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

Smith, Joshua Thomas; Das, Saptarshi; and Appenzeller, Joerg, "Broken-Gap Tunnel MOSFET: A Constant-Slope Sub-60-mV/decade Transistor" (2011). *Birck and NCN Publications.* Paper 781.
http://dx.doi.org/10.1109/LED.2011.2162220

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

Joshua T. Smith, Saptarshi Das, and Joerg Appenzeller, *Fellow, IEEE*

***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 ungated broken-gap heterostructure Esaki region in series with a conventional MOSFET. The gate voltage independence of the energy spacing between the conduction and 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.**

***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 ($V_{\rm GS}$) and supply ($V_{\rm DD}$) 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 ($I_{\rm ON}$) 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 broken-gap TFETs (BG-TFETs) to improve $I_{\rm ON}$ 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.

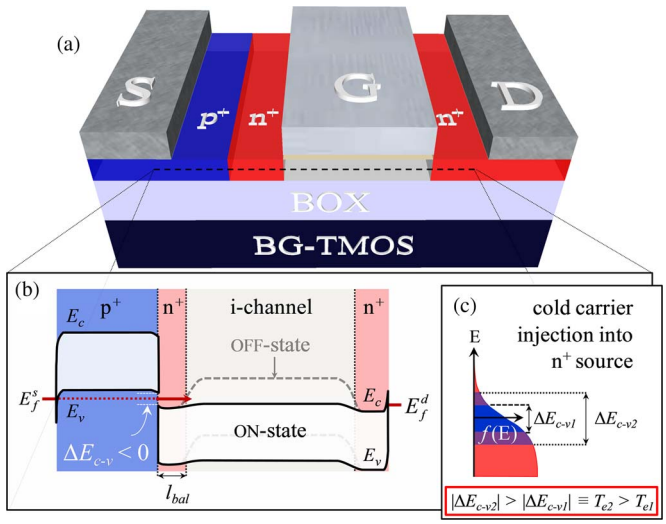

Fig. 1. (a) Planar n-mode version of the proposed BG-TMOS device with (b) a band diagram projection. Proper device operation requires the n^+ region at the tunnel junction to be ballistic and *not* fully depleted to achieve *source-controlled* injection. (c) Effective temperature T_e of the carriers 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^s lies in the middle of the $\Delta E_{c-\nu}$ window.

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

Manuscript received June 8, 2011; revised June 28, 2011; accepted July 11, 2011. Date of publication August 7, 2011; date of current version September 28, 2011. This work was supported in part by the Center for Functional Engineered Nano Architectonics under Grant 104295. The review of this letter was arranged by Editor M. Passlack.

The authors are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: jtsmith@purdue.edu; sdas@purdue.edu; appenzeller@purdue.edu).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LED.2011.2162220

made possible through the use of III–V material systems, such as p-GaSb/n-$In_xGa_{1-x}As$ [10]. A p-mode version of this structure can be envisioned by altering the III–V materials used, e.g., n-InAs/p-$Al_xGa_{1-x}Sb$. 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 ($\Delta E_{c-\nu}$), 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 $\Delta E_{c-\nu}$, 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 $\Delta E_{c-\nu}$ 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., $In_xGa_{1-x}As$ [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 $l_{\rm bal}$ 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 $\Delta E_{c-\nu}$ 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 I_D–$V_{\rm DS}$ for small $V_{\rm DS}$—without a 1-D channel requirement as in the TFET case [6].

Although Fig. 1(a) shows a planar BG-TMOS for illustrative purposes, other structures, such as a vertical double-gate arrangement, may provide a more practical route toward fabrication, wherein epitaxial measures could be used to form the junctions with gate alignment being the primary challenge.

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

$$I_D = \frac{2q}{h} \cdot \gamma \cdot \int_0^{\Phi_f^0} dE \cdot T(E) \left(f_s(E, T_e) - f_d(E, T_e) \right) \qquad (1)$$

$$T_e = \frac{1}{k \cdot \ln 2} \int_0^{\Delta E_{c-\nu}/2} \frac{dE}{1 + \exp(E/kT)}. \qquad (2)$$

In (1), Φ_f^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 defined by $\Delta E_{c-\nu}$, as shown in (2). The screening factor γ in (1) is the ratio of (2) over the same expression but integrated from zero to ∞ to account 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.

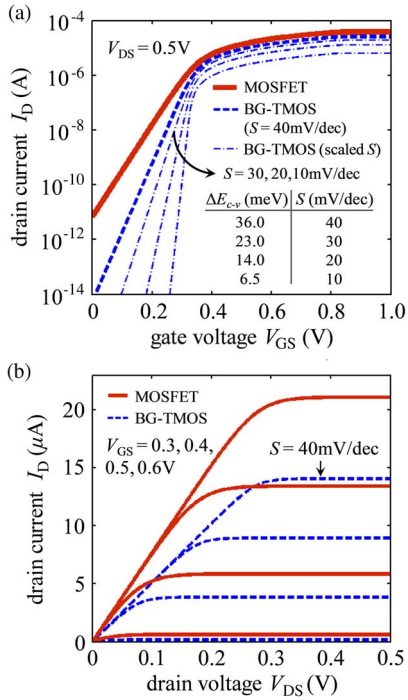

Fig. 2. Simulated (a) transfer and (b) output characteristics comparing BG-TMOS operation to a similarly scaled MOSFET. The constant $\Delta E_{c-\nu}$ in the BG-TMOS achieves MOSFET-like transfer characteristics but with smaller inverse subthreshold slopes. The table inset of (a) shows the $\Delta E_{c-\nu}$ values that yield the constant S curves shown.

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 $\Delta E_{c-\nu}$ 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 $I_{\rm ON}$ for decreasing S, since the number of carriers screened from $f_s(E)$ increases as $\Delta E_{c-\nu}$ is made smaller, i.e., γ 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 $I_{\rm ON}$ in the BG-TMOS for different $V_{\rm DD}$ values. Each of these curves was generated by assuming a fixed $V_{\rm DD}$ (with $V_{\rm DD} = V_{\rm DS}$) and $I_{\rm ON}/I_{\rm OFF} = 10^5$, as is often required for digital applications.

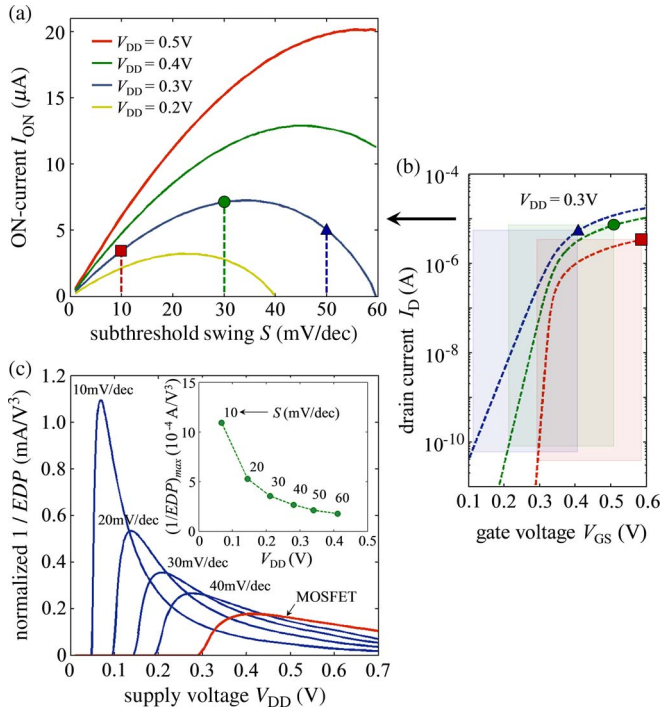

Fig. 3. (a) $I_{\rm ON}$ versus S for a BG-TMOS with $S < 60$ mV/decade for $I_{\rm ON}/I_{\rm OFF} = 10^5$ and $V_{\rm DD} = 0.2, 0.3, \ldots, 0.5$ V. (b) Extraction process for $I_{\rm ON}$ in (a) and (c). (c) Normalized $1/EDP$ versus $V_{\rm DD}$ for the (red) MOSFET and (blue) BG-TMOS with different constant S values. (Inset) $1/EDP$ maxima for different inverse subthreshold slopes.

As shown in Fig. 3(b), $I_{\rm ON}$ was extracted as the maximum I_D for a given $V_{\rm DD}$ window under these conditions. A noteworthy outcome of Fig. 3(a) is that $I_{\rm ON}$ does not drop as substantially as might be expected at very low S. This result is a direct consequence of the fact that a sharper voltage swing for a given $V_{\rm DD}$ allows access to overdrive currents deeper into the ON state [see Fig. 3(b)]. This insight helps to understand the nonmonotonic dependence of $I_{\rm ON}$ on S observed in Fig. 3(a). Furthermore, this demonstrates that a "sweet spot" exists for S, where $I_{\rm ON}$ is maximized at a given supply voltage.

An appropriate benchmark for comparing the BG-TMOS with the MOSFET for low-power applications is the energy-delay product (EDP), which captures the tradeoff between energy dissipation and performance with the expression

$$EDP = E \cdot \tau = \left(C_{\rm par} \cdot V_{\rm DD}^2\right) \cdot \left(\frac{C_{\rm par} \cdot V_{\rm DD}}{I_{\rm ON}}\right) = \frac{C_{\rm par}^2 \cdot V_{\rm DD}^3}{I_{\rm ON}} \tag{3}$$

where $C_{\rm par}$ is the parasitic capacitance. Assuming that $C_{\rm par}$ is dominated by external circuit capacitances such as interconnect capacitance, i.e., that $C_{\rm par}$ is the same for the BG- TMOS and MOSFET, Fig. 3(c) shows the $1/EDP$ normalized to $C_{\rm par}$ using the same $V_{\rm DD}$ and $I_{\rm ON}/I_{\rm OFF}$ criteria used to extract $I_{\rm ON}$ in Fig. 3(b). For $S < 60$ mV/decade, this plot indicates that the BG-TMOS structure has a decided advantage over the MOSFET at smaller supply voltages (a larger $1/EDP$ is preferred). At sufficiently small $V_{\rm DD}$ for a chosen S value, the $I_{\rm ON}$ term in (3) begins to dominate, and a peak in the $1/EDP$ curve appears until, ultimately, S can no longer meet the $I_{\rm ON}/I_{\rm OFF} = 10^5$ requirement as $V_{\rm DD}$ is scaled further and $1/EDP$ goes to zero. Of great importance for the proposed BG-TMOS device is the fact that *the EDP can be optimized for a given $V_{\rm DD}$ requirement by appropriately tuning $\Delta E_{c-\nu}$ to achieve a desired S*. The peaks in Fig. 3(c), or $(1/EDP)_{\rm max}$ points, are roughly described by a $1/V_{\rm DD}$ relationship, as shown in the inset.

IV. Conclusion

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

References

[1] J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris, "Band-to-band tunneling in carbon nanotube field-effect transistors," *Phys. Rev. Lett.*, vol. 93, no. 19, pp. 196 805-1–196 805-4, Nov. 2004.

[2] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-gate strained-Ge heterostructure tunneling FET (TFET) with record high drive currents and $<$ 60 mV/dec subthreshold slope," in *IEDM Tech. Dig.*, 2008, pp. 947–949.

[3] J. Knoch, S. Mantl, and J. Appenzeller, "Impact of the dimensionality on the performance of tunneling FETs: Bulk versus one-dimensional devices," *Solid State Electron.*, vol. 51, no. 4, pp. 572–578, Apr. 2007.

[4] V. Nagavarapu, R. Jhaveri, and J. C. S. Woo, "The tunnel source (PNPN) n-MOSFET: A novel high performance transistor," *IEEE Trans. Electron Devices*, vol. 55, no. 4, pp. 1013–1019, Apr. 2008.

[5] J. T. Smith, C. Sandow, S. Das, R. A. Minamisawa, S. Mantl, and J. Appenzeller, "Silicon nanowire tunneling field-effect transistor arrays: Improving subthreshold performance using excimer laser annealing," *IEEE Trans. Electron Devices*, vol. 58, no. 7, pp. 1822–1829, Jul. 2011.

[6] J. Knoch and J. Appenzeller, "Tunneling phenomena in carbon nanotube field-effect transistors," *Phys. Stat. Sol. (A)*, vol. 205, no. 4, pp. 679–694, Apr. 2008.

[7] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, "I-MOS: A novel semiconductor device with a subthreshold slope lower than kT/q," in *IEDM Tech. Dig.*, 2002, pp. 289–292.

[8] N. Abelé, R. Fritschi, K. Boucart, F. Casset, P. Ancey, and A. M. Ionescu, "Suspended-gate MOSFET: Bringing new MEMS functionality into solid-state MOS transistor," in *IEDM Tech. Dig.*, 2005, pp. 1075–1077.

[9] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano Lett.*, vol. 8, no. 2, pp. 405–410, Feb. 2008.

[10] J. Knoch and J. Appenzeller, "Modeling of high-performance p-type III–V heterojunction tunnel FETs," *IEEE Electron Device Lett.*, vol. 31, no. 4, pp. 305–307, Apr. 2010.

[11] S. O. Koswatta, S. J. Koester, and W. Haensch, "On the possibility of obtaining MOSFET-like performance and sub-60-mV/dec swing in 1-D broken-gap tunnel transistors," *IEEE Trans. Electron Devices*, vol. 57, no. 12, pp. 3222–3230, Dec. 2010.

[12] A. Pal, A. B. Sachid, H. Gossner, and V. R. Rao, "Insights into the design and optimization of tunnel-FET devices and circuits," *IEEE Trans. Electron Devices*, vol. 58, no. 4, pp. 1045–1053, Apr. 2011.

[13] J. Wang and M. Lundstrom, "Ballistic transport in high electron mobility transistors," *IEEE Trans. Electron Devices*, vol. 50, no. 7, pp. 1604–1609, Jul. 2003.

[14] M. Yokoyama, R. Iida, S. H. Kim, N. Taoka, Y. Urabe, T. Yasuda, H. Takagi, H. Yamada, N. Fukuhara, M. Hata, M. Sugiyama, Y. Nakano, M. Takenaka, and S. Takagi, "Extremely-thin-body InGaAs-on-insulator MOSFETs on Si fabricated by direct wafer bonding," in *IEDM Tech. Dig.*, 2010, pp. 46–49.

[15] S. R. Mehrotra, A. Paul, M. Luisier, and G. Klimech, "Atomistic alloy scattering theory in InGaAs: Bulk to ultra-thin-body transistors," unpublished.

[16] S. Datta, *Electronic Transport in Mesoscopic Systems*. Cambridge, U.K.: Cambridge Univ. Press, 1997.