HAL
open science

Energy efficient fJ / spike LTS e-Neuron using 55-nm node

Pietro Maris Ferreira, Nathan de Carvalho, Geoffroy Klisnick, A. Benlarbi-Delai

To cite this version:

Pietro Maris Ferreira, Nathan de Carvalho, Geoffroy Klisnick, A. Benlarbi-Delai. Energy efficient fJ / spike LTS e-Neuron using 55-nm node. 32nd SBC/SBMicro/IEEE/ACM Symposium on Integrated Circuits and Systems Design (SBCCI), Aug 2019, São Paulo, Brazil. 10.1145/3338852.3339852 . hal-02290309

HAL Id: hal-02290309

https://hal.science/hal-02290309v1

Submitted on 5 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.

Energy efficient fJ/spike LTS e-Neuron using 55-nm node

Pietro M. Ferreira
Nathan De Carvalho
maris@ieee.org
nathan.de-carvalho@student-cs.fr
GeePs, UMR CNRS 8507, CentraleSupélec, Univ. Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France

Geoffroy Klisnick
Aziz Benlarbi-Delai
geoffroy.klisnick@sorbonne-universite.fr
aziz.benlarbi_delai@sorbonne-universite.fr
Sorbonne Université, Laboratoire d'Electronique et Electromagnetisme, L2E, F-75005, Paris, France

ABSTRACT

While CMOS technology is currently reaching its limits in power consumption and circuit density, a challenger is emerging from the analogy between biology and silicon. Hardware-based neural networks may drive a new generation of bio-inspired computers by the urge of a hardware solution for real-time applications. This paper redesigns a previous proposed electronic neuron (e-Neuron) in a higher firing rate to reduce the silicon area and highlight a better energy efficiency trade-off. Besides, an innovative schematic is proposed to state an e-Neuron library based on Izhikevichs model of neural firing patterns. Both e-Neuron circuits are designed using 55 nm technology node. Physical design of transistors in weak inversion are discussed to a minimal leakage. Neural firing pattern behaviors are validated by post-layout simulations, demonstrating the spike frequency adaptation and the rebound spikes due to post-inhibitory effect in LTS e-Neuron. Presented results suggest that the time to rebound spikes is dependent of the excitation current amplitude. Both e-Neurons have presented a fF/spike energy efficiency and a smaller silicon area in comparison to Izhikevichs library propositions in the literature.

CCS CONCEPTS

• **Hardware** → **Analog and mixed-signal circuit synthesis**; **Standard cell libraries**; *Emerging architectures*.

KEYWORDS

non-linear electronics, ultra-low-power, neuromorphic circuits

ACM Reference Format:
Pietro M. Ferreira, Nathan De Carvalho, Geoffroy Klisnick, and Aziz Benlarbi-Delai. 2019. Energy efficient fJ/spike LTS e-Neuron using 55-nm node. In *32nd Symposium on Integrated Circuits and Systems Design (SBCCI '19), August 26–30, 2019, Sao Paulo, Brazil.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3338852.3339852

1 INTRODUCTION

Moore's law has been a reliable predictor of the pace of electronic technology advancement [10]. While CMOS technology is currently reaching its limits in power consumption and circuit density, a challenger is emerging from the analogy between biology and silicon. Biomimetic-hardware may drive a new generation of bio-inspired computers driven by the urge of a hardware solution for real-time applications in robotics or in data mining and signal classification. Flexibility, area, power-efficiency, and reliability are some of the most significant challenges addressed by such biomimetic-hardware implementation.

From all biological cells, one of the most intriguing is the neuron, thanks to its memory and computing capabilities. A biological neuron membrane potential (V_m) is excited by a current pulse (I_{ex}) of few hundreds of pico-Amps. Thus, it operates in an average firing rate (f_{spike}) of few Hz with an energy efficiency (E_{eff}) of 2.45 pJ/spike. Such neurons have an average membrane capacitance (C_m) of 245 pF and operate with an action potential (V_d) of 100 mV [5]. A specific class of neurons which draws all attention of this paper is the cortex neurons [8]. The reason of such interest is due to cortex-neuron capabilities in terms of biological sensing and actuation. Biomimetic-hardware characteristics addressed in [5] might turn in favors of cheap and low-power bio-inspired computers if f_{spike} and E_{eff} trade-off is considered. Literature lacks of studies to address area and power-efficiency challenge, while presenting a variety of schematics for Izhikevich's neuron models in low f_{spike} operation and pJ/spike E_{eff} range [7]. Recently, I. Sourikopoulos *et al.* have proposed a fJ/spike E_{eff} solution for such problem, showing one firing pattern of Izhikevich's models [12].

This paper aims at a higher f_{spike} and a lower C_m, to minimize both E_{eff} and the silicon area. Starting from [12] proposal, this paper redesigns such electronic neurons (e-Neuron) and proposes an innovative e-Neuron schematic. Towards a biomimetic fJ/spike e-Neuron library, both e-Neurons are analyzed in order to demonstrate two out of six firing pattern in Izhikevich's models [8]. Circuits are designed using 55 nm technology node to explore a high capacitance density (i.e. 0.2 fF/μm^2), and post-layout simulations are carried out to highlight e-Neuron figure-of-merits.

2 BACKGROUND

2.1 Cortex neurons

Cortex neurons are involved in higher functions such as sensory perception, generation of motor commands, spatial reasoning, conscious thought, and human's language. Actually, most biologists agree with the classification of cortex neurons in six most fundamental classes of firing patterns observed in the mammalian neocortex [8]. These cortex neurons are:

RS Regular spiking neurons fire tonic spikes with adapting frequency in response to injected pulses of DC current.

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.
SBCCI '19, August 26–30, 2019, Sao Paulo, Brazil
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-6844-5/19/08…$15.00
https://doi.org/10.1145/3338852.3339852

IB Intrinsically bursting neurons generate a burst of spikes at the beginning of a strong depolarizing pulse of current, then switch to tonic spiking mode.

CH Chattering neurons fire high-frequency bursts of spikes with relatively short inter-burst periods; hence fast rhythmic bursting.

FS Fast spiking neurons fire high-frequency tonic spikes with relatively constant period.

LTS Low-threshold spiking neurons fire tonic spikes with pronounced spike frequency adaptation (decreasing) and rebound spikes due to post-inhibitory effect.

LS Late spiking neurons exhibit voltage ramp in response to current excitation, resulting in delayed spiking.

Cortex neuron firing patterns are often modeled using the Hodgkin-Huxley [6] or the Moris-Lecar model [9]. Both most accepted models behave as non-linear leak integration and firing by exponential functions of real numbers. In CMOS technology, such behavior is obtained in weak inversion (subthreshold regime) [4]. Besides, transistors operating in such regime also take advantage of ultra-low-power consumption by a supply voltage around $\pm V_d$. Drawbacks of such biomimetic-hardware implementation are obtaining a higher f_{spike} and a lower C_m than biological counterparts.

2.2 Silicon neurons

Since 2010, e-Neuron library proposition of Izhikevich's models has been interested in the scientific community. V. Rangan *et al.* have proposed a current-mirror-based schematic capable to mimic all classes of firing patterns[11]. Rangan's proposal achieves a silicon area of 2980 μm^2 and an E_{eff} of 1 pJ/spike using MOSIS 90 nm technology. Eventually, this work did not consider a V_m (current-mode circuit) which is an indispensable variable to model the integrating phase or the polarizing region in biological cell analogy. Besides, A. Schaik *et al.* have proposed a Izhikevich neuron model implementation using two first-order log-domain low-pass filters and two trans-linear multipliers [14]. Schaik's proposal achieves a silicon area of 0.02 mm^2, f_{spike} = 40 Hz, and an E_{eff} = 1 nJ/spike using AMIS 0.5 μm technology. Both have suggested weak inversion operation to mimic Izhikevich's models behavior.

Later then, a variety of e-Neuron implementations were presented in [7]. Most of implementations have a smaller silicon area than Rangan's or Schaik's work (around 1000 μm^2) but achieving an E_{eff} of few pJ/spike in low f_{spike} (i.e. few Hz range). Demirkol's proposal is one of such examples achieving a chip area of 1100 μm^2, f_{spike} = 10 Hz, and E_{eff} = 165 pJ/spike using AMS 0.35 μm process [3].

Recently, X. Wu *et al.* have included the learning capability based on resistive synapses in their spiking neuron proposal [15]. Using class AB-OTAs, Wu's proposal presented a silicon area of 110 x 110 μm^2 and E_{eff} = 9.3 pJ/spike using 0.18 μm CMOS process. M. Azghadi *et al.* have improved the learning capabilities using memristor-based synapses [1]. Using a non-conventional technology node, their work presented a silicon area of 600 μm^2. M. Azghadi *et al.* affirmed that power consumption was not the design goal. Both have neither detailed f_{spike} range, or Izhikevich's models mathematical behavior (i.e. exponential functions of real numbers) which requires transistors operating in weak inversion.

I. Sourikopoulos *et al.* has innovated by a biomimetic- and a simplified-version of FS e-Neurons. Discussing the trade-off between speed and consumption, biomimetic version achieves an area of 300 μm^2, f_{spike} = 1 kHz, and an E_{eff} = 40 fJ/spike; simplified version obtained an area of 35 μm^2, f_{spike} = 26 kHz, and an E_{eff} = 4 fJ/spike. Both circuits are implemented using TSMC 65 nm technology and considered only dynamic power consumption in figure of merit. This work opens the way to propose an exhaustive library regarding Izhikevich's neuron models operating in a higher f_{spike} with E_{eff} in fJ/spike range.

Towards a large-scale spiking neural network, L. Zhang *et al.* have made an effort in defining a PVT-robust e-Neurons [16]. A lower $C_m = 1\ fF$ is presented to address area and reliability challenges. Eventually, e-Neurons operates at 1 V supply, extremely high f_{spike}, and lower power consumption. Figure of merit values are not available, and results suggest that Izhikevich's models mathematical behavior (i.e. exponential functions of real numbers) were not considered. The area is estimated at 150 μm^2 using 65 nm CMOS technology.

2.3 Weak Inversion

Weak Inversion (WI) model of saturated MOS transistor is obtained by integrating the charge conduction for the channel length considering the minority carriers injected into the channel. Drain-to-source current relationship is then obtained as [4]:

$$\begin{aligned} I_{DS} &= I\left(V_G, V_S\right) - I\left(V_G, V_D\right) \\ &= I_s e^{\frac{V_{GS}}{\eta\phi_t}} - I_s e^{\frac{V_{GD}}{\eta\phi_t}} \\ &= I_s e^{\frac{V_{GS}}{\eta\phi_t}} \left(1 - e^{\frac{-V_{DS}}{\eta\phi_t}}\right), \end{aligned} \tag{1}$$

where ϕ_t is the thermal voltage $\frac{kT}{q}$; η is the slope factor $1 + \frac{C_d}{C_{ox}}$ (i.e. depletion C_d and oxide C_{ox} capacitance ratio); and I_s is the specific current. Considering small variations of V_{DS}, (1) can be approximated by [12]

$$I_{DSi} = G_i e^{\frac{V_{GSi}}{\eta\phi_t}} V_{DSi}, \tag{2}$$

where the device conductance G_i for a i transistor can be estimated as a function of the mobility (μ) and the transistor sizing by

$$G_i = \mu C_{ox} \frac{\phi_t}{2} \frac{W}{L}. \tag{3}$$

Another interesting result presented in [12] is the association of a PMOS (MP_i) and a NMOS (MN_i) transistors biased in WI, having drain node connected as V_{out}, gate node as V_{in}, and source node bias as $\pm V_d$. When applying Kirchhoff's Current Law (KCL) at the transistors' drain node of the resultant transconductor, expressing the currents as in (2), assuming the same η for both transistors; the output voltage of the transconductor in static WI regime can be expressed as:

$$V_{out} = -V_d \cdot \tanh\left(\frac{V_{in}}{\eta\phi_t} + \frac{1}{2}\ln\left(\frac{G_N}{G_P}\right)\right). \tag{4}$$

Equation (4) assumes that the output current of the transconductor MP_i/MN_i is negligible. In the circuit schematics studied in this

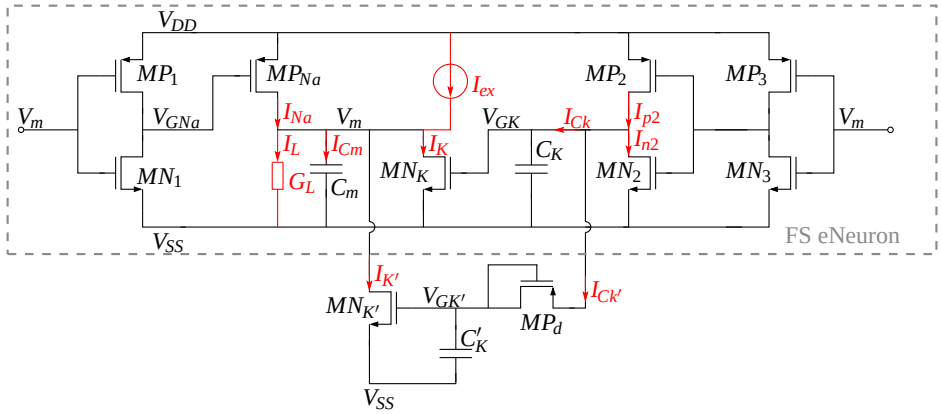

Figure 1: Proposed LTS e-Neuron schematic having a FS e-Neuron core highlighted in the gray box ($V_{DD} = V_d$ and $V_{SS} = -V_d$).

paper, the output current is in fact the gate-leakage of a following transconductor (MP_{i+1}/MN_{i+1}) playing the role of a load for MP_i/MN_i.

State-of-the-art biomimetic-hardware uses capacitors ($I_i = C \cdot \frac{dV_i}{dt}$) associated to transconductor's non-linear behavior expressed in (4), in order to model the leak integration and firing behavior of cortex neurons. Indeed, an e-Neuron is a current-controlled oscillator having I_{ex} as the control current and V_m the spiking output node. This is obtained from positive feedback association of capacitors and transconductors.

3 LTS E-NEURON PROPOSAL

Figure 1 illustrates the schematic topology of the proposed biomimetic neuron. Based on FS e-Neuron highlighted in the gray box [12], MP_{Na} and MP_K transistors biomimic the ionic movement of *Na*-in and *K*-out the membrane C_m. The I_{ex} source models the current pulse excitation and I_L the current leak. Thus, LTS e-Neuron integrates and fires tonic spikes similarly to FS e-Neuron when $I_{ex} \geq I_L$. To biomimic the post-inhibitory phenomenon, a non-linear RC circuit is implemented using a diode-connected MP_d transistor and $C_{K'}$. Following subsections will draw a circuit analysis of both e-Neurons considering transistors in weak inversion.

3.1 FS e-Neuron Analysis

I. Sourikopoulos *et al.* have first demonstrated the fast spiking behavior of biomimetic neuron with relatively constant period in [12]. Later, I. Sourikopoulos *et al.* have demonstrated the stochastic resonance behavior of such a neuron [2]. This phenomenon is observed under a decreased I_{ex} (below a certain critical value), where the neuron fires irregular spikes, switching randomly between resting and spiking states. According to [8], both characteristics are presented only for FS neurons, and thus Sourikopoulos' work is named as FS e-Neuron in this paper.

Applying KCL at nodes V_m and V_{GK} considering the gray-box schematics in Fig. 1, one gets the system of equations

$$\begin{aligned} I_{Cm} &= I_{Na} - I_K + I_{ex} - I_L, \\ I_{Ck} &= I_{P2} - I_{N2}. \end{aligned} \tag{5}$$

Using (2) and (4); assuming transconductors are in static WI regime, $V_{DD} = V_d$ and $V_{SS} = -V_d$; the (5) system leads to the following coupling system between V_m and V_{GK}

$$\begin{aligned} C_m \frac{dV_m}{dt} =& G_{Na} e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N1}}{G_{P1}}\right)\right)\right)}{\eta\phi_t}} \cdot \left(V_d - V_m\right) \\ & - G_K e^{\frac{V_{GK}+V_d}{\eta\phi_t}} \cdot \left(V_m + V_d\right) + I_{ex} - G_L \cdot V_m, \\ C_K \frac{dV_{GK}}{dt} =& G_{P2} e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N3}}{G_{P3}}\right)\right)\right)}{\eta\phi_t}} \cdot \left(V_d - V_{GK}\right) \\ & - G_{N2} e^{\frac{V_d\left(1-\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N3}}{G_{P3}}\right)\right)\right)}{\eta\phi_t}} \cdot \left(V_d + V_{GK}\right); \end{aligned} \tag{6}$$

with the parameters being:

- C_m, C_K are capacitances;
- V_m, V_{GK} are voltages applied to the respective capacitances';
- I_{ex}, I_L are the excitation and leak currents;
- G_L, G_{Na}, G_K are the leak, sodium and potassium conductances;
- G_{Pi}, G_{Ni} $\forall i \in \{1,2,3\}$ are the *i*-transistor conductances.

Thus, (6) is a time-invariant, first-order, non-linear system relating the V_m and the V_{GK} potentials. It depends on ten parameters $I_{ex}, G_L, G_{Na}, G_K, \left(G_{pi}, G_{ni}\right) \forall i \in \{1,2,3\}$ when both temperature (ϕ_t = 26 mV) and action potential (V_d = 100 mV) are fixed. A circuit designer would probably like to solve (6) system using MatLab, for instance. However, no analytical solution neither any analytical equilibrium point can be found for such system. In fact, when considering the exponential argument $\frac{V_d}{\eta\phi_t} \approx 5$ and noticing that experimental data of V_m can have the same order of magnitude as $\eta\phi_t$ [12], one may conclude that the linearization of $\tanh\left(x\right)$ and $e^{(x)}$ functions has no physical meaning for this non-linear circuit.

3.2 LTS e-Neuron Analysis

To biomimic the post-inhibitory phenomenon in LTS neurons, the I'_K current is designed to be a delayed version of I_K. By increasing the total $I_K + I'_K$, a pronounced spike frequency adaptation is mimicked. Furthermore, rebound spikes due to post-inhibitory effects will take place when $I_{ex} - I'_K = I_{ex,eff} \leq I_L$. Applying KCL at nodes

V_m and V_{GK} for the LTS e-Neuron, the (5) system turns into

$$\begin{aligned} I_{Cm} =& I_{Na} - I_K - I_{K'} + I_{ex} - I_L, \\ I_{Ck} + I_{Ck'} =& I_{p2} - I_{n2}. \end{aligned} \tag{7}$$

Thus, the additional currents I'_K and $I_{Ck'}$ are time-dependent model for the spike frequency adaptation. Assuming C'_K in charged phase and $V_{SS} = -V_d$ constant, the MP_d-$C_{K'}$ circuit loop is modeled from

$$\begin{aligned} I_{Ck'} &= C_{K'} \frac{d\left(V_{GK'}\right)}{dt} \\ &\approx e^{\frac{1}{2\cdot\eta}} \cdot G_{Pd} \cdot \left(V_{GK} - V_{GK'}\right), \end{aligned} \tag{8}$$

where $V_{GS} = V_{DS} = V_{GK'} - V_{GK}$ for MP_d transistor and $MN_{K'}$ transistor gives

$$I_{K'} = G_{K'} e^{\frac{V_{GK'}+V_d}{\eta\phi_t}} \cdot \left(V_m + V_d\right). \tag{9}$$

Similarly to the FS e-Neuron, by expressing the (7) system with the parameters of the circuit from (8) and (9), one gets (10). Since $V_{GK'}$ is dependent on time, the (10) is a time-variant, first-order, non-linear system relating the V_m and the V_{GK} potentials. It depends on thirteen parameters $I_{ex}, G_L, G_{Na}, G_K, G_{pi}, G_{ni} \forall i \in \{1,2,3\}, G_{Pd}$, $C_{K'}, G_{K'}$ when both temperature (ϕ_t = 26 mV) and action potential (V_d = 100 mV) are fixed. Thus, the additional branch has introduced a time dependency in the system of equations required to model the variation of the spiking frequency over time.

$$\begin{aligned} C_m \frac{dV_m}{dt} =& G_{Na} e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_N1}{G_P1}\right)\right)\right)}{\eta\phi_t}} \cdot \left(V_d - V_m\right) \\ &- G_K e^{\frac{V_{GK}+V_d}{\eta\phi_t}} \cdot \left(V_m + V_d\right) \\ &- G_{K'} e^{\frac{V_{GK'}+V_d}{\eta\phi_t}} \cdot \left(V_m + V_d\right) \\ &+ I_{ex} - G_L \cdot V_m, \\ C_K \frac{dV_{GK}}{dt} =& G_{P2} e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_N3}{G_P3}\right)\right)\right)}{\eta\phi_t}} \cdot \left(V_d - V_{GK}\right) \\ &- G_{N2} e^{\frac{V_d\left(1-\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_N3}{G_P3}\right)\right)\right)}{\eta\phi_t}} \cdot \left(V_d + V_{GK}\right) \\ &- C_{K'} \frac{d\left(V_{GK'}\right)}{dt} \end{aligned} \tag{10}$$

In fact, simulation experiments could prove that $V_{GK} \approx -V_d$ being almost constant over time during the integration phase, and only spiking in firing phase. Therefore, one may solve (8) verifying that $V_{GK'}\left(t\right)$ depends on previous value of $I_{ex(i-1)}$ according to $C_{K'}$ charge and discharge cycles. The delay characteristic introduced by the additional branch could be given by a time constant

$$\tau = \frac{C_{K'}}{e^{\frac{1}{2\cdot\eta}} \cdot G_{Pd}}. \tag{11}$$

Consequently, the settling time (i.e. 5τ) of MP_d-$C_{K'}$ circuit is the time required to the frequency adaptation (decrease), assuming a discharged $C_{K'}$ (i.e. $I_{ex0} = 0$). Assuming a charged $C_{K'}$ (i.e. $I_{ex}\left(i-1\right) \neq 0$), the settling time of MP_d–$C_{K'}$ circuit is the time required to rebound spikes due to post-inhibitory effect in LTS e-Neuron. To conclude, the voltage-value charged in $C_{K'}$ actually biomimics the memory information of previous excitation ($I_{ex}\left(i-1\right)$) depending on the charge of the capacitance $C_{K'}$ and its ability to discharge itself throughout time.

3.3 e-Neuron physical design

I. Sourikopoulos *et al.* have proposed a first sizing of FS e-Neuron in 65 nm technology node [12]. Silicon area is dominated by the total capacitance value, depending on the capacitance density in the according node (i.e. 0.2 fF/μm^2) [2]. The 55 nm technology node, used in this paper, is only a shrink down node of 65 nm technology using a 0.9 optical-lens reduction factor with similar capacitance density. From [12], f_{spike} and E_{eff} trade-off can be improved by increasing f_{spike}. In this case, static energy consumption is reduced when f_{spike} increases, and dynamic energy consumption is also reduced when total load capacitance decreases (i.e. C_m) to reach a higher f_{spike}. To find such a better trade-off, a designer could use (3) to resize MN_i and MP_i transistors from Fig. 1.

LTS e-Neuron design is carried out aiming a higher-f_{spike}, as also FS e-Neuron proposed in this paper. Considering a I_{ex} pulse width of 1 ms, MP_d-$C_{K'}$ circuit is sized for a $\tau \approx 40$ μs having $G_K = G_{K'}$ for simplicity. After some design-of-experiments, Table 1 presents the final e-Neuron circuit sizing. Indeed, layout constraints were considered in choosing C_m, C_K, $C_{K'}$ to improve circuit fill-factor. To reduce process variability, dummy-capacitors were added in the layout and C_m, C_K, $C_{K'}$ were positioned in a common-centroid to minimize mismatch.

Likewise, one may argue that transistors should be inter-digitized to minimize gate resistance and reduce transistor mismatch. This common sense in circuit layout also increases the area of the depletion region of weak inverted transistors leading to an increasing bulk-leakage current. The G_L conductance, see Fig. 1, is not implemented in both e-Neurons, but it represents the total area of such depletion region. The I_L is the sum in i of bulk-leakage current from all i-transistors. Bottom-line, it will be required a higher I_{ex} in integrating phase, if inter-digitization layout technique is used. Besides, it is expected that LTS e-Neuron requires a higher I_{ex} due to the additional area of the MP_d depletion region.

Still, temperature variation represents a major limitation for transistors operating in WI regime. Studies have proved that leakage current is exponentially dependent on temperature; and it doubles for every 10 °C increase in temperature [13]. Thus, I_L shall increase over temperature enabling e-Neuron firing in a higher $I_{ex,min}$, and increasing static energy consumption. Temperature variation is beyond the scope of this paper.

Reduce the static energy consumption is the key to improve E_{eff} even further and to enable e-Neuron to fire with a smaller $I_{ex,min} \geq I_L$. Since static energy consumption does not scale with f_{spike}, physical design shall not neglect described phenomenon. To the best of our knowledge, these layout considerations for weak inverted transistors is first presented in this paper.

4 POST-LAYOUT RESULTS

The e-Neuron circuit layouts are omitted for lack of space. FS e-Neuron has an area of 85 μm^2 and LTS e-Neuron has 120 μm^2. One might notice that capacitors C_m, C_K, $C_{K'}$ take most of the layout area. By increasing f_{spike}, this paper heads toward a large-scale integration challenge using smaller capacitors. The area of designed e-Neurons are one order of magnitude smaller than [7], and slightly bigger than reported in [12].

Table 1: Sizing of e-Neuron circuits in W x L for MN_i and MP_i and in number of cells x unity capacitance for C_i.

MN_1	180 nm x 55 nm	MP_1	121.5 nm x 55 nm
MN_2	121.5 nm x 55 nm	MP_2	1080 nm x 55 nm
MN_3	121.5 nm x 55 nm	MP_3	180 nm x 55 nm
MN_K	1800 nm x 55 nm	MP_{Na}	720 nm x 55 nm
C_m	9 x 931 aF	C_K	16 x 931 aF
$MN_{K'}$	1800 nm x 55 nm	MP_d	450 nm x 9000 nm
		$C_{K'}$	11 x 931 aF

Post-layout simulations carried out to validate the biomimetic behavior by a fixed pulse width of 5 ms and an increasing I_{ex} amplitude from few pico-amps. When $I_{ex} \leq I_L$, the leak integration is negligible and e-Neuron fires at f_{spike} after that point. Post-layout results reveal an I_L of 22 pA for FS and of 132 pA for LTS. Figure 2 illustrates the f_{spike} pattern for both e-Neurons excited by a current pulse of 30 pA for FS and of 150 pA for LTS. It is observed a tonic $f_{spike} \approx 20$ kHz for both e-Neurons. As expected, f_{spike} of LTS e-Neuron decreases over time, achieving a final value of 5 kHz after 0.2 ms.

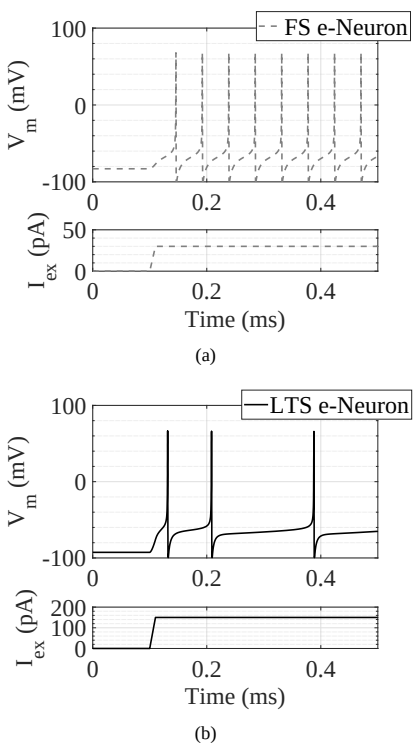

Figure 2: Post-layout transient simulation results for both e-Neurons excited by a current pulse, being: (a) a I_{ex} of 30 pA for FS and (b) of 150 pA for LTS. A tonic $f_{spike} \approx 20$ kHz is observed for both e-Neurons, and spike frequency adaptation is achieved for LTS e-Neuron after 0.2 ms.

To demonstrate the spike frequency adaptation and the rebound spikes due to post-inhibitory effect, a piecewise linear I_{ex} pulse is applied to both e-Neurons. Figure 3(a) presents the estimated f_{spike} response over time for piecewise linear I_{ex} pulses. Four different amplitudes are considered in [30 pA, 45 pA, 60 pA, 75 pA] for FS, and in [150 pA, 225 pA, 300 pA, 375 pA] for LTS e-Neuron. I_{ex} amplitude range under study is considered in increasing (regions from **a** to **d**) and decreasing (regions from **d** to **g**) manners. FS e-Neuron, in dashed-gray line, shows tonic spikes with relatively constant period validating the Izhikevich's model. LTS e-Neuron, in continuous black line, highlights a frequency adaptation (decreasing) among regions from **a** to **d**. However, rebound spikes due to post-inhibitory effect are observed among regions from **d** to **g**.

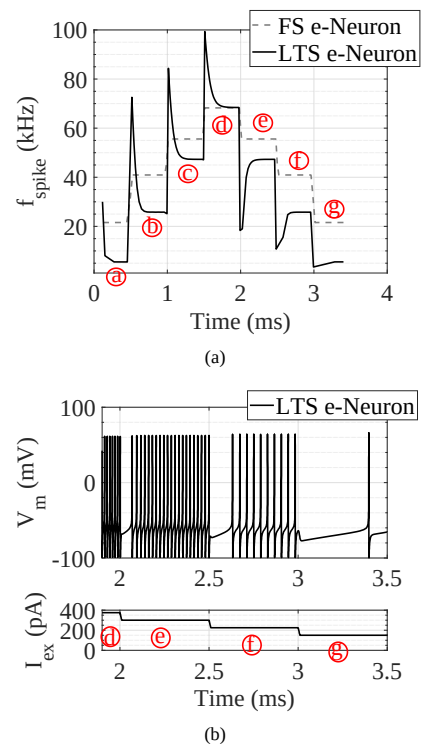

Figure 3: Post-layout simulations demonstrates the spike frequency adaptation and the rebound spikes due to post-inhibitory effect. (a) The f_{spike} response over time for piecewise linear I_{ex} pulses [30 pA, 45 pA, 60 pA, 75 pA, 60 pA, 45 pA, 30 pA] for FS (dashed-gray line) and [150 pA, 225 pA, 300 pA, 375 pA, 300 pA, 225 pA, 150 pA] for LTS e-Neuron (continuous black line). (b) The V_m spiking behavior during I_{ex} decreasing pulse regions for LTS e-Neuron.

Figure 3(b) illustrated the V_m spiking behavior during I_{ex} amplitude decreasing regions from **d** to **g**. Indeed, LTS e-Neuron does not spike due to post-inhibitory effect but starts firing at a low-f_{spike} consistent to value achieved after adaptation. Interesting results are revealed in Fig. 3(b), suggesting that the time to rebound spikes is actually dependent to I_{ex} pulse variation. The solution of (8) considering weak inversion is in fact more complicated than a simple

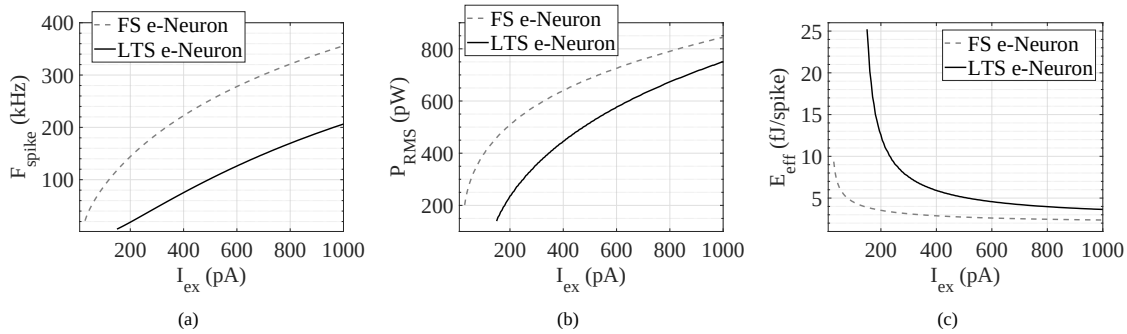

Figure 4: Figure-of-Merits extracted from post-layout simulation results of LTS and FS e-Neurons: (a) f_{spike}, (b) P_{RMS}, and (c) E_{eff}.

MP_d–$C_{K'}$ circuit charge and discharge, and Fig. 3(b) results suggests that GP_d is I_{ex} dependent.

Following the Fig. 3(a) insights, it is expected an increasing f_{spike} as I_{ex} increases. Figure 4 illustrates the figure-of-merits of e-Neurons as I_{ex} increases from estimated I_L to 1 nA using a 1 pA linear step. Figure 4(a) illustrates the growth of f_{spike} versus I_{ex}. The f_{spike} of LTS e-Neuron after frequency adaptation is lower than f_{spike} of FS e-Neuron. By setting a higher f_{spike} than found in the literature, a power consumption increase is expected (i.e. dynamic consumption). Figure 4(b) presents LTS and FS e-Neuron power consumption in the range of few hundreds of pico-Watts. This is a slight increase compared to best (around 100 pW) in [12]. Despite that, the E_{eff} is slightly decreased owing to a higher f_{spike} (around 200 kHz) and specific e-Neuron physical design. Considering static and dynamic energy consumption, figure 4(c) highlights a E_{eff} as low as 2.3 fJ/spike for FS and 3.6 fJ/spike for LTS e-Neurons.

Table 2 summarizes the literature comparison. Obtained area is limited by technology capacitance density as reported in [12]. However, f_{spike} is higher than reported values. This design choice leads to a better trade-off in speed and power consumption. Table 2 highlights a E_{eff} slightly below compared to best in [12], which considers only dynamic consumption. In general, both e-Neurons have presented an E_{eff} below e-Neuron library propositions in the state-of-the-art [7], while ensuring the Izhikevich's mathematical modeling in contrast to simplified neuron in [12].

Table 2: e-Neuron performance comparison.

Ref.	This work	[11]	[14]	[3]	[15]	[12]
Techn.	55	90	500	350	180	65
(nm)	MOS	MOS	MIS	MOS	MOS	MOS
Area	$120^{\dagger}$	2980	20k	1100	10k	300^{*}
(μm^2)	$85^{\ddagger}$					$35^{\star}$
f_{spike}	205 $k^{\dagger}$	7 k	40	10	NA	1 k^{*}
(Hz)	360 $k^{\ddagger}$					26 $k^{\star}$
E_{eff}	2.3 $f^{\dagger}$	1 p	1 n	165 p	9.3 p	40 f^{*}
(J/spike)	3.6 $f^{\ddagger}$					4 $f^{\star}$

† LTS; ‡ FS; * biomimetic; * simplified

5 CONCLUSIONS

A first step towards an ultra-low-power e-Neuron library, using weak inversion CMOS 55 nm node, has been achieved. Two e-Neurons acting as a Fast Spiking (FS) and Low Threshold Spiking (LTS) neurons have been presented. Operating in a higher f_{spike} than found in the literature, these e-Neurons meet the trade-off in terms of silicon area (few tens of μm^2) and energy efficiency (few fJ/spike). For the LTS e-Neuron frequency adaptation and rebound spikes due to post-inhibitory effect were carried out. Future works deal with the four-remaining e-Neurons miming the cortex neuron patterns.

REFERENCES

[1] M. R. Azghadi et al. 2017. A Hybrid CMOS-Memristor Neuromorphic Synapse. *IEEE Trans. Biomed. Circuits and Syst.* 11, 2 (2017), 434–445.
[2] F. Danneville et al. 2017. Ultra low power analog design and technology for artificial neurons. In *Proc IEEE Bipolar/BiCMOS Circuits and Techn. Meeting*. IEEE, Florida, USA, 1–8.
[3] A. Samil Demirkol and Serdar Ozoguz. 2011. A low power VLSI implementation of the Izhikevich neuron model. In *IEEE New Circuits Syst. Conf.* IEEE, Bordeaux, France, 169–172.
[4] Carlos Galup and Marcio Schneider. 2018. The compact all-region MOSFET model : theory and applications. In *IEEE New Circuits Syst. Conf.* IEEE, 166–169.
[5] Jennifer Hasler and Bo Marr. 2013. Finding a roadmap to achieve large neuromorphic hardware systems. *Frontiers in Neuroscience* 7 (sep 2013), 1–29.
[6] A. L. Hodgkin and A. F. Huxley. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiology* 117, 4 (mar 1952), 500–544.
[7] G. Indiveri et al. 2011. Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience* 5, MAY (may 2011), 1–23.
[8] E.M. M Izhikevich. 2003. Simple Model of Spiking Neurons. *IEEE Trans. Neural Networks* 14, 6 (nov 2003), 1569–1572.
[9] C. Morris and H. Lecar. 1981. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal* 35, 1 (jan 1981), 193–213.
[10] James R Powell. 2008. The Quantum Limit to Moore's Law. *Proc. IEEE* 96, 8 (aug 2008), 1247–1248.
[11] V. Rangan et al. 2010. A subthreshold a VLSI implementation of the Izhikevich simple neuron model. In *Int. Conf Eng. Med. Biol.*, IEEE (Ed.). IEEE, Buenos Aires, Argentina, 4164–4167. https://doi.org/10.1109/IEMBS.2010.5627392
[12] I. Sourikopoulos et al. 2017. A 4-fJ/spike artificial neuron in 65 nm CMOS technology. *Frontiers in Neuroscience* 11, 123 (mar 2017), 1–14.
[13] S. M. Sze. 1981. *Physics of Semiconductor Devices* (2nd ed.). Jonh Wiley 'Sons, New York. 868 pages.
[14] A. Schaik et al. 2010. A log-domain implementation of the Izhikevich neuron model. In *Proc. IEEE Int. Symp. Circuits Syst.* IEEE, Paris, France, 4253–4256.
[15] X. Wu et al. 2015. A CMOS Spiking Neuron for Brain-Inspired Neural Networks with Resistive Synapses and in Situ Learning. *IEEE Trans. Circuits and Syst. II: Express Briefs* 62, 11 (2015), 1088–1092.
[16] L. Zhang et al. 2018. A Multi-Mode Silicon Neuron Circuit with High Robustness Against PVT Variation. In *Proc. IEEE Int. Symp. Circuits Syst.* IEEE, Florence, Italy, 1–4.