Design and Implementation of an AlScN-Based FeMEMS Multiplier for In-Memory Computing Applications

Shubham Jadhav[1],Ved Gund[1], Madhav Ramesh[1], Debdeep Jena[1,2], Amit Lal[1]

[1]*SonicMEMS Laboratory, School of Electrical and Computer Engineering*

Cornell University
Ithaca, NY, USA

[2]*Materials Science and Engineering*

Cornell University
Ithaca, NY, USA

2023 IEEE International Symposium on Applications of Ferroelectrics (ISAF) | 978-1-6654-6323-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISAF53668.2023.10265327

***Abstract*—**

This paper reports on the design, fabrication, and experimental validation of an aluminum scandium nitride (AlScN) based Ferroelectric Micro-Electro-Mechanical Systems (FeMEMS) Multiplier—a core component for multiply-accumulate (MAC) operations in next-generation in-memory computing applications. The FeMEMS multiplier leverages ferroelectric polarization switching in AlScN to change the piezoelectric coefficient (d_{31}), facilitating non-volatile, analog memory storage for weights in a neural network. The piezoelectric parameters of the films are then used to change a capacitive gap for readout. The ferroelectric thin films could be partially polarized and reached a peak remnant polarization of 216μC/cm^2 at a voltage of 100V V_P (5MV/cm). Experimental results on optically measured displacements confirmed the AlScN unimorph multiplier's operation. The maximum resonance mode displacement was linearly dependent on the polarization and input voltages. This work provides foundational insights into utilizing AlScN in in-memory computing, opening new avenues for high-speed, low-power, and high-accuracy computing applications.

Keywords— Ferroelectric Micro-Electro-Mechanical Systems (FeMEMS); Aluminum Scandium Nitride (AlScN); In-Memory Computing; Multiply-Accumulate (MAC) Operations

I. Introduction

High-speed, low-power, and high-accuracy computing applications require embedded non-volatile memory elements within the processors to overcome the challenges of exponentially growing latency and power consumption of traditional von-Neumann architectures, where computation capacity is limited by the time and energy required to access memory. High-speed and low-power computations applications span materials discovery with molecular dynamics analysis, large-area high spatial resolution climate modeling, VLSI circuits with transistor-level granularity, and behavior of complex biological systems with microscopic resolution. These applications necessitate alternate approaches such as in-memory computing, to reduce the data acquisition latency for computation and inference in deep neural networks (DNNs). The fundamental building block at the heart of in-memory computing is the multiply and accumulate (MAC) unit, which performs matrix multiplication and summative operations to emulate brain-like synaptic functionality[1]. Synaptic weights require analog control of the computational weights, which has been previously demonstrated by ferroelectric field-effect transistors (Fe-FET) with multi-level channel conductance modulation[2], [3]. A concern with Fe-FETs and resistive memory is the standby leakage power at aggressively scaled technology nodes, which can offset the speed and accuracy benefits of in-memory computation. Moreover, in the future, harsh environment in-memory computing systems for civilian, aviation, and military applications will require MAC units compatible with III-V nitride-based systems, and established platforms for high-power and high-temperature operation [4].

Capacitively readout based Micro and nanoelectromechanical systems (M/NEMS) offer a competitive advantage to transistor-based approaches for in-memory computing due to zero standby leakage power and potential for 500 aJ per operation (2 PFLOPS/Watt) with aggressively scaled hafnium-zirconium oxide NEMS MAC units[5], [6]. The number of programmable analog weights can be further scaled with multi-level conductance modulation using laterally graded potentials across the ferroelectric[7]. The recent discovery and demonstration of ferroelectricity in scandium-alloyed aluminum nitride (AlScN) have made it viable to envision and realize MEMS MAC units in III-V systems[8]–[10]. AlN-based RF MEMS resonator technology has matured substantially over the last 20+ years, making it possible to leverage that learning towards the design, fabrication, and testing of AlScN-based M/NEMS[11]–[13]. This work reports on released AlScN ferroelectric micromachined ultrasonic transducers (FMUTs) with programmable piezoelectric coefficients for analog control of the resonator displacement and frequency as the building block for MAC units and related resonator devices.

II. Device Design

A. Ferroelectric MEMS Multiplier

Figure 1 shows a schematic of the Ferroelectric Micro-Electro-Mechanical Systems (FeMEMS) multiplier. This device is a unimorph consisting of a structural elastic layer and an active ferroelectric material. The unimorph can be actuated by applying electric fields across the ferroelectric. The multiplicative functionality is achieved by the product of two quantities 1) Polarization switching voltage (V_P), and 2) Unimorph actuation voltage (V_{in}). The accumulation or additive functionality is achieved by summing displacement currents from multiple devices with a common electrode to sample the sum of the capacitance changes.

In the context of neural network inference computations, these weights are determined off-chip in advance, and then

stored on-chip. The device demonstrated in this paper enables storage of the multi-bit weights in the FeMEMS as a tunable piezoelectric coefficient (d_{31}), which can be modulated by applying varying V_P voltages, thereby taking advantage of the property of partial polarization. In the following sections, we will denote the V_P-dependent piezoelectric coefficient as $d_{31}(V_P)$.

Once the weights have been programmed, the input voltage, V_{in}, is fed to the device. The output generated is a resonant unimorph displacement (δ_{max}), a function of the scaled product of the two voltages, V_P and V_{in}.

$$\delta_{\max} = \beta . d_{31}(V_P) . V_{in} \tag{1}$$

This relationship is mediated by a displacement proportionality constant, β, which depends on the properties of materials in the unimorph and the device geometry. However, β remains constant only within a specific range of V_P values near the coercive field (E_C). Beyond this range, it begins to behave nonlinearly as polarization saturates.

A gap-changing capacitor can measure this displacement (Figure 1b). The gap is a function denoted by ($g \mp \delta_{max}$), where 'g' represents the initial gap between parallel plates and 'δ_{max}' denotes a resonant beam displacement as a function of $d_{31}(V_P)$ and V_{in}.

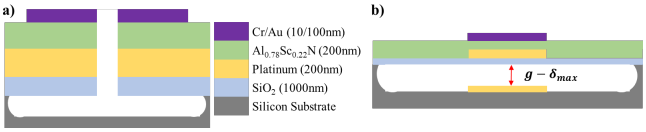

Figure 1: a) Cross-section of the fabricated device featuring a 200nm ferroelectric AlScN layer. b) Conceptual schematic of a parallel plate capacitor sensor designed to detect beam displacement, integral to the functionality of the FeMEMS multiplier.

For small displacement, the output capacitance is represented as:

$$C_{out} = C_0 \left(1 \pm \frac{\delta_{\max}}{g}\right) \tag{2}$$

Here, C_0 is $(\varepsilon_0 . A/g)$

The change in capacitance, δC, can be expressed as:

$$\frac{\delta C}{C_0} = \frac{C_{out} - C_0}{C_0} = \pm \frac{\delta_{\max}}{g} = \pm \left(\frac{\beta}{g}\right) . d_{31}(V_P) . V_{in} \tag{3}$$

B. Device Fabrication and Experimental Setup

A 500 µm Si substrate wafer was used to fabricate the ferroelectric unimorph. A 1 µm thermal SiO_2 layer serves as the elastic layer. Next, a continuous layer of sputtered 200 nm Pt was deposited as the bottom electrode. Ferroelectric AlScN film with a thickness of 200 nm and 22% Sc content was deposited on this Pt electrode. This film stack was achieved through reactive co-sputtering of scandium (Sc) and aluminum (Al) in a nitrogen environment. Next, a lift-off process was employed to evaporate 10nm/100nm Cr/Au metal top electrodes. Reactive ion-etching and ion-milling were then used to etch through the thin-film stack to fabricate etch-holes for releasing the unimorph. This step was followed by a vapor-phase bulk isotropic etching of the underlying silicon substrate using XeF_2 to release the unimorph. This process flow resulted in the successful creation of released cantilevers, clamped-clamped beams, and membranes all on the same wafer.

For electrical characterization of fabricated devices, we utilized the Precision Multiferroic II Ferroelectric Tester from Radiant Technologies Inc. This enabled us to execute various pulse sequences (Figure 2a), detailed in the experimental section, for encoding weight values in the d_{31} coefficient, followed by polarization versus electric field measurements. For vibrational characterization, we employed the Polytec MSA-400 laser doppler vibrometer (LDV). This facilitated the application of an AC actuation voltage (Vin) with variable amplitudes followed by measurement of beam out-of-plane resonant displacement (Figure 2b).

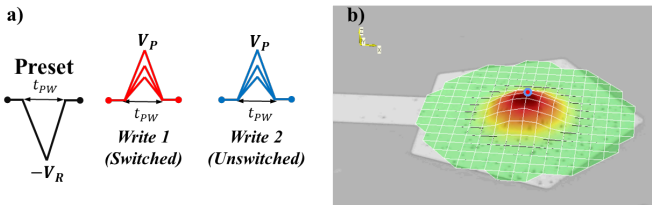

Figure 2: a) Illustration of the switching pulse scheme utilized for weight assignment in the form of piezoelectric coefficients. b) Displacement profile of the fabricated device, as measured by an LDV tool under constant V_p and V_{in} voltages.

III. Experimental Results

A. Electrical Measurements

The switching sequence used involved three triangular pulses, each of 1-ms duration (t_{pw}), specifically termed as reset, switched (write 1), and unswitched (write 2) (Figure 2a). The reset pulse amplitude, denoted by V_R, was consistently maintained at -100V. This value corresponded to an electric field strength of 5MV/cm, sufficient to up-switch all switchable polarization before partial poling experiments. After this, we implemented the first write pulse with amplitude V_P to map the switched polarization half-loop. The second write pulse with same amplitude was subsequently applied to measure the unswitched component of polarization. By subtracting the unswitched polarization from

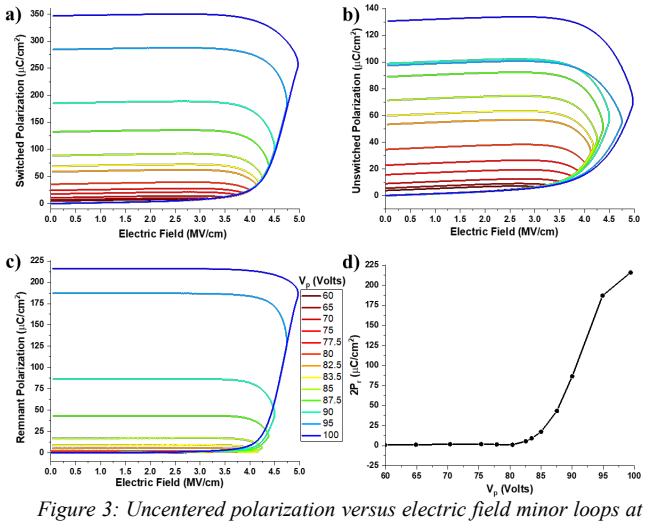

Figure 3: Uncentered polarization versus electric field minor loops at various V_P voltages: a) in the switched state, b) in the unswitched state, and c) corresponding to remnant polarization. d) The remnant polarization value ($2P_r$) is depicted in relation to each applied Vp voltage.

the switched one, we could determine the remnant polarization. We repeated this sequence with progressively increasing values of V_P from 60V to 100V, while maintaining a constant V_R.

Figures 3a-c show the uncentered hysteresis plots for switched (2P*), unswitched (2P^), and remnant hysteresis ($2P_r$=2P*-2P^) polarization minor loops against the electric field at varied V_P ranging from 60V to 100V. Subsequently, Figure 3d illustrates the corresponding remnant polarization ($2P_r$), signifying the polarization value at zero electric field, with respect to the switching voltage V_P. This elucidates the phenomenon of partial polarization with a more explicit visualization. The maximum recorded value of $2P_r$ was 216μC/cm^2 at V_P=100V (5MV/cm), closely aligning with the value on the solid-state MFM capacitor[10]. This result demonstrates that the device's switchable polarization values are retained, even after release.

B. Vibrational Measurements

In order to demonstrate the functionality of the AlScN unimorph as a multiplier, the out-of-plane displacement of the unimorph actuator is measured near its resonance frequency. The quality factor of the resonator in air was 330. This measurement is performed by varying the values of V_P and V_{in}. The displacement is determined using a Polytec LDV. To ensure minimal impact on the unimorph's partial polarization, the V_{in} voltage is three orders of magnitude smaller than V_P to mitigate any polarization switching during actuation.

The transfer characteristics of the FeMEMS multiplier were determined through the following process: The process began by setting up the function generators for both the Ferrotester and LDV. A writing pulse sequence, as described in a previous section, with V_R= -100V and V_P=60 V was applied. Following this, V_P was incrementally increased from 60V to 100V, forming the outer loop of the procedure. After each application of the writing pulse sequence within this outer loop, a trigger signal was sent to the LDV. Next, the LDV function generator was configured to generate a periodic chirp with V_{in}=10mV, spanning the frequency band of 2MHz to 2.15MHz, center at the beam resonance.

The procedure then proceeded to its inner loop. During this phase, V_{in} was increased from 10mV to 100mV in steps of 10mV. At this stage, the magnitude and phase of the unimorph displacement were measured before each increment of V_{in}. Once the inner loop measurements were completed, the function generator was switched back to the Ferrotester, and V_P was increased by 2.5V. This sequence within the outer loop was then repeated until V_P reached 100V.

Figure 4a illustrates a plot of displacement magnitude (on the left y-axis) versus frequency near the beam resonance frequency for V_P =75 V (3.75 MV/cm) (denoted by the black line), and V_P =95 V (4.75 MV/cm) (represented by the red line) at a constant AC actuation (V_{in}=100mV). The displacement magnitudes for both V_P values are nearly identical, but their phase is inverted. Given that all other parameters remained constant and only V_P varied, it can be concluded that the polarization voltage in this range is flipping the ferroelectric domains from being primarily upward to primarily downward. This suggests the signature of polarization reversal at a threshold V_P =83.5 V, which corresponds to the electric field of 4.15 MV/cm. At this value, the orientation of all domains in the films results in the cancellation of the upward and downward piezoelectric moments (associated with positive and negative d_{31} coefficients), yielding a net d_{31} of zero. This causes the unimorph to remain nearly unresponsive to AC actuation voltages, as shown in Figure 4b. This phenomenon is similar to that observed in HZO FeMEMS multipliers [5].

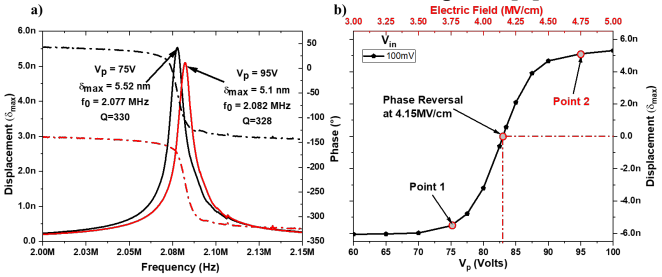

Figure 4: a) Magnitude and phase plot of state 1 (depicted in black) and state 2 (in red) versus frequency, illustrating the characteristic ferroelectric domain reversal phenomenon. b) Displays the measured displacement magnitude at various V_P values for a constant V_{in}=100mV. The two states selected exhibit nearly identical amplitude but with opposite phases.

Figure 5 presents a visual representation showing that the maximum displacement (δ_{max}) is a function of both $d_{31}(V_P)$ and V_{in}. In this figure, it is evident that within a specific voltage range of 75 to 90V, δ_{max} showcases a linear dependence on both V_P and V_{in}. This linear relationship is critical for multiplicative operations as it is represented by the slope of the linear section in the figure. The slope is indicative of constant β , which was defined and discussed in the previous section in relation to (equation (1)) and which is necessary for this multiplicative functionality.

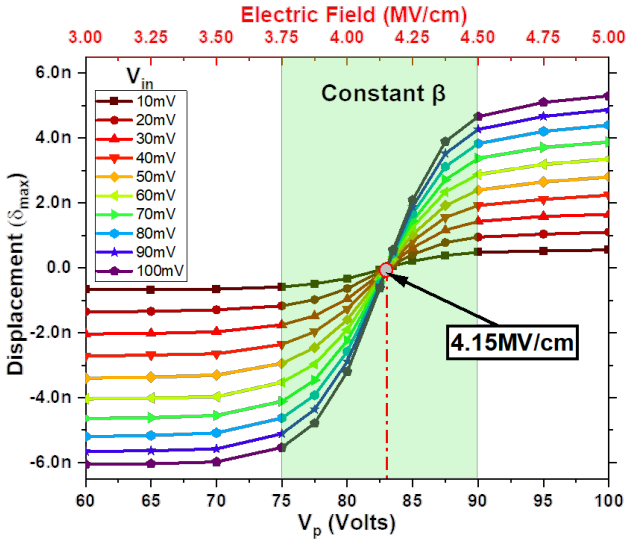

Figure 5: This figure depicts the transfer characteristics of the FeMEMS multiplier device, presenting a plot of the maximum displacement (δ_{max}) versus V_P for various V_{in} voltages.

IV. Conclusions

This paper presents the design and implementation of a ferroelectric AlScN-based programmable NEMS MAC unit, offering a promising approach to in-memory computing. This

approach improves latency and power consumption while maintaining excellent operation using wideband gap III-V platforms in harsh environments. The MAC unit takes advantage of partial polarization switching in AlScN, which allows for a tunable piezoelectric coefficient, effectively storing multi-level computational weights. A multi-step technique was used to measure the ferroelectric/piezoelectric unimorph vibrational characteristics as a function of the applied voltages. The results revealed a direct correlation between maximum displacement, polarization, and input voltages within the 75 to 90V V_P range, validating the device's role as a multiplier.

Moreover, it was observed that at V_P=83.5V (corresponding to 4.15MV/cm), phase reversal occurs, rendering the displacement zero at this point with a complete phase inversion on either side. The FeMEMS multiplier's negligible standby leakage power and compatibility with III-V platforms significantly enhance its efficiency and versatility. This study provides a crucial foundation for future exploration into the innovative use of AlScN, potentially revolutionizing high-speed, low-power, and high-accuracy computing applications. Moreover, it sets the stage for advancements in zero standby power computing and complex solver applications.

V. Acknowledgments

This work was performed in part at the Cornell NanoScale Facility (CNF) , an NNCI member supported by NSF Grant No. NNCI-2025233.

References

[1] C. Mead, "Neuromorphic electronic systems," *Proc. IEEE*, vol. 78, no. 10, pp. 1629–1636, 1990, [Online]. Available: http://ieeexplore.ieee.org/document/58356/.

[2] W. Chung, M. Si, and P. D. Ye, "First Demonstration of Ge Ferroelectric Nanowire FET as Synaptic Device for Online Learning in Neural Network with High Number of Conductance State and G max /G min," *Tech. Dig. - Int. Electron Devices Meet. IEDM*, vol. 2018-December, pp. 15.2.1-15.2.4, Jan. 2019.

[3] M. Jerry *et al.*, "Ferroelectric FET analog synapse for acceleration of deep neural network training," in *2017 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2017, pp. 6.2.1-6.2.4.

[4] D. Jena *et al.*, "The new nitrides: layered, ferroelectric, magnetic, metallic and superconducting nitrides to boost the GaN photonics and electronics eco-system," *Jpn. J. Appl. Phys.*, vol. 58, no. SC, p. SC0801, May 2019.

[5] S. Jadhav, V. Gund, B. Davaji, D. Jena, H. (grace) Xing, and A. Lal, "HZO-based FerroNEMS MAC for in-memory computing," *Appl. Phys. Lett.*, vol. 121, no. 19, p. 193503, Nov. 2022.

[6] S. Jadhav, V. Gund, and A. Lal, "Programmable Ferroelectric HZO NEMS Mechanical Multiplier for in-Memory Computing," *2023 IEEE 36th Int. Conf. Micro Electro Mech. Syst.*, pp. 519–521, Jan. 2023.

[7] V. Gund *et al.*, "Multi-level Analog Programmable Graphene Resistive Memory with Fractional Channel Ferroelectric Switching in Hafnium Zirconium Oxide," 2022.

[8] S. Fichtner, N. Wolff, F. Lofink, L. Kienle, and B. Wagner, "AlScN: A III-V semiconductor based ferroelectric," *J. Appl. Phys.*, vol. 125, no. 11, p. 114103, Mar. 2019, [Online]. Available: http://aip.scitation.org/doi/10.1063/1.5084945.

[9] V. Gund *et al.*, "Temperature-dependent Lowering of Coercive Field in 300 nm Sputtered Ferroelectric Al0.70Sc0.30N," in *IEEE International Symposium on Applications of Ferroelectric (ISAF)*, May 2021, pp. 1–3.

[10] V. Gund *et al.*, "Towards Realizing the Low-Coercive Field Operation of Sputtered Ferroelectric ScxAl1-xN," *21st Int. Conf. Solid-State Sensors, Actuators Microsystems, TRANSDUCERS 2021*, pp. 1064–1067, Jun. 2021.

[11] R. Ruby, "A decade of FBAR success and what is needed for another successful decade," in *Proceedings of the 2011 Symposium on Piezoelectricity, Acoustic Waves and Device Applications, SPAWDA 2011*, 2011, pp. 365–369.

[12] J. Wang, M. Park, S. Mertin, T. Pensala, F. Ayazi, and A. Ansari, "A Film Bulk Acoustic Resonator Based on Ferroelectric Aluminum Scandium Nitride Films," *J. Microelectromechanical Syst.*, vol. 29, no. 5, pp. 741–747, Oct. 2020.

[13] V. Gund, K. Nomoto, H. G. Xing, D. Jena, and A. Lal, "Intrinsically Switchable GHz Ferroelectric ScAlN SAW Resonators," *2022 IEEE Int. Symp. Appl. Ferroelectr. Piezoresponse Force Microsc. Eur. Conf. Appl. Polar Dielectr. ISAF-PFM-ECAPD 2022*, 2022.