Superconductor
Science and
Technology

PAPER • **OPEN ACCESS**

Superconductor bistable vortex memory for data storage and readout

To cite this article: M A Karamuftuoglu *et al* 2025 *Supercond. Sci. Technol.* **38** 015020

OPEN ACCESS
IOP Publishing
Superconductor Science and Technology
Supercond. Sci. Technol. **38** (2025) 015020 (10pp)
https://doi.org/10.1088/1361-6668/ad9863

Superconductor bistable vortex memory for data storage and readout

M A Karamuftuoglu*, B Z Ucpinar, S Razmkhah and M Pedram

Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089, United States of America

E-mail: karamuft@usc.edu, ucpinar@usc.edu, razmkhah@usc.edu and pedram@usc.edu

Received 6 October 2024, revised 14 November 2024
Accepted for publication 28 November 2024
Published 19 December 2024

Abstract

Despite superconductor electronics (SCE) advantages, the realization of SCE logic faces a significant challenge due to the absence of dense and scalable nonvolatile memory. While various nonvolatile memory technologies, including Non-destructive readout, vortex transitional memory, and magnetic memory, have been explored, designing a dense crossbar array and achieving a superconductor random-access memory remains challenging. This work introduces a novel, nonvolatile, high-density, and scalable vortex-based memory design for SCE logic called bistable vortex memory. Our proposed design addresses scaling issues with an estimated area of 10×10 um^2 while boasting zero static power with the dynamic energy consumption of 12 aJ for single-bit read and write operations. The current summation capability enables analog operations for in-memory or near-memory computational tasks. We demonstrate the efficacy of our approach with a 32×32 superconductor memory array operating at 20 GHz. Additionally, we showcase the accumulation property of the memory through analog simulations conducted on an 8×8 superconductor crossbar array.

Keywords: vortex memory, crossbar memory array, SQUID, superconductor electronics

1. Introduction

Superconductor electronics (SCE) has emerged as a promising beyond-CMOS technology [1], characterized by its low power consumption and ultra-fast processing capabilities [2]. Recent works have demonstrated the advantages of the SCE across various domains, such as energy-efficient computing [3], neural networks [4, 5], quantum computing [6], and quantum sensitive detectors [7, 8].

A better approach to utilizing the unique properties of superconductors is using them as a hardware accelerator to complement the well-established CMOS processors. This hybrid approach must include interfaces between CMOS and superconducting logic and transfer large amounts of data across a wide temperature range (e.g. room temperature for conventional CMOS or, at best, 50–70 K range for cryo-CMOS to 4 K for superconducting logic). A robust memory is necessary to enable such systems.

Despite significant advancements in superconductor-based computing technology, creating a dense and reliable superconducting random access memory (SRAM) array with short access times remains a challenge. This difficulty arises partly from the large area footprint of superconducting non-destructive readout (NDRO) memory cells, potential inductive coupling among these cells, which requires precise tuning of inductor values and Josephson junction (JJ) sizes within each memory cell, and the complexity involved in designing peripheral circuits for accessing individual cells within the memory array. Any solution to these challenges must enable the storage and retrieval of vast volumes of data to meet the ever-growing computing and memory demands while working at cryogenic temperatures.

* Author to whom any correspondence should be addressed.

To tackle the challenge of cryogenic memory design, researchers explored several design approaches. One proposed solution was utilizing the same single flux quantum (SFQ) technology found in pulse-based SFQ logic to create a memory array. This technology stores data bits by inducing magnetic flux in superconducting loops, relying on JJ switching. Various cryogenic memory designs have adopted this method to achieve promising initial results. The predominant approach within this framework is shift register-based memory design. Mukhanov *et al* [9] demonstrated the functionality of a 1024-bit shift register memory with a 19 GHz access frequency.

Similarly, Xu *et al* [10] showcased a shift register memory design featuring bit parallel access, enabling the efficient retrieval of multiple bits of data per clock cycle. However, the shift register structure of the memory restricts random access functionality, a crucial requirement in numerous computing applications. Additionally, the attained memory density remains comparatively low. The other approach uses superconducting NDRO flip-flops as storage elements in a RAM design. Multiple DC-powered superconductor RAM implementations [11, 12] utilize transformers for power and data transmission, resulting in a significant layout area. Unfortunately, the area density of such a memory array is relatively low, allowing us to include at most a few kilobits of RAM in a $5 \times 5\,\mathrm{mm}^2$ chip.

Hybrid designs blending CMOS and superconductor circuit characteristics [13, 14] offer high density and scalability. However, the access latency increases due to the need to go back and forth between the CMOS and superconducting logic domains, which operate at very different voltage and temperature ranges. Vernik *et al* [15] presented a memory design using magnetic JJ (MJJ), which is compact and fast. Unlike conventional JJs, MJJs incorporate a magnetic layer to enable distinct behaviors, making them well-suited for memory applications. Yet, it is non-scalable due to the high currents required for data writing and the reliability issues with MJJs. Dayton *et al* [16] demonstrated memory implementation with (also magnetic) π junctions, providing enhanced reliability and scalability. Nevertheless, including transformers in the design causes bulkiness and vulnerability to external noise.

Nanowires like quantum phase slip junctions offer the advantage of a small area [17] and the potential for high-density integration [18]. Nevertheless, a major drawback is the absence of a reliable and mature fabrication process for this technology, resulting in low read and write speeds for such memories. Consequently, despite their low power consumption, they fall short compared to cryo-CMOS memory in terms of access time, reliability, and integration density.

Superconducting spintronic technologies and other nonvolatile superconducting memories rely on the integration of ferromagnetic materials. While there are scientific reports and small-scale empirical studies on such devices, a commercial foundry for VLSI-scale fabrication is currently unavailable. Additionally, these designs require the incorporation of ferromagnetic materials into the barrier that separates the superconducting plates of a JJ, significantly increasing the complexity of the device stack and the associated margins [19, 20]. Ensuring uniform magnetic states across multiple devices on a chip further complicates scalability and consistency.

Another approach to superconductor memories is using a vortex, which stores bit data in the magnetic field generated by a current flow. Unlike pulse-based approaches, this technology does not necessitate JJ switching. It thus eliminates the need for recovery time after each switch, increasing the speed and power efficiency of the circuits. Vortex-based memory has also shown promising results in terms of its scalability. Tahara and Wada [21] demonstrated vortex-based NDRO in one of the earliest accounts of this memory. Numata *et al* [22] demonstrated a high-density memory application with a relatively lower area. Nagasawa *et al* [23] realized a 4 K-bit SFQ hybrid memory using a matrix array of vortex transitional memory (VTM) cells. However, while all the designs mentioned above showed improvement over the speed and integration of conventional SFQ memories, a vortex-based memory cell requires transformers to manage the data inside. That's why the area usage is still large, resulting in the main drawback of this type of memory structure. Komura *et al* [24] introduced a more efficient mutual coupling structure to reduce the area usage of vortex-based memory. Karamuftuoglu *et al* [25] reported a specific optimizer to increase the margin and decrease the area of the vortex-based memory cell. Semenov *et al* [26] demonstrated a large-scale implementation of vortex-based transitional memory cells. The area consumption was decreased by using self-shunted JJs with a higher critical current fabrication process. However, the mutual inductive coupling in the vortex-based memory made it large and susceptible to magnetic noise.

This paper presents a nonvolatile vortex-based memory cell without any transformer. We eliminate the mutual inductive coupling to decrease the memory size and increase its reliability. We demonstrate the correct functionality for both read and write operations. In contrast to superconducting (yet-to-be-developed) memory technologies, the proposed **bistable vortex memory** (BVM) utilizes conventional superconducting components, which guarantees robust memory performance across the chip. It also facilitates simplified read/write operations and has low overhead on peripheral designs, making it highly suitable for integration into superconducting electronics. We showed a 32×32 matrix array of vortex-based memory cells operating at 20 GHz. We also demonstrated analog computing using this design for an 8×8 matrix array at 20 GHz that can be used for fast and efficient artificial neural networks design.

The key contributions of this paper are as follows:

- Eliminating the transformers for reliable operation and higher density
- Nonvolatile Vortex memory cell
- Memory with a high operation frequency of 50 GHz with no static power consumption
- Current readout enabling analog computing on crossbar array

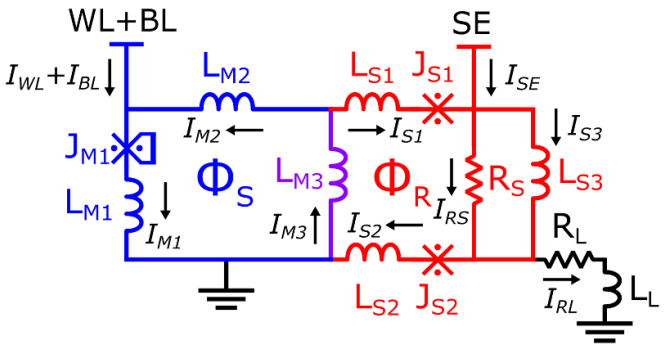


Figure 1. BVM cell model with row and column access lines. Here, the Storage Loop is drawn in blue, whereas the Readout Loop is drawn in red. The junction J_{M1} is the only shunted JJ here, and its symbol's vertical line represents the shunted resistor.

2. BVM

A vortex is a swirling motion of any substance characterized by a rotational flow. In Type-II superconductors, vortices are generated when a magnetic field begins to penetrate the superconducting material in the form of quantized flux. These vortices interact and can exhibit different phases influenced by factors such as the magnetic field, thermal fluctuations, and the pinning effect caused by disorder and defects within the material. Within the realm of superconducting memory design, a vortex transition refers to the movement or rearrangement of these magnetic vortices within the superconducting material. The presence or absence of vortices and their motion serve as the representations of binary data states. Applying a magnetic field or electrical current makes it possible to write, read, and manipulate the data stored within the memory cell based on the behavior of these vortices.

A VTM cell [23] consists of two superconducting loops and a two-junction interferometer that serves as a sense gate and is magnetically coupled to one of its loops. The cell's operation is controlled by two address signals, I_x and I_y, which determine the write and read operations. The control signals must be driven to have positive or negative amplitudes, allowing the data to be encoded with different polarities and values. The cell features a magnetically coupled port for a DC, I_{dc}, which shows that the VTM cell is a volatile memory. Additionally, a sense gate signal, I_s, acts as a clock signal, generating the output. In the VTM design, including transformers for input signals can significantly impact functionality, operating speed, and circuit area due to signal propagation times, physical size, and layout constraints. Therefore, in our proposed BVM cell, we eliminate these design components to simplify the design.

2.1. Analytical phase-current relations

A BVM cell consists of two connected SQUID loops. The first loop (called the Storage Loop) stores the data, whereas the second loop (called the Readout Loop) reads the state of the first loop, as shown in figure 1. Depending on the current direction, the cumulative current from the word line (WL) and the bit line (BL) will trap a state of 1 or 0 in the Storage Loop. The current from the sense-enable (SE) will cause the Readout Loop to generate a current at the load element (L_L).

Let's define the direction of the current flowing in the Storage Loop counterclockwise. Similarly, we define the current flowing in the Readout Loop clockwise. Further, the direction of bias currents is downward into the superconducting loops. Finally, the direction of the current flow in R_S and R_L is to GND. Based on these positive current direction definitions, we can then write the current equations for these loops as follows:

$$\begin{aligned} I_{M2}+I_{\mathrm{WL}}+I_{\mathrm{BL}} &= I_{M1} \\ I_{M3} &= I_{M2}+I_{S1} \\ I_{S1}+I_{\mathrm{SE}} &= I_{\mathrm{RS}}+I_{S3} \\ I_L &= I_{\mathrm{RS}}+I_{S3}-I_{S2}. \end{aligned} \tag{1}$$

For the JJs, we can write the DC Josephson equation to relate the JJs phase and current:

$$\begin{aligned} I_{M1} &= I^0_{M1}\times\sin(\varphi_{M1}) \\ I_{S1} &= I^0_{S1}\times\sin(\varphi_{S1}) \\ I_{S2} &= I^0_{S2}\times\sin(\varphi_{S2}) \end{aligned} \tag{2}$$

where I^0_* and φ_* denote the critical current and phase of the junction J_*. These equations connect the phase of the SQUID loops to their current. The phase should be a multiple of 2π for each SQUID loop. Therefore, we can write the phase equations for the storage and readout superconducting loops as follows:

$$\varphi_{M1}+I_{M1}L_{M1}+I_{M3}L_{M3}+I_{M2}L_{M2}=2M\pi \tag{3}$$

$$I_{M3}L_{M3}+I_{S1}L_{S1}+\varphi_{S1}+I_{S3}L_{S3}+\varphi_{S2}+I_{S2}L_{S2}=2N\pi \tag{4}$$

where M and N are positive or negative integer values. When $M=-1$, the internal flux of the SQUID loop, i.e. $\Phi_S=-\Phi_0$, the memory cell's storage state is zero, indicating that the stored value is zero. Conversely, when $M=1$, i.e. $\Phi_S=\Phi_0$, it represents a storage state of one for the memory cell, with the stored value being one. While a system of nonlinear equations for circuits with more than two JJs does not have a closed-form solution, it can still be solved numerically or approximated [27, 28].

2.2. Circuit design

As stated above, the BVM cell comprises two superconducting loops, denoted as the Storage Loop (or S-Loop for short) and the Readout Loop (or R-Loop for short). Each loop configuration is designed with a pair of JJs and inductance elements, as illustrated in figure 2. In this figure, the inductors L_{P*} within the BVM cell represents the parasitic inductance that is inherently introduced during the physical layout of the cell. L_{PM} component in the ground path is part of the S-Loop. Additionally, we have included other layout-induced parasitic inductances, i.e. L_{PRWL}, L_{PRBL} and L_{PRSE}, in the circuit schematic of figure 2. Note that figure 1 presents the BVM cell

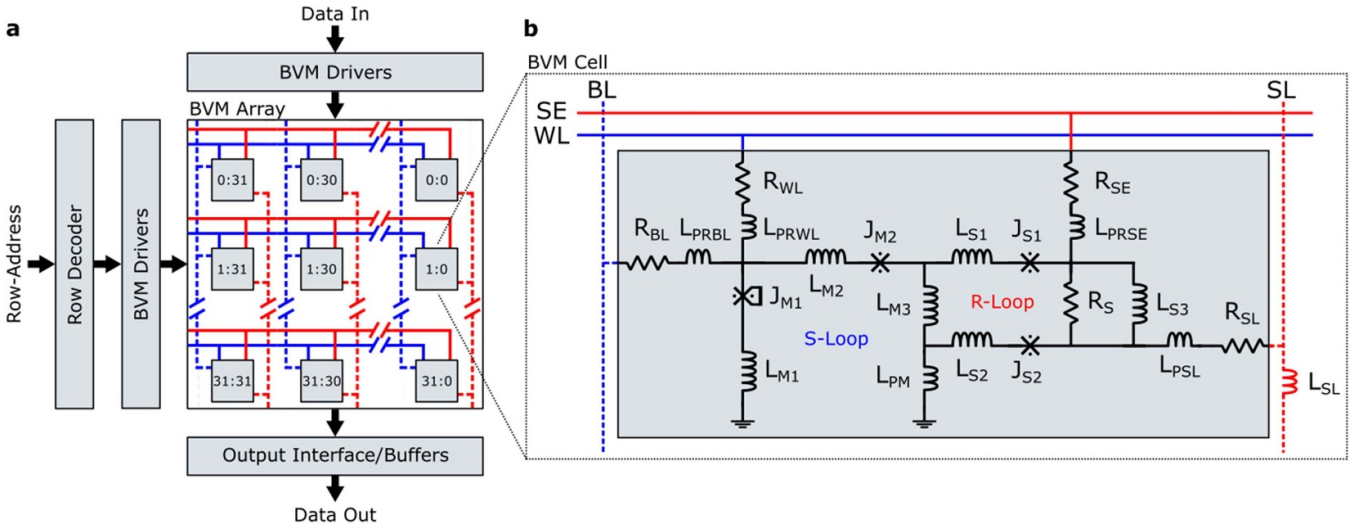


Figure 2. (a) The illustration of the BVM cell array with 32 rows and 32 columns. For a larger range of column bits, a column decoder may be necessary to select the correct address. (b) The BVM cell's schematic consists of 4 JJs with row and column access lines. The inductances named L_P correspond to the parasitics. The S-Loop stores the data, and the read operation is performed with the help of the R-Loop. ($R_{WL} = R_{BL} = R_{SE} = 20.0\,\Omega$, $R_S = 3.0\,\Omega$, $R_{SL} = 12.0\,\Omega$, $L_{PRWL} = L_{PRBL} = L_{PRSE} = L_{PM} = L_{S1} = L_{S2} = L_{S3} = L_{PSL} = 0.5$ pH, $L_{M1} = 12.5$ pH, $L_{M2} = 24.5$ pH, $L_{M3} = 8.5$ pH, $L_{SL} = 0.4$ pH, $J_{M1} = 120\,\mu$A, $J_{M2} = 140\,\mu$A, $J_{S1} = J_{S2} = 74\,\mu$A).

without the aforesaid parasitic inductances and the escape JJ (J_{M2}). The example of the connectivity of the memory array is given with a size of 32 by 32, providing details about the interconnections and data paths.

In the BVM cell, the write and read operations are achieved by the activity of JJs in the loops. By integrating these loops, the memory is designed to operate reliably at cryogenic temperatures (approximately 4 K), which is typical for Nb-based JJs that inherently require low temperatures to maintain their properties. The bistable vortex states are stable and resilient to minor fluctuations in these cryogenic temperatures. In other words, the BVM cells exhibit sensitivity to temperature variations comparable to that of the JJs that comprise them.

The cryogenic temperature requirement is inherently compatible with classical superconductor-based computing systems that function at 4 K, as well as quantum computing systems operating in the millikelvin temperature range, making the BVM a natural fit as a memory solution in such environments. It is important to note that temperature variations of say 1 K or so can alter the critical current of JJs, necessitating adjustments in the input currents for the address lines of the BVM array. To ensure proper read and write operations in the memory array, the drivers must be designed with output current levels that account for local hotspots.

The BVM unit cell stores 1 bit of data; thus, its states are represented by the direction of the circulating current within the S-Loop, either clockwise or counterclockwise. The write and read mechanisms were explained in detail in the previous section. This Storage Loop comprises J_{M1}, J_{M2}, and inductances. The stored data on the S-Loop can be read by checking the state of the R-Loop by switching its JJs without changing the stored data. During the read operation, the output current flows through R_{SL}, and the results can be observed on the sense line (SL).

One of the main advantages of the BVM is the absence of transformers. A transformer with a good coupling factor is bulky. The WL and the BL control currents are applied over R_{WL} and R_{BL} traces, corresponding to the row and column address signals, respectively. These address signals' values determine the write operation's loop direction. In our design, the influences between adjacent cells can be disregarded due to their spatial separation, which effectively isolates each cell from its neighbors. Because there is no mutual coupling, the separation of the SQUID loops with resistors enables us to maintain reliable bistability within each cell without interference from nearby cells. The sense enable (SE) signal provides additional current to the JJs of the R-Loop during the read cycle, which results in switching if the applied and stored currents flow in the same direction.

Moreover, the proposed memory cell is nonvolatile and requires no bias or offset current to maintain the stored data. Therefore, the design does not consume static power in the memory array. The simulations in the following sections are conducted utilizing the SPICE-based Josephson simulator (JoSIM) [29], incorporating inductive parasitics within the circuit designs. The models of our JJs are based on the SFQ5ee fabrication process developed for SCE at MIT Lincoln Laboratory (MITLL) [30]. More details are provided below.

2.3. Write operation

In vortex-based memory implementations, a limitation arises from the necessity of Boolean control signals to choose the write operation modes: write '0' and write '1'. This requirement hinders the simultaneous execution of write operations

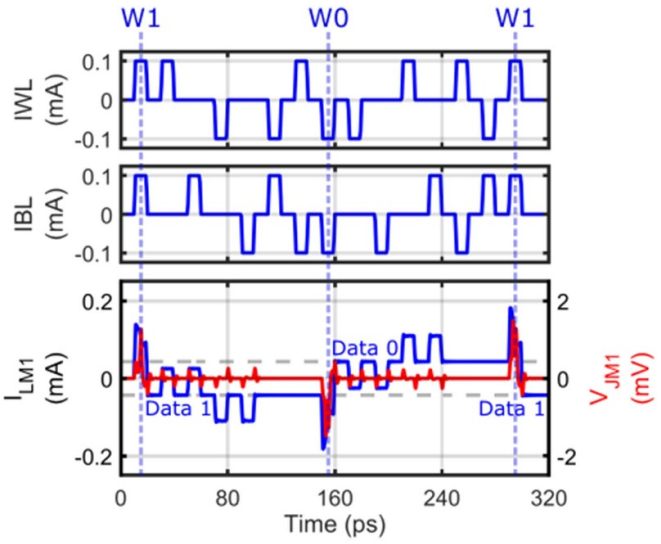


Figure 3. JoSIM result of the BVM cell for the write operation *W* with all possible control signal combinations. The simulation is performed at 50 GHz.

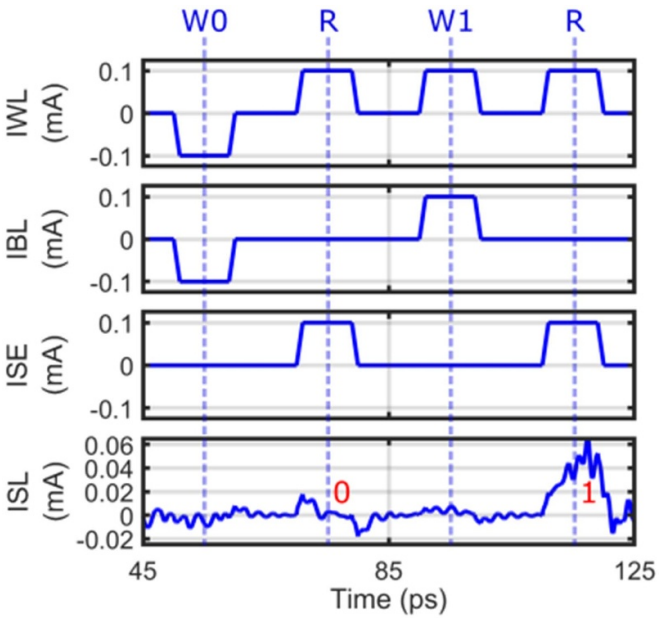


Figure 4. Simulation result of the BVM cell for the read operation *R*. The read is performed for each write operation *W* of '0' and '1', and the output current is observed on the SL. The simulation is performed at 50 GHz, and the load on the sense line is assigned as 12 non-switching junctions with 320 μA critical current value each.

for '0' and '1' within memory arrays. Consequently, the entire write process may need to be completed in two cycles, reducing the write frequency by half.

To change the current direction in the S-Loop, write operations are coded into different signal polarities on WL and BL control lines. When the total applied current of I_{WL} and I_{BL} surpasses the critical current of J_{M1}, the junction switches, leading to a change in the circulating current's direction. The junction J_{M2} in this design creates a DC-SQUID structure and is used as a non-switching element to prevent feedback from the readout to the storage section. This junction can be removed without a change in functionality (which would save area but also lower parameter critical margins) as explained in section 2.1. The induced current circulates clockwise if both I_{WL} and I_{BL} amplitudes are positive. Conversely, if they are negative, the current direction changes to counterclockwise. The rest of the address signal combinations are observed as half-selection, and the related operations are shown in figure 3.

I_{WL} and I_{BL} govern the operations of the memory cell, as discussed in section 2.1. In the analog simulation, the pulse widths of I_{WL} and I_{BL} are set to 10ps with a 50% duty cycle, resulting in a 50 GHz operating frequency. The initial write operation begins at 10ps, with the change in current flow within the S-Loop being monitored on L_{M1}. After the write '1' operation, we test various combinations of address signals corresponding to half-selection cases until 150 ps when the current direction changes due to the negative values of I_{WL} and I_{BL}, causing the write '0' operation. The memory state remains unchanged until the subsequent write '1' operation is executed.

2.4. Read operation

When the direction of the applied current is aligned with the circulating current in the Readout Loop, an output signal is generated. The I_{SE} polarity indicates the signal polarity in this process. A read-0 operation, where I_{SE} has a negative amplitude, results in an output current when the stored data is '0' and no current when the stored data is '1'. Conversely, a read-1 operation, where I_{SE} has a positive amplitude, generates an output current when the stored data is '1' and no output when the stored data is '0'. In this context, we used the polarity of stored data 1 for reading. Therefore, the absence of an output current when SE is applied corresponds to the stored data '0'. The level of SE current cannot be higher than $I_{WL} + I_{BL}$ or it can change the stored data. Maintaining the stored data in the S-Loop is critical to successfully carrying out the NDRO process.

Detecting the direction of the current flow in the S-Loop can be achieved by simply triggering the junctions in the R-Loop without altering the data stored in the S-Loop. To achieve this operation, we apply I_{WL} to select the WL and I_{SE} to sense the flowing current. As a result, the sense junctions J_{S1} and J_{S2} continuously switch, thereby generating output current directed towards the R_{SL}–L_{SL} path. The readout simulation is shown in figure 4.

We execute a series of write operations during the evaluation, alternating between 0 s and 1 s. Concurrently, we perform successful read operations between these write cycles at 50 GHz operating frequency. As a result, we observe the output current generated across the SL inductance, L_{SL}, when the stored data is '1'.

2.5. Design considerations and trade-offs

We analyze various inter-dependencies and trade-offs in the BVM cells. The SQUID structures offer flexibility that allows

modifications on the critical current (I_C) of JJs and loop inductance (L) to maintain the storage behavior in the S-Loop while satisfying the criterion for flux storage $I_C \times L > \Phi_0$. Adjustments may involve increasing the I_C of JJs while decreasing L, or vice versa. In many series cells (rows or columns), reducing the critical current of BVM's switching junction J_{M1} is the way to lower the required current on WL and BL. This approach helps with the scalability and peripheral circuit design.

Utilizing a single JJ (J_{M1}) is adequate for establishing an RF-SQUID Storage Loop structure for data storage. Adding a non-switching junction J_{M2} enhances stability in our design. Introducing the resistor R_S into the R-Loop in parallel with L_{S3} improves the robustness of the cell, albeit with a slight increase in the cell area. For layout design, utilizing multiple metal layers and substituting the inductance in the Storage Loop with kinetic inductance can decrease the area of the circuit to an estimated area of $10\,\mu\text{m} \times 10\,\mu\text{m}$. The proposed memory utilizes superconducting materials such as niobium, which are based on the available fabrication technology and compatible models. The specified cell size is based on MITLL fabrication technology, from which we utilized JJ models to perform the simulations. This approach ensures that the design incorporates essential functional characteristics and performance targets while remaining manufacturable within the MITLL process. It is important to note that parameter values can be adjusted while maintaining the circuit structure for fabrication technologies and foundries other than MITLL. With no mutual coupling required, an all-JJ cell design is feasible, offering significant area reduction. [31].

An alternative modification involves integrating WL and SE to create a unified control signal. This adjustment raises the magnitude of the required current in the control line but reduces the cell area usage. Consequently, during read operations, only the control line WL is activated. In such instances, simultaneous read and write operations can be achieved by activating BL and WL signals. Keeping these signals separate enhances controllability and distinguishes between read and write operations.

In the BVM design, the inductance L_{SL} on the SL column is crucial for signal propagation to the readout circuit. An alternative approach is to use a stack of non-switching JJs on the SL column to replace the inductance. Since the JJ stack does not require shunt resistance, it occupies a smaller area than L_{SL}, thereby reducing the circuit layout area and improving scalability within the BVM array. The time constant for the read operation on each row is determined by the ratio of $\sum_i^n L_{SL}^i$ to R_{SL}, where index i shows the row ID and n denotes the total row count. In this case, the inductance introduced by the JJs must be considered. In our design, these stacks of junctions are referred to as J_{SL}.

2.6. Memory array demonstration of 32 × 32

In the memory design featuring BVM cells, distinct write operations are necessary for storing data corresponding to '0' and '1'. To test the functionality and validate the performance of BVMs, a memory configuration of 2 × 2 has been designed using the same parameter set reported in figure 2, and the simulation result is shown in figure 5.

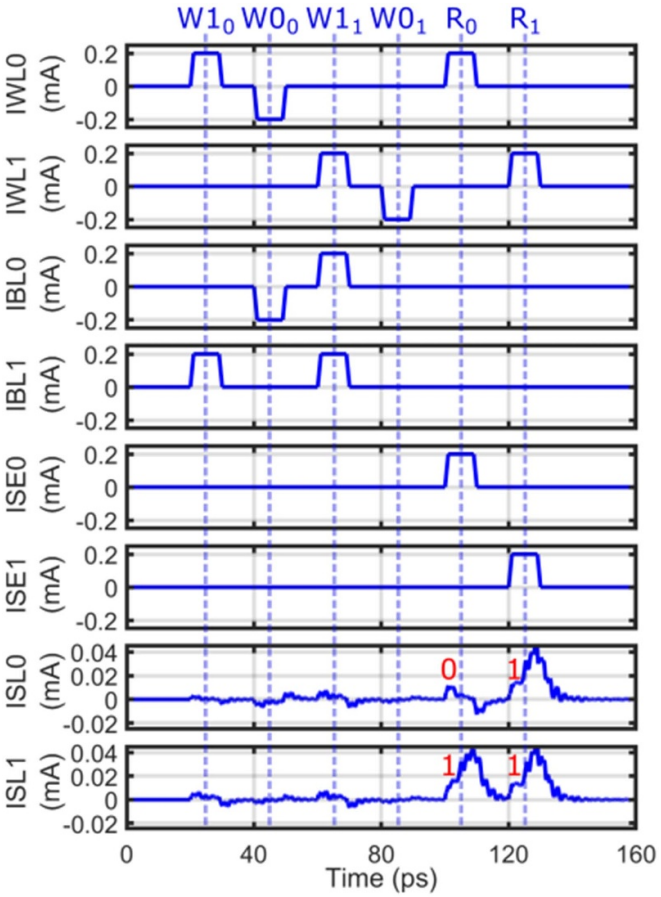


Figure 5. The simulation of the memory array with two rows and two columns. The write operation W is completed with two cycles per row, and the read operation is represented with R. The subscripts in the operation label correspond to the row index. The stored data in the first and second rows are '01' and '11'. The simulation is conducted at the frequency of 50 GHz, with each sense line having a load of 12 non-switching 320 μA junctions.

At the beginning of the simulation in figure 5, the consecutive write operations are performed for each row, collectively requiring four write cycles at 50 GHz. Upon storing the data in the rows, each row is individually read by enabling WL and SE signals. For each stored value of '1', the output current is observed on each SL, confirming the successful read operation and providing the result.

Even though one of the primary characteristics of the BVM is the ultra-fast access time, that is, how fast information can be written and read, it is not feasible to overlap signals with 10 ps width to achieve ultra-high-speed write/read operations. Therefore, the operating frequency was lowered from 50 GHz to 20 GHz for the memory array with 32 rows and 32 columns. Such frequency could enhance overall memory compatibility while facilitating more stable interaction between the

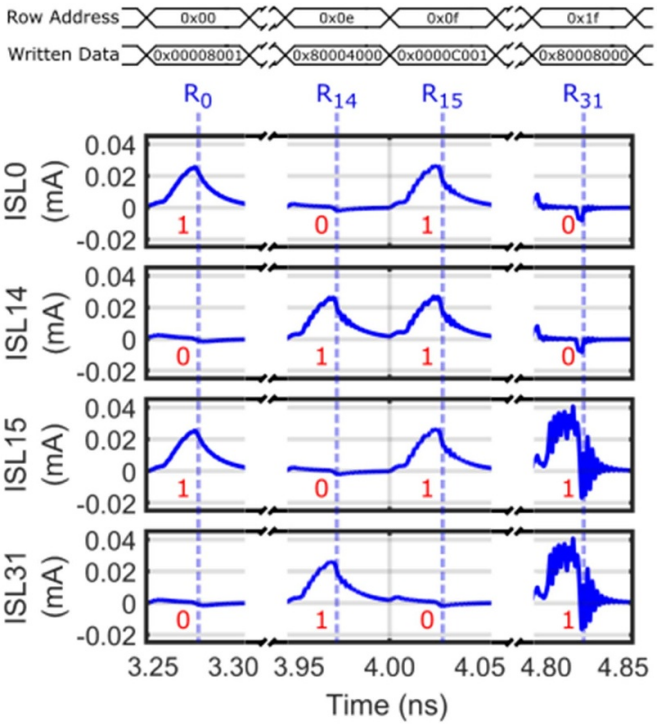


Figure 6. Simulation result for the memory array with 32 rows and 32 columns. R_0 corresponds to the read operation of row 0, and I_{SL0} represents the data on column zero. The simulation is performed at 20 GHz, and the load on the sense line is assigned as a non-switching junction with 320 μA.

memory cell and its peripherals. Moreover, the proposed circuit exhibits strong feasibility for large-scale integration, supporting the development of addressable memory banks. The write/read cycle time is stretched for the simulation shown in figure 6.

The simulation of a memory array with 32 rows and 32 columns operating at 20 GHz is shown in figure 6. In this design, R_{SL}, L_{PSL}, and L_{SL} are 16 Ω, 0.3 pH, and 0.3 pH, respectively. During the read operation, a current appears on the output path when the stored data is '1'; there will be no current when the stored data is '0'. If the time allowed for dissipating this current is insufficient, a residual current remains on the output path. Throughout the simulation, the rise/fall time of the input signals is slightly modified to mitigate the current residue between cycles while maintaining a 50% duty cycle. This approach ensures the integrity of the output data, although it results in a slight drop in the output's current level. Nonetheless, the distinction between '0' and '1' state outputs remains large enough for interface circuits. The operational characteristics and performance are confirmed with the successful completion of the reading process. Although there is some ripple in the output current of rows near the load, the current on the SL surpasses a certain amplitude. Therefore, the resultant current can be detected, and slight differences are ignored at the interface circuit, as explained in the following section.

3. Peripheral circuits

3.1. Driver

The peripheral's operating speed can be a bottleneck in high-speed BVM designs. One promising circuit solution is the amplified SFQ to DC converter [32], which generates the necessary current at high frequencies. This circuit serves as an interface between pulse-based circuits and BVMs. When an SFQ pulse is received, the circuit converts it to a voltage level, causing the intrinsic state of the SQUID to flip and allowing the circuit to generate current for the output peripherals. Each control signal driving the BVM is separately generated by its dedicated SFQ converter. This configuration enables each control signal to be produced by high-frequency SFQ circuits on the same chip.

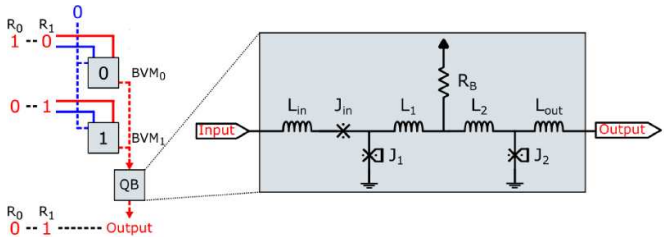
To further improve the scalability, a multistage superconductor voltage amplifier can be interfaced with BVM [32]. The amplifier circuit comprises a series of SQUID loops, wherein the SFQ-to-DC output can be connected to the amplifier input. A common DC bias is applied over the series of SQUIDs, thereby highlighting the cumulative effect of the series arrangement on the overall voltage output. Moreover, the total voltage observed is the product of the voltage across a single SQUID multiplied by the total number of stages in the series configuration. The driver circuit should support positive and negative amplitudes at its output to ensure accurate write operation. The differential nature of the SQUID amplifier allows us to put the ground point in the middle and supply both positive and negative inputs from a single circuit. Consequently, the amplifier circuit enhances the drivability of BVM and increases the number of BVMs that can be integrated into a single row or column.

Given the similarity in control signals between VTM and BVM cells over their respective address ports, the VTM's bipolar signal transmitter can also be utilized as a driver for BVM [25]. This transmitter consists of two identical components, each comprising JJs and resistors, with one being biased by positive current and the other by negative current. As a result, both positive and negative signals can propagate to the BVM cells upon activating either side of the transmitter.

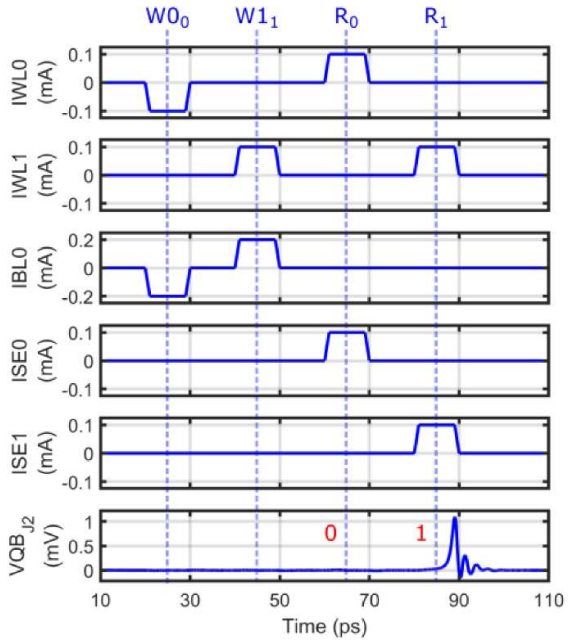
3.2. Readout

A peripheral circuit, such as a comparator, is required for analog-to-digital conversion. A readout quantizer buffer (QB) circuit converts the input current into an SFQ pulse in RSFQ technology [33]. The QB operates similarly to an asynchronous quasi-one junction SQUID circuit [2]. Ensuring the input current meets the SL's required current level is crucial. Once the input current exceeds a predetermined threshold, the circuit generates an SFQ pulse at the output, finalizing the conversion after the read operation on the BVM. The QB circuit is depicted in figure 7.

The QB circuit, which has a simple structure with three JJs, generates an SFQ pulse upon receiving an input current sufficient to excite J1. The QB circuit is modified to detect a



(a) Testbench for the BVM array with the readout QB circuit. The sense line has eight non-switching junction with 500 μA critical current value each.



(b) Simulation result at 50 GHz operating frequency.

Figure 7. Evaluation of the BVM array with a readout QB circuit. Data values 0 and 1 are written into the BVMs in rows 0 and 1, respectively. $W0_0$ represents the write-0 operation in row 0, whereas R_0 corresponds to the read operation in row 0. If the stored value in the BVM cell is 1, the QB cell generates a pulse at the output.

single BVM output in the test setup. Two BVMs are placed in a single column, with the SL connected to a single QB. Data values 0 and 1 are initially written to the BVMs in rows 0 and 1, respectively. Subsequently, each row is read sequentially, generating an SFQ pulse on the QB output when the BVM data is 1. This QB circuit successfully establishes the output interface for pulse-based SCE logic circuits.

In neuromorphic computing, activation functions involve comparing input signals to a specific threshold value to make decisions. These operations are performed by neurons, particularly their soma parts, which are responsible for processing and integrating input signals. Threshold circuits in neuromorphic systems that utilize superconducting components offer promising alternatives for achieving similar functionalities to those found in BVM readout circuits [34, 35]. To generate digital SFQ outputs, the BVM's SL can be connected to the input of a neuron's threshold gate, generating an SFQ pulse in response.

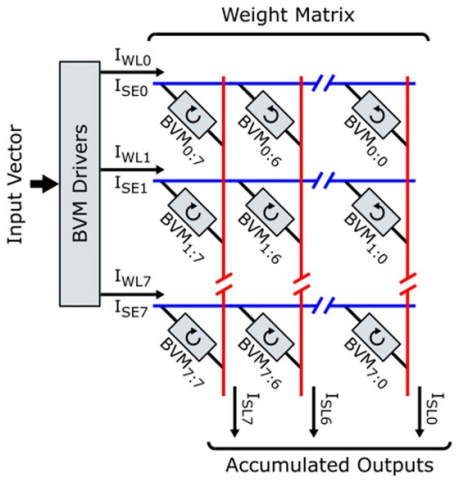


Figure 8. General concept illustration of accumulation operation. The test case incorporates the BVM cell array with eight rows and eight columns. Unlike the general memory purposes, multiple rows are simultaneously read, and as a result, the current accumulation is observed on each SL.

4. Analog computing with BVM crossbar arrays

4.1. Structure and operation

Within the framework of conventional computing systems, performing operations on data involves transferring data to a processing unit and its subsequent return to memory for writing. As a result, the fetch cycle leads to significant latency and energy consumption. For analog computing, the operations are carried out within a computational memory unit by leveraging the intrinsic properties of memory devices to eliminate the requirement for data transfers to the processing unit via data buses.

Each memory unit generates an output current on the SL towards a common ground in the BVM array. Since the SL resistor R_{SL} isolates each memory unit within the same column, there is no interaction among the memory units while the output current is generated. Therefore, by enabling multiple rows with SE signal application, a simultaneous output current generation event occurs, resulting in a summation of multiple row outputs on the SL.

To demonstrate the capabilities of BVM, we designed 8 × 8 memory as a testbench shown in figure 8. The current on the load inductance at the end of the SL is used to dictate the accumulation operation. After writing the data into corresponding rows, the simulation reads individual lines. In the final test case, all rows are simultaneously read, and the current accumulation is observed, confirming the applicability of BVMs for the applications of in-memory computing and neural network computations. The related simulation result showing the read operations is presented in figure 9.

Throughout the simulation at 20 GHz operating frequency, the written data across rows zero to seven exhibits a decreasing number of 1 s as the row index moves from one to the next. Write-0 operations for all rows can be accomplished in just one cycle due to simultaneous access of multiple rows. The

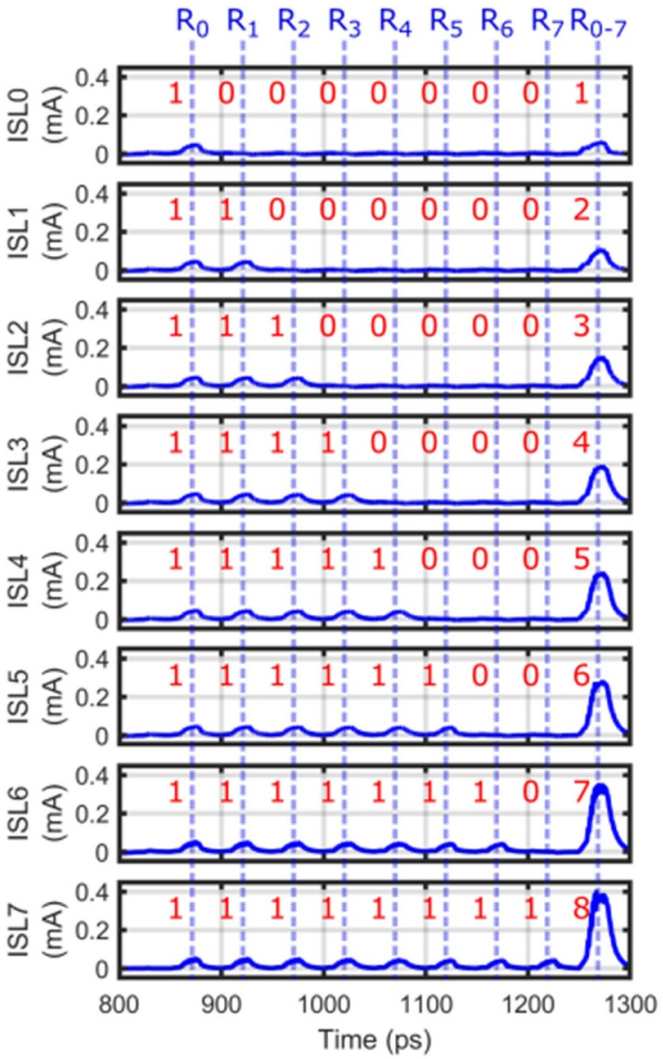


Figure 9. Simulation result of BVM array with enabled multiple rows. R_0 corresponds to the read operation of row zero, and I_{SL0} represents the data on column zero. The written numbers represent the output amplitude level. The simulation is performed at 20 GHz, and the load on each sense line is assigned as 12 non-switching junctions with 500 μA.

remaining write-1 operations can be completed by accessing one row per cycle. Upon performing each read operation R on a single row, a change in the SL current is observed if the stored data is '1'. Subsequently, at the end of the simulation, where all rows are accessed simultaneously for the read operation R_{0-7}, different output current levels, depending on the number of stored '1' s, are observed in the SL. Such functionality highlights BVM's computing capabilities. In this simulation, the BVM parameters are kept the same values reported in figure 2.

With its inherent two-state operation, the BVM cell is naturally suited for bitwise operations. In the previous example, each row is assigned to a single input bit, allowing for simple and efficient bitwise operations. The states can be modulated to store multi-bit weights in a neural network and thus enable multi-bit computations. Indeed, multiple rows in the BVM array can be allocated to a single weight.

4.2. Design considerations for BVM-based crossbars

BVMs for low-power artificial intelligence applications via analog computing create opportunities and pose challenges. Due to fan-in and fan-out limitations on the superconductor hardware, our BVM-based array is much smaller than the standard MAC designs to perform similar operations.

Within the crossbar array, the load on the SL is a crucial factor for each column. Such a design can consist of a single inductor L_{SL} or a stacked set of junctions J_{SL} with a critical current surpassing the maximum level of the achievable accumulated current on the SL. In the case of J_{SL} junctions without an adequate level of critical current, switching events occur during the accumulation, causing an error in the output. The accumulation operation requires a long enough timeframe, depending on the number of accumulated row values. In other words, the accumulation timeframe must increase linearly with the number of accumulating rows, thus decreasing the operating frequency.

5. Conclusion

This paper introduced BVM as a compact, efficient, and high-performance solution for SCEs. BVM offers zero static power consumption, NDRO, and nonvolatile characteristics. Its simple, transformer-free structure contributes to its compactness and scalability. We demonstrated a 32 $\times$ 32 memory array operating at 20 GHz. Additionally, BVM arrays can perform column addition in a single cycle, making the proposed design an efficient in-memory and near-memory computational unit for matrix-vector multiplication, neural networks, and combinatorial logic implementations.

Data availability statement

The data that support the findings of this study are available at the following URL: https://github.com/Karamuft/Bistable_Vortex_Memory.

Acknowledgment

This work has been funded by the National Science Foundation (NSF) under the project Expedition: (Design and Integration of Superconducting Computation for Ventures beyond Exascale Realization) project with Grant No. 2124453.

ORCID iDs

M A Karamuftuoglu https://orcid.org/0000-0002-0951-1697
B Z Ucpinar https://orcid.org/0009-0001-7398-8844
S Razmkhah https://orcid.org/0000-0003-0029-2800
M Pedram https://orcid.org/0000-0002-2677-7307

References

[1] IEEE International Roadmap for Devices and Systems 2023 Beyond CMOS and emerging materials integration, Institute of electrical and electronics engineers (https://doi.org/10.60627/0P45-ZJ55)

[2] Razmkhah S and Febvre P 2023 Superconducting Quantum Electronics *Beyond-CMOS* (ISTE & Wiley) pp 295–391

[3] Holmes D S, Ripple A L and Manheimer M A 2013 Energy-efficient superconducting computing-Power budgets and requirements *IEEE Trans. Appl. Supercond.* **23** 1701610–1701610

[4] Schneider M, Toomey E, Rowlands G, Shainline J, Tschirhart P and Segall K 2022 SuperMind: a survey of the potential of superconducting electronics for neuromorphic computing *Supercond. Sci. Technol.* **35** 053001

[5] Karamuftuoglu M A, Ucpinar B Z, Razmkhah S, Kamal M and Pedram M 2024 Unsupervised SFQ-based spiking neural network *IEEE Trans. Appl. Supercond.* **34** 1–8

[6] Huang H L, Wu D, Fan D and Zhu X 2020 Superconducting quantum computing: a review *Sci. China Technol. Sci.* **63** 180501

[7] Stolz R, Schmelz M, Zakosarenko V, Foley C, Tanabe K, Xie X and Fagaly R 2021 Superconducting sensors and methods in geophysical applications *Supercond. Sci. Technol.* **34** 033001

[8] Natarajan C M, Tanner M G and Hadfield R H 2012 Superconducting nanowire single-photon detectors: physics and applications *Supercond. Sci. Technol.* **25** 063001

[9] Mukhanov O 1993 RSFQ 1024-bit shift register for acquisition memory *IEEE Trans. Appl. Supercond.* **3** 3102–13

[10] Xu W, Ying L, Lin Q, Ren J and Wang Z 2021 Design and implementation of bit-parallel RSFQ shift register memories *Supercond. Sci. Technol.* **34** 085002

[11] Kirichenko A, Sarwana S, Brock D and Radpavar M 2001 Pipelined DC-powered SFQ RAM *IEEE Trans. Appl. Supercond.* **11** 537–40

[12] Nagasawa S, Hinode K, Satoh T, Kitagawa Y and Hidaka M 2006 Design of all-dc-powered high-speed single flux quantum random access memory based on a pipeline structure for memory cell arrays *Supercond. Sci. Technol.* **19** S325

[13] Ghoshal U, Kroger H and Van Duzer T 1993 Superconductor-semiconductor memories *IEEE Trans. Appl. Supercond.* **3** 2315–8

[14] Van Duzer T, Zheng L, Whiteley S R, Kim H, Kim J, Meng X and Ortlepp T 2013 64-kb hybrid Josephson-CMOS 4 Kelvin RAM with 400 ps access time and 12 mW read power *IEEE Trans. Appl. Supercond.* **23** 1700504

[15] Vernik I V, Bol'ginov V V, Bakurskiy S V, Golubov A A, Kupriyanov M Y, Ryazanov V V and Mukhanov O A 2013 Magnetic Josephson junctions with superconducting interlayer for cryogenic memory *IEEE Trans. Appl. Supercond.* **23** 1701208

[16] Dayton I M *et al* 2018 Experimental demonstration of a Josephson magnetic memory cell with a programmable π-junction *IEEE Magn. Lett.* **9** 1–5

[17] Butters B A, Baghdadi R, Onen M, Toomey E A, Medeiros O and Berggren K K 2021 A scalable superconducting nanowire memory cell and preliminary array test *Supercond. Sci. Technol.* **34** 035003

[18] Zhao Q-Y, Toomey E A, Butters B A, McCaughan A N, Dane A E, Nam S-W and Berggren K K 2018 A compact superconducting nanowire memory element operated by nanowire cryotrons *Supercond. Sci. Technol.* **31** 035009

[19] Fert A 2008 Nobel lecture: origin, development and future of spintronics *Rev. Mod. Phys.* **80** 1517–30

[20] Mel'nikov A S, Mironov S V, Samokhvalov A V and Buzdin A I 2022 Superconducting spintronics: state of the art and prospects *Phys. Usp.* **65** 1248–89

[21] Tahara S and Wada Y 1987 A vortex transitional NDRO Josephson memory cell *Jpn. J. Appl. Phys.* **26** 1463

[22] Numata H, Nagasawa S and Tahara S 1997 A vortex transitional memory cell for 1-Mbit/cm/sup 2/density Josephson RAMs *IEEE Trans. Appl. Supercond.* **7** 2282–7

[23] Nagasawa S, Hashimoto Y, Numata H and Tahara S 1995 A 380 ps, 9.5 mW Josephson 4-Kbit RAM operated at a high bit yield *IEEE Trans. Appl. Supercond.* **5** 2447–52

[24] Komura Y, Tanaka M and Nagasawa S Bozbey A and Fujimaki A 2015 Vortex transitional memory developed with Nb 4-layer, 10-kA/cm^2 fabrication process *2015 15th Int. Superconductive Electronics Conf. (ISEC)* (IEEE) pp 1–3

[25] Karamuftuoglu M A, Demirhan S, Komura Y, Çelik M E, Tanaka M, Bozbey A and Fujimaki A 2016 Development of an optimizer for vortex transitional memory using particle swarm optimization *IEEE Trans. Appl. Supercond.* **26** 1–6

[26] Semenov V K, Polyakov Y A and Tolpygo S K 2019 Very large scale integration of Josephson-junction-based superconductor random access memories *IEEE Trans. Appl. Supercond.* **29** 1–9

[27] Akgun A and Razmkhah S 2022 PySQIF, a statistical analysis tool for bi-SQUID magnetometers *J. Phys.: Conf. Ser.* **2323** 012026

[28] Ilin S, Razmkhah S, Yilmaz U, Protsenko O and Febvre P 2019 Static Analysis to predict behaviour of Bi-SQUIDs in external magnetic field *2019 IEEE 9th Int. Conf. Nanomaterials: Applications & Properties (NAP)* (IEEE) pp 02TM09–1

[29] Delport J A, Jackman K, Roux P l and Fourie C J 2019 JoSIM-superconductor SPICE Simulator *IEEE Trans. Appl. Supercond.* **29** 1–5

[30] Tolpygo S K, Bolkhovsky V, Weir T J, Johnson L M, Gouker M A and Oliver W D 2015 Fabrication process and properties of fully-planarized deep-submicron Nb/Al–AlO$_x$/Nb Josephson junctions for VLSI circuits *IEEE Trans. Appl. Supercond.* **25** 1–12

[31] Cong H, Razmkhah S, Karamuftuoglu M A and Pedram M 2024 Superconductor logic implementation with all-jj inductor-free cell library *IEEE Trans. Appl. Supercond.* **34** 1–10

[32] Razmkhah S, Bozbey A and Febvre P 2021 A compact high frequency voltage amplifier for superconductor-semiconductor logic interface *Supercond. Sci. Technol.* **34** 045013

[33] Razmkhah S, Karamuftuoglu M A and Bozbey A 2024 Hybrid synaptic structure for spiking neural network realization *Supercond. Sci. Technol.* **37** 065011

[34] Karamuftuoglu M A, Ucpinar B Z, Fayyazi A, Razmkhah S, Kamal M and Pedram M 2024 Scalable superconductor neuron with ternary synaptic connections for ultra-fast SNN hardware

[35] Crotty P, Schult D and Segall K 2010 Josephson junction simulation of neurons *Phys. Rev.* E **82** 011914