Semiconductor Science and Technology

PAPER

Field emission from graphene sheets and its application in floating gate memories

IOP Publishing Semiconductor Science and Technology
Semicond. Sci. Technol. 33 (2018) 125003 (7pp) https://doi.org/10.1088/1361-6641/aae626

Field emission from graphene sheets and its application in floating gate memories

R Nashed[1], C Pan[1], K Brenner[2] and A Naeemi[1]

[1] Georgia Institute of Technology, Atlanta, GA, 30332, United States of America
[2] Harper Semi, Atlanta, GA, 30332, United States of America

E-mail: r_nashed@gatech.edu

Received 16 March 2018, revised 28 September 2018
Accepted for publication 4 October 2018
Published 24 October 2018

Abstract

The potential performance of 2D NAND flash with a graphene floating gate (FG) layer is presented. The field enhancement factor for patterned CVD graphene sheets atop a tunneling dielectric is experimentally extracted and used to drive higher-level circuit simulations on 64-bit NAND strings. The average field enhancement factor at a barrier height of 3.1 eV was found to be ~2.85 with a maximum value of 4. Our modest extracted β value explains the contradiction in prior experiments that reported a field enhancement factor of few thousands but only 30%–40% improvement in the write voltage of FG memory devices. Design and operational tradeoffs are benchmarked based on these experimental values and it is shown that 2D NAND programming time and/or programming voltage can be suppressed to 10 ns and 5 V, respectively, based on a 65 nm process node. The onset of read disturbs from the more efficient FG layer are identified and shown to be easily mitigated through error correction code.

Keywords: 2D NAND, non-volatile memory, CVD graphene, field emission

(Some figures may appear in colour only in the online journal)

Introduction

Nonvolatile memory (NVM) is a critical component within a variety of essential devices and systems, including mobile devices, satellites, and the data science boom leveraging enterprise solid-state drives. Such applications place tremendous pressure on NVM to continue meeting breakneck cost and performance demands, even in the face of slowed physical scaling. The dominant NVM technology is based upon charge emission and capture, such as floating gates (FGs) and charge traps. This includes 2D and 3D NAND flash. The performance of NAND flash is largely dictated by this underlying charge emission/capture mechanism [1]. Physically, this mechanism is implemented through a material stack consisting a charge emission layer and a tunneling dielectric. Proper engineering of this material stack can induce orders of magnitude swings in the performance of NAND flash and the devices/systems that it drives. While enterprise applications have spawned architectural advancements in 3D scaling and die stacking, there is similar opportunity to induce vast improvements at the device-level through the integration of nontraditional materials. Low-dimensionality materials, such as carbon nanomaterials, exhibit superior field emission properties and can be seamlessly integrated into mature NVM technologies, such as 2D NAND.

The field-driven tunneling of carriers in flash devices can be treated as a Fowler–Nordheim (FN) tunneling current, given in equation (1) [2]. Here, q is the electron charge, h is Planck's constant, m_{ox} is the effective mass of the electron in the tunneling dielectric, and φ_b is the potential barrier that the electrons must overcome to tunnel through the dielectric

$$J_{FN} = a\beta^2 E^2 \exp\left(-\frac{b}{\beta E}\right),$$
$$a = \frac{q^3}{16\pi^2 h\varphi_b}, \quad b = \frac{4}{3}\frac{\sqrt{2m_{ox}}}{qh}\varphi_b^{3/2}. \tag{1}$$

One important figure of merit for the FN tunneling current is the field enhancement factor, β. The value of β, which functions as a scalar multiple to the applied voltage, is dictated by a local field enhancement that stems from the physical geometry of the FG layer and its work function

[3, 4]. At the device-level, increasing the value of β above unity corresponds to reductions in both the programming voltage and/or programming time. This is especially interesting for low-dimensional carbon nanomaterials, with sharp edges and tunable work functions, and has led to both graphene and carbon nanotubes (CNTs) being exploited for field emission sources [5–7]. The atomically-thin and extremely sharp edges of graphene enhance the electric field which, in turn, enhances the value of β. Furthermore, while graphene struggles with defect-limited mobility in transistor and interconnect applications, such defects can play a positive role in further enhancing the field emission [7, 8].

Reported values of β from graphene sheets (both monolayers and multilayer composite films) can reach as high as 25 000 [9]. Eda *et al* deposited a graphene-polystyrene composite on degenerately doped Si substrate and achieved a β of 1200 [10]. Malesevic *et al* grew vertically aligned few-layer CVD graphene on Ti and Si substrates using microwave plasma enhanced CVD and reached a β of 5000 and 7500 for Si and Ti, respectively [11]. Palnitkar *et al* studied the effect of graphene doping on β [9]. They prepared undoped, boron-doped, and nitrogen-doped graphene using arc discharge technique and deposited it on Si substrate using electrophoretic deposition. The extracted β values were 15740, 11879, and 25849 for undoped, boron-doped, and nitrogen-doped graphene, respectively. Nitrogen-doped graphene showed the highest probably due to the upshift of Fermi energy which reduced the barrier that electrons need to overcome to tunnel to the Si substrate. Wu *et al* [7] also used electrophoretic deposition to fabricate single-layer graphene films and reported a β of 3700. From equation (1), it can be seen that for $\beta = 1000$, J_{FN} would be roughly multiplied by a factor of 10^6. This would cause a drastic decrease in the read and write voltage of the FG memory devices, namely, down to the μV range. However, Hong *et al* reported graphene-based FG memory structure with a write voltage of only 7 V [12]. Also, Hossain *et al* reported a write voltage of 12 V for a FG structure using CNTs as the FG and multilayer graphene as the channel [13]. Although the reported write voltage is lower than the current industry standard (~20 V) [12], it is still much higher than the values predicted based on the β values previously mentioned. Furthermore, the reported β values were based on a structure in which graphene was deposited on a metal or degenerately doped Si substrate which might change the electrical properties of graphene, namely its Fermi energy and hence the barrier height. Therefore, a more accurate determination of β is required.

In this paper, we present a detailed benchmarking of β for CVD graphene sheets, at specific and nominal defect densities and doping levels, in a flash-like material stack. We specifically solve the contradiction in prior experiments that reported a field enhancement factor of few thousands but only 30%–40% improvement in the write voltage of FG memory devices. This experimental data is used to drive higher-level circuit simulations on 64-bit NAND strings to identify performance improvements stemming from graphene integration.

Experimental

The field enhancement factor for CVD graphene sheets was extracted by fabricating flash-like capacitor devices. The fabrication process flow is summarized below in figure 1. First, a 5 nm SiO_2 tunnel oxide is thermally grown on top of a degenerately doped Si substrate (resistivity of ~0.001 Ω cm) by placing the substrate in a quartz furnace and passing O_2 gas for 2 h at 750 °C. Second, a single-layer graphene sheet was grown and transferred onto the tunnel oxide layer using the process described in [14]. The quality of the transferred graphene was verified using Raman spectroscopy and is shown in figure 2. The G band, which occurs around 1587 cm^{-1}, corresponds to optical phonons around Γ-point of the Brillouin zone whereas the 2D peak, occurring around 2680 cm^{-1}, corresponds to the double resonance process of optical phonons around K-point. As can be seen from figure 2, the ratio of the 2D to G peaks is around 2, which is the indicator of monolayer graphene. The peak occurring around 2300 cm^{-1} is due to the SiO_2/Si substrate. 100 × 100 μm^2 graphene devices were patterned using a JEOL JBX-9300FS electron beam lithography (EBL) with a dose of 600 $\mu C\,cm^{-2}$ and a 35 nm layer of 2% Hydrogen Silsesquioxane (HSQ) as resist. After exposure, the sample was developed in MF-319 and the pattern was transferred onto the graphene sheet using a 10 s reactive ion etch process with an O_2 plasma at 25 W. Next, 70 × 70 μm^2 vias were patterned on HSQ using a second EBL step with a dose of 700 $\mu C\,cm^{-2}$ and Poly(methyl methacrylate) (PMMA) as resist. The vias were then formed using a 5 s wet etch in 1:1 buffered oxide etchant (BOE) made from six parts of NH_4, and one part of HF. Here we decide to open vias in the HSQ resist instead of stripping it since HSQ was shown to improve the charge carrier mobility in graphene as well as mechanically pin the graphene sheet on the substrate, leading to a better contact [15]. The remaining PMMA is then stripped by placing the sample in Acetone for 30 min. It is worth mentioning here that ZEP resist should be avoided in this step since it reacts with BOE and damages the fabricated EBL pattern.

Adjacent to each device, a reference device was fabricated where the graphene sheet had been etched away (in the first EBL step) as to provide a direct comparison between β with and without graphene. As such, the reference devices consisted of Pd directly contacting the SiO_2 tunneling dielectric. The graphene devices as well as the blank devices are fabricated by spin-coating 500 nm of ZEP520A resist at a speed of 2000 rpm for 60 s and using an EBL dose of 350 $\mu C\,cm^{-2}$. The sample is developed in Amyl Acetate bath for 2 min followed by a 2 min IPA bath. 80 nm thick Pd pads are deposited using e-beam evaporation followed by a liftoff process in 1165 for 6 h. The ground pads (second terminal of the devices) are then fabricated by etching the exposed SiO_2 using reactive ion etching under CHF_3 gas flow for 30 s. A final EBL step is employed to pattern the ground pads using ZEP520A under the same conditions mentioned earlier. A schematic 3D view of the devices is shown in figure 1(b).

Electrical testing of the devices was immediately carried out (post-metallization) under a vacuum of 1 × 10^{-4} Torr at room temperature using a Lakeshore CPX probe station. A sweep of

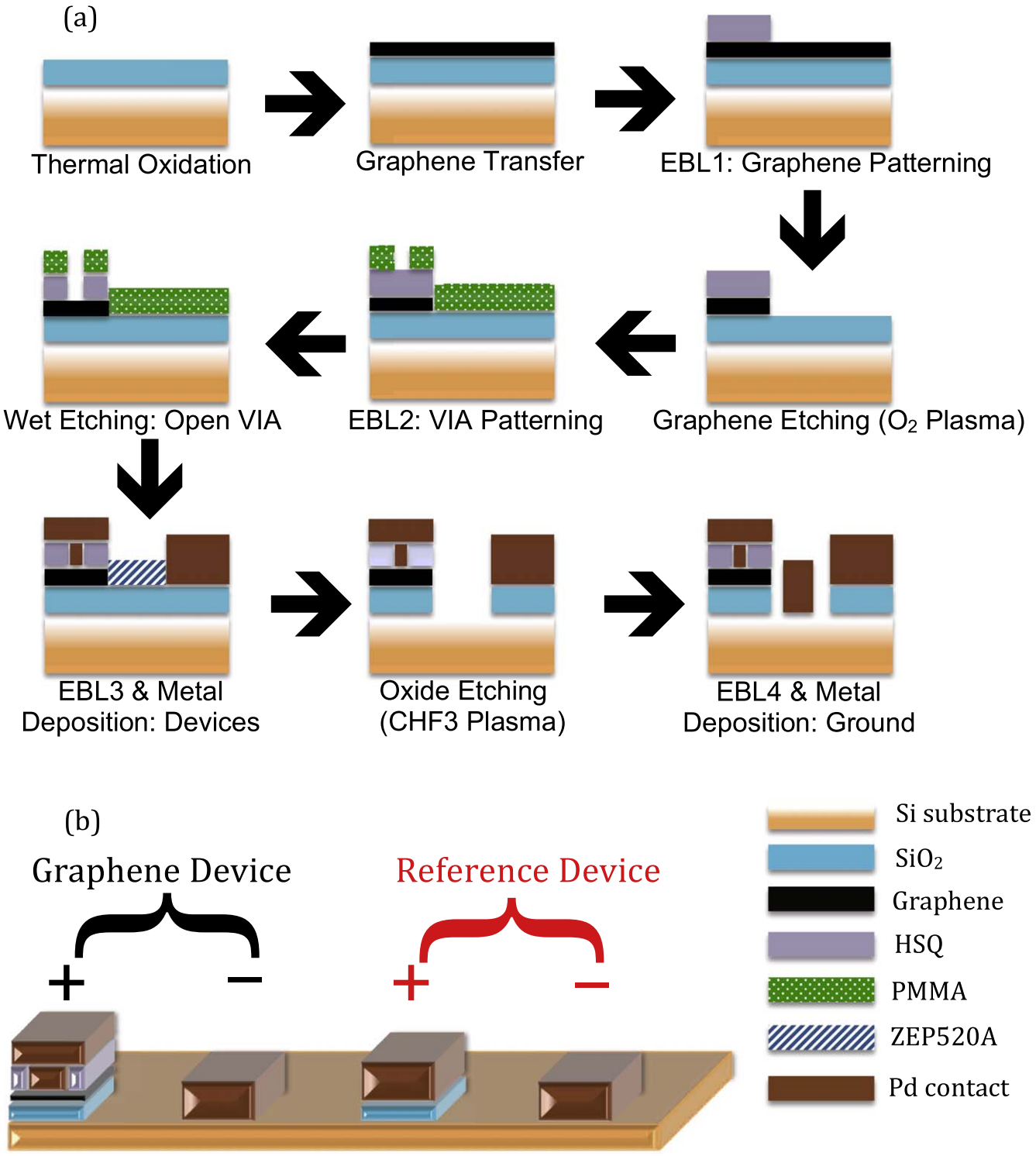

Figure 1. Illustration of fabrication process flow.

DC voltages from 0–12 V was applied to the device using 2 ms pulses to minimize charge trapping and hysteresis [16]. The tunnel current through the device was monitored.

Results and discussion

In this section, we experimentally extract the field enhancement factor from patterned CVD graphene sheets. These values are then used to drive rigorous modeling of CVD graphene as a FG layer in 2D NAND flash devices, including program/erase (P/E) simulations on 64-bit circuits. We identify operational tradeoffs for such graphene-based NAND (G-NAND) that can induce significant reductions in the programming voltage and/or programming time for NAND, which can have a tremendous impact on a variety of NAND-dependent systems. Finally, we detail the onset of potential failure mechanisms

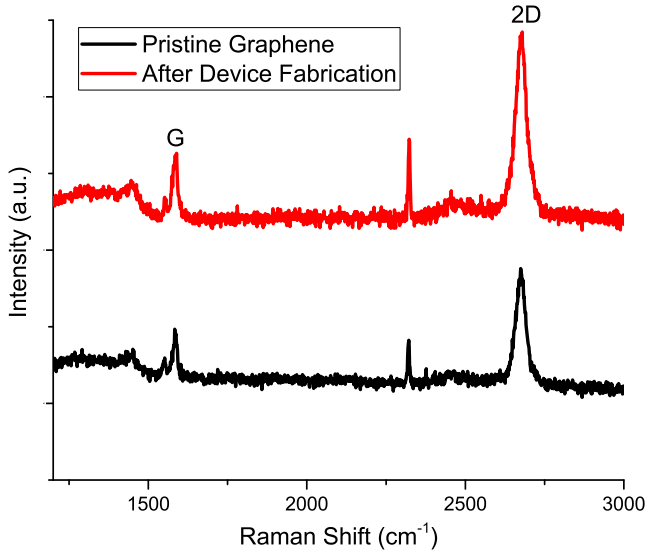

Figure 2. Raman spectra of CVD graphene sheets pristine and prior to metallization.

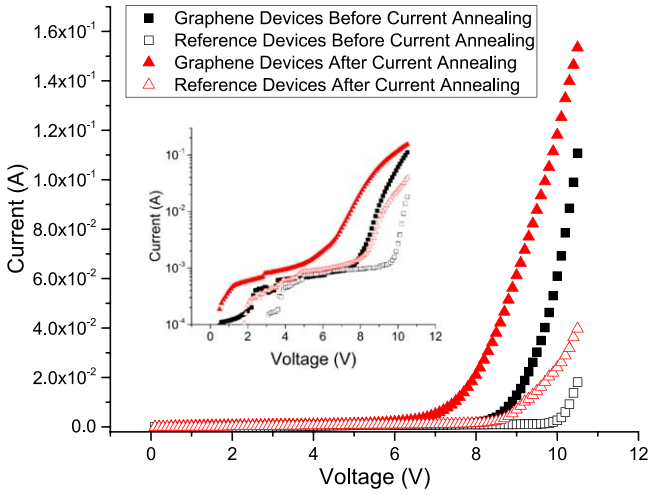

Figure 3. *IV* readings of devices before and after annealing.

associated with high-β NAND, including programming disturbs and how these failures can be successfully suppressed.

Field enhancement factor extraction

The average results from a set of 25 devices (both with and without graphene) are shown below in figure 3. Each of these devices was tested before and after electric current annealing. The current annealing is performed in accordance to the method given in [17]. The graphene devices exhibit a significantly lower turn-on voltage than reference devices, with both devices shifted towards lower voltages after the current anneal. Namely, the turn-on voltage is reduced from ~10 to ~8 V upon incorporation of graphene. Furthermore, the tunneling current for graphene-based devices is significantly higher than the blank devices. This is attributed to the electric field enhancement at the edges of graphene which concentrates the electric field and hence leading to apparent thinning of the tunneling barrier thickness. Since the tunneling current depends exponentially on the barrier thickness, a reduction in the barrier thickness will lead to an exponential increase in the current. Furthermore, current annealing seems to further reduce the turn-on voltage of graphene-based devices to ~6 V. This is partially due to removing the resist residues, H_2O vapor, and O_2 molecules that are introduced with each lithography step, which in turn achieves a robust interface between the metal contact and the underlying layer which, in turn, improves the contact resistance. Current annealing also removes the resist residues from the grain boundaries of the CVD grown polycrystalline graphene layer leading to a better transport of charge carriers and a reduction in graphene's sheet resistance [17]. It is worth mentioning that the reduction in the turn-on voltage does *not* stem from the difference of the work function between graphene and Pd since this difference was measured to be only 0.1 eV [18] whereas the reduction in the turn-on voltage is 2 V. This small difference in the work function was also asserted by Mueller *et al* [19] who measured a potential step of only 0.1 eV at the graphene/Pd interface.

The value of β can be extracted by plotting the *I–V* curve as the $\ln(J/E^2)$ versus $(1/E)$, known as FN plot. Our β is then read from the slope and y-intercept of the FN tunneling plot at a high electric field [2]. Figure 4(A) shows the FN tunneling plot for an average of 25 devices. The distribution of the value of β is shown in figure 4(B) with a median value of 2.85 and a standard deviation of 0.59. This value agrees with the operating voltage of other graphene-based devices [12, 13]. We attribute the overestimation of β in the other (non-device) reports to depositing graphene on a metal instead of a dielectric, which in turn shifts the Fermi energy and lowers the barrier for tunneling. In addition to that, previous reports used air as the tunneling dielectric, which is leakier than SiO_2 due to its low dielectric constant. It is worth mentioning that the reference devices have a β value of 1. This work presents a detailed extraction of the value of β for patterned CVD graphene sheets atop a tunneling dielectric. In addition to extracting the value, the variation of β is also presented for a standard CVD graphene transfer process.

Performance analyses of graphene-based NAND Flash

Regarding the implications of β for NAND flash, the capacitor device is tied to NAND flash operation through the control gate coupling ratio (GCR). The value of the GCR is determined from equation (2) below and dictates the capacitance (or voltage) across the FG as a function of the external voltage applied on the control gate [20]. The value of the GCR should always be maximized to ensure efficient transfer of the program voltage to the FG, which ultimately drives the FN tunnel current

$$\mathrm{GCR} = \frac{\frac{A^*\varepsilon_{\mathrm{CD}}}{\chi_{\mathrm{CD}}}}{C_{\mathrm{FG}} + C_{\mathrm{FS}} + C_{\mathrm{FD}} + C_{\mathrm{FCH}}}. \quad (2)$$

Here, A is the FG area, χ_{CD} is the control dielectric thickness, C_{FG} is the FG capacitance, C_{FS} is the FG to source

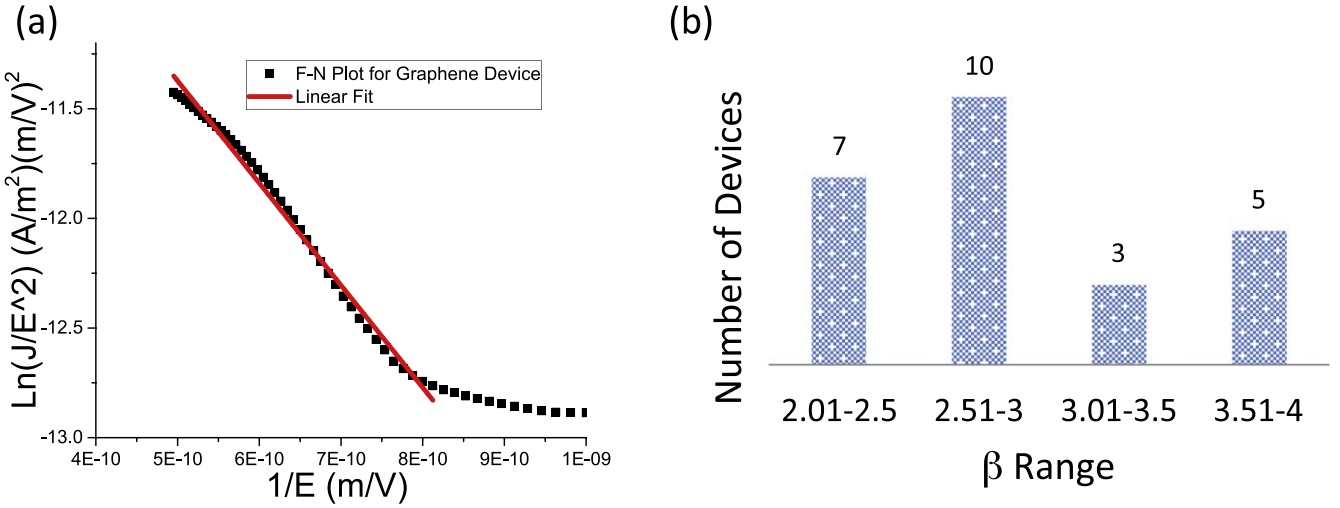

Figure 4. Field emissions of a graphene capacitor device. (a) FN tunneling plot. (b) Histogram showing the distribution of the field enhancement factor of multiple graphene devices.

capacitance, C_{FD} is the FG to drain capacitance, and C_{FCH} is the FG to channel capacitance. Once a model for the FN tunnel current was established, it was implemented as a voltage-controlled current source of a custom SPICE model of our device.

A graphene FG device model was created and HSPICE simulations were performed to identify the potential benefits of G-NAND. The G-NAND device is built by adding a capacitor atop of the gate of a CMOS transistor. Our transistor model is adopted from a predictive technology model (PTM) at the 65 nm processing node [21]. The reason we use a relatively old technology node is to match the process during the fabrication. The capacitance value is properly set to be consistent with the oxide thickness assumed in PTM as well as the targeted thickness of tunneling dielectric of 10 nm and CG dielectric of 20 nm. Two voltage-controlled current sources are added between gate/source and body/source, respectively, to mimic the tunneling current during the write and erase operations. Figure 5 shows a comparison of the ON/OFF resistance ratio for a 64 bit NAND string when the target cell is being read at the ON and OFF states. Our G-NAND is compared against a standard metal- or poly-based 2D NAND at the same 65 nm process node. For a low write voltage of 6 V, even for a long write pulse width of 1 ms, the conventional FG devices are not able to be written properly because of the low tunneling current. With a field enhancement factor of 3, the G-NAND string reaches the target ON/OFF ratio of 2 within 30 μs thanks to the large effective field across the tunneling oxide, leading to a large tunneling current.

Under the same target ON/OFF ratio of 2, the write pulse and write voltage required are shown in figure 6 for three different devices; conventional 2D NAND and G-NAND at $\beta = 2$ and $\beta = 3$. A trend of reduced write pulse width and write voltage is clearly seen as the value of β increase. A clear tradeoff exists between the write pulse width and the write voltage to reach the target ON–OFF ratio. By taking advantage of the large effective field of G-NAND, one can either

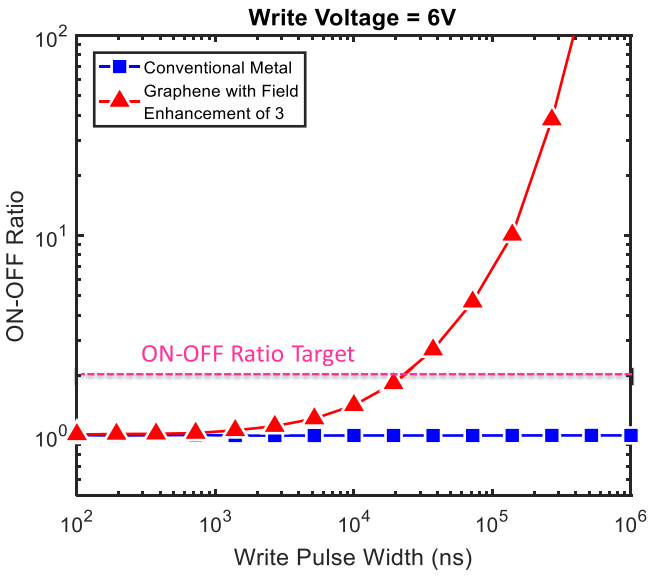

Figure 5. The resistance ratio of the 64-bit NAND string at ON and OFF states versus the write pulse width at a given low write voltage of 6 V.

(1) maintain the speed of the conventional FG devices with reduced write voltage as indicated by the red arrow or (2) keep a relatively large write voltage and significantly improve the write pulse width, hence a fast write operation, as indicated by the blue arrow.

A reliability concern that emerges for G-NAND is the onset of pass disturbs. Due to the more efficient graphene FG layer, the pass voltage applied on nearby cells could induce a large tunneling current that may unintentionally write those pass transistors. To investigate the potential adverse effect of a large β, figure 7 shows the maximum number of allowed read/write cycles before a potential failure for values of $\beta = 2$, 3, 4, and 5. Here, the potential failure is defined as the situation when the ratio of ON and OFF currents flowing through the nearby cells is less than 2 during their read

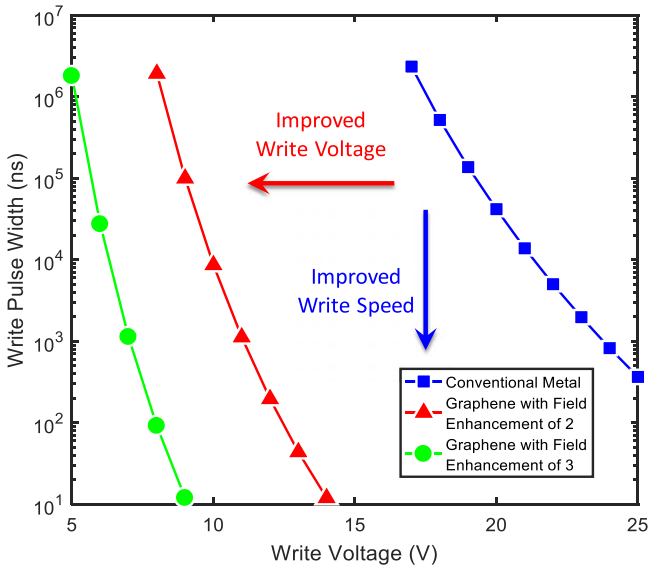

Figure 6. The comparison of write speed versus write voltage between the conventional metal and the graphene floating gate based flash with certain field enhancement factors for a given ON–OFF ratio of 2.

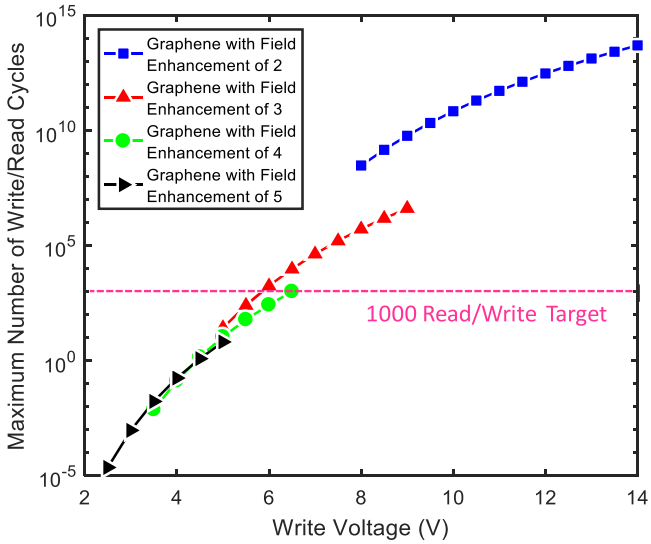

Figure 7. The maximum number of allowed read/write cycles before a potential failure due to the pass disturb versus the write voltage under various field enhancement assumptions of graphene-based flash devices.

operations. One can observe that the maximum cycles increase as the write voltage increases. This is because a large write voltage significantly reduces the write pulse. Since the pass voltage remains the same, a short write pulse leads to fewer charges tunneling through the oxide of pass transistors. Therefore, for a given value of β, more P/E cycles can be achieved without a potential pass disturb. For a P/E cycle target of 1000, one can benefit from a value up to $\beta = 3$ with a write voltage of >6 V. If the field enhancement is too high, due to the large effective field across the tunneling oxide, the nearby cells will be unintentionally programmed within hundreds or even less number of read/write cycles.

Conclusions

The field emission from low-dimensionality materials can play a pivotal role in extending, or developing entirely new, NVM devices to meet the demands of advanced systems. In this work, 2D graphene sheets are experimentally benchmarked in a FG architecture and used to drive circuit-level simulation on a mature NVM technology, 64-bit 2D NAND strings. The field enhancement factor, at a barrier density of 3.1 eV was shown to be 2.85 with a standard device-to-device deviation of 0.59. This modest value solves the contradiction in prior experiments that reported a field enhancement factor of few thousands but only 30%–40% improvement in the write voltage of FG memory devices. The role of this enhancement factor is to expand the operational design window for G-NAND and enable improved programming time and/or programming voltage down to 10 ns and 5 V, respectively, at a 65 nm process node. NAND remains the dominant flavor of NVM and 2D materials, such as graphene, can expand the operational windows of FG and charge trap layers.

Acknowledgments

This work was supported by a Small Business Innovation Research award from the United States Air Force.

ORCID iDs

R Nashed https://orcid.org/0000-0003-1728-0670

References

[1] Kahng D and Sze S M 1967 A floating gate and its application to memory devices *Bell Syst. Tech. J.* **46** 1288
[2] Fowler R H and Nordheim L 1928 Electron emission in intense electric fields *Proc. R. Soc.* A **119** 173–81
[3] Carey J D, Forrest R D and Silva S R P 2001 Origin of electric field enhancement in field emission from amorphous carbon thin films *Appl. Phys. Lett.* **78** 2339–41
[4] Fisher J C and Giaever I 1961 Tunneling through thin insulating layers *J. Appl. Phys.* **32** 172
[5] Deheer W A, Chatelain A and Ugarte D 1995 A carbon nanotube field-emission electron source *Science* **270** 1179–80
[6] Bonard J M *et al* 2002 Field emission of individual carbon nanotubes in the scanning electron microscope *Phys. Rev. Lett.* **89** 197602
[7] Wu Z S *et al* 2009 Field emission of single-layer graphene films prepared by electrophoretic deposition *Adv. Mater.* **21** 1756
[8] Chen I C, Holland S E and Hu C M 1985 Electrical breakdown in thin gate and tunneling oxides *IEEE J. Solid-State Circuits* **20** 333–42
[9] Palnitkar U A *et al* 2010 Remarkably low turn-on field emission in undoped, nitrogen-doped, and boron-doped graphene *Appl. Phys. Lett.* **97** 063102

[10] Eda G *et al* 2008 Field emission from graphene based composite thin films *Appl. Phys. Lett.* **93** 233502
[11] Malesevic A *et al* 2008 Field emission from vertically aligned few-layer graphene *J. Appl. Phys.* **104** 084301
[12] Hong A J *et al* 2011 Graphene flash memory *ACS Nano* **5** 7812–7
[13] Hossain N M and Chowdhury M H 2015 Multilayer graphene nanoribbon and carbon nanotube based floating gate transistor for nonvolatile flash memory *ACM J. Emerg. Technol. Comput. Syst.* **12** 806–9
[14] Li X S *et al* 2011 Large-area graphene single crystals grown by low-pressure chemical vapor deposition of methane on copper *J. Am. Chem. Soc.* **133** 2816–9
[15] Nashed R *et al* 2016 Ultra-high mobility in dielectrically pinned CVD graphene *IEEE J. Electron Devices Soc.* **4** 466–72
[16] Kim W *et al* 2003 Hysteresis caused by water molecules in carbon nanotube field-effect transistors *Nano Lett.* **3** 193–8
[17] Prasad N *et al* 2014 Current induced annealing and electrical characterization of single layer graphene grown by chemical vapor deposition for future interconnects in VLSI circuits *Appl. Phys. Lett.* **105** 113513
[18] Song S M *et al* 2012 Determination of work function of graphene under a metal electrode and its role in contact resistance *Nano Lett.* **12** 3887–92
[19] Mueller T *et al* 2009 Role of contacts in graphene transistors: a scanning photocurrent study *Phys. Rev.* B **79** 245430
[20] Micheloni R 2012 *Inside Solid State Drives (SSDS)* (Dordrecht: Springer)
[21] Sinha S *et al* 2012 Exploring sub-20 nm FinFET design with predictive technology models *2012 49th Acm/Edac/Ieee Design Automation Conf. (Dac)* pp 283–8