Approximation of Hyperbolic Tangent Activation Function Using Hybrid Methods

Maicon A. Sartin - Universidade do Estado de Mato Grosso (UNEMAT), Brazil.
Alexandre C. R. da Silva - Universidade Estadual Paulista (UNESP), Brazil.
mapsartin@unemat.br, acrsilva@dee.feis.unesp.br

Abstract

The contribution of this work is the approximation of a nonlinear function used in ANN, the popular hyperbolic tangent (HT) activation function (AF). The system architecture is composed of several scenarios that provide a tradeoff of performance, precision and area used in FPGA. The results are compared in different scenarios and with current literature on error analysis, area and system performance.

Basic Concepts

The HT and sigmoid functions both produce a curve with an “S” shape. HT contains values (±) and the output is between [-1,1]. The behavior of the AF is nonlinear and has a steeper curve before saturation of the two ends. Four features can be observed on the HT: (i) there is symmetry between the y axis; (ii) For very large values of x, $Th(x) \rightarrow \pm 1$; (iii) the central part is near a linear equation and the variation is greater; (iv) The output (Th (x)) is equal to its input values for x near zero.

$$th(x) = \frac{1 - e^{-\alpha x}}{1 + e^{-\alpha x}}$$

Table 1 - Analyzed Scenarios.

N.	Scenarios
1	HPR of 180 Elements
2	HPR of 212 Elements
3	HPR of 314 Elements
4	HPR of 340 Elements
5	HPR of 404 Elements
6	HPC of 180 Elements
7	HPC of 212 Elements
8	HPC of 340 Elements
9	HPC of 404 Elements

Analysis of the Results

Table 2 - Errors Absolute and Relative Found

Approximations	Error_Max	Error_Ave	Rel_Ave
Sc. 1 (180)	4.18×10^{-3}	1.05×10^{-3}	1.24×10^{-3}
Sc. 2 (212)	4.18×10^{-3}	1.03×10^{-3}	1.20×10^{-3}
Sc. 3 (314)	3.30×10^{-3}	6.80×10^{-4}	8.43×10^{-4}
Sc. 4 (340)	3.30×10^{-3}	7.18×10^{-4}	8.99×10^{-4}
Sc. 5 (404)	3.23×10^{-3}	6.80×10^{-4}	8.69×10^{-4}
Sc. 6 (180)	5.94×10^{-3}	2.75×10^{-3}	3.04×10^{-3}
Sc. 7 (212)	5.94×10^{-3}	2.75×10^{-3}	3.04×10^{-3}
Sc. 8 (340)	5.94×10^{-3}	2.36×10^{-3}	2.72×10^{-3}
Sc. 9 (404)	5.94×10^{-3}	2.28×10^{-3}	2.67×10^{-3}
Comb(Tommiska,03)	3.90×10^{-3}	1.70×10^{-3}	-
Hybrid(Meher,10)	2.00×10^{-2}	1.72×10^{-2}	-
Hybrid(Bajger;Omondi,08)	4.61×10^{-4}	1.35×10^{-4}	1.38×10^{-3}

Conclusion

This paper presented the implementation of a system for approximating a HT AF into a reconfigurable device. Scenarios 1 to 5 are good options in the precision HT AF to the FPGA. The scenarios 6 to 9 don’t need BRAM and the precision is relatively compromised with the previous scenarios, but has the best tradeoff.

Acknowledgements

The authors would like to thank the support of UNEMAT, UNESP and PPGEE, LPSSD, CAPES and CNPQ - process (309023/2012-2).

Proposed Methods for the Approximation of the HT

HPR (Hybrid with PWL and RALUT) and HPC (Hybrid with PWL and Combinational) methods use division into parts (PWL) with the goal of acquire high precision and an amount of elements adapted to the variation of interval. Combinational techniques and RALUT can reduce hardware resources with suitable. In the HPR hardware, the mapping is made from amount of addressing bits related with the scenario interval. To make mapping in the HPC hardware each interval contains a sum of products (SOP) to each output bit. SOP is made based in the amount of elements in interval and simplified with the Quine-McCluskey algorithm.

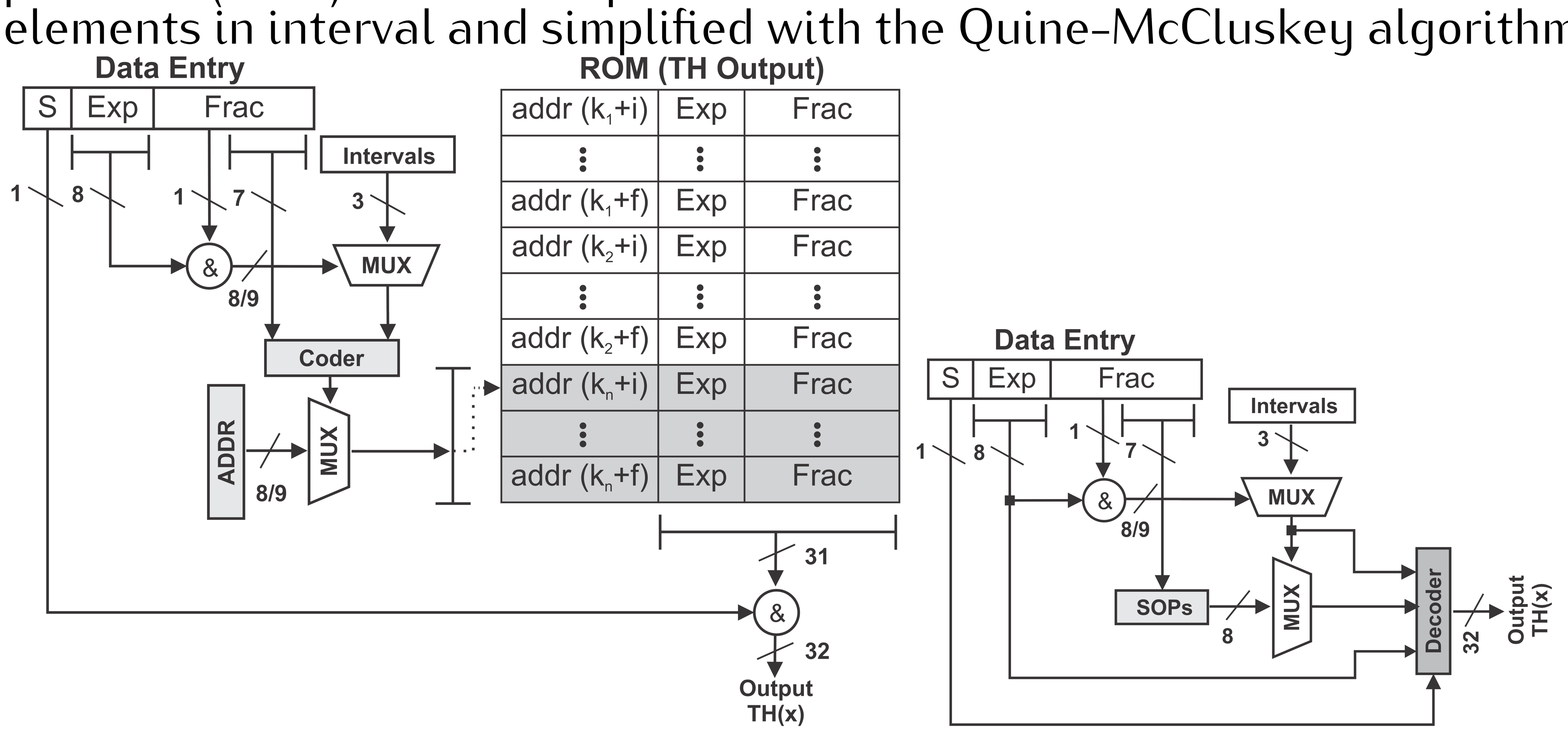

Figure 1 - Hardware for Mapping the Outputs of the AF (HPR and HPC)

Analysis of the Results

The system contains 500 inputs samples for the error analysis, area, and performance in implementing the HT AF in hardware. With all samples the implementation of the scenarios were run in approximately 40 µs for Spartan 3E platform, with clock period of 20 ns.

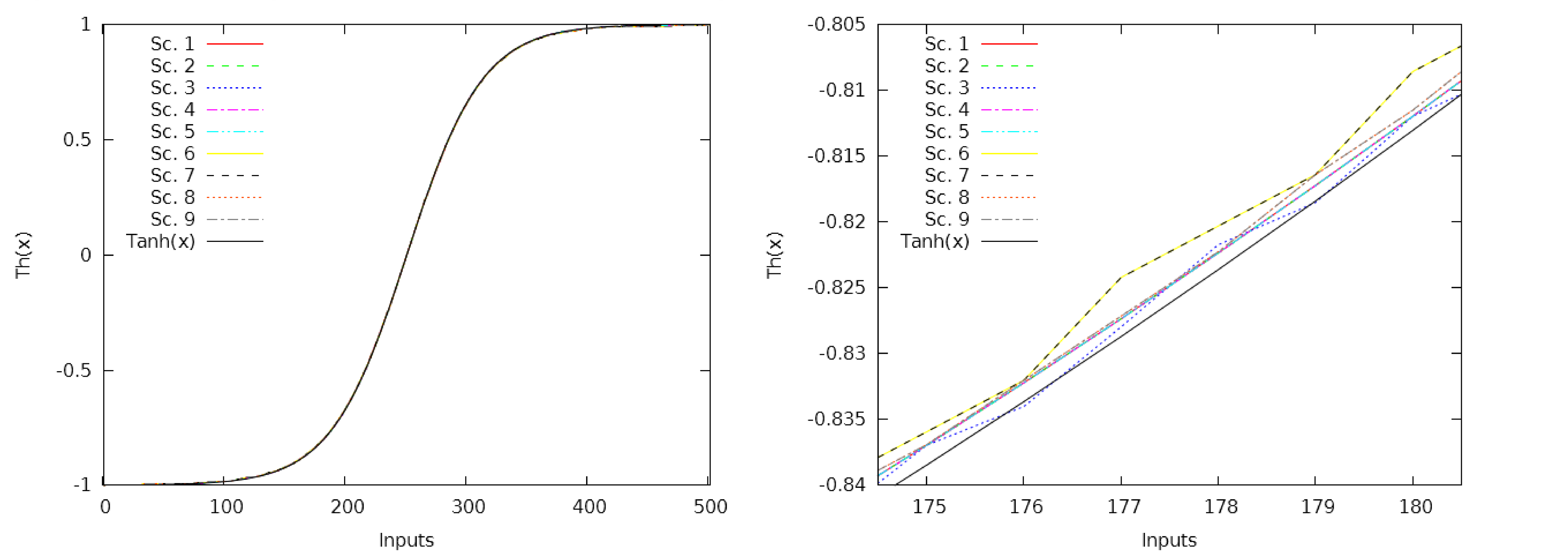

Figure 2 - Approximation of the HT AF with 500 and 6 points for all scenarios

The scenarios (1 to 5) (HPR) present the best accuracy, when compared with combinational methods (Scenarios 6 to 9) (HPC). But the HPR scenarios need of dedicated memory (BRAM), different of HPC scenarios that have just Boolean expressions. Analyzing the tradeoff between area and accuracy the scenario 9 is the best choose between all. In the results of the scenarios present a similarity in the area usage and accuracy. The results in the AF system present most values below 15 slices, just the scenarios 6 and 7 with approximately 135 slices, similar to recent literature.

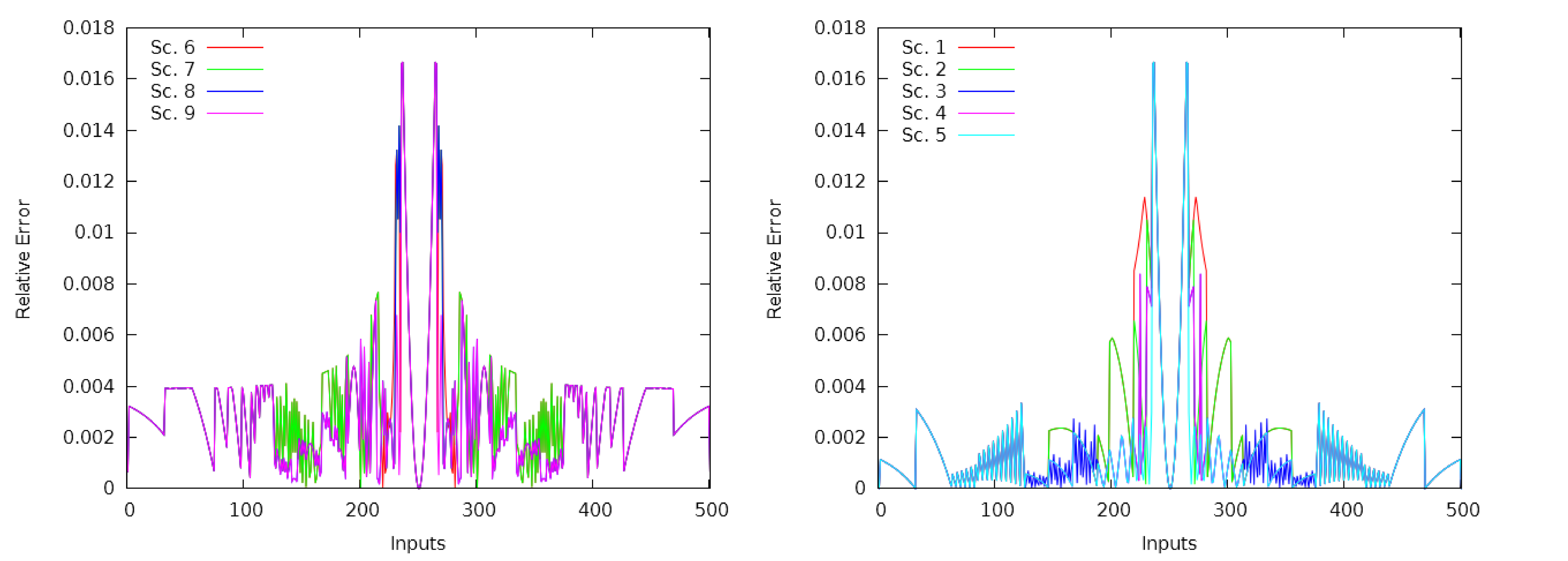

Figure 3 - Relative Error of the HPR and HPC Scenarios with 500 Points.