CHAPTER 2

Physical and Mathematical Setting for Simulation

2.1 Simulate Tool

The chapter 3 is employed by ISE TCAD 10.0 (ISE) which is supplied by National Center of High-Performance Computing (NCHC). The ISE simulation flows are as follows: (1) ISE simulation flow starts from process simulation and lead to electrical device characteristics (2) process flow operations are modeled in LIGAMENT or DIOS and result in device structure and doping distribution (3) MDRAW links process and device simulations providing an optimized high quality grid to DESSIS (4) the electrical characteristics are obtained solving potential and carrier transport equations with the device simulator DESSIS. The LIGAMENT, DIOS, MDRAW, and DESSIS are parts of tool in ISE. By directly drawing device structure, one can omit flows (1) and (2), and MDRAW supplies this function for users. All of the results in chapter 3 are just in device simulation but without in process simulation since ISE still has no germanium (Ge) relating process simulation.

2.2 Modeling

2.2.1 Energy Band Model

The detailed energy band diagrams of silicon (Si) and Ge are shown in Fig. 2.1 [2.1]. The energy band diagrams are frequently simplified when analyzing semiconductor devices. Since the electrons properties of a semiconductor are dominated by the highest partially empty band and the lowest partially filled band, it

is often sufficient to only consider those bands. This leads to a simplified energy band diagram for semiconductors as shown in Fig. 2.2. The energy band gap is located between the two lines, conduction band (E_c) and valence band (E_v), which are separated by the bandgap energy E_g. The distance between the E_c and the energy of a free electron outside the crystal (called the vacuum level labeled E_{vacuum}) is quantified by the electron affinity, χ multiplied with the electron charge q. Fig. 2.3 shows the comparison of Si and Ge energy band levels [2.2].

The E_g of semiconductors tends to decrease as the temperature is increased. This behavior can be better understood if one considers that the interatomic spacing increases when the amplitude of the atomic vibrations increases due to the increasing thermal energy. An increased interatomic spacing decreases the average potential seen by the electrons in the material, which in turn reduces the size of the E_g. The temperature dependence of the E_g has been experimentally determined yielding the following expression for E_g as a function of the temperature, T:

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T+\beta} \quad (2\text{-}1)$$

where $E_g(0)$, α and β are the fitting parameters. These fitting parameters are listed for Si and Ge in Tab. 2.1 [2.3]. The Fig. 2.4 shows the E_g with temperature increasing form 0K to 1000K.

2.2.2 Mobility Model

The mobility is computed in two steps. First, the low field mobility μ_{low} is determined according to section 2.2.2.1 constant mobility model. Second, the final mobility is computed from the formula in section 2.2.2.2 hydrodynamic Canali model or 2.2.2.3 driving force model. The two steps can be written as a function :

$$\mu = f(\mu_{low}, F) \quad (2\text{-}2)$$

In our simulations, we use hydrodynamic Canali model for electrons and driving force model for holes in second step calculating.

2.2.2.1 Constant Mobility Model

The constant mobility model assumes that carrier mobility is only affected by phonon scattering and, therefore, dependent on the lattice temperature. For the purpose of simulation one usually takes a simple power law whose coefficients are obtained by fitting experimental mobility values.

$$\mu_{low,n}{}^{L} = \mu_{low,n}{}^{0}\left(\frac{T}{300K}\right)^{-\alpha_n} \quad (2\text{-}3)$$

$$\mu_{low,p}{}^{L} = \mu_{low,p}{}^{0}\left(\frac{T}{300K}\right)^{-\alpha_p} \quad (2\text{-}4)$$

The published numerical values for the constants in (2-3), (2-4) show some scatter. The coefficients are summarized in Tab. 2.2 [2.4].

2.2.2.2 Hydrodynamic Canali Model

In this model, the driving force F is expressed in terms of the carrier thermal energy w_c. In a homogeneous and stationary situation, the hydrodynamic equations have the simple solution:

$$F_c = \sqrt{\frac{w_c - w_0}{\tau_{e,c} q \mu}} \quad (2\text{-}5)$$

Where $w_c = \frac{3k_B T_c}{2}$ is the average carrier thermal energy, $w_0 = \frac{3k_B T_L}{2}$ gives the equilibrium thermal energy, $\tau_{e,c}$ is the energy relaxation time, T_c denotes the carrier temperature, and T_L denotes the lattice temperature. The index c is e for electrons and h for holes. The Canali model is shown [2.5]:

$$\mu(F) = \frac{\mu_{low}}{\left[1 + \left(\frac{\mu_{low} F}{v_{sat}}\right)^{\beta}\right]^{1/\beta}} \tag{2-6}$$

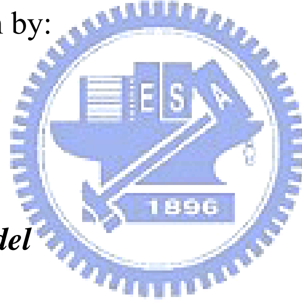
$$\beta = \beta_0 \left(\frac{T}{T_0}\right)^{\beta_{\exp}} \tag{2-7}$$

$$v_{sat} = v_{sat,0} \left(\frac{T_0}{T}\right)^{v_{sat,\exp}} \tag{2-8}$$

where T_0=300K, constants for (2-7) and (2-8) are listed in Tab. 2.3 and Tab. 2.4. Substituting F_c into the (2-6) and solving for μ yields the hydrodynamic Canali model:

$$\mu = \frac{\mu_{low}}{\left[\sqrt{1 + \alpha^2 (w_c - w_0)^{\beta}} + \alpha (w_c - w_0)^{\beta/2}\right]^{2/\beta}} \tag{2-9}$$

where the parameter α is given by:



$$\alpha = \frac{1}{2}\left(\frac{\mu_{low}}{q \tau_{e,c} v^2_{sat}}\right)^{\beta/2} \tag{2-10}$$

2.2.2.3 *Driving Force Model*

We use this model to set the driving field F which is given by the gradient of the Fermi potential φ_c:

$$F = \left|\nabla \varphi_c\right| \tag{2-11}$$

Then the F in (2-6) will be substituted by (2-11).

2.2.3 Impact Ionization Rate Model

Impact ionization is a generation/recombination modeling. We define α_n and α_p as the ionization rates for electrons and holes. They are defined as generated electron-hole pairs per unit length of travel and per electron and hole, respectively. The theoretical results for the ionization rate α_n, α_p are not unique. However, both

theoretical and experimental investigations indicate a good approximation to be an exponential dependence of the ionization rates upon the electric field component E in direction of current flow.

$$\alpha_n = \alpha_n^{\infty} \cdot \exp\left(-\left(\frac{E_n^{crit}}{E} \right)^{\beta_n} \right) \tag{2-12}$$

$$\alpha_p = \alpha_p^{\infty} \cdot \exp\left(-\left(\frac{E_p^{crit}}{E} \right)^{\beta_p} \right) \tag{2-13}$$

The exponents β_n, β_p are found in the range 1 to 2 [2.6-2.8]. Investigations by Baraff [2.9] have predicted that these theories can be interpreted as the two limiting cases of a much more rigorous model. For low fields Sockley's model is more appropriate, whereas for high fields Wolff's model is asymptotically correct. Baraff's results can, unfortunately, not be given in closed form. They have been obtained by a numerical solution of the Boltzmann transport equation, however, restricted to the assumption of an unrealistic band structure. However, a universal plot for both electrons and holes has been presented, which shows:

$$\alpha \cdot \lambda = f\left(\frac{E_r}{E_i}, \frac{E_i}{q \cdot \lambda \cdot E} \right) \tag{2-14}$$

λ is the mean free path between collisions with high energetic phonons; E_r is the average loss of energy defined per such collision; and E_i denotes the ionization energy, which is consumed from the ionizing carrier. Baraff's universal curves have been approximated with compact formulae so that an application for the purpose of simulation is facilitated. Okuto and Crowell [2.10] have proposed an empirical expression which is supposed to fit the theoretical results of Baraff as well as measurement.

$$\alpha = a_{300}\left(1+c\left(\frac{T}{K}-300\right)\right)\cdot E\cdot \exp\left(-\left(\frac{b_{300}\left(1+d\left(\frac{T}{K}-300\right)\right)}{E}\right)^2\right) \quad (2\text{-}15)$$

The coefficients of formula in (2-15), which is temperature dependent, are summarized in Tab. 2.5.

2.2.4 Thermal Conductivity Model

Most currently available models for the thermal conductivity of Si and Ge are based on the early measurements of Glassbrenner and Slack [2.11]. They presented theoretical investigations which led to the following formula for thermal conductivity in semiconductors.

$$k(T) = \frac{1}{a + bT + cT^2} \quad (2\text{-}16)$$

The constants a, b, and c are summarized for Si and Ge in Tab. 2.6. Above sections (2.2.1 to 2.2.4) have shown the models which are used in simulation, but there are still some parameters have set for Si and Ge as Tab. 2.7.

2.2.5 Hydrodynamic Model

With continued scaling into the deep submicron regime, neither internal nor external characteristics of sate-of-the-art semiconductor devices can be described properly using the conventional drift-diffusion transport model. In particular, the drift-diffusion approach cannot reproduce velocity overshoot and often overestimates the impact ionization generation rates. The Monte Carlo method for the solution of the Boltzmann kinetic equation is the most general approach, but because of its high computational requirements, it cannot be used for the routine simulation of devices. The hydrodynamic model provides a very good compromise. In the hydrodynamic

case, current densities are defined as:

$$\overrightarrow{J_n} = q\mu_n(n\nabla E_C + k_B T_n \nabla n + f_n^{td} k_B n \nabla T_n - 1.5 n k_B T_n \nabla \ln m_e) \quad (2\text{-}17)$$

$$\overrightarrow{J_p} = q\mu_p(p\nabla E_V + k_B T_p \nabla p - f_p^{td} k_B p \nabla T_p - 1.5 p k_B T_p \nabla \ln m_h) \quad (2\text{-}18)$$

where k_B is the Boltzmann's constant, T_n and T_p are carriers temperature, and f_n^{td} and f_p^{td} are parameters representing different approach in literature. The first term takes into account the contribution due to the spatial variations of electrostatic potential, electron affinity, and the band gap. The three remaining terms in (2-17) and (2-18) take into account the contribution due to the gradient of concentration, the carrier temperature gradients, and the spatial variation of the effective masses.

2.2.6 Capacitive Coupling Model

As the name of capacitive coupling model, we calculate the floating-gate (FG) voltage by using of capacitors between FG and others. Fig. 2.5 shows the cross section of this model. If we name the total capacitances as:

$$C_T = C_{IPD} + C_{TOX} + C_S + C_D \quad (2\text{-}19)$$

and we define $\alpha_J = \frac{C_J}{C_T}$ the coupling coefficient relative to the electrode J, where J can be CG, substrate, source, and drain. The potential on the FG due to capacitive coupling is given by:

$$V_{fg} = \alpha_g V_g + \alpha_b V_b + \alpha_d V_d + \alpha_s V_s \quad (2\text{-}20)$$

The higher coupling ratio, the lower CG operation voltage is needed to get the same speed.

2.3 Programming and Erasing Mechanisms

2.3.1 Fowler-Nordheim (F-N) Tunneling

2.3.1.1 Programming

One of the most important injection mechanisms used in nonvolatile memory is F-N tunneling. When a large control-gate (CG) voltage is applied during programming, its energy band structure will be influenced as shown in Fig. 2.6. In the figure, the applied CG voltage creates the electric field resulting in a potential barrier. This barrier provides a path for the electrons in the substrate to tunnel through the thin gate oxide (typically less than 12nm) and eventually be collected in the n^+-poly Si FG. The electrons collected at the FG lead to a tunneling current density which is given by the following equations:

$$J = \alpha E_{inj}^{2} \cdot \exp\left(-\frac{\beta}{E_{inj}}\right) \tag{2-21}$$

$$\alpha = \frac{q^3}{8\pi h \phi_b} \frac{m}{m^*} \tag{2-22}$$

$$\beta = 4\sqrt{2m^*} \frac{\phi_b^{\frac{3}{2}}}{3\hbar q} \tag{2-23}$$

where, E_{inj}=Electric field at the injection surface=$\frac{V_{app} - V_{fb}}{t_{ox}}$, V_{app}=Voltage applied across the tunneling oxide (TOX), V_{fb}=Flat band voltage, and t_{ox}=TOX thickness, $\hbar = \frac{h}{2\pi}$, h=Planck's constant, ϕ_b=Energy barrier at the injection surface, and m^*=Effective mass of an electron in the band gap of SiO_2. Fig. 2.7 shows a cross-section of a nonvolatile memory with electrons tunneling uniformly with CG voltage at positive potential while the source, the drain, and the substrate are at ground potential. An optional method that can be used to program Flash is given in Fig. 2.8 which is called drain-side tunneling. Drain-side tunneling is sometimes preferred over the uniform tunneling due to the programming speed as a result of

higher tunneling current density due to smaller injecting area. The simplified model in ISE is:

$$j_{FN} = AF^2 e^{-\frac{B}{F}} \tag{2-24}$$

where j_{FN} is the tunneling current density, F is the insulator electric field at the interface, and A and B are physical constants. The defaults coefficients used for F-N tunneling are listed in Tab. 2.8.

2.3.1.2 Erasing

F-N tunneling can also be used to erase a nonvolatile memory. One of the methods is by applying a large negative voltage at the CG. The energy band structure will be influenced as shown in Fig. 2.9. The applied CG voltage created the electric field resulting in a potential barrier. This barrier provides a path for the electrons to tunnel from the FG to the substrate through the thin gate oxide. Fig. 2.10 and Fig. 2.11 show two choices to erase a Flash. For uniform tunneling, a large negative CG voltage is applied while for drain-side tunneling method. In general, uniform tunneling is slower than drain-side tunneling, but, drain-side tunneling tends to cause reliability issues. The reliability issue is the gate oxide damage that occurs near the drain since a small area is bombarded by electrons and that the tunneling current density as a result of small area is higher. The simplified model in ISE is as (2-24), and the coefficients are listed in Tab. 2.9.

2.3.2 Hot Carriers Injection

The hot carriers refers to either electrons or holes that have gained very high kinetic energy after being accelerated by a strong electric field in areas of high field intensities within a semiconductor device. There are four commonly encountered hot

carrier injection mechanisms found: (1) the drain avalanche hot carrier (DAHC) injection (2) the channel hot electron (CHE) injection (3) the substrate hot electron (SHE) injection (4) the secondary generated hot electron (SGHE) injection.

The DAHC injection is said to produce the worst device degradation under normal operating temperature region. This occurs when a high voltage applied at the drain under non-saturated conditions (V_d>V_g) results in very high electric fields near the drain, which accelerate channel carriers into the drain's depletion region. Studies have shown that the worst effects occur when V_d=2V_g. The acceleration of the channel carriers causes them to collide with Si lattice atoms, creating dislodged electron-hole pairs in the process. This phenomenon is known as impact ionization, with some of the displaced electron-hole pairs also gaining enough energy to overcome the electric potential barrier between the Si substrate and the gate oxide. Under the influence of drain-gate field, hot carriers that surmount the substrate-gate oxide barrier get injected into the gate oxide layer where they are sometimes trapped. This hot carrier injection process occurs mainly in a narrow injection zone at the drain end of the device where the lateral field is at its maximum. Hot carriers can be trapped at the Si-SiO_2 interface or within the oxide itself. These charges shift some of the characteristics of the device, such as its threshold voltage (V_{th}) and its conveyed conductance (g_m). Injected carriers that do not get trapped in the gate oxide become gate current. Excessive substrate current may therefore be an indication of hot carrier degradation. The Fig. 2.12 shows the DAHC injection cross-section view.

The CHE injection occurs when both the gate voltage and the drain voltage are significantly higher than the source voltage, with V_g~V_d. Channel carriers that travel from the source to the drain are sometimes driven towards the gate oxide even before they reach the drain because of the high gate voltage. Fig. 2.13 shows the CHE cross-section view, and Fig. 2.14 shows the energy band diagram during hot-electrons

programming.

The SHE injection occurs when the substrate back bias is very positive or very negative, i.e., $|V_b| >> 0$. Under this condition, carriers of one type in the substrate are driven by the substrate field toward the Si-SiO2 interface. As they move toward the substrate-oxide interface, they further gain kinetic energy from the high field in surface depletion region. They eventually overcome the surface energy barrier and get injected into the gate oxide, where some of them are trapped. The Fig. 2.15 shows the SHE injection cross-section view.

SGHE injection involves the generation of hot carriers from impact ionization involving a secondary carrier that was likewise created by an earlier incident of impact ionization. This occurs under conditions similar to DAHC, i.e., the applied voltage at the drain is high or $V_d > V_g$, which is the driving condition for impact ionization. The main difference, however, is the influence of the substrate's back bias in the hot carrier generation. This back bias results in a field that tends to drive the hot carriers generated by the secondary carriers toward the surface region, where they further gain kinetic energy to overcome the surface energy barrier. The Fig. 2.16 shows the SGHE injection involves hot carriers generated by secondary carriers.

In our simulation during programming the Flash cell, it is going through the DAHC and CHE. In ISE, the total hot carrier injection current according to the Fiegna model [2.12] can be written as:

$$I_g = q\int P_{ins}\left(\int_{E_{B0}}^{\infty} v_{\perp}(\varepsilon) f(\varepsilon) g(\varepsilon) d\varepsilon\right) ds \qquad (2\text{-}25)$$

where ε is the electron energy, E_{B0} is the height of the semiconductor-insulator barrier, $v_{\perp}$ is the velocity normal to the interface, f(ε) is the electron energy distribution, g(ε) is the density of states of the electrons, P_{ins} is the probability of scattering in the image

force potential well as described by (2-29), and ∫ds is an integral along the semiconductor-insulator interface. The following expression for the electron energy distribution was proposed for a parabolic and an isotropic band structure, and equilibrium between lattice and electrons:

$$f(\varepsilon) = A \cdot \exp\left(-\chi \frac{\varepsilon^3}{F^{1.5}{}_{eff}}\right) \tag{2-26}$$

therefore, the gate current can be rewritten as:

$$I_g = q\frac{A}{3\chi}\int P_{ins} n \frac{F^{1.5}{}_{eff}}{\sqrt{E_B}} e^{\frac{\chi E^3{}_B}{F^{1.5}{}_{eff}}} ds \tag{2-27}$$

where n is the electron density and F_{eff} is an effective field, and E_B is shown as insulator field F_{ins} function:

$$E_B = \begin{cases} E_{B0} - \alpha |F_{ins}|^{\frac{1}{2}} - \beta |F_{ins}|^{\frac{2}{3}} & F_{ins} < 0 \\ E_{B0} - \alpha |F_{ins}|^{\frac{1}{2}} - \beta |F_{ins}|^{\frac{2}{3}} + V_{ins} & F_{ins} > 0 \end{cases} \tag{2-28}$$

the E_B is the Si-SiO_2 barrier height, E_{B0} is the zero field barrier height at the semiconductor-insulator interface, the second term is the equation represents barrier lowering due to image potential, and the third term of the barrier lowering is due to the tunneling processes. All hot carrier models contain a probability P_{ins} of scattering in the image force potential well:

$$P_{ins} = \exp\left(-\frac{x_0}{\lambda_{ins}}\right) \tag{2-29}$$

where λ_{ins} is the scattering mean free path in the insulator and the distance x_0 is given as:

$$x_0 = \sqrt{\frac{q}{16\pi\varepsilon_{ins}\varepsilon_0 F_{ins}}} \tag{2-30}$$

in the above expression, ε_{ins} is the dielectric constant of the insulator. The coefficients and their defaults are given in Tab. 2. 10.

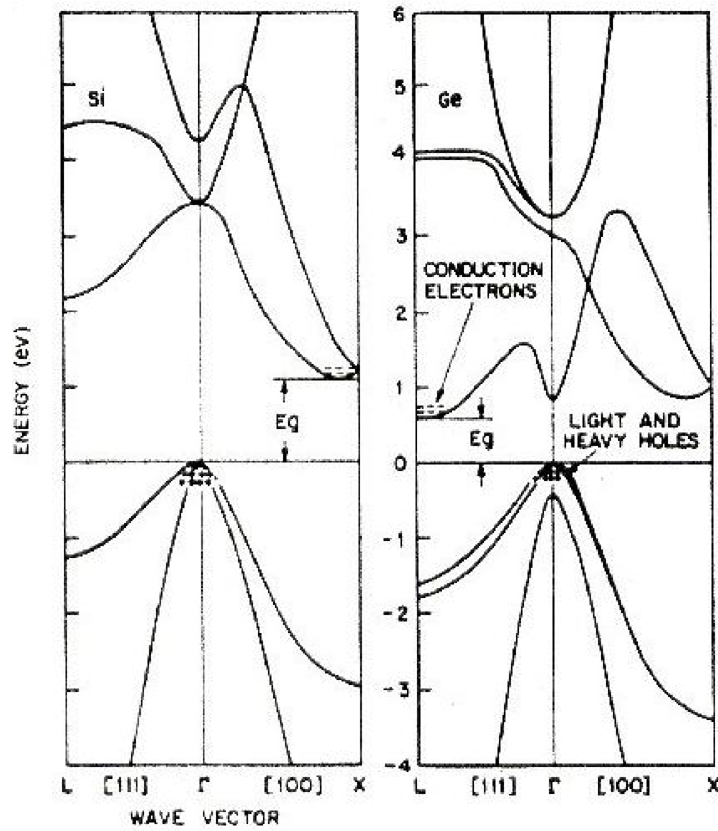


Fig. 2.1 Energy band diagram of Si and Ge.

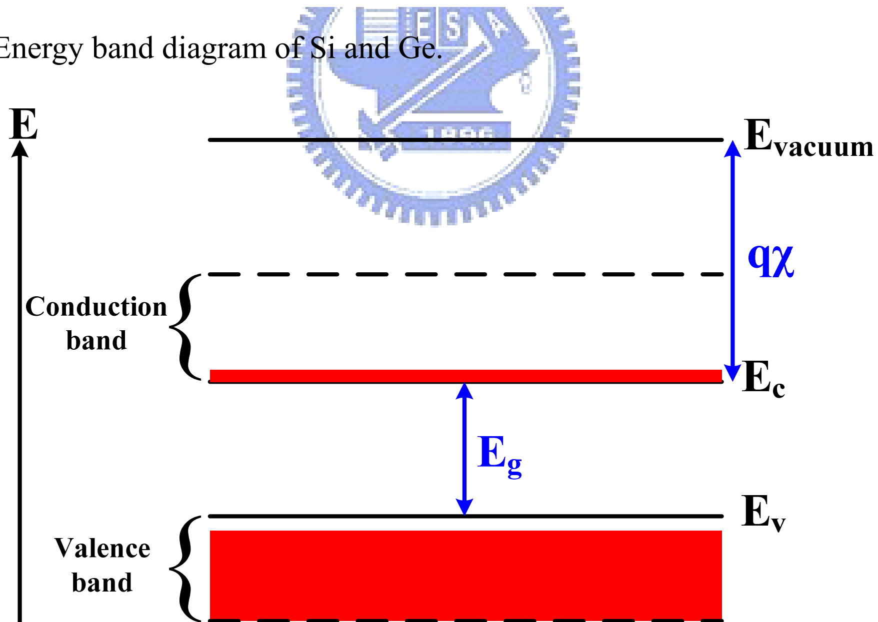


Fig. 2.2 A simplified energy band diagram used to describe semiconductor. Shown are the valence and conduction band as indicated by the valence band edge, E_v, and the conduction band edge, E_c. The vacuum level, E_{vacuum}, and the electron affinity, χ, are also indicated on the figure.

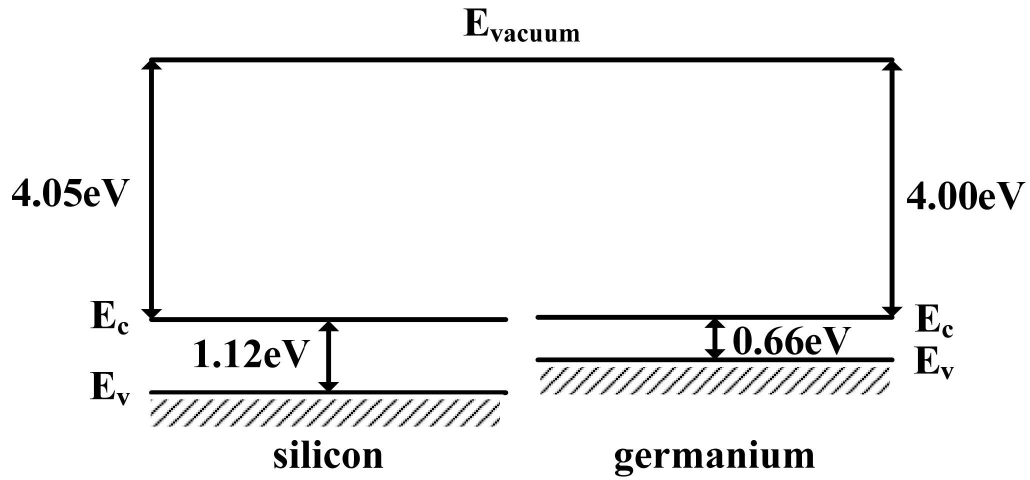


Fig. 2.3 Comparison of energy-band levels in Si and Ge.

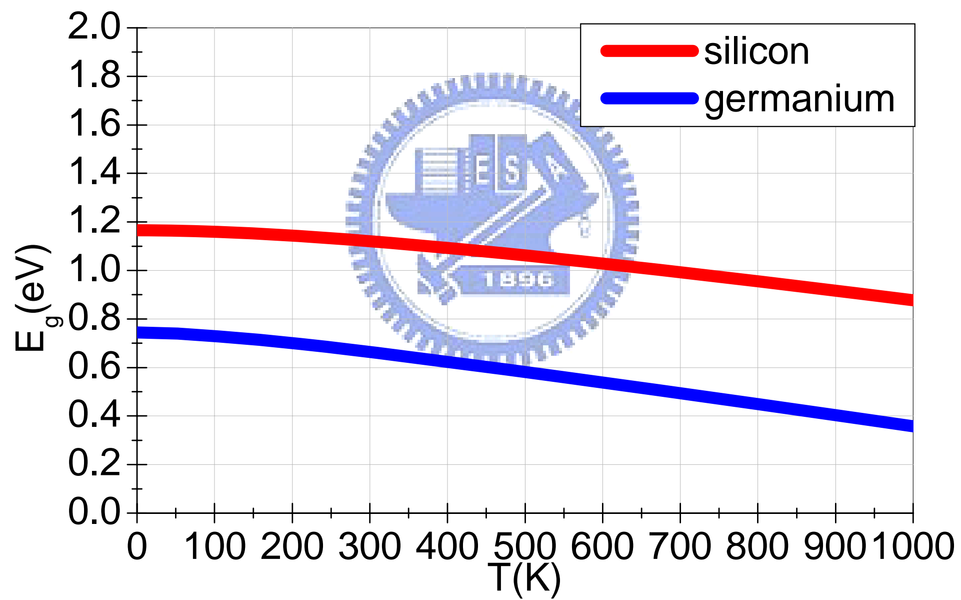


Fig. 2.4 Temperature dependence of the energy bandgap of Si and Ge.

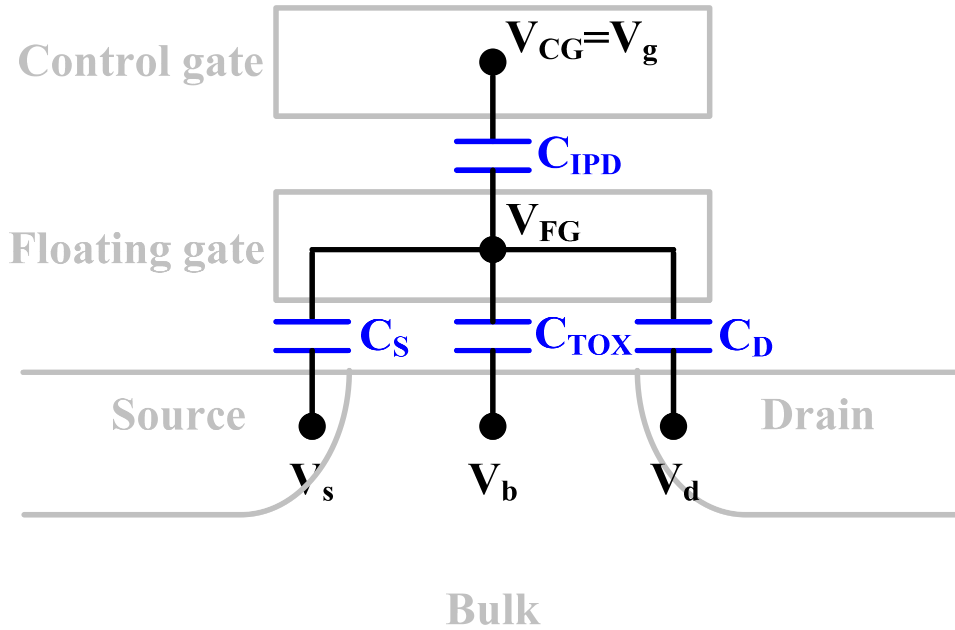


Fig. 2.5 The capacitive coupling model that using the capacitance between the FG and the other electrodes.

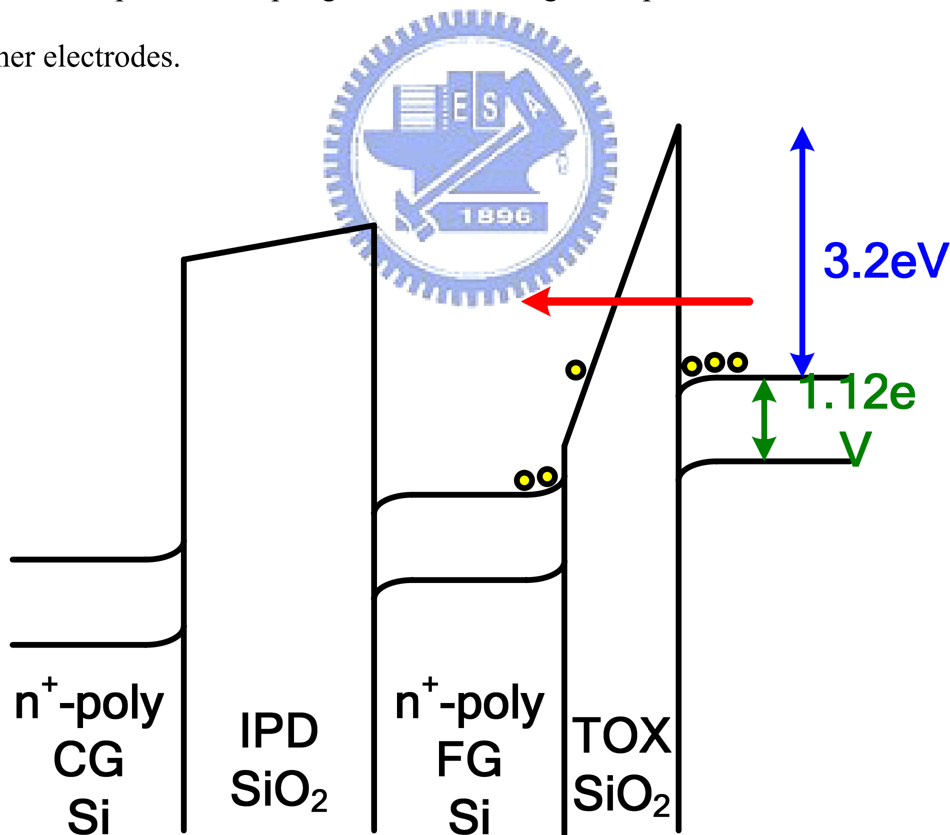


Fig. 2.6 Energy band diagram of a FG memory during programming by F-N tunneling.

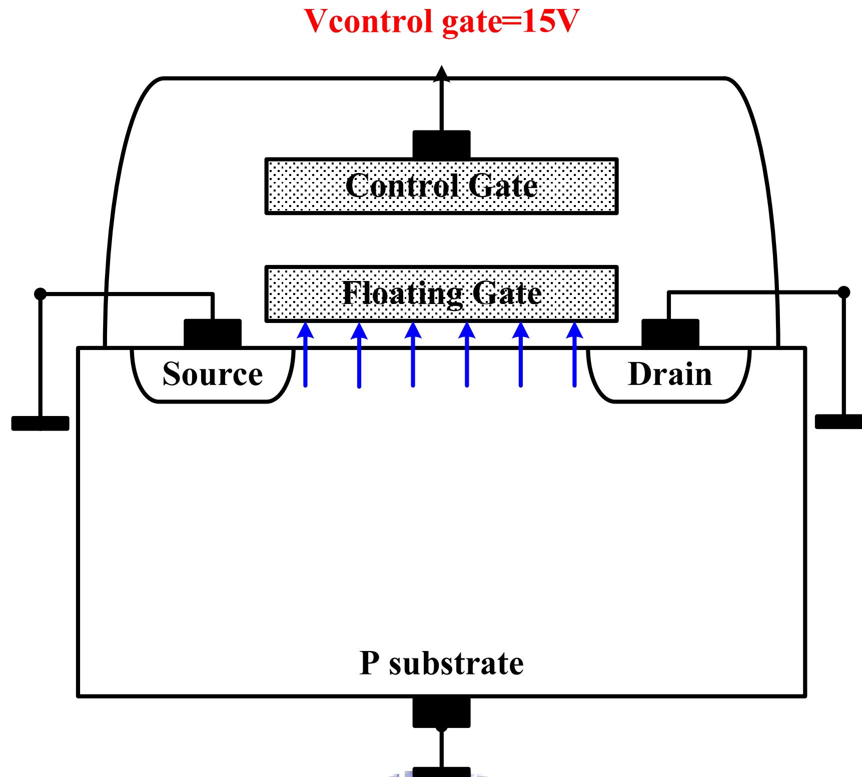


Fig. 2.7 A cross-section of a nonvolatile memory with electrons tunneling uniformly.

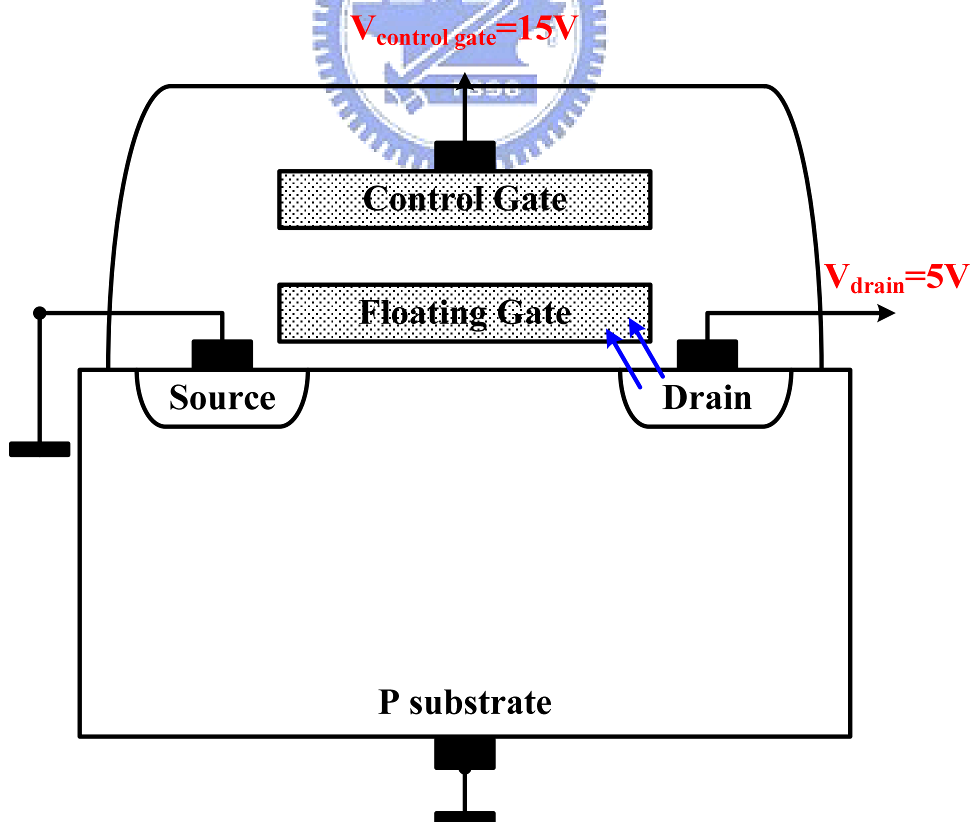


Fig. 2.8 Drain-side tunneling to program Flash.

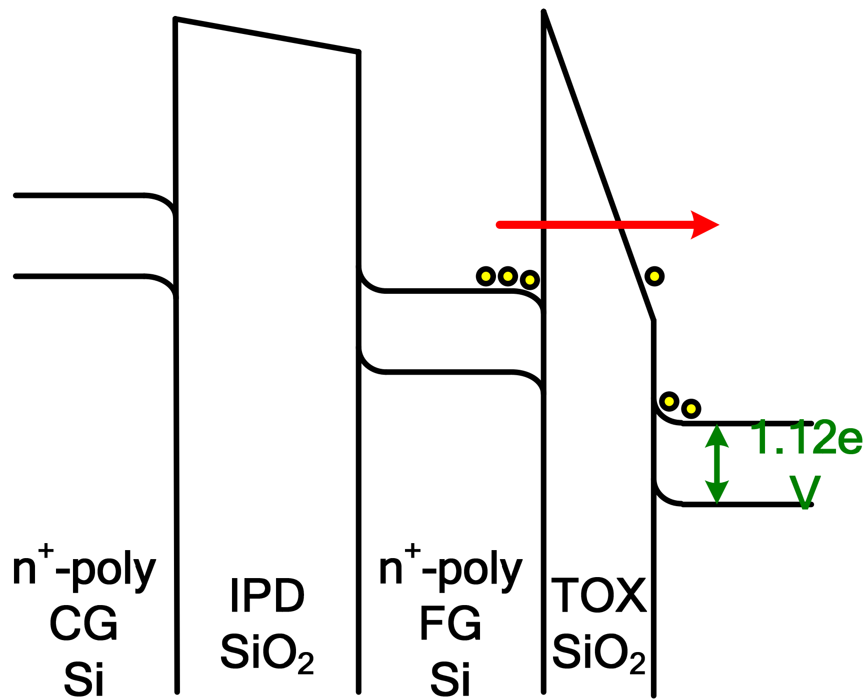


Fig. 2.9 Energy band diagram of a FG memory during erasing by F-N tunneling.

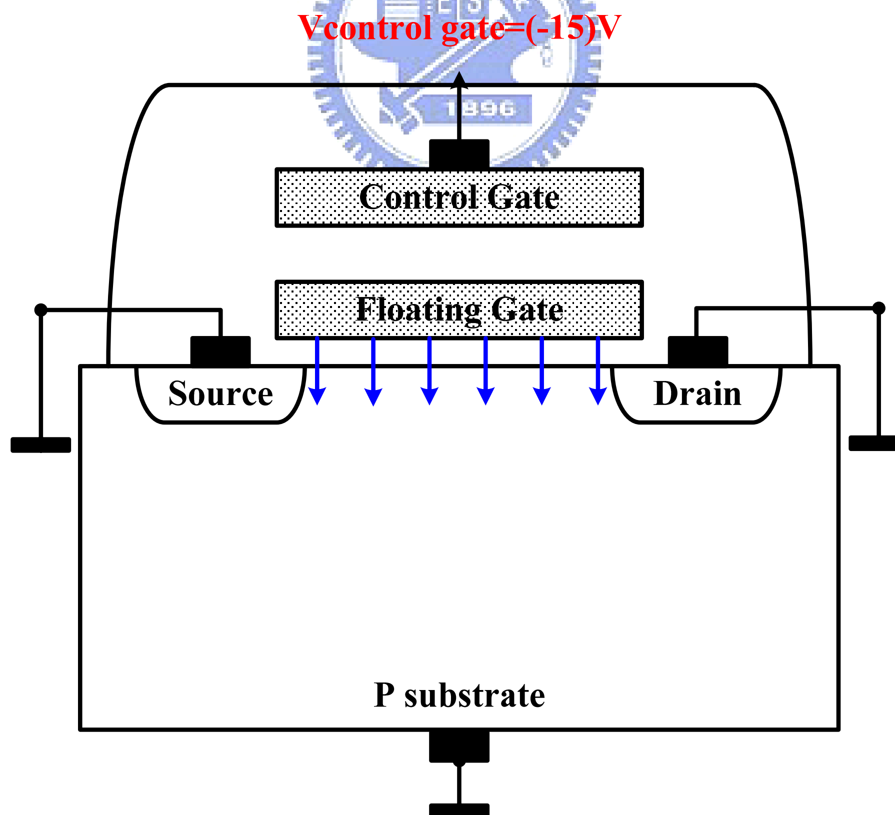


Fig. 2.10 Uniform tunneling to erase Flash.

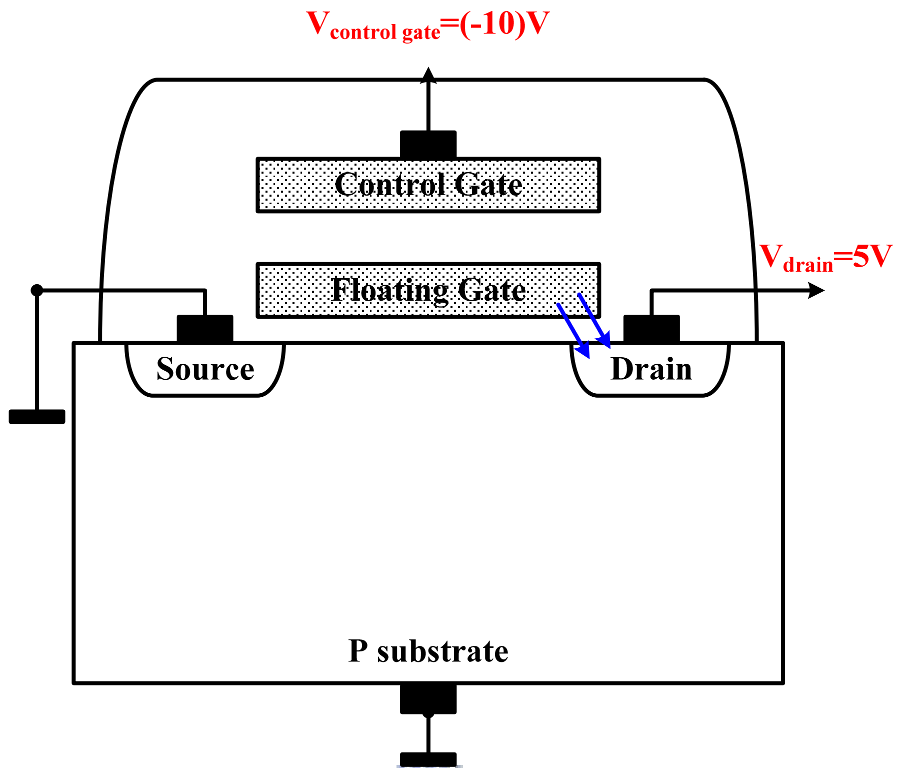


Fig. 2.11 Drain-side tunneling to erase Flash.

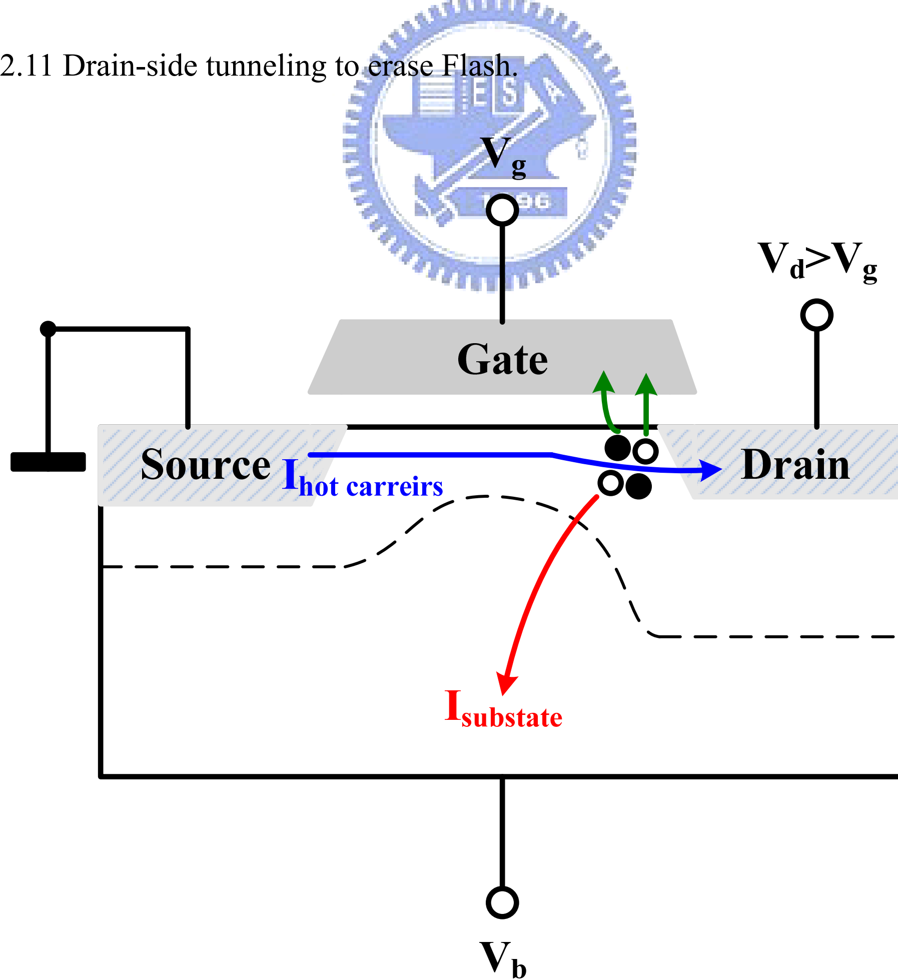


Fig. 2.12 DAHC injection involves impact ionization of carriers near the drain area.

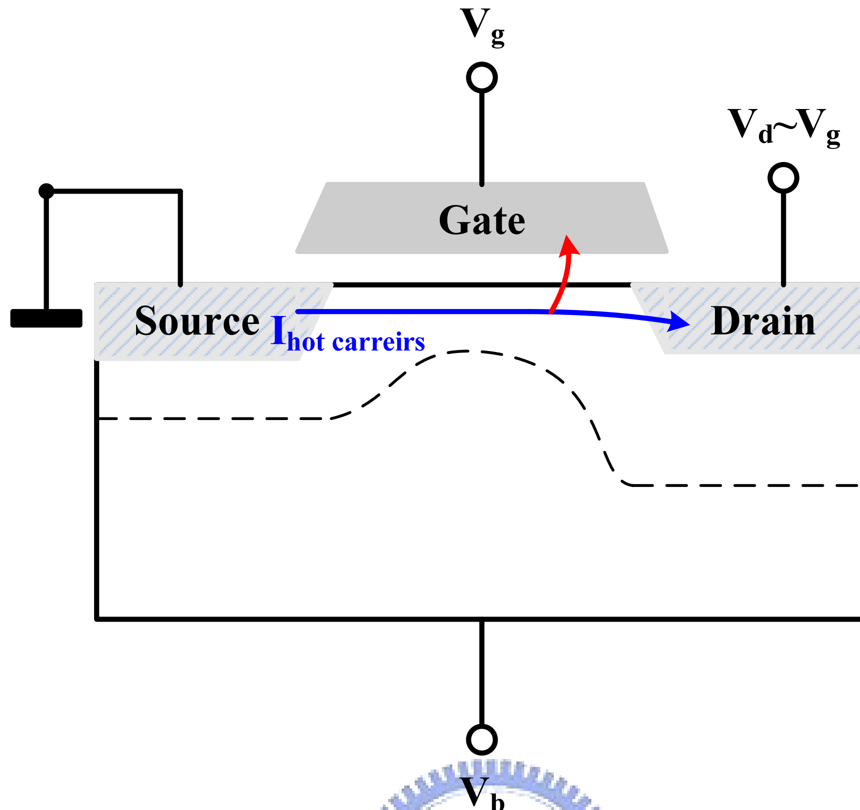


Fig. 2.13 CHE injection involves propelling of carriers in the channel toward the oxide even before they reach the drain area.

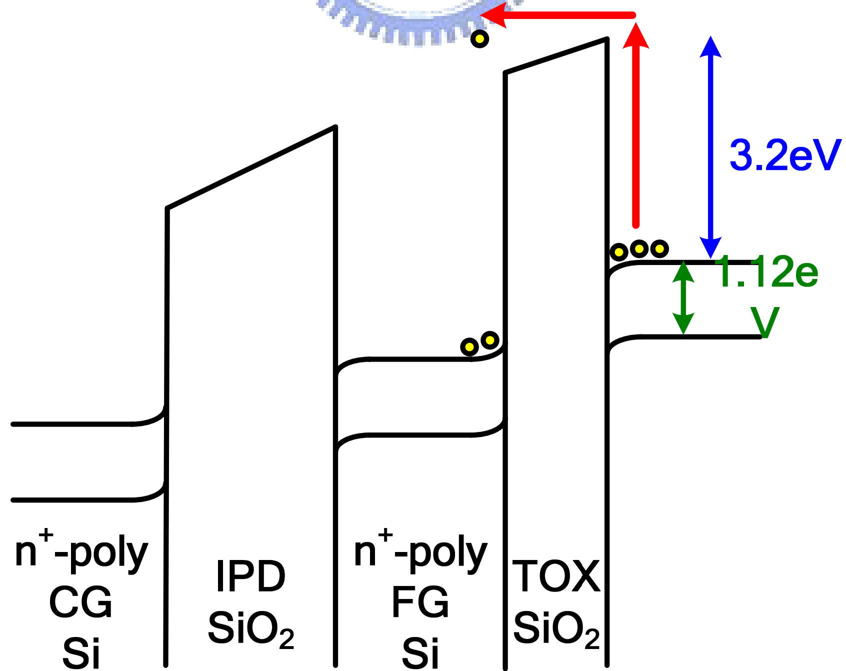


Fig. 2.14 Energy band diagram of a FG memory during programming by hot-electron injection.

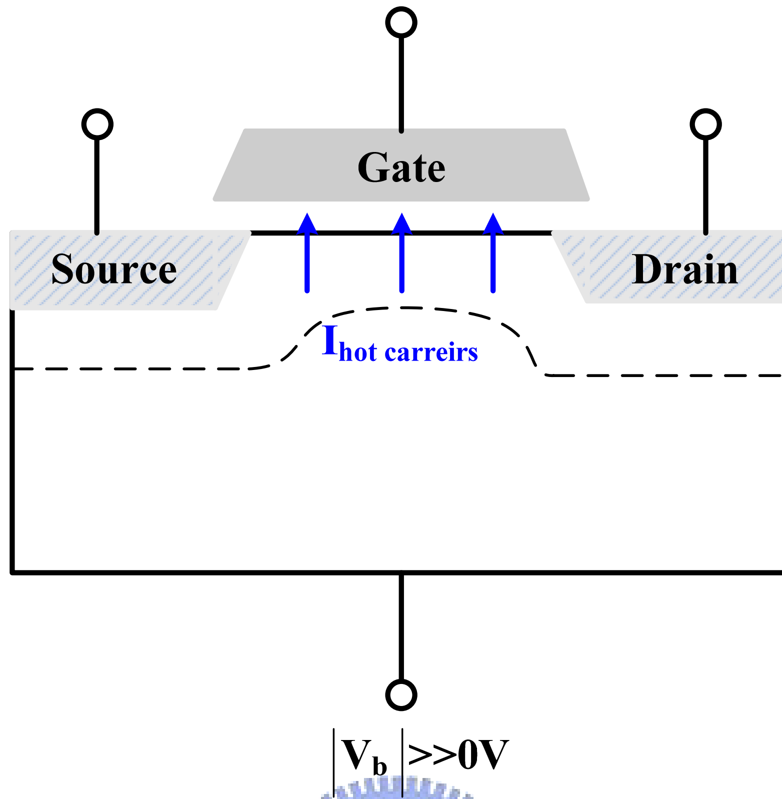


Fig. 2.15 SHE injection involves trapping of carriers from the substrate.

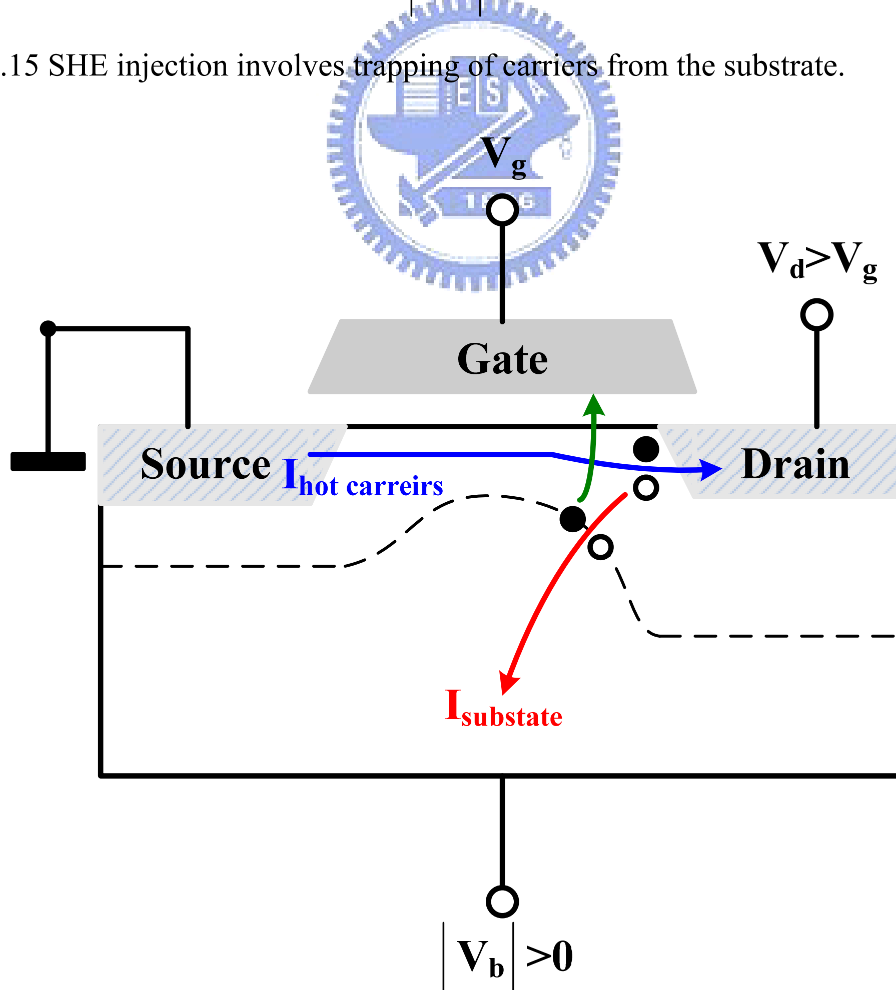


Fig. 2.16 SGHE injection involves hot carriers generated by secondary carriers.

	$E_g(0)$ (eV)	α (eV/K)	B (K)
Silicon	1.166	0.473×10^{-3}	636
germanium	0.744	0.477×10^{-3}	235

Tab. 2.1 Parameters used to calculate the energy bandgap of Si and Ge as a function of temperature.

	μ_n^0 ($cm^2/V\cdot s$)	α_n	μ_p^0 ($cm^2/V\cdot s$)	α_p
silicon	1417	2.5	470.5	2.2
germanium	3900	1.6	1900	2.3

Tab. 2.2 Lattice mobility constants.

		β_0	β_{exp}
silicon	electrons	1.109	0.66
germanium		1.109	0.66

Tab. 2.3 Canali model parameters.

		$v_{sat,0}$ (cm/s)	$v_{sat,exp}$
silicon	Electrons	1.07×10^7	0.87
germanium		7.43×10^6	0.87

Tab. 2.4 Default velocity saturation parameters.

		a_{300} (1/V)	b_{300} (V/cm)	c	d
silicon	electrons	0.426	4.81×10^{5}	3.05×10^{-4}	6.86×10^{-4}
	holes	0.243	6.53×10^{5}	5.35×10^{-4}	5.87×10^{-5}
germanium	electrons	0.569	3.32×10^{5}	6.33×10^{-4}	9.34×10^{-4}
	holes	0.559	2.72×10^{5}	7.87×10^{-4}	8.82×10^{-4}

Tab. 2.5 Coefficients for (2-15).

	a (cm·K/V·A)	b (cm/V·A)	c (cm/K·V·A)
silicon	0.03	1.56×10^{-3}	1.65×10^{-6}
germanium	0.17	3.95×10^{-3}	3.38×10^{-6}

Tab. 2.6 Coefficients for (2-16).

	silicon	germanium
relative epsilon: ε_r	11.7	16.1
refractive index: $n(T)=n_{300}\left[1+2\times10^{-4}(T-300)\right]$	n_{300}=3.45	n_{300}=1
$\frac{m_n^*}{m_0}$	1.18	0.55
$\frac{m_p^*}{m_0}$	0.5	0.3
crystal lattice constant ($\mathring{A}$)	5.43072	5.65754

Tab. 2.7 Parameters used in ISE beside mentioned above.

	A (program) (A/V^2)	B (program) (V/cm)
silicon	1.23×10^{-6}	2.37×10^{8}
Germanium	1.25×10^{-6}	2.71×10^{8}

Tab. 2.8 Coefficients for F-N tunneling to program in ISE.

	A (erase) (A/V^2)	B (erase) (V/cm)
silicon	1.82×10^{-7}	1.88×10^{8}
Germanium	1.84×10^{-7}	2.15×10^{8}

Tab. 2.9 Coefficients for F-N tunneling to erase in ISE.

	electrons	holes
A ($cm \cdot eV^{2.5}/s$)	4.87×10^{2}	4.87×10^{2}
χ $(V/cm \cdot eV)^{1.5}$	1.3×10^{8}	1.3×10^{8}
λ_{ins} (cm)	3.2×10^{-7}	3.2×10^{-7}
E_{B0} (eV)	3.1	4.7
α $(V \cdot cm)^{0.5}$	2.6×10^{-4}	2.6×10^{-4}
β $(V \cdot cm)^{1/3}$	1.5×10^{-5}	1.5×10^{-5}

Tab. 2.10 Coefficients for Fiegna model.