# CHAPTER-4

# *FinFET Fin-Trimming During Replacement Metal Gate for an Asymmetric Device Toward STT MRAM Performance Enhancement*

4.1 Introduction........................................................ 84
4.2 Calibration of Simulation Deck and Description of Approach ..........87
4.3 Selective Fin-Trimming and Gate-Stack Engineering for Asymmetric Performance......... 89
4.3.1. Proposed process flow........................................ 90
4.3.2. Physics of proposed asymmetric device ..........................93
4.4 Results and Discussions..........................................96
4.4.1. Asymmetricity ($I_{ON}$ Ratio) comparison ..........................96
4.4.2. Capacitance comparison ........................................ 98
4.5 Conclusion............................................................99

**The part of the work is adopted from-**

**R. Singh**, S. Verma and S. Mittal, "FinFET Fin-Trimming During Replacement Metal Gate for an Asymmetric Device Toward STT MRAM Performance Enhancement," in ***IEEE Transactions on Electron Devices***, vol. 69, no. 12, pp. 6699-6704, Dec. 2022, **doi: 10.1109/TED.2022.3217206**.

***Abstract:***

STT-MRAM-based non-volatile logic circuits require asymmetric write current to enable high efficiency. This requires methods at the device or circuit level, to meet the asymmetric current requirement of STT-MRAM-based circuits. In this work, we propose a novel asymmetric FinFET-based access device, along with a process flow, at 3-nm node. The proposed device uses selective fin-trimming during replacement metal gate (RMG) process, along with a dielectric deposition in the trimmed fin regions. This process does not require a new mask design and can be easily integrated in standard CMOS technology, and thus is very attractive with respect to previously proposed devices. With 3-D TCAD modeling, we show that up to 26% asymmetricity can be achieved with this device architecture, which meets the requirement of STT-MRAM-based circuits. The proposed device offers 30% lower gate capacitance with respect to previously proposed devices.

## 4.1 Introduction

Memory technology is crucial in present-day computers and other mobile/immobile devices. Memory capable of providing high-density, non-volatile storage, low-power operation, speed, and high endurance can be groundbreaking to achieve in-memory, non-volatile, instant ON-OFF computing. Spin transfer torque and spin-orbit torque (SOT)-based memory technology, commonly known as MRAM, has potentially emerged as one of the best alternatives to conventional memories involved in computing. The non-volatile feature of the MTJ devices and compatibility with the CMOS make STT/SOT MRAMs key technology enabler for logic-in-memory (LiM) or compute-in-memory (CiM) applications [9],[164].

MTJ devices present in STT-MRAMs show an asymmetric threshold switching

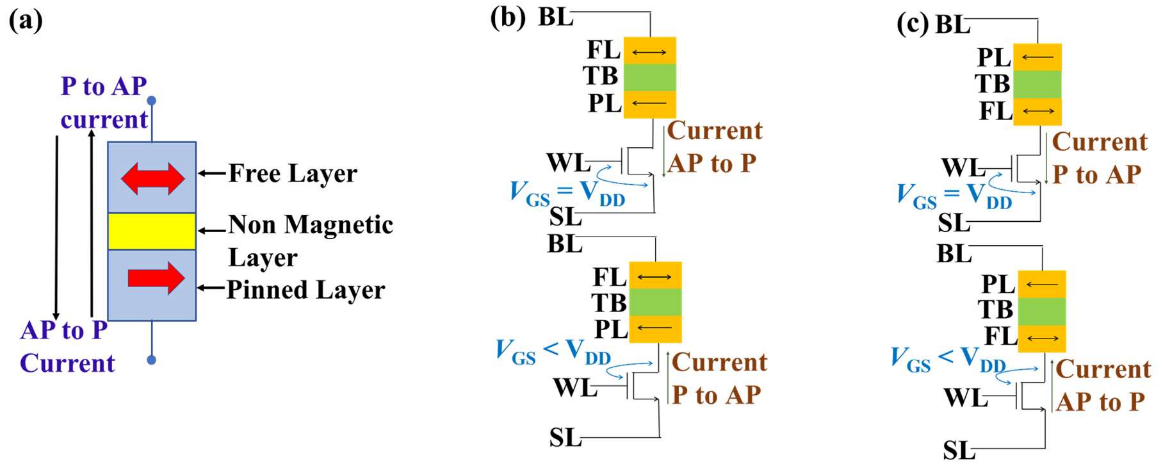


**Figure 4.1** (a) A two-terminal MTJ device, (b) Standard connected STT-MRAM bit-cell, (c) Reverse connected STT-MRAM bit-cell.

current, as shown in **Figure 4.1 (a)** [122]. The anti-parallelizing switching threshold current is around two to three times the parallelizing threshold. On the other hand, source degeneration occurs in both standard and reverse connected (RC) STT-MRAM cells [see **Figure 4.1 (b)** and **(c)**]. In standard connected (SC) STT-MRAM cells, source degeneration occurs with anti-parallelizing current (P to AP switching) during current conduction from source line (SL) to bit line (BL). Similarly, in RC STT-MRAM cells, source degeneration arises with parallelizing current (AP to P switching) during current conduction from SL to BL. Both asymmetric threshold and source degeneration have been shown to increase the write current requirement and power dissipation in STT-MRAM cells [122]. Further, this asymmetry is pronounced in multilevel cells (MLCS) [165] and redundant MLCS [166], which have been proposed to increase the capacity of STT-MRAMs. MLC structure in series or parallel provides an enhanced integration density that reduces the cost per bit. The aforementioned points highlight the increasingly asymmetric requirement of write current in MTJ-based STT-MRAMs and non-volatile logic circuits.

The problem of asymmetric write current requirement has been dealt with by different kinds of device and circuit innovations by researchers in recent years [115],[116],[117],[118],[119],[120]. Circuit techniques for solving the problem of asymmetric write current in STT MRAM cells are explored by Kim *et al.* [120]. These techniques are BL voltage clamping using a pass transistor and 2T-1R dual SL technique. However, it is worth noting that, both these techniques involve an additional transistor. The additional transistor will significantly increase the area of the STT-MRAM bit-cell. Further, at the device level, the concept of asymmetrically doped access devices was utilized in [118] and [119] for writing the current optimization of an STT MRAM bit cell. In AD devices, the source–drain terminal is lightly doped to achieve the asymmetry in drive current during forward and reverse conduction. Further, an asymmetric spacer device is proposed in [122] wherein the asymmetric behavior can be achieved by using different spacer materials. Recently, FinFET devices [167], [168] are being used commercially with faster switching times and higher current density than planar CMOS technology. Some AS FinFET devices have also been reported in [123] and [124]. The architecture of the AS FinFET device is quite challenging from a fabrication point of view. The major fabrication challenge associated with the AS device is to form an asymmetric dielectric spacer along with the symmetrical underlap regions (region below the spacers) on both sides. Both AS and AD devices require an extra mask to create asymmetry in the device, which makes them very costly to integrate into a CMOS process flow.

To address these fabrication challenges, in this article, a novel process flow is proposed which is compatible with the pristine FinFET process flow. Further, the proposed

architecture is better in terms of asymmetricity ratio, drive current, gate capacitance, and robustness as compared to AD and AS FinFET devices for STT-MRAM applications. The low gate capacitance offered by the proposed device further highlights the advancement achieved in this work that can translate into a low write line capacitance in STT-MRAM bit-cell. The WL capacitance is quite significant in achieving high integration density from a memory perspective.

The rest of the article is organized into four sections. Section 4.2 discusses the simulation approach and device architecture. The proposed fabrication methodology and device physics for the asymmetrically trimmed FinFET device is presented in section 4.3. Section 4.4 presents important simulation results from this study. Finally, a comparison between proposed and existing asymmetric devices for STT MRAM applications is kept in section 4.4. We conclude this article in section 4.5.

## 4.2 Calibration of Simulation Deck and Description of Approach

To enable this study, firstly a baseline 3-D FinFET n-MOS TCAD deck was set up at a 3-nm node. The 3–D structure was generated using process emulation in Sentaurus TCAD [169]. The nominal parameters used for the device structure creation are shown in **Table 4.1**. The structure thus generated is shown in **Figure 4.2 (a)**. On this structure, electrical simulations were performed. The validation of our TCAD simulations is done by comparing them with the existing simulation and experimental results. In aggressively scaled FinFETs, quantum confinement of electrons in channel-fin [170],[171],[172] should be captured accurately. To accurately account for quantum confinement, the density-gradient quantum-confinement model is well calibrated against the experiments results in [137], wherein the change in

**Table 4.1** TCAD simulation deck parameters [159],[173].

| Parameter | Value |
|---|---|
| Gate Pitch (nm) | 42 |
| Fin Pitch (nm) | 26 |
| Nominal Fin Width (nm) | 7 |
| Fin Height (nm) | 50 |
| Poly CD (nm) | 18 |
| SD Doping ($cm^{-3}$) | $1e^{20}$ |
| Channel Doping ($cm^{-3}$) | $1e^{16}$ |
| Interface Oxide ($SiO_2$) (nm) | 0.75 |
| Gate Oxide ($HfO_2$) (nm) | 1 |
| Spacer width (nm) | 6 |
| Spacer height (nm) | 30 |
| Gaussian doping from SD to channel (nm/decade) | 2 |
| $V_{DD}$ (V) | 0.7 |
| Gate Work-function (eV) | 4.52 |

saturation threshold voltage is captured, with the variation of fin-width, for fin-width as low as 2 nm [see **Figure 4.2 (b)**]. Further, the device's electrostatics and transport properties were calibrated against Intel's 10-nm experimental results from [159]. The following effects are included in the TCAD Sentaurus deck. First, to capture the doping dependence of mobility, the Masetti model was included. At such small channel lengths, the effect of velocity saturation and overshoot is required to be captured correctly. Toward that, an extended Canali model was used to capture high-field saturation [130]. To capture the interface roughness scattering, the Lombardi model was included. To correctly capture generation-recombination, shockley–read–hall recombination (SRH), trap assisted tunneling (TAT), and band-to-band tunneling (BTBT) models were included. Further, the doping dependence of the recombination lifetime of SRH is captured using the Scharfetter relation. For TAT, the Hurkx model was used. To capture the band-gap narrowing effect, the Old Slotboom model was

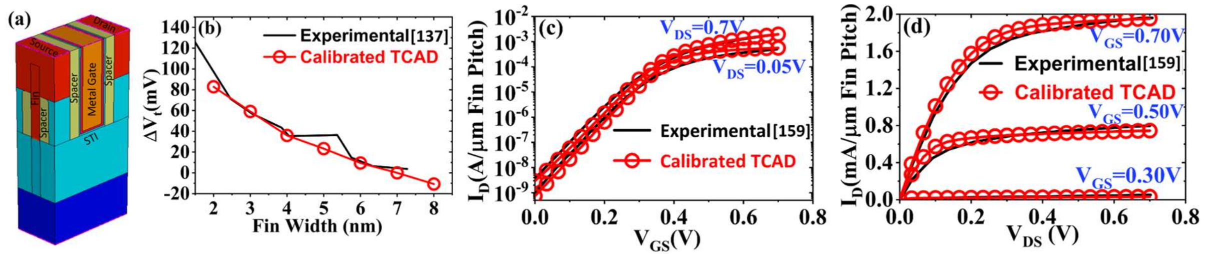


**Figure 4.2** (a) 3-D NMOS FinFET structure, generated using process emulation, (b) Impact of global fin trimming on NMOS and comparison with reported results from [137], (c) $I_D$-$V_G$ characteristics and comparison with experimental data reported in [159], (d) $I_D$-$V_D$ characteristics and comparison with experimental data reported in [159].

used. **Figure 4.2 (c)** shows a comparison of $I_D$–$V_G$ characteristics, whereas **Figure 4.2 (d)** shows $I_D$–$V_D$ characteristics along with the experimental data from [159]. TCAD deck used in this study thus captures the quantum-confinement, transport, and electrostatics accurately. Furthermore, a similar process flow for fabrication had also been shown earlier, the detailed description of the fin-trim process and experimental feasibility has been discussed in detail in [174].

The simulated saturation ($V_{DD}$ = 0.7 V) and linear ($V_{DD}$ = 50 mV) $I_D$–$V_G$ curves are shown in **Figure 4.2 (c)** for the nominal baseline FinFET. The electrical properties of this device are measured as $I_{OFF}$ of 11.16 nA/µm of fin-pitch, and the $I_{ON}$ is 2.22 mA/µm fin-pitch. The device shows a sub-threshold slope of 71.39 mV/decade and a DIBL of 30 mV/V.

## 4.3 Selective Fin-Trimming and Gate-Stack Engineering for Asymmetric Performance

Selective fin-trimming was proposed in [174] as a way to define fin-thickness during the replacement-metal gate process step. In this, after the etching of a dummy metal gate and dummy gate oxide, a silicon fin is etched to define the final fin thickness. As only the fin region is etched and the fin region under the spacer remains protected, the thickness of the

source-drain extension region is more than the fin region. This leads to a 10% – 15% improvement in $I_{ON}$ (at the same $I_{OFF}$) and a 5 KΩ/fin reduction in source-drain resistance [174]. In [174], this process flow is also experimentally demonstrated.

### 4.3.1. Proposed process flow

In this work, we propose an improvement over the process proposed in [174], and it is shown in **Figure 4.3**. The steps leading to asymmetry are highlighted in **Figure 4.3**. In step 1, a shifted polycritical dimension (CD) mask is used to introduce asymmetry. Additionally, instead of a positive photoresist, a negative photoresist is used to partially etch Si in this step. With a shifted poly mask, only the portion uncovered by the mask will be etched, thus this will create an asymmetry in the structure. After the partial etching of the Si fin, in step 2, a dielectric is deposited in the trimmed region of the fin to create further asymmetry in the device *I–V* characteristics.

The usage of this shifted poly-mask is depicted in **Figure 4.4**. In **Figure 4.4 (a)**, the layout of a FinFET-based CMOS inverter is shown, which has no asymmetry. Layers showing fins for NMOS and PMOS, gate (and dummy gate as well), along with the source-drain contact regions, are highlighted. In **Figure 4.4 (b)**, a representative layout as per [122] is shown to introduce asymmetry by introducing AS. As shown, a new mask design is required to enable this fabrication, as one spacer is required to be covered, while another one is open. Similarly, to fabricate an AD device, an additional mask is required (not shown here) because

doping for the source-drain will happen in two separate steps, requiring at least one more mask. In **Figure 4.4 (c)**, we show that the existing poly-CD mask layer [from **Figure 4.4 (a)**], with a lateral shift, can be used to introduce this asymmetry in the device. This laterally

shifted mask layer, along with a negative photoresist, can enable an asymmetric structure as shown in **Figure 4.4**.

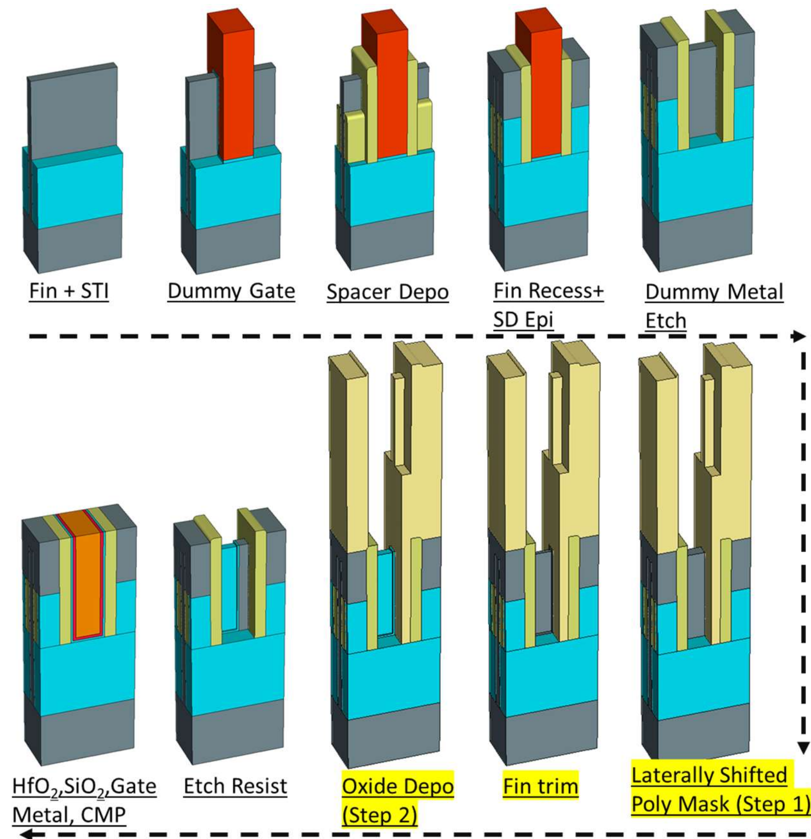


**Figure 4.3** Proposed process flow of laterally-shifted selective fin-trimming and gate stack engineering for the asymmetric performance of FinFET.

The final device structure after fabrication is shown in **Figure 4.5**. **Figure 4.5 (a)** shows the final 3-D structure, which is obtained after following all the steps as shown in **Figure 4.4**. On this structure, two cut-planes are taken to show the internals of the structure, and the cut-planes are visible in **Figure 4.5 (a)**. **Figure 4.5 (b)** shows the device cross-section across cut plane *C1*, wherein source-drain is marked. The central portion of this device is zoomed and shown in **Figure 4.5 (c)**. Herein, the effect of partial Si fin-trim with a lateral shift of mask is apparent. Fin thickness is small in some regions because of this fin-trimming.

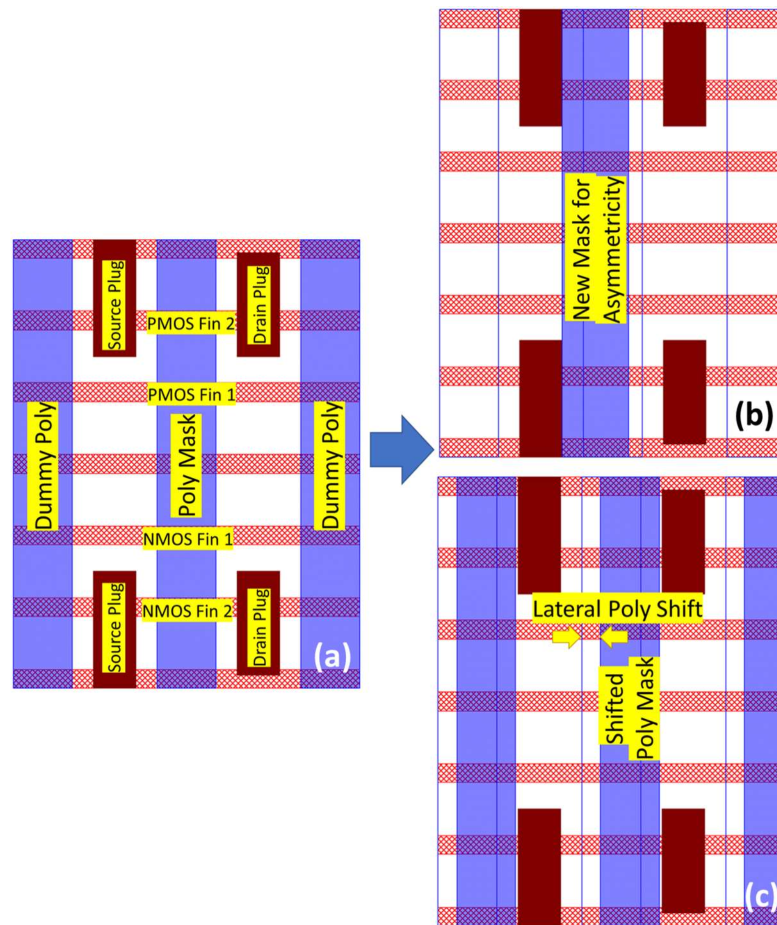


**Figure 4.4** (a) A typical layout for usual symmetric FinFET, depicting position of poly mask, fins and source/drain contact plug, (b) A modified layout to introduce asymmetricity by asymmetric spacers [122] –a new mask layer is required to enable this fabrication, (c) proposed layout to introduce asymmetricity in device. No new mask layer is required – poly mask is used with a lateral shift and a negative photoresist. Usage of this mask produces a structure, as shown in **Figure 4.3**.

This will create asymmetricity in the device *I–V* characteristics. Furthermore, the oxide which is deposited in this trimmed Si region is also highlighted. This oxide causes gate control to reduce in this portion of the device and thus causes more asymmetricity. In the final device, interface oxide, and high-*k* $HfO_2$ oxide are used and are also shown in **Figure 4.5 (c)**. In **Figure 4.5 (d)**, a 2-D device cross section along cut-plane *C2* is shown. The zoomed structure is shown in **Figure 4.5 (e)**, wherein it is visible that fin-trim causes a reduction in

fin-height.

As this process causes a reduction in fin-height and reduced gate control, it is expected that the overall performance of this device is compromised to enable the asymmetric nature of this device.

There are three parameters that one can toggle to modulate asymmetric electrical performance. One is the lateral shift of the mask, the second is fin-trim, and the third is the amount of oxide deposited. The performance comparison with all these parameters is discussed in the next section.

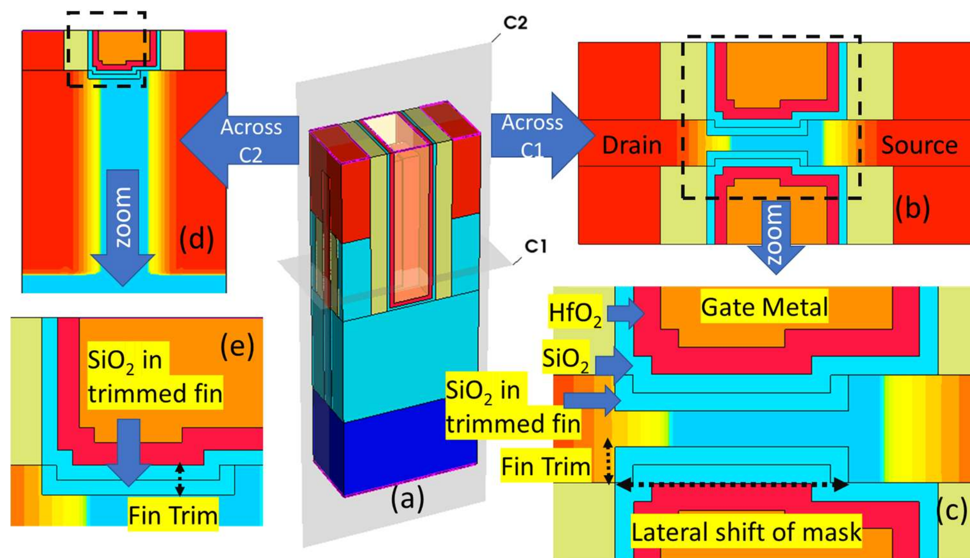


**Figure 4.5** (a) The 3-D asymmetric FinFET as fabricated from the process steps shown in **Figure 4.3** (b) A 2-D cross-section of the 3-D FinFET, along the cut-plane *C1*; (c) The zoomed picture of fin-region of 2D FinFET. Partial fin trim and additional gate oxide in trimmed fin portion are shown; (d) A 2-D cross-section of the 3-D FinFET, along the cut-plane *C2* (a); (e) A zoomed picture of fin-region of 2-D FinFET, showing the reduced fin-height and additional gate oxide.

### 4.3.2. Physics of proposed asymmetric device

This section presents the device physics of the proposed device shown in **Figure 4.5**. The pristine symmetric FinFET device is referred to as *Sym*. *Asym1* and *Asym2* refer to the same

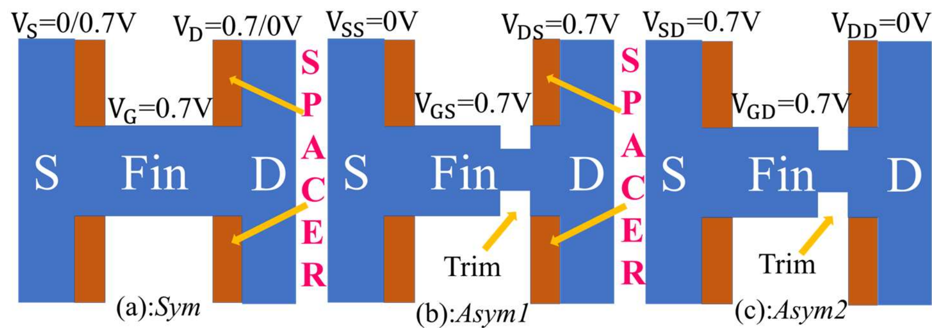


**Figure 4.6** Cross-section of the device (a) *Sym*: source/drain is grounded, (b) *Asym1*: source is grounded, (c) Asym2: drain is grounded.

trimmed fin device with different biasing conditions. The representation *Asym1* and *Asym2* only differ in terms of the biasing condition, as illustrated in **Figure 4.6**. In device *Asym1*, the source is kept at ground potential, and the drain voltage is ramped, while in Asym2, the drain is grounded, while the source voltage is ramped. The proposed device exploits the trimming of fin and oxide fill to achieve asymmetry. In order to show the impact of fin-trimming and oxide fill, the total resistance ($R_T$) offered by the device, which includes the source-drain resistance ($R_S$, $R_D$) and channel resistance ($R_C$) is plotted [see **Figure 4.7 (a)**] with respect to the inverse of overdrive voltage ($V_{GS}–V_T$). Further, we have obtained $R_C$ of the *Asym1* device which has a fin-trimming of 2 nm and oxide fill. The channel resistance of the asymmetric device has a different value when we interchange the source-drain terminal, which is clearly shown in **Figure 4.7 (b)** for *Asym1* and *Asym2*. A significant change in channel resistance was observed when the source-drain terminals are interchanged. The difference in channel resistance of the asymmetric device is around 7.68 KΩ for the two conditions, *Asym1* and *Asym2*.

This difference in the channel resistance can be further understood using **Figure 4.8,** which shows a representational view of the channel or inversion layer under the *Asym1* and

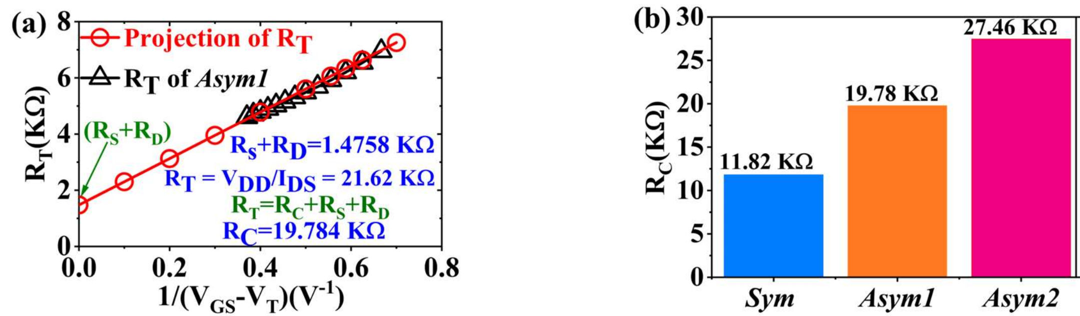


**Figure 4.7** (a) Total channel resistance($R_T$) vs inverse of overdrive voltage ($1/V_{GS}-V_T$) of device *Asym1* for biasing condition is $V_{SS}$ = 0 V and $V_{DS}$ = 0.7 V, (b) Comparison of channel resistance (Rc) of devices *sym*, *Asym1,* and *Asym2*.

*Asym2* biasing conditions. Herein, the pinch-off conditions are also shown. The device is "weak" on the drain side because of increased gate oxide thickness. However, in this case, *i.e.*, *Asym1*, it does not matter as much, because the source-side has "strong inversion" due to thinner gate oxide thickness, and the channel is having more charge carriers at the source side. Thus, the overall channel resistance is smaller for this device. Now consider **Figure 4.8 (b)**, wherein the polarities are reversed, *i.e.*, *Asym2*. Here, the effective channel potential toward the source is 0 V, and toward the drain, it is 0.7 V. However, as the drain side is "weak," because of increased gate-oxide thickness, a much lesser inversion charge density should be seen, leading to increased channel resistance, and eventually resulting in asymmetricity. This was also verified from the TCAD simulation result in **Figure 4.9 (a)** for the proposed device, during the *Asym1* condition, wherein the effective gate electric field is much more, in comparison to what it is for condition *Asym2* in **Figure 4.9 (b)**. A higher electric field in device *Asym1* results in more inversion charge density and, thus, more current and less channel resistance. The structural asymmetry due to trimming of the fin together with the asymmetry in the gate electric field due to oxide fill contribute to different channel resistance. The inversion-charge density for *Asym1* and *Asym2* cases is shown in **Figure 4.10**,

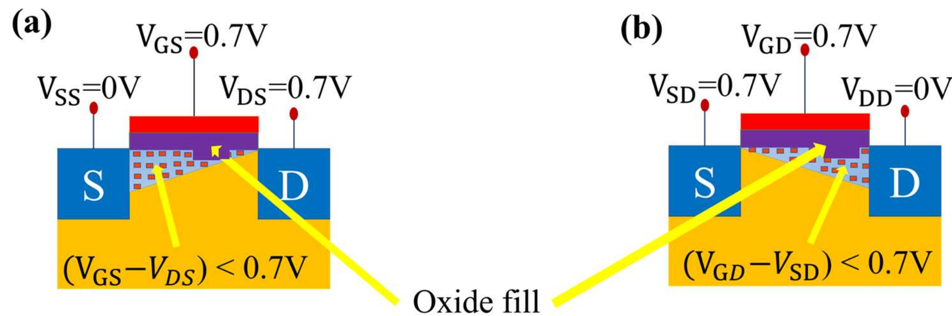


**Figure 4.8** Formation of inversion layer in channel for biasing condition (a) $V_{SS}$ = 0 V & $V_{DS}$ = 0.7 V, (b) $V_{SD}$ = 0.7 V & $V_{DD}$ = 0 V.

wherein *Asym1* shows about an order higher inversion charge carrier density, in comparison to *Asym2* conditions, for the reasons discussed above.

## 4.4 Results and Discussions

### 4.4.1. Asymmetricity ($I_{ON}$ Ratio) comparison

**Figure 4.11 (a)** shows the effect of asymmetricity in the proposed structure as a function of fin-trim and lateral mask shift at an oxide fill of 50% of the fin-trimmed amount. $I_{ON}$ ratio is defined as follows and is used as a measure of asymmetricity in the device's *I–V* characteristics, as also shown in [122].

$$\mathrm{I_{ON}\ Ratio} = \frac{I_{ON}(V_{drain} = V_{DD}, V_{source} = 0)}{I_{ON}(V_{drain} = 0, V_{source} = V_{DD})}$$

With the increase in both fin-trim and lateral shifting of the mask, the amount of asymmetricity increases. At 2 nm of fin trimming and 7.5 nm of lateral mask shift, an asymmetricity of 18.2% is observed in the ON-current. The fin-trim causes a reduction in both fin-width, as well as fin-height, thus causing a reduction in performance as well. Deposition of additional gate oxide also causes a performance hit, and the combined effect is depicted in **Figure 4.11 (b)**. $I_{ON}$ reduces by up to 50%, as more fin-trimming and lateral mask shifts are introduced

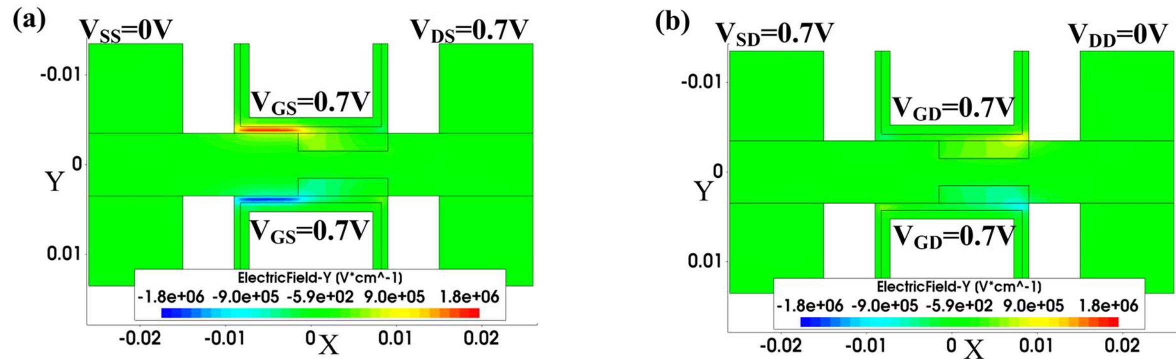


**Figure 4.9** Cross-section of asymmetric device (a) Distribution of electric field when biasing condition $V_{SS}$ = 0 V & $V_{DS}$ = 0.7 V, *i.e.*, *Asym1*, (b) Variation of electric field when biasing condition $V_{SD}$ = 0.7 V & $V_{DD}$ = 0 V *i.e.,Asym2.*

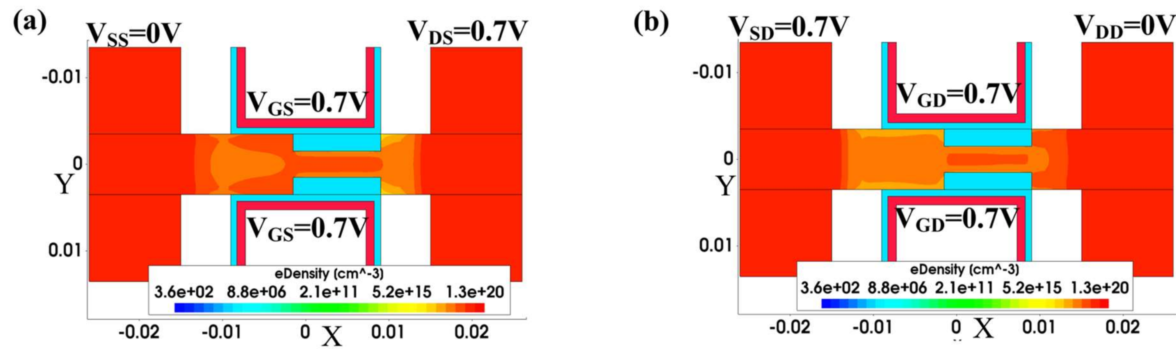


**Figure 4.10** Cross-section of the asymmetric device (a) Describes electron density profile for biasing condition $V_{SS}$ = 0 V & $V_{DS}$ = 0.7 V,*i.e. Asym1,* (b) Describes electron density profile for biasing condition $V_{SD}$ = 0.7 V & $V_{DD}$ = 0 V,*i.e. Asym2.*

to enable asymmetricity. This $I_{ON}$ has the same work function as the original device. Thus, for the final device architecture, in order to meet ON-current requirements, the number of fins of FinFET should be increased. Further, **Figure 4.12 (a)** shows the same plot as shown in **Figure 4.11 (a)**, but for a 100% oxide fill in the trimmed fin device. As expected, a thicker gate oxide leads to more asymmetricity in the device's electrical performance. It is seen that for a lateral shift of 7.5 nm and fin-trim of 2 nm, an $I_{ON}$ ratio of 27% is obtained. However, this comes at a price. The % reduction in $I_{ON}$ increases to 52% for this performance point.

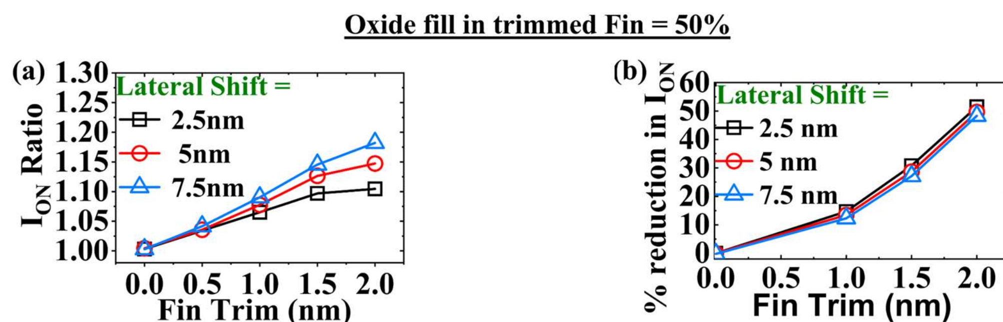


**Figure 4.11** (a) $I_{ON}$ ratio ($I_{ON}$ with drain ramped: $I_{ON}$ with source ramped), vs. Fin trim amount [**Figure 4.5 (c)**], as a function of lateral-mask shift, for oxide fill in trimmed fin = 50%. More the fin-trim and lateral shift, more is asymmetricity; (b) % reduction in $I_{ON}$ with drain ramped to $V_{DD}$, as a function fin-trim and lateral mask shift. At 2 nm drain ramped, even though asymmetricity is more, but % reduction in $I_{ON}$ is also significant.

This $I_{ON}$ is also at the same work function as the original device. Thus, almost two times the fin would be required to meet the $I_{ON}$ criterion.

### 4.4.2. Capacitance comparison

Capacitance–voltage (*C–V*) characteristics of the proposed device is compared to a simple FinFET, without any fin-trim, for the case of lateral shift = 7.5 nm and fin-trim of 2 nm. The plot of total gate capacitance (*Cgg*) for these devices is shown in **Figure 4.13**. Evidently, the proposed asymmetric device offers significantly lower capacitance with respect to a usual FinFET. This reduction in capacitance is nearly 17%, which should positively impact when used in STT-MRAMs. It is worth noting that the write line capacitance in STT- MRAMs determines the size of the drivers.

To compare with the previously proposed asymmetric device architectures, the AD structure is also simulated, which does not have a trimmed fin. The AD structure is the same as a conventional FinFET structure with a different source-drain doping. The results are shown in **Table 4.2**. The proposed device shows nearly 17% lower capacitance with respect to the

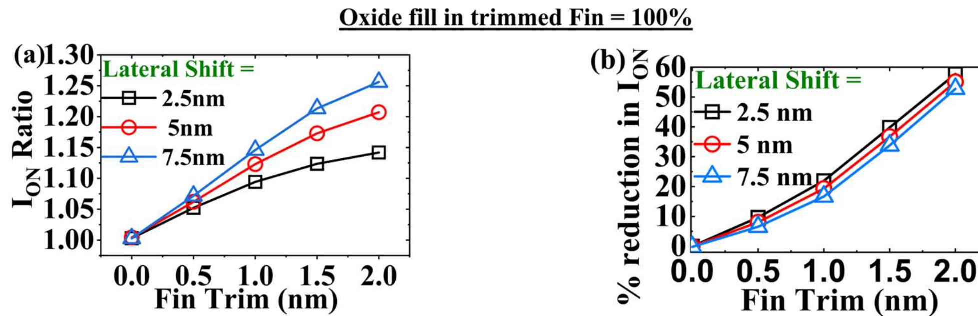


**Figure 4.12** (a) $I_{ON}$ ratio ($I_{ON}$ with drain ramped: $I_{ON}$ with source ramped), vs. Fin trim amount (**Figure 4.5 (c)**), as a function of lateral-mask shift, for oxide fill in trimmed fin = 100%. An attractive asymmetricity of 27% is observed; (b) % reduction in $I_{ON}$ with drain ramped to $V_{DD}$, as a function of fin-trim and lateral mask shift. 27% asymmetricity comes at a price of even more reduced performance.

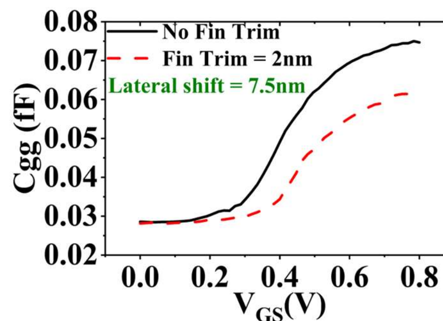


**Figure 4.13** Comparison of *C-V* characteristics, with and without fin-trim. With the fin-trim case, a nearly 17% reduction in capacitance is observed. No fin-trim case capacitance is also representative of asymmetrically doped FinFET devices.

previously proposed AD-metal oxide semiconductor (AD-MOS) asymmetric device architecture.

## 4.5 Conclusion

To design power-efficient STT-MRAM and MTJ-based non-volatile logic devices, asymmetric devices are quite important. In this work, we propose a novel device architecture, along with its CMOS-compatible process flow, for efficient asymmetric devices. The summary of advancements achieved in the proposed device is presented in **Table 4.2**. The $I_{ON}$ (at the same $I_{OFF}$ = 11.16-nA/μm fin-pitch) is comparable to that achieved in literature for

previously reported asymmetric devices [122]. The capacitance of the proposed device is nearly 17% less with respect to the previously reported asymmetric device, which is a significant advancement. Lower write line capacitance is very important to achieve high integration density for STT-MRAM memory technology. Both devices have nearly the same $I_{ON}$ asymmetricity ratio. However, the fabrication of the previously proposed device requires an extra mask to create asymmetricity. The device proposed in this work does not require any extra design of the mask, thus making it very economical and attractive.

**Table 4.2** Comparison of the asymmetric device proposed in [122] and in the present work.

| **Name of Parameter** | **AD MOS** | **This Work** |
|---|---|---|
| Capacitance | 1 | 0.83 |
| $I_{ON}$ ratio | 27% | 27% |
| Extra Mask | Yes | No |