Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI: 10.21474/IJAR01/17587
DOI URL: http://dx.doi.org/10.21474/IJAR01/17587

RESEARCH ARTICLE

ELECTRON TRANSPORT IN GRAPHENE BASED NANOTRANSISTOR AND USE OF NEGATIVE CAPACITANCE FOR STEEPER SUB-THRESHOLD SLOPE

Rahul B. Awale
Department of Electrical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai-400076, India.

Manuscript Info

Manuscript History
Received: 25 July 2023
Final Accepted: 27 August 2023
Published: September 2023

Key words:-
Graphene, GNRFET, Electron Transport, Ballistic FET, Subthreshold Slope, Negative Capacitance

Abstract

Graphene has demonstrated tremendous promise in recent years as a material that, in the future, could replace silicon-based materials because of its exceptional electrical transfer characteristics. The diffusive MOSFETs suffer from short channel effects caused by their shorter channels, however graphene has numerous unusual features. The strongest material yet tested, Graphene exhibits a significant and nonlinear diamagnetism, has great mobility at room temperature, a low atomic thickness, a high current density, and is almost transparent. A single sheet of carbon atoms organized in a hexagonal lattice makes up graphene, an allotrope form of carbon. It is a semimetal with little short channel effects overall and little overlap between the valence and conduction bands. The Graphene nanoribbon has been incorporated into the ballistic nanotransistor as its channel and differentiate it form the diffusive one.In this review , I have presented the report on the survey of Nanotransistor . It has been divided into three sections : (1) Ballistic transport in Nanotransistor (2) Modelling of GNRFET (Graphene Nano-ribbon FET) using NEGF (Non-equilibrium Green's function) (3) Using ferroelectric capacitance into FET to reduce it's sub-threshold regime below 60 mV/decade.The purpose of this review is to understand charge carrier transport phenomenon in the Nanotransistor. This includes the description about the ballistic nanotransister and differentiate it form the diffusive one. The NEGF allows us to characterizing it's transfer characteristics in the real lattice space.Also the Id-Vd, sub-threshold regime , log(Id)-Vg , transmission coefficient and their application to drastically enhance the device performance for low energy digital electronics are being studied and modeled. Furthermore, the negative capacitance due to FE (ferroelectric layer) in FET will enhance the switching speed of FET and reduce the sub-threshold swing (SS) below 60 mV/decade which is impossible in case of traditional MOSFET. Thus, improving I_{ON}/I_{OFF} ratio for low power digital devices.

Introduction:-

In MOSFET , we control the current flow from the source to the drain by controlling the height of this energy barrier between the source and the drain. When we push the barrier down electrons flow across. If they flow across unimpeded just in a straight line and don't encounter anything that they can scatter or bounce off of we call that ballistic transport.

Corresponding Author:- Rahul B. Awale
Address:- Department of Electrical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai-400076, India.

Now if the channel length is longer, or there are impurities or surface roughness, or lattice vibrations the electrons might scatter, they might turn around and go back out the source, they might bounce around a few times and eventually go out the drain. If they scatter many times we call that diffusive transport and that's the domain of the original MOSFET, the traditional MOSFET theory was really worked out assuming diffusive transport. If they scatter a few times then we call that quasi-ballistic transport. It's somewhere between this ballistic and diffusive limits.

The electron flow is caused by the lowering all energy levels at the positive drain contact with respect to the source contact .

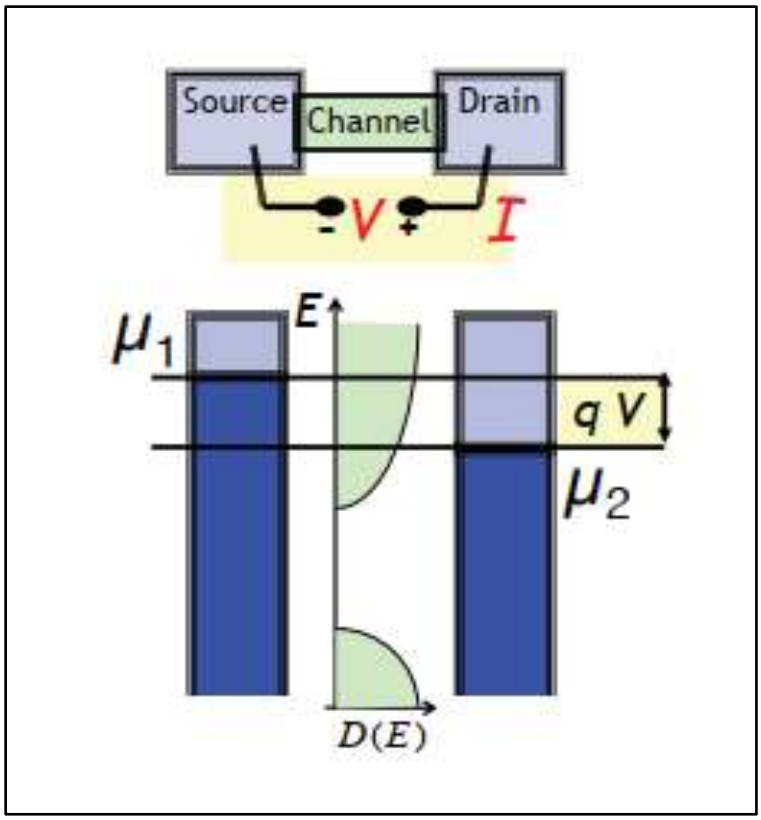


Figure 1.1:- Application of gate voltage and it's impact on electrochemical potential[1].

This happens when gate voltage $V_G > 0$ for the p-type ballistic MOSFET. Thus , the electrons hop over the channel barrier from the source provides a current from the source to drain. The electrons that hop over the barrier from the drain gives a current from the right to the left. And the net current is just the difference between the two i.e

$I(E) \sim f_1(E) - f_2(E)$

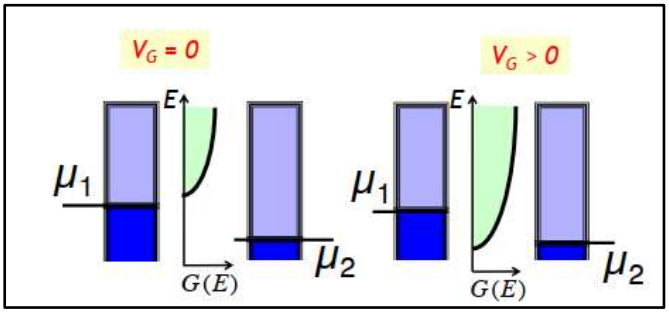


Figure 1.2:- A positive gate voltage (V_G) causes the states to descend in energy, which increases the current in a FET[1].

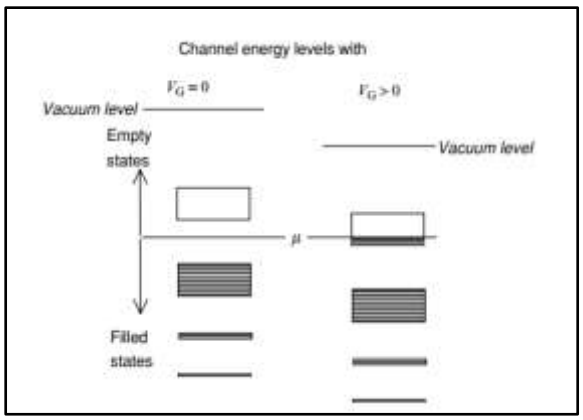


Figure 1.3:- P-type or hole conduction example. The quantity of electrons in the channel is decreased when the gate voltage is negative ($V_G < 0$)[2].

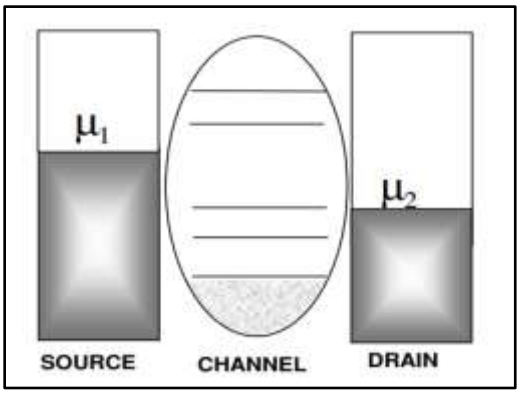


Figure 1.4:- Electrons in the contacts occupy the available states with a probability described by a Fermi function f(E) with electro-chemical potential μ[2].

$\mu_1 - \mu_2 = qV_D$

$f_1(E) = \frac{1}{1+\exp[(E-\mu 1)/kT]} = f_0(E-\mu_1)$

$f_2(E) = \frac{1}{1+\exp[(E-\mu 2)/kT]} = f_0(E-\mu_2)$

$f_1(E)$ and $f_2(E)$ are the Fermi functions or the occupation factors in the source and drain contacts respectively.

The the ratio of current from source to drain and from drain to source is just the ratio of their probabilities to hop over the barrier. So the ratio of those two currents is $\exp(-E_{SB}/K_BT)$. That difference in barrier height comes because of applied voltage to the drain and increased the barrier for electrons from the drain.

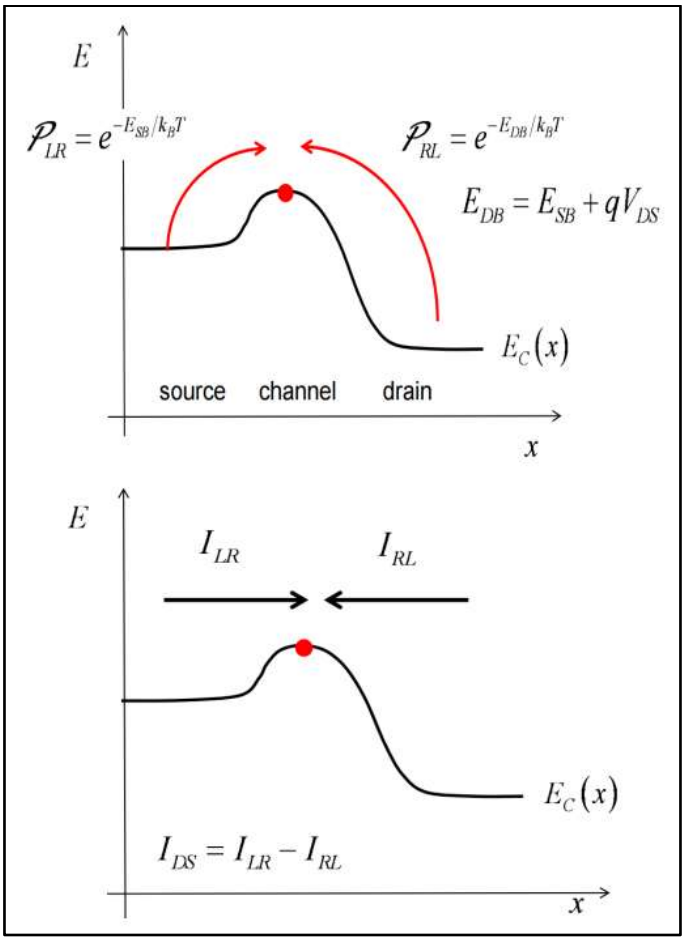


Figure 1.5:-An illustration of a ballistic MOSFET's two fluxes together with the probabilities of electron injection from the source and drain. I_{LR} stands for the current injected from the source, and I_{RL} for the drain. $I_{DS} = I_{LR} - I_{RL}$, where the difference between the two is the net drain current[3].

More specifically ,

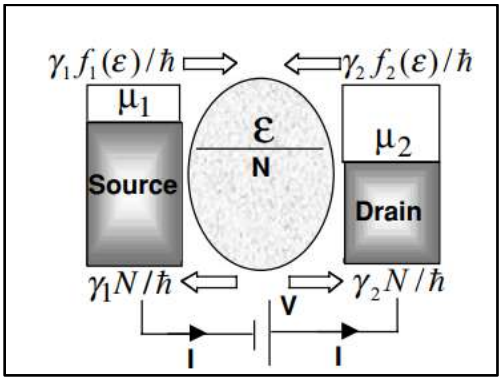


Figure 1.6:- Simple rate equation illustration showing the flow of electrons into and out of a one-level channel at the source and drain ends[2].

The Expressions for I_{LR} or I_1 (current from source to drain) and I_{RL} or I_2 (current from drain to source) can be written as ,

$I_1 = \frac{q\gamma 1}{\hbar}(f_1\text{-}N)$

$I_2 = \frac{q\gamma 2}{\hbar}(f_2 - N)$

$N = \frac{\gamma 1 f 1 + \gamma 2 f 2}{\gamma 1 + \gamma 2}$

$I = \frac{q}{\hbar}\frac{\gamma 1 \gamma 2}{\gamma 1 + \gamma 2} = \frac{q\gamma 1}{2\hbar}$ if $\gamma_1 = \gamma_2$

γ_1/h and γ_2/h are rate constants , the rates at which an electron in the level ε initially will escape into the source and drain contacts respectively.

N is the number density of charge carriers in the barrier under non-equilibrium condition (applied gate voltage V_G).

Furthermore , we can always write current as a product of charge times velocity. So the current from the left to the right is carried by electrons with positive velocities, electrons moving in the plus x direction. And current from right to left is carried by electrons with negative velocities.

$I_1 = I_{LR} = WqN_s^+ v_T$

$I_2 = I_{RL} = WqN_s^- v_T$

W is the width of the MOSFET , v_T is the voltage equivalent of temperature , Ns^+ is Electron density per square centimeter in the inversion layer for electron moving in +x direction and Ns^- is Electron density per square centimeter in the inversion layer for electron moving in -x direction.

Consequently , the expression for net current I_{DS} would be ,

$I_{DS} = I_{LR} - I_{RL} \Longrightarrow I_{DS} = I_{LR}(1 - e^{-\Delta EB/KBT})$ (1.1)

ΔE_B is the difference of source and drain barrier .

$\Delta E_B = E_{DB} - E_{SB} = qV_{DS}$

To get the complete expression for I_{DS} we need to know the charge density at the top of the barrier (x=0) ,

$Q_n(x=0) = \frac{ILR + IRL}{WvT}$

$= \frac{ILR}{WvT}(1 + I_{RL}/I_{LR})$

$= \frac{ILR}{WvT}(1 + \exp(-qV_{DS}/k_BT)$ (1.2)

Then , from the equations 1.1 and 1.2 ,

$I_{DS} = W|Q_n(x=0)|\upsilon_T \frac{(1 - exp(-qVDS/kT))}{(1 + exp(-qVDS/kT))}$ (1.3)

For small drain bias(linear regime) , the expression for I_{DS} becomes as ,

$V_{DS} << K_BT/q$

$I_{DS} = WQ_n(0)v_T \frac{qVDS}{KT}$(by Taylor series , $e^x \sim 1 + x$) (1.4)

$I_{DS\,(diffusive)} = \frac{W}{L}\mu Q_n(0)V_{DS}$(1.5)

By comparing the equation 1.4 with 1.5, we can conclude that the ballistic expression is independent of the channel length. Because it do not matter how long the device is if electrons are going through unimpeded, the length of the channel will not limit the current. In the diffusive approximation the longer the current, the more scattering that occurs and the harder it is for electrons to flow from the source to the drain.

For large Drain bias (saturation regime) , the expression for I_{DS} becomes as ,

$V_{DS} >> K_BT/ q$

$I_{DS} = WQn(0)v_T$ because $e^{-x} \sim 0$(1.6)

$I_{DS\,(diffusive)} = WQn_{(0)}v_{SAT}$(1.7)

By Comparing the equation 1.6 with 1.7 . we can state that net current equation in ballistic nanotransistor is independent of the saturation velocity that's simply because ideally there is no scattering and thus , it is replaced by the thermal equilibrium velocity of electrons.

Ballistic Transport in Transistor
Visualization of ballistic transport
Let 't' be the transport or transfer time for the electron along the channel of length L then,

Ballistic regime : Transfer time $t \sim L$
Diffusive regime: Transfer time $t \sim L^2$

$t_{(ballistic)} = t_B = \frac{L}{<vz>}$(1.8)

$<v_z>$ is the average electron velocity moving along the channel.

But , for diffusive transistor ,

$t = \frac{L}{<vz>} + (L^2/2D)$(1.9)

D is the diffusion constant. $D=<v_z\tau>$, τ is mean free path of the electron travelling along L.

From 1.8 and 1.9 ,

$t = t_B(1+ L<v_z>/2D)$ and mean free path $\lambda = 2D/ <v_z>$

$t=t_B(1 +L/\lambda)$(1.10)

For the expression of ballistic conductance G_B we can use diffusive conductance formula G,

$G = q^2D/2t$(1.11)

Using equation 1.10 and 1.11 , we can drive for G_B

$G = G_B\lambda/(L+\lambda)$ and $G_B = q^2D/2t_B$

$G=\frac{\sigma A}{L+\lambda}$

Where $\sigma A = G_B\lambda$

Also , R and R_B as diffusive and ballistic resistances respectively ,

$R_B = R\left(1 + \frac{L}{\lambda}\right)$ (1.12)

For 3D , 2D and 1D conductors we can express G as,

$G = \frac{\sigma}{L+\lambda}\{1,W,A\}$(1.13)

Where $\sigma = G_B\lambda\{1,\frac{1}{W},\frac{1}{A}\}$

Landauer Approach

The contacts source and drain are assumed to large area contacts.So the population of the energy states in these contacts are given by Fermi functions. But the Fermi functions or the Fermi levels might be different because we might apply a voltage to one contact versus the other.

Thus , the expression derived for the net current I would be ,

$I = \frac{2q}{h}\int T(E)M(E)(f1 - f2)dE$(1.14)

T(E) is the transmission probability with range from 0 to 1 , M(E) is the number of modes in which current could flow .

T(E) = 1 means ballistic transport

T(E) << 1 means diffusive transport

$M(E) = \frac{Wh}{4}\langle v_x^+(E)\rangle D(E)$

$v_x^+(E)$ is the average velocity in the direction of current flow

$D(E) = g_v(m^*/\pi\hbar^2)$ for 2D conductor , g_v is degeneracy factor

$\langle v_x^+(E)\rangle = \frac{2}{\pi}v(E)$

$v(E) = \sqrt{\frac{2E}{m*}}$

Self-Consistent field potential U_{SCF}

Barrier potential U

Increasing the V_G will impact barrier potential U and this will reduce the barrier because $V_G$increases the current in an FET by moving the states down in energy. At the particular point of increasing V_G the saturation current I_{SAT} has

reached. But , practically the current does not saturate completely because the barrier potential is also dependant upon the $V_{DS.}$

The current saturates once once $\mu_2$drops below the band-edge but the states in the channel are also lowered by the V_{DS} as shown below .

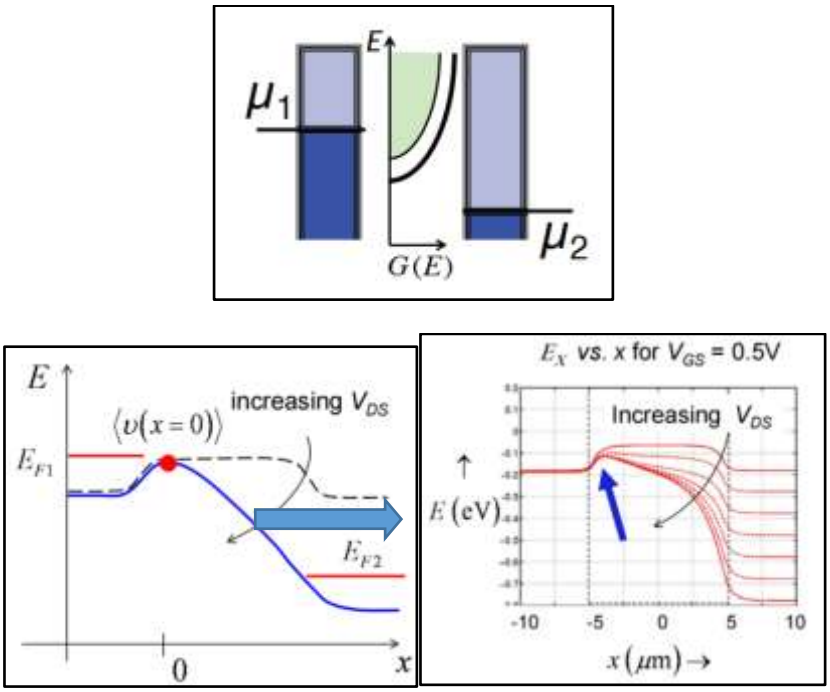


Figure 1.7:-Potential at the top of the barrier U_L (Laplacian potential) due to terminal biases is dependant upon the V_G and V_D , V_S is usually grounded[3].

$U_L = -q(\alpha_G V_G + \alpha_D V_D + \alpha_S V_S)$(1.15)

$\alpha_G=\frac{CG}{C\Sigma}\alpha_D=\frac{CD}{C\Sigma}\alpha_S=\frac{CS}{C\Sigma}$

C_G , C_D and C_S are the gate , drain and source capacitance respectively. C_Σ is the total capacitance.We have to reduce the factor α_D to get total U_L dependence on the $V_{G.}$

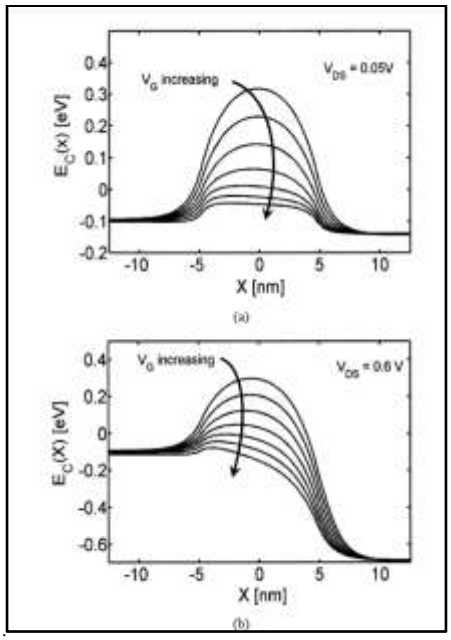


Figure 1.8:- Computed energy band diagrams under (a) low drain bias and (b) high drain bias. The parameter is the gate voltage when potential at the top of the barrier is independent of the V_D[4].

This is analogous to the DIBL (Drain induced barrier lowering)

To ensure that V_G has far greater control over the channel than V_D it is necessary to make the insulator thickness a small fraction of the channel length so that α_D would be as small as possible.

Moreover , under non -equilibrium condition the excess mobile charge carriers occurs due to the change to the number density of the electron ΔN. This gives rise to the new potential U_P. Then , total potential at the top of the barrier is U_{SCF} i.e self consistent field potential.

$U_P = (q^2/C_\Sigma)\Delta N$

$U_{scf} = U_L + U_P = -q(\alpha_G V_G + \alpha_D V_D + \alpha_S V_S) + U_C \Delta N$(1.16)

$U_C = (q^2/C_\Sigma)$

$\Delta N = N - N_0$,

The N_0 is the equilibrium electron density at the top of the barrier (Terminal biases are zero).

$N_0 = \int_{-\infty}^{\infty} D(E)\, f(E - EF) dE$(1.17)

With the bias application: Source terminal is ground (i) The self-consistent potential at the top of the barrier becomes and (ii) the states at the top of the barrier are now populated by two different Fermi levels.

The non-equilibrium electron density at source N_1 and drain N_2

$N_1 = \frac{1}{2} \int_{-\infty}^{\infty} D(E - Uscf)\, f(E - EF1) dE$

$N_2 = \frac{1}{2} \int_{-\infty}^{\infty} D(E - Uscf)\, f(E - EF2) dE$(1.18)

$N = N_1 + N_2$ is the non-equilibrium electron density at the top of the barrier (Terminal biases applied)

Then, the bias induced charge is,

$\Delta N=(N_1 + N_2) - N_0$

From equation 1.14 , we can write the net current expression under non-equilibrium condition,

$$I=\frac{1}{q}\int_{-\infty}^{\infty} dEG(E-U)(f1(E)-f2(E))$$

Using the change of the variables

$$I=\frac{1}{q}\int_{-\infty}^{\infty} dEG(E)(f1(E+U)-f2(E+U)) \quad (1.19)$$

Process to calculate I_D using channel potential U_{scf} (or U) as shown below [1],

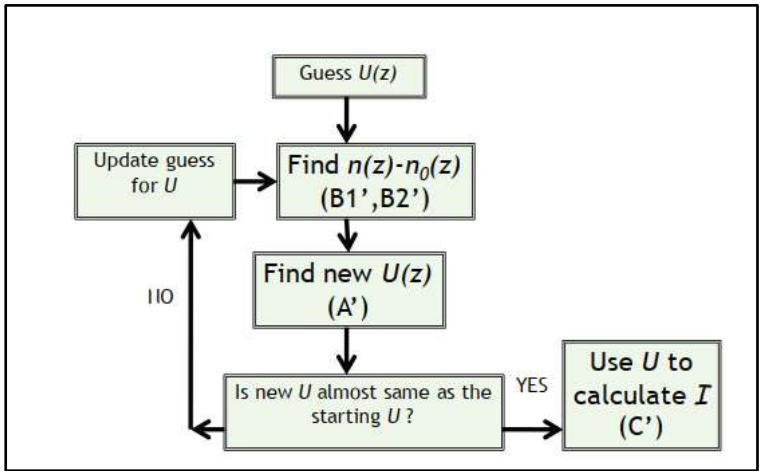


Figure 1.9:- Flow chart for Self Consistent field Potential $U_{scf.}$

(A') $\frac{d}{dz}(\varepsilon\frac{dU}{dz}) = q^2 (n - n_0)$ (1.20)

(B1') $n(z) = \int_{-\infty}^{\infty} dE \frac{D(E-U(z))}{L} \frac{1}{1+exp(\frac{E-\mu(z)}{kT})}$

(B2') $n_0 = \int_{-\infty}^{\infty} dE \frac{D(E)}{L} \frac{1}{1+exp(\frac{E-\mu 0}{kT})}$ (1.21)

This is called extended channel model where barrier potential is considered with respect to the position z. Here , D(E) is varying while moving from source to drain.And electric field is considered. The electric field causes to have more D(E) at the drain contact compared to that at source contact . This is because number of modes at the drain contact is larger than that at source contact. Specifically, the M(E) depends upon how far the given energy E from the bottom of the band defined by U(z) .

M(E) modes would be ,

$$M(E)= g\frac{2W}{h}\sqrt{2m(E-Ec)}\upsilon(E - Ec) \quad (1.22)$$

The equations B1' and B2' has to be solved self consistently with the Poisson equation 1.20 to get the U(z). This model based on the assumption of the electro-chemical potential μ(z).

Ballistic Transport in 2D GNRFET using NEGF

Introduction to GNRFET

The GNRFET and its possible applications are briefly described in this introduction, with an emphasis on the GNRFET's contribution to the development of nanoelectronics. One-dimensional Nanoribbon Graphene field-effect transistors, or GNRFETs, are a potential new class of nanoelectronics that have attracted a lot of attention recently. The structure of this device is a few nanometer wide narrow Graphene ribbon with source and drain electrodes at either end and a gate electrode on top. The behaviour of the GNRFET can be precisely controlled by modulating the electrical conductivity of the Graphene ribbon by the application of a gate voltage.Due to its unique one-dimensional structure, the GNRFET exhibits exceptional electrical and mechanical properties, making it an ideal candidate for use in various electronic applications, such as high-speed logic devices, ultra-sensitive sensors, and low-power circuits.With a GNRFET, a third electrode (the gate) is positioned on top of the Graphene ribbon, which is sandwiched between two metal electrodes (the source and drain). An electric field is produced when a voltage is applied to the gate electrode, altering the energy level of the electrons in the Graphene ribbon. The gadget functions in a "depleted" state if the gate voltage is adjusted at a level below the Graphene ribbon's Fermi energy level. In this mode, the device's electrical conductivity is reduced as a result of the gate voltage's reduction in the quantity of electrons in the Graphene ribbon.In contrast, the device works in a "accumulated" mode if the gate voltage is set higher than the Fermi energy level. In this state, the device's electrical conductivity rises as a result of the gate voltage drawing more electrons to the Graphene ribbon. The GNRFET may be turned on or off by adjusting the gate voltage, enabling fine-grained control of the electron flow through the component.

Characterizing the GNRFET with NEGF

The Non-Equilibrium Green Function (NEGF) method is a powerful computational tool for modeling and characterizing the behavior of one-dimensional Graphene nanostrip field-effect transistors (GNRFETs). The NEGF method is a quantum transport model that can accurately capture the electrical and transport properties of nanoscale devices such as GNRFETs by solving the Schrödinger equation and the corresponding equations of motion for the electron density matrix. In the NEGF method, the GNRFET is modeled as a dissipation region between two semi-infinite electrodes (source and drain). The electronic properties of the Graphene ribbon and metal electrodes are described by the respective Hamiltonians, which are then coupled to the self-energy responsible for electron transport through the GNRFET. The NEGF method can provide valuable information about membrane transport properties. GNRFET such as its conductivity, current-voltage characteristics and transmission spectrum.The method can also be used to study the effects of various physical factors such as temperature, diffusion mechanisms and interference on device performance. In addition, the NEGF method can be used to provide a theoretical framework for understanding the behavior of GNRFETs. to optimize the design of these devices for specific applications. For example, this method can be used to identify the optimal width and length of a Graphene ribbon, as well as the optimal lattice voltage and doping level to achieve desired electrical and transport properties.

Review on Graphene Nanoribbon Field Effect Transistor

Structure of 1-D Armchair Graphene

Graphene, a monolayer of carbon atoms arranged in a two-dimensional honeycomb lattice, has gained attention as a potential material for use in nanoelectronics. based on Graphene.Devices provide superior thermal conductivity, high carrier velocity for quick switching, monolayer thin body for optimal electrostatic scaling, and high mobility for ballistic transport. A notable benefit over carbon nanotubes (CNTs) is the potential to fabricate wafer-scale Graphene films with full planar processing for devices, which promises strong integration potential with traditional CMOS production techniques. Even though Graphene in two dimensions has no band gap, it can be patterned into a few nanometer wide Graphene Nanoribbon (GNR) to create a band gap. A GNR's band-gap is often inversely proportional to its width.

The structure of the device can change, resulting in variable electrical properties, depending on how the nano-ribbon is sliced or nano-tubes are formed. The "chirality" of the device is frequently used to describe its construction.

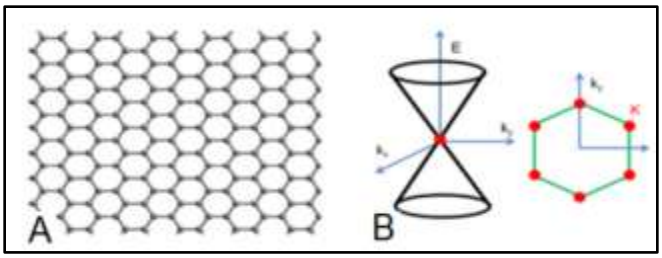


Figure 2.1:- Graphene's atomic and band structure. A) The honeycomb-like arrangement of carbon atoms in Graphene's one layer of atomic structure. B) The 1st BZ of Graphene band structure with 6 Dirac points at the vertices. The E-K linear relationship close to the Dirac Point.

A thin Graphene strip with its edges organized in a way that resembles the shape of an armchair is known as an Armchair Graphene Nanoribbon (AGNR). The 4-atom unit cell AGNR's structure can be summed up as follows:

- ✧ A single layer of carbon atoms organized in a hexagonal lattice makes up the Graphene Nanoribbon.
- ✧ Two pairs of parallel lines of carbon atoms that run the length of the ribbon are what give the ribbon its edges. These two sets of lines are 90 degrees apart from one another.
- ✧ Four carbon atoms are grouped in a rectangle within the unit cell of the AGNR.
- ✧ The carbon atoms in the unit cell are all bonded to each other and to their nearest neighbors in the Graphene lattice, forming a stable and rigid structure.

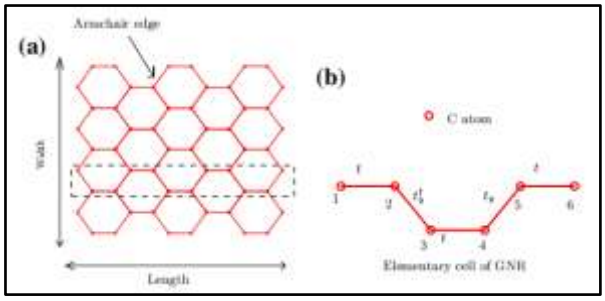


Figure 2.2:- (a) A schematic of an armchair edge GNR with an elementary cell.(b) 1D unit cell and associated binding energy t. And , $t_y=\exp(ik_y a)$[5].

Hamiltonian of the of 1-D Armchair Graphene

The matrix H in these most basic models has a dimension of (NxN), where N is the total number of carbon atoms. The matrix element H_{NXM} equals some value t (hopping parameters or binding energy) if N and M are the nearest neighbors, whereas its diagonal elements have some value U. The U is the self-consistent potential energy on the atomic sites ,it's also called on-site energy.

The complete Hamiltonian H can be written as ,

$$H=\begin{bmatrix} \alpha & \beta & 0 & 0 \\ \beta' & \alpha & \beta & 0 \\ 0 & \beta' & \alpha & . \\ 0 & 0 & . & . \end{bmatrix}$$

Whereas $\alpha = U_i(x)$, $\beta = \text{texp}(ik_ya)$, $\beta' = \text{texp}(-ik_ya)$, a is the lattice constantand $k_y$is the transverse wave vector. Also, t = 3.0 eV , k_y =0 for hopping between atoms on the same plane . k_y is non-zero for hopping between atoms on the different planes.

In this report , I have used $Ui(x) = U_L$ (or U_o) since I am considering only the potential at top of the barrier in the channel (or device). So, This modelling do not involve potential profile calculations using U_{scf} (self-consistent potential) method.

$U_L = -q(\alpha_G V_G + \alpha_D V_D)$ (2.1)

$\alpha_G = \frac{CG}{C\Sigma}$, $\alpha_D = \frac{CD}{C\Sigma}$

U_L is the Laplacian potential and is usually gate voltage dependent for well controlled MOSFET. $\alpha_D \sim 0$, $\alpha_G \sim 1$. C_G , C_D are capacitance of gate and drain respectively .C_Σ is the sum of capacitance of gate , drain and source.

The band-structure of the device can then be determined by utilizing the tight binding technique to make a 1-D array of unit cells using:

$H(k) = \alpha + \beta + \beta' = U_o + \text{texp}(ik_ya) + \text{texp}(-ik_ya)$(2.2)

Thus , the complete Hamiltonian would be shown as ,

$$H = \begin{pmatrix} Uo & t & 0 & 0 & \dots \\ t & Uo & te^{(ikya)} & 0 & \dots \\ 0 & te^{(-ikya)} & Uo & t & \dots \\ 0 & 0 & t & Uo & \dots \\ . & . & . & & \end{pmatrix} \quad (2.3)$$

Self Energies of the Contacts and Green Function

The self-energy defines the coupling of the contacts with the channel. the coupling of the source and drain contacts to the device is described by analyzing the source and drain self-energy matrices (ΣS and ΣD). The elementary unit cell of GNRFET complete device is depicted in Figure 2.3, where it is split into a channel region (unit cells 1, 2,..., M-1, M), a right contact region (unit cells M + 1, M + 2,...), and a left contact region (unit cells 0, -1, -2,..). Only the atoms on the channel's first and last are shown to be linked to the contacts. As a result, the first 1x1 entry for the source self-energy matrix is a non-zero block, and the last 1x1 entry is for the drain self-energy matrix.

$\Sigma S(1X1) = \beta' g_0 \beta$ and $\Sigma D(NXN) = \beta' g_{M+1} \beta$(2.4)

In Matrix form ,

$$\Sigma S = \begin{pmatrix} \beta' g_0 \beta & 0 \ldots\ldots\ldots \\ 0 & 0 \ldots\ldots\ldots \end{pmatrix}$$

$$\Sigma D = \begin{pmatrix} \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ 0 & 0 \\ \ldots\ldots 0\ \beta' g_{M+1} \beta & \end{pmatrix} \quad (2.5)$$

The g_0 and g_{M+1} are the surface Green function for sourceand drain contacts respectively.

The potential inside the contacts is constant so , $U_0 = U_{-1} = U_{-2} = U_{-3}$. Also , I have considered potential inside channel as U_o by neglecting the effect of U_{scf} . Thus,

$$g_0(k_y) = \frac{-(E-U_0)^2+t^2-t_y t^+_y \pm \sqrt{[[(E-U_0)^2-t^2-t_y t^+_y]^2-4(E-U_0)^2 t_y t^+_y}}{2(E-U_0)t_y t^+_y} \quad (2.6)$$

The recursive relation for the surface green function inside the source contact can be calculated by knowing g_0.

$$g_0(k_y) = [(E + i0^+)I - U_0 - t g_{-1} t]^{-1}$$

$$g_{-1}(k_y) = [(E + i0^+)I - U_{-1} - t_y g_{-2} t^+_y]^{-1}$$

$$g_{-2}(k_y) = [(E + i0^+)I - U_{-2} - t g_{-3} t\ \ ]^{-1}$$

$$g_{-3}(k_y) = [(E + i0^+)I - U_{-3} - t_y g_{-4} t^+_y]^{-1} \quad (2.7)$$

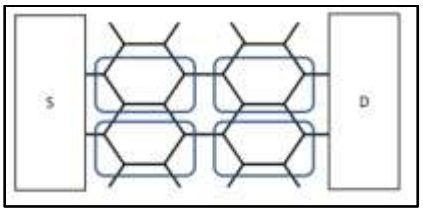
Due to periodicity of the lattice and constant potential inside the contact , we can conclude $g_0 = g_{-2}$ and $g_{-1} = g_{-3}$. And , same goes for the drain contact.

Then, the Retarded Green function is ,

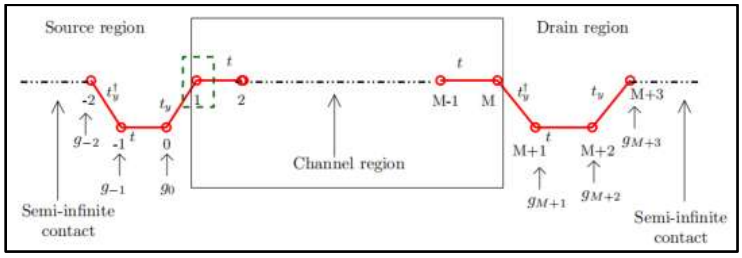
$$G^R = [(E+i\eta^+)I - H - \Sigma S - \Sigma D]^{-1} \quad (2.8)$$

$\eta+$ an infinitesimal value greater than zero.

E is the energy and I is the identity matrix of N X N.



(a)



(b)

Figure 2.3 : (a)The schematic of GNRFET with 4 atom unit cells (in blue rectangles). (b)The elementary unit cell of GNRFET[5].

Modelling using NEGF

Transfer Characteristics of GNRFET

For Modelling in MATLAB , the tight binding parameter 't' and a few basic constants are defined in this section, along with some user input on the device's kind, structure, and width (measured in 4-atom unit cells). Given that the device is operating within the ballistic limit the transport through the device is unaffected by its length, thus in order to computation rate as high as possible,the device's length set equal to 1.

η^{+} =0.01 eV

The number of states per interval of energy at each energy level that are accessible to be inhabited is referred to as a system's state density called DOS. The LDOS due to the source and drain contacts,

$D_{S,D} = G\Gamma_{S,D}G^{+}$(2.9)

$\Gamma_{S,D}$ is the broadening function due the coupling between contact and channel.

$\Gamma_{S,D} = i[\Sigma_{S,D} - \Sigma^{+}_{S,D}]$(2.10)

The transmission function T(E) defining the probability of the transmission vs. Energy is given as ,

$T(E) = \text{trace}(\Gamma_S G \Gamma_D G^{+})$(2.11)

Using the Landaurer-Buttiker formula at finite temperature T , the current equation could be written as ,

$I = \frac{2q}{h} \int \sum_{ky} T(E, ky)[fs(E) - fd(E)]dE$(2.12)

$fs(E)$ and $fd(E)$ are the occupation factors in the source and drain contacts respectively.

$fs(E) = \frac{1}{1+exp\,[(E-\mu s)/KT]}$, $fd(E) = \frac{1}{1+exp\,[(E-\mu d)/KT]}$(2.13)

μs , μd are the electro-chemical potentials at source and drain contacts respectively.

Using the equations 2.10 , 2.11 ,2.12 and 2.13 , we can plot current as the function of drain and gate voltage.From equations 2.1 and 2.2 , we can decide whether to make current profile vs. drain voltage or gate voltage by considering the values for α_G andα_D.

The device performance of GNRFET are defined by the Id-Vds characteristics as given in figure 2.4.I have set the values of Vgs and varied the Id with respect to the Vds. The charge at the top of the barrier should be independent of the Vds for well gate voltage controlled FET. Moreover , the Vgs values are as twice as Vds in this model. Since , only charge at the top of the barrier is being considered , then $Q_{(0)}$ is given by ,

$Q_{(0)} = C_G(V_{GS} - V_T)$(2.14)

V_T is threshold voltage.
Thus, given that the electron's ability to cross the potential barrier beneath the top gate determines the source-drain current, the formula for this current's density is as follows.

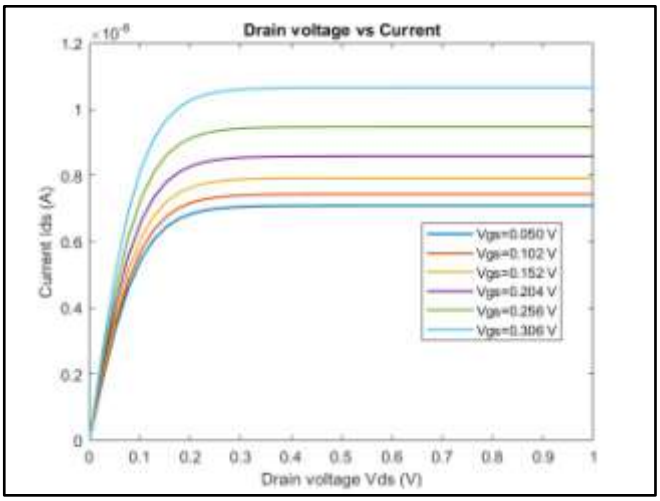


Figure 2.4:- Drain Voltage Vds vs. Drain Current Ids.

$J(E) = \frac{1}{2}q\frac{2}{\pi}\sqrt{\frac{2E}{m*}}D(E)$ (2.15)

Figure 2.4 demonstrates how the average electron velocity increases until it reaches saturation at the top of the barrier. The Fermi-Dirac distribution's thermal equilibrium velocity is the only thing that determines the saturated velocity at the top of the barrier.

$Id = \int_{-\infty}^{\infty} J(E)[fs(E) - fd(E)]dE$(2.16)

From equation 2.14 and 2.15, we can get Id which is analogous to Id from equation 2.12.Under non-equilibrium , fs ≠ fd over the energy range qVds. DIBL is achieved after saturation point . While comparing to conventional MOSFET:

1. When the drain voltage is raised in traditional MOSFETs, the depletion area moves closer to the source, lowering the threshold voltage. This is known as DIBL. As a result, the device's capacity to regulate current flow is diminished. Ballistic GNRFETs, in comparison, have a narrower channel, making DIBL far less of a problem.
2. Conventional MOSFETs are distinguished by a relatively low saturation current, which depends on the channel length, breadth, and charge carrier mobility. Ballistic GNRFETs, on the other hand, function in a ballistic transport regime where the charge carriers flow freely without scattering, leading to a greater saturation current.
3. In typical MOSFETs, impurities or defects in the channel region can scatter charge carriers, increasing the likelihood of backscattering. This lowers the carriers' overall mobility and lowers the transistor's performance. Ballistic GNRFETs, in contrast, have a lesser likelihood of backscattering since the graphene material is free of imperfections and flaws.

4. The transmission probability T(E) is equal 1 for ballistic GNRFET and it's in between 0 to 1 for diffusive MOSFET. The Id is directly proportional to the VT for ballistic case and for diffusive case , Id is dependant on the saturation velocity V_{SAT} profile . Therefore , after pinch off region for Id is consistent for rise in Vds for each Vgs.

The Figure 2.5 depicts the gate voltage control on Ids for different Vds.

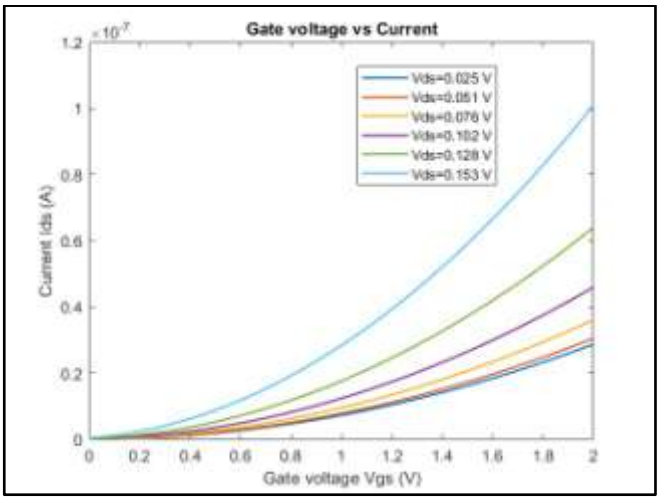


Figure 2.5:- Gate Voltage Vds vs. Drain Current Ids.

Ballistic GNRFETs are ambipolar (figure 2.6), hence Ioff does not naturally happen when Vgs= 0. To produce appropriate I-V curves for PMOS and NMOS transistors, voltage shifting is necessary. For the GNRFET to have the correct I-V characteristics for P-type and N-type transistors, the I-V curve must be shifted. The shifting quantity should ideally be 1/2Vds because Ioff happens when Vgs=1/2Vds depicted in figure2.6.

In figure 2.6 , for Vds =0.025 V and Vds= 0.051 V , the log(Id)-Vgs response for ambipolar behavior is obtained. Regarding the ambipolar behavior in GNRFET, this occurs because graphene has a zero bandgap, which means that the conduction and valence bands touch at the Dirac point. Therefore, the Fermi level can be easily tuned by the gate voltage, allowing both electrons and holes to contribute to the current. This results in the log(Id)-Vg curve showing ambipolar behavior, with both negative and positive gate voltages causing an increase in current, though with different polarity. But , for higher Vds , the log(Id)-Vg curve is shifted to left side slightly which means a leftward shift of the log(Id)-Vg curve for higher drain voltages in GNRFET would be an undesirable effect, indicating a degradation of the device performance.

However , the diffusive MOSFETs can be turned on or off more effectively than ballistic GNRFETs because they have a higher Ion/Ioff ratio. However, GNRFETs have the ability to show a higher Ion following I-V curve shifting. Because of these factors, MOSFETs are typically thought to be more suited for use in digital circuit applications.Moreover, the inherent graphene-metal contact resistance limits the performance of MOSFET circuits. In contrast to MOSFET circuits, the absence of graphene-metal contact resistance in GNRFET circuits may be advantageous. Additionally, process variation may affect the doping level in MOSFET reservoirs, which could have a significant impact on the characteristics of the transistors and the performance of the circuit. GNRFETs don't have this issue because their drain and source are made of undoped metal.

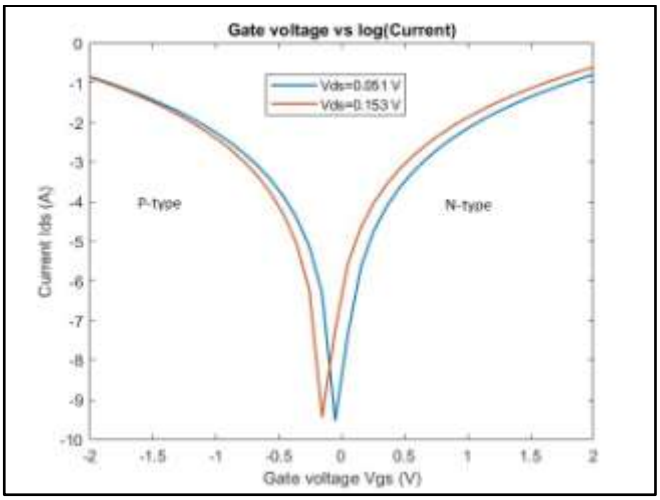


Figure 2.6:-Log(Id) vs. Vgs characteristics for both N and P type GNRFET.

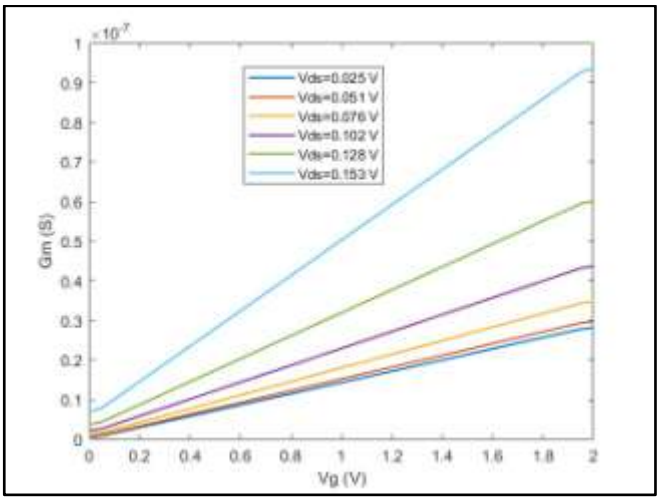


Figure 2.7:- Gate Voltage Vds vs Transconductance Gm.

Subthreshold regime of I_D - V_G curves and Transmission

As per the Figure 2.8, In a ballistic GNRFET , the sub-threshold regime is defined by the onset of quantum tunneling, where the gate voltage is not sufficient to create an inversion layer, but it is sufficient to allow electrons to tunnel through the potential barrier that separates the source and drain.In the sub-threshold regime of a ballistic GNRFET, the drain current increases exponentially with the gate voltage, similar to a conventional GNRFET. However, the slope of the curve in the sub-threshold regime is steeper, due to the fact that the transport is ballistic, and the electrons do not experience scattering events as they move through the channel. The equation governing the drain current in the sub-threshold regime of a ballistic GNRFET is given by:

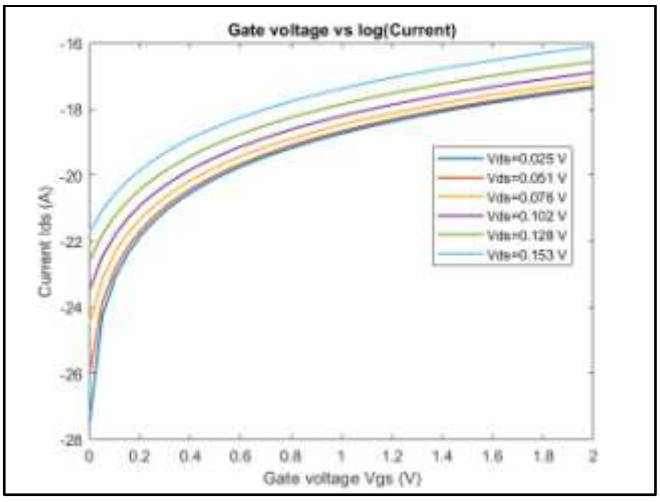


Figure 2.8:- Log(Id) vs. Vgs for various drain voltages Vds.

$I_d = I_{d0}\exp[(V_g - V_T) / (mV_T)]$(2.17)

where Id is the drain current, Id_0 is a constant, Vg is the gate voltage, VT is the thermal voltage, and m is the sub-threshold slope factor (body factor), which is different from the sub-threshold slope factor in a conventional GNRFET or diffusive MOSFET.

The graph in Figure 2.8 shows the logarithm of the drain current (log(Id)) as a function of the gate voltage (Vg) for different drain voltages (Vd) in a ballistic GNRFET. The graph shows that as the drain voltage increases, the drain current increases as well, but the slope of the curve in the sub-threshold regime remains the same. This indicates that the GNRFET is operating in the ballistic regime, as expected in a ballistic GNRFET. The steeper slope of the curve in the sub-threshold regime compared to a conventional GNRFET indicates better control over the transistor due to the absence of scattering events in the channel.

Furthermore, the slope of the curve in the sub-threshold regime is often expressed in terms of mV/decade, which is a measure of the rate at which the drain current changes with respect to the gate voltage on a logarithmic scale.In particular, mV/decade is the change in gate voltage required to produce a tenfold increase in the drain current (on a logarithmic scale). This quantity is often used as a measure of the sub-threshold slope of the transistor, which is an important parameter that determines the sensitivity and energy efficiency of the device.In the figure 2.8, the sub-threshold slope of the ballistic GNRFET is very steep, with a slope of about 60-80 mV/decade. This indicates that a small change in the gate voltage produces a large change in the drain current, which is desirable for applications that require high sensitivity and low power consumption.In conclusion, mV/decade is a measure of the sub-threshold slope of a transistor, which indicates the rate at which the drain current changes with respect to the gate voltage on a logarithmic scale. A smaller value of mV/decade indicates a steeper slope and better control over the transistor.

SS ~ 74.57 mV/dec (Calculated using obtained numerical values in Origin)

The transmission is calculated by using equation 2.11. The transmission for ballistic device is zero for energies below the band gap E_c and above bandgap it increases to one.The transmission coefficient T(E) represents the probability that an electron will pass through the GNRFET device. The transmission coefficient is influenced by the energy of the electron and the geometry of the device.The plateaus in the transmission vs energy graph of a ballistic GNRFET device represent energy levels at which the transmission coefficient remains constant regardless of changes in the electron energy.These plateaus occur because of the quantum confinement effect that occurs in narrow semiconductor channels, such as graphene nanoribbons. In these structures, the electrons are confined to a narrow region and can only take on certain allowed energy levels, which are determined by the geometry of the device.As the energy of the electrons changes, the transmission coefficient will typically vary as well, except at these specific energy levels where the transmission is completely or partially blocked. This results in plateaus in the Transmission vs. Energy graph.

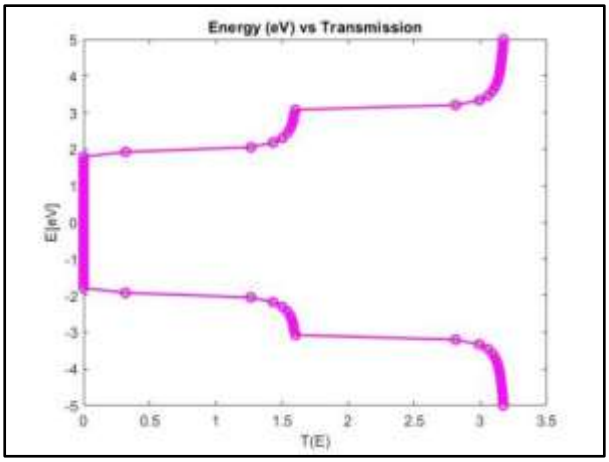


Figure 2.9:- Energy (eV) vs. Transmission.

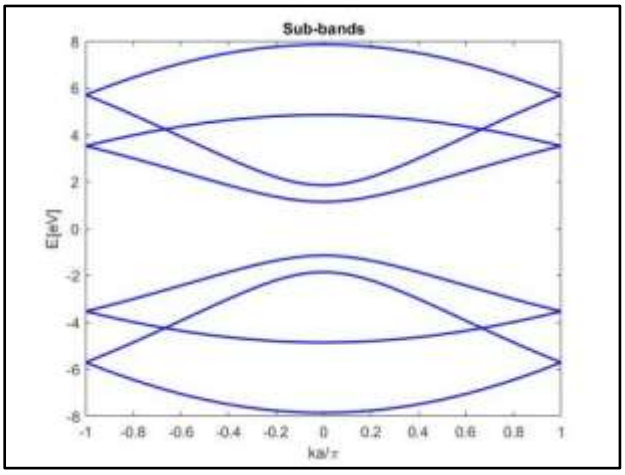


Figure 2.10:- Sub-bands.

Steep-sub threshold (SS) with Negative Capacitance FET

Modern digital circuit power consumption is typically described by $P = fCV_{DD}^2$, where 'C' is the output node capacitance, 'f' is the operating frequency, and 'V_{DD}' is the supply voltage.The Dennard scaling laws have led to a reduction in device size, which has resulted in decreased capacitance and decreased power consumption. The supply voltage V_{DD}, however, was unable to be scaled in accordance with the scaling principles, acting as a barrier to the development of high-speed, low-power devices. This is because the MOSFET's source and drain regions are subject to Boltzmann statistics, which set a 60 mV/dec restriction on the rate at which the drain current increases as a function of voltage.In order to achieve high frequency low power devices, the concept of negative capacitance in the gate stack of MOSFETs was proposed. SS = 60 mV/dec.

Sub-threshold swing (SS) in NCFET

The SS in NCFET can be obtained as inverse of slope of $\log(I_D)$ vs. Vg curve. The equation can be written as below ,

$$SS = \left[\frac{\partial logID}{\partial Vg}\right]^{-1} \quad (3.1)$$

SS in general can determine the heat generated in the switching process and thus , the power dissipation. Moreover, the SS could be written as,

$$SS = \frac{\partial Vg}{\partial logID} = \frac{\partial Vg}{\partial \psi s}\frac{\partial \psi s}{\partial logID} = m \times n \quad (3.2)$$

Whereas , the **m** is the body factor and **n** is the transport factor.

m > 1 because capacitance is always positive so SS > 60 mV/decade.

We require a minimum supply voltage of 360 mV to maintain an I_{ON}/I_{OFF} ratio of greater than 10^6 for acceptable noise margins. In order to get the SS relatively close to 60 mV/dec, electrostatics has been optimized to minimize m to unity. However, the novel FETs using alternative carrier methods, such as Tunnel FETs (TFETs) or Impact Ionization FETs (IIFETs), have been developed in order to push the boundaries of **n**.

However , rather than pushing the limits of **n** , we shall focus on **m** because The second term $\partial\psi s/\partial(\log_{10}I_D)$ linking the change in current to the change in surface potential in the channel cannot be less than 60 mV/decade at room temperature, according to standard FET analysis. Since the capacitive voltage divider connects Vg and ψs, it is clear that the first term $\partial Vg/\partial\psi s$ (the body factor "m") given in SS equation below,

SS = m x n = 60 mV/decade$(1+\frac{Cs}{Cins})$, C_s is the semiconductor capacitance ad C_{ins} is the insulator capacitance which could be oxide capacitance C_{ox}.

So, $\frac{\partial Vg}{\partial\psi s} = 1+\frac{Cs}{Cins}$which cannot be less than 1 since C_{ins} is positive capacitor. Even if thickness of the insulator is reduced while keeping high dielectric-κmaterial still we cannot achieve **m** less than 1. Therefore , to make body factor less than 1 the oxide (insulator) in FET could be replaced by the ferroelectric layer. The ferroelectric layer will introduce negative capacitance which in turn will be responsible for body factor less than 1.

Getting $\partial Vg/\partial\psi s$ less than 1 and a S value less than 60 mV/decade should be doable. The slope of P versus E (which is a scaled version of Q versus V) around the origin is negative, which explains why the ferroelectric capacitor is effectively a negative one.Normally, tests that show hysteretic jumps in the polarization do not directly observe this negative slope section because it is unstable. The negative capacitance segment can, however, be efficiently stabilized if the ferroelectric capacitor is connected in series with a regular capacitor, which allows the channel potential as on an internal node to alter more than the voltage Vg applied externally.The following positive feedback mechanism can be used to explain the negative capacitance in ferroelectrics.Consider a (positive) capacitor C_0 (per unit area), which experiences a terminal voltage $\alpha_f Q$ proportional to the charge on the capacitor Q (per unit area) in addition to the applied voltage V.

$C_{ins} = \frac{C0}{1-\alpha f C0}$ with condition $\alpha_f Q > 1$

This feedback voltage will lead to instability and charge will increase to very high value. So, to prevent this instability of charge increase we must have a series combination of the FE capacitor with ordinary capacitor C_s.

This series combination will make overall capacitance to be positive , $C_{eq} = [C_s^{-1} + C_{ins}^{-1}]^{-1}$. The C_{eq} will be greater than C_s and C_{ins}, and the composite ferroelectric-semiconductor system will function like a typical positive capacitor. As a result, for a given Q, the energy lost during switching will be substantially less than it would be if C_s and C_{ins} were both standard capacitors.

$\frac{\partial Vg}{\partial\psi s} = 1 - \frac{Cs}{C0}(\alpha_f C_0 - 1)$ (3.3)

Furthermore , the voltage across the FE capacitor could be written as follows ,

$V = F^{-1}(Q) - \alpha_f Q$(3.4)

We can expand the FE voltage equation 3.4 to the fifth power,

$V \approx \alpha_0 Q + \beta_0 Q^3 + \gamma_0 Q^5 + \rho_0 \frac{dQ}{dt}$(3.5)

Origin of Negative capacitance in NCFET

When C = dQ/dV > 0, If C < 0 as V↓, Q↑a capacitor has a negative capacitance (in a differential sense). This indicates that when charge across it rises, voltage across it falls. Therefore, when a negative capacitance device is utilized with a FET, a higher charge in the semiconductor would require a lower supply voltage.A different way to quantify capacitance is in terms of the free energy U. The energy landscape for a negative capacitor is an inverted parabola (see figure 3.1). $U = Q^2/2C$ for a linear capacitor. The capacitance can be defined as follows in terms of free energy.

$$C=[d^2U/dQ^2]^{-1} \quad (3.6)$$

The above relation is for capacitance of the non-linear dielectrics.

The figure 3.1 depicts the energy landscape of a ferroelectric material. Two degenerate energy minima exist in it. This indicates that even in the absence of an applied electric field, the ferroelectric material might produce a non-zero polarization. In general, the equation Q = εE +P, where is the ε ferroelectric's linear permittivity, E is the external electric field, and P is the polarization, can be used to express the total charge density in a particular material. P >> εE in common ferroelectric materials results in Q ≈ P.

By comparing the energy landscape of FE capacitor with oridinary capacitor . The curvature around Q = 0 of a ferroelectric is exactly the opposite of that of an ordinary capacitor, as seen by comparing the characteristic ferroelectric energy landscape with that of an ordinary capacitor in figure 3.1 and 3.2 . This opposite curvature already suggests a negative capacitance for the ferroelectric material near Q = 0, keeping in mind that the energy of an ordinary capacitor is supplied by ($Q^2/2C$). A ferroelectric substance could therefore offer a negative capacitance at this location.

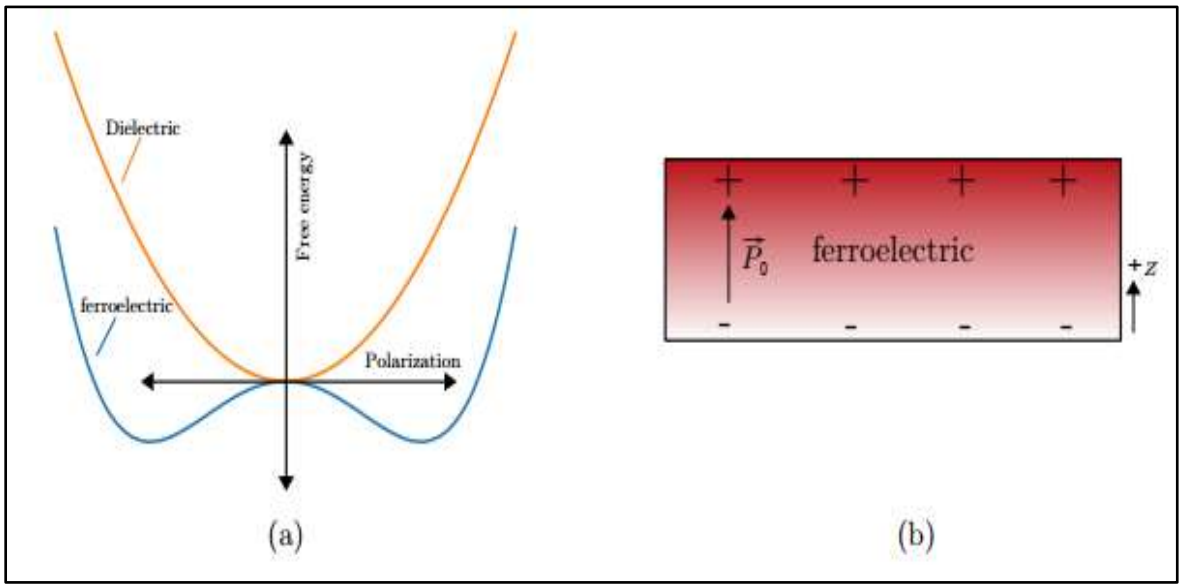


Figure 3.1:- Dielectrics and ferroelectrics (a) Ferroelectric's free energies, in contrast to dielectrics, exhibit a double well shape (b) Ferroelectrics features open surfaces and finite polarization.

Landau Theory of NC

The Landau theory, a symmetry-based phenomenology, connects known macroscopic occurrences to tiny models conceptually. It presumes that local fluctuations will be spatially averaged. In light of this, it is especially well suited to long-range interaction systems like ferroelectrics and superconductors. Landau highlighted that a system cannot transition smoothly between two phases with different symmetries in his seminal 1937 papers. The symmetry of one phase must be higher than that of the other because the thermodynamic states of two symmetry-distinct phases must coincide at their shared transition line.The transition was subsequently described by Landau in terms of an order parameter, a physical entry that starts off at zero in the high symmetry (disordered) phase and gradually increases to a finite value as the symmetry is dropped. The polarization order parameter and the high and low symmetry phases, which correspond to the paraelectric and ferroelectric states, respectively, are used to describe the ferroelectric-paraelectric transition. Then, only symmetry-compatible terms are maintained in the expanded free energy of U, which

is represented as a power series of the order parameter P. The system's state is then determined by obtaining the spontaneous polarization P by minimizing the free energy U(P) with regard to P.

Using Landau-Khalatnikov theory of non-linear dielectrics, the free energy of the non-linear dielectrics is give as ,

$U= \alpha P^2+ \beta P^4+ \gamma P^6- E_{ext}P$(3.7)

P is the polarization charge per unit area, and U is the Gibb's free energy calculated by adding the anisotropy energy an d the energy from the exter-nal field, $E_{ext.}$ α, β and γ are the constants.

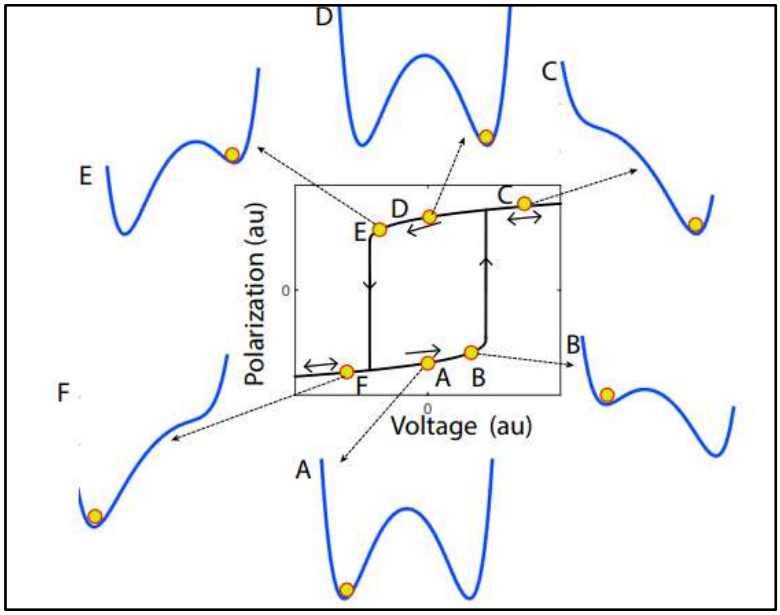


Figure 3.2:- Ferroelectric capacitor polarization-voltage hysteresis characteristics. Additionally, the energy landscapes at various locations along the hysteresis curve are displayed[11].

P-E curve shows negative slope and region where negative slope occurs is the region of the negative capacitance in FE material.

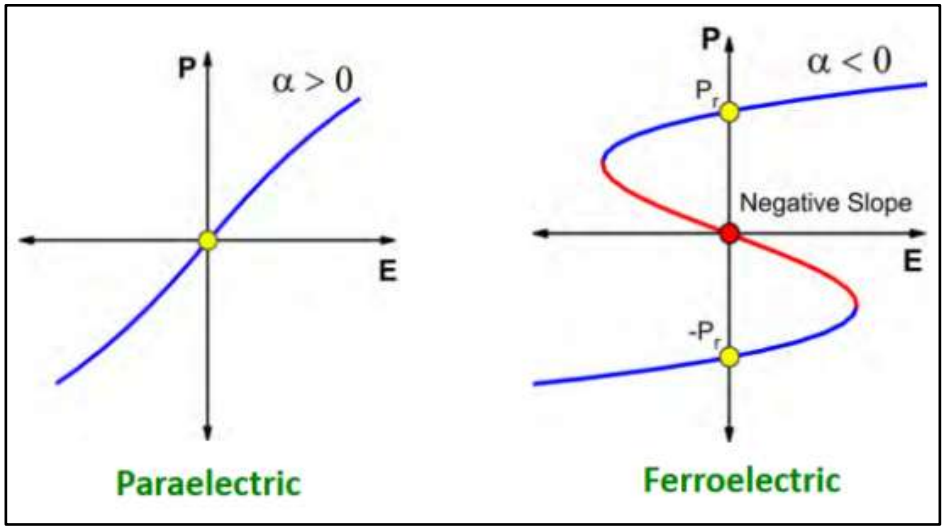


Figure 3.3:- P-E curves for Paraelectric and Ferroelectric Materials[11].

The negative slope in the P-E curve in ferroelectrics is the reason for negative capacitance.

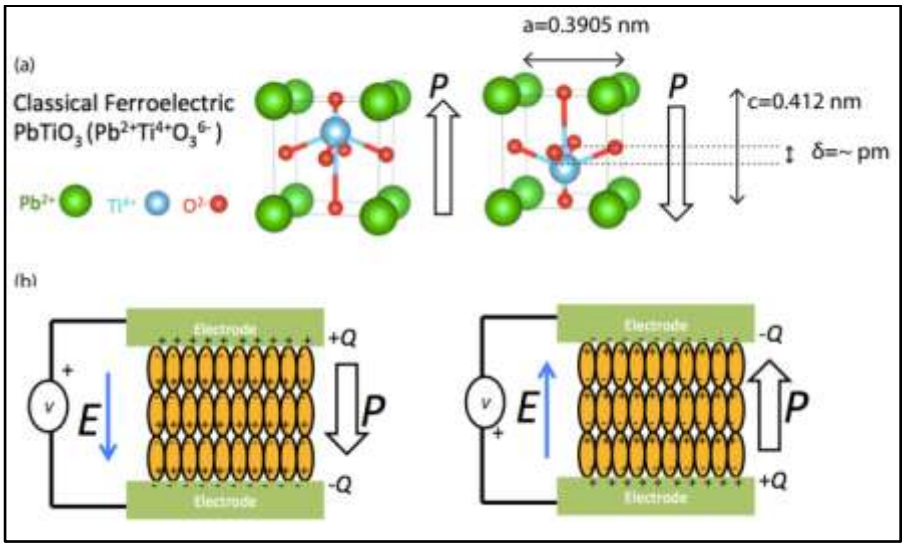


Figure 3.4:- (a) A PbTiO3 classical ferroelectric cell unit. The two different polarization states's respective corresponding opposite off-centerings of the central ion are displayed. (b) When a voltage greater than the coercive voltage is applied, a ferroelectric capacitor's polarization switches[11].

Relationship between C_s (Semiconductor capacitance) and t_{ins} (FE thickness)

At equilibrium, dU/dP = 0with equation 3.7, we obtain E_{ext} the field external to the FE capacitor.

$$E_{ext}=2\alpha P + 4\beta P^3+ 6\gamma P^5+ \rho\frac{dP}{dt} \quad (3.8)$$

Consider as per previous NC physics , Q =P and V =$E_{ext}t_{ins}$in the equation 3.8 . Then , we obtain $\alpha_0=2\alpha t_{ins}$, $\beta_0=4\beta t_{ins}$, $\gamma_0=6\gamma t_{ins}$, and $\rho_0=\rho t_{ins}$.

The channel potential ψs appears on C_s while $V=V_{FE}=Vg-\psi s$, the voltage across the FE. Both C_s and V_{FE} has same charge Q.

$$\psi s=Q/C_s$$

$$V_{FE}=Vg -\psi s = \alpha_0P^2+ \beta_0P^4+ \gamma_0P^6 +\rho_0\frac{dQ}{dt} \quad (3.9)$$

With steady-state condition $\frac{dQ}{dt} = 0$ (or $\frac{d\psi s}{dt} = 0$),

$$V_{FE} = Vg -\psi s = \alpha_0P^2+ \beta_0P^4+ \gamma_0P^6 \quad (3.10)$$

Combining equations 3.10 and $\psi s=Q/C_s$, set P=Q ,

$$Vg=(1+a_1)\psi s + a_2\psi s^3 + a_3\psi s^5 \quad (3.11)$$

Whereas , $a_1= 2\alpha C_s t_{ins}$, $a_2= 4\beta C_s^3 t_{ins}$, $a_3=6\gamma C_s^5 t_{ins}$

a_1 ,a_2 , a_3 are the coefficients corresponding to the material parameters α , β , γ respectively.

Usually , the non-linear terms a_2 and a_3 are discarded because they are very small compared to a_1.

Thus, $\frac{\partial Vg}{\partial \psi s} \approx 1 + a_1$

Given that we are ignoring the nonlinear terms, this is exactly the same as our prior finding in equation 3.3.The important thing to remember is that even if a_2 and a_3 are not negligible, $(1 + a_1)$ reflects the slope at the origin as 0, Vg= 0 and should be positive if we want the origin to represent a stable operating point. It is simple to observe that the ferroelectric insulator must have a thickness smaller than a critical thickness, which is denoted by eq. 3.11 for a_1.

$$t_{ins} \leq \frac{1}{2|a|Cs} \equiv t_c \quad (3.12)$$

The relationship between the C_s and t_{ins} shows that C_s has to be small and FE capacitor has to be large. This means the thickness of the FE capacitor will increase with C_s getting smaller. But , with smaller C_s , t_{ins} will go on increasing causing the steep slope in ψs vs Vg curves.Using the suitable values of α, β, and γ for $BaTiO_3$, in series with a linear capacitance C_s of 100 fF/m^2, figure 3.5 displays ψs against Vg for various values of t_{ins}. This data was derived directly from Eq. 3.11 .The plots are created by sweeping the voltage first toward the positive maximum and then back to the negative maximum, starting from a fully negatively polarized ferroelectric capacitor as the initial condition.

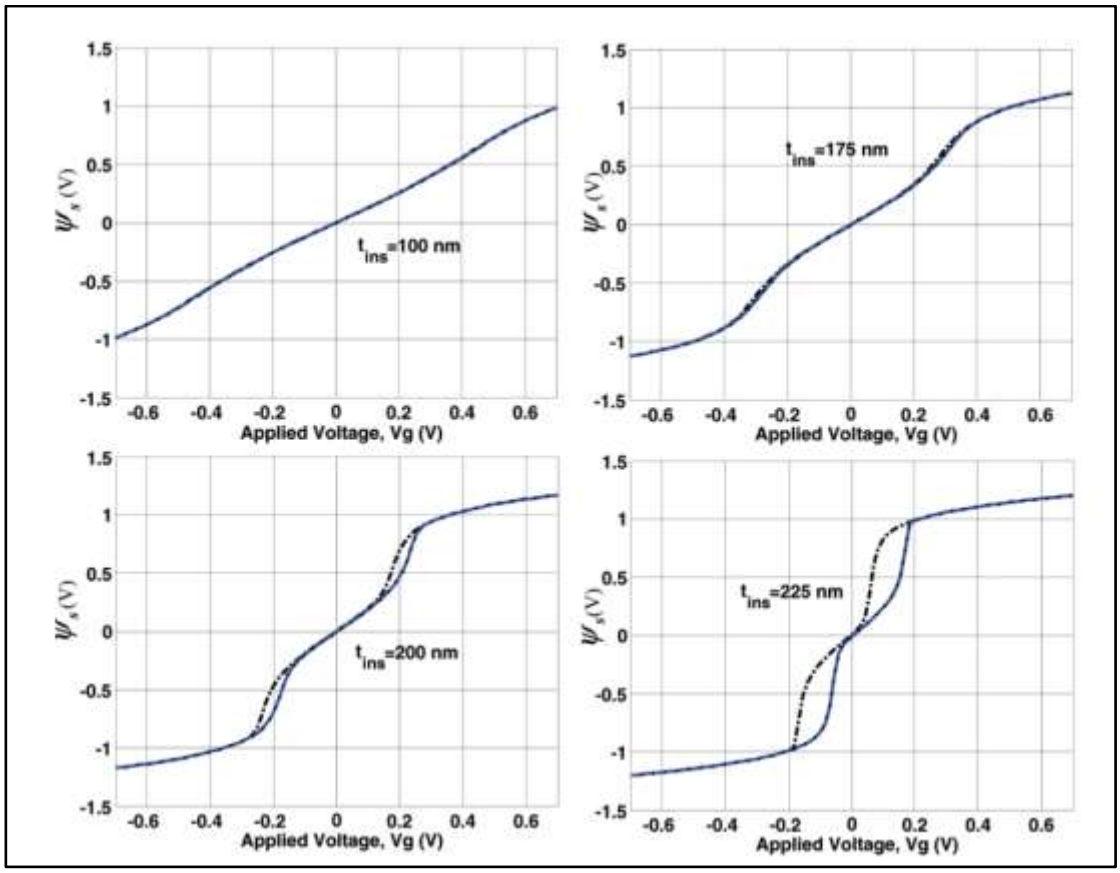


Figure 3.5:- The plots of the ψs vs Vg for various insulator thickness[8].

The black (dashed) curve depicts the sweep from positive to negative voltage, while the blue (solid) lines depict the sweep from negative to positive voltage. The slope of ψs vs Vg steepens with increasing insulator thickness, eventually beginning to open up a hysteresis. The ψs vs Vg graphs will no longer pass through the origin and resemble a typical ferroelectric hysteresis curve at an even greater thickness.

Introduction of Quantum capacitance into NCFET

In the research paper "Use of Negative Capacitance to Provide Voltage Amplification for Low Power Nanoscale Devices" by Sayeef Salahuddin and Supriyo Datta , the C_s is said to be smaller and C_{ins} is said to be larger. However , the semiconductor capacitance C_s is composed of quantum capacitance C_Q along with depletion and trap capacitance i.e C_{dep} and C_{trap} respectively.

Electrostatic (or capacitive) manipulation of the potential of the channel via which current is carried out is used to switch FETs. All pertinent capacitors are shown in a basic NC-FET arrangement in Figure 3.6(a) . If not stated otherwise, the discussion in this work is predicated on an n-type device. The capacitances of the NC layer and gate oxide, respectively, are C_{NC} and C_{OX}. Gate voltage divider capacitance $C_{div,}$ also known as C_Q, C_{trap}, C_{dep}, and $C_{s/d,geo}$, stand for quantum, interface trap induced, depletion, and source/drain geometrical capacitance, respectively.

$$C_{div} = C_Q + C_{trap} + C_{dep} + C_{s,geo} + C_{d,geo} \quad (3.13)$$

C_Q is the modulation of the electron charge density with respect to the channel potential.

$$Qe = q\sum_i \int_{Ec,i}^{\infty} dE \frac{DOSi}{1+exp\,(E-Ef)/kT} = qkT\sum_i DOSi[\ln(1+\exp(E_f-E_{c,i})/kT)] \quad (3.14)$$

Q_e is the electron charge density, $E_{c,i}$ is the energy level of the i^{th} mode, Q is the fundamental charge, i is the index of conduction modes, DOS_i is the density-of-states of the i^{th} mode, and E_f is the Fermi level. Below threshold, C_Q approaches zero, whereas beyond threshold, it approaches $q^2\Sigma_i DOS_i$. The trap state filling induced capacitance is known as C trap ($\sim dQ_{trap}/d\Phi_{Ch}$). The depletion layer's geometric capacitance is known as C_{dep}. The source and drain capacitive connection is where $C_{s/d,geo}$ originates. This capacitance also records the influence of the fringing field along the path inside the device during the short-channel state, from source/drain toward gate. $C_{frin,s/d-g}$is the source/drain-to-gate fringing capacitance.

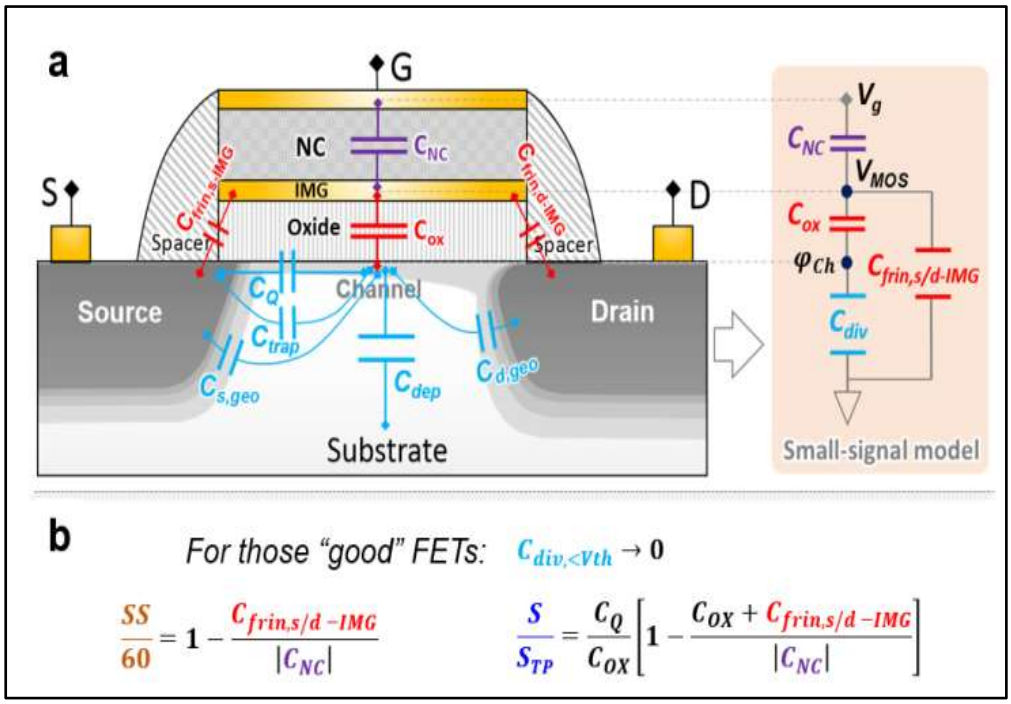


Figure 3.6:- (a) Capacitors with an internal metal gate (IMG) NC-FET.(b) S and SS formula[10].

A I_D-Vg of NCFET curve has a slope of,

$$S_{NC\text{-}FET} = \left(\frac{d[logId]}{dVg}\right)^{-1} = \frac{dVMOS}{d\Phi ch}\frac{dVg}{dVMOS}\frac{d\Phi ch}{d[logId]} \quad (3.15)$$

where V_{MOS} is the voltage at the interface between NC and oxide, V_g is the gate voltage, and Φ_{ch} is the channel potential.

Now , the drain current I_D of the NCFET can be obtained as ,

$I_D = W.v.Q_e$(3.16)

where W is the device's width, v is the carrier velocity,

As per the small-signal model of the NCFET shown in figure 3.6 (b),

$(\Delta V_{MOS} - \Delta\varphi_{Ch})\cdot C_{OX} = \Delta\varphi_{Ch}\cdot C_{div}$, $(\Delta V_g - \Delta V_{MOS})\cdot C_{NC} = -(\Delta Vg - \Delta V_{MOS})\cdot|C_{NC}| = \Delta V_{MOS}\cdot C_{MOS.}$

$$S_{NC\text{-}FET} = \left(1-\frac{CMOS}{|CNC|}\right)\left(1+\frac{Cdiv}{Cox}\right)\frac{IDd\Phi c\,h}{dID}\ln 10 \qquad (3.17)$$

Further derivation ,

$$S_{NC\text{-}FET} = \left(1+\frac{Cdiv}{Cox}\right)\left(1-\frac{1}{|CNC|}\frac{CdivCox}{Cdiv+Cox}\right)\frac{qkTDOSln\,(1+exp\,(Ef-Ec)/kT)d(-Ec/q)}{qkTDOSdln\,(1+exp\,(Ef-Ec)/kT)}\ln 10$$

$$=\left(1-\frac{1}{|CNC|}\frac{CdivCox}{Cdiv+Cox}+\frac{Cdiv}{Cox}-\frac{Cdiv}{Cox}\frac{1}{|CNC|}\frac{CdivCox}{Cdiv+Cox}\frac{(1+exp\,(Ef-Ec)/kT)ln\,(1+exp\,(Ef-Ec)/kT)}{exp\,(Ef-Ec)}\frac{kT}{q}\right)\ln 10$$

$$=\left(1-C_{div}\left[\frac{1}{|CNC|}-\frac{1}{Cox}\right]\right)S_{TP} \qquad (3.18)$$

Whereas ,

$S_{TP} = (1+\exp(Ec-Ef)/kT)\ln(1+\exp(Ef-Ec)/kT)60$

S_{TP} is the the thermionic transport mechanism limiting S.

In the near and above- threshold regimes, the C_Q begins to dominates and thus , $C_Q \approx C_{div}$.

Then , S becomes

$$S_{NC\text{-}FET} = \left(1-C_Q\left[\frac{1}{|CNC|}-\frac{1}{Cox}\right]\right)S_{TP} \qquad (3.19)$$

In the sub-threshold regime, C_Q is negligible and short-channel effects begins, $C_{div} \approx C_{SCE}$, S_{TP}=60. So, In the lowest conduction modes , S_{TP} is reduced to 60 because the $E_f << E_c$.

Thus, any steep-slope device pursuing small SS will engage in combat in the subthreshold regime, when Ef is significantly lower than Ec. As explained mathematically before , STP is constrained to 60 mV per decade at room temperature which is actually the "Boltzmann tyranny". Additionally, C_Q is really small.Then , the C_{MOS} is simplified to $C_{MOS,Vth} = C_{div,Vth}C_{OX}/(C_{div,Vth} + C_{OX})$, and C_{div} is reduced to $C_{div,Vth} = C_{trap} + C_{dep} + C_{s,geo} + C_{d,geo}$. S is transformed into SS as seen in Fig.3.7 , m is referred to as the body factor that captures the FET's under-NC layer gate efficiency, and A_V is the voltage gain that is brought about by the NC layer.
C_{MOS} and A_V can be expressed as ,

$$C_{MOS} = \frac{Cdiv\ .Cox}{Cdiv\ + Cox}$$

$$A_V = \frac{dVMOS}{dVg} = \frac{|CNC|}{|CNC|-CMOS\ ,<Vth}$$

The **m** and A_V defines the SS value.

When $|C_{NC}|$ exceeds C_{OX}, m triumphs and SS exceeds 60 mV per decade (shaded area on the left) in the figure 3.8(a). When $|C_{NC}|$ is less than C_{OX}, A_V rules SS and makes sub-60 SS possible. But C_{NC} shouldn't be any smaller than C_{MOS}. The total gate capacitance $C_g = |C_{NC}|C_{MOS,Vth}/(|C_{NC}| C_{MOS,Vth})$, $|C_{NC}|$ is negative, indicating that the metastable state of

the NC layer is not much stabilized by the FET structure below it. However, | $|C_{NC}|$ should not be smaller than $C_{MOS,Vth}$ because A_V and SS will become smaller than zero .

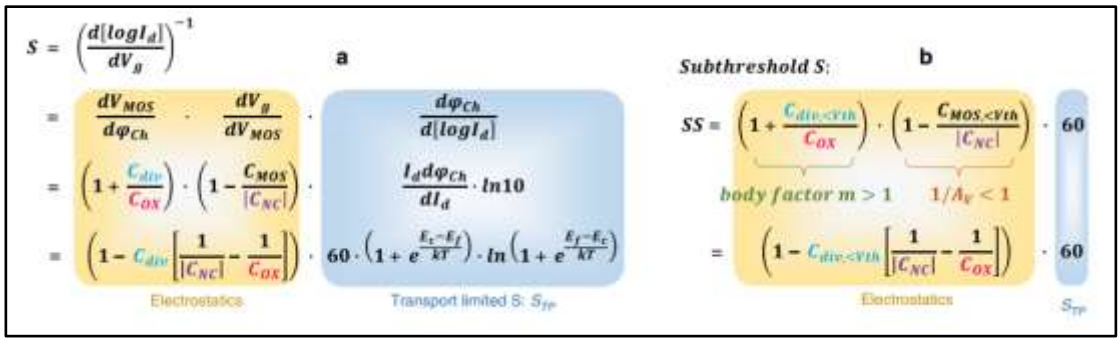


Figure 3.7:- Defining the I_D-V_g slope or swing (S) by separating the contributions from electrostatics and transport. (b) In the sub-threshold regime, S is reduced to sub-threshold S [10] .

Large Quantum capacitance may destroy the NCFET

As gate bias approaches and exceeds threshold, the channel electron density Q_e (or C_Q in Eq. (3.14)), which dramatically increases C_{div}, the slope of S/S_{TP} vs $1/|C_{NC}|$ (see the equation in Fig. 3.8(a), poses the biggest problem for NC-FETs.Increased C_{div} is similar to rotating the S/S_{TP} line in a clockwise direction with respect to the SS/60 line, as shown in Fig. 3.8(d). The lower bound of the forbidden hysteresis region (shaded region on the RHS of Fig. 3.8(d) of $1/|C_{NC}|$ gets extended from $1/C_{MOS,Vth}$ (= $1/C_{div,Vth}$ + $1/C_{OX}$) to $1/C_{MOS,>Vth}$ = ($1/C_{div,>Vth}$ + $1/C_{OX}$ = $1/C_Q$ + $1/C_{OX}$), which is very close to 1/ C_{OX}, since C_Q in In other words, C_Q drastically reduces the design space of NC.

Derivation of Minimum SS

Figure 3.8(c) shows how to get the least SS (SS_{min}) without hysteresis by matching $|C_{NC}|$ to $C_{MOS,>Vth}$, or by putting $|C_{NC}| = C_{MOS,>Vth}$ into the SS formula.

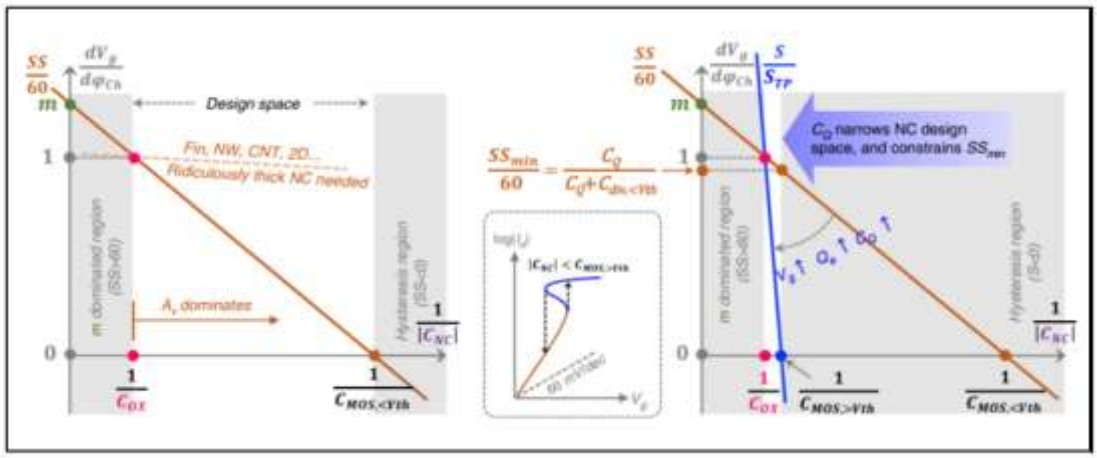
$$SS_{min} = (1 - C_{div,<Vth}[\frac{1}{|CNC|}-\frac{1}{Cox}])60 = (1-C_{div,<Vth}[\frac{1}{CMOS,>Vth}-\frac{1}{Cox}])60 \qquad (3.20)$$

Substituting $C_{MOS,>Vth} = C_{div,>Vth} \cdot C_{OX}/(C_{div,>Vth} + C_{OX})$

$$SS_{min} = (1-C_{div,<Vth}[\frac{1}{Cdiv,>Vth} + \frac{1}{Cox}-\frac{1}{Cox}])60 = (1-\frac{Cdiv,<Vth}{Cdiv,>Vth})60$$

Then , Substituting $C_{div,>Vth} = C_{div,<Vth} + C_Q$

$$SS_{min} = \frac{CQ}{CQ+Cdiv,<Vth}60 \qquad (3.21)$$



(a) (b) (c)

Figure 3.8:- (a) Design space for NC is quite large as C_Q is small enough (b) An example NC-FET's Id-Vg curve, where $1/|C_{NC}|$ is intended to be within the shaded area on the RHS (SS<0) (c) Design space for NC is narrowed as C_Q is large enough [10].

Application of IMG (Internal Metal Gate) to NCFET

It is important to note that, the P-E properties of the FE layer in the NC region are not absolutely linear, meaning that $|C_{NC}|$ does not remain constant with bias or charge density. The source of negative capacitance is shown by the fact that from equation 3.8 is negative. The nonlinearity terms (with and coefficients) start to divert the P-E relation from linearity when the electric field gets closer to the coercive field of FE.

In this paper, the coefficient is set to zero to streamline the analysis. Because of this, $|\varepsilon_{NC}|$ (and thus $|C_{NC}|$) rises with bias or charge density.

This non-linearity adds a damping component to the slopes of the SS/60 and S/S_{TP} lines vs the NC layer thickness T_{NC}. This damping term can be visualized as an anticlockwise rotation of the SS/S_{TP} and SS/60 lines, which results in a larger NC design space at the expense of a higher SS.

It has been suggested that an internal metal gate (IMG) inserted between the NC layer and the underlying MOSFET (figure 3.6(a)) . It has been discovered that IMG-induced fringing capacitance, which encourages more polarization in the NC layer, can enhance SS without hysteresis.Between the source/drain and the floating IMG, two more fringing (overlapping included) capacitors, $C_{frin,s/d\text{-}IMG}$, may be seen. For this structure, S and SS formula are calculated and reviewed.

From equation 3.17 , the modified $S_{NC\text{-}FET}$ would be ,

$$S_{NC\text{-}FET} = \left(1- \frac{CMOS + Cfring,s/d-IMG}{|CNC|}\right)\left(1+ \frac{Cdiv}{Cox}\right)\frac{IDd\,\Phi ch}{dID}\ln 10$$

$$=\left(1-\frac{CMOS + Cfring,s/d-IMG}{|CNC|}\right)\left(1+ \frac{Cdiv}{Cox}\right)S_{TP}$$

In the sub-threshold regime , $C_Q \rightarrow 0$, $C_{MOS} \rightarrow 0$, $C_Q/C_{MOS}=(C_{OX}+C_Q)/C_{OX} \rightarrow 1$, then above equation becomes

$$SS_{NC\text{-}FET} = \left(1 - \frac{Cfring,s/d-IMG}{|CNC|}\right)60 \quad (3.22)$$

In the above -threshold regime ,$C_Q >> C_{OX}$, $C_{MOS} \rightarrow C_{OX}$,

$S_{NC\text{-}FET} \rightarrow (1 - \frac{Cox + Cfring\ ,s/d-IMG}{|CNC|})S_{TP}$ (3.23)

Now , we will have SS_{min} without any hysteresis , the derived SS_{min} is

$SS_{min} = (\frac{Cox}{Cox + Cfring\ ,s/d-IMG})60$ (3.24)

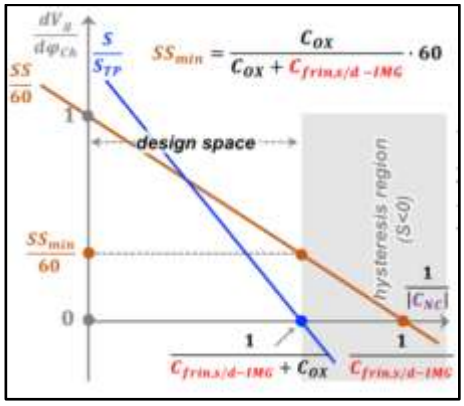


Figure 3.9:- Design Space of NC with IMG in NCFET[10].

IMG can unlock the polarization from C_Q and increase the NC design space for small SS_{min} by borrowing charge from fringing/overlap capacitance in the sub-threshold region.

Conclusions:-

The ballistic transport in Nanotransistor is being studied and compared with diffusive transport in conventional MOSFET. Using the self-consistent field procedure , we could obtain the current equation with transmission modes.
The GNRFET is modeled using NEGF in real space basis. The 4 atom unit cells of 1-D Nanoribbon is being used in the numerical model for 2D GNRFET by considering 2 unit cells in the transverse direction (along the width).The log(Id)-Vg transfer characteristics are plotted which are ambipolar due the significant changes in the I_{ON}/I_{OFF} of the ballistic GNRFET as compared to that in case of diffusive GNRFET or conventional MOSFET. The sub-threshold swing SS is obtained as 74.57 mV/dec which is in good agreement for the steeper slope and modulation of the drain current with respect to the gate voltage. The transmission of the electrons through the device determined as transmission coefficient affected by the number of modes in the channel. The model has 2 modes for 2 one dimensional Nanoribbons along the width of the device.This NEGF modelling provides a cogent understanding for high performance of the GNRFET in the digital circuits and application of the sub- threshold regime to modulate the drain current of the device for 2D low energy electronic devices.The ferroelectric layer above oxide in the FET has reduced SS below 60 mV/decade which is almost impossible to achieve in GNRFET. The NC physics for minimizing SS has been explained with derivation.This work not only clearly explains the interesting physics that prevents NC-FETs from obtaining steep slopes and invalidates all previous claims of "NC-FETs" in the literature, but it also reveals how NC in current FETs can reduce voltage loss. Consequently , the I_{ON}/I_{OFF} ratio is minimized to certain extent by increasing the switching speed of FET which is the building block for the low power 2D Quantum FETs.

Acknowledgment:-

I would like to express my special thanks totheElectrical department , IIT Bombay for their time and efforts that theyprovided . And the CNQT Lab in the Electrical department for the providing resources for the research work regarding Nano-transistor.

References:-

[1] Suriyo Datta , Lessons from Nanoelectronics: A New Perspective on Transport , World Scientific .
[2] Suriyo Datta, Quantum Transport: Atom to Transistor, CAMBRIDGE University press.
[3] Mark Lundstrom, Fundamentals of Nanotransistors , Lessons from Nanoscience: A Lecture Note Series, Purdue X.

[4] Anisur Rahman, Jing Guo, Supriyo Datta,Theory of Ballistic Nanotransistors,IEEE Transactions on Electron Devices, Vol. 50, No. 9, September 2003.
[5] Brajesh Rawat , Roy Paily , Modeling of graphene- based field-effect transistors through a 1-D real-space approach ,J Comput Electron (2018) 17:90–100.
[6] Ying-Yu Chen , Amit Sangai, Morteza Gholipour and Deming Chen, Graphene Nano-Ribbon Field-Effect Transistors as Future Low-Power Devices ,Symposium on Low Power Electronics and Design, 978-1-4799-1235-3/13/$31.00 ©2013 IEEE.
[7] Huei Chaeng Chin, Cheng Siong Lim, Weng Soon Wong,Kumeresan A. Danapalasingam, Vijay K. Arora,and Michael Loong Peng Tan, Enhanced Device and Circuit- Level Performance Benchmarking of Graphene Nanoribbon Field-Effect Transistor against a Nano-MOSFET with Interconnects, Hindawi Publishing Corporation , Journal of Nanomaterials,Volume 2014, Article ID 879813,
[8] Sayeef Salahuddin and Supriyo Datta, Use of Negative Capacitance to Provide Voltage Amplification for Low Power Nanoscale Devices, NANO Letters , Vol.8, No.2,405-410, 2008.
[9] Justin C. Wong and Sayeef Salahuddin, Negative Capacitance Transistors, IEEE Xplore Digital Library , 1558-2256, 2018.
[10] Wei Cao & Kaustav Banerjee , Is negative capacitance FET a steep-slope
logic switch? ,Nature Communications , (2020)11:196.
[11] Yogesh S. Chauhan , Modeling and Simulation of Negative Capacitance Transistors,Nanolab, Department of Electrical Engineering IIT Kanpur, India.