10-2011

Broken-Gap Tunnel MOSFET: A Constant-Slope Sub-60-mV/decade Transistor

Joshua Thomas Smith  
*Purdue University*, jtsmith@purdue.edu

Saptarshi Das  
*Purdue University*, sdas@purdue.edu

Joerg Appenzeller  
*Purdue University*, appenzeller@purdue.edu

Follow this and additional works at: [http://docs.lib.purdue.edu/nanopub](http://docs.lib.purdue.edu/nanopub)  
Part of the [Electronic Devices and Semiconductor Manufacturing Commons](https://docs.lib.purdue.edu/electronics), and the [Nanotechnology Fabrication Commons](https://docs.lib.purdue.edu/nanofab)

---

[http://dx.doi.org/10.1109/LED.2011.2162220](http://dx.doi.org/10.1109/LED.2011.2162220)

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Abstract—We propose a novel low-power transistor device, called the broken-gap tunnel MOSFET (BG-TMOS), which is capable of achieving constant sub-60-mV/decade inverse subthreshold slopes $S$ at room temperature. Structurally, the device resembles an un gated broken-gap heterostructure Esaki region at the heterojunction is the key to producing a constant $S < 60$ mV/decade, which can be tuned by properly engineering the material composition at this interface. In contrast to the tunneling field-effect transistor, the tunnel junction in the BG-TMOS is independent of the electrostatics in the channel region, enabling the use of 2-D architectures for improved current drive without degradation of $S$—attractive features from a circuit design perspective. Simulations show that the BG-TMOS can exceed MOSFET performance at low supply voltages.

Index Terms—Broken gap, constant slope, heterostructure, low power, steep-slope transistor.

I. INTRODUCTION

ONE of the most pervasive challenges within the semiconductor industry today is dealing with the 60-mV/decade voltage scaling limit inherent in the conventional MOSFET, which has led to increasing power dissipation in integrated circuits. To overcome this challenge, theoretical investigation, together with experimental demonstration of several low-power device alternatives, including the tunneling field-effect transistor (TFET) [1]–[6], impact-ionization MOS [7], suspended-gate FET [8], and ferroelectric FET [9], has been the focus of research efforts worldwide. Among these candidates, the TFET has emerged the most widely studied option for logic applications, primarily due to its ability to operate at low gate ($V_{GS}$) and supply ($V_{DD}$) voltages.

While the TFET has shown promise toward remedying the low-power issue, it has so far experimentally suffered from certain obstacles, such as low ON current ($I_{ON}$) due to limited transmission through a finite tunneling barrier and a subthreshold swing $S < 60$ mV/decade over only a narrow gate voltage range at very small drain currents [1]. Recent reports have provided theoretical insights into the use of staggered and extended valence bands at the heterojunction is the key to producing a constant $S < 60$ mV/decade, which can be tuned by properly engineering the material composition at this interface. In contrast to the tunneling field-effect transistor, the tunnel junction in the BG-TMOS is independent of the electrostatics in the channel region, enabling the use of 2-D architectures for improved current drive without degradation of $S$—attractive features from a circuit design perspective. Simulations show that the BG-TMOS can exceed MOSFET performance at low supply voltages.

II. DEVICE STRUCTURE AND PRINCIPLE OF OPERATION

Resembling an Esaki diode in series with a conventional MOSFET, an n-mode representation of the proposed BG-TMOS structure is shown in Fig. 1(a). The $p^+/n^+$ tunnel junction on the source side consists of a type-III, or broken-gap, semiconductor heterojunction that screens the “hot” carriers from the Fermi distribution in the $p^+$ source and permits a tunneling probability near unity [see Fig. 1(b)]—an arrangement broken-gap TFETs (BG-TFETs) to improve $I_{ON}$ and realize $S < 60$ mV/decade [10], [11]; however, just as in the homojunction TFET case, the gate-controlled tunneling mechanism in the BG-TFET yields $S$ values that are dependent on the actual surface potential in the channel region [3]. The required band control in turn favors 1-D channel geometries over a 2-D layout to avoid nonlinearities in the output characteristics and drain-voltage-dependent threshold voltages [6], which are undesirable from a circuit design prospective [12].

In this letter, we propose a broken-gap tunnel MOSFET (BG-TMOS) device with source injection that is independent of the electrostatics in the channel, yielding a constant yet tunable $S < 60$ mV/decade. BG-TMOS operation permits 2-D architectures without the previously discussed disadvantages.

Fig. 1. (a) Planar n-mode version of the proposed BG-TMOS device with (b) a band diagram projection. Proper device operation requires the $n^+$ region at the tunnel junction to be ballistic and not fully depleted to achieve source-controlled injection. (c) Effective temperature $T_e$ of the carriers injected into the $n^+$ source decreases for smaller $\Delta E_{C-\nu}$ windows through increased screening of the high-energy tails of the Fermi distribution $f(E)$, assuming that $E_f^* \approx E_{C-\nu}^*$ lies in the middle of the $\Delta E_{C-\nu}$ window.
made possible through the use of III–V material systems, such as p-GaSb/n-In_xGa_1−xAs [10]. A p-mode version of this structure can be envisioned by altering the III–V materials used, e.g., n-InAs/p-Al_xGa_1−xSb. In both cases, the band lineup at the heterojunction can be tuned by changing the x content and doping concentration in the source regions, which alters the energy spacing of the tunneling window (ΔE_c−v), as well as the position of the source Fermi level E_F^s. In this way, the effective temperature (T_e) of the system decreases with ΔE_c−v, as shown in Fig. 1(c). The transmission of this “cold” carrier distribution into the source is then regulated by the gate-controlled barrier in the channel.

The key to realizing a constant S < 60 mV/decade is maintaining ΔE_c−v at a fixed value when switching between the ON and OFF states—a feat that requires careful engineering of the p^+/n+ junction. Importantly, the n^+ region must simultaneously remain ballistic and not fully depleted during operation. This situation is possible for the high-mobility III–V materials typically considered for this type of heterostructure, e.g., In_xGa_1−xAs [13]. Even in the presence of high doping, scattering mechanisms such as phonon, alloy, and surface roughness scattering tend to dominate [14], [15], yet ballisticity can still prevail over a length scale of tens of nanometers [15]. As shown in Fig. 1(b), the ballistic length scale l_bal requirement of the n^+ region is necessary to prevent rethermalization of “cold” carriers injected from the p^+ source. The not fully depleted constraint is a prerequisite to achieving source-controlled rather than gate-controlled band-to-band tunneling, i.e., where ΔE_c−v is not influenced by the gate and S therefore does not vary with the surface potential in contrast to previous work [4]. The MOSFET-type drain configuration prevents the “slow turn-on” feature—the nonlinearity in MDS—the ratio of (2) over the same expression but integrated from zero to infinity for the smaller number of carriers injected from the source compared to a conventional MOSFET. We assume that current is uniform along the width direction, meaning that the consideration of multiple modes will simply scale the currents in both devices. We emphasize that the simulation and foregoing analysis provide a first level approximation of the expected BG-TMOS behavior since we do not consider effects such as phonon scattering, gate tunneling currents, or standing wave effects due to interference, which may degrade S.

Simulated characteristics for the BG-TMOS and the MOSFET yield similar features with some distinct differences, as shown in Fig. 2. Fig. 2(a) shows the transfer characteristics for the BG-TMOS and the MOSFET, along with the proposed BG-TMOS operation to a similarly scaled MOSFET. The constant ΔE_c−v in the BG-TMOS achieves MOSFET-like transfer characteristics but with smaller inverse subthreshold slopes. The table inset of (a) shows the ΔE_c−v values that yield the constant S curves shown.

### III. Device Simulation and Discussion

Our simulation of the BG-TMOS device considers the Landauer expression for a 1-D ballistic device with one contributing mode [16], given by (1) with T(E) assumed to be unity due to the broken gap and assuming one-to-one band movement in the device’s ON state

\[
I_D = \frac{2q}{h} \cdot \Phi_0 \cdot \int_0^{\frac{\Phi_0}{k \cdot \ln 2}} dE \cdot T(E) \left( f_s(E, T_e) - f_d(E, T_e) \right) \tag{1}
\]

\[
T_e = \frac{1}{k \cdot \ln 2} \cdot \int_0^{\Delta E_c-v/2} \frac{dE}{1 + \exp(E/kT)}. \tag{2}
\]

In (1), \(\Phi_0\) represents the maximum surface potential in the channel and \(f_{s,d}(E, T_e)\) denotes the source and drain Fermi distributions. Note that \(f_{s,d}(E, T_e)\) depends on \(T_e\), which is determined by \(\Delta E_{c-v}\), as shown in (2). The screening factor \(\gamma\) in (1) is the ratio of (2) over the same expression but integrated from zero to infinity for the smaller number of carriers injected from the source compared to a conventional MOSFET. We assume that current is uniform along the width direction, meaning that the consideration of multiple modes will simply scale the currents in both devices. We emphasize that the simulation and foregoing analysis provide a first level approximation of the expected BG-TMOS behavior since we do not consider effects such as phonon scattering, gate tunneling currents, or standing wave effects due to interference, which may degrade S.

Simulated characteristics for the BG-TMOS and the MOSFET yield similar features with some distinct differences, as shown in Fig. 2. Fig. 2(a) shows the transfer characteristics for a conventional MOSFET, along with the proposed BG-TMOS device for several different \(\Delta E_{c-v}\) windows, corresponding to the S values shown. The MOSFET-like constant S of the BG-TMOS is a direct result of the constant \(T_e < 300 K\) during operation. Importantly, the BG-TMOS has the advantage of attaining arbitrarily small and specific values of S < 60 mV/decade. The primary tradeoff is a decrease in \(I_{ON}\) for decreasing S, since the number of carriers screened from \(f_s(E)\) increases as \(\Delta E_{c-v}\) is made smaller, i.e., \(\gamma\) decreases. The output characteristics given in Fig. 2(b) for the MOSFET with S = 60 mV/decade and a BG-TMOS with S = 40 mV/decade help to illustrate this point.

Fig. 3(a) shows further insights into the S dependence of \(I_{ON}\) in the BG-TMOS for different \(V_{DD}\) values. Each of these curves was generated by assuming a fixed \(V_{DD}\) (with \(V_{DD} = V_{DS}\)) and \(I_{ON}/I_{OFF} = 10^5\), as is often required for digital applications.
for sufficiently small $V_{\text{DD}}$ that the BG-TMOS structure has a decided advantage over a given $V_{\text{DD}}$ requirement by appropriately tuning $\Delta E_{\text{g}}$ to achieve a desired $S$. The peaks in Fig. 3(c), or $(1/\text{EDP})_{\text{max}}$ points, are roughly described by a $1/V_{\text{DD}}$ relationship, as shown in the inset.

IV. CONCLUSION

The BG-TMOS has been suggested as a novel device alternative for overcoming the $V_{\text{DD}}$ scaling issue. Further exploration of appropriate III–V material systems and device fabrication are needed to add experimental insight.

REFERENCES