HAL
open science

# Neuromorphic Analog Spiking-Modulator for Audio Signal Processing

Pietro Maris Ferreira, Jamel Nebhen, Geoffroy Klisnick, A. Benlarbi-Delai

## To cite this version:

Pietro Maris Ferreira, Jamel Nebhen, Geoffroy Klisnick, A. Benlarbi-Delai. Neuromorphic Analog Spiking-Modulator for Audio Signal Processing. Analog Integrated Circuits and Signal Processing, 2021, 106, 10.1007/s10470-020-01729-3 . hal-02948290

**HAL Id: hal-02948290**

**https://hal.science/hal-02948290v1**

Submitted on 9 Jun 2021

**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.

Analog Integrated Circuits and Signal Processing (2021) 106:261–276
https://doi.org/10.1007/s10470-020-01729-3

# Neuromorphic analog spiking-modulator for audio signal processing

Pietro M. Ferreira[1,2] · Jamel Nebhen[3] · Geoffroy Klisnick[1,2] · Aziz Benlarbi-Delai[1,2]

Received: 14 April 2020 / Revised: 5 September 2020 / Accepted: 28 September 2020 / Published online: 26 October 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

**Abstract**
While CMOS scaling is currently reaching its limits in power dissipation and circuit density, the analogy between biology and silicon is emerging as a solution to ultra-low-power signal processing. Urgent applications involving artificial vision and audition, including intelligent sensing, appeal original energy efficient and ultra-miniaturized silicon-based solutions. While state-of-the-art is focusing on digital-oriented solutions, this paper proposes a neuromorphic analog signal processor using Izhikevich-based artificial neurons in an analog spiking modulator. A varicap-based artificial neuron is explored reducing the silicon area to 98.6 $\mu m^2$ and the substrate leakage to a 1.95 fJ/spike efficiency. Post-layout simulation results are presented to investigate the high-resolution, high-speed, and full-scale dynamic range for audio signal processing applications. The proposal demonstrates a 9 bits spiking-modulator resolution, a maximum of 8 fJ/conv efficiency, and a root–mean–square error of 0.63 $mV_{RMS}$ .

**Keywords** Artificial neuron · Spiking signal processing · Non-linear electronics · Ultra-low power

## 1 Introduction

Neuromorphic computing appeared in the 90 s as a complementary architecture to von Neumann systems using analog circuits designed to mimic biological neural systems [1]. Since then, digital neuromorphic systems have been often implemented in FPGAs considering its shorter design and manufacturing time, reconfigurability and reusability for different applications [2]. From both circuitries, analog one has often been a good solution to implement the processing components of neurons and synapses, due to its ability to faithfully mimic biological systems. Besides, analog solutions have presented the best energy consumption per unit of information, often represented in J/spike. However, analog solutions have presented several reliability challenges in terms of process, temperature, and voltage variation that are frequently overcome by digital solutions [3]. A mixed circuit approach usually presents a better trade-off between digital and analog solutions [4]. Neuromorphic spiking signal processors are highly energy-efficient, parallel and distributed computing enabled, and require a small silicon area [5].

The immediate applications in neuromorphic spiking processors are artificial vision and audition by mimicking the retina and the cochlea, respectively [6] and [7]. Zaghloul and Boahen have proposed an artificial retina bio-inspired on the signals in the optic nerve in [6]. In their work, four major ganglion cell types that drive the visual cortex were modeled using spiking neural models and optical capabilities rely on local modulation of synaptic strength. Wen and Boahen have proposed the first integrated circuit modeling the cochlea's micromechanics [7]. Such an artificial cochlea demonstrates sound sensitivity, frequency selectivity, and dynamic range. Recent system

✉ Pietro M. Ferreira
maris@ieee.org

Jamel Nebhen
j.nebhen@psau.edu.sa

Geoffroy Klisnick
geoffroy.klisnick@sorbonne-universite.fr

Aziz Benlarbi-Delai
aziz.benlarbi_delai@sorbonne-universite.fr

1 Université Paris-Saclay, CentraleSupélec, CNRS, Lab. de Génie Électrique et Électronique de Paris, 91192 Gif-sur-Yvette, France

2 Sorbonne Université, CNRS, Lab. de Génie Électrique et Électronique de Paris, 75252 Paris, France

3 Prince Sattam bin Abdulaziz University, College of Computer Engineering and Sciences, P.O. Box 151, Alkharj 11942, Saudi Arabia



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

architectures are enabling neuromorphic spiking signal processing for complex and promising future applications [5].

In [8], nanowire microphones of an artificial cochlea exploit piezo resistivity to capture an acoustic vibration and transduces it to an electrical signal of a few millivolts amplitude in the frequency range from 20 Hz to 20 kHz. Traditionally, cochlea circuitry includes an analog front end [9] and multibit $\Delta\Sigma$ modulators in order to analog-to-digital convert the transduced audio signal. Such modulators are well known for its noise shaping capabilities and the resolution and speed trade-off. Co-integrated microphones and mixed-signal processor enables good speech comprehension and low background noise. Recent publications in artificial cochlea applications are focusing on intelligent acoustic sensing that combines the high energy efficiency and the signal processing capabilities such as spiking neural networks [5, 10–12].

The authors have focused on the analog-circuit design of the two cortex neuron models, named eNeuron [13]. Both eNeurons are part of the six most fundamental classes of firing patterns observed in the mammalian neocortex [14]. In a biological neuron for instance, the membrane potential ($V_m$) is excited by an ionic current pulse ($I_{ex}$) of a few hundreds of picoamperes. Thus, it operates in an average firing rate ($f_{spike}$) of few Hertz 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 [15]. Using electronic conduction, Rangan et al. [16] and Schaik et al. [17] have developed current-mirror-based architectures that are capable to mimic all classes of firing patterns. Their works have proved that once transistors are operating in the weak inversion regime, the Izhikevich's models mathematical behavior [14] can lead to energy efficient hardware devices mimicking the neocortex neurons.

Recently, Sourikopoulos et al. have considerably reduced the energy efficiency to 4 fJ/spike using CMOS 65 nm technology to implement a neuromorphic analog circuit solution [18], and in [19], a compact and energy efficient sub-threshold analog synapse and neuron circuits are presented. Using a 28 nm FD-SOI process, Qiao and Indiveri [19] have optimized neuromorphic circuitry to implement massively parallel large-scale neuromorphic computing systems. In [12], Yang et al. have proposed a spiking neuromorphic computing for cochlea-inspired analog front end and digital neural networks.

This paper extends the work already presented in [13] by proposing: (a) a neuromorphic analog spiking processor (NASP) using both eNeurons in a spiking-modulator architecture; and (b) a varicap-based version for both eNeuron to reduce silicon area. Proposed NASP aims at artificial cochlea applications coding an input-signal amplitude in $f_{spike}$ inspired by $\Delta\Sigma$ modulators. The two-output spiking-modulated signal in NASP is also capable to estimate the signal derivative sign from $f_{spike}$ variation over time combining both Izhikevich's neuron model proposed in [13]. Post-layout simulations (PLS) are carried out to highlight both eNeuron figure-of-merits, and to ensure low-power and low-silicon area with varicap-based version. NAPS simulation results demonstrate its capability to process audio signals from an artificial cochlea.

This paper is organized as follows. Section 2 presents a literature review of cortex neuron behavior, firing patterns, and electronic implementations which found the background our study is laid on. In addition, it highlights the principles of weak inversion model of saturated MOS transistors. Section 3 proposes the neuromorphic analog spiking processor implementation and analyses recent works dealing with energy efficient circuits [13] and [18]. The proposed NASP is designed using 55 nm technology node exploring high-density varicap and low supply voltage ($\sim$200 mV). Section 4 highlights PSL results with a focus on multi-spiking NASP operation in intelligent audio sensing. Finally, conclusions are drawn towards a biomimetic fJ/conv NASP for artificial cochlea neuromorphic sensing in Sect. 5.

## 2 Background

In this section, literature is revised. Section 2.1 describes neuromorphic modeling from the biological point of view, introducing Izhikevich's neuron model. Section 2.2 presents a short state-of-the-art review from analog and mixed-signal neuromorphic systems used in literature comparison of Sect. 4. For readers wishing to go further, an interesting survey is available in [1]. Section 2.3 recalls important model equations for transistors biased in weak inversion used in this work during circuit design [22].

### 2.1 Cortex neurons

Cortex neurons are involved in higher functions such as sensory perception, generation of motor commands, spatial



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

reasoning, conscious thought, and human's language. Most biologists agree with the classification of cortex neurons in six most fundamental classes of firing patterns observed in the mammalian neocortex [14]. These cortex neurons are:

| | |
|---|---|
| RS | Regular spiking neurons fire tonic spikes with adapting frequency in response to injected pulses of DC current |
| 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 and Huxley [20] or the Moris and Lecar model [21]. Both most accepted models behave as non-linear leak integration and fire by exponential functions of real numbers. In CMOS technology, such behavior is obtained in weak inversion (subthreshold regime) [22]. Besides, transistors operating in such regime also take advantage of ultra-low-power consumption by a supply voltage around $\pm V_d$ (i.e. $\pm 100$ mV). 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, eNeuron library proposition of Izhikevich's models has caught attention in the scientific community. Rangan et al. [16] have proposed a current-mirror-based schematic capable to mimic all classes of firing patterns. Rangan's proposal achieves a silicon area of 2980 $\mu m^2$ 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, Schaik et al. [17] have proposed the Izhikevich neuron model implementation using two first-order log-domain low-pass filters and two trans-linear multipliers. 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 model's behavior.

Later then, a variety of eNeuron implementations were presented in [23]. Most implementations have a smaller silicon area than Rangan et al. [16] and Schaik et al. [17] (around 1000 $\mu m^2$) 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 $\mu m^2$, $f_{spike} = 10$ Hz, and $E_{eff} = 165$ pJ/spike using AMS 0.35 µm process [24].

Recently, Wu et al. [25] have included the learning capability based on resistive synapses in their spiking neuron proposal. Using class AB-OTAs, Wu's proposal presented a silicon area of $110 \times 110$ $\mu m^2$ and $E_{eff} = 9.3$ pJ/spike using 0.18 µm CMOS process. Azghadi et al. [26] have improved the learning capabilities using memristor-based synapses. Using a non-conventional technology node, their work presented a silicon area of 600 $\mu m^2$. Azghadi et al. affirmed that power consumption was not the design goal. Both have neither detailed $f_{spike}$ range, nor Izhikevich's models mathematical behavior (i.e. exponential functions of real numbers) which requires transistors operating in weak inversion.

Sourikopoulos et al. [18] has innovated by a biomimetic and a simplified version of FS eNeurons. Discussing the trade-off between speed and consumption, biomimetic version achieves an area of 300 $\mu m^2$, $f_{spike} = 1$ kHz, and an $E_{eff} = 40$ fJ/spike; simplified version obtained an area of 35 $\mu m^2$, $f_{spike} = 26$ kHz, and an $E_{eff} = 4$ fJ/spike. Both circuits are implemented using TSMC 65 nm technology and considered only dynamic power consumption in the figure of merit. Sourikopoulos' 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, Zhang et al. [27] have made an effort in defining a PVT-robust eNeurons. A lower $C_m = 1$ fF is presented to address area and reliability challenges. Eventually, eNeurons operates at 1 V supply, extremely high $f_{spike}$, but power consumption is not presented. 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 $\mu m^2$ using 65 nm CMOS technology.

In the previous paper [13], the authors have redesigned Sourikopoulos' eNeuron in a higher firing rate ($f_{spike} \geq 200$ kHz) to reduce the silicon area and highlight a better energy efficiency trade-off. Besides, an innovative schematic is proposed to establish an eNeuron library based on an Izhikevich's neural firing patterns model. Presented results suggest that the time to rebound spikes is dependent of the excitation current amplitude. Both eNeurons have



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

presented a 2.3 and 3.6 fF/spike energy efficiency and a smaller silicon area 357 and 515 μm$^2$ for respectively FS and LTS implementations.

### 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 [22]:

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

where $\phi_t$ is the thermal voltage $kT/q$; $\eta$ is the slope factor $1 + C_d/C_{ox}$ (i.e. depletion $C_d$ and oxide $C_{ox}$ capacitance ratio); and $I_s$ is the specific current. Considering a small variation of $V_{DS}$, then (1) can be approximated by

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

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

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

An interesting result presented by Sourikopoulos et al. [18] is the association of a PMOS ($MP_i$) and a NMOS ($MN_i$) transistors biased in WI, having the drain node connected as $V_{out}$, gate node as $V_{in}$, and source node bias as $\pm V_d$. This circuit schematic is in fact a well-known inverter-based OTA, biased in weak inversion. When applying Kirchhoff's Current Law (KCL) at the transistors' drain node of the OTA, expressing the currents as in (2), assuming the same $\eta$ for both transistors; the output voltage of the transconductor in static WI regimecan 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 $\frac{MP_i}{MN_i}$ is negligible. In the circuit schematics studied in this paper, the output current is in fact the gate leakage of a following transistors' gates (PMOS and NMOS) playing the role of a load for the inverter-based OTA. Details in the inverter-based OTA circuit analysis demonstrating (4) are presented in [18].

State-of-the-art biomimetic-hardware uses capacitors 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 eNeuron is a current-controlled oscillator having $I_{ex}$ as the control current and $V_m$ as the spiking output node. This is obtained from positive feedback association of capacitors and transconductors.

## 3 Neuromorphic analog spiking processor

In this section, the proposed neuromorphic analog spiking processor (NASP) is described. Section 3.1 details the system-level specifications in terms of dynamic range (70 mV), resolution (9 bits), and energy consumption through leakage current minimization. Sections 3.2 and 3.3 present the FS and LTS eNeurons design from transistor-level point of view. Section 3.4 presents in detail the NASP physical design and the leakage current minimization from layout-level point of view.

### 3.1 System-level NASP proposal

The novel NASP is illustrated in Fig. 1(a); it has a sample and hold building block, two current-mirror synapse implementation and both FS and LTS eNeurons. In this paper, the sample is assumed ideal ($r_{on} \to 0$) for simplicity and $C_{in}$ is chosen high enough to a negligible parasitic capacitance (total $C_{GS}$ of $MP_{in}$, $MP_{FS}$, and $MP_{LTS}$). In fact, the hold capacitor $C_{in}$ should be designed considering the sample and hold speed performance. Indeed, the short-circuit resistance of the sampler ($r_{on}$) and $C_{in}$ play an important role in the sample-and-hold speed performance. A circuit design technique that minimizes the switch $r_{on}$ variation in the presence of large input ($V_{in}$) and output voltage swings ($V_m$) is detailed in [29]. Considering this technique, the aforementioned assumption is thus valid.

Bartolozzi and Indiveri have been one of the first to propose a silicon synaptic circuit in [28]. Synapses emulate the function of biological synaptic connections. They modulate the incoming events from the axon array with synaptic weights and generate destination post-synaptic signals addressed of the following neurons. Silicon synapses are designed using a pulsed current-source circuit activated by a presynaptic $V_{GS}$ signal. Complex synapse blocks are an array implementation of current mirrors [19], where current-mirror gains represent the synaptic weights.

The excitation current $I_{ex}$ of FS eNeuron is obtained through $MP_{in}$-$MP_{FS}$, and $I_{ex}$ of LTS eNeuron is obtained through $MP_{in}$–$MP_{LTS}$ current mirrors. The spiking modulator generates two spiking signals in $V_{m,FS}$ and $V_{m,LTS}$ outputs having a $f_{spike}$ proportional to excitation. The $f_{spike}$ can be estimated in a window of observation bounded in a time of $1/f_s$, where $f_s$ is the sampling frequency. Following



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

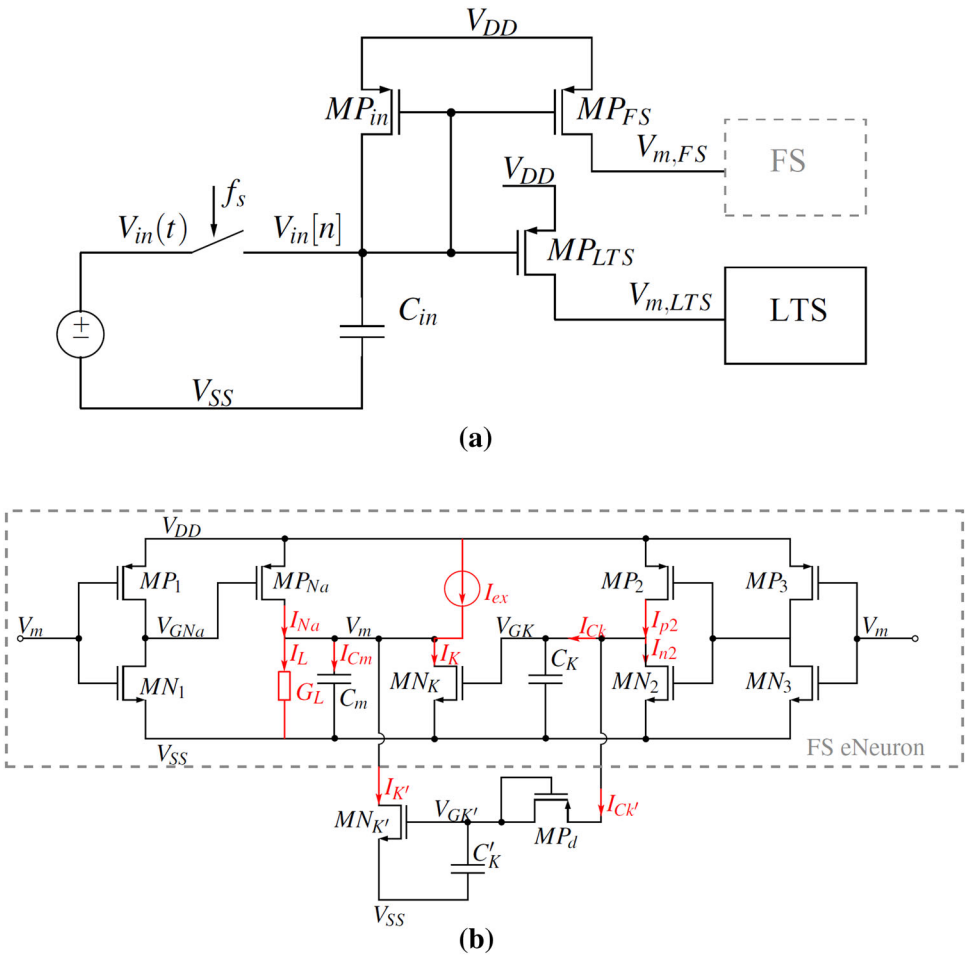


**Fig. 1** Neuromorphic Analog Spiking Processor proposal: **a** spiking-modulator architecture, and **b** the LTS e-Neuron schematic having the FS e-Neuron core highlighted in the gray box ($V_{DD} = V_d$ and $V_{SS} = -V_d$)

this constraint, the NASP resolution in number of bits could be defined as

$$N(\text{bits}) = \log_2 \frac{f_{spike,max}}{f_s}, \tag{5}$$

where $f_{spike,max}$ is the full-scale maximum $f_{spike}$. Thus, $N(\text{bits})$ depicts the resolution and speed trade-off of the modulator. Indeed, one might increase the NASP resolution by setting a longer holding time by decreasing $f_s$.

Current mirror gains in synapse implementation are weighted to maximize the full-scale dynamic range in terms of $f_{spike,max}$, which is similar to an offline spiking neural network learning process. Current mirror design focused on minimum leakage, and silicon area for $MP_{in}$. Following the results of a previous paper if a supply voltage of $V_{DD} = V_d$ and $V_{SS} = -V_d$ is chosen ($V_d = 100$ mV) the minimum $V_m$ of an eNeuron is $-70$ mV [13] mimicking the biological membrane rest potential [14]. Thus, $MP_{FS}$ and $MP_{LTS}$ sizing is obtained for an optimal $I_{ex}$ to $f_{spike,max}$, aiming a full-scale input dynamic range of $-70\text{ mV} \leq V_{in}[n] \leq 70\text{ mV}$. However, different strategies could be considered in the learning process to favor other NASP performance. This paper points out the synaptic weights for a maximum resolution in terms of $N(\text{bits}) \approx 9$, see Sect. 4 for details.

A two-neuron spiking neural network completes the proposed NASP system-level, having a FS and an LTS eNeuron. FS eNeuron generates high-frequency tonic spikes $f_{spike}$ proportional to the sampled $V_{in}[n]$ at $V_{m,FS}$ output. LTS eNeuron generates a pronounced spike frequency adaptation (decreasing) and rebound spikes due to post-inhibitory effect proportional to the sampled $V_{in}[n-1]$ at $V_{m,LTS}$ output. FS eNeuron schematic is depicted in the Fig. 1(b) highlighted gray-dashed box [18]. Both $MP_{Na}$ and $MP_K$ transistors mimic 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. LTS eNeuron, as shown in the Fig. 1(b) [13], integrates and fires tonic spikes similarly to FS eNeuron when $I_{ex} \geq I_L$. To mimic 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 eNeurons considering transistors in weak inversion [13].



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

### 3.2 FS eNeuron analysis

Sourikopoulos et al. have first demonstrated the fast-spiking behavior of biomimetic neuron with relatively constant period in [18]. Later, Sourikopoulos et al. have demonstrated the stochastic resonance behavior of such a neuron [30]. According to [14], both characteristics are presented only for FS neurons, and thus Sourikopoulos' work was named as FS eNeuron in [13].

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

$$\begin{aligned} I_{Cm} &= I_{Na} - I_K + I_{ex} - I_L, \\ I_{Ck} &= I_{P2} - I_{N2}. \end{aligned} \tag{6}$$

Using (2) and (4); assuming transconductors are in static WI regime, $V_{DD} = V_d$ and $V_{SS} = -V_d$; the (6) system leads to the following coupling system between $V_m$ and $V_{GK}$

$$\begin{aligned} C_m \cdot \frac{dV_m}{dt} = G_{Na} \cdot e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N1}}{G_{P1}}\right)\right)\right)}{\eta\phi_t}} \cdot (V_d - V_m) - G_K \\ \cdot e^{\frac{V_{GK}+V_d}{\eta\phi_t}} \cdot (V_m + V_d) + I_{ex} - G_L \cdot V_m, \end{aligned} \tag{7a}$$

$$\begin{aligned} C_K \cdot \frac{dV_{GK}}{dt} = G_{P2} \cdot e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N3}}{G_{P3}}\right)\right)\right)}{\eta\phi_t}} \cdot (V_d - V_{GK}) \\ - G_{N2} \cdot e^{\frac{V_d\left(1-\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N3}}{G_{P3}}\right)\right)\right)}{\eta\phi_t}} \cdot (V_d + V_{GK}); \end{aligned} \tag{7b}$$

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;
- $G_L, G_{Na}, G_K$ are the excitation and leak currents;
- $G_{Pi}, G_{Ni} \forall i \in \{1, 2, 3\}$ are the $i$-transistor conductances.

Thus, (7) is a time-invariant, first-order, non-linear system connecting the $V_m$ and the $V_{GK}$ potentials. It depends on ten parameters $I_{ex}, G_L, G_{Na}, G_K, (G_{Pi}, G_{Ni}) \forall i \in \{1, 2, 3\}$ when both temperature ($\phi_t = 26$ mV) and action potential ($V_d = 100$ mV) are fixed. A circuit designer would probably like to solve (7) system using MatLab, for instance. However, neither analytical solution nor any analytical equilibrium point can be found for such a system. In fact, when considering the exponential argument $V_d/\eta\phi_t \approx 5$ and noticing that experimental data of $V_m$ can have the same order of magnitude as $\eta\phi_t$ [18], one may conclude that the linearization of $\tanh(x)$ and $e^x$ functions have no physical meaning for this non-linear circuit.

### 3.3 LTS eNeuron analysis

To mimic 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 eNeuron, the (6) system turns into

$$\begin{aligned} I_{Cm} &= I_{Na} - I_K - I_{K'} + I_{ex} - I_L, \\ I_{Ck} - I_{Ck'} &= I_{P2} - I_{N2}. \end{aligned} \tag{8}$$

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

$$I_{Ck'} = C_{K'} \cdot \frac{dV_{GK'}}{dt} \approx e^{\frac{1}{2\eta}} \cdot G_{Pd}(V_{GK} - V_{GK'}), \tag{9}$$

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

$$I_{K'} = G_{K'} e^{\frac{V_{GK'}+V_d}{\eta\phi_t}} (V_m + V_d) \tag{10}$$

By expressing the (8) system with the parameters of the circuit from (9) and (10), one gets (11) like the FS eNeuron solution. Since $V_{GK'}$ is dependent on time, the (11) is a time-variant, first-order, non-linear system relating the $V_m$ and the $V_{GK}$ potentials. It depends on thirteen parameters

$$I_{ex}, G_L, G_{Na}, G_K, (G_{Pi}, G_{Ni}) \forall i \in \{1, 2, 3\}, G_{Pd}, C_K, C_{K'}$$

when both temperature ($\phi_t = 26$ 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.

$$\begin{aligned} C_m \cdot \frac{dV_m}{dt} = G_{Na} \cdot e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N1}}{G_{P1}}\right)\right)\right)}{\eta\phi_t}} \cdot (V_d - V_m) \\ - G_K \cdot e^{\frac{V_{GK}+V_d}{\eta\phi_t}} \cdot (V_m + V_d) \\ - G_{K'} \cdot e^{\frac{V_{GK'}+V_d}{\eta\phi_t}} \cdot (V_m + V_d) + I_{ex} \\ - G_L \cdot V_m, \end{aligned} \tag{11a}$$



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

$$C_K \cdot \frac{dV_{GK}}{dt} = G_{P2} \cdot e^{\frac{V_d\left(1+\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N3}}{G_{P3}}\right)\right)\right)}{\eta\phi_t}} \cdot (V_d - V_{GK}) - G_{N2} \cdot e^{\frac{V_d\left(1-\tanh\left(\frac{V_m}{\eta\phi_t}+\frac{1}{2}\ln\left(\frac{G_{N3}}{G_{P3}}\right)\right)\right)}{\eta\phi_t}} \cdot (V_d + V_{GK}) - C_{K'} \cdot \frac{dV_{GK'}}{dt}. \tag{11b}$$

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 the firing phase. Therefore, one may solve (9) verifying that the discrete time ($t = n/f_s$) $V_{GK'}[n]$ depends on previous sample of $I_{ex}[n-1]$ according to $C_{K'}$ charge and discharge cycles. The delay characteristic introduced by the additional branch could be given by a time constant

$$\tau = \frac{C_{K'}}{e^{\frac{1}{2\eta}} \cdot G_{Pd}} \tag{12}$$

Consequently, the settling time (i.e. $5\tau$) of $MP_d$-$C_{K'}$ circuit is the time required to the frequency adaptation (decrease), assuming a discharged $C_{K'}$, i.e. $I_{ex}[n-1] \leq I_{ex}[n]$. Assuming a charged $C_{K'}$ ($I_{ex}[n-1] \geq I_{ex}[n]$), the settling time of $MP_d$-$C_{K'}$ circuit is the time required to rebound spikes due to post-inhibitory effect (does not spike) in LTS eNeuron. To conclude, the voltage value charged in $C_{K'}$ actually mimic the memory effect in LTS neurons from previous excitation $I_{ex}[n-1]$, depending on the $C_{K'}$ value and its ability to discharge itself throughout time. The LTS eNeuron $f_{spike}$ therefore depends on $I_{ex}[n-1]$ and $I_{ex}[n]$ values:

- *High* $f_{spike}$ is followed by a frequency adaptation (decrease) is observed if $I_{ex}[n-1] \leq I_{ex}[n]$;
- *Low* $f_{spike}$ according to rebound spikes of post-inhibitory effect is observed if $I_{ex}[n-1] \geq I_{ex}[n]$.

### 3.4 NASP physical design

Sourikopoulos et al. have proposed a first sizing of FS eNeuron in 65 nm technology node [18]. The silicon area is dominated by the total capacitance value, depending on the capacitance density in the technology node (i.e. 0.2 fF/μm$^2$) [30]. 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. In previous paper [13], eNeuron physical design only considers metal–insulator-metal (MIM) capacitors as proposed in [30]. In this paper, a varicap-based version is also considered for silicon area minimization, while the capacitance value ($C_m, C_K, C_{K'}$) is kept unchanged.

**Table 1** Sizing of $MN_i$ and $MP_i$ transistors in W × L (nm) and in number of cells × unity capacitance for $C_i(fF)$

| | | | |
|---|---|---|---|
| $MN_1$ | 180 × 55 | $MP_1$ | 121.5 × 55 |
| $MN_2$ | 121.5 × 55 | $MP_2$ | 1080 × 55 |
| $MN_3$ | 121.5 × 55 | $MP_3$ | 180 × 55 |
| $MN_K$ | 1800 × 55 | $MP_{Na}$ | 720 × 55 |
| $C_m$ | 9 × 0.931 | $C_K$ | 16 × 0.931 |
| $MN_{K'}$ | 1800 × 55 | $MP_{Na}$ | 450 × 9000 |
| | | $C_{K'}$ | 16 × 0.931 |
| $C_{in}$ | 1 × 197 | $MP_{in}$, | 121.5 × 55 |
| $MP_{FS}$, | 486 × 55 | $MP_{LTS}$ | 1458 × 55 |

From [18], $f_{spike}$ and $E_{eff}$ trade-off can be improved by increasing $f_{spike}$. In this case, static energy consumption is reduced when $f_{spike}$ increases. In addition, 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(b). FS eNeuron design is carried out aiming at a higher-$f_{spike}$, as well LTS eNeuron proposed in [13]. Considering a $I_{ex}$ pulse width of 1 ms, $MP_d$-$C_{K'}$ circuit is sized for a $\tau \approx 40$ ms having $G_K = G_{K'}$ for simplicity. After some design-of-experiments, Table 1 presents the final eNeuron circuit sizing. Indeed, layout constraints were considered in choosing $C_m, C_K, C_{K'}$ to improve the 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. Dummy-capacitor area is not considered in eNeuron core area depicted in Fig. 2(a) and (b).

Likewise, one may argue that transistors should be interdigitized 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(b), is not implemented in both eNeurons, 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. Thus, it will be required a higher $I_{ex}$ in the integrating phase if the inter-digitization layout technique is used. Besides, it is expected that LTS eNeuron requires a higher $I_{ex}$ due to the additional area of the $MP_d$ depletion region.

The MIM-based eNeuron circuit layout is shown in Fig. 2(a) for FS, and in Fig. 2(b) for LTS. The area of FS eNeuron is 18.9 × 18.9 μm$^2$. The area of LTS eNeuron is 22.7 × 22.7 μm$^2$. One might notice that capacitors $C_m, C_K, C_{K'}$ take most of the layout area. By increasing



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

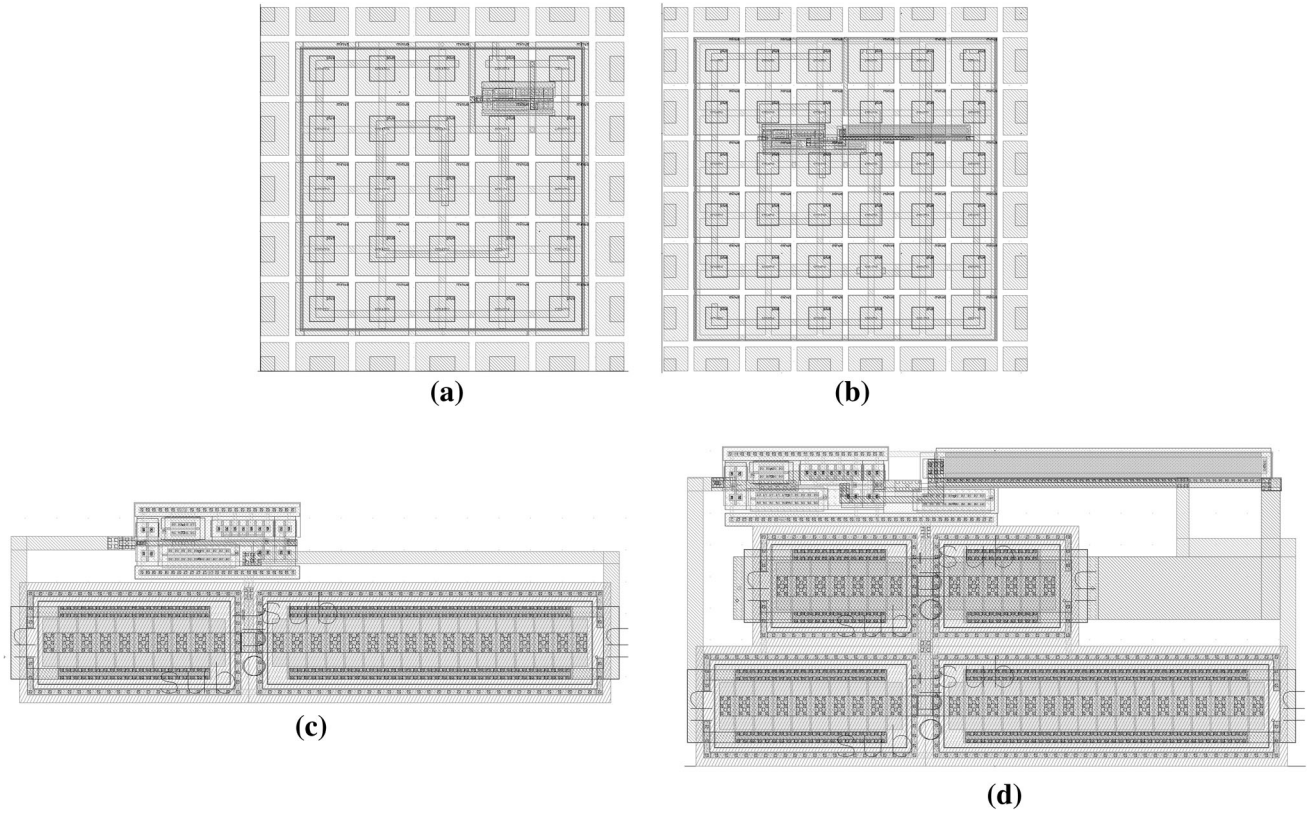


**Fig. 2** eNeuron physical design: **a** FS eNeuron having 18.9 × 18.9 μm$^2$, **b** LTS eNeuron having 22.7 × 22.7 μm$^2$, **c** varicap-based FS eNeuron having 17.3 × 5.7 μm$^2$, **d** varicap-based LTS eNeuron having 17.3 × 9.1 μm$^2$

$f_{spike}$, the previous work [13] heads toward a large-scale integration challenge using smaller capacitors. The area of designed eNeurons are one order of magnitude smaller than [23], and slightly bigger than reported in [18].

While MIM-based eNeurons are favored by a null leakage in $C_m, C_K, C_{K'}$, this is hardly true for both varicap-based eNeuron. Since the bottom plate is physically connected to the substrate, it becomes intrinsically subjected to leakage. To circumvent this issue, the varicap bottom plate is wisely chosen to be connected to $V_{SS}$, which is the same substrate potential. By this strategy, one may consider a null leakage in $C_m, C_K, C_{K'}$ for varicap-based eNeuron. PLS results in Sect. 4 shall demonstrate such assumption. The varicap-based eNeuron circuit layout is shown in Fig. 2(c) for FS, and in Fig. 2(d) for LTS. The area of FS eNeuron is 17.3 × 5.7 μm$^2$. The area of LTS eNeuron is 17.3 × 9.1 μm$^2$. As expected, both varicap-based eNeuron depicted in Fig. 2(c) and (d) present a smaller silicon area three times smaller to [18] and [13] implementations. Varicap-based eNeuron is also a very competitive solution for integration and firing eNeuron implementations in the state-of-the-art [5].

Still, temperature variation represents a major limitation for transistors operating in WI regime. Studies have proved that leakage current is exponentially dependent on temperature; it doubles for every 10 °C increase in temperature [31]. Thus, $I_L$ shall increase over-temperature enabling eNeuron firing in a higher $I_{ex,min}$, and increasing static energy consumption. Zhang et al. have suggested a solution for temperature sensitivity in eNeurons [27]. State-of-the-art artificial neurons are often considered in a fixed temperature point [5]. Temperature variation is beyond the scope of this research and all PLS results consider $T = 27$ °C. Future studies should consider analog eNeuron reliability limitations to cope with digital eNeuron implementations [5].

Reduce the static energy consumption is the key to improve $E_{eff}$ even further and to enable eNeuron 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 the described phenomenon. To the best of our knowledge, these layout considerations for WI transistors are first presented in [13]. Besides, a varicap-based implementation dealing with substrate leakage and silicon area reduction is first presented in this paper.

According to PLS results of eNeurons, leakage current of MIM- and varicap-based solutions are comparable and as low as 22 pA for FS and 132 pA for LTS eNeuron. For



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

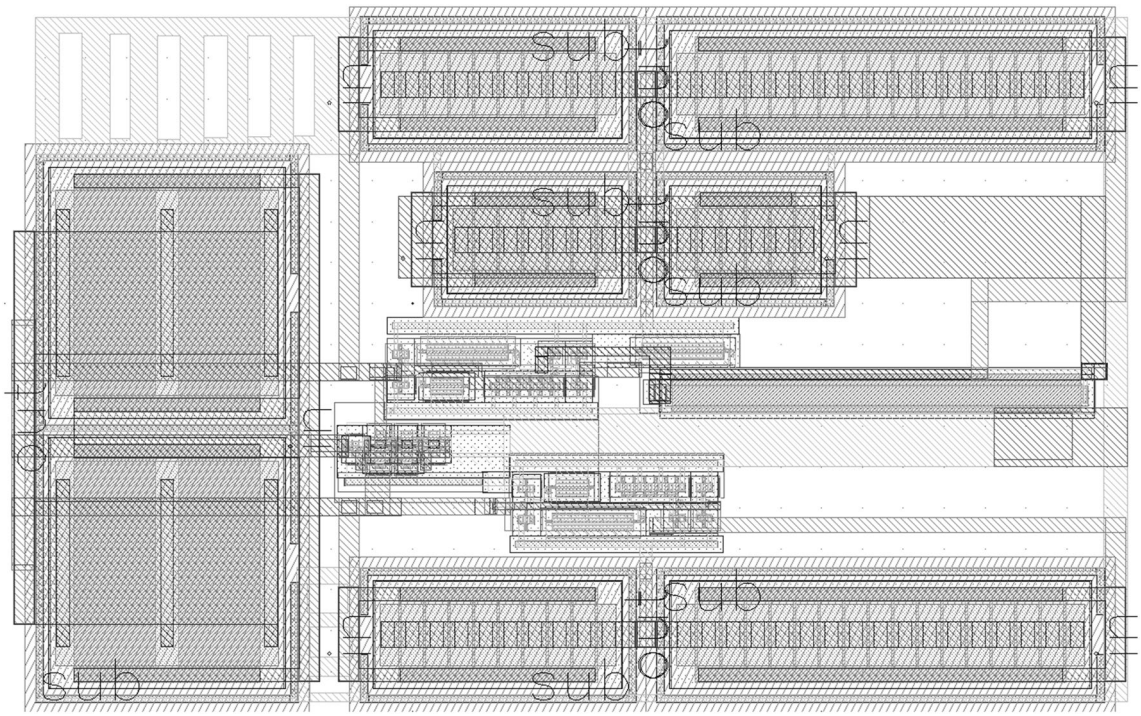


**Fig. 3** NASP physical implementation using the varicap-based eNeurons FS and LTS, having 24.5 × 15.5 μm$^2$

this reason, varicap-based eNeuron is the adopted solution for NASP physical design. To maximize the full-scale dynamic range, the $MP_{FS}$ aspect ratio is four times the $MP_{in}$; the $MP_{LTS}$ aspect ratio is twelve times the $MP_{in}$. Table 1 presents the NASP circuit sizing and a compact layout is then obtained. Figure 3 illustrated the NASP physical implementation having an area of 24.5 × 15.5 μm$^2$.

## 4 Post-layout simulation results

In this section, post-layout simulations (PLS) are carried out to demonstrate the spiking modulator characteristics for audio signal processing. Section 4.1 presents both eNeuron PLS results and a state-of-the-art comparison. Section 4.2 presents PLS results for NASP operation considering: (a) figure of merits of the spiking modulator; (b) a reduced input-signal dynamic (±10 mV) for $f_{in} = 440$ Hz is sampled at $f_s = 1$ kHz and $f_{in} = 20$ kHz is sampled at $f_s = 44$ kHz to highlight audio bandwidth signal; and (c) a large input-signal dynamic(±70 mV) for a fixed audio frequency of $f_{in} = 440$ Hz, and $f_s = 1$ kHz.

### 4.1 eNeuron simulations

The eNeuron PLS results are carried out to validate the biomimetic behavior by a fixed pulse width of 5 ms and an increasing $I_{ex}$ amplitude from a few pico-amps. When $I_{ex} \geq I_L$, the leak integration is negligible and eNeuron fires at $f_{spike}$ after that point. MIM- and varicap-based eNeuron PLS results reveal a current $I_L$ of 22 pA for FS and of 132 pA for LTS. However, varicap-based eNeuron presents a slightly higher frequency $f_{spike}$ compared to MIM-based implementation. Such differences could be explained by its compact layout which leads to a reduction in parasitic capacitances.

Figure 4 illustrates the $f_{spike}$ pattern for both MIM-based eNeurons excited by a current pulse of 30 pA for FS and of 150 pA for LTS. From Fig. 4(a), a tonic $f_{spike} \approx 20$ kHz for FS eNeuron is observed. As expected, the $f_{spike}$ of LTS eNeuron starts at 20 kHz, but it decreases over time, achieving a final value of 5 kHz after 0.2 ms, see Fig. 4(b) for details. Varicap-based eNeuron achieved tonic a $f_{spike} \approx$ 25.2 kHz when excited by a current pulse of 30 pA for FS and of 150 pA for LTS. Likewise, $f_{spike}$ of varicap-based LTS eNeuron decreases over time, achieving a final value of 5.8 kHz after 0.2 ms. Transient simulation results for both varicap-based eNeurons demonstrate a similar behavior as observed in Fig. 4.

To demonstrate the spike frequency adaptation and the rebound spikes due to the post-inhibitory effect, a piecewise linear $I_{ex}$ pulse is applied to all eNeurons. Figure 5(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



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

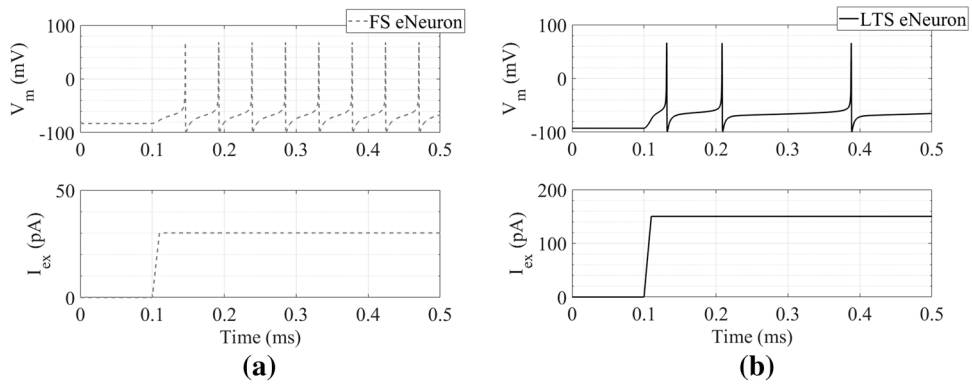


**Fig. 4** Post-layout transient simulation results for both MIM-based eNeurons excited by a current pulse, being: **a** a $I_{ex}$ of 30 pA for FS and **b** of 150 pA for LTS. A tonic $f_{spike} \approx 20$ kHz is observed for both eNeurons, and spike frequency adaptation is achieved for LTS eNeuron after 0.2 ms. Similar results are found for varicap-based eNeurons

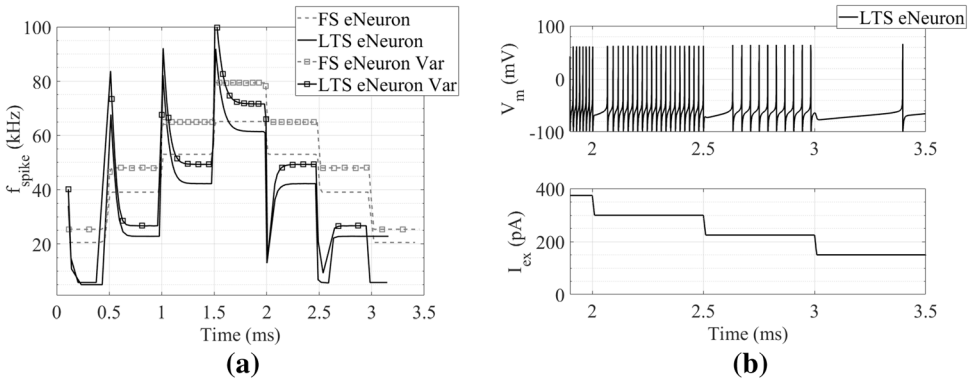


**Fig. 5** Post-layout simulations demonstrate the spike frequency adaptation and the rebound spikes due to post-inhibitory effect. **a** The $f_{spike}$ response over time for piecewise linear $I_{ex}$ pulses [30 pA, 45 pA, 60 pA, 75 pA, 60 pA, 45 pA, 30 pA] for FS (dashed-gray line) and [150 pA, 225 pA, 300 pA, 375 pA, 300 pA, 225 pA, 150 pA] for LTS eNeuron (continuous black line). Varicap-based eNeurons are square-marked in (**a**). **b** The $V_m$ spiking behavior of LTS eNeuron during $I_{ex}$ decreasing pulse. Similar behavior is identified for varicap-based LTS eNeuron

pA, 225 pA, 300 pA, 375 pA] for LTS eNeuron. $I_{ex}$ amplitude range under study is considered in increasing (from 0 to 1.7 ms) and decreasing (from 1.7 to 3.5 ms) manners. As expected, Fig. 5(a) highlights an increasing $f_{spike}$ as $I_{ex}$ increases. FS eNeuron, in dashed-gray line, shows tonic spikes with a relatively constant period validating the Izhikevich's model. LTS eNeuron, in continuous black line, highlights a frequency adaptation (decreasing) among amplitudes from 0 to 1.7 ms. However, rebound spikes due to the post-inhibitory effect are observed among regions from 1.7 to 3.5 ms.

MIM-based eNeuron PLS results were previously presented in [13]. Varicap-based eNeuron PLS results are illustrated in square marked lines in Fig. 5(a). One may observe the $f_{spike}$ increase for both eNeurons FS and LTS. Nevertheless, FS tonic spikes shape remains unchanged, as LTS spike frequency adaptation and the rebound spikes due to the post-inhibitory effect. Using varicap instead of MIM capacitance does not change the circuit behavior because the voltage variation ($\Delta V_m \leq 200$ mV) is negligible if compared to the required bias voltage changes ($V_C \approx 2$ V) according to measurement results [32]. The varicap of eNeuron circuit remains biased near the membrane rest potential ($\sim -70$ mV), and under spike variations, the varicap filters its high-frequency components (short to $V_{SS}$).

Figure 5(b) illustrated the $V_m$ spiking behavior of MIM-based LTS eNeuron during $I_{ex}$ amplitude decreasing from 1.7 to 3.5 ms. Indeed, LTS eNeuron does not spike due to the post-inhibitory effect but starts firing at a low-$f_{spike}$ consistent to value achieved after adaptation. Interesting results are revealed in Fig. 5(b), suggesting that the time to rebound spikes is actually dependent to $I_{ex}$ pulse variation. The solution of (9) considering weak inversion is in fact more complicated than a simple $MP_d$-$C_{K'}$ circuit charge and discharge. Moreover, Fig. 5(b) results suggest that $G_{Pd}$ is $I_{ex}$ dependent. Similar behavior is identified for varicap-based LTS eNeuron.

Spiking eNeuron is often compared in the state-of-the-art by its spiking frequency $f_{spike}$, power consumption $P_{RMS}$ and energy efficiency $E_{eff} = P_{RMS}/f_{spike}$, where $P_{RMS}$ is the root-mean-squared power consumption including static and dynamic powers [5]. Figure 6 illustrates theses figure-of-merits of eNeurons as $I_{ex}$ increases from estimated $I_L$ to 1



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

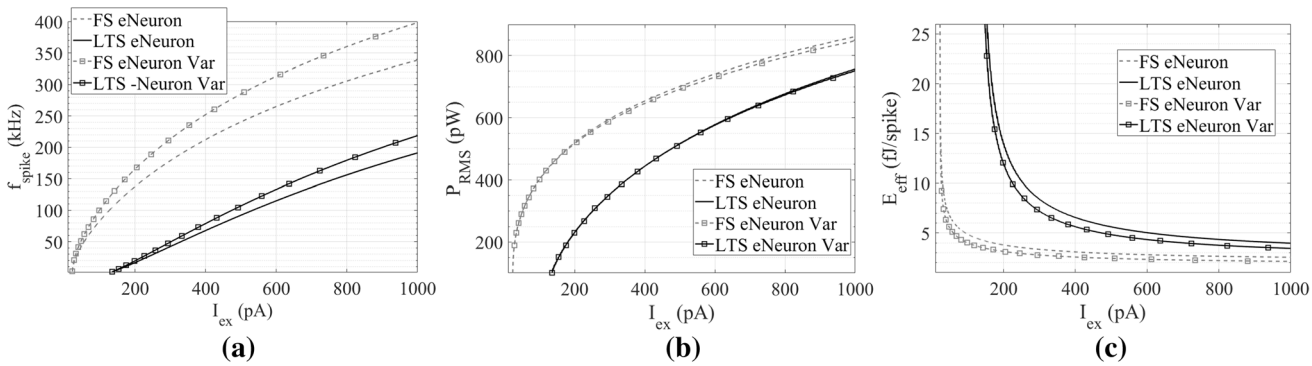


**Fig. 6** Figure-of-Merits extracted from post-layout simulation results of LTS and FS eNeurons: **a** $f_{spike}$ in kHz, **b** $P_{RMS}$ in pW, and **c** $E_{eff}$ in fJ/spike

nA using a 1001-points linear DC sweep. MIM-based implementation is shown in Fig. 6 by lines without marks; varicap-based implementation is depicted using square marks. Figure 6(a) illustrates the growth of $f_{spike}$ versus $I_{ex}$. The $f_{spike}$ of LTS eNeuron after frequency adaptation is lower than $f_{spike}$ of FS eNeuron, for both implementations. Varicap-based FS eNeuron demonstrates an average $f_{spike}$ increase of 20% and a standard deviation of 2.6%. Varicap-based LTS eNeuron demonstrates an average $f_{spike}$ increase of 14% and a standard deviation of 1.9%.

By setting a higher $f_{spike}$, a power consumption increase is expected (i.e. dynamic consumption). Figure 6(b) presents LTS and FS eNeuron power consumption in the range of few hundreds of pico-Watts. This is a slight power consumption increase compared to [18] (around 100 pW). Varicap-based implementation demonstrated an average $P_{RMS}$ decrease of 0.6% and a negligible standard variation for both FS and LTS eNeurons in comparison to MIM-based ones. These results prove that varicap-based solution achieves a higher $f_{spike}$ performance for a likewise $P_{RMS}$. Despite the overall $P_{RMS}$ slight increase, the $E_{eff}$ is slightly decreased owing to a higher $f_{spike}$ (around 200 kHz) and specific eNeuron physical design. Considering static and dynamic energy consumption, Fig. 6(c) highlights a $E_{eff}$ as low as 2.3 fJ/spike for FS and 3.6 fJ/spike for LTS eNeurons implemented with MIM capacitances [13]. Moreover, a higher $f_{spike}$ for a similar $P_{RMS}$ favors varicap-based implementations in terms of $E_{eff}$. Figure 6(c) highlights a $E_{eff}$ as low as 1.95 fJ/spike for FS and 2.83 fJ/spike for LTS eNeurons, see square-marked data.

Table 2 compares the silicon neurons literature (see Sec. 2.2 for details) with the PSL results of this proposal. The obtained area in analog solutions for artificial neurons is limited by technology capacitance density and the number of required capacitors. Previous reported MIM-based implementation found a 357 $\mu m^2$ for FS [13] likewise the results reported for biomimetic FS having 300 $\mu m^2$ in [18]. However, varicap-based solution offers a better area trade-off, being 98.61 $\mu m^2$ for FS and 157 $\mu m^2$ for LTS, if compared to the mixed solution (200 $\mu m^2$) proposed in [2] and the digital solution (20 $\mu m^2$) proposed in [4]. The $f_{spike}$ is higher than reported values for MIM- and varicap-based implementations. This design choice leads to a better trade-off in speed and power consumption. Table 2 highlights a $E_{eff}$ slightly below compared to the best (4 fJ/spike simplified FS) in [18], which considers only dynamic

**Table 2** Literature eNeuron performance comparison: MIM-based from [13] and varicap-based implementation from this work

| Ref. | [16] | [24] | [25] | [2] | [18] | [4] | [13] | This Work |
|---|---|---|---|---|---|---|---|---|
| Techn. (nm) | 90 | 350 | 180 | 130 | 65 | 28 | 55 | 55 |
| | MOS | MOS | MOS | CMOS | MOS | FD-SOI | MOS | MOS |
| Area ($\mu m^2$) | 2980 | 1100 | 10 k | 200 | 300[c] | 20 | 515[a] | 157.43[a] |
| | | | | | 35[d] | | 357[b] | 98.61[b] |
| $f_{spike}$ (Hz) | 7 k | 10 | NA | 135 k | 1 k[c] | 100 | 205 k[a] | 220 k[a] |
| | | | | | 26 k[d] | | 360 k[b] | 400 k[b] |
| $E_{eff}$ (J/spike) | 1 p | 165 p | 9.3 p | 48 p | 40 f[c] | 2.8 p | 3.6 f[a] | 2.83 f[a] |
| | | | | | 4 f[d] | | 2.3 f[b] | 1.95 f[b] |

[a]LTS, [b]FS, [c]biomimetic, [d]simplified



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

consumption. MIM- and varicap-based implementations consider both dynamic and static consumption. Such a low power consumption is obtained due to a specific physical design for minimal leakage. Previous reported MIM-based implementation found a 2.3fJ/spike for FS and 3.6fJ/spike for LTS [13]. Moreover, varicap-based solution highlights an even lower $E_{eff}$, being 1.95 fJ/spike for FS and 2.83 fJ/spike for LTS. Depicted $E_{eff}$ is among the lowest in the state-of-the-art [5, 23], while ensuring the Izhikevich's mathematical modeling in contrast to simplified neuron in [18]. Obtained $E_{eff}$ in analog solutions for artificial neurons is three orders of magnitude lower than mixed (48 pJ/spike) [2] and digital (2.8 pJ/spike) [4] solutions. High efficiency is found due to a better trade-off in speed and power consumption analog solutions.

### 4.2 NASP simulations

The NASP post-layout simulation is done using a 64-point transient noise for a fixed time window. The estimated $f_{spike}$ resolution is 1 kHz, under a holding time of 1 ms Input signal $V_{in}[n]$ is swept from −70 to 70 mV using 512-point parametric simulation. For each $V_{in}[n]$ value, the estimated $f_{spike}$ and $P_{RMS}$ are then averaged over all 64-points transient noise simulations. Figure 7(a) highlights a $f_{spike}$ variation from 542 to 39.9 kHz for FS membrane output ($V_{m,FS}$), and a variation from 522 to 1 kHz for LTS membrane output ($V_{m,LTS}$). Indeed, the synaptic connection was designed for $f_{spike,LTS} \approx f_{spike,FS} \geq 512$ kHz by current-mirror aspect ratio sizing. From (5), full-scale dynamic rage is estimated as $N \approx 9$ bits.

The total power consumption of NASP is illustrated in Fig. 7(b). The $P_{RMS}$ varies over $V_{in}[n]$ sweep from 4.2 to 0.46 nW. Energy efficiency in analog signal processing circuits is often estimated using the Walden's figure-of-merit as [33]

$$FoM = \frac{P_{RMS}}{f_s \cdot 2^N}. \tag{13}$$

Figure 7(c) illustrates $FoM$ variation over $V_{in}[n]$ sweep from 8.2 fJ/conv to 0.89 fJ/conv. This result is in the same order of the magnitude of state-of-the-art analog-to-digital $\Sigma\Delta$ converters [34]. Towards a neuromorphic computing, one could use the NAPS proposal combined to a time-to-digital converter and digital-based spiking neural networks like [4] in the perspective of a mixed-signal solution as [2], but still ensure a $FoM$ in fJ/conv range.

The signal processing capabilities of the proposed NASP are evaluated using transient PLS for 50 samples. Input signal $V_{in}(t)$ has an amplitude of 10 mV. Its input frequency $f_{in}$ is sampled at $f_s$ using an ideal sample and hold circuit at NASP input voltage $V_{in}[n]$. Results, shown in Fig. 8, have total transient simulation time of $50/f_s$. Both membrane-voltage spikes in the vicinity of 200 kHz, according to foreseen $f_{spike}$ from Fig. 7(a). Since power consumption is only $V_{in}[n]$ value dependent, the average NASP $P_{RMS} = 1.73$ nW, and considering the amplitude of 10 mV, it is in accordance with results presented in Fig. 7(b).

To explore the NASP high-resolution, the $V_{in}(t)$ with $f_{in} = 440$ Hz is sampled at $f_s = 1$ kHz and results are depicted in Fig. 8(a). Spiking frequency from $V_{m,FS}$ output is illustrated in dashed-gray line; and $V_{in}[n]$ can be estimated from the $f_{spike}$ found and the system-level transfer function. Spiking frequency from $V_{m,LTS}$ output is illustrated in continuous black line; and $V_{in}[n-1] - V_{in}[n]$ (discrete differential) sign can be estimated from the $f_{spike}$ variation over time. The frequency adaptation is presented when the sign is negative, and the post-inhibitory effect is observed when the sign is positive. Both spiking frequency behavior is presented in the beginning of the sampling in $n/f_s$.

To explore the NASP high-speed, the $V_{in}(t)$ with $f_{in} =$ 20 kHz is sampled at $f_s = 44$ kHz and results are depicted in Fig. 8(b). Spiking frequency from $V_{m,FS}$ output is illustrated in dashed-gray line; and the $f_{spike}$ found is likewise the PLS results in Fig. 8(a). Spiking frequency from $V_{m,LTS}$

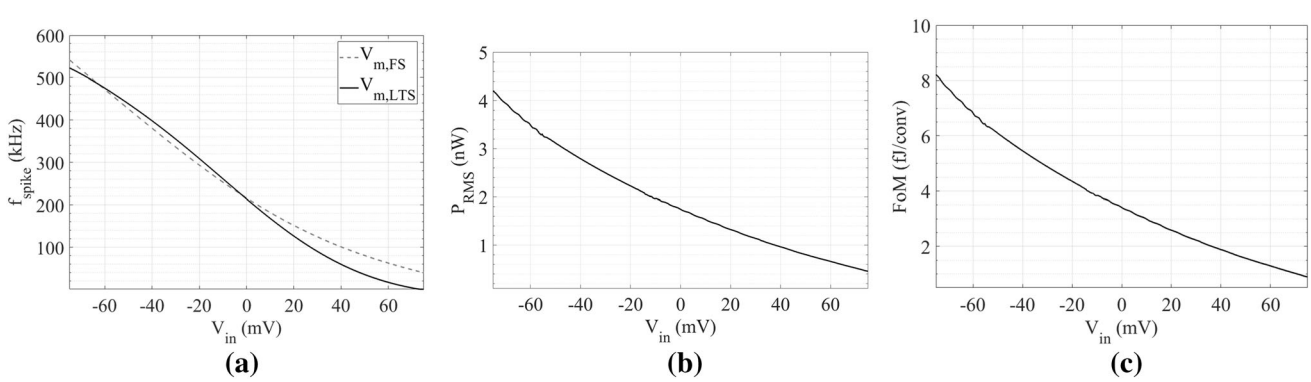


**Fig. 7** Figure-of-Merits extracted from post-layout simulation results of NASP: **a** $f_{spike}$ in kHz, **b** $P_{RMS}$ in pW, and **c** $FoM$ in fJ/conv



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

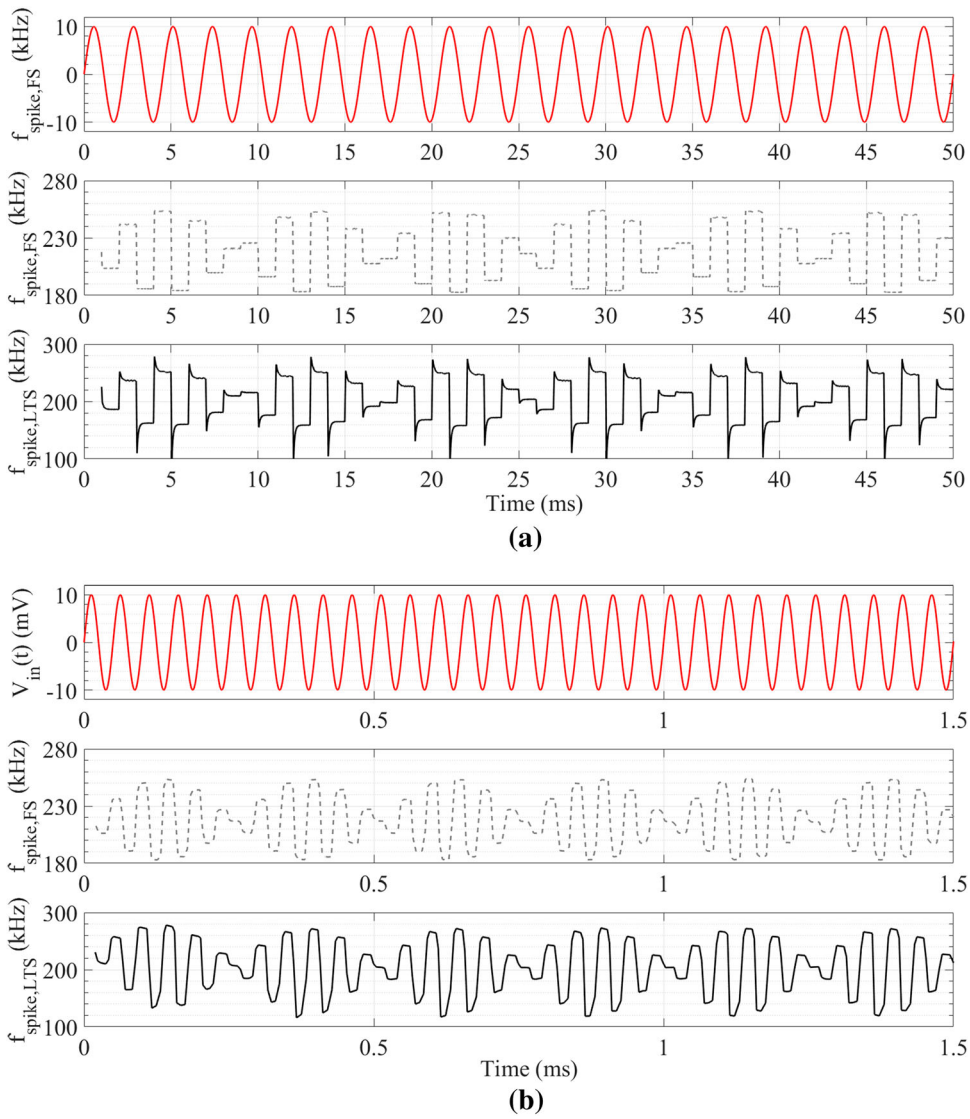


**Fig. 8** Transient PLS results of NASP membrane $f_{spike}$ being $V_{m,LTS}$ in continuous-black line and $V_{m,FS}$ in dashed-gray line for: **a** a sinusoidal 440 Hz input signal (red line) sampled at 1 kHz, **b** a sinusoidal 20 kHz input signal (red line) sampled at 44 kHz

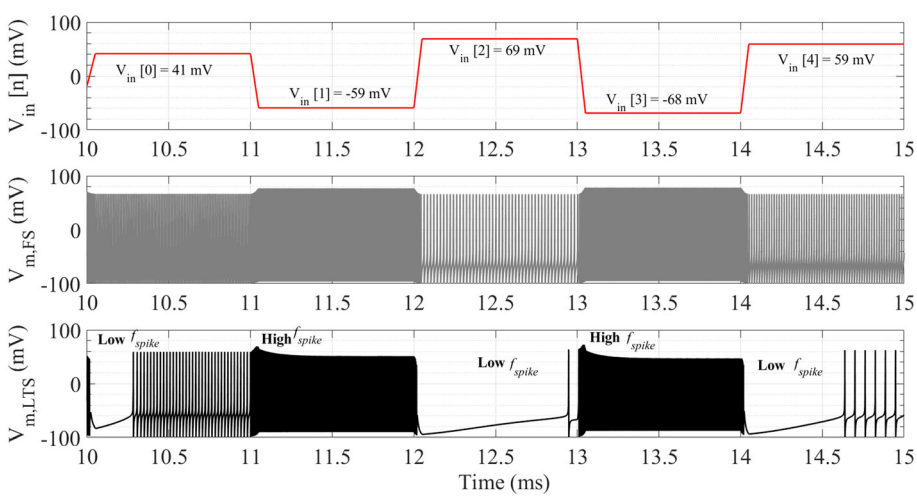


**Fig. 9** Transient PLS results of NASP membrane full-scale dynamic range as $-70\ \text{mV} \le V_{in}[n] \le 70\ \text{mV}$: being $V_{m,LTS}$ in continuous black line and $V_{m,FS}$ in dashed-gray line



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

output is illustrated in continuous black line; and $f_{spike}$ variation over time is less pronounced than Fig. 8(b). Thus, the audio-sensing is demonstrated in frequency range of an artificial cochlea application.

To investigate the NASP full-scale dynamic range, a transient PLS is considered for a $V_{in}(t)$ amplitude of 70 mV, $f_{in} = 440$ Hz, and $f_s = 1$ kHz. The estimated NASP power consumption is 2.26 nW over 50 samples. Figure 9 illustrates the PLS results from 10 to 15 ms highlighting five samples of $V_{in}[n]$ in the red-continuous line. One may estimate $V_{in}[n]$ from $V_{m,FS}$ spiking modulated signal in the gray-dashed line in Fig. 9 from Fig. 7(a). Comparing the estimated $V_{in}[n]$ to the input samples $V_{in}[n]$, it is found a root-mean-square error of 0.63 $mV_{RMS}$. One may estimate $V_{in}[n-1] - V_{in}[n]$ sign from $V_{m,LTS}$ spiking modulated signal in black-continuous line comparing them with highlighted samples in red-continuous line in Fig. 9. It is observed:

- *High* $f_{spike}$ followed by a frequency adaptation (decrease) is observed if $V_{in}[n-1] - V_{in}[n] \leq 0$;
- *Low* $f_{spike}$ according to rebound spikes of the post-inhibitory effect is observed if $V_{in}[n-1] - V_{in}[n] \geq 0$.

## 5 Conclusions

A neuromorphic analog spiking processor using Izhikevich-based artificial neurons is proposed in this paper. Using a varicap-based eNeuron implementation, the first novelty of this proposal compared to previous work is a better trade-off in area and energy efficiency. The eNeuron silicon area reduced to one-third of recent analog solutions (98.6 $\mu m^2$), but it is still five times greater than the digital solutions. Dealing with substrate leakage, energy efficiency is reduced to 1.95 fJ/spike, which is three orders of magnitude lower than mixed and digital solutions. The proposed spiking modulator for audio signal processing applications is the second novelty presented in this work. The NASP post-layout simulation results are presented to investigate the high-resolution, high-speed, and full-scale dynamic range considering audio signal bandwidth and input voltage. The proposal achieves a 9 bits spiking-modulator resolution. NASP demonstrates a maximum of 8 fJ/conv. efficiency while ensuring a root–mean–square error of 0.63 $mV_{RMS}$.

## References


1. Schuman, C. D., Potok, T. E., Patton, R. M., Birdwell, J. D., Dean, M. E., Rose, G. S., & Plank, J. S. (2017). *A survey of neuromorphic computing and neural networks in hardware*. Retrieved from http://arxiv.org/abs/1705.06963
2. Park, J., Yu, T., Joshi, S., Maier, C., & Cauwenberghs, G. (2017). Hierarchical address event routing for reconfigurable large-scale neuromorphic systems. *IEEE Transactions on Neural Networks and Learning Systems, 28*(10), 2408–2422. https://doi.org/10.1109/TNNLS.2016.2572164.
3. Liu, S. C., & Delbruck, T. (2010). Neuromorphic sensory systems. *Current Opinion in Neurobiology, 20*(3), 288–295. https://doi.org/10.1016/j.conb.2010.03.007.
4. Moradi, S., Qiao, N., Stefanini, F., & Indiveri, G. (2018). A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (DYNAPs). *IEEE Transactions on Biomedical Circuits and Systems, 12*(1), 106–122. https://doi.org/10.1109/TBCAS.2017.2759700.
5. Thakur, C. S., Molin, J. L., Cauwenberghs, G., Indiveri, G., Kumar, K., Qiao, N., et al. (2018). Large-scale neuromorphic spiking array processors: A quest to mimic the brain. *Frontiers in Neuroscience, 12,* 1–37. https://doi.org/10.3389/fnins.2018.00891.
6. Zaghloul, K. A., & Boahen, K. (2006). A silicon retina that reproduces signals in the optic nerve. *Journal of Neural Engineering, 3*(4), 257–267. https://doi.org/10.1088/1741-2560/3/4/002.
7. Wen, B., & Boahen, K. (2009). A silicon cochlea with active coupling. *IEEE Transactions on Biomedical Circuits and Systems, 3*(6), 444–455. https://doi.org/10.1109/TBCAS.2009.2027127.
8. Nebhen, J., Savary, E., Rahajandraibe, W., Dufaza, C., Meillere, S., Kussener, E., Barthelémy, H., Czarny, J., & Walther, A. (2013). Low-noise smart sensor based on silicon nanowire for MEMS resistive microphone. In *Proceedings of the IEEE Sensors* (pp. 4–7). Baltimore, MD, USA. https://doi.org/10.1109/icsens.2013.6688605
9. Nebhen, J., Ferreira, P. M., & Mansouri, S. (2020). A chopper stabilization audio instrumentation amplifier for IoT applications. *Journal of Low Power Electronics and Applications, 10*(13), 1–14. https://doi.org/10.3390/jlpea10020013
10. Xiao, F., & Weibei, D. (2016). A biologically plausible spiking model for interaural level difference processing auditory pathway in human brain. In *IEEE Proceedings of the International Joint Conference on Neural Networks* (pp. 5029–5036). Vancouver, BC, Canada. https://doi.org/10.1109/ijcnn.2016.7727862
11. Gao, X., Grayden, D. B., & McDonnell, M. D. (2017). Modeling electrode place discrimination in cochlear implant stimulation. *IEEE Transactions on Biomedical Engineering, 64*(9), 2219–2229. https://doi.org/10.1109/TBME.2016.2634461.
12. Yang, M., Liu, S.-C., Seok, M., & Enz, C. (2019). Ultra-low-power intelligent acoustic sensing using cochlea-inspired feature extraction and DNN classification. In *IEEE Proceedings of the International Conference on ASIC (ASICON)* (pp. 1–4). Chongqing, China. https://doi.org/10.1109/asicon47005.2019.8983619
13. Ferreira, P. M., Carvalho, N. De, Klisnick, G., & Benlarbi-Delai, A. (2019). Energy efficient fJ/spike LTS e-Neuron using 55-nm node. In *Proceedings of the ACM IEEE Symposium on Integrated Circuits and Systems Design* (pp. 1–6). São Paulo, SP, Brazil. https://doi.org/10.1145/3338852.3339852
14. Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transaction on Neural Networks, 14*(6), 1569–1572. https://doi.org/10.1109/TNN.2003.820440.
15. Hasler, J., & Marr, B. (2013). Finding a roadmap to achieve large neuromorphic hardware systems. *Frontiers in Neuroscience, 7,* 1–29. https://doi.org/10.3389/fnins.2013.00118.




Content courtesy of Springer Nature, terms of use apply. Rights reserved.

16. Rangan, V., Ghosh, A., Aparin, V., & Cauwenberghs, G. (2010). A subthreshold a VLSI implementation of the Izhikevich simple neuron model. In IEEE (Ed.), *International Conference Engineering in Medicine and Biology* (pp. 4164–4167). Buenos Aires, Argentina. https://doi.org/10.1109/iembs.2010.5627392
17. Schaik van A., Jin, C., McEwan, A., & Hamilton, T. J. (2010). A log-domain implementation of the Izhikevich neuron model. In *Proceedings of the IEEE International Symposium Circuits System* (pp. 4253–4256). Paris, France: IEEE. https://doi.org/10.1109/iscas.2010.5537564
18. Sourikopoulos, I., Hedayat, S., Loyez, C., Danneville, F., Hoel, V., Mercier, E., et al. (2017). A 4-fJ/spike artificial neuron in 65 nm CMOS technology. *Frontiers in Neuroscience, 11*(123), 1–14. https://doi.org/10.3389/fnins.2017.00123.
19. Qiao, N., & Indiveri, G. (2017). Analog circuits for mixed-signal neuromorphic computing architectures in 28 nm FD-SOI technology. In *IEEE SOI-3D-Subthreshold Microelectronics Unified Conf.* (pp. 1–4). Burlingame, CA, USA. https://doi.org/10.1109/s3s.2017.8309203
20. Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology, 117*(4), 500–544.
21. Morris, C., & Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal, 35*(1), 193–213. https://doi.org/10.1016/S0006-3495(81)84782-0.
22. Galup, C., & Schneider, M. (2018). The compact all-region MOSFET model: theory and applications. In *IEEE Proceedings of the International New Circuits and Systems Conference* (pp. 166–169). Montreal, QC, Canada. https://doi.org/10.1109/NEWCAS.2018.8585657
23. Indiveri, G., Linares-Barranco, B., Hamilton, T. J., van Schaik, A., Etienne-Cummings, R., Delbruck, T., et al. (2011). Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience, 5*(May), 1–23. https://doi.org/10.3389/fnins.2011.00073.
24. Demirkol, A. S., & Ozoguz, S. (2011). A low-power VLSI implementation of the Izhikevich neuron model. In *IEEE New Circuits and Systems Conference* (pp. 169–172). Bordeaux, France. https://doi.org/10.1109/newcas.2011.5981282
25. Wu, X., Saxena, V., Zhu, K., & Balagopal, S. (2015). A CMOS spiking neuron for brain-inspired neural networks with resistive synapses and in situ learning. *IEEE Transactions on Circuits and Systems II: Express Briefs, 62*(11), 1088–1092. https://doi.org/10.1109/TCSII.2015.2456372.
26. Azghadi, M. R., Linares-Barranco, B., Abbott, D., & Leong, P. H. W. (2017). A hybrid CMOS-memristor neuromorphic synapse. *IEEE Transactions on Biomedical Circuits and Systems, 11*(2), 434–445. https://doi.org/10.1109/TBCAS.2016.2618351.
27. Zhang, L., Zhang, J., Wang, Y., Zhang, X., & Huang, R. (2018). A multi-mode silicon neuron circuit with high robustness against PVT variation. In *Proceedings of the IEEE International Symposium on Circuits and Systems* (pp. 1–4). Florence, Italy. https://doi.org/10.1109/iscas.2018.8351313
28. Bartolozzi, C., & Indiveri, G. (2007). Synaptic dynamics in analog VLSI. *Neural Computation, 19*(10), 2581–2603. https://doi.org/10.1162/neco.2007.19.10.2581.
29. Razavi, B. (2015). The bootstrapped switch. *IEEE Solid-State Circuits Magazine, 7*(3), 12–15. https://doi.org/10.1109/MSSC.2015.2449714.
30. Danneville, F., Sourikopoulos, I., Hedayat, S., Loyez, C., Hoel, V., & Cappy, A. (2017). Ultra-low power analog design and technology for artificial neurons. In *Proc IEEE Bipolar/BiCMOS Circuits and Technology Meeting* (pp. 1–8). Florida, USA: IEEE. https://doi.org/10.1109/bctm.2017.8112899
31. Sze, S. M. (1981). *Physics of semiconductor devices* (2nd ed.). New York: Wiley.
32. Ferreira, P. M., Donche, C., Avignon-Meseldzija, E., Quémerais, T., Gianesello, F., Gloria, D., et al. (2018). Fully integrated interferometry-based reflectometer for high-impedance instrumentation. *IEEE Transactions on Microwave Theory and Techniques, 66*(8), 3901–3908. https://doi.org/10.1109/TMTT.2018.2831699.
33. Walden, R. H. (1999). Analog-to-digital converter survey and analysis. *IEEE Journal on Selected Areas in Communications, 17*(4), 539–550. https://doi.org/10.1109/9780470546444.ch2.
34. Muntal, P. L., Harald, I., & Jørgensen, H. (2018). 34.3 fJ/conv.-step 8-MHz bandwidth fourth-order pseudo-differential ring-amplifier-based continuous-time delta—sigma ADC in 65 nm. *IEEE Solid-State Circuits Lett., 1*(10), 198–201. https://doi.org/10.1109/LSSC.2019.2910468

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Pietro M. Ferreira** (S'03-SG'06-M'12-SM'18) received the B.Sc. cum lauda and the M.Sc. degrees from the Federal University of Rio de Janeiro (UFRJ), Brazil in 2006 and 2008, respectively; and the Ph.D. degree from the TélécomParis, IXP, France, in 2011, all in electronic engineering. Researching high-performance high-reliability circuits and systems, he joined IM2NP lab. (UMR CNRS 7334) for one year and IEMN lab. (UMR CNRS 8520) for two years during his tenure track. Since 2014, he has been with GeePs lab. (UMR CNRS 8507) as Associated Professor of the Dep. of Electronic Systems of CentraleSupélec at Université Paris-Saclay, France. His research interest is design methodologies for harsh environments and microwave integrated circuits.

**Jamel Nebhen** received the M.Sc. in Microelectronics from the National Engineering School of Sfax, Tunisia in 2007, and the Ph.D. degrees from the Aix-Marseille University, France, in 2012, all in Microelectronics. From 2012 to 2018, he worked as a Postdoctoral Researcher in France in LIRMM-Lab Montpellier, IM2NP-Lab Marseille, ISEP Paris, LE2I-Lab Dijon, Lab-Sticc Telecom Bretagne Brest, and IEMN-Lab Lille. Since 2019, he joined the Prince Sattam bin Abdulaziz University in Alkharj, Saudi Arabia, as an Assistant Professor. His research interests are mainly in the design of analog signal processing, analog and RF integrated circuits for IoT, and analog sensors instrumentation.



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

**Geoffroy Klisnick** received the Ph.D. degree in 1995 from the Paris VI University in the field of analog and mixed CMOS circuits design. Since 1998, he is Assistant Professor in the L2E Lab of the Paris VI University which is now Lab. de Génie Electrique et Electronique de Paris of Sorbonne Université and CNRS, France. Since then, he has extended his research topics to low-noise cryogenic analog circuits design in both CMOS and SiGe BiCMOS technologies, electronic instrumentation technics. These works allow him to be involved in cryogenic instrumentations for characterization of superconducting devices and terahertz detection. Nowadays, he's involved in the design of analog neuromorphic integrated circuits and analog read-out circuits for opto-electronics sensors.

**Aziz Benlarbi-Delaï** received the Ph.D. degree in electrical engineering and the Habilitation Diriger des Recherches (HDR) in Sciences Physiques from the University of Lille I, Villeneuve-d'Ascq, France, in 1992 and 2002, respectively. From 1992 to 2006, he was an Assistant Professor with the University of Lille I and was involved as a Researcher with the Institute of Electronic Microelectronic and Nanotechnology (IEMN), Villeneuve-d'Ascq, France, in the field of microwave and microfluidic devices and systems for connected objects, and also in the field of ultrafast sampling using micro and nanostructures. He is currently a Full-time Professor with the Department of Electrical Engineering, University Pierre et Marie Curie (UPMC Paris 06), Paris, France, and Head of the Laboratory of Electronics and Electromagnetism (L2E). He has authored 94 publications and communications and is the holder of two patents. He is a Visiting Professor or Member of the External Evaluation team of several foreign universities. His research interests include millimeter wave communication and localization for green radio and intelligent ambient issues. Dr. Benlarbi-Delaï participated in several Technical Program Committees of international conferences.



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH ("Springer Nature").

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users ("Users"), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use ("Terms"). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com