Ultra-Low Energy and High Speed LIF Neuron using Silicon Bipolar Impact Ionization MOSFET for Spiking Neural Networks

Alok Kumar Kamal and Jawar Singh, Senior Member, IEEE

Abstract—Silicon bipolar impact ionization MOSFET offers the potential for realization of leaky integrated fire (LIF) neuron due to the presence of parasitic BJT in the floating body. In this work, we have proposed an L shaped gate bipolar impact ionization MOS (L-BIMOS), with reduced breakdown voltage ($V_B = 1.68$ V) and demonstrated the functioning of LIF neuron based on positive feedback mechanism of parasitic BJT. Using 2-D TCAD simulations, we manifest that the proposed L-BIMOS exhibits a low threshold voltage (0.2 V) for firing a spike, and the minimum energy required to fire a single spike for L-BIMOS is calculated to be 0.18 pJ, which makes proposed device 194× more energy efficient than PD-SOI MOSFET silicon neuron (MOSFET silicon neuron) and 5 × 10$^3$ times more energy efficient than analog/digital circuit based conventional neuron. Furthermore, the proposed L-BIMOS silicon neuron exhibits spiking frequency in the GHz range, when the drain is biased at $V_{DG} = 2.0$ V.

Keywords—Leaky integrated fire (LIF), L shaped Gate Bipolar Impact Ionization MOS (L-BIMOS), Impact ionization MOS (I-MOS).

I. INTRODUCTION

The resemblance of the spiking neural network (SNN) with a biological neural network has captivated much attention for the development of neuromorphic computing. The SNN is a third generation artificial neural network (ANN), where the neuron model can convey information by the generated spikes. Moreover, the functionality of spiking neuron is similar to the biological neuron, and one such model is known as the leaky integrated fire (LIF) neuron model. For the implementation of the LIF neuron, conventional analog/digital circuits are widely used [1]–[6]. However, such circuit neurons have two significant limitations (a) extensive circuitry (or silicon overhead), and (b) large power consumption.

It is therefore necessary to address these limitations, recently in [7] the authors have demonstrated a silicon neuron that captures the functioning of the LIF neuron using PD (partially depleted) - SOI (silicon on insulator) MOSFET under impact ionization biasing which utilizes the charge-discharge mechanism of excess carriers in MOSFET silicon neuron. Two mechanisms can enable these excess carriers (a) the impact ionization [2], and (b) the band-to-band tunneling (BTBT) [8]. The PD-SOI MOSFET silicon neuron is energy efficient in contrast to circuit neuron implementation. However, higher supply (avalanche breakdown) voltage requirements of PD-SOI MOSFET silicon neuron to enable impact ionization (or generate excess carriers) for firing a spike consumes significant power and makes it less energy efficient, also slow floating body charging transients of the MOSFET silicon neuron exhibits slow transients behavior of the order of microseconds [9].

Therefore, to address the above mentioned limitations of the analog/digital circuit based neurons and the PD-SOI MOSFET silicon neuron, in this paper, first, we investigated a silicon bipolar impact ionization MOS (BIMOS) device which utilizes the positive feedback mechanism of parasitic BJT present in the floating body to realize the “leaky-integration” and “fire” modes needed for LIF neuron. However, this conventional BIMOS exhibits an avalanche breakdown voltage of $V_B = 2.85$ V [10], which is extremely high and it may consume huge amount of power to realize a silicon neuron. Thus, we modified the conventional BIMOS structure into ‘L’ shape gate, which results in the crowding of the electric field around the corner of the gate that reduces the avalanche breakdown voltage, referred as L-BIMOS. This ‘L’ shape gate structure of the proposed device also facilitate higher impact generation.
The proposed L-BIMOS silicon neuron emulates the LIF neuron leaky behavior without any additional circuitry, unlike circuit neurons. It also exhibits significantly low avalanche breakdown voltage in contrast to conventional BIMOS. Thus, the proposed L-BIMOS silicon neuron is compatible with CMOS process, exhibits fast transient behavior, and 194× more energy efficiency as compared to its counterpart PD-SOI MOS silicon neuron.

II. Device Structures and Simulation Parameters

Fig. 1 shows the cross-sectional views of (a) PD-SOI MOSFET silicon neuron [2], (b) conventional BIMOS, and (c) proposed L gate shaped BIMOS (L-BIMOS) silicon neuron. In the proposed L-BIMOS, the portion of the channel region which is remained uncovered with the gate towards drain terminal is referred as open-body (LOB) region, as shown in Fig. 1 (c), whereas in PD-SOI MOSFET silicon neuron whole channel region is covered with gate, as shown Fig. 1 (a). This open-body region facilitate the maximum applied potential \( V_{DG} \) across open-body region that results in a high electric field \( E \). The higher electric field in this region further enhances the impact ionization rate (IIR) in the proposed L-BIMOS. The enhanced IIR increases the rate of accumulation of holes in the floating body that leads to turn on the parasitic BJT action, hence, fast (order of nanoseconds) transient behavior of LIF neuron. In addition, the generated electrons due to high electric field flow towards drain region without passing under gate oxide, which results in less carrier injection in gate oxide as compared to PD-SOI MOSFET neuron, as shown in Fig. 1. Further, the gate length of the L-BIMOS is scaled down to 45 nm by keeping \( L_G/LOB \) ratio = 1. We have considered the silicon film thickness underneath the gate and the source region is 25 nm. Remaining simulation parameters and dimensions for both devices are tabulated in Table-I.

The Atlas Silvaco version 3.10.18R is used to simulate the behavior of the proposed device. The device is designed in 2-D background after including parallel electric-field-dependent model, concentration dependent Shockley-Read-Hall model, Fermi-Dirac carrier statistic, standard band to band tunneling (BTBT) model and Masetti mobility models. To initiate impact ionization in the proposed device SELB model is incorporated. In addition, to simulate breakdown phenomenon we have used CURVETRACE algorithm [12].

III. Parasitic BJT Action in L-BIMOS

The snapback observed in the output characteristic confirms the parasitic BJT action contributing to breakdown [13], as shown in Fig. 2(a). We can see that at \( V_{SG} = 0 \) V the breakdown voltage of conventional BIMOS is \( V_B = 2.85 \) V. The essential parasitic BJT action that describes the charge integration and leaky functionality required for spiking neural network (SNN) to be mimicked artificially by the silicon neuron. However, to recognize conventional BIMOS as a LIF neuron, we have to supply \( V_{DG} > 2.85 \) V which is high enough and consumes large amount of power. For the energy efficient use of conventional BIMOS as an LIF neuron, supply voltage should be kept as low as possible, therefore, the breakdown voltage of the conventional BIMOS needs to be reduced further.

To attain the lower value of breakdown voltage, we have proposed a BIMOS by transforming the shape of gate to L shape and named it as L-BIMOS. The proposed device (L-BIMOS) exhibits crowding of electric field at the corner as well as at the edges of gate, and hence, increases electric field at the open-body (LOB) region when compared to the conventional BIMOS, as shown in Fig. 2(b). The increased electric field in open-body (LOB) region of L-BIMOS lowers the breakdown voltage to \( V_B = 1.68 \) V. Moreover, from Fig. 2(b) we can deduce that the electric field density in the open body region (LOB) of the L-BIMOS is 1.6 times higher than that of conventional BIMOS. The high electric field in L-BIMOS tends to further increases the impact ionization rate (IIR). The increased IIR results in rise of total drain current \( I_D \), this occurs due to accumulation of generated excess holes in floating body region which turns on the parasitic BJT action [10] followed by forward biasing the source channel junction resulting in increase of the total drain current of L-BIMOS (positive feedback mechanism). On the other hand, the generated electrons due to impact ionization escapes towards the highly doped drain side in the absence of gate oxide above it. Therefore, the injection of electrons in gate oxide is minimized in the conventional BIMOS and the proposed L-BIMOS.

IV. L-BIMOS as a LIF Neuron: Results and Discussion

To replicate biological neuron artificially, the proposed L-BIMOS silicon neuron exhibits charge integration through positive feedback mechanism, and leaky via leakage current that allows L-BIMOS to function as a silicon neuron for SNN. In biological neuron the effect of \( Na^+ \) channels activation and deactivation is same as positive feedback mechanism in L-BIMOS silicon neuron which occurs due to parasitic BJT.

| Parameters                        | BIMOS | L-BIMOS |
|-----------------------------------|-------|---------|
| Source doping \((N_D) (cm^{-3})\) | \(10^{19}\) | \(10^{20}\) |
| Drain doping \((N_D) (cm^{-3})\)  | \(10^{20}\) | \(10^{20}\) |
| Channel doping \((N_A) (cm^{-3})\)| \(5\times10^{11}\) | \(5\times10^{12}\) |
| Gate Work Function (eV)          | 4.6   | 4.6     |
| Gate length \((L_G) (nm)\)        | 75    | 45      |
| Source length \((L_S) (nm)\)      | 100   | 100     |
| Drain length \((L_D) (nm)\)       | 100   | 100     |
| Open body length \((L_{OB}) (nm)\)| 75    | 45      |
| Channel Length \((L_{CH} = L_G+L_{OB}) (nm)\) | 150 | 90      |
| Trench depth \((T_D) (nm)\)       | -     | 25      |
| Open body depth \((L_{OD}) (nm)\) | -     | 25      |
| Channel thickness at drain side \((T_{CD}) (nm)\) | 50 | 50      |
| Channel thickness at source side \((T_{CS}) (nm)\) | 50 | 25      |
| Box thickness \((T_{BOX}) (nm)\)  | 100   | 100     |
| EOT \((\varepsilon_{O}) (nm)\)    | 3     | 3       |
Fig. 2. (a) Output characteristics comparison of the proposed device (L-BIMOS) with the conventional BIMOS at $V_{SG} = 0$ V, and (b) electric field comparison between L-BIMOS and conventional BIMOS.

action. Another important property of biological neuron is its leaky behavior and in L-BIMOS silicon neuron that is incorporated by hole leakage current ($I_{\text{leaky-hole}}$) which is produced at source-channel junction. Further, in biological neuron $K^+$ channel activation plays major role in resetting the neuron and to reproduce same reset mechanism in L-BIMOS silicon neuron, we reduce $V_{DG}$ to zero volts. Detailed description and analysis of the proposed L-BIMOS silicon neuron is provide in the following sub-sections.

A. Operation of L-BIMOS silicon neuron

In this Section, we have discussed the functioning of the proposed device (L-BIMOS) as a LIF neuron with the help of energy band diagrams and carrier concentration, as shown in Fig. 3. The biasing of the proposed device (L-BIMOS) under equilibrium ($V_{SG} = V_{DG} = 0$ V) and its energy band diagram is shown in Fig. 3(a). From energy band diagram, we can observe that the potential barrier height at the source channel region is very high, as a result there is no thermionic emission of the electron from source side to the drain side. Therefore, from Fig. 3(a-i) we can infer that under this condition the proposed device does not conduct. So in order to start conduction, the impact ionization is initiated with the help of the biasing scheme as shown in Fig. 3(a-ii). In accordance to the biasing scheme the L-BIMOS is biased at $V_{SG} = -0.2$ V and $V_{DG} = 2.0$ V due to which the potential barrier at the source channel region reduces, as shown in Fig. 3(b-i). As a result, the thermionic emission of the electron starts from the source side to drain side of the device. In addition, when these electrons move towards the open-body region, the electrons get heated up by the applied field. Some of them acquires enough energy to produce impact on covalent bond, thereby new electrons and holes are created. The generated electron moves towards drain, and holes are swept into floating body which results in the electrostatic lowering of the potential barrier at the source-channel junction and due to which the hole leakage current ($I_{\text{leaky-hole}}$) starts flowing from the channel region to the source region that defines the “leaky” behavior of the proposed L-BIMOS silicon neuron [7].

As electron current is greater than the hole leakage current ($I_{\text{leaky-hole}} >> I_{\text{II-electron}}$) shown in Fig. 3(b). The hole concentration and the total current density of silicon neuron is shown in Fig. 3(b-ii) and Fig. 3(b-iii). Further, with the progress of time, the source-channel junction shows additional potential barrier lowering as shown in Fig. 3(c) due to excess generation of electron and hole pair which results in higher concentration of holes in floating body, as shown in Fig. 3(c-i). Due to this additional source-channel potential barrier lowering
higher rise in total current is observed, as shown in Fig. 4(a-c-iii) as compared to Fig. 4(b-iii). With the passage of time, further lowering of barrier stops as soon as \(I_{\text{leaky-hole}}\) becomes equal to \(I_{\text{II-hole}}\) which leads to total current \((I_D)\) saturation in L-BIMOS. Before total current \((I_D)\) touches saturation limit, a pre-defined threshold \((I_{th} = 0.5 \text{ mA/\text{um}})\) is achieved that mimics the “fire” mode.

At \(I_D = I_{th}\) the proposed device resets by supplying \(V_{DG} = 0 \text{ V}\). This results in recombination of holes present in channel region and the device is completely at reset condition, as shown in Fig. 4(d-i). In this state, the \(I_{\text{leaky-hole}}\) and \(I_{\text{II-hole}}\) currents are zero and the neuron is said to be at reset condition. When the reset is complete the device is again biased with \(V_{DG} = 2.0 \text{ V}\) after certain refractory time to re-initiate the LIF process. Thus, the input voltage \(V_{SG}\) produces a response in terms of spike as shown in Fig. 4(d-ii).

**B. Impact Ionization in L-BIMOS Silicon Neuron**

The proposed device (L-BIMOS) works on the principle of impact ionization (II) that is responsible for the generation of excess carriers required to work as silicon neuron. The II is a phenomenon in which the electrons from the source side traverses towards the drain side through the channel region. A small amount of these electrons that produces drain current, tends to scatter because of high electric field in open body region, due to which electron-hole pairs generate. The generated electrons moves towards the drain whereas the holes gets accumulated in the floating body resulting in lower potential barrier at source-channel junction.

To initiate the II in L-BIMOS silicon neuron the required bias \((V_{DG} = 2.0 \text{ V} \text{ and } V_{SG} = -0.2 \text{ V})\) is applied. Due to the applied bias there is instantaneous potential barrier lowering which generates the drain current instantly, as shown in inset in Fig. 4(a). As the potential barrier further lowers due to hole accumulation in floating body that further increases the drain current of L-BIMOS as compared to conventional BIMOS, as shown in Fig. 4(a) reaches towards the saturation current.

The increase in drain current can be well understood from Fig. 4(b-ii-i) in which at 1 ns we can see that the total current density of L-BIMOS is higher as compared to conventional BIMOS because the impact ionization rate (II) of L-BIMOS is greater than conventional BIMOS, as shown in Fig. 4(b-iii-iv). Hence, due to high II in the L-BIMOS, we can observe the fast transient as compared to conventional BIMOS. The fast transients in L-BIMOS results in higher spike frequency as compared to that of conventional BIMOS which makes L-BIMOS more suitable to realize as spiking neural networks (SNN) using silicon neurons.

**C. Impact of \(V_{SG}\) on Positive Feedback**

For mimicking the neuron behavior, positive feedback plays an important role during integration of drain current over time, under different bias voltages \((V_{SG})\) and \(V_{DG}\) is fixed at 2 V is shown in Fig. 5(a). It can be observed that for certain \(V_{SG}\) the \(I_D\) reaches a threshold current, \(I_{th}\) (fixed at 0.5 mA/um). The \(V_{SG}\) corresponding to \(I_{th}\) is considered as the threshold voltage \((V_{th})\) of L-BIMOS silicon neuron and it is \(|V_{th}| = 0.2 \text{ V}|\). As far as \(|V_{SG}|\) is less than \(|V_{th}|\) the positive feedback will not be establish (Fig. 5(b)). Under this condition the proposed L-BIMOS is unable to generate a sufficient amount of excess holes, which further fails to provide sufficient positive feedback and hence, the total drain current saturates (Fig. 5(c)) before it reaches \(I_{th} = 0.5 \text{ mA/\text{um}}\), thus no spike is generated. Beside this as we increase \(|V_{SG}| \geq |V_{th}|\) the total drain current of the proposed device reaches very fast to \(I_{th} = 0.5 \text{ mA/\text{um}}\) (“leaky integration”) generating a spike as shown in Fig. 5(c) and this is due to the presence of sufficient positive feedback (i.e. \(|V_{SG}| \geq 0.2 \text{ V}|\). On the other hand, in biological neurons this positive feedback mechanism is controlled by \(Na^+\) channels activation and deactivation. Hence, we can observe that the strength and resemblance of positive feedback mechanism in our proposed device with the biological neuron that depends on the input bias voltage \((V_{SG})\).
D. Spiking Behavior of L-BIMOS Silicon Neuron

For realization of spiking neural network, spiking behavior of the proposed device under different biasing conditions ($V_{SG}$ and $V_{DG}$) is shown in Fig. 6. However, to get spikes $|V_{SG}|$ should be greater than 0.2 V and $V_{DG}$ must be above breakdown voltage (i.e., 1.68 V). For $V_{SG}$ = -0.22 V and $V_{DG}$ = 2.0 V, the drain current attains $I_{th}$ and as it exceeds $I_{th}$, a spike is fired followed by forced reset i.e., $V_{DG}$ is grounded for 700 ps to flush the stored holes from the floating-body. Now again $V_{DG}$ is set back to 2 V for 140 ps and by doing this the current again starts to integrate and fires an identical spike, as shown in Fig. 6(a). As a result, train of spikes can be generated after every 840 ps that dictates the spiking frequency. The number of spikes increases as we increase the $|V_{SG}|$, which can be seen in Fig. 6(b-d). This increase in number of spikes is due to faster accumulation of excess holes in floating-body with increases in $|V_{SG}|$ and corresponding to that the $I_{th}$ reaches $I_{th}$ faster. Hence we can achieve higher frequency of spikes with increase in $|V_{SG}|$.

As long as $|V_{SG}| < 0.2$ V zero spiking frequency ($f_o$) is observed at the output of proposed L-BIMOS silicon neuron, as shown in Fig. 7(a). This is due to the insufficient electron supply from the source region to the drain region which results in zero spikes. However, for $|V_{SG}| \geq 0.2$ V, the drain current of the device achieves the threshold current i.e., $I_{th} = 0.5$ mA/um, results generation of spikes. As we increases the $|V_{SG}|$ beyond 0.2 V, the $f_o$ increases monotonically, as shown in Fig. 7(a). Thus, the proposed device offers the signature of a biological neuron, i.e., for $|V_{SG}| < 0.2$ V the device gives zero $f_o$ whereas for $|V_{SG}| \geq 0.2$ V $f_o$ increases linearly.

We have also compared the performance of the proposed device with the previous published CMOS process compatible works, as mentioned in Table-II. The CMOS circuit based neuron [1, 2] consist of leaky integrate circuit which increases the complexity of SNN as well as the energy consumption per spikes. So, in this work we have replaced the leaky integration function of neuron with the proposed L-BIMOS silicon neuron. One can see that the proposed L-BIMOS silicon neuron requires 0.2 V for firing a spike, which is 60 mV less when compared to previous reported LIF neuron [7] and hence makes the proposed LIF neuron energy efficient. For estimating the energy per spike, we have adopted the following expression [16]:

$$E_{spike} = V_{spike} \times t_{spike} \times I_{th}$$  \hspace{1cm} (1)

| References | Device Type | $V_{th}$ | $V_{sd}$ | Energy/spike |
|------------|-------------|---------|---------|-------------|
| 1          | CMOS        | -       | >1 V    | 900 pJ      |
| 2          | CMOS        | -       | -       | 8.5-9 pJ    |
| 3          | CMOS        | -       | -       | 41.3 pJ     |
| 4          | PCM         | -       | -       | 5 pJ        |
| 5          | SOI MOS     | 100 nm  | 0.26 V  | 35 pJ       |
| 6          | PCMO RRAM   | -       | -       | 4.8 pJ      |
| This Work  | L-BIMOS     | 90 nm   | 0.2 V   | 0.18 pJ     |

V. CONCLUSION

In this paper, we have demonstrated the functioning of LIF neuron based on CMOS process compatible L shaped gate bipolar I-MOS (L-BIMOS). The proposed device exhibits...
reduced breakdown voltage $V_{BS} = 1.68$ V, and consumes $0.18$ pJ/spike energy that makes it $194 \times$ more efficient in energy in contrast to its counterpart PD-SOI silicon neuron for same functionality. Further, we manifest that the proposed L-BIMOS silicon neuron offers the signature of a biological neuron that can be exploited for the complex spiking neural networks (SNN). The spiking frequency increases linearly with increase in $\left| V_{SG} \right|$ above threshold voltage, and the proposed device exhibits higher spiking frequency of the order of GHz at a lower supply voltage as compared to PD-SOI MOS silicon neuron (MHz). Thus, L-BIMOS silicon neuron has a potential for realization of massive parallel processing required for SNN.

**REFERENCES**

[1] G. Indiveri, E. Chicca, and R. Douglas, “A VLSI Array of Low-Power Spiking Neurons and Bistable Synapses With Spike-Timing Dependent Plasticity,” *IEEE Trans. Neural Networks*, vol. 17, no. 1, January 2006.

[2] J. V. Arthur and K. A. Boahen, “Silicon-Neuron Design: A Dynamical Systems Approach,” *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 58, no. 5, pp. 1034-1043, May 2011.

[3] J. Wijekoon and P. Dudek, “Compact silicon neuron circuit with spiking and bursting behaviour,” *Neural Networks*, doi:10.1016/j.neunet.2007.12.037, 2008 page no.:524-534.

[4] A. Joubert, B. Belhadj, O. Temam and R. Hliot, “Hardware spiking neurons design: Analog or digital?,” The 2012 International Joint Conference on Neural Networks (IJCNN), Brisbane, QLD, 2012.

[5] P. Merolla, J. Arthur, F. Akopyan, N. Imam, R. Manohar and D. S. Modha, “A digital neuromorphic core using embedded crossbar memory with 45pl per spike in 45nm,” *IEEE Custom Integrated Circuits Conference (CICC)*, San Jose, CA, 2011.

[6] R. Emery, A. Yakovlev and G. Chester, “Connection-centric network for spiking neural networks,” *ACM/IEEE International Symposium on Networks-on-Chip*, San Diego, CA, pp. 144-152, 2009.

[7] S. Dutta, V. Kumar, A. Shukla, N. R. Mohapatra, and Udayan Ganguly, “Leaky Integrate and Fire Neuron by Charge-Discharge Dynamics in Floating-Body MOSFET,” *Scientific reports*, August 2017.

[8] T. Chavan, S. Dutta, N.R. Mohapatra and U. Ganguly, “An Ultra Energy Efficient Neuron enabled by Tunneling in Sub-threshold Regime on a Highly Manufacturable 32 nm SOI CMOS Technology,” 76th Device Research Conference (DRC). June 2018.

[9] A. Wei, M. J. Sherony, and D. A. Antoniadis, “Transient Behavior of the Kink Effect in Partially-Depleted SO1 MOSFET,” *IEEE Electron Device Letters*, vol. 16, no. 11, November 1995.

[10] M. J. Kumar, M. Maheedhar, and P. P. Varma, “Bipolar I-MOS An Impact-ionization MOS with Reduced Operating Voltage using the Open Base BJT Configuration,” *IEEE Trans. Electron Devices*, vol. 62, no. 12, pp. 4345-4348, Dec. 2015.

[11] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, “Impact ionization MOS (I-MOS) Part I: Device and circuit simulations,” *IEEE Trans. Electron Devices*, vol. 52, no. 1, pp. 6976, Jan. 2005.

[12] A. Lahgere and M. J. Kumar, “1-T Capacitorless DRAM Using Bandgap-Engineered SiliconGermanium Bipolar I-MOS,” *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1583-1590, April 2017.

[13] Y. Tsividis and C. McAndrew, “Operation and Modeling of the MOS Transistor,” Vol. 2. Oxford: Oxford university press, 1999.

[14] T. Tuma, A. Pantazi, M. Le Gallo, A. Sebastian, and E. Eleftheriou, “Stochastic phase-change neurons,” *Nature Nanotechnology*, vol. 11, p. 693, May 2016.

[15] S. Lashkare, S. Chouhan, T. Chavan, A. Bhat, P. Kumbhare and U. Ganguly, “PCMO RRAM for Integrate-and-Fire Neuron in Spiking Neural Networks,” *IEEE Electron Device Letters*, vol. 39, no. 4, pp. 484-487, April 2018.

[16] B. Das, J. Schulze and U. Ganguly, “Ultra-Low Energy LIF Neuron Using Si NIPIN Diode for Spiking Neural Networks,” *IEEE Electron Device Letters*, vol. 39, no. 12, pp. 1832-1835, Dec. 2018.