A spiking neuron implemented in VLSI

P Stoliar 1, I Akita 1, O Schneegans 2, M Hioki 1 and M J Rozenberg 1

1 National Institute of Advanced Industrial Science and Technology (AIST), 305-8565 Tsukuba, Japan
2 Laboratoire Génie électrique et électronique de Paris, CentraleSupélec, UMR8507 CNRS - Sorbonne Université, Université Paris-Saclay, 91192 Gif-sur-Yvette Cedex, France
3 Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, Orsay, France
E-mail: marcelo.rozenberg@universite-paris-saclay.fr

Keywords: artificial spiking neurons, neural networks, memristor, neuromorphic circuits

1. Introduction

We have recently introduced a new artificial spiking neuron circuit implemented by using a few discrete out-of-the-shelf electronic components [1]. The core of this neuron is a silicon-controlled rectifier (SCR), also known as a thyristor. It is a very common component used in power electronics and ESD protection.

This SCR-based neuron circuit already demonstrated the capability of cascading and reproducing many neuronal behaviors of biological relevance [2]. Nevertheless, there is still an open question whether it can be integrated into VLSI circuits. The main issue to migrate the circuit to VLSI is the SCR.

SCRs can be integrated with ad-hoc CMOS processes [3, 4]. However, to develop a fully neuromorphic chip, engineers require SCRs to be fully integrated into the design libraries of the main foundries. This would require a significant effort.

Thus, here we take a different approach to the problem. We look for emulating the SCR functionality by using standard MOS transistors. This should improve the appeal of the SCR-based circuit as a competitive option for VLSI implementations of neural networks and generate enough momentum to drive the introduction of SCRs into high-end processes.

Different circuits emulating SCRs functionality were already proposed, each focusing on specific application constraints. For example, Kim et al introduced an SCR emulating circuit focusing on linearity [5]. Schell et al focused on precise matching between cells [6]. Saft et al introduced a circuit using extra transistors to increase the on/off ratio and threshold voltage [7]. Li et al focused on the behavior in temperature [8]. Here, we focus on the simplest possible implementation (minimum number of transistors) that seamlessly fit into our neuron circuit.

Park et al have recently reported a comprehensive overview of the state of the art when it comes to compact neuron designs [9]. Designs using novel electronic devices in general use less components and are more energy efficient, down to six transistors and 100 pJ/pulse. Both power and transistor count tend to increase when standard CMOS technologies are used. Nevertheless, it should be noted that many novel designs that claim lower-component count cannot operate in a full neuromorphic system without the use of aditional inter-stage amplifier/buffering circuitry. Here, we are introducing a design implemented with standard VLSI technology, using nine transistors, that is capable to operate as stand alone, and with a power consumption of 100 fJ/pulse.
In previous work, we have identified the key properties of the SCRs that are relevant to our circuits, namely, that its high conduction state can be triggered by a control pin (the gate), and then it remains in that state until the (holding) current is decreased to close to 0, therefore is has hysteresis [1]. Here we demonstrate that the same functionality as the SCR can be implemented with three MOS transistors, and, provide a fully VLSI-compatible circuit.

2. Background

2.1. LIF model

Our SCR-based neuron realizes a leaky-integrate and fire (LIF) neuron. The LIF model is a basic, yet powerful, model of a neuron behavior [1, 2, 10]. Figures 1(a) and (b) depicts the basic concept behind the LIF model. Signals arrive from pre-neurons in the form of ionic currents. The neuron membrane behaves like a leaky capacitor (C with a leak resistor R in parallel). It accumulates these ions, building up the membrane potential $V_C$. This is the leaky-integration, LI.

If the membrane potential $V_C$ reaches a threshold voltage $V_{TH}$, the LIF model predicts the firing of a spike. The idea is that at the threshold the membrane channels suddenly open (not shown), and the ionic charges compensate. Therefore, the accumulated charge rapidly discharges, the neuron generates an output signal, and the integration starts over again.

Notice that in this work we are dealing with many parameters that specify threshold voltages; we use $V_{TH}$ for the threshold of the LIF behavior (figure 1), $V_{BF}$ for the voltage threshold of the SCRs, and, $V_{TH1}, \ldots, V_{TH6}$ for the threshold of the MOSFETs.

For the sake of clarity, in most discussions below we shall neglect second-order or parasitic effects such as, for example, gate leak in MOS transistors (nevertheless they are considered in all the numerical results). For the simulations presented in sections 2–4, we use the software LTspice [11]. The short channel models for the VLSI transistor are reported in [12].

A basic implementation of the LIF model is presented in figure 1(c). We assume an input current that in general is variable in time. This current is integrated by the capacitor C, while the leaky resistor R slowly discharges it. The voltage in C is continuously monitored by the amplifier A1. When it reaches a threshold $V_{TH}$, the neuron ‘fires’. An output pulse is generated, and, the capacitor is rapidly discharged through the closed switch S. Then, the output buffer amplifies the fire signal so that it can be transmitted to the following neurons.

There is plenty of circuits of LIF neurons, some implementing variations of the basic model we just described [13]. Moreover, there is plenty of VLSI implementations [14–16] and references therein. The key issue is that
2.2. Silicon-controlled rectifier

The central component in our neuron is a silicon-controlled rectifier, which we describe next. Figure 2(a) presents the symbol and structure of an SCR. They are four-layer devices [19], that are standard in power electronics; they typically handle tens of amperes and hundreds of volts.

Figure 2(b) presents the schematic current-voltage characteristics (IA versus VAK) for increasing IG [19]. Starting from VAK = 0, the SCR blocks the IA current until VAK reaches the switching voltage VBF. At this point, the SCR turns on, and IA rapidly increases with VAK (i.e., high conduction state). The SCR remains in this high conduction state until IA decreases below a holding current Ih. The SCR can also be triggered by IG. Indeed, gate and cathode form a pn junction similar to a diode. When VGK reaches ~0.6 V, IC flows, and the SCR fires.

2.3. SCR-based neuron

Figure 3(a) presents the SCR-based circuit as reported in [1]. We have, similarly to figure 1(a), the three basic blocks for a LIF neuron, namely, the 'leaky integrator' block, the 'firing & discharge' block, and the output buffer, plus, an additional subcircuit block to detect firing.

In the SCR-based circuit [1], the implementation of the 'leaky-integrator' and the 'output buffer' blocks are conventional and straightforward. However, the key aspect of that work that is relevant to us now is that the core of the 'firing and discharge' circuit is a SCR triggered by the gate. This feature is illustrated for reference in the simulation data of figure 3.

Figures 3(b)–(d) presents a simulation of the SCR-based circuit for a constant input current of 2 mA. The input current is integrated by C1, making V_A to increase until it reaches ~5 V (figure 3(b)). At that point, the voltage at the SCR gate reaches a threshold value V_GK ~ 0.6 V, and firing occurs (figure 3(c)).

We should point out here that the value of 2 mA adopted for the input current is orders of magnitude larger than the corresponding used in the VLSI implementation that we shall describe later on. In fact, the replacement of one circuit by the other only concerns their qualitative function and not the actual magnitudes of the electric parameters.

The firing sets the SCR in high conduction state. IA jumps to ~50 mA, limited by R3. The charge in C1 starts to discharge in R3 with a fast time constant with respect to the charge (notice that R3 << R1 + R2). The high conduction state remains until IA drops below the holding current IH = 2 mA. The circuit thus resets and then the integration stats over again. We finally point out that at the beginning of the discharge, VK exceeds the threshold voltage of Q1 (~1 V), turning Q1 and Q2 on, and generating an output pulse (figure 3(d)).

3. Replacing sub-circuits

Our strategy to convert the circuit in figure 3(a) into the VLSI circuit in figure 4, is to rethink the implementation of each block separately. We shall consider first the most important ‘firing and discharge’ block, which involves the main issue that we address in this work, namely the VLSI replacement of the SCR. Therefore, we provide next a brief description of the implementation of the neuron using discrete components.

From a general perspective, the SCR is a device that realizes an instance of S-type negative differential resistance (S-NDR). In an early work in 1985, Chua et al proposed a catalog of circuits to implement S-NDR by using bipolar and field-effect transistors, and resistors [20]. However, implementations relying solely on three...
enhancement-mode short-channel MOSFETs are still missing. Below, we shall describe how to obtain such an implementation.

### 3.1. Firing and discharge sub-circuit

The 4-layer structure of the SCR (figure 2(a)) can be interpreted as two 3-layer structures (figure 5(a)), leading to the well-known 2-bipolar transistor analogy of the SCR (figure 5(b)) with an npn and a pnp components. These bipolar transistors cannot be simultaneously replaced by MOS transistors. Still, revisiting this 2-transistor equivalent circuit is useful for our goal.

We assume a positive bias $V_{AK}$, say $V_{AK} > 1\, V$, and $V_S = 0$ (figure 5(b)). Both transistors are off. We now start increasing $V_S$, leading to $I_G$ eventually flowing into the base of $T_2$. $I_G$ gets amplified to $I_{C_2} = \beta_2 I_{B_2}$, where $\beta_2$ is the current gain of $T_2$. We see that $T_2$ is a current amplifier (see below).

$I_{C_2}$ flows into the base of $T_1$ and gets amplified as $I_{C_1} = \beta_1 I_{B_1}$. Now, $I_{C_1}$ adds up to $I_{G}$; we have a runaway condition where the current of both transistors continuously increases.

Notice that $r_S$ should be relatively high compared to the dynamical input resistance of $T_2$. For example, if $r_S$ is zero, $I_{C_1}$ would not add-up to increase the current into $T_2$ because $V_{G_2}$ is fixed to $V_S$.

We do not expect a very high effective $\beta$ for the $T_1$ when the transistors enter in this runaway condition (i.e., the current gain, $\beta$, of bipolar transistors depends on the collector current. $\beta$ remains almost constant for
nominal currents, nevertheless, for relatively high collector currents it drops as $I_{C1} \sim I_{B1}$ [21]). We can think of $T_1$ behaving like a loose current mirror circuit [21]. At this point, even if we turn $V_S$ off, both transistors remain in a high conduction state. To turn the transistors off, we should force $I_{C1}$ and $I_{C2}$ to drop down close to zero, hence equivalent to going below $I_h$ in an SCR. We also see that the gate node is a current summing point, $I_{B2} = I_{C1} + I_G$.

We have just dissected the SCR into (i) a current amplifier, (ii) a current mirror, and (iii) a current summing point. This sets the stage for the introduction of our VLSI equivalence of the SCR that is presented in figure 5(c).

As before, we first assume $V_S = 0$ and all transistors ($M_2$–$M_4$) are off. As $V_S$ approaches $V_{TH4}$, the current $I_{D4}$ becomes significant. Initially, the voltage across $r_S$ is 0 because no current flows. $M_2$ and $M_3$ are a current mirror with ratio $1/10$ since the area ratio of both transistors is 10 (we shall describe the determination of such a ratio value later in section 4). Thus, $I_{D4}$ flows out of the drain of $M_3$ and, therefore, $M_2$ injects a current $I_{D2} = I_{D3}/10$ into $r_S$. The voltage drop on $r_S$ adds to $V_S$ to increase the value of $V_{GS4}$, so $I_{D4}$ increases further too (i.e., the current flowing through the source–drain of $M_4$). Hence, $I_{D3}$ and $I_{D2}$ further increase as well. We have, similarly as before, a runaway condition.

Importantly, notice that at this point, even if $V_S$ is turned off, $I_{D2} r_S > V_{TH4}$. To turn off the circuit, we should decrease the currents until $V_{GS4}$ goes well below $V_{TH4}$. Therefore, we also realized the current holding property, which provides the hysteresis behavior that is a key feature of the SCR.

In figure 5(d) we may observe that our qualitative discussion is indeed realized in the simulated I–V characteristic of the circuit in figure 5(c). A sharp threshold is observed for $V_A$ depending on $V_S = 190$ mV…230 mV. We shall see later in section 4 that for the final optimized circuit $I_A$ jumps from $<100$ nA to $>1$ μA when firing occurs, at $V_A \sim 0.6$ V.

### 3.2. Firing detector

The VLSI ‘firing detection’ circuit (see figure 4) is conceptually different from the one used in the SCR-based circuit (figure 3(a)).

$M_4$ and $M_5$ (respectively in yellow and orange blocks of figure 4) are equal and share the same $V_{GS}$. Therefore, $I_{DS} \approx I_{D4}$.

From the optimized circuit data that we shall discuss later on (cf figure 6(b) red curve $[W/L]_3 = 100$), the current source of the firing detector block may be conveniently chosen at the intermediate $I_2 = 100$ nA, which lays between the off-current $\sim 5–50$ nA and the discharge current $\sim 0.2–10$ mAmp. Thus, before firing the voltage at the gate of $M_4$ and $M_5$ is low and $I_{DS} \approx I_{D4} < 100$ nA. Hence, the current source $I_2$ is in compliance. The output of the firing circuit block (drain of $M_5$) is in a high state ($\sim VDD$) and $VOUT$ is at $\sim GND$ (see below).

When firing occurs, $C_1$ gets discharged mainly through the channels of $M_3$ and $M_4$, as a high voltage is now present at the gates of $M_4$ and $M_5$, this would lead to large currents $I_{DS} \approx I_{D4}$. As $I_{DS}$ would greatly exceed the set...
value of 100 nA of the current source $I_2$, then, the output of the firing detector (drain of M5) goes to a low state ($\sim \text{GND}$), and $V_{OUT}$ is now at $\sim V_{DD}$ (see below).

3.3. Leaky integrator and output buffer

The implementation of these two blocks is straightforward. The integration is still performed in a capacitor. We replaced the two resistors ($R_1$ and $R_2$) that bias the gate of the SCR (figure 3(a)), by a diode in series with a current source, M1 and $I_1$. The output buffer is a standard CMOS inverter.

4. Results

The design of the circuit mainly involves defining the [W/L] ratio of the transistors [22], selecting $C_1$, designing $M_1$, and specifying $I_1$ and $I_2$.

In the previous section, we have already discussed $I_2$, $M_5$ (which is equal to $M_4$), and $C_1$. For the output buffer, we can simply choose $[W/L]_b = [W/L]_1 = 1$; however, this may eventually need to be redesigned depending on the circuit load.

The diode $M_1$ is implemented as a gate-drain shorted NMOS. The determination of this transistor geometry is done in the next sub-section.

The final selection of the components for the ‘leaky integrator’ and the ‘firing and discharge’ blocks involved the optimization process described next. Such optimization is necessary to tune the component specifications (designed using standard design rules) to operate with short-channel transistors.

4.1. Optimization

Figure 6(a) presents the circuit that we used to design the values for $M_1$–$M_4$, and $I_1$. In figure 6(b) we show one example of the procedure we used. We estimated the evolution of $I_{\text{disch}}$ for $V_C$ ramping up from 0.2 V to 0.7 V (rate 0.5 V ms$^{-1}$) and then going back to 0.2 V. Two of these curves present hysteresis in $V_C$, while the other two show no hysteresis at all. Before, we had identified the hysteresis as one of the key features of the SCR.

Thus, we then proceed to iteratively adjust the values of the six parameters defined in table 1 with the goals to: (i) maximize the hysteresis in $V_C$, (ii) maximize the jump of $I_{\text{disch}}$ when firing, and (iii) reduce the whole area of the circuit (smaller $W$ and $L$).

In the example of figure 6(b), we changed the $[W/L]_3$ parameter in factors of 10. After this iteration, we decided to adopt $[W/L]_3 = 100$ because it maximizes hysteresis.

| Symbol | Description | Optimal Value |
|--------|-------------|---------------|
| $S$    | Scale factor. $W_1 = L_2 = L_3 = L_4 = S$ | 130 nm |
| $[W/L]_1$ | Geometry of $D_1$, $I_1 = S/[W/L]_1$ | 0.1 |
| $[W/L]_2$ | Geometry of $M_2$, $W_2 = S \times [W/L]_2$ | 10 |
| $[W/L]_3$ | Geometry of $M_3$, $W_3 = S \times [W/L]_3$ | 100 |
| $[W/L]_4$ | Geometry of $M_4$, $W_4 = S \times [W/L]_4$ | 1 |
| $I_1$ | Current sink | 10 nA |
4.2. Simulations

Figure 7 presents simulation results of the VLSI circuit upon application of the two paradigmatic types of excitation inputs: train of pulses and DC current.

The analysis for input pulses (figure 7(a)) is interesting because we can see the leaky integration. The initial train of pulses have an amplitude of 750 nA. We observe that for our choice of parameters, 8 pulses with a frequency of 500 kHz are required to fire the neuron. The capacitor voltage increases ∼ 75 mV per pulse, and it decays between pulses because of the leaky term.

As expected for a LIF neuron, when the amplitude of the pulses is increased, the number of pulses required for it to fire decreases. As seen in the figure, the circuit requires 4 pulses of 1 µA, and, 3 pulses of 1.25 µA to reach the threshold.

The analysis for the case of DC input (figure 7(b)) is particularly important for our LIF neuron because it might evidence a potential issue. Specifically, one could worry that if the input current $I_{IN}$ is set higher than the holding current, the circuit might never recover from the firing condition.

From the data of panel (b), we see that this issue does not arise. In fact, from the results of figure 6(b) and the choice $[W/L]_3 = 100$, we established that the holding current is ∼ 200 nA, so we just need to remain beneath that threshold. Thus, for an input of 65 nA the capacitor charges until the threshold voltage 0.57 V, and then, the circuit starts to fire. A train of output pulses is generated as the capacitor discharges down to 0.34 V during each fire event. The output frequency reaches 140 kHz. Moreover, increasing the amplitude of the input, the output frequency increases as expected for LIF behavior. It reaches a value of 260 kHz for an input current of 95 nA, and 330 kHz for 130 nA.

Generally, silicon neurons operate at frequencies much higher than biological ones (∼ kHz), which allows fast neuromorphic computation.

4.3. Stability/limits/performance

In figure 7(b) we observe that $V_{C1}$ oscillates between a resting potential $V_{REST} = 340$ mV and $V_{TH} = 757$ mV. We studied the stability of these two parameters with respect to temperature and change in $V_{DD}$. The temperature drift of $V_{REST}$ is $-0.17$ mV/°C ($<7\%$ in the range $-25 \sim 125$ °C), but $V_{TH}$ drift is $-0.74$ mV/°C$^{-1}$, getting as low as 0.5 V for 125 °C. The sensitivity to $V_{DD}$ is 32 mV V$^{-1}$ for $V_{REST}$ and 5.6 mV V$^{-1}$ for $V_{TH}$ (both shift <10% in the range 0.8 ~ 1.2 V).

The operating frequency is set by $C_1$ and $I_{IN}$. There is no upper limit for $C_1$ nor lower limit for $I_{IN}$. The minimum $C_1$ is 10 fF and $I_{IN}$ max is 140 nA, resulting in a max frequency of ~1 MHz. The $I_{IN}$ limit can be extended by reducing $W_1$ and increasing $I_2$.

We estimated the power consumption by connecting the input to a 1 V source with a 7 MΩ resistor and loading it with another 7 MΩ resistor to $(V_{REST} + V_{TH})/2$. The 7 MΩ resistors simulate pre- and post-synapses. The energy drawn from $V_{DD}$ was 100 fJ/pulse at 1 MHz ($C_1 = 0.1$ pF).
5. Post-layout simulations

Figure 8 shows the VLSI layout and post layout simulation results where TSMC 65 nm CMOS process is assumed. As shown in figure 8(a), a part of the 1 pF capacitor, in figure 4, is implemented by metal-oxide-metal (MOM) capacitor of 610 fF which consists of metals from fourth to seventh layers. The rest 390 fF is realized by MOS capacitor as shown in figure 8(b). This hybrid implementation of C1 enables to save die area. The current sources in figure 4 are implemented with transistors of a longer channel length of 500 nm to relax short-channel effect, and these are included in the layout. The designed VLSI circuit occupies 33 $\mu m \times 10 \mu m$.

It is not a surprise that the C1 takes most of the area since the miniaturization of the ‘membrane’ capacitor remains an issue for all VLSI implementations of spiking neuron circuits.

The post-layout simulation results are shown in figures 8(c) and (d), where the parasitic capacitors and resistors are extracted and applied. In these simulations, input conditions are the same as those in figures 7(a) and (b), respectively, and we can observe the same qualitative results as in the LTspice simulations reported in figure 7.

6. Conclusions

In conclusion, we have introduced a VLSI implementation of our SCR-based neuron circuit, which is a realization of the leaky-integrate-and-fire (LIF) neural model. The original circuit was based on discrete components and was built around the concept of the memristive behavior of a silicon-controlled rectifier (SCR). A main part of the present work was therefore devoted to implementing the functionality of an SCR in VLSI technology. This goal was achieved using a three MOS transistor circuit within 130 nm technology simulated with openly available LTSpice simulation models. We also present post-layout simulation of the circuit designed by using TSMC 65 nm CMOS process, with a total area of 330 $\mu m^2$. This may motivate large scale implementations of artificial intelligence circuits of unprecedented simplicity.
Acknowledgments

This work was partially conducted within the framework of the AIST-CNRS-CentraleSupélec Joint Research Agreement 'Bioinspired electronic systems'. MJR acknowledges support from the French ANR project 'MoMA' ANR-19-CE30-0020. Author contributions: PS, OS, and MJR designed the architecture of the circuit in figure 4, obtained the results presented in section 4, and wrote the manuscript (except section 5). PS, IA and MH adjusted the circuit to VLSI standars. IA did the CMOS implementation in TSMC 65 nm CMOS process and obtained the results presented in section 5 and wrote that section. All the authors contributed to the final revision of the manuscript and discussion of the results.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

MJ Rozenberg @ https://orcid.org/0000-0001-9161-0370

References

[1] Rozenberg M J, Schneegans O and Stoliar P 2019 Scientific Reports 9 1
[2] Stoliar P, Schneegans O and Rozenberg M J 2020 Frontiers in Neuroscience 14 421
[3] Tong X, Wu H, Liang Q, Zhong H, Zhu H, Chen D and Ye T 2012 Proc. of the SISPAD (Denver, CO) 316
[4] Ker M D and Hsu K C 2005 IEEE Trans. Device Mater. Reliabil. 5 235
[5] Kim G, Kim M K, Chang B S and Kim W 1996 IEEE J. Solid-State Circuits 31 966
[6] Schell B and Tavilidis Y 2008 IEEE J. Solid-State Circuits 43 1227
[7] Saft B, Schafer E, Jager A, Rolapp A and Henning E 2014 European Solid-State Circuits Conf. (ESSCIRC), ESSCIRC 2014–40th (Piscataway, NJ) (IEEE) 123–6
[8] Li J, Lin Y, Ye S, Wu K, Ning N and Yu Q 2020 Micromachines 11 124
[9] Park Y-S, Woo S, Lim D, Cho K and Kim S 2021 Front. Neurosci. 15 309
[10] Abbott L F 1999 Brain Res. Bull. 50 303
[11] LTSpice® 2020 Analog Devices Inc., USA. (Freeware) Available: (www.analog.com/en/design-center.html)
[12] Zhao W and Cao Y 2006 IEEE Transactions on Electron Devices 53 2816
[13] Indiveri G et al 2011 Front. Neurosci. 5 73
[14] Chicca E, Stefanini F, Bartolozzi C and Indiveri G 2014 Proc. of the IEEE 102 1367
[15] Joubert A, Belhadj B and Héliot R 2011 9th Int. New Circuits and Systems Conf. (Piscataway, NJ) (IEEE) 9–12
[16] Kornijcuk V, Lim H, Seok J Y, Kim G, Kim S K, Kim I, Choi B J and Jeong D S 2016 Front. Neurosci. 10 212
[17] Izhikevich E M 2004 IEEE Transactions on Neural Networks 15 1063
[18] Markram H 2006 Nat. Rev. Neurosci. 7 153
[19] Sze S M and Ng K K 2006 Physics of Semiconductor Devices 3rd edn (New York: Wiley)
[20] Chua L O, Yu J and Yu Y 1985 IEEE Trans. Circuits Syst. 32 46
[21] Gray P R, Hurst P, Meyer R G and Lewis S 2001 Analysis and Design of Analog Integrated Circuits (New York: Wiley)
[22] Baker R J 2008 CMOS: Circuit Design, Layout, and Simulation 2nd edn (New York: Wiley)