HZO-Based Ferro NEMS MAC for In-Memory Computing

CNF Project Number: 1121-03

Principal Investigator(s): Amit Lal

User(s): Shubham Jadhav, Ved Gund

Affiliation(s): School of Electrical and Computer Engineering, Cornell University

Primary Source(s) of Research Funding: Defense Advanced Research Projects Agency (DARPA), Tunable Ferroelectric Nitrides (TUFEN)

Contact: amit.lal@cornell.edu, saj96@cornell.edu

Website: http://www.sonicmems.ece.cornell.edu/

Primary CNF Tools Used: SÜSS MA-6 Contact Aligner, CVC SC-4500 Odd-Hour Evaporator, Zeiss SEM, OEM Endeavor M1, Plasma-Therm Takachi HDP-CVD, Arradiance ALD, AJA Sputter Deposition, Oxford PECVD, Oxford 81/82, Primaxx Vapour HF Etcher, UN770 Etcher, YES EcoClean Asher, Xactix Xenon Difluoride Etcher, AJA Ion Mill, Heidelberg Mask Writer-DWL2000, P7 Profilometer, Zygo Optical Profilometer, Flexus Film Stress Measurement

Abstract:

This work details the fabrication of a hafnium zirconium oxide (HZO)-based ferroelectric NEMS beam as the fundamental building block for very low-energy capacitive readout in-memory computing. The demonstration device consists of a 250 × 30 µm unimorph cantilever of 20 nm thick ferroelectric HZO on 1 µm SiO_2. The displacement of the piezoelectric unimorph was measured by actuating the device with different input voltages V_{in}. The resulting displacement was measured as a function of the ferroelectric programming voltage V_p. The beam displacement scale factor was measured to be 182.9 nm/V for -8V V_p and -90.5 nm/V for 8V V_p, demonstrating that the programming voltage can be used to change the direction of motion of the beam. The resultant output beam displacement from AC actuation is in the range of -15 to 18 nm and is a scaled product of the input voltage and programmed d_{31} (governed by the poling voltage). The multiplication function serves as the fundamental unit of MAC operations with the ferroelectric NEMS beam.

Summary of Research:

Neuromorphic computation is of great interest to computing theory and practical implementations due to the potential for low-power, high efficiency, and small form factor information processing with deep neural networks (DNN) [1,2]. With the ever-increasing number of variables required for neuromorphic computation with high accuracy, there is an urgent need to develop highly energy-efficient device architectures [2-4]. FET-based in-memory computation architectures are susceptible to large read and write energy consumption and high leakage currents in idle mode, particularly with gate dielectric thickness scaling down to < 5 nm [5]. NEMS switches and beams offer an alternate pathway to zero-leakage in-memory compute synaptic functionality, provided that the beam actuation has embedded programmable weights in the form of tunable capacitive or piezoelectric coupling.

While analog in-memory computing has been demonstrated using different architectures that use transistors or memristors, few prior works have used a NEMS-based approach that takes advantage of released beam structure to eliminate energy leakage in an idle state [6].

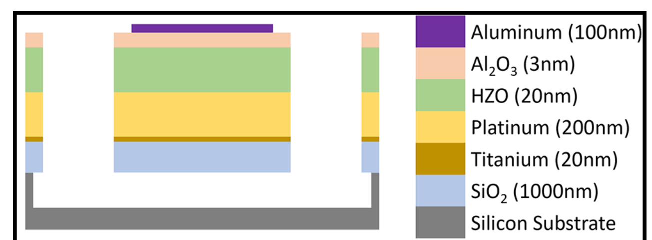

Figure 1: Schematic cross-section view of the ferroelectric beam.

In this work, we present a ferroelectric/piezoelectric beam transducer to enable the multiplication that can be read out capacitively eliminating any DC currents.

Figure 1 shows the cross-section view of the clamped-clamped unimorph ferroelectric beam used to

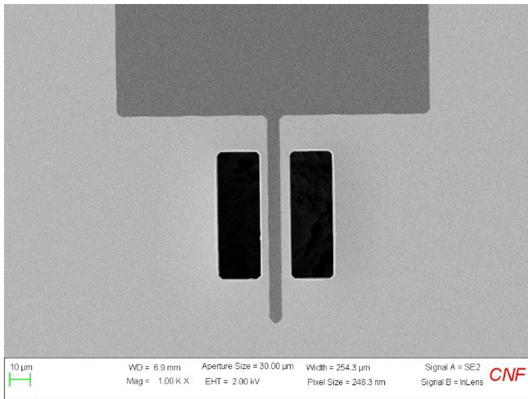

Figure 2: SEM of the beam showing the released structure.

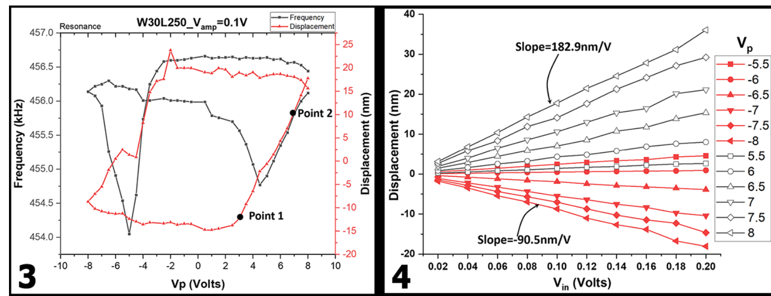

Figure 3, left: Central beam displacement δ_{max} and resonant frequency f_0 vs. poling voltage V_p. Figure 4, right: The displacement vs. V_{in} for different values of the V_p.

demonstrate piezoelectric coefficient programming. 1 µm thick thermal SiO_2 forms the passive elastic layer underneath the 20 nm ferroelectric HZO. The HZO is capped by 3 nm of alumina (Al_2O_3), followed by annealing at 400°C to crystallize the HZO in its ferroelectric orthorhombic phase. Then 200 nm platinum (Pt sputtering) and 100 nm Al (Al evaporation) were deposited to form the bottom and top metal electrode contacts for the HZO respectively. Pt was etched using AJA ion mill and the liftoff technique was used to pattern the Al electrode. The beam was released by isotropic etching of the silicon substrate using Xactix XeF_2.

A Zeiss scanning electron microscope image of the fabricated device is shown in Figure 2.

After release, the beams were observed to be buckled due to residual film stress generated during microfabrication. A 3D optical profilometer (Zygo™ system) was used to measure the beam buckling profile. The maximum displacement for a 250 × 30 µm was 4.98 µm.

Figure 3 shows the central beam displacement δ_{max} and resonant frequency f_0 vs. poling voltage V_p plot in the frequency range of 450 to 460 kHz for the nominal 250 × 30 µm ferroelectric clamped-clamped beam presented here. Peak displacement δ_{max} (red line) is modulated for different values of V_p and traces a hysteresis loop. The net effect of the number of upward and downward pointing dipoles that control the macroscopic polarization in the beam (induced due to V_p) also changes the beam stiffness resulting in resonance frequency modulation (black line), which presents a separate modality of memory storage in the beam. The two dips in f_0 correspond to the positive and negative coercive fields of the HZO film where the net polarization is almost zero.

To show the effect of V_{in} on δ_{max}, we swept V_{in} from 0.02V to 0.20V for different values of the V_p. Figure 4 shows a plot of δ_{max} vs. V_{in} at different values of V_p, which emulates the transfer characteristics of the analog multiplier. For constant poling voltages, if we sweep the inputs, we get a linear increase in displacement value. Similarly, for constant inputs, different poling voltages outputs different displacement values.

In conclusion, we have demonstrated an HZO-based NEMS multiplier weight storage functionality. The device was fabricated and characterized, showing the dependence of beam displacement on the poling voltage of the ferroelectric film. The frequency tunability was also demonstrated that can be further explored to realize unique features such as an electrically tunable filter. Future work will build on this demonstration for device scaling towards high device density.

References:

[1] C. Mead, "Neuromorphic electronic systems," Proc. IEEE, vol. 78, no. 10, pp. 1629-1636, 1990.

[2] A. Keshavarzi, K. Ni, W. Van Den Hoek, S. Datta, and A. Raychowdhury, "FerroElectronics for Edge Intelligence," IEEE Micro, vol. 40, no. 6, pp. 33-48, Nov. 2020.

[3] W. Haensch, T. Gokmen, and R. Puri, "The Next Generation of Deep Learning Hardware: Analog Computing," Proc. IEEE, vol. 107, no. 1, pp. 108-122, 2019.

[4] M. Horowitz, "Computing's Energy Problem (and what we can do about it)," pp. 10-14, 2014.

[5] S. S. Cheema, et al., "Enhanced ferroelectricity in ultrathin films grown directly on silicon," Nature, vol. 580, no. 7804, pp. 478-482, Apr. 2020.

[6] A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh, and E. Eleftheriou, "Memory devices and applications in in-memory computing," Nat. Nanotechnol., vol. 15, no. 7, pp. 529-544, 2020.