Double-Gate W-Doped Amorphous Indium Oxide Transistors for Monolithic 3D Capacitorless Gain Cell eDRAM

H. Ye1*, J. Gomez1*, W. Chakraborty1, S. Spetalnick2, S. Dutta1, K. Ni3, A. Raychowdhury2, and S. Datta1
1University of Notre Dame, Notre Dame, IN 46545, USA; 2Georgia Institute of Technology, Atlanta, GA 30332, USA; 3Rochester Institute of Technology, Rochester, NY 12623 USA
*Equal Contribution; Email: hye1@nd.edu

Abstract—We experimentally demonstrate W-doped amorphous In$_2$O$_3$ double-gate field-effect transistors (DG IWO FET) with 5nm channel thickness and 50nm channel length exhibiting (a) excellent subthreshold slope (SS) of 73mV/dec, (b) record $I_{ON,SAT}$ of 550μA/μm at $V_{GS}$=$V_{TH}$=2V, $V_{DS}$=1V, and (c) high on-off ratio> 1x10$^{12}$. 20nm gate length IWO DG FET was also fabricated to demonstrate scaling potential. We experimentally demonstrate IWO FET based capacitorless 2T gain cell embedded DRAM (eDRAM) ideal for monolithic 3D (M3D) integration exhibiting (a) cell level leakage current of ~1x10$^{-12}$A/μm and ~1x10$^{-14}$ A/μm at 25C and 85C, (b) retention time 1s and 300ms at 25C and 85C respectively with effective storage node capacitance of 1 fF. Array level analysis of IWO capacitorless 2T eDRAM shows that access time < 2ns is achievable with further scaling and moderate outside-the-rail voltages. M3D 2T eDRAM based on IWO FETs offers lower write time than embedded non-volatile random access memory (eNVRAM) and consumes 100x lower stand-by power and 1000x lower refresh power than conventional SRAM and eDRAM, respectively, making it an excellent candidate for fast and embedded memory with unlimited endurance.

I. INTRODUCTION

Recent years have witnessed a rise in domain specific accelerators designed for graphics, deep learning, bioinformatics, image processing and other tasks [1-4]. In addition to the obvious benefit of specialization of logic cores and high level of parallelism, a key source of the acceleration comes from availability of local memories for individual cores and large shared buffer memory on a monolithic chip. Accelerators have created an increasing need of high-density, high-performance, low stand-by power memory alternatives to traditional SRAM. Capacitor-less two-transistor gain cell embedded DRAM is a promising candidate due to its potential for higher density, lower power consumption, higher endurance than eNVRAM, 2-port functionality, non-ratioed circuit operation and their scaling potential to advanced nodes. Gain cells are dynamic memory cells made of two logic transistors, where the second transistor is used not only to increase the in-cell storage capacitance but also to amplify the readout charge via the transconductance gain resulting in a non-destructive read. In this work, we fabricate BEOL compatible W-doped Indium Oxide (In$_2$O$_3$) FETs with ultra-low leakage of 1fA/um to demonstrate a monolithic 3D capacitor-less 2T gain cell eDRAM (Fig. 1). Compared to existing 2T eDRAM [6] in Fig. 2, the fabricated M3D IWO 2T eDRAM shows 1000x higher retention time, comparable access time of 1-10ns and 1000x lower refresh power. Finally, we perform a design-space optimization of the prototype M3D IWO 2T eDRAM to determine its potential for high-speed, high-density and low power embedded memory applications.

II. DEVICE FABRICATION

The low-thermal budget (<250°C) process flow and schematic structure of the fabricated tungsten (W)-doped amorphous In$_2$O$_3$ (IWO) FET with 5nm ALD HfO$_2$ gate oxide is shown in Fig. 3. First, 20nm thick palladium (Pd) metal gate was deposited via e-beam evaporation and liftoff as back-gate followed by 5nm-HfO$_2$ back gate oxide deposition using thermal ALD at 250°C. Next, 5-nm amorphous W-doped (1 wt. %) In$_2$O$_3$ channel was deposited by RF magnetron sputtering in the presence of 0.02 Pa excess O$_2$ at room temperature. 30nm thick Ni was deposited as the source and drain electrode followed by 10-min anneal at 150°C under N$_2$ to improve the contact resistance. Finally, 5nm thick HfO$_2$ top gate oxide was deposited using thermal ALD at 120°C, followed by deposition and patterning of Pd as the top gate.

III. IWO FET CHARACTERIZATION

Fig. 4 (a) shows the measured transfer characteristics of 50nm and 100nm channel length ($L_G$) IWO Double Gate (DG) FETs displaying $SS_{AVG}$ of 73 mV/dec and 68 mV/dec respectively, under $V_{DS}$ 1V. Output characteristics of the DG IWO FET with $L_G$=50nm, shown in Fig. 4 (b), exhibit a high saturation current of 550 μA/μm at $V_{GS}$=$V_{TH}$=2V and $V_{DS}$=1V, thanks to the low contact resistance between IWO film and Ni electrodes. As the off-state leakage current of the IWO FET is extremely low [7], instrumentation with the current detection limit at approximately 10$^{-15}$A cannot directly measure the $I_{OFF}$ leakage. Hence, an ultra-wide device (W=100μm) was used to directly measure the off-state leakage of ~1x10$^{-15}$ A/μm in DG IWO FET, as shown in Fig. 5, resulting in high $I_{ON}/I_{OFF}$ ratio of 1x10$^{12}$. $I_{OFF}$ was found to be limited by the gate-to-drain leakage and can be reduced by increasing the effective oxide thickness (EOT). Figs. 6(a,b) show the cross-sectional TEM image and STEM mode EDX profile respectively, of fabricated IWO FET with $L_G$ = 20 nm. Fig. 7(a,b) show the measured transfer and output characteristics, respectively, of $L_G$=20nm DG IWO FET exhibiting well-tempered electrostatics with $SS_{AVG}$ of 81mV/dec. The performance of $L_G$=50nm DG IWO FET was benchmarked against other BEOL compatible FETs (Fig. 8), where $I_{ON}$ is taken at $V_{DS}$=1V over a 1.8V $V_{GS}$ swing from the reported $I_{MIN}$ point. DG IWO FETs in this work display the highest $I_{ON}$ over a fixed swing, maintaining excellent $I_{ON}/I_{OFF}$ ratio of 1x10$^{12}$ compared to other BEOL compatible FETS. Hence, IWO DG FET is an
excellent candidate for M3D capacitor-less 2T eDRAM application.

IV. VIRTUAL SOURCE MODEL FOR IWO FET

We extend the semi-empirical physics-based Virtual Source (VS) model for IWO transistors [12]. We incorporate gate-voltage ($V_{GS}$) modulated Schottky diodes in series with an intrinsic FET to capture the $V_{GS}$ dependent source-drain contact resistance in IWO FET [7]. According to the VS model, the intrinsic FET drain-to-source current ($I_{DS, FET}$) is computed as the product of the mobile charge density ($Q_{inv}$) and injection velocity ($v_{inv}$) as: $I_{DS, FET} = WQ_{inv}v_{inv}$, where $W$ is the geometric device width. The current through source-drain Schottky junction ($I_{CS, ICD}$ respectively) is empirically modeled as shown in Fig. 8(a). Current continuity equation for the gated Schottky source-drain contacts and the intrinsic IWO FET is solved iteratively to match the experimental transfer (Fig. 8(b)) and output characteristics (Fig. 8(c)). The model parameters are listed in Fig. 8(a). The voltage-dependent capacitances are estimated from the derivative of the terminal charges with respect to the terminal voltages, following channel-charge partitioning as [12]. Fig. 8(d) shows that the VS charge model captures the $V_{GS}$ dependence of gate-to-channel capacitance ($C_{GG}$), gate-to-source capacitance ($C_{GS}$) and gate-to-drain capacitance ($C_{GD}$) at $V_{DS}=0$V. Fig. 8(e) shows excellent agreement between measured and simulated $C_{GG}$, $C_{GS}$, $C_{GD}$ characteristics with varying $V_{DS}$.

V. CHARACTERIZATION OF IWO CAPACITORLESS 2T eDRAM CELL

Fig. 10 (a) shows an optical image of the fabricated capacitor-less IWO 2T DRAM gain cell with four signal lines: read-bitline (RBL), write-bitline (WBL), write-wordline (WWL) and read-wordline (RWL), along with the corresponding circuit schematic shown in Fig. 10(b). The state of the memory cell is given by the charged state in the gate capacitance of the read transistor. Fig. 10 (c) illustrates the cell bias conditions used during the write, read, and hold modes. Write time and retention time are improved by holding WWL at $V_{BOOST}$ above $V_{DD}$ and $V_{HOLD}$, below $V_{SS}$, during write and hold phases respectively. Fig 10 (d) shows how the cell is integrated into an eDRAM array. We characterize the performance of the IWO eDRAM cell, particularly during the hold operation in which the storage node is discharged over time. A characteristic retention time ($\tau_r$) is extracted when the node has discharged 80% of the total charge. We track the voltage of the storage node by continuously measuring the drain current of the read transistor. The storage node voltage ($V_{storage}$) can be obtained from the $I_D-V_G$ of the read transistor. Fig. 11 (a) shows the measured discharge dynamics of the eDRAM cell for different hold voltages. By plotting the characteristic retention time for each hold voltage (Fig. 11 (b)), an optimal hold voltage ($V_{HOLD}$) is extracted at which max retention is achieved before it is reduced by gate leakage. At the optimal $V_{HOLD}$, we study the impact of varying temperature on $\tau_r$ for three different temperatures: 25°C, 50°C and 85°C. The discharge dynamics are shown in Fig. 12 (a). Fig. 12 (b) shows the dependence of $\tau_r$ on temperature. Fig. 13 (a) shows the optical images of various eDRAM cells with different node capacitances that range from 960 fF to 7fF. While the fabricated capacitor-less IWO eDRAM cell has the lowest node capacitance of 7fF, we project it to be ~1fF at a further scaled technology. Figs. 12 (b, c) show the discharge dynamics for different in-cell node capacitances and the retention time as a function of node capacitance, respectively. Extrapolating these results, retention time for 1fF node capacitance at 25°C was found to be ~10s. Projected retention time at 50°C and 85°C are 2s and 0.3sec, respectively.

VI. IWO 2T eDRAM ARRAY

For array level simulation of IWO 2T eDRAM, we use a resistive-capacitive (RC) delay model considering a 128x128 eDRAM array. Interconnect resistance, peripheral (BL/WL driver) output resistance and memory cell transistor channel resistance are treated separately, with a series of cells in a row or column forming a distributed RC line. To compute the read, write, and retention times the large-signal channel resistance of the pertinent transistor, calibrated to the projected transistor (reduced contact resistance and $V_{TH}$ engineered), is estimated at the bias points defined by the midpoints of successive voltage segments. The total time-delay is the sum of the delay for the segments covering the voltage slew range and the limiting interconnect delay time in the case of read and write. The write and retention timing results across a range of $V_{BOOST}$ and $V_{HOLD}$, as shown in Fig. 15, shows <1ns write time with >1s retention are feasible within compact cell dimensions and moderate outside-the-rails voltages. Fig. 16 shows that these timing results describe a class of memory that is much faster (~10x) than emerging nonvolatile memories (eNVM) while requiring significantly less (~100x) standby power than conventional SRAM and eDRAM [18-20]. Fig. 16 (b) shows an array density scaling path through 3D stacking of BEOL memory layers, where cost-effectiveness and density eventually saturate with increasing no. of stacking layers. For these plots, we conceive of a 10% compounding increase in cost while requiring that the equivalent peripheral area at 60% efficiency be present in the scaled CMOS logic layer [21]. Benchmarking against other memory technologies (Fig. 17) the IWO FET based M3D capacitor-less 2T eDRAM emerges as a promising alternative to fill the gap between eNVRAM and traditional Si SRAM and eDRAM.

VII. CONCLUSIONS

We fabricated and characterized BEOL compatible W-doped amorphous In$_2$O$_3$ (IWO) FETs with record Ion (550µA/µm), ultra-low leakage (~1fA/µm) and scaled Lg (20nm). Using IWO FETs we fabricated capacitorless 2T DRAM cell with high retention time and low refresh power, which bridges the gap between eNVRAM and traditional SRAM, eDRAM.


Acknowledgment: This work was supported in part by ASCENT, one of six centers in JUMP, sponsored by DARPA and the Semiconductor Research Corporation (SRC).
**Fig. 1:** Monolithic-3D integration of Tungsten-doped Indium Oxide (IWO) channel FET based 2T-eDRAM allows high memory density, high retention time, ultra-low refresh power and >90% cell-availability.

**Fig. 2:** Comparison of different embedded DRAM options demonstrating IWO 2T eDRAM to exhibit the desired characteristics of high retention time and low access time due to its high \( I_{ON}/I_{OFF} \) ratio. The BEOL compatibility allows high density and the 2T topology allows a non-destructive read.

**Fig. 3:** Schematic device structure and process flow of Dual-Gate IWO FET.

**Fig. 4:** (a) Transfer characteristics ofDual-Gate (DG) IWO FET with \( L_{G}=100 \text{nm} \) and 50nm (b) Output characteristics of DG IWO FET with \( L_{G}=50 \text{nm} \).

**Fig. 5:** Direct measurement of ultra-low \( I_{OFF} \) (~1fA/µm) OFF-state leakage in ultra-wide DG IWO FET showing \( I_{OFF} \) limited by gate-oxide.Current (I).

**Fig. 6:** (a) Cross-sectional TEM image and (b) STEM-EDX elemental map showing fabricated structure of ultra-wide Dual Gate IWO FET with \( L_{G}=20 \text{nm} \).

**Fig. 7:** (a) Transfer and (b) Output characteristics of \( L_{G}=20 \text{nm} \) DG IWO FET with well tempered electrostatics, showing \( I_{OFF} \) limited by contact resistance (~1kΩ-µm).

**Fig. 8:** Benchmarking shows advantage of Dual Gate IWO (Ni S/D) FET with highest \( I_{ON} \) among oxide-semiconductor FETS.

**Fig. 9:** (a) IWO FET VS model equations and parameters, Measured (sym.) and modeled (line) (b) Transfer, (c) Output Characteristics of DG IWO FET, \( C_{GS}, C_{GD}, C_{D} \) versus (d) \( V_{GS} \) at \( V_{DS}=0 \text{V} \) and (e) \( V_{DS} \) at \( V_{GS}=2 \text{V} \).
IWO Capacitor-less 2T-eDRAM Cell Operation

(a) Optical image and (b) corresponding circuit schematic of capacitor-less 2T-eDRAM, (c) timing diagram showing voltage waveforms for Write, Hold, and Read operations, including V_{BOOST} (above VDD) and V_{HOLD} (below VSS) and (d) schematic of the eDRAM array with individual cells.

Experimental Characterization of IWO 2T eDRAM Cell

Fig.11: (a) Discharge dynamics of storage node voltage (V_{storage}) at different hold voltages (V_{hold}) and (b) dependence of retention time (t_r) on different V_{hold}.

Fig.12: (a) Node voltage discharge characteristics at different temperatures and (b) dependence of retention time on operating temperature.

Fig.13: (a) Array of fabricated node capacitances, (b) Node voltage discharge dynamics for different storage-node capacitances (C_{StorageNode}) and (c) dependence of retention time on C_{StorageNode}. Projected retention times for C_{StorageNode}=1F at different operating temperatures show 300ms retention time at 85°C.

IWO 2T eDRAM Array Simulation and Benchmarking

Fig.15: (a) Write time and (b) standby retention across access transistor width scaling and V_{BOOST} and V_{HOLD} respectively. > 1s retention and < 3ns write time are achievable with minimum access device width and moderate outside-the-rails voltages (~2V).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V_{DD} (V)</td>
<td>1.2</td>
<td>1.2</td>
<td>1.05</td>
<td>1.2</td>
<td>1.0</td>
</tr>
<tr>
<td>Density (Mb/mm²)</td>
<td>80</td>
<td>80</td>
<td>17.5</td>
<td>4</td>
<td>180</td>
</tr>
<tr>
<td>BEOL Compatible</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cell Cap. (fF/cell)</td>
<td>1.2</td>
<td>3.5</td>
<td>13.8</td>
<td>*</td>
<td>1</td>
</tr>
<tr>
<td>Retention (ms@85°C)</td>
<td>10³</td>
<td>10⁶</td>
<td>0.3</td>
<td>1</td>
<td>10³</td>
</tr>
<tr>
<td>Access Time (ns)</td>
<td>30</td>
<td>30</td>
<td>5</td>
<td>1.6</td>
<td>3</td>
</tr>
<tr>
<td>Destructive Read</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Fig.16: (a) Comparison of embedded storage technologies’ timing and standby power performance and (b) demonstrating projected density and cost scalability as the number of eDRAM layers increases.

Fig.17: Monolithic-3D capacitorless 2T eDRAM based on W-doped In2O3 FETs emerges as a promising alternative to fill the gap between eNVM and traditional Si SRAM and eDRAM.

*Assuming limited thickness budget ²At similar out of the rails voltages