



**HAL**  
open science

# Deep Neural Network Feasibility Using Analog Spiking Neurons

Thomas Soupizet, Zalfa Jouni, Joao Frischenbruder Sulzbach, A.  
Benlarbi-Delai, Pietro Maris Ferreira

► **To cite this version:**

Thomas Soupizet, Zalfa Jouni, Joao Frischenbruder Sulzbach, A. Benlarbi-Delai, Pietro Maris Ferreira. Deep Neural Network Feasibility Using Analog Spiking Neurons. 35th SBC/SBMicro/IEEE/ACM Symposium on Integrated Circuits and Systems Design (SBCCI), Aug 2022, Porto Alegre, Brazil. 10.1109/SBCCI55532.2022.9893216 . hal-03689837

**HAL Id: hal-03689837**

**<https://hal-centralesupelec.archives-ouvertes.fr/hal-03689837>**

Submitted on 7 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Deep Neural Network Feasibility Using Analog Spiking Neurons

Thomas Soupizet, Zalfa Jouni, Joao Frischenbruder Sulzbach, Aziz Benlarbi-Delai, Pietro M. Ferreira,  
 Université Paris-Saclay, CentraleSupélec, CNRS, Lab. de Génie Électrique et Électronique de Paris, 91192, Gif-sur-Yvette, France.  
 Sorbonne Université, CNRS, Lab. de Génie Électrique et Électronique de Paris, 75252, Paris, France.  
 Email: tsoupizet@gmail.com, maris@ieee.org

**Abstract**—Novel non-Von-Neumann solutions have raised based on artificial intelligence (AI) such as the neuromorphic spiking processors in either analog or digital domain. This paper proposes to study the deep neural network feasibility using ultra-low-power eNeuron. The trade-offs in terms of deep learning capabilities and energy efficiency are highlighted. A linear fit model is found in the region of high energy efficiency of neuromorphic components. Thus, deep learning and energy efficiency mutually exclusive if those neuromorphic components are used.

**Index Terms**—neuromorphic circuits, spiking neural network, deep learning, low power

## I. INTRODUCTION

The advent of industry 4.0 and the Internet of Things have pushed the requirements for smart processing capabilities in electronics. Along this advance, the smartness is also in a high energy efficient devices, which challenges common solutions. Since the 90's, neural networks and artificial intelligence solutions answered such need of high processing capability, and are thus widely implemented on classic computing [1]. Those implementations, however, tend to be power-hungry using Von-Neumann architectures and cloud computing. While reduced pace in Moore's law has further disclosed architecture limitations, demanded alternative solutions [2].

In this context, novel non-Von-Neumann solutions have raised based on artificial intelligence (AI). One of the most researched nowadays is the neuromorphic spiking processors in either analog or digital domain [3]. In contrast of Von-Neumann architecture, neuromorphic computing has appeared as an exciting alternative which provides edge computing capabilities for ultra-low-power applications [4]. Moreover, software and hardware solutions have brought the computational complexity of neural networks on energy efficient system with deep learning capabilities [5]. Both software and hardware solutions have pointed to the most widely used solution, which is the Feed-Forward Neural Networks (FNNs) [4]. Software implementations of FNNs are often used along with deep learning algorithms in order to solve highly complex problems. However, recent neuromorphic hardware is often single neuron circuits [6], [7], [8] or small neural networks [9], [10]. It is thus natural to ask if FNNs algorithms are usable on analog spiking neurons (eNeuron), to try and narrow the gap between hardware and software AI.

This paper objective is to study the deep neural network feasibility using ultra-low-power eNeuron. Previous published eNeurons from [9], [10] have been redesigned using in-house BiCMOS 55 nm technology from ST Microelectronics. Literature synapses are revised to propose a simple, linear, and weight reconfigurable one. Figure 1 illustrates the eNeuron and synapse model build up in this work from post-layout simulations. Further study of mathematical FNNs reveals set necessary conditions for deep learning using spiking neural networks. To the best of the author's knowledge, the trade-offs highlighted in terms of deep learning capabilities and energy efficiency are first proposed in this paper.

This paper is organized as follows. Section II revises literature in both hardware and software point of view. Finally, conclusions are drawn towards deep learning neural networks using ultra-low-power eNeurons.

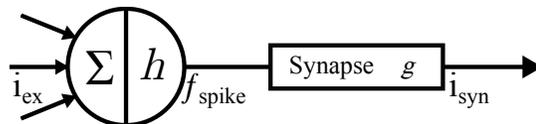


Fig. 1: Model spiking eNeuron and synapse. A neuron convert a current inputs to a spiking frequency. Synapse convert a spiking input to an output current.

## II. ANALOG SPIKING NEURAL NETWORKS

Neuromorphic hardware aims to faithfully mimic biological systems in artificial neurons and synapses, illustrated in Fig. 1. Those devices have presented the best energy consumption per unit of information (i.e.  $E_{eff}$  in J/spike) and a competitive area trade-off. In the other hand, FNNs have been dominant in software AI research aiming deep learning capabilities and interested in mathematical network properties. Following subsections revise both hardware and software point of view.

### A. Neurons

To design a neural network, one should first select a mathematical model for the neurons. The eNeuron model sets the degrees of complexity and biological accuracy as explained in [4]. Many popular models have seen widespread use as they offer different trade-offs among energy efficiency and computational capabilities. The simplicity of McCulloch-Pitts

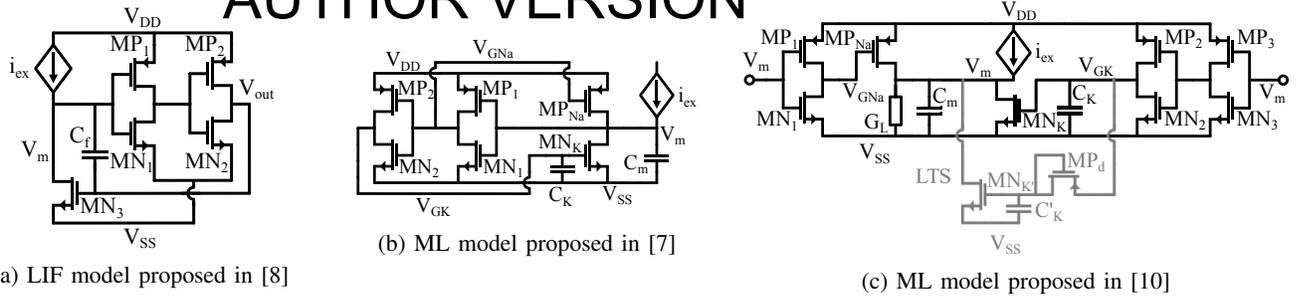


Fig. 2: eNeuron circuits studied. By using low supply-voltages, low membrane capacitance, and subthreshold transistors, authors have minimised power-consumption and surface area.

eNeurons makes them dominant in software implementations, as they only use an activation function to associate an output to the sum of their inputs.

More biologically inspired models mimic the spiking behavior of biological neurons. Among them, the (Leaky) Integrate-and-Fire (LIF), and Morris Lecar (ML) models stand out due to their widespread adoption in the literature. LIF eNeuron offers a spiking behavior at a low complexity by modeling the charge of a neuron. Figure 2(a) illustrates a LIF eNeuron proposed by Danneville et al., which is an Axon-Hillock topology. By using parasitic capacitors as the membrane capacitance ( $C_m$ ), Danneville’s eNeuron achieved both low energy consumption and low surface area, which could lead to the design being used in SNNs [8]. The LIF eNeuron has a mathematical model described in the literature as

$$V_m = \frac{1}{C_m} \int i_{ex} \left( e^{\frac{V_{DD}-V_m}{\eta\phi_t}} - e^{\frac{-2V_m}{\eta\phi_t}} \right) dt, \quad (1)$$

where  $V_m$  is the spiking voltage signal;  $i_{ex}$  is the input excitation current;  $V_{DD}$  is the eNeuron supply voltage;  $\eta$  is the subthreshold slope;  $\phi_t$  is the thermal voltage ( $k_bT/q$ ). A detailed modeling is presented in [11].

The ML model provides behavior close to biological neurons by using multiple complex non-linear differential equations. Such eNeuron has recently been used to implement artificial neural networks, achieving tasks such as edge recognition [9] and audio signal processing [10]. Two common concerns in such edge computing applications are circuit surface and energy consumption. The popular figure of merit  $E_{eff}$  is used to assess the later. The  $E_{eff}$  is depicted considering the spike as a unit of information. State-of-the-art eNeurons [9], [10] are most efficient when they function at high spiking frequency, and they become increasingly  $E_{eff}$ -inefficient as spiking frequency lowers. ML eNeuron has a mathematical model described in the literature, being a non-linear function

$$V_m = -V_{DD} \cdot \tanh \left( \frac{G_L \cdot \sum i_{ex}}{\eta\phi_t} + \frac{1}{2} \ln \left( \frac{G_N}{G_P} \right) \right), \quad (2)$$

where  $G_L$  is the equivalent conductance between eNeuron membrane and ground;  $G_N/G_P$  are the transistors’ aspect ratio (NMOS conductance over PMOS conductance). A detailed modeling is presented in [7].

To address large-scale applications where energy dissipation is critical, Sourikopoulos et al. have proposed a simplified ML eNeuron illustrated in Fig. 2(b). By minimizing supply voltage and membrane capacitance, the authors achieved both low power consumption and high-spiking frequency to a high  $E_{eff}$  [7]. As a step towards SNNs, Ferreira et al. have implemented a Neuromorphic Analog Spiking-Modulator, which translated a sample input into a spiking output. The output layer is composed of two neurons with different spiking behavior, named ‘LTS’ and ‘FS’ eNeuron. Figure 2(c) illustrates both eNeuron, which are differentiated by the additional branch in blue, only present in the ‘LTS’ one. Recent ML eNeuron implementations notably make use of subthreshold transistors to reduce surface and power consumption [7], [10].

Literature results highlight the widely different performances in different areas and  $f_{spike}$ . For this reason, eNeurons depicted in Fig. 2 shall be implemented, in this work, with in-house tools. Thus, this work may propose a mathematical model from post-layout simulations. Moreover, the use of the same technology and simulator for eNeurons will prevent the result from being skewed by technology improvements, which will make the comparison between eNeuron models fairer.

### B. Synapses

In spiking neural networks, neuron input (post-synaptic signals) and output (pre-synaptic signals) do not have the same dimensions. It is thus necessary to link neurons using synapses, which convert spiking information in  $i_{syn}$  post-synaptic signal.

Bartolozzi and Indiveri have proposed a synapse design adequate for VLSI implementation. By using transistors operating in subthreshold, the proposed differential pair integrator synapse response is

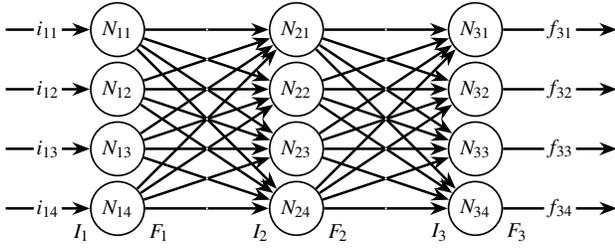
$$I_{syn}(t) = I_0 \cdot e^{\frac{V_{syn}(t)-V_{DD}}{\eta\phi_t}}, \quad (3)$$

where  $I_0$  is the leakage current;  $V_{syn}$  is the pre-synaptic signal;  $I_{syn}$  is the post-synaptic signal. This model fits closer to the Destexhe mathematical model for synapses [12]. It is remarkable that the mean current output of the synapse is a linear function of the pre-synaptic  $f_{spike}$  [13].

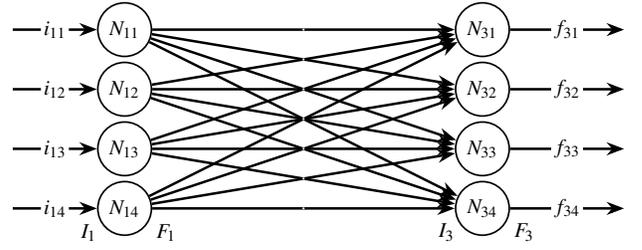
In order to build an SNN, Danneville et al. have proposed two new synapses in [9]. The first proposition uses inverters along with a transimpedance amplifier and a transistor biasing



# AUTHOR VERSION



(a) An example software neural network.



(b) An equivalent two layer neural network.

Fig. 4: Feed-Forward Neural Networks. In the case of linear activation functions, a network with no hidden layer such as (b) can be found with the same transfer function as any (a).

The output  $F_N$  of the FNN is given by a linear combination of the input  $I_0$ . It is possible to find an equivalent two layers neural network, as represented in Fig. 4(b), where those two FNNs are equivalent. Having a linear activation functions, it thus hinders FNN to have the properties of a deep neural network, even with multiple hidden layers. For deep learning purposes, a non-linear activation function is needed. In literature, the non-linear functions *ReLU* and *tanh* are often used as activation functions. The latter is often related to eNeuron models (1) and (2).

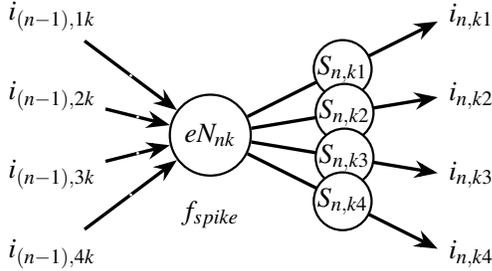


Fig. 5: Electronic Neural Network

## B. Analysis of electronic spiking neural networks

While (5) and FNN property are widely known in the case of mathematical neural networks [14], it is not the case for electronic spiking neural networks (eNN). Figure 5 illustrates a system model for eNeurons ( $eN_{nk}$ ) connected to a synapse array ( $S_{n,ki}$ ). Considering eNNs, the presence of a synapse linking the neurons modifies the proof in (5). An eNeuron takes as an input current  $i_{ex}$ , which is the sum of the current fed ( $\sum_{i=1}^4 i_{(n-1),ik}$ ) in the node  $V_m$ . It outputs a spiking voltage in  $V_m$  which  $i_{ex}$  amplitude is coded in the  $f_{spike}$ . A function  $h$  can thus be defined for the eNeuron such that  $f_{spike} = h(i_{ex})$ . The synapse does the opposite, it converts  $f_{spike}$  pre-synaptic information to a post-synaptic current  $i_{syn}$ . From  $i_{syn}$  and  $f_{spike}$  relationship, a function  $g$  can be defined such that  $i_{syn} = g(f_{spike})$ .

An eNN such as the one modeled in Fig. 4(a), now considering eNeuron models from Fig. 5, can be studied with some formalisms. The  $k^{th}$  eNeuron on the  $n^{th}$  layer will be referred to as  $eN_{nk}$ , with the transfer function  $h_{nk}$ , input  $i_{nk}$ , and spiking

frequency  $f_{nk}$ . The synapse linking together  $eN_{nk}$  and  $eN_{(n+1)l}$  will be referred to as  $S_{n,kl}$ , with the transfer function  $g_{n,kl}$ , input  $f_{n,kl}$ , output  $i_{n,kl}$ . Let  $M_n$  be the number of neurons in layer  $n$ . Assuming the transfer functions of the synapses and neurons are linear, i.e. for all  $n, k, l$  such that  $eN_{nk}, eN_{(n+1)l} \in eNN$ :

$$h_{nk}(i_{ex}) = k_{nk} \cdot i_{ex} + b_{nk} \quad (6)$$

$$g_{n,kl}(f_{spike}) = w_{n,kl} \cdot f_{spike} + c_{n,kl}. \quad (7)$$

The spiking frequency of  $eN_{(n+1)l}$  is then

$$\begin{aligned} f_{(n+1)l} &= h_{(n+1)l} \left( \sum_{k=1}^{M_n} i_{n,kl} \right) \\ f_{(n+1)l} &= h_{(n+1)l} \left( \sum_{k=1}^{M_n} g_{n,kl}(f_{nk}) \right) \\ f_{(n+1)l} &= k_{(n+1)l} \cdot \left( \sum_{k=1}^{M_n} w_{n,kl} \cdot f_{nk} + c_{n,kl} \right) + b_{(n+1)l}. \quad (8) \end{aligned}$$

This relation is linear, which means that if we express the spiking frequencies of a layer of eNeurons as a vector, we can find matrices such that

$$F_n = (f_{n1}, \dots, f_{nM_n})^T \quad (9)$$

$$F_n = A_n F_{n-1}. \quad (10)$$

By induction, one may obtain

$$\begin{aligned} I_1 &= (i_{11}, \dots, i_{1M_1})^T \\ F_N &= \prod_{n=1}^N A_n \cdot I_1. \quad (11) \end{aligned}$$

This proves that the previous result still applies in the case of an eNN. Provided neurons and synapses have linear transfer functions, any multi-layer feed-forward neural network is equivalent to a two layers neural network.

## IV. CIRCUIT DESIGN EXAMPLE

A  $i_{ex}$  parametric sweep is PLS simulated between 1 pA and 1 nA to extract the  $f_{spike}$  response of the eNeurons as a function of  $i_{ex}$ . At  $i_{ex} > 1nA$ , spiking frequency begins to fall due to circuit stability due to subthreshold biasing (i.e. spiking oscillations stops). Thus, the pointed  $i_{ex}$  range corresponds to the estimated dynamic of the eNeurons. This work will use the

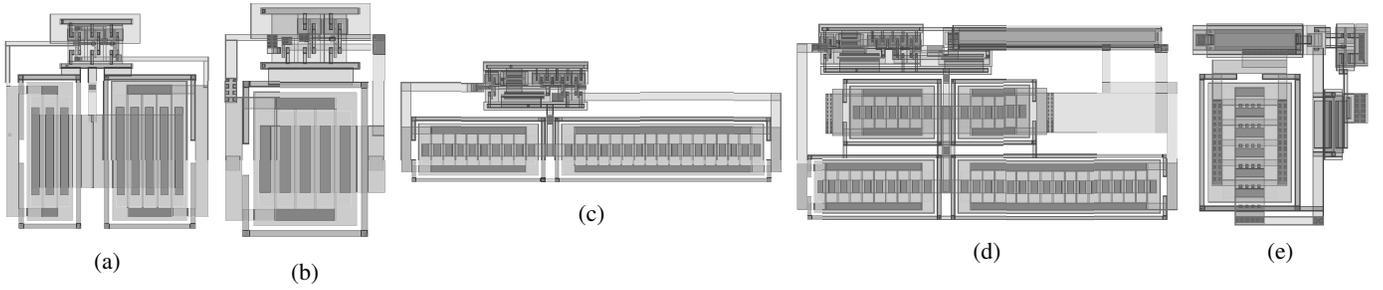


Fig. 6: In-house proposed layouts for eNeurons under test: (a) Simplified ML using  $8.210 \times 7.525 \mu m^2$  [7], (b) LIF using  $6.565 \times 4.335 \mu m^2$  [8], (c) "FS" biomimetic ML using  $5.695 \times 17.335 \mu m^2$  [10], (d) "LTS" biomimetic ML using  $9.110 \times 17.335 \mu m^2$  [10]. Studied synapse layout (e) using  $6.60 \times 7.55 \mu m^2$ .

least mean squares method to evaluate the transfer function linearity of aforementioned eNeurons. If an accurate linear fit is found for eNeurons under test, they should not be used in deep learning applications. Fitting accuracy is estimated from  $r^2$  method from MatLab tools. Same PLS provide the relation between energy efficiency and spiking frequency, similarly tests are carried out as proposed in [9], [10].

TABLE I: Sizing of  $MN$  and  $MP$  transistors in  $W \times L$  (nm) and in capacitance for  $C$  (fF) for the neurons

LIF model proposed in [8]			
$MP_1, MN_1, MP_2, MN_2, MN_3$	135 × 65		
$C_f$	5.038		
ML model proposed in [7]			
$MP_1, MN_1, MP_2, MN_2, MP_{Na}, MN_K$	135 × 65		
$C_m$	3.996	$C_K$	7.991
ML model proposed in [10]			
$MP_1$	135 × 60	$MN_1$	200 × 60
$MP_2$	1200 × 60	$MN_2$	135 × 60
$MP_3$	200 × 60	$MN_3$	135 × 60
$MP_{Na}$	800 × 60	$MN_K$	2000 × 60
$C_m$	5.534	$C_f$	9.838
$MP_d$	500 × 10,000	$MN_{K'}$	2000 × 60
$C_K'$	6.763	$G_L$	Leakage
Proposed synapse			
$MP_R$	500 × 3000	$MN_{Trans}$	135 × 480
$MP_1$	270 × 60	$MP_2$	324 × 60
$C$	11.9		

In-house design is proposed for eNeurons illustrated in Fig. 2. eNeuron transistor sizing is summarized in Tab. I. Implementing every neuron in BiCMOS 55 nm technology, eNeuron layouts are depicted in Fig. 6. Post-layout simulations (PLS) are carried out using Spectre Virtuoso from BSIM4 models, which are measured-based foundry models validated in [16].

A basic synapse, illustrated in Fig. 3, was designed to assess the linearity characteristic of the Destexhe mathematical model in (3). A diode-connected transistor is used to have a high load, which makes the RC filter operate at a frequency in the range of spiking frequencies. The current mirror allows for choosing a weight for the synapse, and makes the output current positive. The last diode connected transistor allows for controlling the common-mode offset. Transistor and capacitor sizing are obtained so that the input and output ranges of the

synapse match those of eNeurons. Table I summarizes the final sizing. Figure 6(e) illustrates the layout for the synapse under study. For PLS results, the synapse is used to connect two neurons, thus assuring the correct operation of the synapse.

### A. eNeuron PLS

Figure 7 shows that the relations between  $i_{ex}$  and  $F_{spike}$  are non-linear, with energy efficiency significantly increasing with increased input current. As energy efficient is an important concern when designing and using eNeurons, it is expected for the neurons to be used in operating areas of high energy efficiency. Thus, a linear fit was realized for those areas in particular, omitting the less efficient low-spiking frequencies. Having a smaller range of operation, the LIF eNeuron proposed in [7] is significantly more non-linear than the other eNeurons. No satisfying linear fit could be found for the transfer function in Fig. 7(a), for any significantly large  $i_{ex}$  range.

For the eNeuron proposed in [8] and the FS eNeuron proposed in [10], a satisfying linear fit was found ( $r^2 = 0.99$ ) for  $i_{ex} > 200 pA$ . While the transfer functions are still non-linear in this area, they can be closely approximated by a linear function. The difference between the linear fits and real transfer functions on average of 1.58% of spiking frequency for Fig. 7(b), with the fit  $f(i) = 0.10 \times i + 158.5$ , and 3.09% for the FS eNeuron in Fig. 7(c), with the fit  $f(i) = 0.28 \times i - 14.0$ . For the LTS eNeuron proposed in [10], a linear fit can be found for the whole area of operation, with  $r^2 = 0.99$  for  $i_{ex} > 100 pA$  with the fit  $f(i) = 0.30 \times i + 133.9$ .

Figure 7 also shows the energy efficiency of the neurons for a sweep of their input range. Figure 7(b) shows that the neuron becomes most energy efficient above 100 pA, with energy efficiency strongly degrading for small  $i_{ex}$ . Figure 7(c) show that the neurons become energy efficient at  $i_{ex} > 200 pA$ , with performance degrading for smaller input. The neurons are thus expected to be used in ranges above 100 pA and 200 pA respectively.

This results highlights a trade-off that has to be made between energy efficiency and deep-learning capabilities while using state-of-the-art eNeurons. If a designer chooses to use the full range of frequency of the eNeurons, he will face poor energy efficiency. If instead one chooses to limit the

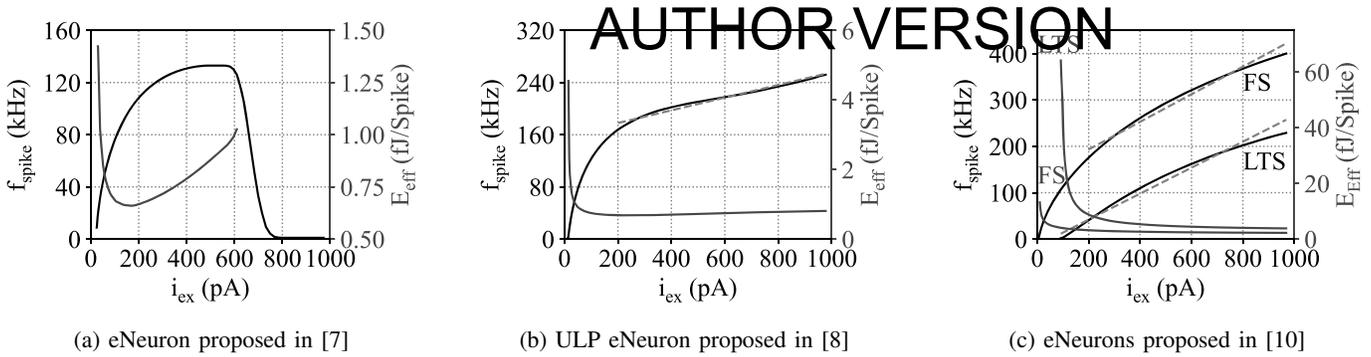


Fig. 7: Results of in-house PLS using BiCMOS 55nm technology. In black, spiking frequencies of eNeurons as a function of input current. In red, linearisation at high  $F_{Spike}$ . In blue, energy efficiency of the neurons as a function of the input current.

operating range to higher  $f_{spike}$ , the neurons will have a linear characteristic. In this second case, he will either be limited to the processing capabilities of a shallow neural network, or have to use a non-linear synapse. A designer could also use the Sourikopoulos eNeurons, which would provide nonlinearity at the cost of dynamic range.

### B. Synapse PLS

The synapse PLS results exhibit the necessary functions: output current grows with input spiking frequency, and frequency injection is prevented. The excitation current of the input neuron was swept from 0 to 1 nA, to generate a sweep of the input spiking frequency of the synapse.

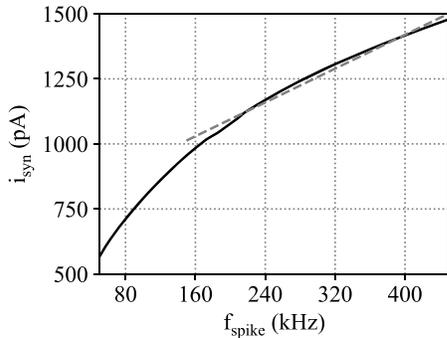


Fig. 8: Synapse transfer function in black, with linear fit in red.

Figure 8 shows that the output current of the synapse increases with the input frequency, which is consistent with the expected behavior for an excitation synapse. The frequency range explored is 0 to 450 kHz, as it is the output range of the FS Neuron used here [10]. The output is non-linear, however, as with the neurons, a linear fit can be found for a specific input range. For  $f_{spike In} > 150$  kHz, the linear fit  $i(f) = 1.63 \times f + 769.2$  was found with  $r^2 > 0.99$ , with an average difference between linearization and output of 1.09%.

As a linear fit can be found for the synapse in areas of high energy efficiency of the eNeurons, it will not introduce non-linearity for networks functioning in those areas of operation.

This means that the transfer function between the outputs of two layers will be a linear combination. This makes deep learning and energy efficiency mutually exclusive if those neuromorphic components are used.

## V. CONCLUSIONS

This paper studied the deep learning capabilities of eNeurons, exposing trade-offs for using FNN architecture. PLS results showed that linear fits were possible at  $i_{ex} > 200$  pA, with  $r^2 \geq 0.99$ , which makes deep networks useless in those areas of high energy efficiency. A linear fit could also be found with  $r^2 > 0.99$  for the synapse, showing that using such components at  $i_{ex} > 200$  pA would not introduce any non-linearities. To the best of the author's knowledge, the trade-offs revealed in this paper are the first to highlight the counter-intuitive choice of high power consumption that has to be made to allow for deep learning capabilities.

## REFERENCES

- [1] D. Orrey, D. Myers, and J. Vincent, "A high performance digital processor for implementing large artificial neural networks," in *Proceedings of the IEEE Custom Integrated Circuits Conference*. Orlando, FL, USA: IEEE, jul 1991, pp. 16.3/1–16.3/4.
- [2] *The Quantum Limit to Moore's Law*, vol. 96, no. 8, aug 2008.
- [3] C. S. Thakur *et al.*, "Large-Scale Neuromorphic Spiking Array Processors: A Quest to Mimic the Brain," *Front. Neurosci.*, vol. 12, pp. 1–37, dec 2018.
- [4] C. D. Schuman *et al.*, "A survey of neuromorphic computing and neural networks in hardware," 2017. [Online]. Available: <https://arxiv.org/abs/1705.06963>
- [5] K. Guo, S. Han, S. Yao, Y. Wang, Y. Xie, and H. Yang, "Software-Hardware Codesign for Efficient Neural Network Acceleration," *IEEE Micro*, vol. 37, no. 2, pp. 18–25, mar 2017.
- [6] G. Indiveri *et al.*, "Neuromorphic silicon neuron circuits," *Frontiers in Neuroscience*, vol. 5, 2011. [Online]. Available: <https://doi.org/10.3389/fnins.2011.00073>
- [7] I. Sourikopoulos *et al.*, "A 4-fj/spike artificial neuron in 65 nm cmos technology," *Frontiers in Neuroscience*, vol. 11, 2017. [Online]. Available: <https://doi.org/10.3389/fnins.2017.00123>
- [8] F. Danneville, C. Loyez, K. Carpentier, I. Sourikopoulos, E. Mercier, and A. Cappy, "A sub-35 pw axon-hillock artificial neuron circuit," *Solid-State Electronics*, vol. 153, pp. 88–92, 2019. [Online]. Available: <https://doi.org/10.1016/j.sse.2019.01.002>
- [9] F. Danneville, K. Carpentier, I. Sourikopoulos, M. Paindavoine, and C. Loyez, "Sub-0.3V CMOS neuromorphic technology and its potential application," in *Int. Conf. Content-Based Multimed. Index*. Lille, France: IEEE, jun 2021, pp. 1–6.

- [10] P. M. Ferreira, J. Nebhen, G. Klisnick, and A. Benlarbi-Delai, "Neuro-morphic analog spiking-modulator for audio signal processing," *Analog Integr. Circuits Signal Process.*, vol. 106, no. 1, pp. 261–276, jan 2021.
- [11] M. Daliri, P. M. Ferreira, G. Klisnick, and A. Benlarbi-Delai, "A comparative study between E - neurons mathematical model and circuit model," *IET Circuits Devices Syst.*, vol. 15, pp. 175–192, feb 2021.
- [12] A. Destexhe, Z. Mainen, and T. Sejnowski, "Kinetic models of synaptic transmission," *Methods in Neuronal Modelling, From Ions to Networks*, vol. 2, pp. 1–26, 01 1998.
- [13] C. Bartolozzi and G. Indiveri, "Synaptic dynamics in analog vlsi," *Neural computation*, vol. 19, pp. 2581–603, 11 2007.
- [14] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2, no. 4, pp. 303–314, Dec 1989. [Online]. Available: <https://doi.org/10.1007/BF02551274>
- [15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, J. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: <https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>
- [16] P. Chevalier *et al.*, "A 55 nm triple gate oxide 9 metal layers SiGe BiCMOS technology featuring 320 GHz  $f_T$  / 370 GHz  $f_{MAX}$  HBT and high-Q millimeter-wave passives," in *Int. Electron Devices Meet. IEDM*, San Francisco, CA, USA, dec 2014, pp. 3.9.1–3.9.3. [Online]. Available: 10.1109/IEDM.2014.7046978

AUTHOR VERSION

AUTHOR