Chapter 2

Literature Survey on Tunnel FETs

In this chapter, we present a detailed survey of literature on tunnel FET. Prior to that, the limitations of the CMOS technology for further extending the Moore's law, particularly for low-power operation, are discussed. Based on the literature review, a justification of the work undertaken for this thesis is made at the end of this chapter.

2.1 Limitations of CMOS Scaling

Silicon Complementary Metal Oxide Semiconductor (CMOS) technology has dominated the microelectronics industry (specially for the digital logic and semiconductor memories) over the past few decades for its main advantages of low power dissipation, excellent noise immunity, high packing density, a wide range of supply voltages, and ease to scale down the device dimensions for improved performance. Due to aggressive down-scaling, the device dimensions have surpassed the micrometer scale and are now in the nanometer regime. This continuous downward scaling for improved performance of CMOS technology has reached the fundamental limits of material and fabrication technology. In the nanometer regime, the power management has become the most crucial issue for further scaling of this technology. Due to dimension scaling, severe short-channel effects such as Drain Induced Barrier Lowering (DIBL) causes a substantial increase in the leakage current. Furthermore, the dimension scaling necessitates scaling of the supply voltage $V_{DD}$ in order to reduce power density. This, in turn, demands a reduction in the threshold voltage $V_T$ in order to achieve commensurate device performance. The subthreshold swing $SS$, defined as the gate voltage required to change the drain current by one order of magnitude (one decade) has, however, the fundamental physical limit of $60mV/\text{decade}$ ($kT/q$). In 10) for conventional MOSFETs at room temperature. Due to this limitation, any effort to reduce $V_{DD}$ will further increase the leakage current
and, therefore, the $V_{DD}$ cannot be scaled further for conventional CMOS technology. This fundamental subthreshold-swing limitation stems from the thermal injection mechanism that responsible for carrier injection from the source to the channel in the MOSFET.

Dennard's scaling rules [1] no longer work as they did in the past. The reason can be seen in Fig. 2.1, which shows the scaling trend from the 1.4 μm node to the 65 nm node. Decrease of supply voltage $V_{DD}$ to about 20% of its original value is accompanied by the decrease of threshold voltage $V_T$ to approximately half of its starting value. That threshold voltage decrease did not happen naturally from Dennard's scaling; rather it had to come about in another ways, such as varying the doping concentration of the channel region under the gate. The electric fields inside a MOSFET stay almost constant when we go after the scaling rules correctly; and so the threshold voltage stays almost constant as well, if we do not make other changes.

![Figure 2.1: The trend of supply voltage and threshold voltage scaling vs. technology generation. $V_{DD}$ decreases with device dimensions, but $V_T$ does not [2].](image)

With $V_{DD}$ scaling by keeping $V_T$ nearly constant, the gate overdrive, as shown in Fig. 2.1, reduces. Decrease in gate overdrive leads to decreases in on-current, which has adverse effects on device performance in terms of switching speed ($C_g V_{DD}/I_{ON}$). The two possible solutions to this problem are: (a) $V_{DD}$ can be made higher than it
should with constant field scaling, or (b) $V_T$ can be scaled down more aggressively.
But, both of these options have their own consequences, which will be discussed.

But, the formerly-followed scaling trends of $1/\kappa = 0.7$ every 2 or 3 years no longer hold true for $V_{DD}$ [3]. To maintain the acceptable levels of gate overdrive voltage, $V_{DD}$ scaling has slowed down considerably. Decrease in supply voltage along with device dimensions, results in constant power density $I_{ON}V_{DD}/A$ (on-current times supply voltage divided by surface area), which implies that the energy required driving the chip, and the heat produced by the chip, remain almost constant. So chip size is not decreasing with scaling down of devices, rather more complexity and functionality are added with each generation, and size of the chip remains relatively constant.

But power density would increase unless $V_{DD}$ is scaled down. The dynamic and static power consumption, for a MOSFET, can be expressed as [4].

$$P_{dynamic} = fC_LV_{DD}^2 \quad (2.1)$$

where $f$ is the frequency and $C_L$ is the total switched capacitive load, and

$$P_{static} = \text{I}_{\text{leak}}V_{DD} \quad (2.2)$$

where $I_{\text{leak}}$ is the sum of the leakage currents in the device when the MOSFET is in the off-state.

If $V_{DD}$ is not scaled down proportionately with device dimensions, and more devices are included within a chip such that chip size is not reduced much, then the obvious result is considerable increase in power consumption, and the modern trend of that is depicted in Fig. 2.2.
Figure 2.2: Trends of dynamic and static CMOS power, showing that static power consumption has become a greater problem than dynamic power consumption [4].

The impact of reducing $V_T$ on $I_{OFF}$ (i.e., $I_D$ at $V_{GS} = 0 \, V$) can be visualized in Fig. 2.3. A reduction in $V_T$ by 60 mV results in increase in $I_{OFF}$ by one order of magnitude.

Figure 2.3: A typical $I_{DS}$-$V_{GS}$ curve for a highly-optimized conventional MOSFET, showing the subthreshold swing limited to 60 mV/decade at room temperature. If we want to decrease $V_T$ by shifting the curve left, we pay a price in leakage current. Solid curve’s data is from an optimized asymmetrical double-gate conventional MOSFET in [5], which is then shifted three times. Such a shift could come from engineering the gate work function, for example.
The ICs will be even more usable if they consume low power and low energy. The reasons behind this are as follows. As far as the global energy conservation and benefit of environment is concerned, we would like our electronic gadgets to consume less power. Again it’s less expensive to use less electricity. For battery-operated gadgets it is better as their batteries will last longer before needed to be recharged. Also, for laptops and hand-held gadgets having a lower power density means generation of less heat. The trend of increasing power for Intel computer chips [6] can be observed from Fig. 2.4. An ITRS presentation predicted that the power density for the 14 nm node would be greater than 100 W/cm² [7]. The right axis of Fig. 2.4 suggests about the heat management issue which tells that, to manage the increasing power density, the heatsink must grow in volume. Again this also has a limit, since we prefer our electronics gadgets to retain their original size or rather to shrink, but not to become larger to accommodate a large heatsink for managing the dissipated power and heat of the chips inside. It is because; the portability issue would then be a concern in terms of weight and size of the gadgets.

Figure 2.4: Computer chip power trends, along with the accompanying increase in heatsink volume [6].
2.2 Device-Level Solutions

To circumvent the problem, researchers are looking for architecture-level and circuit-level innovations. Besides, they are also trying to find a device-level solution by changing the devices themselves. Here, we only discuss about the device-level solution, because our work is completely based on that. Novel device structures and material options are explored to overcome the problem and improve performance. The objective is to search for an alternative but conventional MOSFET-compatible device with a low subthreshold swing SS (i.e. SS < 60 mV/decade at room temperature), required to maintain high $I_{\text{ON}}$ with an acceptable $I_{\text{OFF}}$ for further extending the Moore's law for digital and memory applications. The subthreshold swing is inherently related to the physical mechanism of carrier injection inside the device in subthreshold region. As electrons go over a potential barrier by thermal injection for MOSFETs, the leakage current is determined by the Boltzmann distribution and that limits SS to 60 mV/decade at room temperature. To overcome this limit for swing, some groups also started to think that it would be necessary to alter the physical mechanism of carrier injection within the device.

The devices that are demonstrated to have sub-60-mV/decade at room temperature are tunnel FET (TFET) and impact ionization MOSFET (IMOS). A TFET is essentially a gated p-i-n diode that exploits a band-to-band tunneling mechanism to overcome the subthreshold swing limitation of a conventional MOSFET. On the other hand, the IMOS structure consists of a p-i-n diode where the intrinsic region is partially gated, so that a very high electric field exists in the non-gated portion of the i-region when the device is on, leading to avalanche breakdown [8]. The impact ionization process means that the IMOS can have a very small subthreshold swing and high on-current [9]. Experimental transfer characteristics also reveal SS of about 4 and 9 mV/decade for n-channel and p-channel devices, respectively [10]. The main drawback of IMOS device include, (i) the scalability issue, due to the presence of a gated and an ungated region between the source and drain, (ii) reliability issue, as the hot carriers, created by impact ionization, can go into the gate oxide, and (iii) the difficulty of low-voltage operation, as high voltages
are required to induce breakdown. Due to these severe problems of IMOS, the novel and simple device structure of the TFET is explored and has become a strong contender to replace the present state-of-the-art MOSFET for further extending the Moore's law.

### 2.3 The Tunnel Field-Effect Transistor

The potential advantage for such devices is that they produce very low off-current with a small subthreshold swing. The presence of quantum tunneling barrier for Tunnel FETs makes these devices interesting for low-power application. When the devices are turned on, current starts flowing from source to drain as the electrons move through the energy barrier (i.e. tunneling), but not over it like in MOSFETs. When the devices are turned off, the tunnel-width becomes large making the off-current extremely small, which can be several orders of magnitude lower than that of a conventional MOSFET.

A typical Tunnel FET, as shown in Fig. 2.5, is basically a gated p-i-n diode. The tunneling takes place in this device between p+ source-region and the intrinsic-region, when the diode is reverse biased by application of a voltage to the gate terminal. For device operation, the p-i-n diode is reverse-biased, the source is grounded, and a positive voltage is applied to the drain, and a voltage is applied to the gate. Without a gate voltage, the width of the energy barrier between the intrinsic region and the p+ region is much wider than 10 nm (the approximate minimum for significant tunneling probability), and the device is in the OFF-state, as shown in Fig. 2.6(a) [11]. As the positive gate voltage increases, the bands in the intrinsic region are pushed down in energy, narrowing the tunneling barrier and allowing tunneling current to flow, as shown in Fig. 2.6(b) [11]. In order to be consistent with MOSFET technology, the names of the device terminals are chosen such that voltages are applied in a similar way for Tunnel FET operation. Since a reverse bias is needed across the p-i-n structure in order to create tunneling, and since an NMOS operates when positive voltages are applied to the drain and gate, the n-region of a Tunnel FET is referred to as its drain, and the p+ region as its source for an n-type device [11]. For the n-
channel Tunnel FET of Fig. 2, we have a p⁺ source and an n⁻ drain. In a p-channel Tunnel FET, we have an n⁺ source and a p⁻ drain.

![Cross-sectional view of a simple n-channel TFET with single gate. The intrinsic substrate/well is tied to source contact. While the drain extension (left) is a highly doped n⁻-region, the source extension (right) is a highly doped p⁺-region.](image1)

![Figure 2.6: (a) Schematic of energy-band diagram of the OFF-state of the Tunnel FET at V_{DS} = 1 V and V_{GS} = 0 V. In this state, the only current is p-i-n diode leakage current. (b) Schematic of energy-band diagram of the ON-state of the Tunnel FET at V_{DS} = 1 V and V_{GS} = 1.8 V. In this state, the energy barrier is thin enough that electrons can tunnel from the valance band of the p⁺ region to the conduction band of the intrinsic region [11].](image2)

A symmetric structure of TFET (similar doping levels, similar gate alignment, etc. between the n- and p-sides) shows ambipolar behavior, and thus the transfer characteristics resemble those of a p-TFET when a negative voltage is applied to the gate, and those of an n-TFET when a positive voltage is applied to the gate. The energy bands in the intrinsic region under the gate are pushed up in energy value, and
the band-to-band tunneling take place between the valence band of the intrinsic region and the conduction band of the n⁺-region. On the other hand, when a positive voltage is applied to the gate, the energy bands in the intrinsic region are pushed down, as in Fig. 2.6(b), and the energy barrier is now small enough for tunneling to take place between the valence band of the p⁺-region and the conduction band of the intrinsic region. The energy barrier width for band-to-band tunneling is the main important factor for determining the amount of drain current for a Tunnel FET.

The dependence of the energy barrier width on the gate voltage for several different gate dielectric constants for n-TFET can be seen from Fig. 2.7(a). The barrier width starts to saturate at high VGS. It can also be seen from Fig. 2.7(b) that the on-current of an n-type Tunnel FET increases exponentially with a reduction in the barrier width between the intrinsic and p⁺ region.

![Figure 2.7](image)

**Figure 2.7:** (a) Width of energy barrier for band-to-band tunneling, vs. VGS, for different values of the gate dielectric constant. The value of the barrier width was extracted from the narrowest location on the simulated band diagrams at the tunnel junction (as shown in the inset), at a distance of 2.5 nm from the dielectric surface, with 1 V applied to the drain and the source grounded. (b) Drain current vs. energy barrier width for different values of the gate dielectric constant. Double-gate Tunnel FET with Lg = 50 nm, ss = 10 nm, tsubstrate = 3 nm, VDS = 1 V [11].

The dependence of subthreshold swing on gate voltage up to the threshold voltage (taken at IDSS = 10⁻⁷ A/μm) is shown in Fig. 2.8, demonstrating two important things [11]. Firstly, the subthreshold swing of Tunnel FETs is not constant, but rather is a
function of gate voltage. Secondly, at low gate voltages, Tunnel FETs can have a subthreshold swing less than 60-mV/dec.

![Figure 2.8: Dependence of the Tunnel FET subthreshold slope on gate voltage for different dielectric constants, from numerical simulation. Each curve goes up to the threshold voltage of that device. $L_g = 50 \text{ nm}$, $t_{\text{dielectric}} = 3 \text{ nm}$, $t_{Si} = 10 \text{ nm}$, $V_{DS} = 1 \text{ V}$. The points were generated by taking the swing value ($dV_{GS}/d(\log I_{DS})$) at each point on the $I_{DS}$-$V_{GS}$ curves [11].](image)

The benefits of a DG TFET over a DG MOSFET is shown in Fig. 2.9 [5], which compares the transfer characteristics of an optimized asymmetrical DG MOSFET with that of a simulated DG Tunnel FET. The two devices have the same dimensions for dielectric thickness (3 nm), channel length ($i$-region length in the Tunnel FET, equal to 50 nm), and body thickness (10 nm). The optimized DG TFET uses a high-$\kappa$ gate dielectric with a dielectric constant of 29. It is important to notice the difference between the subthreshold regions in these two types of devices. A MOSFET has a constant slope between the OFF-state and threshold. A Tunnel FET, however, demonstrates a slope that is steeper (smaller swing) closer to the OFF-state and less steep closer to threshold and varies as a function of the gate voltage, as can be noted from Fig. 2.8. Since the threshold voltage of a TFET cannot be extracted using certain standard MOSFET techniques, Boucart et al. [11] used the constant-current method, with a threshold current of $10^{-7} \text{ A/}\mu\text{m}$. It is clear that the swing for the TFET in Fig. 2.9 is lower than that of the MOSFET, whether we look at the point slope or the average slope. The point value of SS is defined as the minimum swing value at any
point on the $I_{DS}-V_{GS}$ curve. The average SS value is calculated between the voltage at which the current begins to increase with increasing gate voltage, and the threshold voltage. These two values are different because SS is a function of $V_{DS}$. Their extraction method is shown in the inset of Fig. 2.9.

![Graph](image-url)

**Figure 2.9:** Qualitative comparison of $I_{DS}-V_{GS}$ for an optimized asymmetrical conventional DG MOSFET [5] and a DG Tunnel-FET. While the subthreshold slope is constant for a MOSFET in the subthreshold region, it is a function of $V_{GS}$ at $V_{DS} = 1$ V for a Tunnel FET. Both devices have $L_g = 50$ nm, $t_{body} = 5$ nm, and $t_{dielectric} = 3$ nm. In general, Tunnel FETs have a much lower off-current and a lower on-current than conventional MOSFETs. Inset: Extraction of the average slope, and the point at which the point slope is measured [11].

### 2.4 History and Status of Tunnel FET

Quinn et al. at the Brown University [12] in 1978 first proposed the formation of a surface-tunnel transistor, by replacing the n-type source of an n-MOSFET with a p-type degenerate source, for the use of this device for spectroscopy. In 1987, Banerjee et al. at the Texas Instruments [13] studied the behavior of a three-terminal silicon tunnel device using a p-region instead of the i-region under the gate. The paper discussed a closed-form analytical model for the tunneling current in the trench transistor cell (TTC). Takeda et al. at Hitachi [14] in 1988 proposed a band-to-band tunneling MOS device on silicon, named as the B$^2$T-MOSFET, which shows greatly suppressed $V_T$ roll-off and reduced temperature dependence of the device characteristics.
In the initial years, the TFET was more viewed [C. Seabaugh [15], A. M. Ionescu [16], Leburton et al. [17]] as a high-speed transistor in which the gate was used to control the negative differential resistance (NDR). In 1992, T. Baba at the NEC [18] independently demonstrated lateral TFET structure, named as surface tunnel transistors (STT), using MBE technique to create mesa structures in III-V materials. In 1994, Uemura and Baba at the NEC [19] observed, for the first time, NDR at room temperature due to interband tunneling in a STT made of GaAs, under forward-biased conditions. The STT was demonstrated in 1995 on bulk Si by Kawaura et al. [20], and in 1996 on SOI substrate by Y. Omura [21]. None of them, however, showed NDR at room temperature. The room temperature NDR in silicon STT was first demonstrated by Koga and Toriumi at the Toshiba [22], [23].

In 1995, Reddick and Amartunga at the Cambridge [24] proposed gating of the reverse Zener-tunneling current of the STT, which is named as TFET later on, as a mean to achieve better scaling due to the absence of punch-through. In 1996, Zhang et al. [25] proposed a design for TFET with greatly suppressed the short-channel effects. The most important feature in the design of the transistor was the use of a heterojunction quantum well as the (internal) drain. In 1997, Wang et al. [26] experimentally demonstrated the TFET with an ultrashort channel-length of 25 nm, using InAs/AlSb single-QW heterostructures. The paper also predicted that, TFETs with nanometer-scale channel length would bring a 100-fold increase in VLSI circuit complexity over that of the then 0.1 mm MOSFET technology. Hansch et al. [27] at the University of the German Federal Armed Forces in Munich in 2000 fabricated a Si vertical TFET and noted its potential for low off-current relative to the MOSFET. The necessary sharp doping profile structure was created by means of MBE and also the saturation behavior in the $I_D-V_G$ characteristics was noted. In 2001, Hansch et al. [28] showed an improvement in the characteristics of a fabricated vertical TFET, made of silicon, by a boron surface phase. In 2002, S. Sedlmaier et al. in Germany [29] investigated phonon assisted tunneling controlled by gate bias and confirmed by output characteristics in vertical TFET. The paper also showed that, the simulated transfer characteristics qualitatively agreed well with electrical measurements and
promised device improvement regarding threshold voltage, subthreshold slope and gate-controlled current gain for sharper doping profiles and reduced gate oxide thickness. In 2004, Aydin et al. [30] demonstrated experimental characteristics of a lateral SOI TFET without an intrinsic region, instead placing the gate over a p-n junction, claiming that this would reduce gate capacitance and therefore increase speed. The authors also claimed that there should be no current saturation for these devices, which did not have an inversion channel and so were not subject to scaling rules of standard Si transistors. In the same year Wang et al. [31] at the TUM in Munich, Germany demonstrated experimental characteristics for complementary Si TFET (CTFET), which were fabricated on the same wafer using a CMOS compatible process.

2.4.1 Reports on SS less than 60mV/dec

The first discussion of sub-60-mV/decade subthreshold swing in a TFET appeared in 2004 by different groups [Wang et al. [32], Bhuwalka et al. [33], and Appenzeller et al. [34]]. Zhang et al. [35] derived an analytic expression for the sub-60-mV/decade SS at room temperature. Experimental demonstration of sub-60-mV/decade SS have been reported in TFETs based on carbon nanotubes (CNT) [34], [36], Si [37]–[40], Ge [41], and p⁺Ge-n⁺Si [42] channels. In 2007, Knoch et al. [43] in Germany investigated with simulation the influence of the dimensionality on the performance of TFETs. They showed that in a three-dimensional tunneling FET it is possible to achieve inverse subthreshold slopes smaller than 60mV/dec; but however, there was a trade-off between high on-currents and small values for the subthreshold swing. In 2008, N. Patel et al. at the IISC-Bangalore, India, proposed a new n-type classical-MOSFET-alike TFET architecture, which offered sub-60mV/decade SS along with a significant improvement in $I_{ON}$ [44]. The enhancement in $I_{ON}$ was achieved by introducing a thin strained SiGe layer on top of the silicon source. Through 2D-simulations they observed that the device was nearly free from short channel effect (SCE) and its immunity towards drain induced barrier lowering (DIBL) increased with increasing germanium mole fraction.
2.4.2 Reports on Length Scaling

Bhuwalka, et al. [45] discussed scaling rules for silicon Tunnel FETs, and pointed out that constant field scaling would not apply. It is reported in [46] that the TFET characteristics are nearly independent of channel length scaling. A detailed length scaling in DG TFETs was also carried out in [47], where it had been reported that the improved control of the tunnel junction provided by high-k dielectrics improves the gate length scaling trends of the DG TFET, than those with a SiO₂ gate dielectric. It was also reported that \( g_m \) and \( g_a \) for Tunnel FETs show a different dependence on applied voltages and gate length than those seen in a conventional MOSFET.

2.4.3 Reports on Si TFET

A detailed device optimization of TFET has been reported by Boucart and Ionescu [11]. Impact of various device structural parameters (such as single or double gate, doping levels in the source, drain, and intrinsic regions, gate dielectric permittivity, silicon body thickness, and band gap at the tunnel junction) on the characteristics was looked at. A double-gate (DG) architecture with a high-k gate dielectric was reported to improve the device performance. It had also been shown that, an asymmetrical lateral strain profile, or a heterojunction, would boost on-current by reducing the band gap only at the tunnel junction side of the device, while keeping off-current low with a large band gap at the drain side. In 2007, Toh et al. at the National University of Singapore studied the impact of body thickness optimization on silicon DG TFET, and explored the device physics through two dimensional device simulations [48]. It was established that band-to-band tunneling at the surface is very strong and accounts for a large part of the total drain current. Detailed potential distributions showed that the coupling of two gate electrodes in the DG TFET could effectively reduce the tunneling width \( \omega_T \) at the center of the silicon film up to an optimum \( T_{Si} \) where maximum drain current is obtained. In 2007, Verhulst et al. at the IMEC showed by simulation that shortening Tunnel FET gate length, so that the gate covers the source-side junction where tunneling takes place, but does not cover the majority of the intrinsic region, has the benefits of decreasing off current (tunneling through the drain-side junction) and reducing speed, with a
small or no reduction in the on-current, depending on the device design [49]. In 2009, Schlosser et al. [50] at the University of the German Federal Armed Forces in Munich made a simulation-based study on the influence of using a high-κ gate dielectric in the TFET compared to a standard SiO₂ with same equivalent oxide thickness (EOT), which exhibited a quite different behavior compared to a conventional MOSFET due to its totally different working principle. The fringing field effect, while deteriorating conventional MOSFET characteristics, improves the characteristics for TFET significantly. This also gives rise to technological development of high-κ materials with permittivities > 30, which has been regarded as impractical until then due to the FIBL effect in conventional MOSFETs. Sandow et al. [51] at the Forschungszentrum Jüllich reported experimental data for p-type TFET on SOI, showing the effects of varying source and drain doping levels, gate dielectric thickness, and device length. He also reported the typical output characteristics for the device that exhibited an exponential onset followed by distinct saturation, and this saturation currents also showed no dependence on the gate length, as can be seen in Fig. 2.10.

Figure 2.10: Typical output characteristics of a p-TFET with gate length $L = 6\mu m$, gate oxide thickness $d_{ox} = 4.5$ nm, and implantation doses $D_{As} = 3 \times 10^{15}/cm^2$ and $D_{b} = 3 \times 10^{15}/cm^2$. The exponential onset is expected due to the large quantum capacitance in the channel, which for small $V_{DS}$ prevents the bands in the channel from following the external gate potential. The total current level reflects the small band-to-band tunneling transmission due to the thick gate oxide. Inset: saturation currents measured at $V_{DS} = -3.5$ V for different gate lengths $L$ and $V_{GS}$. No dependence of the saturation currents on the gate length is observed. The threshold voltage spread is negligible.
In 2010, Anghel et al. at the Institut Superieur d'Electronique de Paris, France, proposed for an improved DG TFET structure (low-k spacer combined with a high-k gate dielectric) with superior performance [52]. The proposed structure increases $I_{ON}$ by a factor of 3.8 and reduces the subthreshold slope by a factor of 2 compared to other structures reported at that time. In 2010, Choi and Lee at the Sogang University, Korea, proposed hetero-gate-dielectric TFETs, incorporating different gate dielectric materials at the drain and the source ends [53]. The proposed device enhances ON-current, suppresses ambipolar behavior, and makes abrupt ON-to-OFF transition by replacing the source-side gate insulator with a high-k material, which induces a local minimum of the conduction band edge at the tunneling junction. In 2010, Virani et al. [54] at the IIT-Bombay, India, proposed a dual-k spacer for both nonunderlap and underlap n-TFETs, the impact of which is shown in Fig. 2.11. The dual-k spacer consists of an inner layer made of a high-k material and an outer layer made of a low-k material. Fig. 2.12 shows that the dual-k spacer improves the performance of n-TFETs and more so for the underlap structures. The structure was also optimized for the $I_{ON}$ without degrading the $I_{OFF}$ or the SS.

Figure 2.11: n-TFET structures with the dual-k spacer investigated in [54]; (a) Nonunderlap structure and (b) underlap structure.
Figure 2.12: $I_D - V_{GS}$ characteristic for (a) a nonunderlap n-TFET structure with a SiO$_2$ spacer, (b) a nonunderlap n-TFET structure with a dual-k spacer made of HfO$_2$ and SiO$_2$, and (c) an underlap n-TFET structure with a dual-k spacer made of HfO$_2$ and SiO$_2$. The gate dielectric is 1.1-nm SiO$_2$ and $V_{DS} = 1$ V.

In 2011, Jhaveri et al. [55] at the University of California, first investigated the source-pocket TFET (through simulation and experiment). The source-pocket TFET showed a higher $I_{ON}$ (~10 times) and steeper subthreshold swing as compared to an ordinary p-i-n TFET. The ambipolar conduction was also reduced due to the use of a low-doped drain extension. Saurabh and Kumar [56] in 2011 at the IIT-New Delhi, India, reported the application of a dual material gate (DMG) in a TFET to simultaneously optimize $I_{ON}$, $I_{OFF}$, $V_T$, SS etc. It was shown that, if appropriate work functions are chosen for the gate materials on the source side and the drain side, the TFET shows a significantly improved performance. Heteromaterial gate TFET had also been proposed in [57] that demonstrated, through 2D numerical simulations, that a local minimum of the conduction band in the channel was formed by work function mismatch, which resulted in the abrupt transition between the on- and the off states and significant drive current enhancement. In 2011, Anghel et al. [58] at the Institut Superieur d'Electronique de Paris, France, showed for a heterodielectric gate structure that $I_{ON}$ can be boosted by correctly positioning the source with respect to the gate edge. He also demonstrated that the ambipolar character of the TFET is
completely inhibited by using only one spacer of 30-nm length to separate the drain and the gate. In 2011, Guo et al. at the University of California, Berkeley, investigated the effect of back biasing on the performance of a planar TFET, implemented on silicon-on-insulator (SOI) substrate [59]. It was found that reverse back biasing reduced SS and increases the range of drain current over which SS is less than \((kT/q)\ln(10)\).

2.4.4 Reports on SiGe TFET

In 2006, Bhuwalka et al. at the University of the German Federal Armed Forces in Munich proposed a lateral Tunnel FET on SiGe on insulator [60], and showed through simulation that on-current would increase with increased percentage of Ge in the SiGe. It was also shown that \(I_{on}\) was nearly independent of gate length scaling and temperature. In 2007, Toh et al. [61] at the National University of Singapore studied the device physics and guiding principles for the design of the SiGe heterojunction in a DG TFET. The tunneling junction for their device, at high drain and gate biases, was located at a distance of \(\sim 4\) nm to the left of the gate edge. It had been suggested that, to harness the benefits of the higher BTBT rate of SiGe having a narrower bandgap, the SiGe heterojunction should be located at \(\sim 2\) nm to the right of the gate edge, as it would allow for the maximum band-to-band tunneling rate and hence the on-state current. They also pointed out that, too much overlap of SiGe with the gate would result in decrease in on-state current due to a decrease in the electric field. In 2008, again the same group studied a DG TFET with silicon–germanium (SiGe) source to overcome the scaling limits of CMOS technology to further extend Moore’s law [62]. Less than 60mV/decade SS with extremely low off-state leakage current was achieved by optimizing the device parameters and Ge content in the source. With extensive simulations, they showed for the first time that, such a technology would proved to be viable to replace CMOS for high performance, low standby power, and low power technologies. In 2009, Virani and Kottantharayil at the IIT-Bombay, India, proposed the use of high-k spacers to boost the ON state current of SiGe-Si heterojunction TFETs [63]. It was shown that the fringing fields through the spacer enhance the ON-state current without modifying the OFF-state current or the subthreshold
swing. In 2009, Khatami and Banerjee at the University of California, Santa Barbara, proposed a novel n- and p-type TFETs based on heterostructure Si/intrinsic-SiGe channel layer, which exhibit very small subthreshold swings, as well as low threshold voltages [64]. The performance of the TFET-based inverter was compared with the 65-nm low-power CMOS-based inverter, and an improvement in static power consumption by four orders of magnitude was achieved for the TFET-based inverter with smaller gate delay. In 2010, Virani et al. [65] at the IIT-Bombay, India, presented optimization techniques for 20nm channel length novel Si/SiGe heterojunction p–i–n p-channel TFETs using extensive device simulations. It was shown that depending on the Ge mole fraction in SiGe and the gate voltage, the tunneling could be from the channel to source or within the source only.

### 2.4.5 Reports on Ge TFET

Investigation of the device physics and electrical characteristics of the germanium (Ge) TFET for high performance and low power logic applications using two dimensional device simulations was done in 2008 [66]. It was shown for the first time that the high off-state leakage due to the drain-side tunneling in the Ge TFET can be effectively suppressed by controlling the drain doping concentration. A lower drain doping concentration reduces the electric field and increases the tunneling barrier width in the drain side, giving a significantly reduced off-state leakage. In 2009, Kazazis et al. at the Brown University, France, fabricated TFET on thin GeOI that showed very high leakage, with $I_{ON}/I_{OFF} < 100$ [67]. In 2010, Kim et al. at the UC-Barkley, USA, investigated the impact of body doping and thickness on the performance of a germanium-source TFET in which band-to-band tunneling occurs entirely within the source region [68]. It was found that the dominant leakage mechanism varies depending on the body design parameter values. In 2010, the same group proposed [69] a TFET with a raised germanium (Ge) source to achieve steep switching behavior and higher $I_{ON}$ for low-voltage ($V_{DD} < 0.5$ V) operation by carefully suppressing the lateral carrier tunneling in the device.
2.4.6 Reports on Strained TFET

A simulated result of strained DG n-TFET using single-layer strained-silicon-on-insulator (SSOI) technology was reported in [70]. SSOI is a novel SiGe-free material system that has the advantage of strained silicon while improving the scalability of thin-film SOI. Due to uniform strain throughout the device, the $I_{ON}$ improved along with increase in $I_{OFF}$. To improve this increase in $I_{OFF}$, Boucart and Ionescu at the EPFL, Switzerland, along with Riess at the IBM Zurich Research Laboratory, Switzerland proposed a lateral asymmetric strain profile in a silicon nanowire or ultrathin silicon film as a key technology booster for the performance of all-silicon TFETs, in 2009 [71]. They also claimed that inherent finite drain threshold voltage of the Tunnel FET, which could be a disadvantage for logic design with Tunnel FETs, is exponentially reduced with the strain-induced bandgap shrinkage at the source side. It was shown in [72] that incorporation of pseudomorphic strained-Si$_{1-x}$Ge$_x$ layers leads to a significant performance increase.

2.4.7 Reports on Novel Structures

In 2008, Nagavarapu et al. [73] at the UCLA proposed the concept of a novel tunnel source (PNPN) n-MOSFET based on the principle of band-to-band tunneling, in which a narrow region of the opposite doping is introduced into the TFET source under the gate edge. This narrow region acts as a source of electrons, and increases the band bending and the electric field at the tunnel junction, thus increasing on-current with SS below 60 mV/dec at 300K. Formation of a tunneling junction at the source under the gate was confirmed by experimental demonstration of the device using spike rapid thermal anneal. It had been pointed out that, abruptness of the tunneling junction profile was an important area of concern and novel annealing methods can be potential candidates for successful device realization. It was also shown that, tighter control of pocket width can further reduce the SS below the 60 mV/dec. Woo et al. at the Stanford University, in 2008, reported [74] with simulation based investigation that, Si TFETs would possibly be unable to achieve the necessary ON-current to compete with MOSFETs, due to the exponential drain characteristic, a problem fundamental to TFET structures. It was suggested that overcoming this
problem would require device improvements and more efficient tunneling materials to achieve higher ON-currents. Besides, it was prescribed to engineer novel BTBT based devices without exponential drain voltage characteristics. In 2011, Cao et al. at the Fudan University, Shanghai, China, showed that the insertion of a thin n-layer into the tunneling junction of the p-i-n TFET (p-n-i-n TFET) not only enhances its drive current, as has been previously reported [55], but also improves its reliability, such as the direction of the electric field near the tunneling junction, the threshold-voltage shift induced by dielectric charge, and the variation of the threshold voltage due to device size fluctuations as compared with the p-i-n TFET [75]. In 2011, Tura et al. [76] at the Intel Corporation, Hillsboro, reported an experimental study where very sharp (~1.2 nm/dec) optimized tunneling-junction dopant profile for the silicon p-n-p-n TFET was demonstrated by molecular beam epitaxial growth. It was shown that, as compared with a p-i-n TFET, the p-n-p-n TFET has 30% lower subthreshold swing and three times higher $I_{ON}$. In 2011, Asra et al. [77] at the IIT-Bombay, Mumbai, India, proposed a novel structure for TFET, called the sandwich tunnel barrier FET (STBFET), as shown in Fig. 2.13. The proposed STBFET shows a high $I_{ON}$, exceeding 1 mA/$\mu$m at $I_{OFF}$ of 0.1 pA/$\mu$m with a subthreshold swing below 40 mV/dec (Fig. 2.14). The device also showed better static and dynamic performances for sub-1-V operations.

![Figure 2.13](image-url)

**Figure 2.13:** (a) A cross-sectional view of the STBFET with metal gate, high-k gate dielectric, high-k spacer, and a 2-nm-thick Si channel layer. Unlike conventional TFETs, P' source is under the gate below the epitaxial layer, and the drain is on both sides of the spacer. (b) A cross-sectional view of the STBFET at $V_{DS} = V_{GS} = 1$ V is shown along with the electron density and the depletion width (white line). The source doping of $2 \times 10^{20}$ cm$^{-3}$ and drain doping of $5 \times 10^{19}$ cm$^{-3}$ has been used for the simulation.
Figure 2.14: (a) Transfer characteristics of the n-channel STBFET. (Shown on the log scale) $I_{off}$ of the device is 0.1 pA/µm and (also shown on the linear scale) $I_{on} \sim 1.4$ mA/µm. The STBFET shows an SS below 40 mv/dec. (b) The Simulation result shows that the output characteristics of an n-channel STBFET has very good saturation like long-channel MOSFET.

Shih and Chien [78] at the National ChiNan University, Nantou, Taiwan, in 2011, studied a new sub-10-nm TFET with bandgap engineering using a graded Si/Ge heterojunction. Both the height and width of the tunneling barrier were highly controlled by applying gate voltages to ensure a near ideal sub-5-mV/dec switching of sub-10-nm TFETs at 300 K. The breakthrough in subthreshold swing and short-channel effect make the graded Si/Ge TFET highly promising as an ideal green transistor into sub-10-nm regimes. In 2011, Leonelli et al. at the IMEC, Belgium, proposed a novel architecture to boost the performance of TFET, named hybrid TFET [79]. The main feature of this architecture was to boost the vertical tunneling while suppressing the lateral tunneling path responsible for the degradation of the subthreshold swing. The improvement of drive current and subthreshold swing had been achieved by inserting the channel between the source region and the gate dielectric, confining the electric field in the undoped channel region. In 2012, Lattanzio and his group at the EPFL, Switzerland presented another novel device structure for Ge TFET, named electron–hole (EH) bilayer TFET, which exploits carrier tunneling through a bias-induced EH bilayer [80]. The proposed architecture provides a quasi-ideal alignment between the tunneling path and the electric field.
controlled by the gate. This device allows interesting features in terms of low operating voltage (< 0.5 V), due to its supersteep subthreshold slope (SSAVO ~ 13 mV/dec over six decades of current), $I_{ON}/I_{OFF}$ ratio of ~ 10$^9$, and drive current of $I_{ON}$ ~ 10 $\mu$A/µm at $V_{DD}$ = 0.5 V. Low et al. [81] at the National University of Singapore, also investigated a different type of novel TFET with an L-shaped Ge source to produce larger $I_{ON}$, in 2012. The device comprises a Ge source that extends underneath a Si-channel region and separated/isolated from the drain by SiO$_2$ layer. In 2012, Kao et al. reported a promising TFET configuration having a gate on the source only, which simultaneously exhibits a steeper subthreshold slope and a higher ON-current than the lateral tunneling configuration with a gate on the channel [82]. It was shown that the two disadvantages of the structure, namely, the sensitivity to gate alignment and the physical oxide thickness, were mitigated by placing a counter-doped parallel pocket underneath the gate-source overlap. The pocket was found to significantly reduce the field-induced quantum confinement (QC). In 2012, Kim et al. at the Seoul National University, Korea, proposed another L-shaped TFET [83] to get higher Ion and lower SS than conventional TFETs.

2.4.8 Reports on Modeling

In 2008, Shen et al. at the National University of Singapore reported a new approach to treat the 2-D nonlinear Poisson’s equation in the context of MOS devices and discuss its application in the modeling of TFET [84]. It was revealed that the narrowing of tunneling barrier in TFET had different mechanisms before and after inversion layer is formed. Closed-form equation was obtained to describe the barrier narrowing in the presence of inversion layer. Vandenberghe et al. [85] at the IMEC, Belgium, in 2008, showed an approximating analytical description of the TFET potential profile in the channel and in the region under the gate. Using this potential profile and adopting Kane’s Model, they derived an approximating compact formula for the point tunneling (tunneling from the source into the channel in a direction perpendicular to the gate field) current, as a function of the gate-source voltage only, complementing the analytical model for line tunneling (tunneling in a direction that is in line with gate field). In both the case of point and line tunneling, they found that, a
lower bandgap would lead to an improved on-current and a lower onset voltage. In 2009, Mookerjea et al. at the Pennsylvania State University [86] estimated the delay of TFET inverters by CV/I technique through mixed-mode device and circuit simulation. It was shown that unlike MOSFET inverters, the $C_{\text{EFF}}$ can be as high as 2.6 times the gate capacitance in TFET inverters. The real-time drive current trajectory for Si TFET inverters, superimposed on its dc $I_D$-$V_{DS}$ characteristics, was also shown (Fig. 2.15). The transient response of double-gate thin-body-silicon TFET with its MOSFET counterpart has been compared in [87]. It was demonstrated that, due to the presence of source side tunneling barrier, the silicon TFETs exhibited enhanced Miller capacitance, resulting in large voltage overshoot/undershoot in its large-signal switching characteristics, which deteriorated the performance of Si TFETs for digital logic applications.

Kao et al. [88] at the IMEC, Belgium, in 2012, theoretically calculated the parameters $A$ and $B$ of Kane’s direct and indirect BTBT models for different tunneling directions ([100], [110], and [111]) for Si, Ge and unstrained Si$_{1-x}$Ge$_x$. A relatively pronounced
gain in the indirect BTBT generation rate between Si_{0.2}Ge_{0.8} and Ge was attributed to an abrupt decrease in band gap while Si_{1-x}Ge_x transforms from Si- to Ge-like material at x = 0.85. It was indicated that, the direct BTBT contribution in unstrained Si_{1-x}Ge_x would be equally important to the indirect BTBT when the Ge concentration is above 80%. For 100% Ge, the calculation of the BTBT generation rate in the uniform electric field limit revealed that direct tunneling always dominates. Michielis et al. at the EPFL, Switzerland, in 2012, reported that the source and the channel Fermi–Dirac distributions in interband-tunneling-controlled transistors play a fundamental role on the modulation of the injected current [89]. They explained the superlinear onset of the output characteristics based on the occupancy function modulation.

### 2.4.9 Reports on Circuit-Level Analysis

In 2008, Kam et al. at the UC Berkeley [90] reported a circuit-level energy-performance analysis for TFET. TFET was found to be more compelling for sub-100MHz applications. In 2009, Koswatta et al. [91] at the Purdue University presented a detailed performance comparison between conventional n-i-n MOSFET and p-i-n TFET, using semiconducting carbon nanotubes as the model channel material. He found that despite lower on-current, TFETs can switch faster in a range of on/off-current ratios. Switching energy for TFETs was observed to be fundamentally smaller than that for MOSFETs, leading to lower dynamic power dissipation. In 2010, Yang et al. [92] at the National University of Singapore reported a numerical simulation study of gate capacitance components in a TFET, showing key differences in the partitioning of the gate capacitance between the source and drain as compared with a MOSFET. The dependence of gate–drain capacitance $C_{gd}$ on drain design and gate length was also investigated for the reduction in switching delay in TFETs. Pal et al. at the IIT-Bombay, India, in 2011 suggested that, to improve its circuit performance, mere increase in $I_{ON}$ is not sufficient; the output characteristics should also be improved [93]. A design space comprising of $I_{ON}$, a drain saturation voltage, and a drain threshold voltage for minimizing the propagation delay of circuits using TFETs was also presented. Cho et al. [94] at the Stanford University, USA, in 2011, reported an investigation where the small-signal
parameters of gate-all-around tunneling field-effect transistors (GAA TFETs) with different gate lengths were extracted and analyzed in terms of their gate capacitance, source–drain conductance, transconductance, distributed channel resistance, and inversion layer length. Based on understanding these parameters, the high-frequency performances of GAA TFETs were investigated using a TCAD simulation. Lattanzio et al. [95], in 2012, showed that, through appropriate optimization of the Ge electron–hole bilayer TFET (EHBTFET), it would be possible to achieve superior static characteristics at low supply voltages, when compared with a double-gate Ge MOSFET with similar geometry. The dynamic behavior of the devices was also investigated by transient simulations of simple circuits based on complementary inverters. It was reported that, due to the increased total EHBTFET capacitance, the fanout-of-1 delay is larger than that in MOSFET, with 11 ns versus 4 ns at $|V_{DD}| = 0.25$ V.

2.5 Justification of the Work Chosen

Degradation of device performance has been reported in [50] for an n-TFET with a high-k spacer as compared to that without a spacer. It has also been reported in [52] that a low-k ($\text{SiO}_2$) spacer and a high-k ($\text{HfO}_2$) gate dielectric improves the ON-state current of a double-gate TFET by a factor of 3.8 as compared to that when high-k is used both as spacer and gate dielectric. On the other hand, it is reported in [63] that a high-k spacer causes improvement in both the SS and the ON-current for a heterojunction n-TFET with SiGe source which seems to have a conflict with that reported in [50], [52]. It is neither clear why such a conflict arises, nor a detailed study on the effects of the spacer on the device performance of TFET is available in the literature. We have, therefore, chosen to investigate the impact of not only a spacer but also several other structural parameters, such as gate overlap/underlap, varying source doping concentration, etc. on the performance of an n-TFET.

To develop a CTFET, we need both n-TFET and p-TFET. Until now, not much work has been reported on p-TFET, unlike its n-channel counterpart. It is well-known that the fringing field arising out of a high-$\kappa$ dielectric causes fringing-induced barrier lowering (FIBL) [96], which has been reported to improve the device performance of
an n-TFET [50]. However, when used as a spacer for an n-TFET, the same
deteriorates the device performance [50, 52], [97]. Very little is, however, reported in
the literature on the impact of either a high-κ gate dielectric or a spacer on the device
performance of a p-TFET. This motivated us to investigate this effect for p-TFET.
Again, a comprehensive study on the impact of several structural parameters on the
performance of a p-TFET has also been done, similar to that done for an n-TFET.

Until now, the influence of drain potential \(V_{DS}\) on the device operation of a TFET
has not been clearly understood. In the early reported papers, the drain current \(I_D\) does
not show saturation even for higher \(V_{DS}\) [24]. Some reports show good saturation
behavior of \(I_D\) when the same is plotted in the logarithmic scale [98]. Recent works
both experimental [42] and simulation [64, 86] show saturation of \(I_D\) when the same
is plotted in the linear scale. Although some understanding has emerged with regard
to why a TFET can have a good output current saturation, based on both the
analytical modeling [84] and the simulation studies of capacitance–voltage
characteristics [86, 92], it is still not clear why some other reported devices [24, 98]
do not show such saturation. We have, therefore, focused our study on resolving this
issue, as this would be helpful not only for development of analytical model but also
for the designing of a device for mixed-signal applications.

Some earlier studies [99, 100] have reported about the sudden increase in drain
current for higher values of \(V_{DS}\) for graphene nanoribbon FETs. This type of
phenomenon at higher \(V_{DS}\) is obvious for these devices made of zero/small band-gap
material, like graphene (acts as semi-metal) [99, 100]. But these graphene nanoribbon
FETs are not at all suitable for low-power digital applications for their very unsuitable on/off current ratio [99, 100] and are not very good for many analog
applications also as their voltage-gain are very low due to poor output current
saturation (related to small band-gap) [99, 100]. Likewise we should also have a
detailed knowledge of the characteristics, even for a large range of applied voltage,
just like as standard silicon MOSFETs, of TFETs, made of relatively larger band-gap
semiconductor material, like silicon; especially when we are aiming to use them for
the future state-of-art mixed-signal SoC applications. We have, therefore, made an
investigation to understand the entire characteristics for Si-TFET, including at large $V_{DS}$.

Until recently, the CMOS technology has also been very attractive for analog and mixed-signal system-on-chip (SoC) applications where the analog circuits are realized with the digital circuits and memories in the same integrated circuit in order to reduce the cost and to improve the performance. Although TFET shows promise for miniaturized digital and memory applications, finding a compatible device for implementation of the analog elements in a SoC is a big challenge. It is, therefore, important to find out a suitable TFET structure for analog components, which should be realized without much change in the fabrication process used for implementing the digital components. So, we have tried to put some light on the comparison based study of analog performance of DG TFET with that of DG MOSFET.