



OPEN

# Enhanced read resolution in reconfigurable memristive synapses for Spiking Neural Networks

Hritom Das<sup>✉</sup>, Catherine Schuman, Nishith N. Chakraborty & Garrett S. Rose

The synapse is a key element circuit in any memristor-based neuromorphic computing system. A memristor is a two-terminal analog memory device. Memristive synapses suffer from various challenges including high voltage, *SET* or *RESET* failure, and *READ* margin issues that can degrade the distinguishability of stored weights. Enhancing *READ* resolution is very important to improving the reliability of memristive synapses. Usually, the *READ* resolution is very small for a memristive synapse with a 4-bit data precision. This work considers a step-by-step analysis to enhance the *READ* current resolution or the read current difference between two resistance levels for a current-controlled memristor-based synapse. An empirical model is used to characterize the HfO<sub>2</sub> based memristive device. 1st and 2nd stage device of our proposed synapse design can be scaled to enhance the *READ* current margin up to ~4.3x and ~21%, respectively. Moreover, *READ* current resolution can be enhanced with run-time adaptation techniques such as *READ* voltage scaling and body biasing. The *READ* voltage scaling and body biasing can improve the *READ* current resolution by about 46% and 15%, respectively. TENNLab's neuromorphic computing framework is leveraged to evaluate the effect of *READ* current resolution on classification, control, and reservoir computing applications. Higher *READ* current resolution shows better accuracy than lower resolution even when facing different levels of read noise.

**Keywords** Current-controlled, Low power, Stochastic computing, Approximate computing, Spiking Neural Network

The development of Artificial Neural Networks (ANNs) and Deep Neural Networks (DNNs) is inspired by the remarkable information processing abilities of mammalian brains, while also achieving low power consumption and minimal latency. Due to their exceptional classification accuracy, DNNs are attracting considerable interest as the preferred classifier in numerous machine learning and computer vision applications<sup>1</sup>. However, DNNs are typically executed on Von Neumann machines and are hence limited by the separation of memory and processing units, also known as the von Neumann bottleneck<sup>2</sup>. Besides that, the extensive computational requirements, high power consumption, and memory bandwidth associated with DNNs make them less attractive for mobile applications, where limitations in area and power are significant constraints<sup>3</sup>.

In order to address the limitations in power and memory capacities of traditional computing architectures mentioned above, coupled with further inspiration drawn from the efficiency of the biological nervous system, a novel concept of neuromorphic architectures has emerged. These architectures typically employ Spiking Neural Networks (SNNs), which aim to mimic the intricacies of the biological nervous system with greater fidelity by employing binary pulses as a means of communication. These architectures represent a distinct paradigm from the conventional Von Neumann architecture in terms of the co-location of memory and processing unit, demonstrating promising results in specific application domains<sup>4-6</sup>. Neuromorphic architectures not only offer better energy efficiency, but also promise parallel signal processing, fault tolerance, and reconfigurability. Furthermore, they can be implemented using diverse silicon-based technologies, large-scale architectures, and computational models of neural components<sup>7-10</sup>.

Neuromorphic systems, also referred to as Neuroprocessors, leverage the co-location of memory and processing units, where neurons act as the computational units, interconnected by synaptic memory elements. Synapses

Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA.  
<sup>✉</sup>email: hdas@utk.edu

contain the weighted connections between neurons and can be implemented using digital<sup>11,12</sup> or analog<sup>13–19</sup> circuits. Nevertheless, incorporating a multitude of synapses presents several obstacles, including the efficient handling of storage space needed for weighted connections and the accommodation of diverse synaptic learning techniques that demand adaptable weight storage<sup>9,20</sup>. Memristors are potential candidates to address these issues. Memristors have proven to be more compact and power efficient for synaptic implementation compared to SRAM<sup>21</sup> and capacitor-based implementations of the same resolutions<sup>9,22</sup>. Memristive synapses have also been shown to have extended memory retention time<sup>22–24</sup>.

First postulated by Leon Chua, the memristor is described as the fourth fundamental passive circuit element<sup>25</sup>. A memristor is a two-terminal device with analog memory properties that originate from its ability to switch resistance levels. When a voltage beyond a certain threshold is applied to its terminals, the resistance is modified. A memristor resistance state is also non-volatile making it a promising candidate for weight storage. Due to their compatibility with CMOS technology and non-volatile properties, memristors are well-suited for analog computation<sup>18,26</sup>.

In this work, we use a current-controlled synapse designed using a TiN TE (TE = Top Electrode)/HfO<sub>2</sub>/TiN BE (BE = Bottom Electrode) memristor<sup>27</sup>. To use this synapse, several operations need to be performed on the memristor, such as *FORMING*, *RESET*, *SET*, and *READ*. This memristor can vary its resistance in a range of a few k $\Omega$  to over 150 k $\Omega$ <sup>13,14</sup>. However, for our design, we exclusively use the low resistance states (LRS) of the memristors due to improved reliability in this operating region, and to avoid a high degree of variability encountered near the high resistance states (HRS)<sup>28</sup>. The memristor can be programmed into different low-resistance states by precisely controlling the compliance current during the *SET* operation that overcomes the issues of variability and limited resolution<sup>29</sup>.

Although the synapse is designed while taking the reliability concerns into consideration, the synapse is vulnerable to limited resolution due to the use of a narrow resistance range in the low-resistance regions<sup>30</sup>. This occurs when the current generated by the different resistance states is not easily differentiable. This limitation can cause a reduction in the learning performance of SNNs<sup>9,30</sup>. Overlapping synaptic currents also make the synapse susceptible to noise and process variation<sup>9</sup>. Another disadvantage of the small difference between synaptic states is that it complicates the analog-to-digital converter (ADC) design significantly<sup>9</sup>. For the ADC to recognize different resistance states for digital conversion, the current output difference between the resistance states needs to be high enough for a compact, yet power efficient design<sup>9</sup>. This work aims at improving the current resolution of the memristive synaptic circuit using several techniques.

The key contributions of this paper are as follows.

1. *READ* current resolution of a current compliance memristive synapse is enhanced,
2. *READ* current resolution is enhanced with proper device scaling,
3. *READ* current resolution is made re-configurable at run time with *READ* voltage scaling,
4. *READ* current resolution is adaptable at run time with body biasing, and
5. the TENNLab neuromorphic software framework<sup>31</sup> is utilized to observe the effect of *READ* current resolution on SNNs. Higher *READ* current resolution illustrates better accuracy with a lower possibility of a read error.

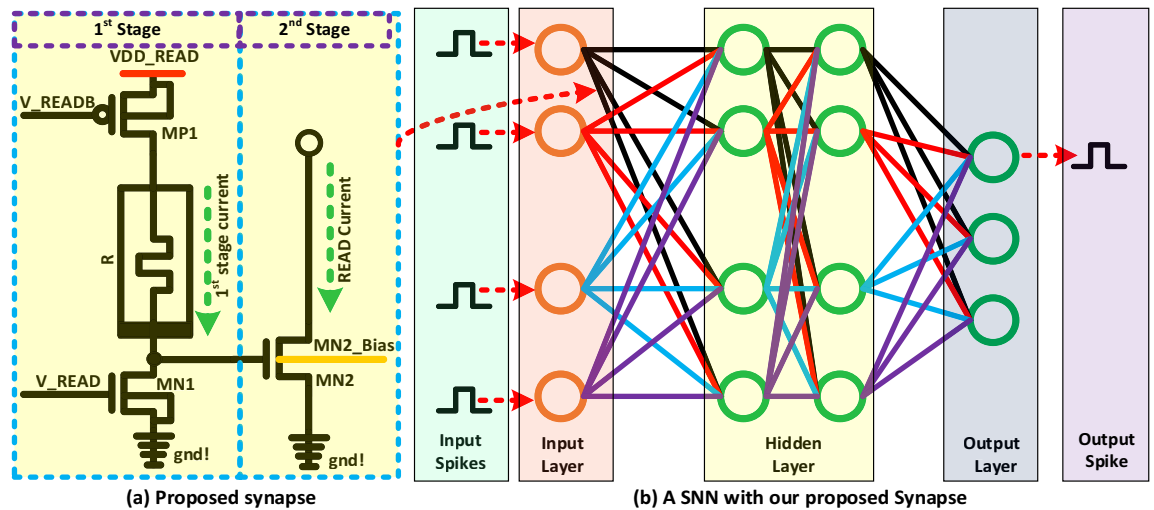
The remainder of this paper is organized as follows. The following section briefly describes a Verilog-A model for the hafnium oxide-based memristor device, the synaptic circuit built from this memristor, and a description of its *READ* operation. The next section shows the proper device sizing to enhance the *READ* current resolution. After that, a section illustrates two techniques for improving *READ* current resolution re-configurable at run time. The next section will evaluate the design performance based on different test cases. The next section exhibits the effect of the *READ* current resolution or weight resolution on Spiking Neural Networks (SNNs). A detailed comparison with prior works is analyzed in the next section. Finally, the paper is concluded with prospective future work.

## Current-controlled memristive synapse HfO<sub>2</sub> based device modeling

A Verilog-A model is utilized to simulate the HfO<sub>2</sub> based memristive devices<sup>32</sup>. In this model, mathematical equations are derived based on the memristance state and the required time to switch the states between HRS to LRS or LRS to HRS. The I–V characteristics of this device are taken under consideration to derive the empirical model of the memristor. The threshold voltage and switching time are two sets of important measured parameters for this model. There are some fitting constants in the model, which are utilized to fit the device's I–V characteristics to observe the simulation behavior as closely as measured data. The sigmoid window function is also considered to be incorporated with different patterns of switching time between HRS to LRS and vice versa. This Verilog-A model is also capable of detecting the *RESET* failure if the *RESET* voltage has crossed its functional window. All mathematical equations, I–V curves, measured, and simulation details are available in a prior work<sup>32</sup>.

## Proposed synapse

Memristors are widely used to construct brain-inspired synapses, which is the key element for neuromorphic computing. There are different synapse flavors based on the memristor's recipe (combination of materials). Various materials are utilized to build the memristors such as HfO<sub>2</sub>, Ta<sub>2</sub>O<sub>5</sub>, NbO<sub>2</sub>, and so on. Our proposed architecture is designed using a TiN TE/HfO<sub>2</sub>/TiN BE memristor. Figure 1a and b show the proposed synapse with *READ* devices and a simple SNN based on our synapse respectively. The first step of this synapse is the one-time *FORM* operation. Thick-oxide transistors are used for this design to take care of high voltage around 3.3 V



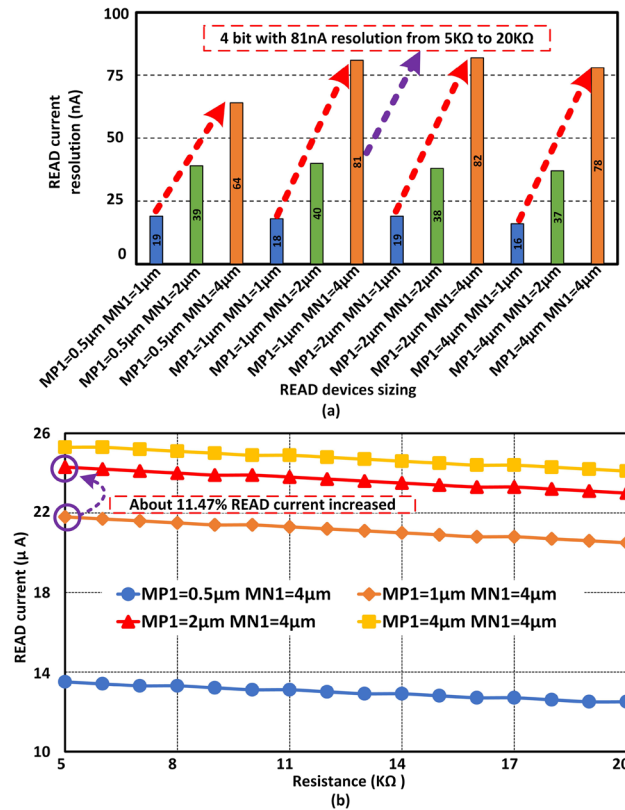
**Figure 1.** Memristor with *READ* circuitry is illustrated. (a) *READ* operation requires 1-PMOS and 2-NMOS. *READ* operation folded into two parts. 1st stage current is generated with  $M_{P1}$  and  $M_{N1}$ . This current initiates a voltage to operate  $M_{N2}$  in the linear region. Finally, the *READ* current will be sensed from the drain of  $M_{N2}$ . The body of the  $M_{N2}$  is considered as a dedicated signal to control the threshold of this device. (b) A spiking neural network (SNN) is constructed with our proposed synaptic circuitry. Input spikes are fed into the input layer of neurons. Finally, the output spike indicates the class which is determined by the network.

for forming. A thick oxide transistor is also useful to reduce flicker noise. Unformed memristors usually exhibit resistance in the range of  $\sim 8$  to  $\sim 10$  M $\Omega$ . After forming, the memristor's resistance level will be in a few k $\Omega$ . Hence, the synapse needs to *RESET* to a higher resistance state (HRS), which is typically hundreds of k $\Omega$ . Finally, our device is ready to *SET* /write /program to a specific low resistance state from an HRS. Due to less variability, the programming region is selected in the LRS region. The targeted LRS for this design is from 5 to 20 k $\Omega$ . Here, the LRS region is considered to program the synapse with a low inherent variation of HfO<sub>2</sub>-based memristor. At the same time, we are sacrificing the low-power operation by eliminating the HRS region for programming. Here, we are targeting 4-bit precision with  $\sim 1$  k $\Omega$  resistance resolution. This design needs a set of *RESET* and *SET* to program in a new LRS value. After a successful *SET* operation, the synapse is ready for a *READ* operation.

Figure 1a shows the proposed current-controlled synapse with *READ* circuitry.  $V_{READ}$  and  $V_{READB}$  signals are utilized to access the memristor (R) during a *READ* operation.  $VDD_{READ}$  and  $V_{READ}$  are 1.2 V and 0.6 V respectively during a *READ* operation. In addition, 0 V is provided to the  $V_{READB}$  node to access the memristor for a read operation. Due to the *READ* signal assertion, there will be a small 1st stage current through the memristor. This current will create a voltage to operate transistor  $M_{N2}$ . Finally, a *READ Current* is sensed from the drain of the  $M_{N2}$ . The body of this MOSFET is utilized as a separated signal to control the threshold voltage of  $M_{N2}$ . *READ* is a very sensitive operation for memristor-based synapses. Especially, the *READ* margin between two resistance levels (e.g. 5 k $\Omega$  and 6 k $\Omega$ ) needs to be good enough to read the data or weight properly. Most of the time the difference between the two resistance levels is a few nA for low-power design, which is very hard to sense properly. A research paper shows the *READ* current between 5 and 6 k $\Omega$  is 20 nA<sup>28</sup>. A few techniques can be utilized to overcome this low *READ* margin /resolution between two resistance levels. All the techniques are explained below with proper analysis.

### READ current resolution enhancement with proper device scaling

At first, *READ* device sizing is considered to observe the effect on *READ* current resolution. A 65 nm 10LPe CMOS process from IBM is utilized to construct and conduct Cadence Spectre simulations. A Verilog-A model is utilized to characterize the HfO<sub>2</sub> based memristor device<sup>32</sup>. Here, 9% of memristive variation is considered for the simulation, which is based on the testing results of memristive devices<sup>28</sup>. Figure 2a shows effect on *READ* current resolution with the sizing of  $M_{P1}$  and  $M_{N1}$ . Here, the width and length of  $M_{N2}$  are set at 0.5  $\mu\text{m}$ . In addition, the width of the  $M_{P1}$  is varied from 0.5 to 4  $\mu\text{m}$  and the width of the  $M_{N1}$  is varied from 1 to 4  $\mu\text{m}$ . In addition, the length of  $M_{P1}$  and  $M_{N1}$  (both are thick oxide transistor) are set at 0.5  $\mu\text{m}$ . When the width of both  $M_{N1}$  and  $M_{P1}$  is minimal, the *READ* current resolution (one memristive level to another, e.g. 5–6 k $\Omega$ ) is at least 19 nA. Due to an optimized read procedure, the *READ* current shows very stable resolutions compared to prior work<sup>28</sup>. In this work, a regular pfet is utilized to control the read voltage at the drain of  $M_{N2}$ . Whereas a diode-connected pfet was connected in prior work. Hence, the width of the  $M_{P1}$  set at 0.5  $\mu\text{m}$  and the width of the  $M_{N1}$  varies from 2 to 4  $\mu\text{m}$ . The *READ* current resolution is 64 nA, when the width of the  $M_{N1}$  and  $M_{P1}$  are 4  $\mu\text{m}$  and 0.5  $\mu\text{m}$  respectively. The larger size of  $M_{N1}$  and smaller size of  $M_{P1}$  allow suitable gate voltage for the  $M_{N2}$  to provide a high-resolution *READ* current. Twelve different sizing combinations are observed for *READ* current resolution. When the width of the  $M_{P1}$  and  $M_{N1}$  are 1  $\mu\text{m}$  and 4  $\mu\text{m}$ , the *READ* current resolution is about 81 nA for 4-bit data precision. According to the last test case, if the width of  $M_{P1}$  increases significantly and its size becomes the same as  $M_{N1}$ , then the *READ* current resolution does not show significant benefit on sizing. A better sizing combination is observed, when the  $M_{N1}$  and  $M_{P1}$  are not the same and  $M_{P1}$  is smaller than  $M_{N1}$ . According to



**Figure 2.** READ simulation results are illustrated based on the sizing of  $M_{P1}$  and  $M_{N1}$ . The length and width of  $M_{N2}$  are fixed at  $0.5\ \mu\text{m}$ . (a)  $M_{P1}$  is varied from  $0.5$  to  $4\ \mu\text{m}$ . In addition,  $M_{N1}$  is varied from  $1$  to  $4\ \mu\text{m}$ . Larger  $M_{N1}$  shows a higher impact on the READ current resolution. (b) Shows the READ current scale with different size of  $M_{P1}$ , when the width of the  $M_{N1}$  is fixed at  $4\ \mu\text{m}$ .

our sizing analysis, the READ resolution provides best performance when the width of  $M_{N1}$  and  $M_{P1}$  are  $4\ \mu\text{m}$  and  $1\ \mu\text{m}$  respectively. According to Fig. 2b, there is about 11.47% READ current overhead to improve only 1 nA current resolution. Next, the effect of length and width of the  $M_{N2}$  is observed, with the width of the  $M_{N1}$  and  $M_{P1}$  transistors are set at  $4\ \mu\text{m}$  and  $1\ \mu\text{m}$  respectively.

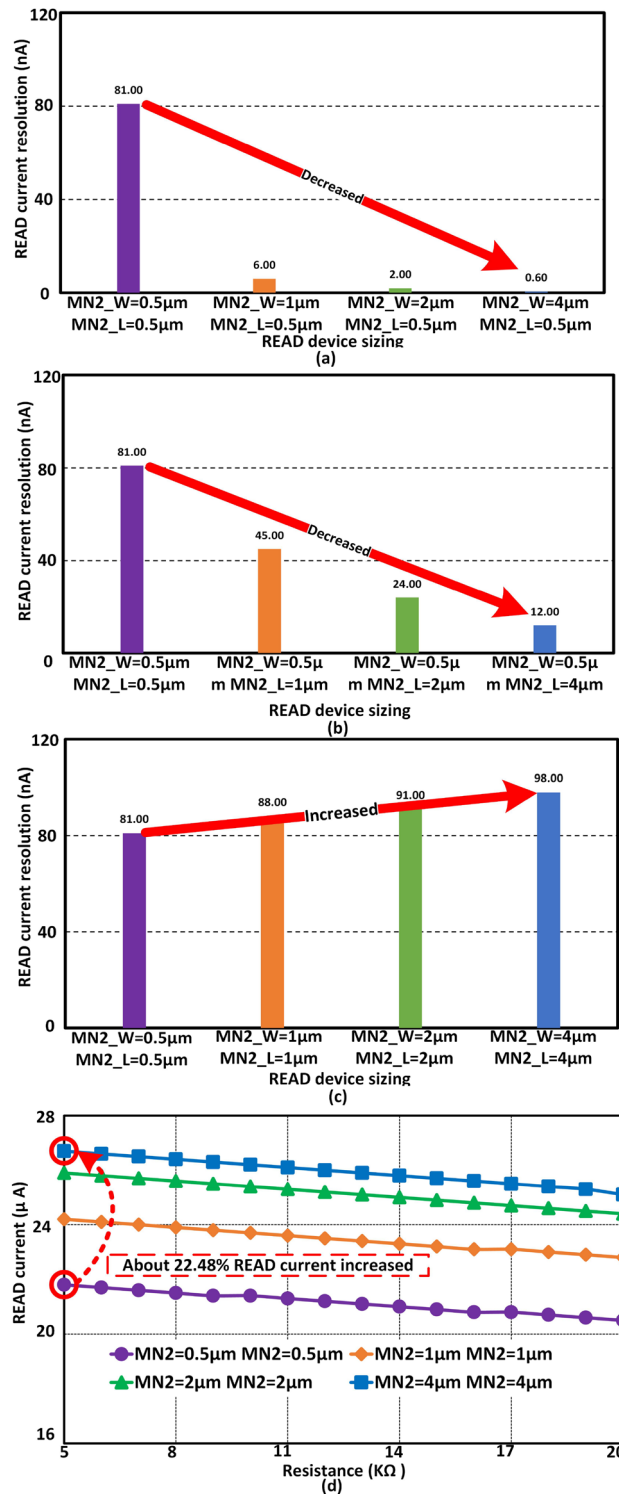
The length and width of  $M_{N2}$  are varied to observe the effect on READ current resolution. Figure 3a illustrates the READ current resolutions, when the width of the device is varied from  $0.5$  to  $4\ \mu\text{m}$  and the length is fixed at  $0.5\ \mu\text{m}$ . When the length and width of the device are set at a minimal size, the READ current resolution is about 81 nA. The first test case shows better READ current resolutions with lower power consumption.

Next, the width of the  $M_{N2}$  is set at  $0.5\ \mu\text{m}$  and the length is varied from  $0.5$  to  $4\ \mu\text{m}$ . According to Fig. 3b, the READ current resolution is decreased with the increment of length. At this point, the overall READ current will be decreased by up-sizing the length of  $M_{N2}$ , with higher latency. Finally, both length and width are increased simultaneously. Figure 3c shows the READ current resolutions when both length and width are increased. READ current resolutions are increased as the length and width are up-scaled at the same time. About 21% improvement in READ current resolutions can be achieved by up-sizing the length and width of  $M_{N2}$  simultaneously. Figure 3d shows the READ current level with different sizing combinations. READ current resolution can be increased with the overhead of area and READ current. Figure 3d shows, at  $5\ \text{k}\Omega$  memristive weight, the READ current is  $21.8\ \mu\text{A}$  with minimal length and width of the  $M_{N2}$ . On the other hand, when both length and width of  $M_{N2}$  are increased to  $4\ \mu\text{m}$ , the READ current is increased by 22.48%. At the same time, the READ current resolution is increased by 21%. There is a clear trade-off between READ current resolution and READ current overhead. In addition, the overall design area is also influenced by a larger length of  $M_{N2}$ .

Table 1 shows the optimized sizing configuration to enhance the READ current resolution. The width of  $M_{P1}$  and  $M_{N1}$  are set at  $1\ \mu\text{m}$  and  $4\ \mu\text{m}$  respectively. Both MOSFET's length are fixed at  $0.5\ \mu\text{m}$ . In addition, the length and width of  $M_{N2}$  is considered as  $4\ \mu\text{m}$ . The READ current resolution is about 98 nA with this optimized sizing. In the next section, the READ current resolution will be adapted dynamically with  $V_{\text{READ}}$  and  $MN2_{\text{Bias}}$  signals.

### Reconfigurable READ current resolution

READ current resolution can be adapted at run time. Various applications can perform better at enhanced READ current or weight resolution. To enhance the application's performance, a reconfigurable or run-time adaptation of READ current resolution is proposed with different circuit techniques. Specifically, READ voltage ( $V_{\text{READ}}$ ) scaling is a useful technique to influence the READ current resolution at run time.



**Figure 3.** Cadence simulation results for *READ* current resolutions with different width and length of  $M_{N2}$  device. (a) Shows the *READ* current resolution when the length of the  $M_{N2}$  is fixed at  $0.5\ \mu\text{m}$  and the width is varied from  $0.5$  to  $4\ \mu\text{m}$ . The current resolution is drastically decreased with the increment of  $M_{N2}$ 's width. (b) shows the *READ* current resolution when the width of the  $M_{N2}$  is fixed at  $0.5\ \mu\text{m}$  and the length is varied from  $0.5$  to  $4\ \mu\text{m}$ . The *READ* current resolution is also decreased with the increment of length of  $M_{N2}$ . Finally (c) Shows the *READ* current resolution when the length and width of  $M_{N2}$  are increased simultaneously. *READ* current resolution is increased when the length and width of the  $M_{N2}$  are increased at the same time. (d) Shows the *READ* current level with different  $M_{N2}$  sizing. About 22.48% *READ* current overhead is observed to improve 21% *READ* current resolution.

Transistor name	Width ( $\mu\text{m}$ )	Length ( $\mu\text{m}$ )
MP1	1	0.5
MN1	4	0.5
MN2	4	4

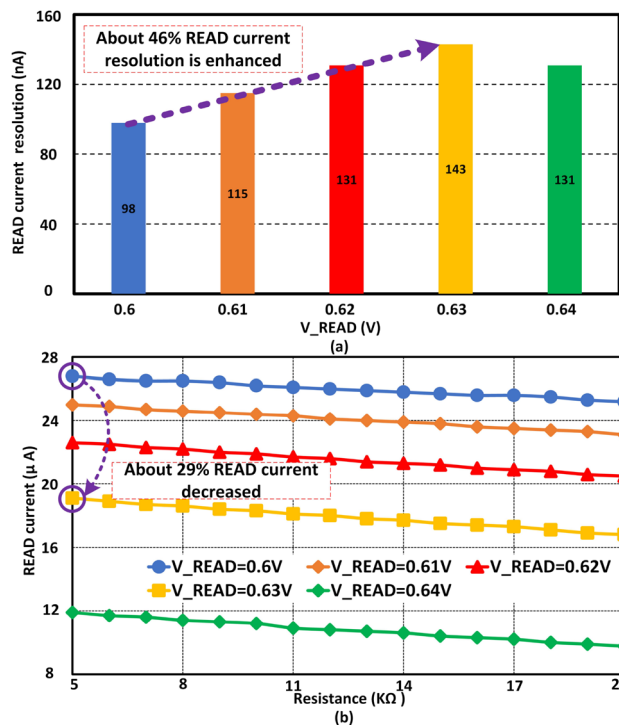
**Table 1.** Transistor scaling to enhance *READ* current resolution.

### READ voltage scaling

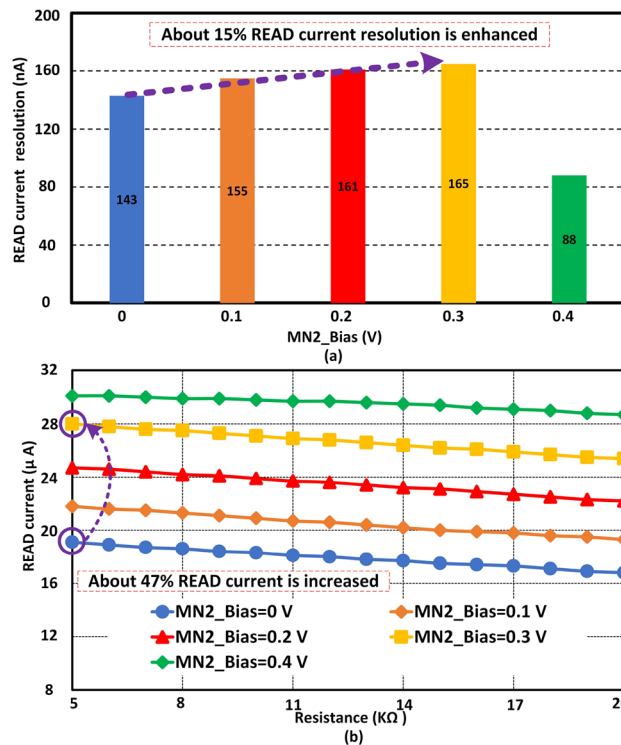
Figure 1 shows  $V_{\text{READ}}$  is the gate voltage of  $M_{N1}$ . Initially, the applied amplitude of this signal was 0.6 V. Figure 4a shows the *READ* current resolution at 0.6 V is 98 nA for 4-bit data. Here, the optimized sizing is utilized from Table 1. The *READ* current resolution is about 143 nA at 0.63 V as  $V_{\text{READ}}$ . About 46% of *READ* current resolution can be enhanced with  $V_{\text{READ}}$  scaling at run time. If the  $V_{\text{READ}}$  is increased more than 0.63 V for this particular sizing or configuration, then the resolution is decreased. A stronger turn-on of  $M_{N1}$  influences the final *READ* current negatively. As a result, the resolution is decreased with excessive  $V_{\text{READ}}$ . Figure 4b shows *READ* current with different gate voltages of  $M_{N1}$ . When the *READ* voltage is 0.6 V, the *READ* current is about 26.8  $\mu\text{A}$  at 5 k $\Omega$ . On the other hand, if the *READ* voltage is scaled up to 0.63 V, the *READ* current will be decreased to about 19.1  $\mu\text{A}$ . About 29% *READ* current can be optimized when the *READ* voltage is scaled from 0.6 to 0.63 V. *READ* current shows a significantly lower value at 0.64 V. However, at 0.64 V the *READ* current resolution is reduced significantly compared to the value at 0.63 V. In addition, at 0.64 V the std. dev. of *READ* current is  $\sim 7\%$  higher than std. dev. at 0.63 V. Thus, the *READ* current at 0.63 V is more reliable than at 0.64 V. According to our design optimization, the *READ* current resolution is higher at 0.63 V among all the *READ* voltages. In addition, the *READ* current resolution can be varied from 98 to 143 nA with  $V_{\text{READ}}$  signal scaling. In the next sub-section, another device technique is utilized to enhance the *READ* current resolution at run time.

### READ device $M_{N2}$ biasing

*READ* device  $M_{N2}$  biasing is another technique to manage synaptic *READ* current at run time. Here, the *READ* voltage  $V_{\text{READ}}$  is set at 0.63 V with optimized sizing. Figure 5a shows the *READ* current resolution at different bias voltages at the body of  $M_{N2}$ . As we know, body biasing can change the threshold voltage of a MOSFET



**Figure 4.** *READ* voltage ( $V_{\text{READ}}$ ) has a significant effect on *READ* current resolution. (a) Shows the *READ* current resolution at different  $V_{\text{READ}}$  voltage. *READ* current resolution is increased with the increment of the gate voltage of  $M_{N1}$ . After a certain level of gate voltage increment, the *READ* resolution starts decreasing. (b) Exhibits the *READ* current level at different *READ* voltages. As we increase the *READ* voltage at the gate of  $M_{N1}$ , the *READ* current level starts decreasing. Due to a weak turn-on of  $M_{N2}$ , the *READ* current level is decreased. Here is an interesting thing to notice, as we increase the gate voltage of  $M_{N1}$  the overall *READ* current level is decreased. As a result, the *READ* current resolution is increased with overall *READ* current optimization.



**Figure 5.** *READ* device  $M_{N2}$  plays an important role to make *READ* current resolution adaptable at run time. The body of  $M_{N2}$  is scaled to enhance the resolution with *READ* power overhead. (a) Shows the body biasing effect on *READ* current resolution. About 15% resolution can be enhanced with body biasing. (b) Illustrates the *READ* current level at different biasing voltages. The *READ* current is increased with the increment of body biasing of  $M_{N2}$ .

to control the current flowing through it. Here, a positive body voltage is applied on  $M_{N2}$  to improve the *READ* current resolution. When the body voltage is 0 V, the *READ* current resolution of our proposed synapse is about 143 nA. If the body voltage is scaled up to 0.3 V, then the *READ* current resolution is further up to 15%. In addition, if the body biasing voltage is increased more, then a reverse phenomenon is observed with *READ* current resolution. Due to the channel effect of the NMOS, the reverse phenomenon is observed with higher body bias voltage. At 0.4 V the *READ* current resolution is decreased to 88 nA. As a result 0.3 V is an optimized body bias voltage for this design scenario. Figure 5b exhibits the *READ* current of our synapse at different body biasing scenarios. Overall *READ* current is increased with the increment of body biasing voltage. About 47% *READ* current is increased with body biasing to enhance the *READ* current resolution at run time. Finally, it can be said that the *READ* current resolution of our proposed synapse can be adapted at run time with different body biasing voltages.

### READ current resolution and design performance evaluation

Table 2 shows an evaluation of device sizing, *READ* current resolution, and *READ* power. There are five test cases considered for the evaluation. The first test case is constructed with base device sizing. Here, the  $V_{READ}$  and body bias ( $M_{N2\_Bias}$ ) are 0.6 V and 0 V respectively for first three test cases. The first test case shows 19 nA *READ* current resolution with 18.87  $\mu W$  as a max *READ* power. Here, both stages' power (1st and 2nd) are considered for the max *READ* power. The second test case is the 1st stage device sizing. Here, the *READ* current resolution is 81 nA with 4-bit precision which is about 4.3 $\times$  higher than the base sizing test case. Only 5% of power overhead is observed compared to the base test case. The next test case is focused on 2nd stage device sizing with 1st stage sizing. Here, the *READ* current resolution is 98 nA, which is 5.16 $\times$  more improved than the base test case (Base sizing). Here, the *READ* power improvement is only 0.16% compared to the base test case. When the 2nd stage device is scaled the overall *READ* current is optimized slightly compared to the base test case. The fourth test case is to increase the  $V_{READ}$  to 0.63 V from 0.6 V. In this scenario, the *READ* current resolution is 143 nA, which is 7.53 $\times$  enhanced than the base test case. In addition, the max *READ* power shows 1.43% overhead compared to the base test case. Finally, the body biasing is applied to the body of  $M_{N2}$ . After applying 0.3 V, the *READ* current resolution is about 165 nA, which is 8.68 $\times$  improved compared to the base test case. Here, the max *READ* power overhead is about 3.60% compared to the base test case. The *READ* voltage on the source of the  $M_{N2}$  is automatically adjusted based on the gate voltage of the  $M_{N2}$ . Due to that, the *READ* power is quite stable in this design. Here, the Monte Carlo simulation is observed with 1000 samples in a Cadence Spectre environment to analyze the *READ* current variation. The *READ* current shows about 0.65 $\times$  variations with 8.68 $\times$  resolution improvement.

Test case	Device size	READ current resolution	READ current resolution evaluation	Max READ power [both stage power]	Max READ power evaluation
Base sizing	MP1 = (0.5/0.5) $\mu\text{m}$ MN1 = (1/0.5) $\mu\text{m}$ MN2 = (0.5/0.5) $\mu\text{m}$	19 nA	–	18.87 $\mu\text{W}$	–
1st stage device sizing	MP1 = (1/0.5) $\mu\text{m}$ MN1 = (4/0.5) $\mu\text{m}$ MN2 = (0.5/0.5) $\mu\text{m}$	81 nA	4.26 $\times$ improved compared to base sizing	19.82 $\mu\text{W}$	5% overhead compared to base sizing
2nd stage device sizing	MP1 = (1/0.5) $\mu\text{m}$ MN1 = (4/0.5) $\mu\text{m}$ MN2 = (4/4) $\mu\text{m}$	98 nA	5.16 $\times$ improved compared to base sizing	18.84 $\mu\text{W}$	0.16% improved compared to base sizing
V_READ @0.63V	MP1 = (1/0.5) $\mu\text{m}$ MN1 = (4/0.5) $\mu\text{m}$ MN2 = (4/4) $\mu\text{m}$	143 nA	7.53 $\times$ improved compared to base sizing	19.6 $\mu\text{W}$	1.43% overhead compared to base sizing
Body bias @0.3 V	MP1 = (1/0.5) $\mu\text{m}$ MN1 = (4/0.5) $\mu\text{m}$ MN2 = (4/4) $\mu\text{m}$	165 nA	8.68 $\times$ improved compared to base sizing	18.19 $\mu\text{W}$	3.60% improved compared to base sizing

**Table 2.** READ current resolution enhancement and design evaluation.

### Impact of READ current resolution on applications

To specifically investigate the impact of READ current resolution, we assessed the performance of spiking neural networks under varying probabilities of synapse read failures. Our hypothesis is that when the current resolution is higher, there will be fewer or no read errors, while when the current resolution is lower, read errors are significantly more likely. To perform this evaluation, we leveraged the TENNLab neuromorphic computing framework<sup>31</sup>, which allows for evaluation of neuromorphic processors using different applications and algorithms. Within the framework, we used the RISP neuromorphic simulator<sup>33</sup>, with integrate and fire neurons and synapses with 4-bit weight resolution.

To specifically study the impact of READ current resolution, we evaluated how spiking neural networks with different likelihoods of read failures for each synapse. In particular, for a particular network evaluation, we defined a likelihood of read failures for each synapse read, wherein the weight read would be one level off (either the level above or level below), which would be more likely to happen for low current resolution. We trained the networks with these read failures using an evolutionary optimization training approach for spiking neural networks and neuromorphic system called EONS<sup>34</sup>. EONS evolves the parameters and topology of the network simultaneously. We trained for the iris dataset, the wine dataset, and the breast cancer dataset, three commonly used toy datasets in machine learning that are available in the UCI machine learning repository<sup>35</sup>, as well as the EEG motion dataset, a timeseries dataset<sup>36,37</sup>. We trained 100 networks for each of six different likelihoods of read errors for the synaptic weight values: 0, 10, 20, 30, 40, and 50 percent for each dataset. Figure 6 shows the results for these simple datasets on both training and testing performance. This figure shows that, in general, the best overall testing performance was achieved by no read errors at all; however, on average, some noise on the read errors does not necessarily hurt performance significantly, either in training or testing. For the results in Fig. 6, it is worth noting that the networks were trained and tested using read errors. Figure 7 shows the results for when networks are trained without read errors and then tested with varying likelihoods of read errors per synapse, which would likely be the case for networks trained in simulation and then deployed to hardware. In this case, we can see that read errors cause a decrease in testing performance for each dataset.

We conducted similar tests for four test applications: two from the OpenAI gym control environments (BipedalWalker-v3 and LunarLander-v2)<sup>38</sup> and two from the TENNlab suite of control applications (bowman and polebalance)<sup>39</sup>. The results for different likelihoods of read errors for these control tasks are shown in Fig. 8. In this case, we do see a significant downward trend in performance for most of the applications when training and testing with noise.

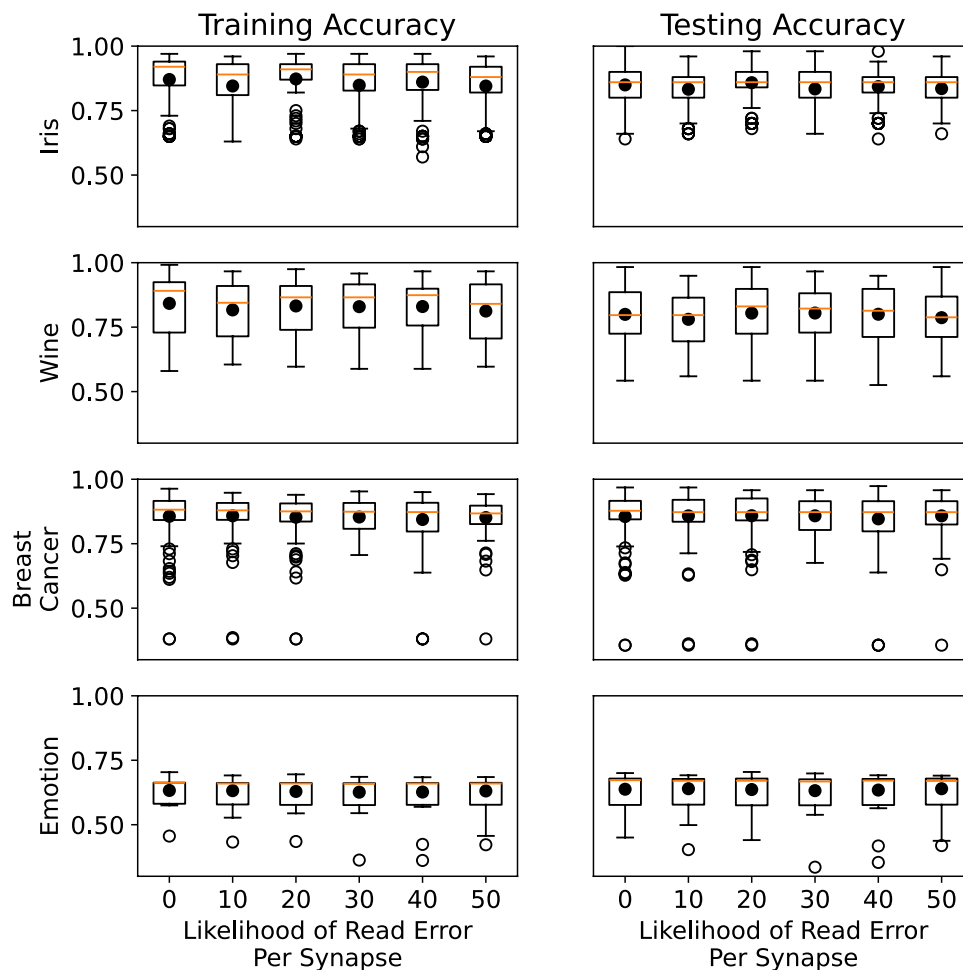
Finally, in Fig. 9, we see the results of a reservoir computing approach when encountering read errors on synapses. We use reservoirs of 100 neurons with 10% randomly initialized connectivity. We once again evaluate the three toy classification datasets, but we omit the timeseries EEG dataset. In this case, we see a significant downward trend, more pronounced and more consistent than in the other cases. As such, this indicates that the reservoir approach may not handle noise well.

Because this work enables higher-precision weights to be used on synapses with more reliability, we investigated the impact of precision on performance. Figure 10 shows the results for different levels of bit-precision (2, 3, and 4) on the synaptic weights. Here, we can see that 2-bit weight synapses perform significantly worse across all three datasets than 3- and 4-bit precision synapses, as expected. We see similar results for the control tasks (Fig. 11) and the reservoir computing tasks (Fig. 12).

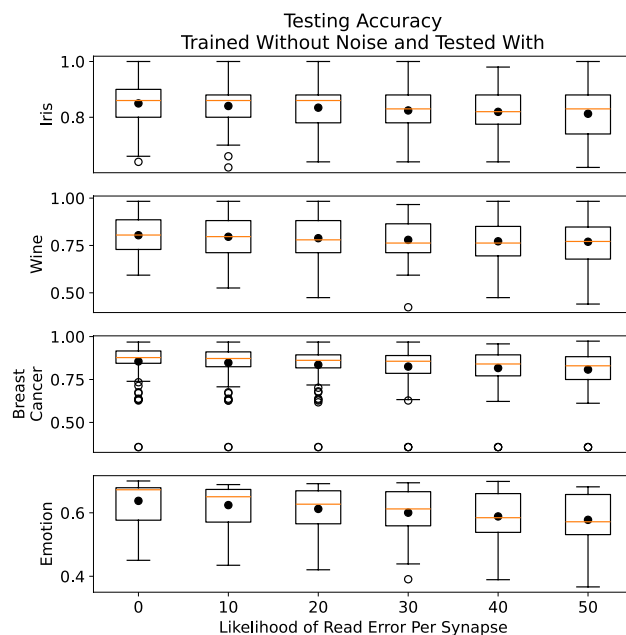
### Comparison with prior works

Reliable READ current resolution is a big challenge for the memristive-based synapse or memory design. Usually, the minimum current resolution is a few nA. Due to that, it is challenging for the circuit designer to sense the current level properly with ADC or CMOS neurons. In our neural network analysis, we consider a CMOS neuron to observe the charge accumulation and fire<sup>46</sup>. In this work, a memristor-based synapse is optimized to enhance the READ current resolutions. At first, the device sizing is considered to optimize the resolution. Hence, READ voltage and body bias are considered to enhance the resolution at run time. Here, the optimal size is adopted from Table 1. In addition, the READ voltage and body bias voltages are selected to 0.63 V and 0.3

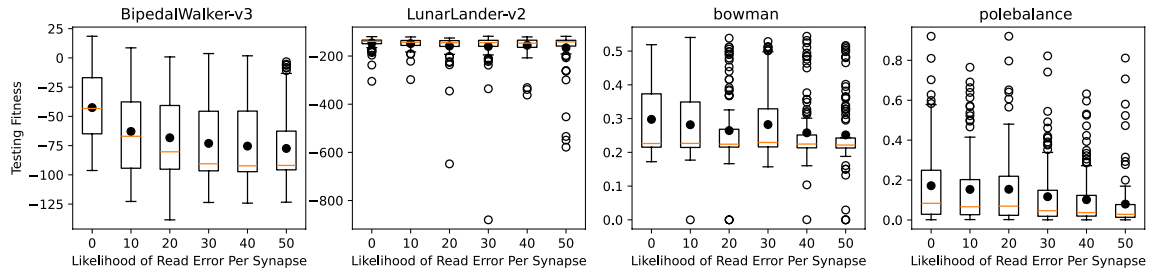




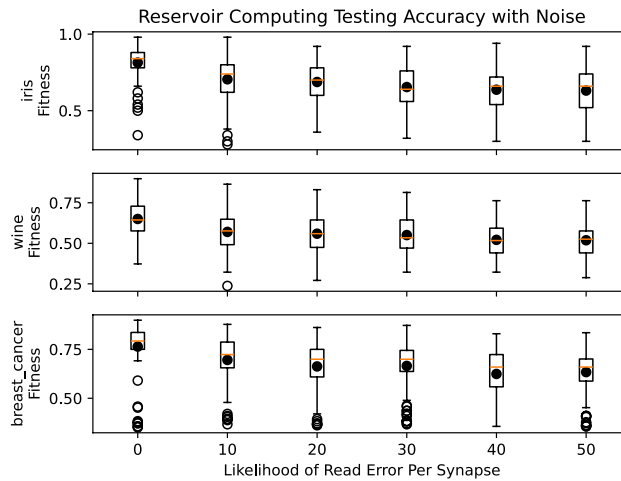
**Figure 6.** Training and testing accuracy for iris, wine, and breast cancer datasets. In these results, the networks were trained and tested with read errors.



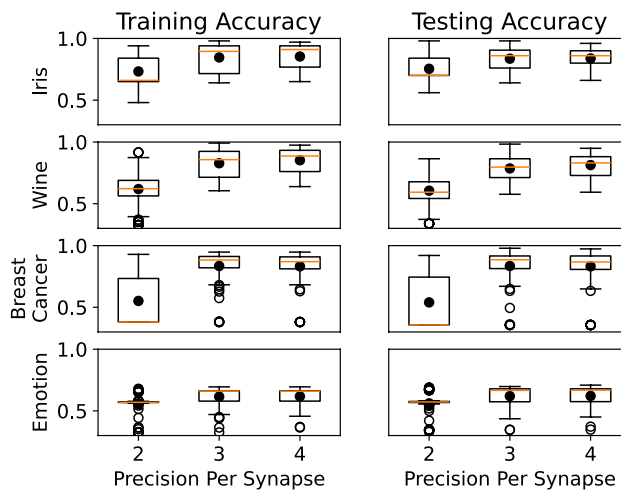
**Figure 7.** Testing accuracies for networks trained without read errors and tested with read errors.



**Figure 8.** Results for control with read errors on synapses.



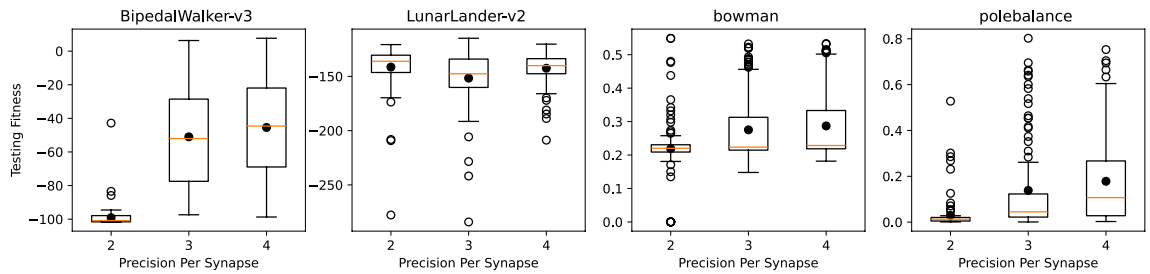
**Figure 9.** Results for reservoir computing with read errors on synapses.



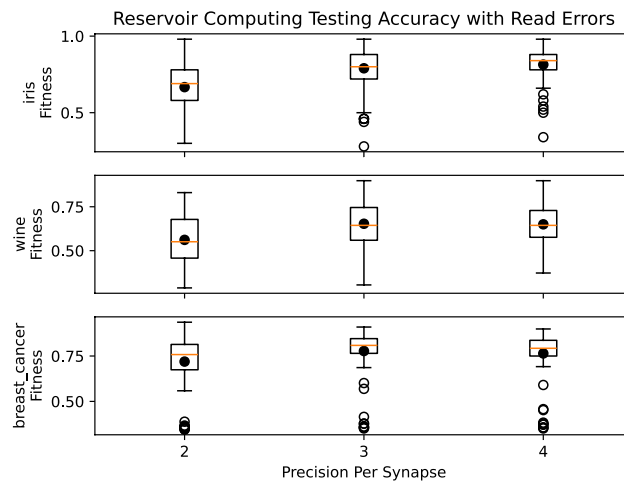
**Figure 10.** Results for reduced precision on the synaptic weights.

V from 0.6 and 0 V respectively. According to Table 3, the minimum *READ* current resolution is 165 nA with 4-bit data precision. As we know, the larger resistance level shows a lower current resolution. In our case, the current level difference between 19 and 20 kΩ is considered to determine the minimum current resolution. At room temperature the value is at least 165 nA. The maximum *READ* current ~ 28 μA is captured at 5 kΩ. Our synapse is programmed at LRS to avoid inherent process variations.

Another research article shows the minimum *READ* current resolution of a 3T1R synapse is 20 nA for 4-bit data precision. Our proposed design shows 8.25× resolution improvement compared to their work<sup>28</sup>. Energy-efficient and high-performance synapse is presented in another research paper, where the minimum *READ* current resolution is 20 nA<sup>30</sup>. This design is also based on HfO<sub>2</sub> based memristor and 65 nm CMOS process. The programming region is between 3 to 18 kΩ, which provides low inherent process variation with 4-bit data



**Figure 11.** Results for control with reduced precision on the synaptic weights.



**Figure 12.** Results for reservoir computing with reduced precision on the synaptic weights.

density. Our proposed design also illustrated  $8.25\times$  current resolution improvement compared to their work<sup>30</sup>. A  $\text{TiO}_2$  based memristive memory is presented in<sup>40</sup>, which is programmed from  $1\text{ k}\Omega$  to  $1\text{ G}\Omega$ . Due to the utilization of HRS, the inherent process variation is higher for this design compared to our proposed design. This design only covers four programming states. As a result, our proposed design shows a  $4\times$  more dense data storage capacity compared to their design<sup>40</sup>. The max *READ* current is about  $3.21\times$  higher than our proposed design. The minimum *READ* current resolution is about  $90\text{ nA}$ , which is  $1.83\times$  lower than our proposed design.

A research group is presented a  $\text{TiO}_2$  based multi-level resistive memory, which is programmed from  $10\text{ k}\Omega$  to  $100\text{ M}\Omega$ <sup>41</sup>. Due to programming in HRS, the inherent process variation is higher than our design. Their design is only programmed in six different states, which exhibits  $2.67\times$  lower memory density compared to our proposed design. The maximum *READ* current of their design is  $100\text{ }\mu\text{A}$ , which is  $3.57\times$  higher than our proposed design. Moreover, the minimum *READ* current resolution of their design is  $\sim 17\text{ nA}$ , whereas our proposed design shows  $\sim 9.7\times$  enhanced resolution compared to their design<sup>41</sup>. A  $\text{HfO}_2$  based multi-level cell is presented with 2-bit memory density, which is  $4\times$  lower memory density than our proposed synapse<sup>42</sup>. Their device is programmed from  $10\text{ k}\Omega$  to  $1\text{ M}\Omega$ , which shows higher inherent process variation compared to our design. Their device draws lower maximum *READ* current than our design. However, the minimum *READ* current resolution is  $2.04\times$  lower than our proposed design.

Another multi-level resistive memory is presented using a-ZnO material, which is programmed between  $24$  and  $176\text{ M}\Omega$ <sup>43</sup>. Their design shows higher process variation and  $3.2\times$  lower memory density than our proposed design. Their maximum *READ* current is lower than our design. However, the minimum *READ* current resolution is extremely low compared to our proposed design. A  $\text{HfO}_2$  based multi-level RRAM is presented in a research article, where authors programmed their device between  $0.8\text{ k}\Omega$  and  $100\text{ M}\Omega$  with 8 programming states<sup>44</sup>. Our proposed design shows lower process variation and  $2\times$  higher memory density than their design. The maximum *READ* current of their design is extremely higher than our proposed design. Our proposed design also shows a  $3.67\times$  improved minimum *READ* current resolution than their design<sup>44</sup>. Another research group presented their device with  $\text{TiO}_2$ , which is programmed between  $1\text{ M}\Omega$  and  $10\text{ G}\Omega$  with six programming states<sup>45</sup>. This design shows a  $2.67\times$  lower memory density than our proposed design. Due to their HRS programming region, the process variation is higher than our design. The maximum *READ* current of their design is lower than our proposed design. However, their minimum *READ* current resolution is extremely lower than our proposed design. Moreover, our design shows power savings compare to prior works<sup>40,41,44</sup>. In addition, our proposed design shows more linear *READ* current compared to other designs<sup>40–45</sup>. Due to a compact programming range, our design shows better linearity compared to prior works. A wide programming range causes non-linear behavior of *READ* current.

References	JETCAS,2023 <sup>28</sup>	MWCAS,2023 <sup>30</sup>	Electron Device Lett.,2010 <sup>40</sup>	Appl. Phys.,2017 <sup>41</sup>	Semicond. Sci., 2015 <sup>42</sup>	RSC. Adv.,2016 <sup>43</sup>	Nanoscale, 2014 <sup>44</sup>	Phys. Status Solidi A, 2017 <sup>45</sup>	This work
Technology	65 nm CMOS	65 nm CMOS	–	–	–	–	–	–	65 nm CMOS
Memristor material	HfO <sub>2</sub>	HfO <sub>2</sub>	TiO <sub>2</sub>	TiO <sub>2</sub>	HfO <sub>2</sub>	a-ZnO	HfO <sub>2</sub>	TiO <sub>2</sub>	HfO <sub>2</sub>
Programming region	5 kΩ–20 kΩ	3 kΩ– 18 kΩ	1 kΩ–1 GΩ	10 kΩ–100 MΩ	10 kΩ– 1 MΩ	24 MΩ–176 MΩ	0.8 kΩ–100 MΩ	1 MΩ– 10 GΩ	5 kΩ–20 kΩ
Impact of process variation	Lower	Lower	Higher	Higher	Higher	Higher	Higher	Higher	Lower
Number of programming states	16	16	4	6	4	5	8	6	16
Storage density improvement compared to prior work	1×	1×	4×	2.67×	4×	3.2×	2×	2.67×	1×(base)
Max READ current	~5.44μA	~2.92μA	~90μA	100μA	10μA	4.166 nA	~625μA	0.5μA	28μA
Max READ power	~8.24μW	~3.50μW	~45μW	100μW	1μW	0.4116 nW	~312.5μW	0.25μW	19.6μW
Max READ power improvement compared to prior work	No	No	Yes	Yes	No	No	Yes	No	–
Minimum READ current resolution	20 nA	20 nA	90 nA	~17 nA	81 nA	0.462 nA	45 nA	0.45 nA	165 nA
READ current resolution improvement compared to prior work	8.25×	8.25×	1.83×	~9.7×	2.04×	357×	3.67×	366.67×	1×
READ current linearity	Linear	Linear	Non-linear	Non-linear	Non-linear	Non-linear	Non-linear	Non-linear	Linear

**Table 3.** Comparison with prior works.

After comparing our proposed design with prior works from different research groups, it can be said that our proposed design shows lower inherent process variation with higher memory density. In addition, our proposed design shows an enhanced *READ* current resolution compared to others' designs.

### Conclusions and future work

In this paper, a HfO<sub>2</sub> based current-controlled memristive synapse is optimized for *READ* operation. At first, the *READ* devices are optimized to enhance the *READ* current resolution. About 4.3× and 21% *READ* current resolution is enhanced with 1st and 2nd stage device sizing respectively. *READ* voltage scaling and body biasing are applied to enhance the *READ* current resolution at run time. About 46% and 15% *READ* current resolution is improved with *READ* voltage scaling and body biasing. A neuromorphic framework EONS shows that a higher *READ* current resolution exhibits better accuracy compared to a lower resolution on classification and control applications. Lower resolutions are more likely to be affected by reading failures with higher noise. As a result, a higher *READ* current resolution makes the neuromorphic system more reliable.

### Data availability

The datasets (circuit simulation) generated during this study are available from the corresponding author (H.D.) upon reasonable request.

### Code availability

The codes for application evaluation are available from the author (C.S.) upon reasonable request.

Received: 26 January 2024; Accepted: 4 April 2024

Published online: 17 April 2024

### References

- Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (eds Pereira, F. *et al.*) (Curran Associates Inc., 2012).
- Adnan, M. M. *et al.* Design of a robust memristive spiking neuromorphic system with unsupervised learning in hardware. *J. Emerg. Technol. Comput. Syst.* **17**, 1–26. <https://doi.org/10.1145/3451210> (2021).
- Lee, K., Park, J. & Yoo, H.-J. A low-power, mixed-mode neural network classifier for robust scene classification. *J. Semicond. Technol. Sci.* **19**, 129–136. <https://doi.org/10.5573/JSTS.2019.19.1.129> (2019).
- Kim, S. *et al.* Analog synaptic behavior of a silicon nitride memristor. *ACS Appl. Mater. Interfaces* **9**, 40420. <https://doi.org/10.1021/acsami.7b11191> (2017).

5. Rashvand, P., Ahmadzadeh, M. & Shayegh, F. Design and implementation of a spiking neural network with integrate-and-fire neuron model for pattern recognition. *Int. J. Neural Syst.* **31**, 2050073. <https://doi.org/10.1142/S0129065720500732> (2020).
6. Rathore, M. *et al.* Reliability analysis of memristive reservoir computing architecture. In *Proceedings of the Great Lakes Symposium on VLSI 2023, GLSVLSI '23* 131–136 (Association for Computing Machinery, 2023). <https://doi.org/10.1145/3583781.3590210>.
7. Merolla, P. *et al.* A digital neurosynaptic core using embedded crossbar memory with 45pj per spike in 45nm. In *2011 IEEE Custom Integrated Circuits Conference (CICC)* 1–4. <https://doi.org/10.1109/CICC.2011.6055294> (2011).
8. Asghar, M., Arslan, S. & Kim, H. A low-power spiking neural network chip based on a compact lif neuron and binary exponential charge injector synapse circuits. *Sensors* **21**, 4462. <https://doi.org/10.3390/s21134462> (2021).
9. Kim, Y., Zhang, Y. & Li, P. A reconfigurable digital neuromorphic processor with memristive synaptic crossbar for cognitive computing. *J. Emerg. Technol. Comput. Syst.* **11**, 1–25. <https://doi.org/10.1145/2700234> (2015).
10. Foshie, A. Z. *et al.* Benchmark comparisons of spike-based reconfigurable neuromorphic architectures for control applications. In *Proceedings of the Great Lakes Symposium on VLSI 2022, GLSVLSI '22* 383–386 (Association for Computing Machinery, 2022). <https://doi.org/10.1145/3526241.3530381>.
11. Foshie, A. Z. *et al.* A multi-context neural core design for reconfigurable neuromorphic arrays. In *2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)* 67–72. <https://doi.org/10.1109/ISVLSI51109.2021.00023> (2021).
12. Schuman, C. D., Das, H., Plank, J. S., Aziz, A. & Rose, G. S. Evaluating neuron models through application-hardware co-design. In *2023 57th Asilomar Conference on Signals, Systems, and Computers (ACSSC)* 537–542 (2023). <https://doi.org/10.1109/IEEECONF59524.2023.10477027>
13. Chakraborty, N. N., Das, H. & Rose, G. S. A Mixed-Signal Short-Term Plasticity Implementation for a Current-Controlled Memristive synapse, *GLSVLSI '23* 179–182 (Association for Computing Machinery, 2023). <https://doi.org/10.1145/3583781.3590283>.
14. Weiss, R., Das, H., Chakraborty, N. N. & Rose, G. S. Stdp based online learning for a current-controlled memristive synapse. In *2022 IEEE 65th International Midwest Symposium on Circuits and Systems (MWSCAS)* 1–4. <https://doi.org/10.1109/MWSCAS54063.2022.9859294> (2022).
15. Chakraborty, N. N., Das, H. & Rose, G. S. Spike-timing-dependent plasticity for a hafnium-oxide memristive synapse. In *2023 IEEE 66th International Midwest Symposium on Circuits and Systems (MWSCAS)* 463–467. <https://doi.org/10.1109/MWSCAS57524.2023.10406099> (2023).
16. Chakraborty, N. N., Das, H. & Rose, G. S. Spike-driven synaptic plasticity for a memristive neuromorphic core. In *2023 IEEE 66th International Midwest Symposium on Circuits and Systems (MWSCAS)* 644–648. <https://doi.org/10.1109/MWSCAS57524.2023.10406136> (2023).
17. Chakraborty, N. N., Das, H. & Rose, G. S. Homeostatic plasticity in a leaky integrate and fire neuron using tunable leak. In *2023 IEEE 66th International Midwest Symposium on Circuits and Systems (MWSCAS)* 738–742. <https://doi.org/10.1109/MWSCAS57524.2023.10406066> (2023).
18. Das, H. *et al.* An efficient and accurate memristive memory for array-based spiking neural networks. *IEEE Trans. Circ. Syst. I Regul. Pap.* **70**, 4804–4815. <https://doi.org/10.1109/TCSI.2023.3301020> (2023).
19. Tushar, S. N. B., Das, H. & Rose, G. S. HfO<sub>2</sub>-based synaptic spiking neural network evaluation to optimize design and testing cost. In *2024 IEEE International Symposium on Circuits and Systems (ISCAS)* (2024).
20. Cruz-Albrecht, J. M., Derosier, T. & Srinivasa, N. A scalable neural chip with synaptic electronics using cmos integrated memristors. *Nanotechnology* **24**, 384011. <https://doi.org/10.1088/0957-4484/24/38/384011> (2013).
21. Xu, Y., Das, H., Gong, Y. & Gong, N. On mathematical models of optimal video memory design. *IEEE Trans. Circ. Syst. Video Technol.* **30**, 256–266. <https://doi.org/10.1109/TCSVT.2018.2890383> (2020).
22. Brivio, S. *et al.* Extended memory lifetime in spiking neural networks employing memristive synapses with nonlinear conductance dynamics. *Nanotechnology* **30**, 015102. <https://doi.org/10.1088/1361-6528/aae81c> (2018).
23. Das, H., Manisankar, S., Cheng, W. & Chung, Y. Experimental n-style two-transistor edram in logic cmos technology. In *2015 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC)* 75–78. <https://doi.org/10.1109/EDSSC.2015.7285053> (2015).
24. Chung, Y., Cheng, W. & Das, H. Built-in parasitic-diode-based charge injection technique enhancing data retention of gain cell dram. *Electron. Lett.* **51**, 1854–1855 (2015).
25. Chua, L. Memristor-the missing circuit element. *IEEE Trans. Circ. Theory* **18**, 507–519. <https://doi.org/10.1109/TCT.1971.1083337> (1971).
26. Mannan, Z. I., Kim, H. & Chua, L. Implementation of neuro-memristive synapse for long-and short-term bio-synaptic plasticity. *Sensors* **21**, 644. <https://doi.org/10.3390/s21020644> (2021).
27. Hazra, J., Liehr, M., Beckmann, K., Rafiq, S. & Cady, N. Improving the memory window/resistance variability trade-off for 65nm cmos integrated hfo2 based nanoscale rram devices. In *2019 IEEE International Integrated Reliability Workshop (IIRW)* 1–4. <https://doi.org/10.1109/IIRW47491.2019.8989872> (2019).
28. Das, H. *et al.* Optimizations for a current-controlled memristor-based neuromorphic synapse design. *IEEE J. Emerg. Sel. Top. Circ. Syst.* **13**, 889–900. <https://doi.org/10.1109/JETCAS.2023.3312163> (2023).
29. Payvand, M., Demirag, Y., Dalgaty, T., Vianello, E. & Indiveri, G. Analog weight updates with compliance current modulation of binary rram for on-chip learning. In *2020 IEEE International Symposium on Circuits and Systems (ISCAS)* 1–5. <https://doi.org/10.1109/ISCAS45731.2020.9180808> (2020).
30. Chakraborty, N. N., Tushar, S., Das, H. & Rose, G. S. Energy efficient and high-performance synaptic operating point evaluation for snn applications. In *2023 IEEE 66th International Midwest Symposium on Circuits and Systems (MWSCAS)* 918–922. <https://doi.org/10.1109/MWSCAS57524.2023.10406062> (2023).
31. Plank, J. S., Schuman, C. D., Bruer, G., Dean, M. E. & Rose, G. S. The tennlab exploratory neuromorphic computing framework. *IEEE Lett. Comput. Soc.* **1**, 17–20 (2018).
32. Das, H. *et al.* Rfam: Reset-failure-aware-model for hfo2-based memristor to enhance the reliability of neuromorphic design. In *Proceedings of the Great Lakes Symposium on VLSI 2023, GLSVLSI '23* 281–286 (Association for Computing Machinery, 2023). <https://doi.org/10.1145/3583781.3590211>.
33. Plank, J. S. *et al.* The case for risp: A reduced instruction spiking processor. [arXiv:2206.14016](https://arxiv.org/abs/2206.14016) (2022).
34. Schuman, C. D., Mitchell, J. P., Patton, R. M., Potok, T. E. & Plank, J. S. Evolutionary optimization for neuromorphic systems. In *Proceedings of the Neuro-inspired Computational Elements Workshop* 1–9 (2020).
35. Asuncion, A. & Newman, D. *Uci Machine Learning Repository* (Springer, 2007).
36. Bird, J. J. *et al.* A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction. *Complexity* **2019**, 145 (2019).
37. Bird, J. J., Ekart, A., Buckingham, C. D. & Faria, D. R. Mental emotional sentiment classification with an eeg-based brain-machine interface. In *Proceedings of the International Conference on Digital Image and Signal Processing (DISP'19)* (2019).
38. Brockman, G. *et al.* Openai gym. [arXiv:1606.01540](https://arxiv.org/abs/1606.01540) (2016).
39. Plank, J. S. *et al.* The TENNLab suite of LIDAR-based control applications for recurrent, spiking, neuromorphic systems. In *44th Annual GOMACTech Conference* (Albuquerque, 2019).
40. Terai, M., Sakotsubo, Y., Kotsuji, S. & Hada, H. Resistance controllability of Ta<sub>2</sub>O<sub>5</sub>/TiO<sub>2</sub> stack rram for low-voltage and multilevel operation. *IEEE Electron. Dev. Lett.* **31**, 204–206. <https://doi.org/10.1109/LED.2009.2039021> (2010).

41. Bousoulas, P., Giannopoulos, I., Asenov, P., Karageorgiou, I. & Tsoukalas, D. Investigating the origins of high multilevel resistive switching in forming free ti/tio<sub>2</sub>-x-based memory devices through experiments and simulations. *J. Appl. Phys.* **121**, 094501 (2017).
42. Chen, W. *et al.* Switching characteristics of w/zr/hfo<sub>2</sub>/tin reram devices for multi-level cell non-volatile memory applications. *Semicond. Sci. Technol.* **30**, 075002 (2015).
43. Huang, Y. *et al.* Amorphous zno based resistive random access memory. *RSC Adv.* **6**, 17867–17872 (2016).
44. PhilipáWong, H.-S. *et al.* Multi-level control of conductive nano-filament evolution in hfo<sub>2</sub> reram by pulse-train operations. *Nanoscale* **6**, 5698–5702 (2014).
45. Tsigkourakos, M., Bousoulas, P., Aslanidis, V., Skotadis, E. & Tsoukalas, D. Ultra-low power multilevel switching with enhanced uniformity in forming free tio<sub>2</sub>-x-based rram with embedded pt nanocrystals. *Phys. Status Solidi A* **214**, 1700570 (2017).
46. Chakma, G. *et al.* Memristive mixed-signal neuromorphic systems: Energy-efficient learning at the circuit-level. *IEEE J. Emerg. Sel. Top. Circ. Syst.* **8**, 125–136. <https://doi.org/10.1109/JETCAS.2017.2777181> (2018).

## Acknowledgements

This material is based in part on research sponsored by Air Force Research Laboratory under agreement FA8750-21-1-1018. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory or the U.S. Government. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under award number DE-SC0022566.

## Author contributions

H.D. proposed the dynamically reconfigurable synaptic device to take advantage of the stochasticity, reliability, and security features of the proposed device. He prepared the initial draft for all the circuit simulations and analysis. N.N.C. helps with the writing. Drs. G.S.R. and C.S. were helping with overall supervision. Moreover, C.S. conducted the application evaluation experiments.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to H.D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024