Demonstration of a Fast, Low-voltage, III-V Semiconductor, Non-volatile Memory

[1]Dominic Lane, [1]Peter Hodgson, [2]Richard Potter and [1]Manus Hayne
[1]Department of Physics, Lancaster University, Lancaster LA1 4YB, UK
[2]Department of Mechanical, Materials and Aerospace Engineering, University of Liverpool, Liverpool L69 3GH, UK

Abstract

ULTRA**RAM**™ is a III-V semiconductor memory technology which exploits resonant tunneling to allow ultra-low-energy memory logic switching (per unit area), whilst retaining non-volatility. Single-cell memories developed on GaAs substrates with a revised design and atomic-layer-deposition Al_2O_3 gate dielectric demonstrate significant improvements compared to prior prototypes. Floating-gate (FG) memories with 20-µm gate length show 0/1 state contrast from 2.5-V program-read-erase-read (P/E) cycles with 500-µs pulse duration, which would scale to sub-ns switching speed at 20-nm node. Non-volatility is confirmed by memory retention tests of 4×10^3 s with both 0 and 1 states completely invariant. Single cells demonstrate promising endurance results, undergoing 10^4 cycles without degradation. P/E cycling and disturbance tests are performed using half-voltages (±1.25 V), validating the high-density random access memory (RAM) architecture proposed previously. Finally, memory logic is retained after an equivalent of $>10^5$ P/E disturbances.

Introduction

A "universal memory" should possess superior performance characteristics compared to current technologies without any major drawbacks. Fundamentally, it should have very robust (non-volatile) logic states, which are nevertheless easily changed at high speed and low energy. These seemingly contradictory requirements have led to the widely-accepted view that such a technology is unfeasible or impossible [1]. ULTRA**RAM**™ is a novel memory concept based on III-V semi-conductors that utilizes the unique band-offsets from the 6.1-Å family (InAs, AlSb and GaSb) [2]. The extraordinarily large (2.1 eV) InAs/AlSb band-offset delivers electron barriers akin to dielectrics, thus forming the basis of a non-volatile memory (NVM). Like flash, ULTRA**RAM**™'s logic states are defined by the storage of electrons within a floating gate (FG). However, they are transported into and out of the FG through a triple-barrier resonant tunneling (TBRT) structure formed from a series of InAs/AlSb heterojunctions specifically engineered for this purpose [3]. This resolves the paradox of universal memory, as resonant tunneling occurs at 2.5 V (approximately 10 times lower than flash), but the InAs/AlSb heterostructure provides a high-energy barrier when voltages are removed.

The ULTRA**RAM**™ memory concept was previously demonstrated on single cell devices [2]. However, they exhibited limited endurance and slow switching speeds, despite operating at low voltages. Although the InAs/AlSb conduction band-offset is 2.1 eV, the valence band offset is just 0.1 eV, allowing significant hole currents to flow through the devices, which degrades performance. Here, the memory structure is revised such that electrons tunnel into (and out of) the FG from the channel side of the cell, allowing us to incorporate an Al_2O_3 gate dielectric with the necessary band-offsets to prevent unwanted leakage from both electrons and holes (Fig. 1).

Device Fabrication

The ULTRA**RAM**™ memory structure (Fig. 1), based on InAs/AlSb/GaSb heterojunctions, was grown on 2-inch n-doped GaAs wafer by molecular beam epitaxy in a Veeco GENxplor system. The lattice mismatch between the substrate and 6.1-Å semiconductors was resolved by use of an interfacial misfit array between the substrate and GaSb buffer layer prior to the growth of the all-important memory layers [2]. Full details of the TBRT structure and how it works are given in [2, 3].

Memory devices were processed using a top-down approach. Standard photolithography techniques were used with inductively-coupled-plasma etching to define the device mesa and access the back gate (BG). *In-situ* reflectance monitoring was used to stop

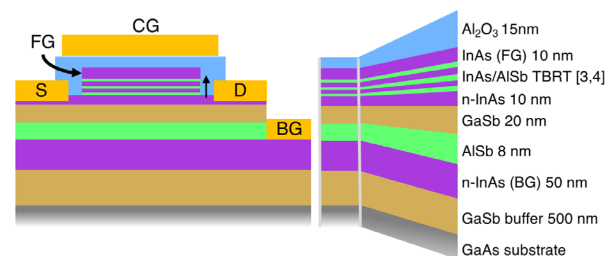

Fig. 1. Schematic depiction of a processed ULTRA**RAM**™ memory cell with corresponding target material layer thicknesses.

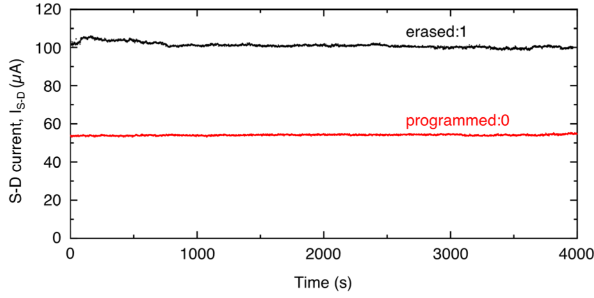

Fig. 2. Retention test for memory in states 1 (erased, black) and 0 (programmed, red). Readout is performed every second with a 0.5 V S-D bias.

the etch in the desired layer. An alternating selective wet etch was used to access the channel, whereby each layer is etched in succession. A citric-based etchant was used to selectively etch InAs over AlSb, and Microposit MF-319 (tetramethylammonium hydroxide) was used to selectively etch AlSb over InAs [4]. Source-drain (S-D) terminals were added by Ti-Au sputtering with a lift-off resist procedure. The Al_2O_3 gate dielectric was then deposited over the sample by thermal atomic layer deposition at 150˚C using 150 cycles of trimethyl-aluminium and water in an Oxford OpAL reactor. A metallization step similar to the S-D formation was used to add the control gate (CG) terminal of the memory cell before the surface was covered by 120 nm of plasma-enhanced-chemical-vapour-deposited SiO_2 for further device isolation. Device terminals were revealed once more via buffered-HF etching of SiO_2 and Al_2O_3 layers, after which final Ti-Au bond pads were added.

Low-voltage P/E

Fig. 2 presents the S-D channel current ($I_{S\text{-}D}$) under 0.5 V S-D bias (*i.e.* a read operation) after both program (P) and erase (E) cycles, depicted in red and black respectively, for a 20-μm gate length ULTRA**RAM**™ cell. Here, the drain and BG terminals are grounded throughout the tests. The n-InAs channel is depleted when electrons are present in the FG, corresponding to a reduction in $I_{S\text{-}D}$ for the programmed state (0). P/E cycling is carried out using -2.5 V program cycles and +2.5 V erase cycles applied to the S terminal with the CG grounded. P/E switching energy per unit area for ULTRA**RAM**™ is 100× lower than dynamic RAM (DRAM) and 1000× lower than flash [3, 5]. Both cycles are conducted with a 500 μs pulse duration, 2000× faster than

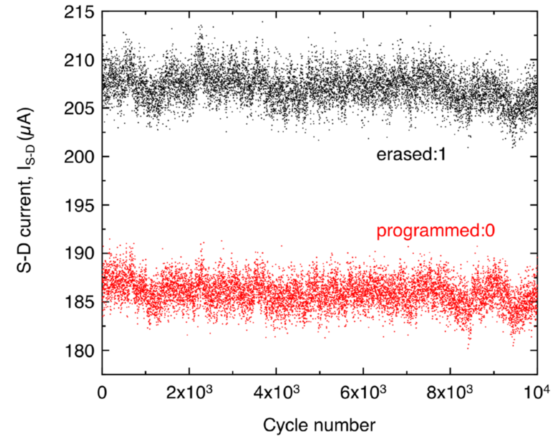

Fig. 3. Continuous endurance cycling for 2.5 V P/E cycles of 500 μs duration. Black dots are $I_{S\text{-}D}$ values after E cycles, and red dots are $I_{S\text{-}D}$ measurements after P cycles. Current is measured during 0.5 V S-D pulses.

previous devices [2]. As a FG-memory, ULTRA**RAM**™ benefits from Dennard's scaling law [6]. Consequently, an ideal scaling from this P/E speed gives sub-ns duration for 20 nm gate lengths, significantly faster than DRAM and flash.

Retention of the logic state was confirmed for 4000 s, where readout took place using a 0.5-V S-D bias every second (Fig. 2). The programmed (0) and erased (1) memory states were observed to remain remarkably stable throughout this test. This evidences ULTRA**RAM**™ as a non-volatile memory, although longer tests at elevated temperature are required to confirm its full retention capabilities. State contrast is sufficient for single cell readout, but is not adequate for large memory-array implementation. However, this is a result of the simplicity of the readout procedure and channel design, not logic-state weakness. Indeed, work is ongoing to implement the high-contrast readout scheme proposed in [3].

Further reliability tests involved repeated P/E cycling of the memory cell. As shown in Fig. 3, it endured 10^4 cycles without degradation. This is a significant improvement compared to previous iterations [2], owing to the elimination of hole leakage currents.

Half-voltage cycling

RAM applications necessitate the fast access of individual memory cells. Previously [3], we introduced a half-voltage architecture for this purpose, where memory cells are arranged in a NOR-type architecture and the necessary voltage for P/E cycles is divided across the word-line (CG) and bit-line (S)

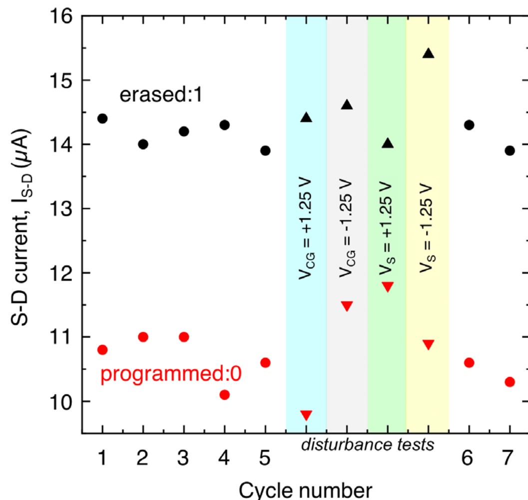

Fig. 4. Half-voltage cycling of ULTRA**RAM**™ using ±1.25 V voltages simultaneously on CG and S terminals (circles) before testing disturbance from individual half-voltages (triangles).

within the architecture. Crucially, these simulations predicted that this half-voltage would not produce a disturbance on surrounding cells which share a word-line or bit-line in common with the target device, which is tested here.

As expected, the simultaneous application of +1.25 V to the CG and -1.25 V to the S of the memory device produces a similar outcome to +2.5 V to the S with the CG at 0 V. The situation is the same for the program cycle, albeit with voltage polarities reversed. Fig. 4 demonstrates this principle, where cycles 1 to 5 are manual P/E cycles, performed with the half-voltage scheme on a 20-µm-gate-length device and show good 0/1 contrast. The important tests here are the "disturbance tests" where an individual half-voltage is applied to the cell. This was done by setting the state (1 or 0) and then applying a constant bias on the CG or S corresponding to the half-voltage disturbance a cell would encounter in this P/E scheme. These are V_{CG} = +1.25 V, V_{CG} = -1.25 V, V_S = +1.25 V and V_S = -1.25 V, and were tested in this order for both 0 and 1 logic states (Fig. 4). Half voltages were sustained on the state for 120 s, equivalent to the disturbance from >10^5 500-µs P/E cycles. The half-voltages are seen to slightly shift the logic of the memory state. Moreover, the direction of the shift is in accordance with loss or gain of electrons from the FG during the disturb test. However, the perturbation is not sufficient to overcome the held memory state during the disturbance test, demonstrating that the rate of electron loss/gain from individual half-voltage cycles is extremely small. This confirms the validity of the P/E scheme for RAM as described in [3], where simulations predicted one electron loss per 4000 10-ns pulses at the 20-nm node. Further manual cycling confirms the switching capability is retained after the disturbance test (Fig. 4, cycles 6 and 7).

Conclusions

We have demonstrated a novel III-V memory technology with ultralow switching energy and non-volatility confirmed by data retention tests of 4000 s at room temperature. A 500-µs switching speed at 20-µm gate length reveals the potential for an ultrafast NVM at small feature sizes. Memory cells remained stable for 10^4 P/E cycles without degradation. Finally, a previously-proposed high-density RAM architecture using half-voltage P/E cycles is demonstrated in principle on single cells, and is found to withstand the equivalent of >10^5 P/E disturbs.

Acknowledgments

This work was supported by the EPSRC, UK, via the 2017-2020 IAA funding allocation to Lancaster University under grant EP/R511560/1, a scholarship under grant EP/N509504/1 and the Future Compound Semiconductor Manufacturing Hub grant EP/P006973/, by the ATTRACT project funded by the EC under Grant Agreement 777222 and by the Joy Welch Educational Charitable Trust.

References

[1] H.S. Wong and S. Salahuddin, "Memory leads the way to better computing," Nat. Nanotechnol., vol. 10, 2015, pp. 191–194.

[2] O. Tizno, A.R.J. Marshall, N. Fernández-Delgado, M. Herrer, S.I. Molina and M. Hayne, "Room-temperature Operation of Low-voltage, Non-volatile, Compound-semiconductor Memory Cells," Scientific Reports volume 9, 2019, 8950.

[3] D. Lane and M. Hayne, "Simulations of Ultralow-Power Nonvolatile Cells for Random-Access Memory," in IEEE Transactions on Electron Devices (Early Access), 2020.

[4] C. Gatzke, S.J. Webb, K. Fobelets and R.A. Stradling, "In-situ monitoring of the selective etching of antimonides in GaSb/AlSb/InAs heterostructures using Raman spectroscopy," Proceedings of the IEEE Twenty-Fourth International Symposium on Compound Semiconductors 337-340 (1997).

[5] K. Prall, "Benchmarking and Metrics for Emerging Memory," 2017 IEEE International Memory Workshop (IMW), Monterey, pp. 1-5 (2017).

[6] R. H. Dennard, F. H. Gaensslen, H. Yu, V. L. Rideout, E. Bassous and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," in IEEE Journal of Solid-State Circuits, vol. 9, no.5, pp. 256-268 (1974). doi: 10.1109/JSSC.1974.1050511.