Enabling Low-Power Signature Recognition for the IoT with SLIF neurons

Guillaume Marthe, Claire Goursaud, Laurent Clavier

Enabling Low-Power Signature Recognition for the IoT with SLIF neurons

Guillaume Marthe
CITI, UR3720 INSA Lyon, Inria
69621 Villeurbanne, France
guillaume.marthe@insa-lyon.fr

Claire Goursaud
CITI, UR3720 INSA Lyon, Inria
69621 Villeurbanne, France
claire.goursaud@insa-lyon.fr

Laurent Clavier
CNRS, UMR 8520 - IEMN IMT Nord Europe,
Université de Lille 59000 Lille, France
laurent.clavier@imt-nord-europe.fr

***Abstract*—Energy constraints are still a significant challenge in numerous IoT applications, particularly due to the excessive power consumption of microcontrollers. To overcome this limitation, novel circuit designs have been introduced, with the integration of spiking neurons and analog computing emerging as a promising solution, facilitating substantial reductions in power consumption. However, the operation within the analog domain introduces complexities in managing the sequential processing of incoming signals, a critical requirement for diverse applications.**

This study employs the Saturating Synapses Leaky Integrate and Fire (SLIF) model, a bio-inspired neuron model, to develop a signature recognition system based on a Spiking Neural Network, without the need of non-biological techniques such as synaptic delays. SLIF neurons exhibit spiking behavior exclusively in response to two consecutive spikes with an Inter Spike Timing (IST) within a specific range, remaining unresponsive to other ISTs. We present the joint design of IST-based signatures and the corresponding network. Subsequently, we evaluate the system's efficiency in recognizing its specific sequence and discriminating against alternative sequences. The novelty of this paper lies in the proposition of a new type of temporal sequence recognition networks based on ISTs, offering significantly lower energy consumption compared to conventional approaches.

***Index Terms*—Spiking Neural Network, Low power design, Spiking Neurons, Synaptic interactions, Temporal integration**

I. Introduction

The sustainability of the Internet of Things (IoT) and its widespread adoption [1] hinges on enhancing the energy efficiency of devices. While current cellular networks, including 5G, aim to optimize performance and reduce core energy consumption, the objects longevity becomes a critical factor in various applications [2]. The efforts to reduce consumption focus on the employed algorithms, the PHY layer, and the protocols. Notably, the usage of Wake-Up Radios (WuRs) in the IoT has permitted to reduce the consumption of the nodes. Such device allows the IoT node to be asleep most of the time, and awake only if needed. However, current WuRs rely on microcontrollers for pattern recognition, consuming around $200\mu W$ [3], which is too high for the intended battery lifespan [2].

A promising avenue for achieving ultra-low-power circuits draws inspiration from biology, particularly the brain, through spiking networks. Analog spiking communication, as proposed in [4], offers energy savings compared to conventional methods [5], [6], where analog-to-digital conversion is a major energy-consuming step.

Taking this approach further, exploring analog reception solutions is crucial. Spiking Neural Networks (SNNs) have emerged as low-power solutions for signal processing, finding applications in image activity recognition [7], [8] and neuroscience studies [9]. These bio-inspired systems mimic neuronal behavior using transistors, achieving comparable computational speed with significantly lower power consumption, around $100pW$ per neuron [10], [11].

A fully analog, brain-inspired design has the potential to reduce WuR consumption by several orders of magnitude, contributing to prolonged battery life. Nevertheless, managing time in the analog domain poses a challenge, as seen in [12]. In this paper, inspired by biology [13], [14], we propose to exploit the synapse saturation effect, akin the impact of dendrites on cortical pyramidal neurons. We demonstrated that this saturation enables the recognition of specific time intervals between two spikes [15]. By parameterizing the system, we target to choose and identify desired time intervals or sequences of spikes, as expected in the development of wake-up radios (WuRs) [16], [17]. The paper presents several noteworthy contributions:

- Proposition of a neural network model based on the Saturating Leaky Integrate and Fire (SLIF) architecture, designed to emulate the impact of dendrites on cortical pyramidal neurons in biological systems.
- In-depth analysis of the behavior of the proposed network when exposed to a sequence of eight spikes, providing a detailed understanding of its performance in a more complex scenario.
- Evaluation of the system's capability to identify specific sequences based on the temporal intervals between spikes, offering insights into its pattern recognition abilities.

These findings enable the definition of sequences that can be reliably detected by the proposed neurons with a saturating effect, potentially serving as a crucial component in the development of an analog ultra-low power Wake-up Radio (WuR).

The paper is organized as follows: Section II provides an introduction to the neuron model along with its corresponding equations. In Section III, the topology of our network is described, elucidating its functioning and the co-design ap-

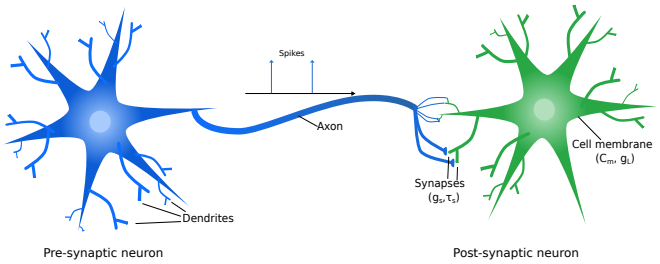

Fig. 1: Schematic representation of a biological neuron

proach for the targeted sequence. The evaluation of recognition performances is conducted in Section IV. Finally, Section V serves as the conclusion of the paper.

II. Neuron model

In cognitive processes, the biological brain employs neurons, i.e. specialized cells serving as messengers. Communication among neurons occurs through electrical pulses, or "spikes," transmitted across synapses. Those spikes are emitted by a neuron when its membrane potential reaches its threshold voltage. This phenomenon initiates the transmission of messages, constituting the fundamental process of neural communication as illustrated in Fig. 1.

The artificial spiking current replicates this behavior in response to incoming spike trains. The membrane voltage of the postsynaptic neuron is described by the Saturating Leaky Integrate and Fire (SLIF) model (1) :

$$C_m \frac{dv}{dt} = g_s(t)(E_s - v(t)) - g_L(v(t) - v_{rest}) \qquad (1)$$

where C_m is the membrane capacity, g_s the synapse conductance, E_s the synaptic reversal potential, set to $0mV$, $v(t)$ the membrane potential, g_L the membrane conductance and $v_{rest} = -65mV$ the resting potential.

There are two parts in (1), one corresponding to the behavior in response to the incoming spike, ruled by $g_s(t)$ which increases the neuron voltage, and the neuron leaky behavior ruled by g_L, which decreases the voltage. There is thus a competition between the 2 parts. The first one prevails just after the incoming spike while the second one predominates in a second time. As can be deduced from the second part, the leaky behavior leads to an exponential decrease of the membrane voltage to the resting potential v_{rest} tending to reduce the voltage as we can see on the grey curve of Fig. 2.

Meanwhile, the first part describes the way an incoming spike permits to increase the neuron voltage. By default $g_s(t) = 0$, cancelling the first term. However, an incoming spike will make $g_s(t)$ increase instantly to its saturating value g_s^{max}.

This term $g_s(t)(E_s - v(t))$ increases and becomes higher than 0 tending to increases the voltage.

Then, $g_s(t)$ decreases exponentially :

$$\frac{dg_s(t)}{dt} = \frac{-g_s(t)}{\tau_s} \qquad (2)$$

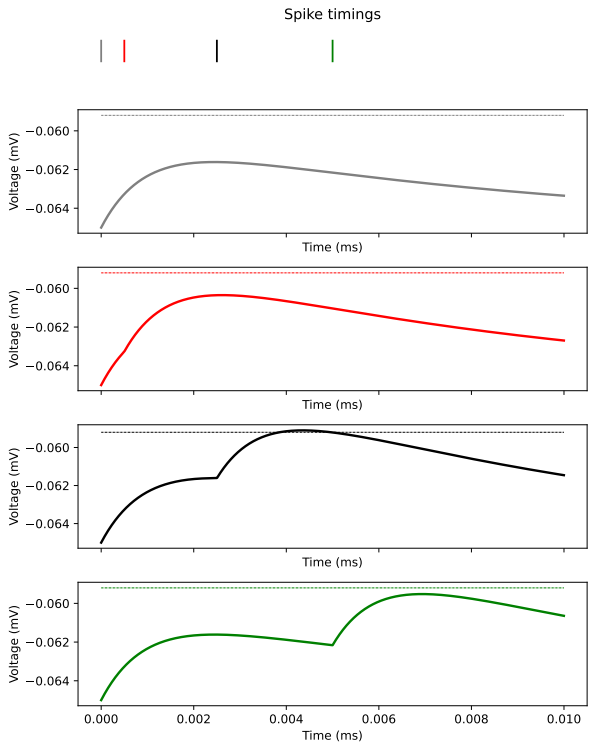

Fig. 2: Response of the SLIF neurons to different ISTs

where τ_S is the synaptic time constant which determines the dynamic.

The interest of this model is apparent during the successive transmission of multiple spikes. The restricted growth of $g_s(t)$ slows the rise to the target membrane voltage, contributing to the observed behavior depicted in the red curve in Fig. 2.

When we send two spikes, the maximal amplitude reached by the membrane potential will depend on the delay between those two spikes, called Inter-Spike Timing (IST). If those two spikes are sent too close together, then $g_s(t)$ has only slightly decreased. So the second spike, which sets $g_s(t)$ back to its maximum, has only a small impact. Indeed, the amplitude gained with the second spike will be reduced compared to the first one, as we can see on the red curve in Fig. 2. On the contrary, if these two spikes are sent too far apart, then the leaky phenomenon of the neuron will decrease the voltage between the reception of those two spikes, leading to a voltage lower than what we could have expected, as depicted on the green curve in Fig. 2. But, in between we can see on the black curve that there is a better IST that realises a trade-off between the saturation and the voltage leak and leads to the optimal membrane voltage. We can thus find an optimal IST that we call the "Favorite IST" of the neuron.

Usually, when a neuron voltage reaches its threshold, it instantly fall to its resting potential, as we can see in Fig. 3, and send a spike to the following neurons. But the thresholds represented in dashed lines on Fig. 2 are fictitious and were not implemented in the shown neurons, so that we can see

the entire dynamic of the neurons. The black curve on Fig. 2 represents the response to the Favorite IST leading to the highest amplitude. The threshold is placed arbitrarily $0.1mV$ under this amplitude. However, there are other ISTs than the favorite one which permit to reach an amplitude higher than the threshold. We call this range of ISTs the "Timewidth" (TW).

III. Network model

We have constructed our neural network as a fully feed-forward network comprising SLIF neurons. In this study, we focus on evaluating a network consisting of three neurons, each with a distinct Favorite IST, as depicted in the left part of Fig. 3. The selection of both the network architecture and the corresponding sequence is an intertwined process. Generally, an n-neuron network is associated with a 2^n sequence, structured similarly to the one illustrated in Fig. 3. Each neuron N_i is dedicated to IST_i, its favorite IST. Consequently, when it receives two consecutive spikes separated by its IST, it emits a spike to the next neuron in the network. In Fig. 3 our 3-neuron network processes an 8-spikes sequence. Examining the first line, we observe four instances where two spikes are spaced by IST_1, resulting in the transmission of four spikes to neuron N_2. Upon reception, neuron N_2 interprets these spikes as two pairs, received with an interval of time corresponding to IST_2, subsequently dispatching two spikes to neuron N_3, each spaced by IST_3. This intricate process ensures that the final neuron fires a single spike.

IV. Performances

A key constraint in this model is the requirement for ISTs to be at least double than the duration of the preceding one. This constraint arises from the potential overlapping of spikes and ISTs, leading to undesirable outcomes. Thus, we chose to design an example network with 3 neurons as defined in Table I whose ISTs are 4.57×10^{-6}, 7.41×10^{-5} and 2.31×10^{-4} respectively and verify the previous constraint.

The time-varying synapse conductance $g_s(t)$ is influenced by incoming spikes, remaining within the range of 0 to a saturation value g_s^{max}, set at $100pS$ for our simulations.

TABLE I: Neurons parameters and metrics

Neuron	$C_m(F.cm^{-2})$	$g_L(S.cm^{-2})$	$\tau_S(s)$	$IST(s)$	$TW(s)$
Neuron 1	10^{-10}	10^{-5}	3×10^{-6}	4.57×10^{-6}	1.47×10^{-6}
Neuron 2	10^{-9}	10^{-5}	10^{-4}	7.41×10^{-5}	2.81×10^{-5}
Neuron 3	10^{-8}	10^{-4}	10^{-3}	2.31×10^{-4}	1.64×10^{-4}

In our previous work, we showed that the Timewidth (TW) of our neurons is not zero, implying that a minor shift could potentially trigger a spike in the network. Thus, a single neuron is capable of recognising IST even if there is a little shift in a spike timing. In this paper, we extend this study to a neuron network to provide a comprehensive evaluation and quantification of this selectivity and to determine the overall performance metrics of the network.

This study focus on evaluating the probabilities of False Alarms (FA) and Misdetections (MD).

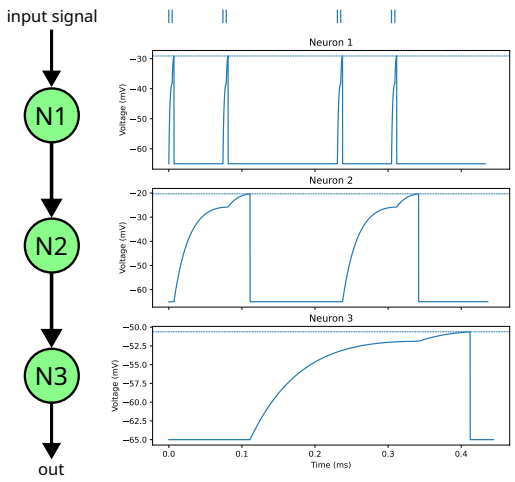

Fig. 3: System operation

Misdetection: For MD evaluation, we now introduce a frequency jitter. Rather than sending the unaltered sequence to the corresponding network, we randomly draw numbers from a normal distribution centered around 1. These randomly drawn factors are then multiplied to all spike timings, effectively introducing variability in the sequence pace. This process resulted in either accelerating or decelerating the sequence. Subsequently, we observed whether the altered sequence reached the membrane threshold for different threshold values. This analysis allowed us to gauge the network's ability to still detect the sequence subject to the variations in timing induced by frequency jitter.

The measured probabilities are presented in Fig. 4 for five distinct threshold values, impacting the selectivity of our neurons. The chosen thresholds for our neurons fall within the range of 10^{-5} to 10^{-3} Volts below the amplitude that the neuron reaches when it receives two spikes with the Favorite IST.

The observation reveals that a smaller threshold margin value leads to reduced network selectivity. This is attributed to the narrow range of amplitude values that can trigger the neuron to reach its threshold. Consequently, only a minuscule range of jitter values permits this, resulting in high Misdetection probabilities. Conversely, a higher threshold margin value increases the likelihood of reaching the neuron threshold for numerous jitter values, thereby reducing selectivity. Thus, the network becomes less discerning, detecting almost every sequence but potentially incurring a higher rate of False Alarms.

False Alarms with totally random sequence: To evaluate the False Alarms probability, we consider random spike timings within a range corresponding to the total duration of the sequence, effectively representing the transmission of a completely random sequence. We introduced various threshold margin values to quantify the system's performance. The threshold margin represents the amplitude range within which the neuron can initiate a spike to the subsequent neurons, and impacts TW. A larger margin corresponds to a lower threshold

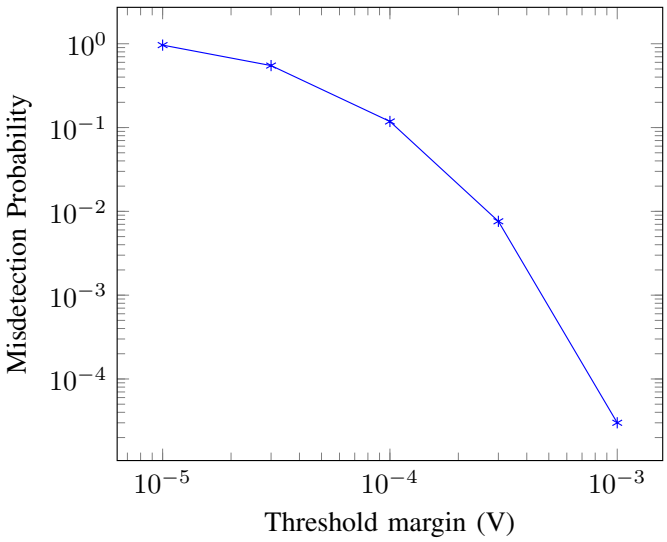

Fig. 4: Probability of Misdetection as a function of the number of neurons for a jitter of N(1,0.01)

and a higher TW. In this approach, we set the first spike at $t = 0s$ and then draw $n_s = 2^n - 1$ spikes. The FA probability can be approximated by :

$$P(FA) = \frac{\prod_{i=1}^{n} TW_i \times n!}{(T_{tot})^n} \tag{3}$$

This equation evaluates the likelihood of each spike being drawn within the correct time window around the intended IST. If a spike should occur IST_i seconds after the previous spike, it must fall within a specific range defined by the time width TW_i within the total time duration T_{tot}. Consequently, the probability of accurately drawing the spike timing within this window is calculated as TW_i/T_{tot}.

The probability of False Alarms, considering a random drawing of $n_s = 2^n - 1$ spikes, for a varying number of neurons and threshold margin, is illustrated in Fig. 5 for the neurons described in table I. This range of threshold margin values allows us to explore different levels of sensitivity and selectivity in our network's response to spike sequences. The graph demonstrates a rapid decline of the probability towards zero when we reduce the threshold margin and thus, when the TW is reduced. This is attributed to the necessity of drawing each IST within the range of the TW of the neurons, centered on the original IST value to get a FA. The first neuron has the smallest TW, so if the spike timing is outside of this thin TW, the first neuron will not emit a spike. As this TW is relatively small compared to the overall sequence range, even a slight shift of one spike timing outside this TW renders the occurrence of False Alarms nearly impossible, even if all other spikes are perfectly chosen. Then, if a second neuron is added, 3 spikes are needed to be well placed in the sequence instead of only one. The second one must be around IST1 after the beginning, the third spike must be placed around IST2, and the last one must be IST1 after the penultimate. This is why

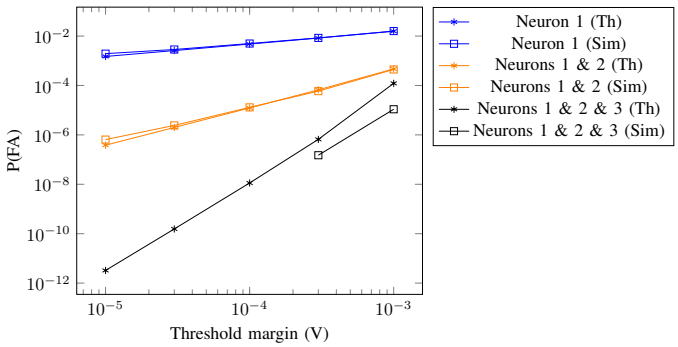

Fig. 5: Theoretical probability of False Alarm as a function of the number of neurons

the FAs are rarer when we use more neurons.

We can observe that the simulations and the theory get almost the same probabilities, but the simulation is a bit worse than what we computed when we use only one or two neurons. This is due to the fact that, in practice, for a pair of spikes, when one is drawn at the very left of the TW, and the next one is drawn at the very right of its TW, the IST will be too large to make the neuron spike, even if there are both in their predetermined TW. This phenomenon leads to a smaller risk of getting FAs when we add neurons and spikes. One may note that, when there are 3 neurons and high threshold margins, the third TW exceeds the range in which we draw our spike timings. Consequently, our simulations leads to smaller FAs than the number we should have had from (3). We thus assume that our theoretical expression acts as an upper approximation for the chances to get a FA in simulation.

Besides, some simulation points are missing. Indeed, for $n_s = 7$ spikes and the smallest threshold margin, an occurrence is exceedingly rare, standing at $P(FA) = 3.25 \cdot 10^{-12}$. Given that running 10^6 iterations takes approximately 1 hour, simulating enough iterations to encounter a False Alarm would take around 100 years. We tested and did not get a single False Alarm in 10^8 iterations for the 10^{-4} threshold margin with 3 neurons, which is consistent with the theoretical approach.

False Alarms with partially random sequences: However, the non desired sequences might not be completely random, but rather designed with the same logic. In particular, the ISTs might be chosen in different orders of magnitude. We thus define 3 different ISTs within each of 3 different ranges (between $2\mu s$ and $10\mu s$, between $20\mu s$ and $100\mu s$ and between $200\mu s$ and $1ms$). Codes are constructed by choosing one IST in each range. Thus, $3^3 = 27$ different codes can be obtained. We consider one of these as the targeted sequence. Then, we create new sequences by selecting another set of ISTs. Each one is either the reference one (with probability of 1/2) or another one within the corresponding range.

The results presented in Fig. 6 show the False Alarm probability for five distinct threshold margin values: $10^{-5}V$, $3 \times 10^{-5}V$, $10^{-4}V$, $3 \times 10^{-4}V$, and $10^{-3}V$ verifying the last constraint. The graph illustrates an increase in the probability as the threshold margin increases, indicating a decrease in

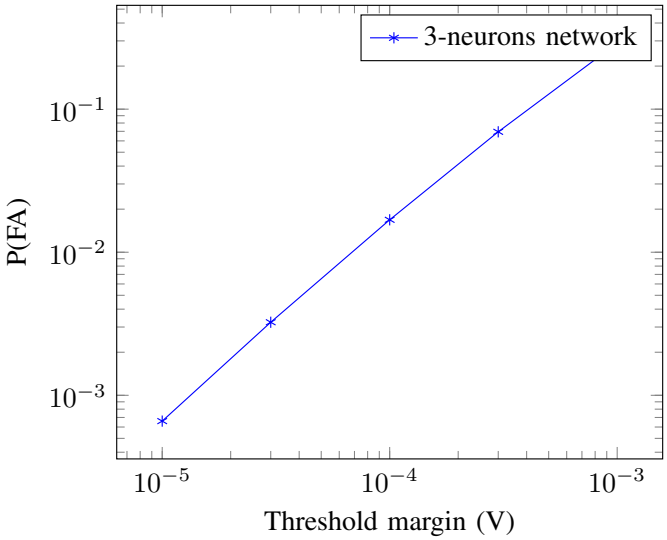

Fig. 6: Simulation probability of false alarm as a function of the threshold margin

the threshold value. This observation aligns with expectations, as a larger threshold margin permits the neuron to fire for ISTs that are further from the favorite one. Consequently, this broader range allows the triggering of False Alarms for a larger spectrum of ISTs, leading to a decrease in the overall performance of our network. Thus, our preference leans towards employing a lower threshold margin value, as it enhances network selectivity. However, it comes with the trade-off of increasing the probability of Misdetection in the presence of spike timing shifts.

As demonstrated earlier, the adjustment of this threshold margin value allows for a strategic choice between minimizing Misdetection and False Alarms. Alternatively, a compromise can be struck by selecting a value in between, yielding to moderate results for both metrics.

False Alarms with constrained sequences: In our simulations, we are currently able to use 6 ISTs which TW doesn't overlap the other ISTs, for each of range. Following our sequence design, we can recognize 6^n orthogonal sequences, with n the number of neurons and ranges in the network. For this example with 3 neurons, it allow us to use 216 different sequences. But, we are able to do the same with any other time ranges and number of neurons in other simulations. We created these 216 receivers and for each of them we presented all the sequences except for their own, to verify that the neurons won't response for other Favorite ISTs in our list. We indeed get zero FA in the simulations, implying that these 216 sequences are orthogonal with respect to this receiver.

V. Conclusion

In this paper, we proposed a SNN architecture based on SLIF neurons for pattern recognition. We showed that this type of SNN is able to discriminate an IST based pattern from others and only spike for the targeted signature. We used a simple but effective topology which consists of a totally feed forward architecture. We also jointly created the corresponding sequence and showed that it is indeed the one that makes the network spikes. We studied the performances of this network in terms of False Alarms and Misdetection, and discussed the impact of the neurons threshold margin on this metrics, and the trade-off we need to make. Currently, n neurons in the networks implies that we are able to create 6^n orthogonal sequences for the IoT. But, it necessitates 2^n spikes. However, more spikes means more consumption, this is why one of our next research focus will be to study other topologies that allowed us to use less spikes for the same number of neurons. We will also study the interest of multi-IST for each neuron to potentially allow us to get more selectivity and less neurons for the same number of spikes.

References

[1] Rishika Mehta, Jyoti Sahni, and Kavita Khanna. Internet of things: Vision, applications and challenges. *Procedia Computer Science*, 132, 2018.

[2] Halil Yetgin, Kent Tsz Kan Cheung, Mohammed El-Hajjar, and Lajos Hanzo Hanzo. A survey of network lifetime maximization techniques in wireless sensor networks. *IEEE Communications Surveys & Tutorials*, 19(2):828–854, 2017.

[3] Nafiseh Seyed Mazloum and Ove Edfors. Performance analysis and energy optimization of wake-up receiver schemes for wireless low-power applications. *IEEE Transactions on Wireless Communications*, 13(12):7050–7061, 2014.

[4] Florian et Al. Roth. Spike-based sensing and communication for highly energy-efficient sensor edge nodes. In *2022 2nd IEEE International Symposium on Joint Communications & Sensing (JC&S)*, 2022.

[5] Paul A. Merolla et Al. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, 345, 2014.

[6] Mike et Al. Davies. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 38(1):82–99, 2018.

[7] Thomas Barbier, Céline Teulière, and Jochen Triesch. Unsupervised Learning of Spatio-Temporal Receptive Fields from an Event-Based Vision Sensor. In *29th International Conference on Artificial Neural Networks,*, pages 622–633, Bratislava, Slovakia, 2021.

[8] Peter Diehl and Matthew Cook. Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Frontiers in Computational Neuroscience*, 9, 2015.

[9] Xu Zhang, Z Xu, C Henriquez, and S Ferrari. Spike-based indirect training of a spiking neural network-controlled virtual insect. 12 2013.

[10] Ilias Sourikopoulos, Sara Hedayat, Christophe Loyez, François Danneville, Virginie Hoel, Eric Mercier, and Alain Cappy. A 4-fj/spike artificial neuron in 65 nm cmos technology. *Frontiers in Neuroscience*.

[11] F. Danneville, C. Loyez, K. Carpentier, I. Sourikopoulos, E. Mercier, and A. Cappy. A sub-35 pw axon-hillock artificial neuron circuit. *Solid-State Electronics*, 153:88–92, 2019.

[12] Ilias Sourikopoulos, Antoine Frappé, Andreia Cathelin, Laurent Clavier, and Andreas Kaiser. A digital delay line with coarse/fine tuning through gate/body biasing in 28nm fdsoi. In *ESSCIRC Conference 2016: 42nd European Solid-State Circuits Conference*, pages 145–148, 2016.

[13] Cazé RD. All neurons can perform linearly non-separable computations. *F1000Research*, 2022.

[14] Therese Abrahamsson, Laurence Cathala, Ko Matsui, Ryuichi Shigemoto, and David A. Digregorio. Thin Dendrites of Cerebellar Interneurons Confer Sublinear Synaptic Integration and a Gradient of Short-Term Plasticity. *Neuron*, 73(6):1159–1172, March 2012.

[15] Guillaume Marthe, Claire Goursaud, Romain Cazé, and Laurent Clavier. On exploiting the synaptic interaction properties to obtain frequency-specific neurons. In *2023 IEEE 16th Dallas Circuits and Systems Conference (DCAS)*, pages 1–6, 2023.

[16] Ilker Demirkol, Cem Ersoy, and Ertan Onur. Wake-up receivers for wireless sensor networks: benefits and challenges. *IEEE Wireless Communications*, 16(4):88–96, 2009.

[17] Guillaume Marthe, Claire Goursaud, and Laurent Clavier. Wake-up radio receiver based on spiking neurons for detecting activation sequence. In *2023 IEEE Wireless Communications and Networking Conference*, 2023.