ARTICLE IN PRESS

Microelectronics Journal ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Microelectronics Journal

journal homepage: www.elsevier.com/locate/mejo

Scattering-limited and ballistic transport in a nano-CMOS circuit

Ismail Saad [a,b,*], Michael L.P. Tan [b,c], Aaron C.E. Lee [b], Razali Ismail [b], Vijay K. Arora [b,d]

[a] School of Engineering & IT, Universiti Malaysia Sabah, 88999 Kota Kinabalu, Sabah, Malaysia
[b] Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia
[c] Electrical Engineering Division, Engineering Department, University of Cambridge, 9 J.J. Thomson Avenue, Cambridge CB3 0FA, UK
[d] Department of Electrical and Computer Engineering, Wilkes University, Wilkes-Barre, PA 18766, USA

ARTICLE INFO

Keywords:
Ballistic mobility
Nano MOSFET
Saturation velocity
Quantum confinement

ABSTRACT

The mobility and saturation velocity in the nanoscale metal oxide semiconductor field effect transistor (MOSFET) are revealed to be ballistic; the former in a channel whose length is smaller than the scattering-limited mean free path. The drain-end carrier velocity is smaller than the ultimate saturation velocity due to the presence of a finite electric field at the drain. The current–voltage characteristics of a MOSFET are obtained and shown to agree well with the experimental observations on an 80 nm channel. When scaling complementary pair of NMOS and PMOS channels, it is shown that the length of the channel is proportional to the channel mobility. On the other hand, the width of the channel is scaled inversely proportional to the saturation velocity of the channel. The results reported may transform the way the ULSI circuits are designed and their performance evaluated.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The metal oxide semiconductor field effect transistor (MOSFET) is a vehicle for the design of integrated circuits both for digital and analog applications. In a nanoscale channel, the ballistic mobility and saturation velocity play an important role in limiting the MOSFET performance. There is ongoing debate over the interdependence of mobility that is controlled by momentum-randomizing scattering events and saturation velocity that depends on the streamlined motion of electrons [1–5]. Most investigators tend to converge on the theme that the saturation velocity does not sensitively depend on the low-field mobility. However, there does not appear to be a consensus on the factors limiting the current and velocity. In the work reported, the mobility degradation due to quantum confinement effect is described and utilized in evaluating nanoscale MOSFET performance. Consequently, the ultimate saturation velocity due to intense longitudinal electric field is shown to be ballistic with magnitude comparable to the Fermi velocity in strongly inverted channel. This ballistic paradigm, both for the ohmic mobility and drain velocity limited by the ultimate intrinsic velocity is utilized in modeling current–voltage (I–V) characteristics. The theory replicates well all the observed features in experiments on an 80 nm channel. When applied to complementary pair of NMOS and PMOS, new design rule is proposed.

2. Quantum confinement

The channel in a nanoscale MOSFET is indeed a quantum one that is constrained by the gate electric field forming an approximately linear quantum well [4]. Contrary to the belief of many naïve researchers, no heating of the electrons by the gate is possible. Rather, the gate confines the electron to length z_{QM} comparable to the de Broglie wavelength. The average distance of an electron from the interface ($z = 0$) is

$$z_{QM} = \frac{2}{3}\frac{(2.338E_0)}{q\varepsilon_t} \tag{1}$$

with

$$E_0 = \left(\frac{\hbar^2 q^2 \varepsilon_t^2}{2m_3^*}\right)^{1/3} \quad \text{and} \quad \varepsilon_t \approx \frac{V_{GS} + V_T}{6t_{ox}}, \tag{2}$$

where $2.338 \times E_0$ is the ground state energy, ε_t is the transverse electric field in strong inversion regime. 2.338 is the first zero of the Airy function appropriate for a triangular quantum well, V_T is the threshold voltage, t_{ox} the thickness of the gate oxide and $m_3^* = m_l = 0.916m_o$ is the longitudinal effective mass. The electrons do not reside at the Si/SiO_2 interface as wave-function vanishes there due to the quantum-confinement effect. In the nano-MOSFET, the gate oxide is a few nm ($t_{ox} = 1.59$ nm in our case). The distance

* Corresponding author at: School of Engineering & IT, Universiti Malaysia Sabah, 88999 Kota Kinabalu, Sabah, Malaysia.
E-mail addresses: ismail_saad03@yahoo.com, ismail_s@ums.edu.my (I. Saad).

0026-2692/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mejo.2008.06.049

ARTICLE IN PRESS

z_{QM} of the electrons that is also a few nm cannot be neglected. The gate capacitance C_G per unit area is smaller than C_{ox} and is given by

$$C_G = \frac{\varepsilon_{ox}}{t_{oxeff}} = \frac{C_{ox}}{1 + (1/3)(z_{QM}/t_{ox})}. \tag{3}$$

3. Ballistic saturation velocity

The carrier drift due to the channel electric field driving the electrons has been explored in a number of works (see Ref. [1] for a review and Ref. [6] for sample experimental results). In the archival work of Arora [1], the ballistic transport, although not specifically mentioned by that name, was predicted in the presence of a high electric field. The theory developed was for non-degenerate bulk semiconductors that gave saturation velocity comparable to the thermal velocity. Recently, the theory has been extended to embrace all dimensions under both degenerate and non-degenerate conditions [5]. The ballistic nature of the velocity is apparent from Fig. 1 that streamlines the randomly oriented intrinsic velocity $\vec{v}_i$ vectors in a very high electric field.

The saturation velocity is the weighted average of the magnitude of the carrier velocity $|v|$ with averaging weight equal to the probability given by the Fermi–Dirac distribution multiplied by the two-dimensional density of states. This results in saturation velocity being intrinsic velocity v_{i2} for a two-dimensional distribution given by

$$v_{i2} = v_{th2}\frac{\Im_{1/2}(\eta_{F2})}{\Im_0(\eta_{F2})} \tag{4}$$

with

$$v_{th2} = \sqrt{\frac{\pi k_B T}{2m_t^*}}, \tag{5}$$

where $\Im_j(\eta_F)$ is the Fermi integral of order j with $\eta_F = (E_F - E_c)/k_BT$ as the normalized Fermi energy with respect to the quantized band-edge $\varepsilon_c = \varepsilon_{co} + 2.338 \times E_0$, ε_{co} is the bulk conduction band-edge that is lifted by the zero-point energy $2.338 \times E_0$ due to the quantum confinement (QC) effect, T is the temperature and m_t^* is the transverse effective mass in ellipsoidal band structure of silicon. In (1 0 0) configuration, only the lower two valleys with conductivity effective mass are occupied in the quantum limit. Eq. (4) shows that the ultimate intrinsic saturation velocity v_{i2} in a semiconductor is the two-dimensional thermal velocity v_{th2} only in a non-degenerate regime. In a degenerately induced MOSFET, the intrinsic saturation velocity is limited to the Fermi velocity v_{F2} that in turn depends on surface carrier concentration n_s:

$$v_{i2Deg} = \frac{2}{3}v_{F2}. \tag{6}$$

The numerical work of Eq. (4) indicates that the electrons make a transition from non-degenerate to degenerate statistics as gate voltage is increased.

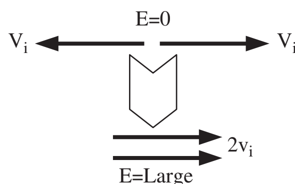

Fig. 1. Random intrinsic velocity vectors $\vec{v}_i$ in a zero-field transforming to streamlined vectors in a very high electric field.

4. *I–V* characteristics of a nano-MOSFET

The current–voltage (*I–V*) characteristics of a nano-MOSFET can be obtained from the velocity-field characteristics as follows:

$$I_D = \frac{C_G\mu_{lf}W/L[V_{GT} - (1/2)V_D]V_D}{1 + (V_D/V_c)}, \quad 0 \leqslant V_D \leqslant V_{Dsat}, \tag{7}$$

where μ_{lf} is the low-field mobility, W the channel width, L the channel length and $V_{GT} = V_{GS} - V_T$. The drain voltage at the onset of saturation V_{Dsat} is less than V_{Dsat1}, appropriate for carrier reaching full saturation ($\alpha = 1$ below). V_{Dsat1} is given by

$$V_{Dsat1} = V_c(\sqrt{1 + 2V_{GT}/V_c} - 1), \quad V_c = \frac{v_{sat}}{\mu_o}L \tag{8}$$

The drain-to-saturation velocity ratio $\alpha = v_D/v_{sat}$ is plotted in Fig. 2 as a function of drain voltage V_D starting from the saturation point V_{Dsat} to $V_D = 1.0$ V. As the drain electric field becomes larger due to increased drain voltage, α tends to advance towards its ultimate value of 1. In the scenario with $\alpha = 1$, the drain velocity is the ultimate saturation velocity. However, when $\alpha < 1$, the drain velocity v_D is smaller than v_{sat} due to the presence of the finite electric field at the drain [5]. The curve at $V_{GS} = 0.7$ V (upper one in Fig. 2) is representative of a long-channel behavior as the ratio V_{GS}/V_c is smaller for a long channel. That is the reason that $\alpha = 1$ model works well for a long channel and current in the saturation region has a flat slope, making channel conduction equal to zero. However, in a nanoscale channel, as seen in the *I–V* characteristics of Fig. 3 for a 80 nm channel, the channel conductance is non-zero due to the increase in the value of α as the drain voltage is increased. The current at the onset of saturation region is

$$I_{Dsat} = \alpha C_G(V_{GS} - V_T - V_{Dsat})v_{sat}W, \quad V_D \geqslant V_{Dsat}. \tag{9}$$

I_{Dsat} depends only on the gate width and not sensitively on the channel length. Hence, channel length modulation due to pinchoff point moving in the channel is a mistaken identity that does not exist either for the long or short channel. In fact, if the binomial expansion of V_{Dsat1} is performed under the long-channel approximation, the carrier concentration at the drain is always finite [5,8].

The *I–V* characteristics for an 80 nm channel of Fig. 3 with $\alpha = 1$ (solid lines) terminate at the velocity-saturation limited current giving zero transconductance. However, when $\alpha < 1$, solid lines branch out at $V_D = V_{Dsat}$ to dotted line with α changing from

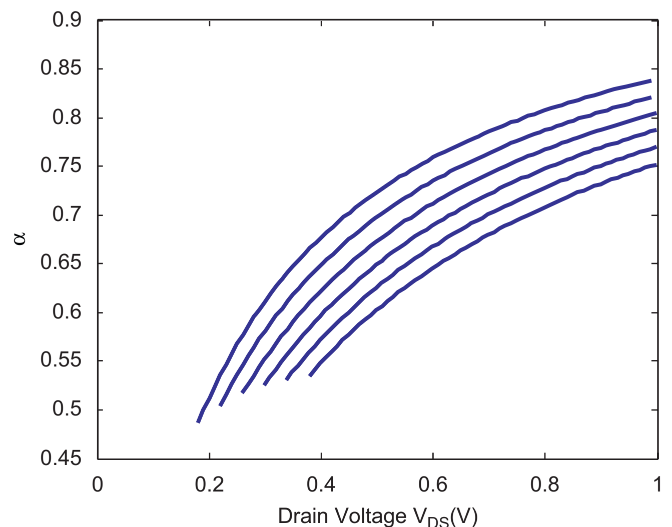

Fig. 2. The ratio $\alpha = v_D/v_{sat}$ of the drain velocity as a function of drain voltage beyond the onset of saturation $V_D > V_{Dsat}$ for $V_{GS} = 0.7$ (topmost curve), 0.8, 0.9, 1.0, 1.1 and 1.2 (bottom curve).

Please cite this article as: I. Saad, et al., Scattering-limited and ballistic transport in a nano-CMOS circuit, Microelectron. J (2008), doi:10.1016/j.mejo.2008.06.049

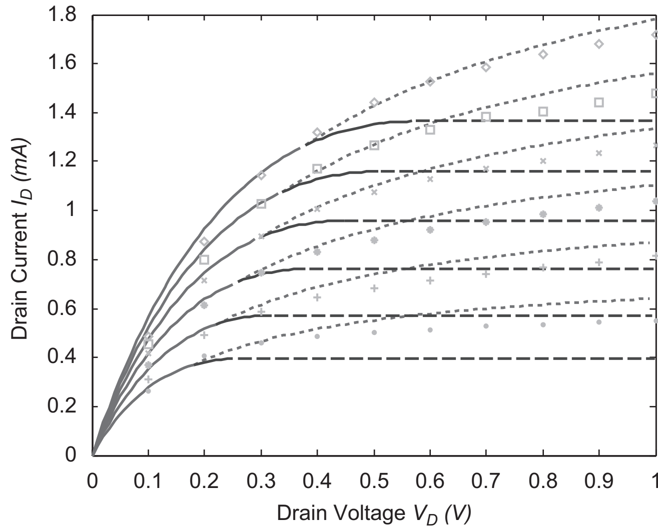

Fig. 3. *I–V* characteristics of an 80 nm MOSFET for gate voltage $V_{GS} = 0.7$, 0.8, 0.9, 1.0, 1.1 and 1.2. Solid lines are from Eq. (7) in the range $0 \leqslant V_D \leqslant V_{Dsat}$. The dotted lines are from Eq. (9) and are extensions for $V_D \geqslant V_{Dsat}$. The dashed lines are for $\alpha = 1$.

its value as described in Fig. 2. The experimental results [7] shown by various markers agree well with the finite electric field model. V_{Dsat} is always smaller but approaches V_{Dsat1} as channel length is increased (not shown). This is indicative of the fact that even for a long-channel, the carrier drift at the drain end is limited by the saturation velocity, not by the Ohmic transport. The current at the source is still ohmic. That is in direct contrast to the consensus of many investigators that are of the opinion that the current in the channel is limited by the source electric field.

5. Nano-CMOS design

In a CMOS circuit design, switching speed is enhanced when the current $I_{Dn(p)}$ in NMOS and PMOS transistors are the same as given in Eq. (7) with

$$V_{GTn(p)} = V_{GSn(p)} - V_{Tn(p)} \tag{10}$$

$$V_{cn(p)} = \frac{v_{satn(p)}}{\mu_{on(p)}} L_{n(p).} \tag{11}$$

In the fabrication process, the oxide thickness t_{ox} can be made the same for both transistors and threshold voltage can be adjusted so that $V_{Tn} = V_{Tp}$. As length is the smallest dimension in the photolithographic process, the designers tend not to sacrifice the smallest dimensions attained and hence keep the channel length same for both channels. In a traditional design based on Ohm's law ($V_{DS} \ll V_c$), the width is scaled so $(W\mu_{lf})_{n(p)}$ product is same for current to be equal in both channels. In order to account for the lower mobility of PMOS, the width ratio of p-channel to n-channel is scaled as

$$\frac{W_p}{W_n} = \frac{\mu_{lfn}}{\mu_{lfp}}. \tag{12}$$

The above design is based on the validity of Ohm's law that obviously does not hold good for nano-scale channels. When velocity saturation is considered, the saturation current of Eq. (9) requires (Wv_{sat}) to be matched for both n- and p-channels. In the triode region $0 \leqslant V_D \leqslant V_{Dsat}$, the current expression of Eq. (7), including V_c, must also match to complement *I–V* characteristics of both channels. This leads to

$$\frac{W_p}{W_n} = \frac{v_{satn}}{v_{satp}}, \quad \frac{L_p}{L_n} = \frac{\mu_{lfp}}{\mu_{lfn}}. \tag{13}$$

The length of the channel is scaled proportionally to the channel mobility and the width of the channel is scaled inversely proportional to the saturation velocity of the channel.

6. Conclusions

In a nanoscale channel, it is highly likely that the mobility and saturation velocity are scattering-independent and hence ballistic. The pseudo channel length modulation in the absence of quasi-pinchoff and finite output conductance is attributed to the drain velocity increasing towards the ultimate saturation velocity as electric field increase due to increase in the drain voltage. New CMOS scaling law that scales width with the saturation velocity and the length with the low-field Ohmic mobility is proposed.

Acknowledgments

The authors would like to acknowledge the financial support from E-Science grant of the Ministry of Higher Education (MOHE). I.S. and V.K.A. are thankful to the Universiti Teknologi Malaysia (UTM) for wonderful hospitality. V.K.A. acknowledges the UTM grant 77506 as a distinguished visiting professor. M.L.P.T. thanks the University of Cambridge for access to its facilities in completion of his PhD work.

References

[1] V.K. Arora, High-field distribution and mobility in semiconductors, Jpn. J. Appl. Phys. 24 (1985) 537–545.
[2] F.F. Fang, A.B. Fowler, Hot electron effects and saturation velocities in silicon inversion layers, J. Appl. Phys. 41 (4) (1970) 1825–1831.
[3] A.M.T. Fairus, V.K. Arora, Quantum engineering of nanoelectronic devices: the role of quantum confinement on mobility degradation, Microelectron. J. 32 (2000) 679–686.
[4] V.K. Arora, Quantum engineering of nanoelectronic devices: the role of quantum emission in limiting drift velocity and diffusion coefficient, Microelectron. J. 31 (11/12) (2000) 853–859.
[5] V.K. Arora, M.L.P. Tan, I. Saad, R. Ismail, Ballistic quantum transport in a nanoscale metal-oxide-semiconductor field effect transistor, Appl. Phys. Lett. 91 (2007) 103510.
[6] D.R. Greenberg, J.A. del Alamo, Velocity saturation in the extrinsic device: a fundamental limit in HFET's, IEEE Trans. Electron. Devices 41 (8) (1994) 1334–1339.
[7] M.L.P. Tan, R. Ismail, Modeling of nanoscale MOSFET performance in the velocity saturation region, Jurnal Elektrika 9 (1) (2007) 37–41.
[8] V.K. Arora, M.B. Das, Effect of electric-field-induced mobility degradation on the velocity distribution in a sub- mu m length channel of InGaAs/AlGaAs heterojunction MODFET, Semicond. Sci. Technol. 5 (9) (1990) 967–973.

Please cite this article as: I. Saad, et al., Scattering-limited and ballistic transport in a nano-CMOS circuit, Microelectron. J (2008), doi:10.1016/j.mejo.2008.06.049