approximation Practice*, SIAM, Philadelphia, Pa, USA (2013).

[3] Z. Chen, F. Cao, The approximation operators with sigmoidal functions, *Comput. Math. Appl.*, **58** (2009), 758–765.

[4] D. Elliott, Sigmoidal transformations and the trapezoidal rule, *The Journal of the Australian Mathematical Society Series B: Applied Mathematics*, **40** (1998/99), pp. E77–E137.

[5] N. Guliyev, V. Ismailov, A single hidden layer feedforward network with only one neuron in the hidden layer san approximate any univariate function, *Neural Computation*, **28** (2016), 1289–1304.

[6] D. Costarelli, R. Spigler, Approximation results for neural network operators activated by sigmoidal functions, *Neural Networks*, **44** (2013), 101–106.

[7] D. Costarelli, G. Vinti, Pointwise and uniform approximation by multivariate neural network operators of the max-product type, *Neural Networks*, **81** (2016), 81–90.

[8] D. Costarelli, R. Spigler, Solving numerically nonlinear systems of balance laws by multivariate sigmoidal functions approximation, *Computational and Applied Mathematics*, **37**, No. 1 (2018), 99–133.

[9] D. Costarelli, G. Vinti, Convergence for a family of neural network operators in Orlicz spaces, *Mathematische Nachrichten*, **290**, No. 2-3 (2017), 226–235.

[10] D. Costarelli, R. Spigler, Constructive Approximation by Superposition of Sigmoidal Functions, *Anal. Theory Appl.*, **29** (2013), 169–196.

[11] B. I. Yun, A smoothening method for the piecewise linear interpolation, *J. Appl. Math.*, (2015), 376362.

[12] B. I. Yun, A Neural Network Approximation Based on a Parametric Sigmoidal Function, *Mathematics*, **7** (2019), 262.

[13] F. Hausdorff, *Set Theory*, 2nd ed., Chelsea Publ., New York (1962).

[14] B. Sendov, *Hausdorff Approximations*, Kluwer, Boston (1990).

[15] Sendov, Bl., V. Popov, The exact asymptotic behavior of the best approximation by algebraic and trigonometric polynomials in the Hausdorff metric, *Math. USSR-Sb.*, Moscow, **89**, No. 131 (1972), 138–147. (in Russian)

[16] N. Kyurkchiev, S. Markov, On the Hausdorff distance between the Heaviside step function and Verhulst logistic function, *J. Math. Chem.*, **54**, No. 1 (2016), 109–119.

[17] N. Kyurkchiev, A Note on the Volmer's Activation (VA) Function, *C. R. Acad. Bulg. Sci.*, **70**, No. 6 (2017), 769–776.

[18] N. Kyurkchiev, S. Markov, *Sigmoid functions: Some Approximation and Modelling Aspects*, LAP LAMBERT Academic Publishing, Saarbrucken (2015), ISBN 978-3-659-76045-7.

[19] N. Kyurkchiev, A. Iliev, S. Markov, *Some techniques for recurrence generating of activation functions*, LAP LAMBERT Academic Publishing, (2017), ISBN: 978-3-330-33143-3.

[20] N. Kyurkchiev, A. Iliev, A. Rahnev, *Some Families of Sigmoid Functions: Applications to Growth Theory*, LAP LAMBERT Academic Publishing, (2019), ISBN: 978-613-9-45608-6.

[21] N. Kyurkchiev, A. Iliev, A. Rahnev, A new class of activation functions based on the correcting amendments of Gompertz-Makeham type, *Dynamic Systems and Applications*, **28**, No. 2 (2019), 243–257.

[22] N. Kyurkchiev, A. Andreev, *Approximation and antenna and filter synthesis: Some moduli in programming environment Mathematica*, LAP LAMBERT Academic Publishing, Saarbrucken (2014), ISBN 978-3-659-53322-8.

[23] N. Pavlov, A. Iliev, A. Rahnev, N. Kyurkchiev, *Some software reliability models: Approximation and modeling aspects*, LAP LAMBERT Academic Publishing, (2018), ISBN: 978-613-9-82805-0.

[24] N. Pavlov, A. Iliev, A. Rahnev, N. Kyurkchiev, *Nontrivial Models in Debugging Theory (Part 2)*, LAP LAMBERT Academic Publishing, (2018), ISBN: 978-613-9-87794-2.

[25] A. Iliev, N. Kyurkchiev, A. Rahnev, T. Terzieva, *Some models in the theory of computer viruses propagation*, LAP LAMBERT Academic Publishing, (2019). (to appear)

[26] S. Markov, A. Iliev, A. Rahnev, N. Kyurkchiev, A note on the Log-logistic and transmuted Log-logistic models. Some applications, *Dynamic Systems and Applications*, **27**, No. 3 (2018), 593–607.

[27] S. Markov, N. Kyurkchiev, A. Iliev, A. Rahnev, On the approximation of the generalized cut functions of degree p+1 by smooth hyper-log-logistic function, *Dynamic Systems and Applications*, **27**, No. 4 (2018), 715–728.

[28] M. Pavlov, A. Iliev, A. Rahnev, N. Kyurkchiev, On Some Nonstandard Software Reliability Models, *Dynamic Systems and Applications*, **27**, No. 4 (2018), 757–771.

[29] S. Markov, A. Iliev, A. Rahnev, N. Kyurkchiev, A Note On the Three-stage Growth Model, *Dynamic Systems and Applications*, **28**, No. 1 (2019), 63–72.

[30] N. Kyurkchiev, A. Iliev, A. Rahnev, A new class of activation functions based on the correcting amendments of Gompertz-Makeham type, *Dynamic Systems and Applications*, **28**, No. 2 (2019), 243–257.

[31] A. Antonov, S. Nenov, Ts. Tsvetkov, Impulsive controllability of tumor growth, *Dynamic Systems and Applications*, **28**, No. 1, (2019), 93–109.

[32] S. Nenov, A. Antonov, Ts. Tsvetkov, Impulsive models: Maximum yield of some biological systems, I, *Dynamic Systems and Applications*, **28**, No. 2 (2019), 317–328.