

RESEARCH ARTICLE OPEN ACCESS

Enabling Quantum-Compatible Nonvolatile Memory: Cryogenic Evaluation of 1T1R HfO₂-Based RRAM From 300 to 1.5 K

Emilio Pérez-Bosch Quesada¹  | Alberto Mistrone¹  | Ruolan Jia² | Keerthi Dorai Swamy Reddy¹  | Markus Fritscher^{1,3}  | Felix Reichmann¹  | Oliver Skibitzki¹  | Juan Bautista Roldán⁴  | Helena Castán⁵  | Salvador Dueñas⁵  | Christian Wenger^{1,3}  | Eduardo Pérez^{1,3} 

¹IHP - Leibniz Institute for High Performance Microelectronics, Frankfurt (Oder), Germany | ²Micro- and Nanosystems Technology, Technical University of Munich, Munich, Germany | ³BTU Cottbus-Senftenberg, Cottbus, Germany | ⁴Department of Electronics and Computer Technology, University of Granada, Granada, Spain | ⁵Department of Electronics, University of Valladolid, Valladolid, Spain

Correspondence: Emilio Pérez-Bosch Quesada (quesada@ihp-microelectronics.com)

Received: 5 November 2025 | **Revised:** 17 April 2026 | **Accepted:** 29 July 2026

Keywords: cryogenic memory | cryogenic | in-memory computing | multi-bit memory | multilevel-cell | nonvolatile memory | RRAM

ABSTRACT

To enable compact quantum computing systems, cryogenically compatible CMOS-based non-volatile memories are required. In this work, we investigate 1-transistor–1-resistor (1T1R) resistive random-access memory (RRAM) devices over an extended temperature range, including previously unexplored cryogenic regimes. To enhance storage density, we explore a multilevel-cell approach via multilevel programming. As temperature decreases, the low-resistance state (LRS) effective conductance increases, improving state separation but also leading to higher LRS currents and increased reset voltages required for reliable switching to the high-resistance state (HRS), resulting in incomplete filament rupture and a broader HRS distribution. Through independent characterization of the access nMOS transistor, we identify transistor temperature dependence as the dominant factor governing LRS evolution, while extraction of the intrinsic metal–insulator–metal (MIM) behavior reveals an opposite trend, with decreasing conductivity at lower temperatures. Based on these insights, we implement an optimized programming strategy for cryogenic operation by tuning the nMOS gate bias during set, reducing LRS currents, particularly in intermediate and high conductance states, while preserving state separation, narrowing the HRS distribution, and lowering reset voltages. Overall, this work establishes a robust and scalable approach for reliable 1T1R operation at cryogenic temperatures and provides a key step toward nonvolatile memory integration in quantum computing architectures.

Abbreviations: 1T1R, 1-Transistor–1-Resistor; ALD, Atomic layer deposition; ANN, Artificial neural network; BEOL, Back-end-of-line; BL, Bit line; CF, Conductive filament; CMOS, Complementary metal–oxide–semiconductor; ELT, Enclosed layout transistor; ESD, Electrostatic discharge; HRS, High-resistance state; IMC, In-memory computing; KDE, Kernel density estimation; LRS, Low-resistance state; MIM, Metal–insulator–metal; MLC, Multilevel cell; NISQ, Noisy intermediate-scale quantum; NVM, Nonvolatile memory; PDF, Probability density function; RRAM, Resistive random-access memory; SCLC, Space-charge-limited conduction; SL, Source line; SMU, Source-measure unit; TE, Top electrode; TMA, Trimethylaluminum; WL, Word line; IQR, Interquartile range.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Advanced Electronic Materials* published by Wiley-VCH GmbH

1 | Introduction

The growing interest in quantum computing is driving transformative advancements across multiple scientific and technological domains. This novel approach leverages the principles of quantum mechanics, such as superposition and entanglement, to process information, potentially revolutionizing fields such as material science, drug discovery, and climate modeling. However, while near-term quantum computers currently operating in the noisy intermediate-scale quantum (NISQ) era [1] have demonstrated greater efficiency than classical computers for specific tasks [2], they cannot yet perform complex computations independently due to their noise levels, limited qubit counts, and short coherence times. Thus, hybrid classical-quantum solutions arise to combine the strengths of both computational frameworks to solve complex problems [3–5]. That is, quantum computers are proficient at solving specific tasks related to optimization, cryptography, and molecular simulations, whereas classical computers are well-suited for general-purpose computation, robust data storage, and error correction.

This hybrid approach demands efficient complementary metal-oxide-semiconductor (CMOS)-based circuits to bridge the classical domain, where most of the information is processed and errors are corrected, and the quantum domain, where the qubits reside. These circuits typically operate at room temperature (300 K), whereas the qubits are maintained at near-absolute zero temperatures and are commonly controlled individually using high-density cryogenic flex I/O wiring across a wide temperature gradient [2, 6]. Each qubit requires precise tuning and recalibration to tune them into a good operating regime, a process that can take up to 3 h per device when done manually, adding significant complexity and cabling overhead [7–9].

To address crucial challenges in quantum chemistry or code-breaking, current estimates suggest that the number of required physical qubits lies in the range of 10^6 – 10^8 , based on existing quantum algorithms and error correction techniques [10, 11]. Given these constraints, scalability remains one of the major hurdles in quantum computing, as individual and manual qubit control becomes impractical at scale. Furthermore, as the number of qubits increases, so do challenges such as error rates, decoherence, and connectivity, making the development of large, fault-tolerant systems increasingly difficult.

Bringing memory and processing units closer to the qubits, ideally within the same cryogenic environment, would help mitigate data transfer bottlenecks, reduce computational power consumption, and enable on-chip control of large-scale qubit arrays, improving the scalability of quantum computers. In-memory computing (IMC) emerges as a particularly well-suited solution for this challenge [12]. Inspired by the brain's architecture, IMC overcomes the energy inefficiencies of classical systems caused by costly data transfers between physically separate memory and processor units. By leveraging emerging nonvolatile memories (NVMs), the IMC paradigm enables in situ computing, performing operations directly within memory arrays and reducing the data volume exchanged between cryogenic qubits and room-temperature processors [13]. Among emerging NVM technologies, HfO₂-based resistive random-access memory

(RRAM) stands out due to its fast switching (< 10 ns), long retention (> 10 years), large endurance ($> 10^{12}$ operations), CMOS compatibility, and small footprint [14]. In contrast to other memory technologies, its nonvolatile operation eliminates the need for refresh cycles, significantly reducing standby power consumption, while its low switching energy and simple two-terminal structure enable dense integration with reduced thermal budget. These features make RRAM particularly attractive for cryogenic and quantum-compatible hardware, where strict constraints on power dissipation, proximity to qubits, and system scalability must be met. The widely adopted 1-transistor-1-resistor (1T1R) structure enhances RRAM performance by suppressing sneak-path currents and improving access control [15, 16], while multilevel-cell (MLC) operation allows multiple bits per cell storage, boosting potential for AI acceleration and memory integration [17]. We have previously demonstrated the capability of IHP's RRAM for storing multiple bits per cell via the MLC approach [16, 17] and their suitability for IMC tasks, including efficient vector–matrix multiplications [18, 19], a key operation in AI models.

RRAM-based IMC techniques are especially promising for hardware-based automatic qubit tuning performed near the physical qubits, enabling fast, localized decision-making with minimal latency [20]. Demonstrations of machine learning techniques for charge autotuning in quantum dots are provided by Czischek et al. [21], who explored compact feed-forward neural network architectures that are simple enough to be realistically implemented in memristor crossbar arrays, making them a promising candidate for on-chip artificial neural network (ANN) applications. Furthermore, Yon et al. [22, 23] propose integrating their automatic calibration method, based on pattern recognition in stability diagrams, near the quantum dots at the 4 K stage of the cryogenic environment, and transferring the trained ANN to cryo-compatible hardware such as arrays of memristive devices.

State-of-the-art CMOS memories and non-superconducting memories in general are typically deployed at higher temperatures (77–120 K) to balance integration and operability costs [24, 25]. Prior demonstrations of HfO₂-based RRAM switching at cryogenic temperatures support the integration of RRAM-based memory modules at even lower cryostat stages, closer to the quantum devices [26, 27]. For instance, Ram et al. [28] demonstrated that operating RRAM devices at 77 K strongly suppresses read noise and enhances the analog dynamic range by more than an order of magnitude, owing to reduced stochasticity and improved heat confinement in the conductive metal oxide layer. However, achieving bi-level resistance switching at even lower temperatures, such as 4 K, has been reported in only a few studies [28–30]. Although these studies contribute significantly, a comprehensive understanding of device behavior at near-zero temperatures remains to be established, leaving important challenges in ensuring consistent device performance at qubit-relevant temperatures.

In our previous work, we demonstrated for the first time the forming and resistive switching of CMOS-compatible RRAM devices at 1.45 K [31], establishing a new operative temperature limit for HfO₂-based RRAMs in standard 130 nm CMOS technology. That study primarily focused on a proof-of-concept demonstration

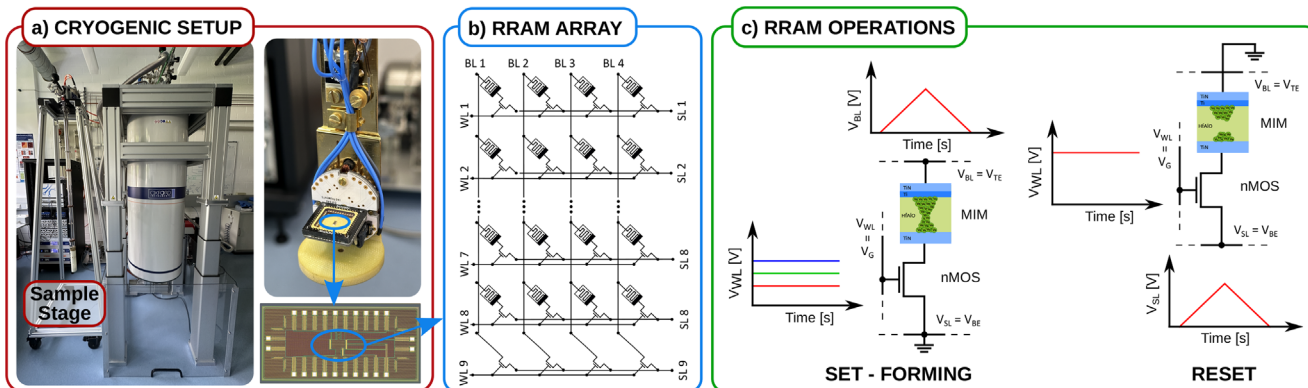


FIGURE 1 | (a) Cryogenic setup used to characterize the RRAM devices under study, (b) RRAM array containing a total of 32 1T1R devices individually addressable, and (c) schematic description of the forming, set, and reset operations carried out during this work.

using a limited dataset and devices based on thicker, undoped HfO_2 layers. In contrast, the present work investigates a new generation of devices employing 5% Al-doped HfO_2 [17], which exhibit significantly improved performance, including reduced variability and enhanced switching stability at both room and cryogenic temperatures. In addition, the experimental scope has been substantially expanded by increasing the number of devices under test and the number of switching cycles per device, enabling a more statistically meaningful evaluation of device behavior.

Building on this, we perform a comprehensive and systematic analysis of temperature-dependent resistive switching across a wide range from 300 K down to 1.5 K. Measurements are carried out at 10 distinct temperature points, allowing a quantitative assessment of the evolution of multilevel resistance states, complemented by dedicated resistive state evolution experiments over 20 intermediate temperature steps. This work further includes a deeper and more statistically robust investigation of the underlying conduction mechanisms of the metal–insulator–metal (MIM) stack, reinforcing and extending the findings of the previous study. Overall, the present manuscript provides a significantly broader and more detailed perspective on 1T1R operation under cryogenic conditions, highlighting the dominant role of transistor-related effects in limiting performance and guiding the optimization of programming strategies at ultra-low temperatures.

2 | Methods

In this section, we provide a detailed description of the methods employed to obtain the results presented above. We begin with the fabrication and structure of the samples, followed by the initialization procedure of the 1T1R devices. Finally, we describe the experimental setup used to cool down the devices under test and to carry out the electrical characterization.

2.1 | Device Fabrication and Structure

The RRAM devices investigated in this work feature a 1T1R configuration fabricated on 200 mm wafers using IHP's 130 nm

BiCMOS technology node [32]. Each cell integrates an n -channel enclosed layout transistor (ELT) with equivalent dimensions $W = 8 \mu\text{m}$ and $L = 0.66 \mu\text{m}$ connected in series with a MIM stack with $600 \times 600 \text{ nm}^2$ area, located between metal layers 2 and 3 in the back-end-of-line (BEOL) process. The bottom electrode of the MIM structure consists of a 150 nm TiN layer deposited by magnetron sputtering. The insulator layer is then deposited by atomic layer deposition (ALD) and consists of a 4 nm-thick $\text{Hf}_{1-x}\text{Al}_x\text{O}_y$ with Al doping level of 5%. This is achieved by co-depositing HfO_2 and Al_2O_3 in a controlled manner: a single trimethylaluminum (TMA) pulse is introduced after every 5 ALD cycles of HfO_2 to obtain 5% doping. The incorporation of Al into HfO_2 has been shown to improve device performance, yielding higher forming efficiency, reduced variability, and enhanced endurance and retention characteristics. The selected Al concentration of 5% represents a trade-off between performance and process complexity, as previous studies [17] investigating higher doping levels (e.g., 7%) did not demonstrate significant additional improvements.

The RRAM structures are arranged in a 9×4 array, where the 1T1R configuration allows for individual cell selection and precise control of the compliance current through the gate voltage (V_G), which is essential for reliable MLC operation since it governs filament formation and thus the reproducibility of intermediate resistance states. To enable characterization of the access transistor response at cryogenic temperatures, the last row of the arrays was intentionally left without MIM structures. To minimize temperature-related effects introduced by additional components, the array was fabricated without peripheral circuitry. As a result, the bit lines (BLs), word lines (WLs), and source lines (SLs) are directly routed to external pads (see Figure 1b). For clarity, we denote the voltage applied to the individual 1T1R cell as V_{TE} , the voltage applied to the TE through the addressed BL, V_S as the voltage applied to the source terminal through the addressed SL, and V_G as the voltage applied to the gate terminal through the corresponding WL. To protect the thin insulating layer of the MIM structures from electrostatic discharge (ESD) damage, anti-ESD diodes were incorporated into each pad connection. The arrays were then diced and bonded onto an LCC-44 sample carrier, which was subsequently mounted on the sample probe and loaded into an Oxford Instruments TeslatronPT system (see Figure 1a).

TABLE 1 | Electrical parameters used to characterize the I - V response of the RRAM devices under study.

Operation	V_{TE} (step) (V)	V_s (step) (V)	V_G (step) (V)	V_G - optimized (V)
I - V transistor	0.0–3.0 (0.01)	0.0	0.0–3.0 (0.1)	—
Forming	0.0–3.0 (0.01)	0.0	1.4	1.3
Reset	0.0	0.0–1.5 (0.01)	3.3	3.3
Set	0.0–1.5 (0.01)	0.0	1.4	1.3
MLC	0.0–1.5 (0.01)	0.0	1.2, 1.4, 1.6	1.2, 1.3, 1.5

Note: The optimized parameters make reference to the V_G adjustment to improve the stability of the resistive switching at cryogenic temperature.

2.2 | Device Initialization and Resistive Switching Protocol

We enabled resistance switching in the RRAM devices by performing the forming operation, which consisted of a voltage ramp applied to the BL terminal of the addressed device while simultaneously biasing the corresponding WL terminal. This process triggered a soft electrical breakdown, driving oxygen ions out of the HfAlO_2 layer and forming an oxygen-vacancy-based conductive filament (CF), thus switching the device to the LRS.

This was followed by the first reset operation, where a voltage ramp was applied to the SL terminal of the addressed device. The corresponding WL terminal was set to its maximum nominal value (3.3 V) to allow maximum current flow through the cell, disrupting the CF and switching the device to the HRS [33].

The set operation reverses this process by applying a voltage ramp to the BL terminal of the addressed device, re-establishing the CF, and returning the device to the LRS. By tuning the WL voltage bias during this operation, an MLC approach is enabled: different compliance limits imposed by the nMOS transistor set the RRAM devices into distinct LRS levels. The higher the voltage bias applied to the transistor gate via the WL terminal, the more conductive the resulting LRS, due to the different current compliance levels imposed by the transistor. The RRAM electrical characterization was conducted applying positive voltage sweeps with a voltage ramp rate of 300 mV s^{-1} . The electrical parameters are summarized in Table 1 and schematically illustrated in Figure 1c.

2.3 | Experimental Conditions and Instrumentation

For the temperature-dependent resistive switching tests, measurements were conducted at 10 temperature points ranging from 300 K down to 1.5 K, following a logarithmic distribution within the operational range of the TeslatronPT cryogenic measurement system (Oxford Instruments). For the readout characterization, 20 temperature steps were selected using the same logarithmic spacing approach.

Electrical characterization was performed using a two-channel Keysight B2902B source-measure unit (SMU), a Keithley 2230g-30-1 power supply, and a Keysight 34980A switching matrix to individually address RRAM devices within the array.

3 | Results

The primary objective of the electrical characterization conducted in this work is to evaluate and optimize MLC RRAM operations at cryogenic temperatures previously unexplored with this type of memory technology. This study lays the groundwork for the development of cryogenically compatible NVM modules, which are essential for applications such as compact quantum computing systems, especially given the current lack of reliable memory technologies suitable for such environments [25]. First, we investigated the temperature effects on 1T1R switching operations and how the nMOS transistor influences the programmed resistance state. This is followed by an investigation isolating the intrinsic temperature dependence of the MIM resistance, performed by measuring the programmed devices during controlled temperature steps with minimal disturbance to their resistance states. Together, these tests provide a comprehensive understanding of how temperature affects both the transistor and the resistive switching elements within the 1T1R structure, enabling the implementation of an optimized switching protocol for cryogenic operation by adjusting the nMOS gate voltage during the set process.

3.1 | Temperature Effects on 1T1R Switching

We initially performed a comprehensive electrical characterization over a temperature range from 300 K down to 1.5 K to evaluate the MLC performance of the 1T1R cells under cryogenic conditions. A total of 15 RRAM devices were first initialized (i.e., formed, reset, and set) at 300 K and subsequently cycled 10 times in each of the three LRS levels ($V_G = 1.2, 1.4, 1.6 \text{ V}$). Following this, we began the temperature descent, executing 10 MLC cycles at each temperature point according to the logarithmic grading shown in Figure 2a. Figure 2b–d displays the median I - V response at each temperature point, calculated from the 10 MLC cycles across all devices under test. Figures 3a,b,c, and 3d depict the box plot representation of the main figures of merit extracted from each I - V curve obtained at each temperature step, namely LRS (I_{LRS}) and HRS (I_{HRS}) current values measured at $V_{TE} = 0.2 \text{ V}$, and set and reset switching voltages, respectively.

The measured LRS current values shown in Figure 3a reveal that the separation between resistance states achieved through the MLC approach becomes more pronounced at cryogenic temperatures. Specifically, the conductivity of LRS2 and LRS3 gradually increases as the temperature decreases from 300 K down to 92.4 K, after which it stabilizes. In contrast, a slight

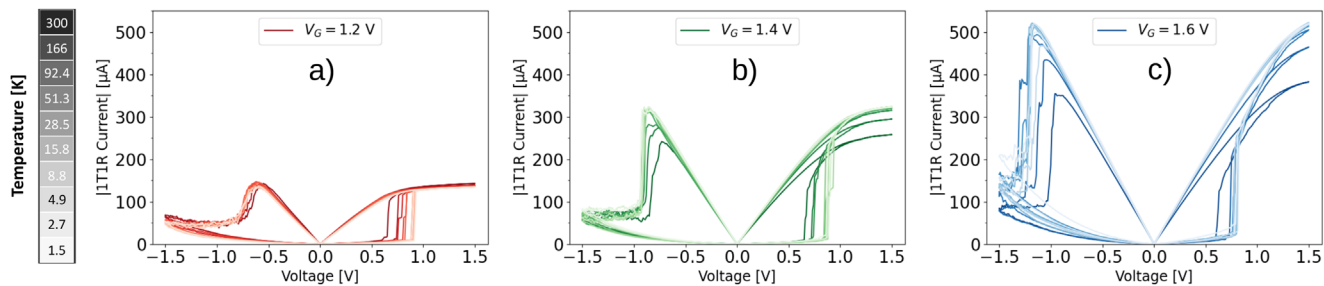


FIGURE 2 | Current–voltage characteristics showing the median response of 15 RRAM devices, each cycled 10 times, for (a) LRS1, (b) LRS2, and (c) LRS3. The color gradient indicates the temperature steps used during characterization.

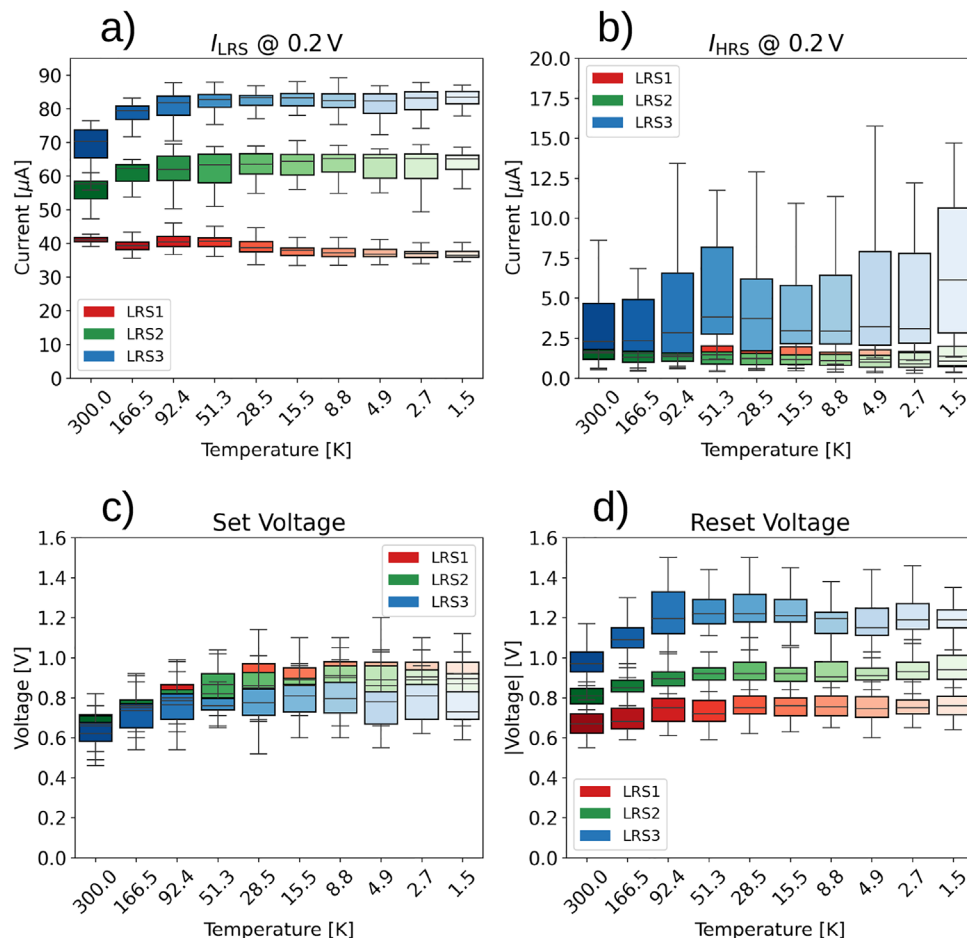


FIGURE 3 | Box plot representations of the extracted figures of merit: (a) LRS current (I_{LRS}), (b) HRS current (I_{HRS}), (c) set, and (d) reset switching voltages.

decrease in the conductivity of LRS1 is observed at temperatures below 51.3 K. As discussed in our previous work [31], the current drift observed at low temperatures may be attributed to variations in the current compliance limit imposed by the nMOS transistor during the set operation in the 1T1R structure, which change with decreasing temperature. In this work, as further discussed below, the cause of the observed LRS current drifts in the 1T1R structures across the temperature range is consistent with previous reports. However, a deeper understanding of the nMOS response within this temperature window is required to quantify its influence on the overall 1T1R behavior. Figure 3b shows the HRS current

obtained after resetting the 1T1R cells from the different LRS states. The I_{HRS} distribution following reset from LRS3 broadens as the temperature decreases, reaching its highest median values at 1.5 K. In contrast, the I_{HRS} distributions for LRS1 and LRS2 remain narrow and stable across all temperature steps.

Concerning the switching voltages, the set voltage (Figure 3c) gradually increases as the temperature decreases, stabilizing at 51.3 K regardless of the target LRS. Similarly, the reset voltage increases with decreasing temperature (Figure 3d), reaching a stable value at 92.4 K. The reset voltage depends on the

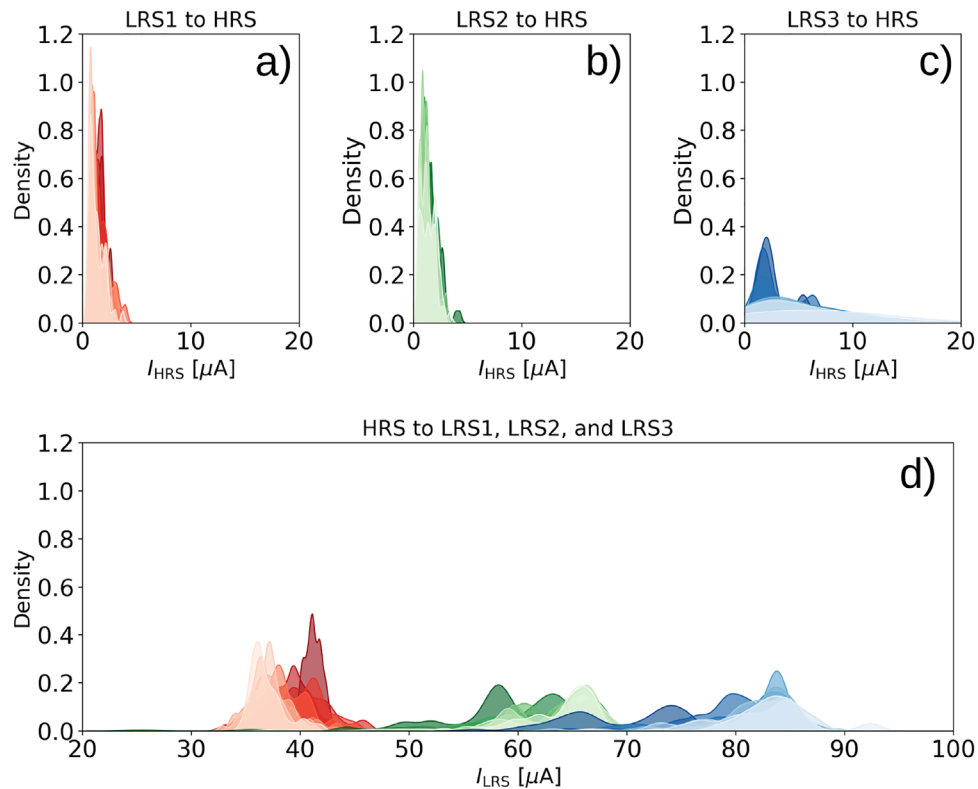


FIGURE 4 | KDE representation of I_{HRS} values after performing the reset operation from (a) LRS1, (b) LRS2, and (c) LRS3, and (d) of I_{LRS} current values after setting the RRAM devices from HRS to the respective LRS, with brighter colors (as in Figure 2) indicating lower test temperatures.

programmed LRS, with higher LRS levels requiring larger reset voltages [34]. This effect is most pronounced during the transition from LRS3 to HRS at cryogenic temperatures. The associated increase in reset voltage likely explains the observed decrease in HRS effective conductance, which can be attributed to an incomplete reset process that leaves a partially formed filament.

To more accurately represent the statistical distribution of I_{LRS} and I_{HRS} shown in Figure 3a,b, Figure 4 presents their kernel density estimation (KDE), providing a smooth empirical probability density function (PDF) based on all sample points and following the color scheme used in Figure 2. Unlike histograms, whose shape is sensitive to the subjective choice of binning parameters, KDE avoids discretization artifacts, and it is particularly advantageous here since not all the distributions necessarily follow a Gaussian profile [35]. Figure 4a–c show the I_{HRS} distributions obtained after resetting the 1T1R devices from LRS1, LRS2, and LRS3, respectively. In particular, Figure 4a,b exhibits narrow HRS current distributions, indicating successful and uniform reset operations from LRS1 and LRS2. In contrast, the broader and smoother I_{HRS} distribution in Figure 4c, corresponding to reset transitions from LRS3 at lower temperatures, suggests the occurrence of incomplete reset, as noted above. This behavior is attributed to the higher LRS3 currents at low temperature, which hinder the complete rupture of the conductive filament during reset. This interpretation is also consistent with previous reports in the literature, where a broader HRS distribution following reset from highly conductive LRS states (e.g., LRS3) has been attributed to increased filament robustness and stochastic partial rupture processes during reset, leading to residual or narrowed conductive paths and enhanced device-to-device variability [36,

37]. Finally, Figure 4d displays the KDE of the LRS states, confirming the shift toward higher current values for LRS2 and LRS3, while LRS1 instead shifts toward lower current values.

In the 1T1R structure, the shift in the nMOS characteristics directly affects the final LRS programmed state, as the set transition depends on the compliance limit imposed by the transistor [31]. The observed LRS current drifts align with the effective conductance variations measured in individual nMOS transistors at cryogenic temperatures, as shown in Figure 5a–c. At $V_G = 1.2$ V, the saturation current decreases with temperature, stabilizing below 92.4 K, which explains the shift of LRS1 toward lower current values. In contrast, at $V_G = 1.4$ and 1.6 V, the saturation current surpasses its nominal value below 92.4 K, accounting for the increased conductivity observed in LRS2 and LRS3 (see Figures 3a and 4d). This phenomenon occurs because the reduction in lattice vibrations (phonons) leads to fewer collisions (scattering), which in turn significantly increases charge carrier mobility. Nevertheless, when the transistor operates at $V_G = 1.2$ V, that is, in the weak inversion regime, this mobility enhancement is masked by the freeze-out of thermally activated carriers, leading to a drop in the compliance current [38].

Additionally, Figure 5d shows the transistor transconductance response, revealing a steeper slope and a higher threshold voltage for temperatures under 92.4 K [39], which lead to the current drop reported in Figure 5a. Below this temperature point the curves overlap. As reported in our previous cryogenic tests [31], this threshold voltage increase at low temperatures reduces the available current budget for low gate biases, thereby preventing the reliable implementation of a fourth LRS or LRS0 at $V_G = 1.0$ V.

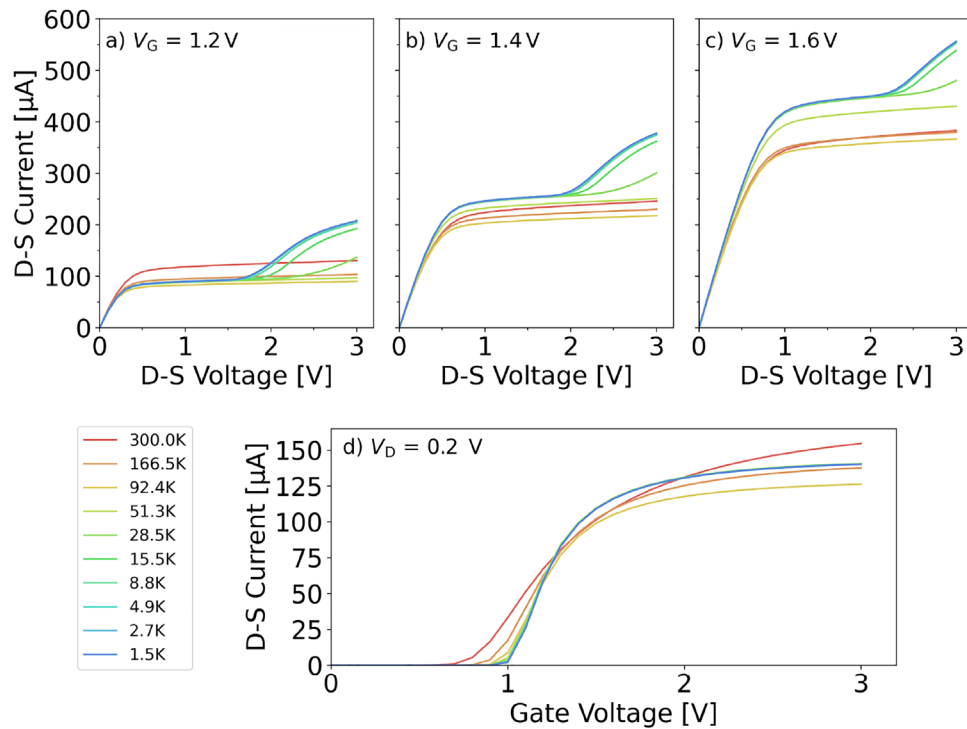


FIGURE 5 | Drain-to-source current measurements of a single nMOS transistor at different cryogenic temperatures, with the gate terminal set to (a) 1.2 V, (b) 1.4 V, and (c) 1.6 V. (d) Corresponding transistor transconductance response. All sub-figures follow the same color grading displayed in the legend.

For all V_G values, the transistor kink effect [40] emerges at 28.5 K and becomes more pronounced at lower temperatures. This effect manifests as a sudden increase in transistor current when the drain-to-source voltage exceeds 1.75, 2.0, and 2.75 V for V_G values of 1.2, 1.4, and 1.6 V, respectively, playing a major role during the forming operation at cryogenic temperatures (which is beyond the scope of this study).

3.2 | Temperature Effects on 1T1R Conduction Mechanisms

The resistance switching study carried out in 1T1R structures reveals that the programmed resistance state is strongly influenced by the nMOS characteristics at the given test temperature, which can mask intrinsic temperature-related effects on the conduction of the MIM structure. To isolate and analyze the dominant conduction mechanisms in a static regime along the cryogenic range under study, we programmed individual 1T1R devices into various LRS and HRS states at room temperature and subsequently cooled the samples down through a series of defined temperature steps. At each step, consecutive readout operations were performed using a sufficiently low readout voltage to avoid disturbing the programmed resistance state of the MIM structure throughout the experiment. This approach enables direct extraction of the MIM resistance (R_{MIM}) as a function of temperature in static regime, allowing us to clearly observe how temperature affects the intrinsic conductivity of the MIM layer independently from the transistor behavior.

At room temperature, sixteen 1T1R structures were programmed, four for each resistance level (HRS, LRS1, LRS2, and LRS3),

by setting V_G to 1.2, 1.4, and 1.6 V during the set operation to achieve the desired LRS states. Consecutive readout operations at different temperatures were executed by setting $V_G = 1.4$ V and sweeping V_{TE} from 0 to 0.3 V in steps of 0.01 V. The temperature was varied from 300 K down to 1.5 K in 20 logarithmically spaced steps.

Figure 6a illustrates the resistance measured over the set of 1T1R structures (R_{1T1R}) at each temperature step for $V_{TE} = 0.2$ V. It can be observed that the resistance of all the LRS levels gradually decreases as the temperature is lowered, reaching minimum values at 98.3 K. A smooth increase of R_{1T1R} is observed at temperatures below this point, stabilizing at 42.6 K, where the measured resistance remains relatively constant down to 1.5 K. Concerning HRS, no major fluctuations are observed along the temperature grading.

Additionally, in purple, we show the equivalent resistance measured from isolated nMOS transistors, facilitating an easy comparison of the temperature-dependent trends between the 1T1R structures and the nMOS devices. Notably, the nMOS resistance at the readout conditions, that is, $V_D = 0.2$ V and $V_G = 1.4$ V, follows a similar trend to that observed in the 1T1R measurements for LRS1, LRS2, and LRS3: a gradual decrease in the resistance with minimum values at 98.3 K, followed by a smooth increase until 42.6 K, after which a stable response is obtained.

Figure 6b represents the nodes within the 1T1R structure taken into account to derive R_{MIM} at the various temperature points based on the current values measured in Figure 6a. Using the isolated nMOS measurements, the transistor resistance at each

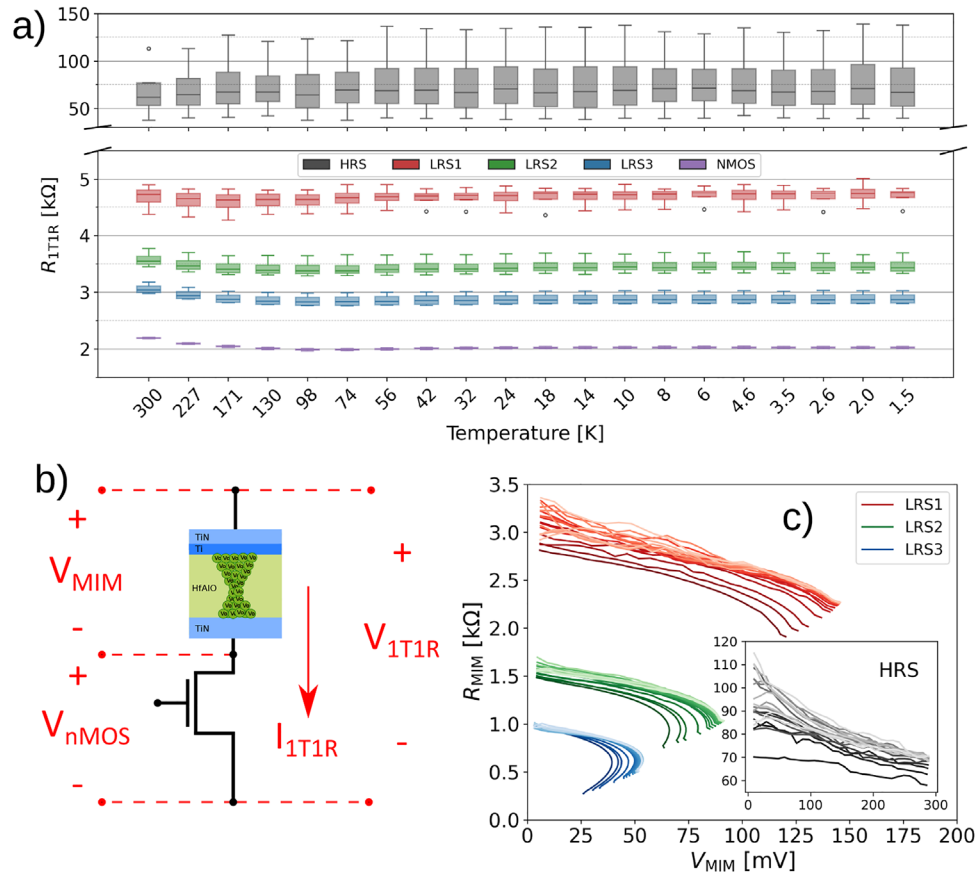


FIGURE 6 | (a) Measured R_{1T1R} values from the readout tests at $V_{1T1R} = 0.2\text{ V}$ across 20 temperature points. (b) Schematic representation of the nodes within the 1T1R structure used to derive the R_{MIM} . (c) Evolution of R_{MIM} for the four resistive states studied at the same temperature points. The brighter the color, the lower the temperature.

measurement point was estimated using the derivative method, that is, computing $R_{DS} = \Delta V_{DS} / \Delta I_{DS}$ from the slope of the I_{DS} – V_{DS} curve, since the applied drain-to-source voltage range lies within the linear region of the transistor operation. Subsequently, we estimated the MIM resistance at each temperature step for each V_{1T1R} point applied to the 1T1R structure during the test. Figure 6c shows the median MIM resistance derived from all 1T1R measurements per level at each temperature point. Although decrease in the 1T1R resistance is observed (Figure 6a), the MIM resistance continues to increase as the temperature is reduced because thermal activation mechanisms that assist carrier transport in the oxide become progressively suppressed, leaving fewer available conduction paths and increasing the effective barrier for charge transport. Ram et al. [28] observe a similar trend in HfO_x -based RRAM devices when cooled down to 4 K.

A deeper understanding of the conduction mechanisms governing the MIM stack may explain the experimental observation that a reduction in temperature results in an increase in MIM resistance. Based on previous studies [31, 41], it can be concluded that under cryogenic conditions the dominant conduction mechanisms in this technology are space-charge-limited conduction (SCLC) and Schottky emission [42, 43]. Figure 7a schematically illustrates the typical I – V characteristics expected for the SCLC mechanism, where three distinct regions can be identified: a linear region ($I \propto V$) when the applied voltage is below the transition voltage (V_{tr}), and two quadratic regions ($I \propto V^2$) for

voltage levels below and above the trap-filled limit voltage (V_{TFL}). This behavior can also be observed in Figure 7b–e, where the quadratic and linear regions (insets) are depicted for HRS, LRS1, LRS2, and LRS3, respectively.

For the experimental curves, several additional observations can be highlighted. First, the linear region of the I – V characteristics does not exhibit a strong dependence on cryogenic temperature, and consequently V_{tr} does not shift with temperature for a given resistance state. However, when comparing across resistance levels, V_{tr} shifts toward lower values for more conductive states. Second, the quadratic region is observed in all levels once $V > V_{tr}$, with temperature influencing the slope of the curve. This effect becomes more pronounced the more conductive the resistance state is. Finally, the V_{TFL} threshold increases as the temperature decreases and, in parallel, shifts towards lower values for more conductive resistance states. In other words, at lower temperatures the transition from the trap-limited regime to the trap-free SCLC regime occurs at higher voltages, implying that the device maintains a higher resistance since the current jump is delayed to larger applied voltages.

A brief analysis of the temperature-dependent transport mechanisms indicates that the experimental data are consistent with a Schottky emission process. This trend was already suggested in our preliminary work [31] and is now further confirmed through an expanded dataset including a larger number of temperature

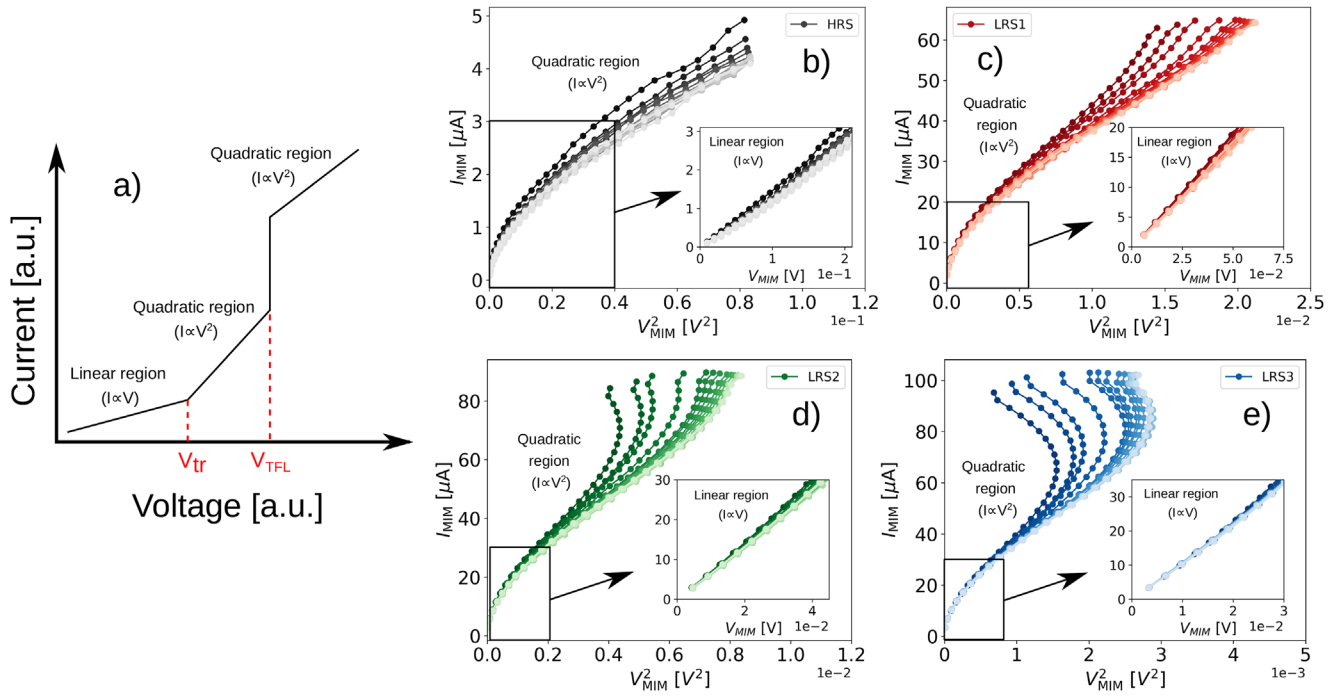


FIGURE 7 | (a) Current–voltage representation of SCLC current based on Chiu, Fu-Chien [42], where V_{tr} is the transition voltage and V_{TFL} is the trap-filled limit voltage. SCLC plots of I_{MIM} – V_{MIM} across 20 temperature points for (b) HRS, (c) LRS1, (d) LRS2, and (e) LRS3. The inset figures represent the linear region of SCLC.

points. Although the detailed fitting plots are not explicitly shown in this manuscript, the underlying analysis consistently reveals this behavior across the full measurement set. In contrast, the Poole–Frenkel model does not show a consistent linear behavior across the investigated bias and temperature range, suggesting that it is not the dominant conduction mechanism under the present experimental conditions.

Therefore, the increase in the 1T1R current observed in the three LRSs, as shown in Figure 3a during programming and in Figure 6b during readout, can be entirely attributed to temperature-induced effects in the nMOS transistor. Specifically, as the temperature is reduced, the nMOS channel resistance decreases, which enhances the current flowing through the 1T1R structures. This effect manifests both during the programming process, where the applied V_{1T1R} controls the filament formation in the $HfAlO_2$ layer, and during the subsequent readout operations, where V_{TE} biases the devices to measure their conduction. Importantly, this behavior indicates that the observed current variations in the LRSs are not due to intrinsic changes in the MIM resistance of the $HfAlO_2$ layer, but rather are dominated by the temperature-dependent response of the series nMOS transistor.

3.3 | Optimizing Electrical Biasing During Resistive Switching in 1T1R Structures at 1.5 K

Given that the operating conditions of the 1T1R structure are primarily governed by the nMOS transistor behavior at cryogenic temperatures, we performed an optimization process aimed at achieving better control of the LRS current and reducing both cycle-to-cycle and device-to-device variability [44]. As reported in the previous section, the pronounced increase in transistor

conductivity at low temperatures is the main cause of elevated LRS effective conductance and the broad distribution of HRS effective conductance, the latter arising from incomplete reset processes. To mitigate these effects, we optimized the gate voltage bias applied during the set operation, specifically when switching the 1T1R cells from HRS to LRS2 and LRS3. The optimized gate voltage levels used in the subsequent tests are summarized in Table 1.

The optimization strategy is based on reducing the gate voltage bias of the nMOS transistor at cryogenic temperatures, in order to maintain comparable D–S current levels at 300 and 1.5 K. This ensures that the compliance conditions during programming remain similar, thereby preventing over-programming and reducing the probability of incomplete reset operations. As shown in Figure 8, the D–S current–voltage characteristics of the transistor overlap when applying $V_G = 1.6$ V at 300 K and $V_G = 1.5$ V at 1.5 K, confirming that equivalent compliance levels are achieved at both temperatures.

For LRS2, the gate bias was also reduced at cryogenic temperature, which leads to a lower D–S current response compared to its counterpart at room temperature. This adjustment is required because the gate bias for LRS1 also decreases with temperature. By doing so, the relative separation between LRS1 and LRS2 is preserved at 1.5 K, ensuring a similar distinction between resistance states as observed at 300 K.

In total, 32 1T1R devices were initialized at 300 K following the procedure described in the *Methods* section, which included device forming and 10 reset/set cycles. Subsequently, the devices were cooled to the minimum operating temperature and switched fifty times per LRS state. Of these, 16 devices were switched using

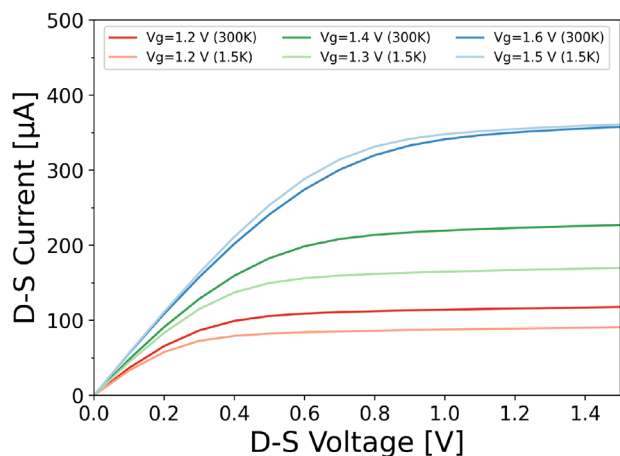


FIGURE 8 | Comparison of the D-S current response of a single nMOS transistor using the initial V_G bias values at 300 K (darker colors) and the optimized V_G bias values at 1.5 K (brighter colors).

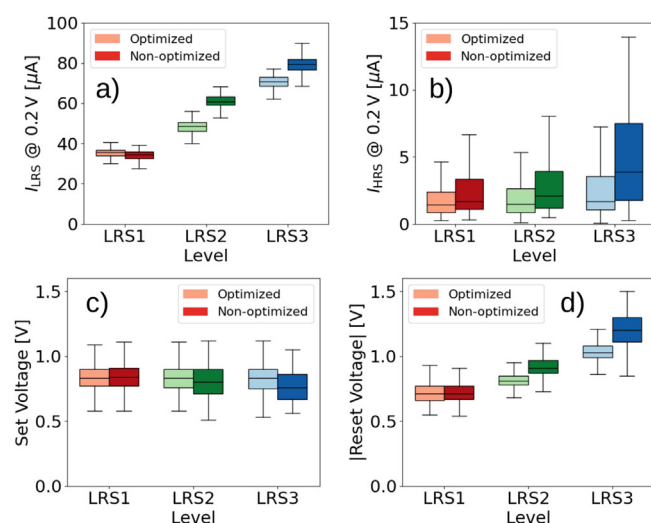


FIGURE 9 | Comparison of key figures of merit obtained from 50 reset/set cycles per LRS level at 1.5 K, using optimized electrical parameters (light colors, 16 devices) and non-optimized parameters (dark colors, 16 devices). (a) LRS and (b) HRS currents measured at $V_{TIR} = 0.2$ V, and (c) set and (d) reset switching voltages.

the optimized V_G biases (1.2, 1.3, and 1.5 V), while the remaining 16 were operated with the non-optimized biases (1.2, 1.4, and 1.6 V). The I - V response of each 1T1R device was measured across the 50 switching cycles at each LRS, and the main figures of merit were extracted to enable a comparative analysis of the switching parameters for both biasing schemes. The main results are presented in Figure 9.

Figure 9a shows the reduction of I_{LRS} for LRS2 and LRS3, resulting from the lower V_G bias applied during the set operation of the 1T1R devices. For LRS2, the median current decreases by 12.3 μ A for the optimized approach, while for LRS3 the reduction is 8.6 μ A. In terms of the interquartile range (IQR), no significant differences are observed between both approaches, with the maximum reduction occurring in LRS3 (approximately 1 μ A). Subsequently, Figure 9b demonstrates the effective reduction of I_{HRS} when resetting the 1T1R devices from LRS2 and, especially,

from LRS3, the latter exhibiting the broadest current distribution in the non-optimized approach. In the case of LRS3, the median I_{HRS} is reduced by 2.2 μ A, together with a reduction in the IQR of 3.24 μ A.

Regarding the set switching voltages (Figure 9c), no significant improvements were observed, as the optimization process primarily involved modifying the compliance current during the set operation, a parameter that does not strongly influence the set voltage.

Finally, the optimization process led to a clear reduction of the reset switching voltage when transitioning the devices from LRS2 and LRS3 to HRS (Figure 9d). Specifically, the median reset voltage is reduced by 100 mV for LRS2 using the optimized approach, while for LRS3 the reduction reaches 200 mV. In terms of variability, the IQR is also reduced, with a decrease of 90 mV for LRS2 and 100 mV for LRS3. Overall, the optimization process reduced I_{LRS} while maintaining clear separation between LRS levels, improved the reset operation by narrowing the I_{HRS} distribution, and lowered the reset voltage required to switch the devices from LRS to HRS, enhancing the stability and reducing the energy consumption of future RRAM-based cryogenic memory modules. In this context, the optimization of V_G plays a central role under DC operation, as it directly controls the current compliance and strongly impacts switching stability and variability at cryogenic temperatures. Additional parameters, such as voltage ramp rates or alternative multilevel programming strategies (e.g., tuning the reset stop voltage), as well as pulsed operation conditions, may provide further improvements, although these aspects fall beyond the scope of the present work and will be addressed in future studies.

4 | Discussion

The resistance switching behavior of 1T1R devices across a wide temperature range, down to cryogenic conditions, reveals that the temperature dependence of the transistor strongly affects the programmed cell state during the resistive switching process. A key observation is the increase of the LRS effective conductance as the temperature decreases. This behavior enhances the separation between resistance states, which could be beneficial for readout fidelity, but it also leads to higher maximum LRS currents. Consequently, the reset operation requires larger voltages to fully disrupt the CF, leading to a broader HRS current distribution after reset from LRS3. This broadening is a signature of incomplete reset, driven by the higher conductivity of LRS2 and LRS3 at low temperatures. Analysis of stand-alone nMOS devices across the same temperature range confirmed that transistor-related effects strongly influence the observed LRS behavior during the resistive switching process, in particular through an increased compliance current limit imposed by the transistor at cryogenic temperatures.

To further analyze the intrinsic contribution of the MIM stack from the transistor effects, we conducted a second set of experiments focused on the static regime (readout without switching) by reading out pre-programmed resistance levels across the temperature range. We observed that the MIM resistance increases as the temperature is reduced. The underlying conduction mechanism analysis shows shifts in both the transition voltage and

the trap-free limited voltage. At cryogenic temperatures, the transition from the trap-limited regime to the trap-free SCLC regime occurs at higher voltages. This implies that the device maintains a higher resistance, since the current jump is delayed until larger voltages are applied. However, at the 1T1R cell level, this effect is masked by the contribution of the transistor to the overall readout operation.

Building on these insights, we adjusted the nMOS gate bias during the set process to regulate the compliance current at cryogenic temperature without compromising the distinction between resistance states. In practice, this adjustment reduced the LRS currents, especially in the most conductive level, while preserving the separation between intermediate states such as LRS1 and LRS2. Additionally, this tuning mitigated the broad HRS distribution by reducing the reset voltage required to transition from LRS2 and LRS3 to HRS, promoting a more uniform reset, enhancing switching efficiency and lowering power demands at low temperatures.

5 | Conclusions

Overall, our results demonstrate the optimal operation of RRAM devices across a wide range of cryogenic temperatures, extending into regimes that have not been previously explored. Although the increase in LRS conductivity is predominantly governed by the temperature-dependent behavior of the series nMOS transistor rather than intrinsic changes in the MIM stack, careful gate bias optimization effectively limits LRS currents, mitigates HRS variability, and reduces resistance switching voltages. This optimization enables reliable cryogenic switching without compromising memory state separation. Importantly, these improvements were achieved in devices fabricated using IHP's standard 130 nm BiCMOS process, highlighting their direct compatibility with conventional fabrication flows and reducing integration costs for future heterogeneous cryogenic in-memory systems. Looking forward, extending gate bias optimization to dynamic pulsed operation and further investigating conduction mechanisms within the MIM stack, together with the integration of peripheral circuitry, will be essential steps toward accurate modeling and practical deployment of 1T1R arrays for cryogenic memory applications. Furthermore, although this study establishes a solid foundation for the development of RRAM-based cryogenic-compatible memories, additional reliability aspects must be thoroughly evaluated at ultra-low temperatures, including resistive state retention, read-disturb phenomena, and resistive switching endurance. In parallel, comprehensive assessments of energy consumption and heat dissipation are required to address key quantum computing constraints, such as proximity to qubits, power dissipation limits, and control latency.

Acknowledgments

This work was financially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Project 551916740.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project Number: 551916740.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

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

References

1. K. Bharti, A. Cervera-Lierta, T. H. Kyaw, et al., "Noisy Intermediate-Scale Quantum Algorithms," *Reviews of Modern Physics* 94, no. 1 (2022): 015004.
2. F. Arute, K. Arya, R. Babbush, et al., "Quantum Supremacy Using a Programmable Superconducting Processor," *Nature* 574, no. 7779 (2019): 505–510.
3. S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, "Hybrid Quantum-Classical Algorithms and Quantum Error Mitigation," *Journal of the Physical Society of Japan* 90, no. 3 (2021): 032001.
4. F. Phillipson, N. Neumann, and R. Wezeman, "Classification of Hybrid Quantum-Classical Computing," in *International Conference on Computational Science* (Springer, 2023), 18–33.
5. B. Patra, R. M. Incandela, J. P. Van Dijk, et al., "Cryo-CMOS Circuits and Systems for Quantum Computing Applications," *IEEE Journal of Solid-State Circuits* 53, no. 1 (2017): 309–321.
6. D. B. Tuckerman, M. C. Hamilton, D. J. Reilly, et al., "Flexible Superconducting Nb Transmission Lines on Thin Film Polyimide for Quantum Computing Applications," *Superconductor Science and Technology* 29, no. 8 (2016): 084007.
7. T. Botzem, M. D. Shulman, S. Foletti, et al., "Tuning Methods for Semiconductor Spin Qubits," *Physical Review Applied* 10, no. 5 (2018): 054026.
8. J. P. Zwolak, J. M. Taylor, R. W. Andrews, et al., "Data Needs and Challenges for Quantum Dot Devices Automation," *npj Quantum Information* 10, no. 1 (2024): 105.
9. H. Moon, D. T. Lennon, J. Kirkpatrick, et al., "Machine Learning Enables Completely Automatic Tuning of a Quantum Device Faster Than Human Experts," *Nature Communications* 11, no. 1 (2020): 4161.
10. L. Vandersypen, H. Bluhm, J. Clarke, et al., "Interfacing Spin Qubits in Quantum Dots and Donors—Hot, Dense, and Coherent," *npj Quantum Information* 3, no. 1 (2017): 34.
11. P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM Review* 41, no. 2 (1999): 303–332.
12. D. Ielmini and H.-S. P. Wong, "In-Memory Computing With Resistive Switching Devices," *Nature Electronics* 1, no. 6 (2018): 333–343.
13. S. Resch, H. Cilasun, and U. R. Karpuzcu, "Cryogenic PIM: Challenges & Opportunities," *IEEE Computer Architecture Letters* 20, no. 1 (2021): 74–77.
14. D. Ielmini and G. Pedretti, "Resistive Switching Random-Access Memory (RRAM): Applications and Requirements for Memory and Computing," *Chemical Reviews* 125, no. 12 (2025): 5584–5625.
15. V. Milo, A. Glukhov, E. Pérez, et al., "Accurate Program/Verify Schemes of Resistive Switching Memory (RRAM) for In-Memory Neural Network Circuits," *IEEE Transactions on Electron Devices* 68, no. 8 (2021): 3832–3837.
16. E. Perez, M. K. Mahadevaiah, E. P.-B. Quesada, and C. Wenger, "Variability and Energy Consumption Tradeoffs in Multilevel Programming of RRAM Arrays," *IEEE Transactions on Electron Devices* 68, no. 6 (2021): 2693–2698.
17. A. Baroni, E. Pérez, K. D. S. Reddy, et al., "Enhancing RRAM Reliability: Exploring the Effects of Al Doping on HfO₂-Based Devices,"

- IEEE Transactions on Device and Materials Reliability* 25, no. 3 (2025): 379–387.
18. A. Baroni, A. Glukhov, E. Pérez, et al., “An Energy-Efficient In-Memory Computing Architecture for Survival Data Analysis Based on Resistive Switching Memories,” *Frontiers in Neuroscience* 16 (2022): 932270.
 19. E. Perez-Bosch Quesada, M. K. Mahadevaiah, T. Rizzi, et al., “Experimental Assessment of Multilevel RRAM-Based Vector-Matrix Multiplication Operations for In-Memory Computing,” *IEEE Transactions on Electron Devices* 70, no. 4 (2023): 2009–2014.
 20. J. D. Teske, S. S. Humpohl, R. Otten, et al., “A Machine Learning Approach for Automated Fine-Tuning of Semiconductor Spin Qubits,” *Applied Physics Letters* 114, no. 13 (2019): 133102.
 21. S. Czischek, V. Yon, M.-A. Genest, et al., “Miniaturizing Neural Networks for Charge State Autotuning in Quantum Dots,” *Machine Learning: Science and Technology* 3, no. 1 (2021): 015001.
 22. V. Yon, B. Galaup, C. Rohrbacher, et al., “Robust Quantum Dots Charge Autotuning Using Neural Network Uncertainty,” *Machine Learning: Science and Technology* 5, no. 4 (2024): 045034.
 23. V. Yon, B. Galaup, C. Rohrbacher, et al., “Experimental Online Quantum Dots Charge Autotuning Using Neural Networks,” *Nano Letters* 25, no. 10 (2025): 3717–3725.
 24. F. Ware, L. Gopalakrishnan, E. Linstadt, et al., “Do Superconducting Processors Really Need Cryogenic Memories? The Case for Cold DRAM,” in *Proceedings of the International Symposium on Memory Systems* (ACM, 2017), 183–188.
 25. S. Alam, M. S. Hossain, S. R. Srinivasa, and A. Aziz, “Cryogenic Memory Technologies,” *Nature Electronics* 6, no. 3 (2023): 185–198.
 26. J. Shang, G. Liu, H. Yang, et al., “Thermally Stable Transparent Resistive Random Access Memory Based on All-Oxide Heterostructures,” *Advanced Functional Materials* 24, no. 15 (2014): 2171–2179.
 27. C. Ahn, S. Kim, T. Gokmen, O. Dial, M. Ritter, and H.-S. P. Wong, “Temperature-Dependent Studies of the Electrical Properties and the Conduction Mechanism of HfO_x -Based RRAM,” in *Proceedings of Technical Program-2014 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)* (IEEE, 2014), 1–2.
 28. M. Ram, D. Lombardo, E. Zaccaria, et al., “Cryogenic Analog 1T-ReRAM with Enhanced Dynamic Range and Suppressed Noise for Cold Neural Networks,” in *2024 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2024), 1–4.
 29. S. Blonkowski and T. Cabout, “Bipolar Resistive Switching from Liquid Helium to Room Temperature,” *Journal of Physics D: Applied Physics* 48, no. 34 (2015): 345101.
 30. R. Fang, W. Chen, L. Gao, W. Yu, and S. Yu, “Low-Temperature Characteristics of HfO_x -Based Resistive Random Access Memory,” *IEEE Electron Device Letters* 36, no. 6 (2015): 567–569.
 31. E. Perez-Bosch Quesada, A. Mistrone, R. Jia, et al., “Forming and Resistive Switching of HfO_2 -Based RRAM Devices at Cryogenic Temperature,” *IEEE Electron Device Letters* 45, no. 12 (2024): 2391–2394.
 32. K. D. S. Reddy, E. Pérez, A. Baroni, et al., “Optimization of Technology Processes for Enhanced CMOS-Integrated 1T-1R RRAM Device Performance,” *European Physical Journal B* 97, no. 11 (2024): 181.
 33. E. Perez, M. K. Mahadevaiah, C. Wenger, C. Zambelli, and P. Olivo, “The Role of the Bottom and Top Interfaces in the 1st Reset Operation in HfO_2 Based RRAM Devices,” in *2018 Joint International EUROSIOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSIOI-ULIS)* (IEEE, 2018), 1–4.
 34. E. Perez-Bosch Quesada, R. Romero-Zalaz, E. Perez, et al., “Toward Reliable Compact Modeling of Multilevel 1T-1R RRAM Devices for Neuromorphic Systems,” *Electronics* 10, no. 6 (2021): 645.
 35. S. Węglarczyk, “Kernel Density Estimation and its Application,” in *ITM Web of Conferences*, Vol. 23 (EDP Sciences, 2018), 00037.
 36. R. Degraeve, A. Fantini, N. Raghavan, et al., “Causes and Consequences of the Stochastic Aspect of Filamentary RRAM,” *Microelectronic Engineering* 147 (2015): 171–175.
 37. S. Yu, X. Guan, and H.-S. P. Wong, “On the Stochastic Nature of Resistive Switching in Metal Oxide RRAM: Physical Modeling, Monte Carlo Simulation, and Experimental Characterization,” in *2011 International Electron Devices Meeting* (IEEE, 2011), 17–3.
 38. M. B. Yelten, “Cryogenic DC Characteristics of Low Threshold Voltage (VTH) n-Channel MOSFETs,” *Balkan Journal of Electrical and Computer Engineering* 7, no. 3 (2019): 362–365.
 39. A. Beckers, F. Jazaeri, A. Ruffino, C. Bruschini, A. Baschiroto, and C. Enz, “Cryogenic Characterization of 28 nm Bulk CMOS Technology for Quantum Computing,” in *2017 47th European Solid-State Device Research Conference (ESSDERC)* (IEEE, 2017), 62–65.
 40. G. Ghibauda and F. Balestra, “Low Temperature Characterization of Silicon CMOS Devices,” *Microelectronics Reliability* 37, no. 9 (1997): 1353–1366.
 41. G. Vinuesa, H. García, E. Pérez, et al., “On the Asymmetry of Resistive Switching Transitions,” *Electronics* 13, no. 13 (2024): 2639.
 42. F.-C. Chiu, “A Review on Conduction Mechanisms in Dielectric Films,” *Advances in Materials Science and Engineering* 2014, no. 1 (2014): 578168.
 43. E. W. Lim and R. Ismail, “Conduction Mechanism of Valence Change Resistive Switching Memory: A Survey,” *Electronics* 4, no. 3 (2015): 586–613.
 44. J. B. Roldán, E. Miranda, D. Maldonado, et al., “Variability in Resistive Memories,” *Advanced Intelligent Systems* 5, no. 6 (2023): 2200338.