

Desarrollo de sinapsis memristiva en una red neuronal pulsante implementada con transistores CMOS del nodo tecnológico SKY130

1st Ricardo Aldair Tirado Torres
Centro de Investigación en Computación
Instituto Politécnico Nacional
Ciudad de México, México
rtiradot2023@cic.ipn.mx

2nd Ricardo Barrón Fernández
Centro de Investigación en Computación
Instituto Politécnico Nacional
Ciudad de México, México
rbarron@cic.ipn.mx

Resumen—El cómputo neuromórfico, con sus redes neuronales pulsantes, el aprendizaje STDP y los memristores, representa un paradigma revolucionario en la informática. La integración de estos conceptos en la tecnología CMOS no solo promete un salto significativo en la eficiencia y capacidad de las máquinas, sino que también nos acerca un paso más a replicar la asombrosa complejidad y eficiencia del cerebro humano. Esta tecnología abre un abanico de posibilidades en el campo de la inteligencia artificial y más allá, con el potencial de transformar la forma en que interactuamos con las máquinas y el mundo que nos rodea. En este trabajo se utiliza al nodo tecnológico de SKY130, un proceso de fabricación que integra los elementos necesarios para diseñar una red neuronal pulsante que integra el concepto del aprendizaje STDP no supervisado.

Index Terms—Redes neuronales pulsantes, sinapsis memristiva, plasticidad dependiente del tiempo de pulso (STDP), SKY130.

I. INTRODUCCIÓN

El cómputo neuromórfico es un campo emergente en la ingeniería y la informática que se inspira en la arquitectura y funcionamiento del cerebro humano para diseñar sistemas de procesamiento de información más eficientes y adaptativos. A diferencia de las computadoras tradicionales, que siguen el modelo de von Neumann, los sistemas neuromórficos emplean una arquitectura distribuida y paralela, emulando las redes neuronales del cerebro. Esta aproximación no solo promete una mayor eficiencia energética y velocidad de procesamiento [?], sino también una capacidad de aprendizaje y adaptación superior.

En el corazón del cómputo neuromórfico se encuentran las redes neuronales pulsantes, que son modelos de redes neuronales que procesan información en forma de pulsos de actividad eléctrica, o pulsos. Estos modelos son más biológicamente realistas que las redes neuronales tradicionales utilizadas en el aprendizaje profundo. Las redes neuronales pulsantes permiten un procesamiento temporal y espacial más natural de la infor-

mación, similar al que ocurre en el cerebro, lo que resulta en un rendimiento superior en tareas que requieren reconocimiento de patrones y procesamiento sensorial en tiempo real.

El aprendizaje dependiente del tiempo de los pulsos (STDP, por sus siglas en inglés) es un mecanismo clave en las redes neuronales pulsantes. El STDP es una forma de aprendizaje sináptico que ajusta la fuerza de las conexiones entre neuronas en función del momento en que ocurren los pulsos. Si una neurona emisora envía un pulso justo antes de que una neurona receptora envíe el suyo, la conexión entre ambas se fortalece. Este principio, que se observa en las sinapsis biológicas, permite a las redes neuromórficas aprender y adaptarse de manera más eficiente, logrando un aprendizaje no supervisado y dinámico. Esta regla temporal puede ser implementada directamente en unos elementos de memoria, llamados memristores, debido a su capacidad para cambiar su resistencia en función de la historia temporal de los pulsos que los atraviesan [?] [?].

El memristor, un componente eléctrico cuya resistencia varía en función de la cantidad y la dirección de la carga eléctrica que lo atraviesa, ha emergido como una pieza fundamental en el campo del cómputo neuromórfico. Descubierta teóricamente por Leon Chua en 1971 [?] y desarrollada prácticamente en los laboratorios de HP en 2008, el memristor es a menudo considerado el cuarto elemento de los circuitos pasivos junto con el resistor, el capacitor y el inductor. Su capacidad para recordar.^{el} último estado de resistencia hace que el memristor sea especialmente adecuado para imitar las sinapsis biológicas en sistemas neuromórficos.

En el contexto del cómputo neuromórfico, los memristores se utilizan como sinapsis artificiales que conectan las neuronas en redes neuronales pulsantes. Estas redes, que procesan información en forma de pulsos de actividad eléctrica similares a los impulsos neuronales del cerebro, se benefician enormemente de la naturaleza adaptable y no volátil del memristor [?]. La resistencia del memristor cambia en respuesta a

los pulsos eléctricos, permitiendo la modulación sináptica de manera analógica, lo cual es crucial para replicar los procesos de aprendizaje y memoria del cerebro humano [?].

Para implementar estos conceptos, se utiliza la tecnología CMOS (Complementary Metal-Oxide-Semiconductor), ampliamente conocida por su uso en la fabricación de chips de computadora tradicionales. La tecnología CMOS ofrece una plataforma robusta y escalable para desarrollar hardware neuromórfico, facilitando la creación de sistemas que son tanto energéticamente eficientes como capaces de realizar cálculos complejos en paralelo [?]. La integración de memristores en la tecnología CMOS ha permitido avances significativos en la construcción de circuitos neuromórficos eficientes y escalables. Los memristores, debido a su tamaño reducido y bajo consumo de energía, se pueden integrar en matrices densas junto con transistores CMOS para crear chips neuromórficos que emulan el funcionamiento del cerebro humano de manera más realista y eficiente. Estos chips, como los desarrollados en proyectos como IBM TrueNorth e Intel Loihi [?], incorporan miles de millones de sinapsis memristivas capaces de aprendizaje autónomo y procesamiento paralelo, replicando la complejidad y eficiencia del cerebro. A medida que la tecnología CMOS sigue evolucionando, se espera que las capacidades del cómputo neuromórfico se expandan, abriendo nuevas posibilidades en la inteligencia artificial y más allá.

El cómputo neuromórfico, con sus redes neuronales pulsantes, el aprendizaje STDP y la reciente fabricación de memristores representa una revolución en cómo concebimos y desarrollamos sistemas de procesamiento de información. La integración de estos conceptos en la tecnología CMOS no solo promete un salto significativo en la eficiencia y capacidad de las máquinas, sino que también nos acerca un paso más a replicar la asombrosa complejidad y eficiencia del cerebro humano. Es por ello que, en este trabajo, se realizaron las siguientes contribuciones:

1. Un circuito de neurona tipo LIF, implementada en el dominio analógico en un nivel de transistores CMOS. A través de una señal de corriente, la neurona realizará un pulso a su salida si se supera un cierto umbral y entrará en un periodo refractario.
2. Una sinapsis del tipo memristiva, que puede cambiar su resistencia en respuesta a las señales eléctricas, lo que permite imitar la plasticidad sináptica del cerebro humano. Esta capacidad de adaptación es crucial para el aprendizaje y la memoria en sistemas neuromórficos.
3. Una red neuronal sencilla, con una capa de entrada de 2 neuronas y una capa de salida de 1 neurona. El nodo tecnológico empleado es el de SKY130, un proceso de fabricación de código abierto, que reduce costos asociados con el diseño y producción de chips, facilitando la creación de sistemas neuromórficos avanzados.

II. SISTEMA PROPUESTO

II-A. Neurona LIF

El circuito de una neurona LIF comprende varias secciones, como integración de corriente, fuga, generación de pulsos, reinicio del potencial de membrana y período refractario.

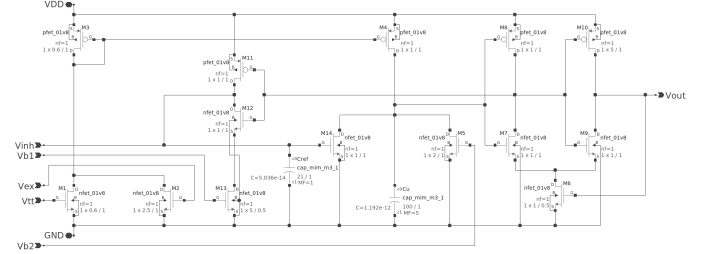


Figura 1: Diagrama esquemático del circuito de neurona implementado por [1]. Consta de varias secciones para integración de corriente, fugas, generación de pulsos, reinicio y período refractario.

La Figura 1 muestra el circuito neuronal LIF propuesto originalmente, que incluye un condensador C_u para la integración de corriente, un transistor M_5 para fugas, un disparador Schmitt para la generación de pulsos, $M_{11} - M_{14}$ para reinicio y un condensador C_{ref} para el período refractario. Las terminales V_{tt} y V_{ex} se utilizan para aplicar un voltaje de entrada a la neurona, ya sea los proporcionados desde las sinapsis de retroalimentación o desde las sinapsis excitadoras laterales. La terminal inhibidora lateral V_{inh} se utiliza para restablecer el potencial de membrana. Los transistores M_1 y M_2 controlan el rango de la corriente de entrada y el transistor M_4 controla la corriente inyectada en C_u . Además, la capacitancia de C_u y C_{ref} se puede alterar para controlar la tasa de pulsos de salida. Cuando se aplica el voltaje de entrada a la neurona, la corriente I_{in} se inyecta en la sección del integrador con fugas a través del espejo de corriente. Luego, la corriente inyectada se integra con C_u y se escapa a través de M_5 . A través de un disparador Schmitt de baja potencia se generan los pulsos. Cuando el potencial de membrana alcanza el voltaje de conmutación del disparador Schmitt, $M_{11} - M_{14}$ intervienen para cambiar el voltaje de alto a bajo. Simultáneamente, el transistor M_{11} se enciende y el potencial de membrana se restablece a cero a través de la sección de reinicio. En consecuencia, se genera un pulso en la salida de la neurona. Además, es posible restablecer el potencial de membrana a través del terminal inhibidor lateral V_{inh} . El circuito fue diseñado originalmente para una tecnología CMOS de 90 nm.

Con el objetivo de reducir el consumo de potencia, se redujo el número de transistores, tomando como base la implementación de [?] y se modificó la etapa de inyección del corriente del circuito, para que varié con respecto a una corriente de entrada y no a un voltaje. La implementación final, vista en la Figura 2, resultó en un circuito con 9 transistores

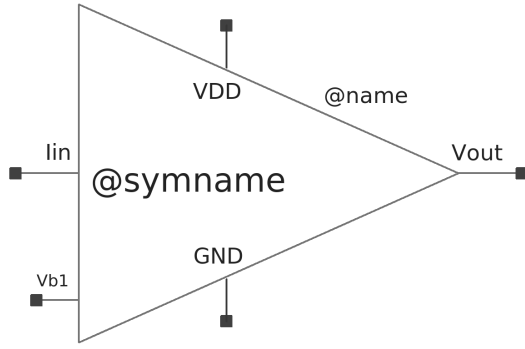


Figura 2: Neurona LIF modificada.

(5 menos que el modelo original) y únicamente 2 entradas: La corriente de entrada a la neurona, I_{in} y un voltaje de polarización V_{b1} , capaz de modificar la frecuencia de pulso de la neurona. De igual forma, se cambiaron los parámetros de diseño originales, para adaptar el modelo final al nodo tecnológico de SKY130, cuyo ancho de canal mínimo es de 130nm.

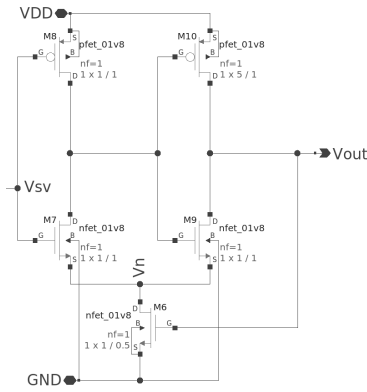


Figura 3: Diagrama esquemático de la sección correspondiente al disparador Schmitt.

El funcionamiento de la neurona resulta muy parecido al original, por lo que el voltaje de conmutación, V_{SV} , del disparador Schmitt (Ver Figura 3) es el umbral de activación de la neurona [?] y se obtiene de la siguiente manera:

$$I_{D6} = \beta_6 (V_n - V_{TN})^2 \quad (1)$$

$$I_{D7} = \beta_7 (V_{DD} - V_{SV} - V_{TP})^2 \quad (2)$$

$$I_{D8} = \beta_8 (V_{SV} - V_n - V_{TN})^2 \quad (3)$$

En donde V_{TN} y V_{TP} son los voltajes de umbral de los transistores de canal N y P, respectivamente y β_N es el parámetro de transconductancia que equivale a $\mu C_{OX} \frac{W}{L}$. Igualando la ecuación 2 y 3 para despejar a V_{SV} , se obtiene:

$$V_{SV} = \frac{R_{7/8}(V_{DD} - V_{TP}) + V_n + V_{TN}}{R_{7/8} + 1} \quad (4)$$

En donde $R_{7/8}$ representa a $\sqrt{\frac{\beta_7}{\beta_8}}$. Igualando la ecuación 1 y 3 para despejar a V_n , se obtiene:

$$V_n = \frac{V_{SV} + (R_{6/8} - 1)V_{TN}}{R_{6/8} + 1} \quad (5)$$

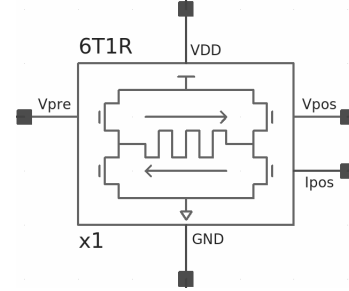


Figura 4: Sinapsis memristiva en configuración 6T1R.

En donde $R_{6/8}$ representa a $\sqrt{\frac{\beta_6}{\beta_8}}$. Sustituyendo la ecuación 5 en 4, se obtiene:

$$V_{SV} = \frac{R_{7/8}(V_{DD} - V_{TP}) + \left(\frac{R_{6/8} - 1}{R_{6/8} + 1} + 1 \right) V_{TN}}{(R_{7/8} + 1) - \frac{1}{R_{6/8} + 1}} \quad (6)$$

II-B. Sinapsis memristiva

La configuración 6T1R, empleada en la sinapsis memristiva de la Figura 4, fue tomada en base a la idea planteada en [2], donde se implementan 4 transistores en configuración de puente H para controlar el sentido de la corriente que pasa por un memristor. El modelo de simulación pertenece a la tecnología SKY130 de [3][4], pero los parámetros del dispositivo son los tomados por [5][6]. Este memristor tiene voltajes de umbral de $V_{ON} \approx 0,9V$ y $V_{OFF} \approx -0,6V$. El voltaje que consumen los transistores entre la terminal de drenador y de fuente es de $V_{Dsat} \approx 0,35V$, lo cual proporciona una suficiente alimentación para que exista el cambio en la resistencia del memristor, debido a que se está utilizando un $V_{DD} = 1,8V$.

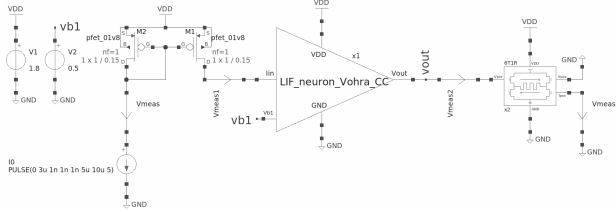
La sinapsis memristiva cumple un rol importante en la red neuronal pulsante, ya que el aprendizaje dependiente del tiempo de los pulsos puede ser implementado de manera efectiva utilizando memristores. En STDP, la fuerza de una sinapsis se ajusta en función de la sincronización relativa de los pulsos pre y post-sinápticos. Si un pulso en la neurona presináptica precede a un pulso en la neurona postsináptica,

la sinapsis se fortalece (potenciación). Por el contrario, si el pulso presináptico sigue al postsináptico, la sinapsis se debilita (supresión). Esta regla temporal puede ser implementada directamente en memristores debido a su capacidad para cambiar su resistencia en función de la historia temporal de los pulsos que los atraviesan. El modelo básico de la regla STDP esta determinado por la ecuación 7:

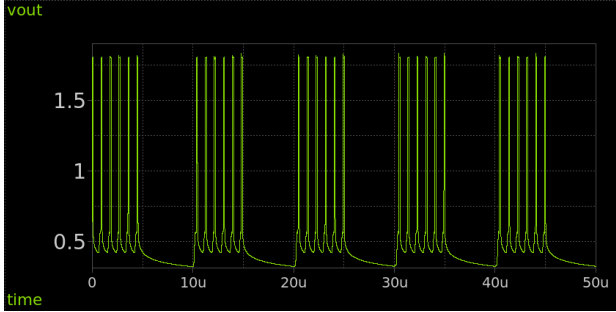
$$\Delta w \begin{cases} A^+ e^{-\Delta t / \tau^+}, & \text{si } \Delta t > 0 \\ -A^- e^{\Delta t / \tau^-}, & \text{si } \Delta t < 0 \end{cases} \quad (7)$$

III. RESULTADOS DE LA SIMULACIÓN

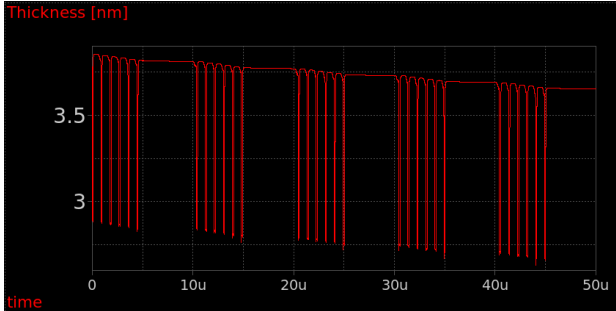
III-A. Simulación de una neurona y una sinapsis



(a) Circuito de una neurona LIF cuya salida esta conectada una sinapsis memristiva.



(b) Generación de pulsos a la salida de la neurona LIF.



(c) Cambio de la anchura del filamento del memristor.

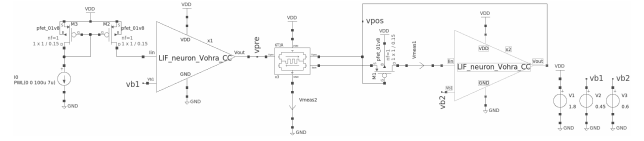
Figura 5: Descripción general de las dos imágenes.

Para el primer caso de simulación, se utilizó una neurona LIF conectada en su entrada a un espejo de corriente que suministra la señal de entrada. A la salida de la neurona se conecta la sinapsis de tal forma que se debe modificar el peso

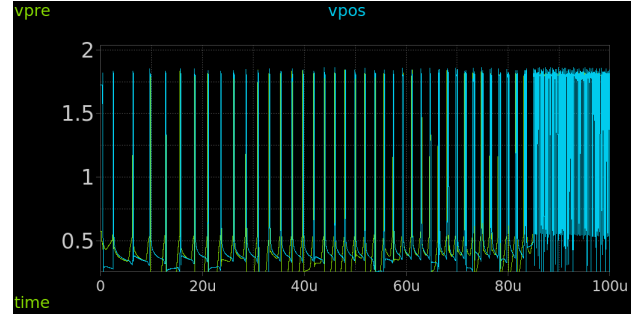
sináptico en forma de supresión. En la Figura 5a se muestra el circuito simulado.

En la Figura 5b se observa que la neurona funciona de manera correcta, generando trenes de pulsos, únicamente cuando se polariza la entrada. En la Figura 5c se visualiza el cambio del ancho del filamento en el memristor. Como este filamento esta disminuyendo conforme recibe pulsos por parte de la neurona, se genera un cambio en la resistencia del memristor, haciendo que aumente y por consecuencia, disminuya el peso sináptico, provocando una supresión, de acuerdo con la regla STDP.

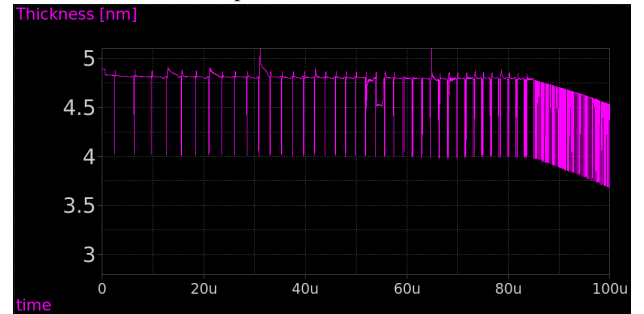
III-B. Simulación de dos neuronas y una sinapsis



(a) Circuito de una neurona LIF cuya salida esta conectada una sinapsis memristiva.



(b) Generación de pulsos a la salida de la neurona LIF.



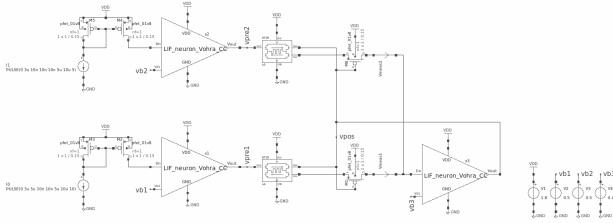
(c) Cambio de la anchura del filamento del memristor.

Figura 6: Descripción general de las dos imágenes.

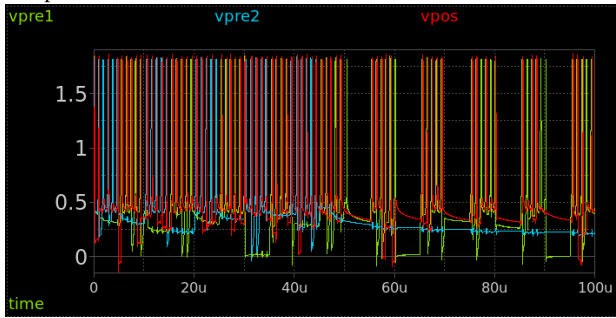
Para el segundo caso de simulación, se utilizaron dos neuronas LIF interconectadas entre sí por medio de una sinapsis memristiva. Se observa en la Figura 6a que la neurona presináptica esta conectada de la misma forma que en el primer caso, no obstante, la neurona postsináptica requiere de un transistor PMOS entre su entrada y la salida de la sinapsis, debido a que solo recibirá una señal de excitación cuando la propia neurona no pulse, si sucede un pulso a su salida, se debe aislar a la entrada para evitar oscilación en el circuito.

En la Figura 6b se observa que los pulsos entre la neurona presináptica y postsináptica suceden de manera similar, salvo en los últimos momentos de la simulación en donde la primera neurona pulsa más rápido que la segunda, provocando que la estabilidad que había en el filamento del memristor cambie, provocando que el peso sináptico disminuya y por lo tanto, la resistencia aumente, como se observa en la Figura 6c.

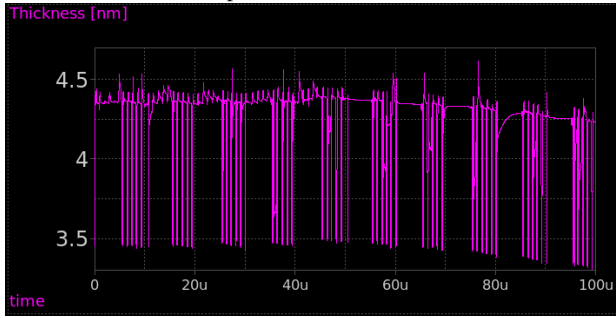
III-C. Simulación de tres neuronas y dos sinapsis



(a) Circuito de una neurona LIF cuya salida esta conectada una sinapsis memristiva.



(b) Generación de pulsos a la salida de la neurona LIF.



(c) Cambio de la anchura del filamento del memristor.

Figura 7: Descripción general de las dos imágenes.

Para el último caso de simulación, se utilizaron tres neuronas LIF, dos en la capa de entrada y una en la capa de salida. Como se observa en la Figura 7a, al ser dos conexiones sinápticas, se requieren de dos sinapsis, junto con dos transistores PMOS, para evitar el fenómeno de oscilación, previamente mencionado.

En la Figura 6b se observa que para esta prueba se generó solo la mitad de pulsos en la segunda neurona presináptica para observar que la neurona postsináptica depende de esta excitación para pulsar de manera más frecuente. En la Figura

6c se visualiza que los pulsos entre las neuronas mantienen un equilibrio en el filamento, sin embargo, cuando termina de pulsar la segunda neurona, los pulsos de la neurona postsináptica no son suficientes, provocando que la resistencia aumente y el ancho del filamento promedio disminuya junto con los pesos sinápticos.

En el Cuadro I, se muestran los parámetros utilizados para la implementación de las neuronas y de las sinapsis de este trabajo.

Cuadro I: Parámetros $\frac{W}{L}$ del sistema

Neurona		Sinapsis	
Parámetro	Valor	Parámetro	Valor
W_1	$0,2\mu m$	W_1	$5\mu m$
W_2	$0,2\mu m$	W_2	$5\mu m$
W_3	$1\mu m$	W_3	$1,5\mu m$
W_4	$0,2\mu m$	W_4	$15\mu m$
W_5	$0,2\mu m$	W_5	$1,5\mu m$
W_6	$0,2\mu m$	W_6	$15\mu m$
W_7	$0,2\mu m$		
W_8	$0,2\mu m$		
W_9	$1\mu m$		

* $L_{min} = 0,1\mu m$ en ambos circuitos.

CONCLUSIÓN

Este trabajo presenta una implementación de neuronas LIF interconectadas entre sí, a través de un modelo de sinapsis 6T1R, cuyo componente principal era el memristor, el cual proporcionó diferentes pesos sinápticos de acuerdo con la regla de aprendizaje STDP. La construcción del sistema se hizo por medio de la herramienta de diseño esquemático Xschem y la simulación con NGSPICE. Cabe señalar que se utilizó la tecnología de SKY130, desarrollado por SkyWater Technology en colaboración con Google, que es un proceso de fabricación de semiconductores que ha sido puesto a disposición de la comunidad de diseño de circuitos de manera abierta y gratuita. Esto facilita la innovación y reduce significativamente los costos asociados con el diseño y la producción de chips, haciendo más accesible la experimentación y el desarrollo de nuevas tecnologías. Dentro de los elementos proporcionados por este nodo tecnológico, se encuentra el memristor, una tecnología clave que facilita la implementación de sinapsis artificiales en sistemas neuromórficos, soportando redes neuronales pulsantes y el aprendizaje STDP. Su capacidad para cambiar de estado de manera no volátil y su integración en la tecnología CMOS permiten la creación de sistemas de cómputo avanzados que no solo son más eficientes energéticamente, sino también más apto de emular las capacidades adaptativas y de aprendizaje del cerebro humano. La convergencia de estas tecnologías promete revolucionar el campo de la inteligencia artificial, llevando el rendimiento y la eficiencia del cómputo a nuevos niveles.

REFERENCIAS

- [1] J. Shamsi, K. Mohammadi, and S. B. Shokouhi, “A hardware architecture for columnar-organized memory based on cmos neuron and memristor crossbar arrays,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 12, pp. 2795–2805, 2018.
- [2] M. Tian, J. Lu, H. Gao, H. Wang, J. Yu, and C. Shi, “A lightweight spiking gan model for memristor-centric silicon circuit with on-chip reinforcement adversarial learning,” in *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2022, pp. 3388–3392.
- [3] “Skywater sky130 pdk,” <https://sky130-fd-pr-reram.readthedocs.io/en/latest/>, [Accessed 27-05-2024].
- [4] “skywater-pdk,” [https://github.com/google/skywater-pdk-libs-sky130_{fd_pr_reram}](https://github.com/google/skywater-pdk-libs-sky130_fd_pr_reram), [Accessed 27 – 05 – 2024].
- [5] T. Wang and J. Roychowdhury, “Well-posed models of memristive devices,” 05 2016.
- [6] A. Alshaya, Q. Han, and C. Papavassiliou, “Rram, device, model and memory,” in *2022 International Conference on Microelectronics (ICM)*, 2022, pp. 117–120.