Computational Analysis of Ballistic Saturation Velocity in Low-Dimensional Nano-MOSFET

Ismail Saad, Khairul A. M., Nurmin Bolong, Abu Bakar A.R. and Vijay K. Arora

1Nanoelectronics Device & Material (NiCER) Research Group, School of Engineering & IT, Universiti Malaysia Sabah, 88999, Kota Kinabalu, Sabah

2Division of Engineering and Physics, Wilkes University, Wilkes-Barre, PA 18766, U. S. A.

ismail_s@ums.edu.my / ismailsaad07@gmail.com

Abstract - The computational analysis of ballistic saturation velocity for low-dimensional nano-devices was presented. The ballistic transport is predicted in the presence of high electric field for non-degenerate and degenerate regime. The saturation velocity is found to be ballistic regardless of the device dimensions. The intrinsic velocity limits this saturation velocity. It’s does not sensitively depend on the ballistic or scattering-limited nature of the mobility. In the degenerate realm, the saturation velocity is shown to be the Fermi velocity that is independent of temperature but strongly dependent on carrier concentration. In the non-degenerate realm, the intrinsic velocity is the thermal velocity that depends only on the ambient temperature.

Keywords - Ballistic transport, Nano-devices, Low-Dimensional devices, Saturation velocity

I. INTRODUCTION

The Metal Oxide Semiconductor Field Effect Transistor (MOSFET) is created when the electric field between the gate and the semiconductor in such that an inverted carrier population is created and forms a conducting channel. This channel extends between the source and drain regions, and the transport through this channel is modulated by the gate potential. As the channel length has gotten smaller, there has been considerable effort to incorporate a variety of new effects into the simple (as well as the more complex) models. These include short-channel effects, narrow width effects, degradation of the mobility due to surface scattering, hot carrier effects, and velocity overshoot. However, as gate lengths have become less than 100 nm, the issue is becoming one of ballistic transport rather than these other problems [1-4]. The speed is determined by the ease with which the carrier (electron or holes) can propagate through the channel of the device. In the earlier designs, the mobility of the carrier was believed to be of paramount importance. However, as development of the devices to nanoscale dimensions continued it became clear that the saturation velocity plays a predominant role [5-7]. The higher mobility brings an electron closer to saturation as a high electric field is encountered, but saturation velocity remaining the same no matter what the mobility. Until today, there is no clear consensus on the interdependence of saturation velocity on low-field mobility that is scattering-limited. There are a number of theories of high-field transport to answer this interdependence. Among them are Monte Carlo simulations, energy-balance theories, path integral methods, green function and many others. Rigor of mathematics and a number of clandestine parameters that are used in these simulations present a foggy picture of what controls the ultimate saturation of drift velocity.

Ballistic transport is referring to the situation in which the channel length is less than the mean-free path of the carriers, so that very little scattering occurs within the channel itself [8-10]. For example if the thermal velocity of a carrier in silicon taken as 2.5 × 10^7 cm/s at room temperature, a channel mobility of 300 cm^2/V.s leads to a relaxation time of 5 × 10^-14 second and a mean-free path of the order of 12 × 10^-7 cm, or 12 nm. Thus, only a few scattering events are expected to occurs in a channel length of 20 to 30 nm. While this is a very crude approximation, it points out that the properties of the carriers in these very small devices will be quite different than those in larger devices. In this regards, the paper discussed and elaborated the fundamental theory of ballistic saturation velocity for giving a meaningful interpretation of nanoscale MOSFET.

II. NATORI VS. LUNDSTROM MODEL OF BALLISTIC TRANSPORT

The principle of possibility that MOSFET would operate in ballistic regime as the gate length been scaled towards sub-100nm was purportedly introduced by Natori (1994). However, a simple treatment of ballistic transport as MOSFETs are scaled to their limits and as new devices structure are explored were firstly examined by Lundstrom (1997). Natori developed his expression of I-V characteristics with a full quantum mechanical (QM) basis. The current is expressed in terms of elementary parameters only without depending on the carrier mobility, µ. It is
independent of the channel length $L_g$ and is proportional to the channel width $W$ and oxide capacitance $C_{ox}$. Due to the quantum effects the energy bands in the channel split and the MOSFET current can be described in terms of energy at each of the split bands using the Fermi Dirac statistics. Thus, the total current can be found by summing the current over all the one-dimensional sub-bands. The sub-channel current component flowing in a direction is given by the product of the unit charge, the number of carriers flowing into the sub-channel per unit time, the transmission coefficient of the sub-channel and the probability that the destination is empty. This sub-channel current component must be integrated over the carrier energy to find the total current. The number of carriers flowing into the sub-channel is further expressed by the product of the input carrier group velocity, the density of states and the probability that the state is occupied by the carrier. The probability of carrier occupancy is given by the Fermi distribution function with the source Fermi level on the source side and that with the drain Fermi level on the drain side of the sub-channel. Both current directions the one from the source to the drain and that in the opposite direction are considered. Thus the total current in the sub-channel is expressed as

$$I = e \sum_{\text{sub}} \sum_{\nu} \left[ (\nu \cdot \mathbf{v}) \mathcal{D}(E) f(\phi_{FS}, E)[1 - f(\phi_{FD}, E)] ight. \\
- e \mathcal{D}(E) f(\phi_{FD}, E)[1 - f(\phi_{FS}, E)] T(E) dE$$

(1)

where $\mathbf{v}$ is the carrier group velocity of 1D wave propagating towards the drain though a certain sub-channel, $\mathcal{D}(E)$ is the density of states (DOS) for that wave, both evaluated near the source side edge of the sub-channel. $\nu$ and $\mathcal{D}(E)$ are the group velocity and the DOS for carriers propagating from drain to source near the drain side edge of the sub-channel. $f(\phi, E)$ is the Fermi distribution function with the Fermi level $\phi$. The source and drain regions are assumed to be ideal reservoirs with the Fermi energies $\phi_{FS}$ and $\phi_{FD}$ respectively. They feed carriers in thermal equilibrium to the channel and also absorb carriers from the channel without reflection (Natori 2006). $T(E)$ is the transmission coefficient of the sub-channel at energy $E$ that represent the probability that a carrier with certain energy can be transmitted over the barrier by the thermionic emission process. In ballistic transport its implies that $T(E)=1$.

Based on transmission view of the device [11], there is an energy barrier between the source and the drain that prevent the current to flow. This potential energy barrier height is modulated indirectly by the gate voltage $V_{GS}$ which is defined to begin at the top of the barrier. A positive $V_{GS}$ pushes the energy barrier down and allows current to flow. This operational principle works similarly as bipolar transistor which has a similar energy band diagram except that a positive base-emitter voltage $V_{BE}$ lowers the height of the barrier and allows current to increase exponentially [12]. In Lundstrom theory, two major scattering regions can be identified. The scattering in the barrier between the channel and the source, which gives a reflection $r_s$, and the scattering within the channel, which gives a reflection $r_c$. In both cases, the reflection coefficients are related to transmission coefficients $t$ by

$$r_s = 1 - t_s, \quad r_c = 1 - t_c$$

(2)

The steady-state flux which reaches the drain $a_d$ can now be written in terms of the injecting flux $a_s$ from source to the source-channel barrier (which is a function of the depth $y$) as

$$a_d = a_s t_s t_c$$

(3)

At the entrance to the channel (which is taken to be $x = 0$, with $x$ the axis aligned from source to drain), the density of carriers can be written as

$$n(0, y) = \left[ \frac{t_s a_s + r_s t_s a_s}{v_T} \right] = \frac{t_s a_s (1 + r_s)}{v_T}$$

(4)

where, $v_T$ is the velocity of the positively and negatively directed fluxes, and $y$ is the direction of the channel depth (normal to the oxide-semiconductor interface). Solving for $t_s$ in these equation yields

$$a_s = n(0, y)v_T\left( \frac{1}{1 + r_s} \right)$$

(5)

The sheet carrier density is given by integrating over the $y$ coordinate, as

$$n_s = \int_{0}^{y_{max}} n(0, y) dy = \frac{C_{ox}}{q}(V_{GS} - V_T)$$

(6)

With this result, the drain saturation current can be written as

$$I_{Dsat} = C_{ox}WV_T\left( \frac{1 - r_c}{1 + r_c} \right)(V_{GS} - V_T)$$

(7)
Equation (7) may be compared with a simple 1D model of a MOSFET. However here the reverse current is represented by the \( r_c \) term in the equation, but the form is quite similar to that of the simple theory. Also here the mobility \( \mu \) is not defined, but instead the carrier transport is discussed in terms of the velocity \( \nu_f \) and the transmission and reflection coefficients within the devices. In the ballistic limit the back scattering coefficient \( r_c = 0 \) and the transmission coefficient \( t_c = 1 \). Thus the maximum current is controlled by the injection velocity at the thermal source.

### III. Ballistic Intrinsic Velocity

Velocity response to the electric field results in velocity saturation in a high electric field. The current in a resistive (channel) is limited by this saturation value \( I_{sat} = n q \nu_{sat} A \) (3D), \( I_{sat2} = n_s q \nu_{sat} W \) (2D), \( I_{sat} = n_l q \nu_{sat} \) (1D) which in turn depends on the doping concentration \( n_d \) \((d=1, 2, 3)\) and the saturation velocity \( \nu_{sat} \). It is, therefore, essential to assess the magnitude of this saturation velocity that results in the current saturation.

In the archived work of Arora (1985) [8], the ballistic (B) 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 [9]. In equilibrium, the band diagram is flat and randomly oriented velocity vectors cancel each other. As the applied electric field tilts the band diagram, an electron traveling in the direction of electric field finds it difficult to surmount the barrier. An electron traveling in the opposite direction accelerates in a mean free path and collides, randomizing its velocity and restarting its journey for another mean free path.

\[
\nu_f \quad \text{E = 0} \quad \nu_f
\]

Figure 1 Random intrinsic velocity vectors \( \nu_f \) in a zero-field transforming to streamlined vectors in a very high electric field.

The net result is that the random vectors \( \nu_f \) streamline in a very high electric field. In the presence of a very high electric field, all electrons are streamlined opposite to the direction of the applied electric field. The ballistic (B) nature of the velocity is apparent from fig. 1 that streamlines the randomly oriented velocity vectors in a very high electric field.

In the absence of external stimulation, the carrier velocity vectors are randomly oriented. This means that their dipole energy \( \pm qE_x \ell_o \) in a mean free path \( \ell_o \) is also zero. However, as an electric field is applied, the electron energy decreases by \( qE_x \ell_o \) for electrons drift opposite to the electric field and increases for those drifting parallel to the electric field. This creates two quasi Fermi levels that so-called electrochemical potentials \( E_F \pm qE_x \ell_o \) within one mean free path on each side of the location of a drifting electron (or a hole). \( E_F \) is the Fermi energy related to carrier concentration.

As conduction and valence bands and associated Fermi and intrinsic levels tilt in an electric field, electrons in antiparallel directions are favored over those in the parallel direction finding it difficult to surmount the potential barrier with the net result that all carriers are drifting with the ballistic velocity in an infinite electric field as shown in fig. 2.

![Figure 2 Partial streamlining of random motion of the drifting electrons on a tilted energy band diagram in an electric-field.](image)

From traveling quantum wave perspective, the waves are reflected by the apparently insurmountable barrier in an intense driving field. In this scenario, the drifting electrons are in a relay race, passing on their drift velocity to next electron at each collision, hence the velocity of the electrons is unaffected by the collision and thus called as Ballistic (B) transport of a carrier. This is in direct contrast to the work of those researchers who believe in enhanced scattering for velocity to saturate. Published literature is inconclusive in predicting the dependence of high-field saturation velocity to low-field mobility. Often, the higher saturation velocity is indicated to arise from higher low-field mobility. This
confuses the issue as to what scattering controls the mobility and what controls the saturation velocity. Arora (1985) [8] framework includes all complicated scattering interactions into a single effective mean free path $\ell_o$ scalable under the strength of carrier scattering that leaves the saturation velocity unaltered and hence $B$. The mean path may be affected by the onset of quantum emission. The saturation velocity on the other hand is scaled by band structure parameters, unaffected either by the electric field or by the scattering that controls low-field mobility.

The ultimate unidirectional drift velocity is the saturation velocity that is the average of its absolute value $|v| = \sqrt{2E_i/m^*}$, where $E_i$ is the kinetic energy for a given dimensionality and $m^*$ is the carrier effective mass. When this averaging is taken by including the Fermi-Dirac distribution and density of states as a weight for a given dimensionality, the intrinsic $B$ velocity for a semiconductor is obtained as

$$v_{id} = \frac{\zeta_{d-1}(\eta_{Fd})}{\zeta_{d-2}(\eta_{Fd})}$$

(8)

with

$$v_{had} = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}$$

(9)

$$v_{ih} = \sqrt{\frac{2k_B T}{m^*}}$$

(10)

Figure 3 Relative intrinsic velocity versus normalized carrier density.

Figure 3 gives a plot of intrinsic $B$ velocity normalized to the thermal velocity of equation (10). In the non-degenerate regime, this ratio is $2/\sqrt{\pi} = 1.128$ for a 3D semiconductor, it drops to $\sqrt{\pi}/2 = 0.886$ for a 2D nanostructure and further drops to $1/\sqrt{\pi} = 0.564$ for a 1D nanostructure.

IV. BALLISTIC VELOCITY IN NON-DEGENERATE & DEGENERATE REGIME

In the non-degenerate and degenerate regime, the $B$ intrinsic velocity is found to be dependence on the temperature and the carrier concentrations. For the non-degenerate regime, the Fermi factor becomes one since both the Fermi-Dirac integral of order $3_{1/2}$, $3_0$ and $3_{-1/2}$ is following the Maxwell Boltzmann approximation. The $B$ intrinsic velocity for non-degenerate regime can be derived from equation (8) to (10) for each system dimensionality and respectively given as

$$V_{i3} = \frac{8k_B T}{\pi m^*}$$

(11)

$$V_{i2} = \frac{\pi k_B T}{2m^*}$$

(12)

$$V_{i1} = \frac{2k_B T}{\pi m^*}$$

(13)

The normalization of non-degenerate $B$ intrinsic velocity to thermal velocity of equation (10) for $d = 3$, 2, 1 is as depicted in figure 2.3. It is found that the ratio of thermal velocity is decreasing from 3D to 1-dimensional nanostructure system and is a factor of temperature. The limiting intrinsic velocity for three-dimensional (bulk) non-degenerate electrons is shown to be $2\sqrt{2k_B T/\pi m^*}$ as shown in equation (4) after being simplified. The ballistic velocity reported by Lundstrom and Guo (2005) [10] is half of this value for the reason that the authors are considering only the average over one-half of the Maxwellian distribution. Another difference is that the ballistic velocity $\sqrt{2k_B T/\pi m^*}$ quoted by Lundstrom and Ren (2002) [12] actually applies to bulk case not to the Q2D gas being considered in this study.

For degenerate case, the weighted average $|v|$ of the magnitude of the velocity with Fermi Dirac distribution is
also given by equation (1). However, in the degenerately doped regime, the B intrinsic velocity depends on Fermi velocity \( v_F \) instead of the thermal velocity \( v_{th} \) given as

\[
v_F = \sqrt{\frac{2(E_F - E_C)}{m^*}}
\]  

(14)

The Fermi Dirac distribution part can be simplified for the \( d = 3, 2 \) and 1 system as shown by

\[
\frac{\mathcal{F}_{d-1}^{(d-1)}}{\mathcal{F}_{d-2}^{(d-2)}}(\eta_d) = \frac{d}{d+1} v_{Fd}
\]  

(15)

Thus, the generic B intrinsic velocity can finally be expressed as

\[
v_{id} = \frac{d}{d+1} v_{Fd}
\]  

(16)

The Fermi velocity is derived by using carrier concentrations equation in each dimensionality. Then, the resulted B intrinsic velocity for degenerately doped regime for \( d = 3, 2, 1 \) of a low-dimensional system is given respectively as

\[
v_{i3} = \frac{3}{4} \frac{\hbar}{m} \left(3n_i \pi^2\right)^{1/3}
\]  

(17)

\[
v_{i2} = \frac{2}{3} \frac{\hbar}{m^*} \sqrt{2/m_2}
\]  

(18)

\[
v_{i1} = \frac{1}{4} \frac{\hbar}{m^*} n_1 \pi
\]  

(19)

Figure 4 indicates the B intrinsic velocity as a function of temperature. The velocity for low carrier concentration follows \( T^{1/2} \) behavior independent of carrier concentration. However for high concentration (degenerate carriers) the velocity depends strongly on concentration and becomes independent of the temperature. Meanwhile, fig. 5 shows the graph of B intrinsic velocity as a function of carrier concentration for three temperatures \((T=4.2 \text{ K}, 77 \text{ K}, \text{ and } 300 \text{ K})\). As expected, at a low temperature, carriers follow the degenerate statistics and hence their velocity is limited by an appropriate average of the Fermi velocity that is a function of carrier concentration.

V. CONCLUSION

Using the distribution function that takes into account the asymmetrical distribution of drifting electrons in an electric field is presented. This distribution function
transforms the random motion of electrons into a streamlined one that gives the ultimate saturation velocity that is a function of temperature in nondegenerate regime and a function of carrier concentration in the degenerate regime. The ultimate drift velocity is found to be appropriate thermal velocity for a given dimensionality for no degenerately doped samples. However, the ultimate drift velocity is the appropriate average of the Fermi velocity for degenerately doped samples.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support from FRGS (FRG0248-TK-2/2010) and ERGS funds (ERGS0002-TK-1/2011) of Minister of Higher Education Malaysia (MOHE). The author is thankful to the Universiti Teknologi Malaysia (UTM) and Universiti Malaysia Sabah (UMS) for providing excellent research environment in which to complete this work.

REFERENCES