國立交通大學
電子工程學系電子研究所

碩士論文

閘極高介電值絕緣層穿隧電流的模擬

Modeling of Tunneling Currents in High-K Gate Stacks

研究生：張華罡　　**Hua-Gang Chang**

指導教授：陳明哲　　**Prof. Ming-Jer Chen**

中華民國 九十九 年 九 月

閘極高介電值絕緣層穿隧電流的模擬

Modeling of Tunneling Currents in High-K Gate Stacks

研究生：張華罡　　Hua-Gang Chang

指導教授：陳明哲　　Prof. Ming-Jer Chen

國立交通大學

電子工程學系電子研究所

碩士論文

A Thesis

Submitted to Department of Electronics Engineering &

Institute of Electronics

College of Electrical and Computer Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Electronics Engineering

September 2010

Hsinchu, Taiwan, Republic of China

中華民國 九十九 年 九 月

閘極高介電值絕緣層穿隧電流的模擬

研究生：張華罡　　指導教授：陳明哲博士

國立交通大學

電子工程學系電子研究所

摘要

因為高介電值絕緣層可以達到抑制閘極漏電流的目的，所以傳統的閘極氧化介電層正逐漸被高介電值絕緣層所取代。雖然高介電值絕緣層的電性特性已經被發表過了，但是高介電值絕緣層的物理基礎模型還沒有被完全的研究透徹。

在此論文中，建立一個針對 n 型金氧半電晶體的高介電值絕緣層電子穿隧模型。針對高介電值絕緣層穿隧電流的機制將會逐步地一一介紹。首先我們用複矽晶閘極單氧化層的簡易結構來解釋其操作原理。這個原理包含四個關鍵的參數：反轉層的電荷密度，電子對於氧化層/矽介面的撞擊頻率，WKB 的傳輸機率，以及反射波的修正因子。

接著，將只有一層的閘極絕緣層模型擴展到適用於高介電值絕緣層的二層結構。對於電子穿隧模型變數的影響將會詳細地分析。並且在對於高介電值絕緣層穿隧電流的模擬及量測上得到卓越的符合。重要的是，穿隧變數以及能帶圖能夠相應的得到，這可以更加地了解電子在高介電值絕緣層中的閘極穿隧機制。

Modeling of Tunneling Currents in High-K Gate Stacks

Student: Hua-Gang Chang Advisor: Prof. Ming-Jer Chen

Department of Electronics Engineering and Institute of Electronics

National Chiao Tung University

Abstract

Because the high-K stacks could achieve the target in suppressing the gate leakage current, conventional SiO_2 gate dielectrics is being gradually replaced by high-K materials. Although electrical characteristics have been demonstrated for high-K stacks, a physically based model of tunneling currents through high-K stacks has not been thoroughly investigated.

In this thesis, an electron tunneling model of high-K stacks will be constructed for nMOSFETs. The mechanisms responsible for the tunneling current will be introduced step by step. First, we explain the operational principle in a simplified framework of one oxide layer with poly gate. This principle comprises four key physical parameters: the inversion layer charge density, the electron impact frequency on SiO_2/Si interface, the WKB transmission probability, and the reflection correction factor.

Then, the structure to treat the case of only one gate dielectric layer model is augmented to that of a two-layer high-K stacks. The impact of the model parameters on the electron tunneling is analyzed in detail. Excellent agreements between simulated and measured tunneling currents are achieved. Importantly, the tunneling parameters and energy band diagrams are obtained accordingly, leading to a better understanding of the tunneling mechanism of electrons in high-K gate stacks.

誌　　謝

首先，要感謝指導教授陳明哲博士的細心教導。當初的我對於半導體物理毫無基礎，即便如此老師依然持續給予我鼓勵、支持。此外在研究方面老師的指導更是令我獲益良多。

接著要感謝的是實驗室的各位，在這二年的相處裡我過得相當有趣、充實。其中，特別是許智育學長，由於他的指導及幫助，我的研究才漸漸步入正軌。最終論文能夠完成，也是學長不厭其煩地指點之故。研究所的二年裡因為有大家的陪伴，讓我的回憶裡充滿著歡笑、溫暖。

最後感謝我的家人。我的任性給家人們帶來了許多的困擾、麻煩，但是他們以包容以及諒解來支持著我。因為有他們的鼓勵和陪伴，我才能一步步往前邁進，真心地感謝他們。

Contents

Abstract (Chinese) ……I

Abstract (English)……... II

Acknowledgement…… III

Contents…… IV

Figure Captions ……..V

Chapter 1 Introduction

1.1 Origin……...….1

1.2 Arrangement of this Thesis……...1

Chapter 2 Experimental……3

Chapter 3 Physical Model of Direct Tunneling across Oxide Dielectric

3.1 Critical Factors……4

3.2 Inversion Layer Charge……4

3.3 Impact Frequency……5

3.4 Transmission Probability……..6

Chapter 4 Simplified Method to Calculate Subband Energy…..10

Chapter 5 Tunneling Model for High-K Gate Stacks

5.1 MOS Electrostatics……13

5.2 WKB Probability for High-K Stacks……...…..14

5.3 Correction factor for High-K Stacks……..16

Chapter 6 Application and Discussion…….…...17

Chapter 7 Conclusion……..……20

References……..22

Figures…….....…24

Figure Captions

Fig. 1. The gate leakage current is composed of conductance electron tunneling current (I_{SD}) and valence hole current (I_B)…..…24

Fig. 2. Gate current separation shows the conductance electron tunneling current is dominant the gate leakage current in inversion region……25

Fig. 3. Measured gate current versus gate voltage for T = 373K and T=300K of (a) TaC/HfSiON/IL gate stacks and (b) TiN/HfO_2/IL gate stacks of nMOSFETs……..…26

Fig. 4. Comparison of simulated (line) gate current change of T = 373K with respect to T = 300K versus gate voltage with our measured current (scatters)……...…27

Fig. 5. The energy band diagram schematically shown for the n^+poly/ SiO_2/p-substrate system……...……..28

Fig. 6. The measured gate current versus gate voltage and its fitting for n^+poly gate MOSFETs…….....29

Fig. 7. The measured drain currents versus gate voltage for n^+poly gate

MOSFETs……..30

Fig. 8. The energy band for the metal gate/high-K/interface layer/Si substrate system diagram for two level high-K stacks……..31

Fig. 9. Three different cases for electron tunneling in varying gate voltage……32

Fig. 10. The demonstration of three different tunneling regions…….33

Fig. 11. The influence of (a) φ_{IL} and (b) φ_{high-K} on the calculated gate tunneling current……34

Fig. 12. The influence of (a) m_{IL} and (b) m_{high-K} on the calculated gate tunneling current……....35

Fig. 13. The influence of (a) t_{IL} and (b) t_{high-K} on the calculated gate tunneling current……36

Fig. 14. The influence of (a) ε_{IL} and (b) ε_{high-K} on the calculated gate tunneling current……37

Fig. 15. The influence of (a) N_{sub} and (b) φ_{metal} on the calculated gate tunneling current……38

Fig. 16. The fitting results by our model in comparison with the measured tunneling current in the literature [7]. The tunneling parameters are the same as those in [7]……39

Fig. 17. The measured capacitance versus gate voltage for two different gate stacks. And it is also shown the fitting lines……..40

Fig. 18. The measured drain currents versus gate voltage for (a) TaC/HfSiON/IL and (b) TiN/HfO_2/IL gate stacks……41

Fig. 19. Comparison of the measured and calculated gate currents versus gate voltage for (a) TaC/HfSiON/IL and (b) TiN/HfO_2/IL gate stacks……..42

Chapter 1 Introduction

1.1 Origin

Ultrascaled metal oxide semiconductor (MOS) technology requires the replacement of conventional SiO_2 gate dielectrics by high-K materials. The high-K gate stacks are able to reduce the gate leakage current. Besides, in order to shrink MOSFET devices into the deep-submicrometer regime, there are two factors taken into account: One is the level of increasing channel dopants and the other is gate dielectric thickness reduction. Both lead to a significant increase of the surface electric field, this imposing high demands on the advanced technology and on the understanding of the device physics involved. The combination of extremely thin gate dielectric material and high channel doping level produces a sufficiently large silicon electric field to confine the mobile carriers.

The MOSFET gate oxide thickness is rapidly approaching the direct tunneling limit that ultimately leads to intolerably large standby power and impractical applications. Thus, high-K stacks exhibit lower gate tunneling current. Although electrical characteristics have been demonstrated for high-K stacks, a physically based modeling of tunnel current through high-K stacks has not been thoroughly investigated. Accurate characterization and modeling of high-K stacks in the tunneling regime is essential and crucial. The latter can provide more transparent understandings since it is made up of four key physical parameters: the inversion layer charge density, the electron impact frequency on interface, the WKB transmission probability, and specially, the reflection correction factor.

1.2 Arrangement of this Thesis

First of all, this thesis is organized based on the following arrangement. Chapter 2 is the discussion of current separation and temperature effect. In Chapter 3, a

physical model for electron direct tunneling current in poly-gate nMOSFETs with ultrathin gate oxides is derived. This model comprises four key physical parameters: the inversion layer charge density, the electron impact frequency on SiO_2/Si interface, the WKB transmission probability, and the reflection correction factor.

In Chapter 4, in order to build a model of the high-K stacks electron tunneling current accounting for electron's subbands in the quantized inversion layer explicitly, a simplified method to calculate the subband energies will be introduced. In particular, the proposed model points out that the calculated secondary subbands and beyond, despite occupying few electrons, indeed contribute substantially to the direct tunneling conduction due to lower effective barrier heights.

In addition, high-K dielectrics with equivalent oxide thickness but larger physical thickness have been investigated as a replacement for SiO_2 one. An ultrathin interfacial oxide layer is required and can be formed during the fabrication below the high-K materials to improve interface quality with Si. Thus, accurate characterization and modeling of high-K gate stacks for the electron tunneling current is essential. The model built in Chapter 2 will be augmented to incorporate the high-K stacks. The process of parameter adjustment will be detailed in Chapter 5.

In Chapter 6, a physical model of tunneling currents through high-K gate stacks will be described. The effect of the model parameters on the electron tunneling will be analyzed in detail. Excellent agreements between simulated and measured tunneling currents will be demonstrated for MOSFETs with high-K gate stacks. We will also quote the literature value for comparison. The model is shown to enable in-depth understanding of the tunneling mechanisms.

Finally, Chapter 7 is the dedicated to conclusion of the proposed electron tunneling model as well as the potential application of the model.

Chapter 2 Experimental

Fig.1 shows the bias condition and gate leakage current separation, the gate leakage current is composed of conductance electron tunneling current (I_{SD}) and valence hole current (I_B). Fig.2 displays that the conductance electron tunneling current is dominant the gate leakage current in high gate voltage. Therefore the gate leakage current can be described by conductance electron direct tunneling model in inversion region.

To determine the mechanisms responsible for gate current, the gate current are also measured at 100^0C, as shown in Fig. 3.Gate current measurement in high temperature (100^oC) was also done for determing the dominant physical mechanisms of the gate current, as shown in Fig. 3. The measured gate current is insensitive with temperature in inversion region. This indicates that the direct and F-N tunneling mechanisms dominate the transmission process of gate current in inversion region. Moreover, the measured gate current change of T = 373K with respect to T = 300K versus gate voltage of Fig. 3(a) is shown in Fig. 4.

The simulated gate current change of T = 373K with respect to T = 300K versus gate voltage is also presented in Fig. 4 and then a good reproduction of experimental data is done. This specific characteristic of temperauture effects on gate current over the gate voltage can serve as the evidence that the physical mechanisms of the mesured gate current are dominated by direct and F-N tunneling. The trap-assist-tunneling is neglected in this study by the support of the temperature characteristic of our mesured gate current. Hence, in this work, considering the direct and F-N tunneling only in fitting gtae curent of the metal gate high-k deice is reasonable.

Chapter 3 Physical Model of Direct Tunneling across Oxide Dielectric

3.1 Critical Factors

First of all, we explain the fundamental principle of electron direct tunneling in a n^+poly-SiO_2-p-substrate structure, and expand it to include the case of high-K stacks. We employ quantum mechanical calculations for the inversion layer in p-type substrate along with a modified WKB approximation for the transmission probability across the gate oxide [1], [2].

The direct tunneling electron current from each subbands can be formulated in a similar way. Electron current density contributed by j-th subband for two or four fold valleys with energy E to $E+dE$ can be written as

$$dJ_{i,j}(E) = qf_j g(E) P_t(E) dE \tag{3-1}$$

where

q is the elemental charge;

$g(E)$ is the inversion layer charge density per unit area magnitude of energy E associated with j-th subband;

f_j is the impact frequency of electron's wave packet on SiO_2/Si interface; and

$P_t(E)$ is the transmission probability through SiO_2 layer.

Then we will make a physical description of the key parameters above.

3.2 Inversion Layer Charge

In order to calculate inversion-layer charge density $g(E)$, we must turn to the

inversion charge increment,

$$dN_{inv} = \int_{E}^{E+dE} g_{2D} F_e(E)\, dE \simeq g_{2D} F_e(E) dE \tag{3-2}$$

where $F_e(E)$ indicates Fermi-Dirac distribution function of electrons and $F_e(E)$ associated with conduction electrons is $1/\left(1+\exp\left(\left(E_f - E\right)/k_B T\right)\right)$. In addition, g_{2D} represents density of states per unit area over electron energy for 2DEG (2-dimensional electron gas). The m_d is the density of states effective mass

$$g_{2D} = \frac{m_d}{\pi \hbar^2} \tag{3-3}$$

From the variation of inversion-layer charge, $g(E)$ could be obtained accordingly,

$$g(E) = \frac{dN_{inv}}{dE} = g_{2D} F_e(E) \tag{3-4}$$

3.3 Impact Frequency

The impact frequency f_j can be expressed as,

$$f_j = \left[2\int_0^{z_j} \frac{dz}{v_{si\perp}(z)} \right]^{-1} = \frac{q\varepsilon |F_{ox}|}{2\varepsilon_{Si}} \left(2 m_{z,i} E_{i,j}\right)^{-\frac{1}{2}} \tag{3-5}$$

where z_j is the classical turning point at the j-th subband edge and $m_{z,i}$ means the effective mass along $\langle 100 \rangle$ direction for two or four fold valleys. The field strength of oxide is designated as F_{ox}. Furthermore, $v_{si\perp}(z)$ is the interface-normal velocity of electron wave packet, which can be written as

$$v_{si\perp}(z) = \sqrt{\frac{2(E_{i,j} - qV(z))}{m_z}} \tag{3-6}$$

Here $E_{i,j}$ is the quantized energy level of the j-th subband for two or four fold valleys. A triangle-like electrostatic potential is a good approximation for $V(z)$ [3].

Thus, we are able to get quantized energy along z-direction or $\langle 100 \rangle$ direction in reciprocal space by using Sommerfeld-Wilson's quantization rule :

$$E_{i,j} = \left(\frac{\hbar^2}{2m_z} \right)^{\frac{1}{3}} \left[\frac{3\pi q \varepsilon_{ox} |F_{ox}| \left(j - \frac{1}{4} \right)}{2\varepsilon_{Si}} \right]^{\frac{2}{3}} \tag{3-7}$$

3.4 Transmission Probability

Before we discuss the transmission probability through SiO_2 layer, the underlying electrostatics must be made clear. The energy band diagram as shown in Fig. 5 is constructed for the n^+poly/SiO_2/p-substrate system. Under the inversion condition, it is easy to write

$$|V_G - V_{FB}| = t_{ox} |F_{ox}| + \phi_s + \phi_{poly} \tag{3-8}$$

where V_G means gate voltage, V_{FB} represents flat-band voltage, and ϕ_s and ϕ_{poly} are the surface potential near the SiO_2/p-substrate and SiO_2/poly gate, respectively. In addition, once the surface potential are quantified, then we can get the quasi Fermi level E_f from the following formula

$$E_f = \phi_s - \phi_{Bp} \tag{3-9}$$

where

$$\phi_{Bp} = \frac{\left[E_g - k_B T \ln \left(\frac{N_V}{N_A} \right) \right]}{q} \tag{3-10}$$

where ϕ_{Bp} is the potential difference between the quasi Fermi level $\left(E_f / q \right)$ and conduction band $\left(E_C / q \right)$ in the charge neutrality range of p-substrate, and, k_B is

the Boltzmann's constant.

Finally, based on the citation [3], $P_t(E)$ can be modeled by

$$P_t(E) = T_{WKB}(E)T_R(E) \tag{3-11}$$

where T_{WKB} is the usual WKB tunneling probability valid for smooth varying potential, and T_R is a factor corrected for the reflections due to the potential discontinuities:

$$\begin{aligned} T_{WKB}(E) &= \exp\left(-2\int_0^{t_{ox}} \kappa(z)\,dz\right) \\ &= \exp\left[-\frac{2}{\hbar}\left|\int_0^{t_{ox}} \sqrt{2m_{ox}\left(E - qV(z)\right)}dz\right|\right] \\ &= \exp\left(\frac{4\sqrt{2m_{ox}}\left(\phi_{an}^{\frac{3}{2}} - \phi_{cath}^{\frac{3}{2}}\right)}{3q\hbar\left|F_{ox}\right|}\right) \end{aligned} \tag{3-12}$$

where ϕ_{an} is the barrier height for tunneling electrons with total energy E at anode side (poly gate/SiO_2 interface), and ϕ_{cath} is that at cathode side (SiO_2/p-substrate interface). In this model, φ_{ox} represents the SiO_2/Si barrier height. The barrier height for tunneling electrons can be described by

$$\phi_{an} = \varphi_{ox} - q\left|F_{ox}\right|t_{ox} - E \tag{3-13}$$

$$\phi_{cath} = \varphi_{ox} - E \tag{3-14}$$

The total energy E consists of the transversal and longitudinal energies:

$$E = \frac{\hbar^2\left(k_x^2 + k_y^2\right)}{2m_t} + E_{i,j} \tag{3-15}$$

where m_t is the in-plane effective mass, $E_{i,j}$ is quantized energy in the out-of-plane direction .

On the other hand, T_R is a justified factor concerning wavefunction's reflection phenomenon occurring at SiO_2/Si interfaces [4]. This correction factor has the band-structure independent form. This relation is important in our model. T_R is obtained by considering reflections between the interfaces,

$$T_R(\mathrm{E})=\frac{4v_{si\perp}(\mathrm{E})\cdot v_{ox}(\phi_{cath})}{v_{si\perp}^2(E)+v_{ox}^2(\phi_{cath})}\times\frac{4v_{si\perp}(\mathrm{E}+\mathrm{q}|\mathrm{F_{ox}}|t_{\mathrm{ox}})\cdot v_{ox}(\phi_{an})}{v_{si\perp}^2(E+\mathrm{q}|\mathrm{F_{ox}}|t_{\mathrm{ox}})+v_{ox}^2(\phi_{an})} \tag{3-16}$$

where $v_{si\perp}(\mathrm{E})$ and $v_{si\perp}(\mathrm{E}+\mathrm{q}|\mathrm{F_{ox}}|t_{\mathrm{ox}})$ are the group velocities of the electron incident and leaving the oxide. The group velocity of electrons with energy E along $\langle 100\rangle$ direction within the j-th subband at SiO_2/Si interface is independent of E, as shown below

$$v_{si\perp}(\mathrm{E})=v_{si\perp}(z=0)=\sqrt{\frac{2E_{i,j}}{m_z}} \tag{3-17}$$

The aforementioned $v_{ox}(\phi_{cath})$ and $v_{ox}(\phi_{an})$ are the magnitudes of the purely imaginary group velocities of electrons at the cathode and anode side within the oxide. Consequently, the imaginary group velocity which is dependent on E within oxide can be described by

$$v_{ox}=\frac{1}{\hbar}\frac{d\phi_{an}}{dk_{ox}}=\sqrt{\frac{2\phi_{an}}{m_{ox}}} \tag{3-18}$$

This factor is significant, because if the electrostatic potential $V(x)$ changes acutely with respect to position or the intensity of incident wave, it cannot be treated to be equal to that of reflection wave. In our model, we adopt a parabolic dispersion relation,

$$\phi_{an}=\varphi_{ox}-E-qV_{ox}(x)=\frac{\hbar^2k_{ox}^2}{2m_{ox}} \tag{3-19}$$

Recalling (3-1), (3-4), (3-5), and (3-11), the tunneling current density contributed by the j-th subband with energy E ranging from $E_{i,j}$ to infinity can be expressed

$$\begin{aligned} J_{i,j} &= \int_{E_j}^{\infty} dJ_{i,j}(E) \\ &= \int_{E_j}^{\infty} qf_j g(E) P_t(E) dE \\ &= qf_j g_{2D} \int_{E_j}^{\infty} F_e(E) P_t(E) dE \end{aligned} \tag{3-20}$$

The above electron direct tunneling model is verified by comparing experimental data from n^+poly gate MOSFETs. The results are given in Fig. 6. It can be seen that excellent agreements between the experimental and simulated results are achieved. The drain currents versus gate voltage for n^+poly gate/SiO_2/p-substrate is shown in Fig. 7, the threshold voltage V_{th} is obtained by the maximum trans-conductance method. In the calculation, the doping concentrations of poly gate and p-type substrate used (N_{poly} and N_{sub}) are 1×10^{20} and $6\times10^{17}\ cm^{-3}$, respectively. The thickness of oxide t_{ox} is extracted by C-V fitting. These parameters are all known when apply the model, except the effective mass m_{ox} that is the only fitting parameter.

Chapter 4 Simplified Method to Calculate Subband Energy

In this chapter, we will introduce a simplified method to calculate the subband energies. The simplified method is to estimate the inversion and depletion charge densities. The band bending in depletion and inversion condition is also treated. In this work, we use the assumption to facilitate the model: The quantum confinement phenomenon of the MOS system can be treated in a triangular well approximation. The characteristic parameters of P-substrate under the inversion situation comprise of the followings:

Interfacial inversion charge N_{inv};

Depleted space charge N_{depl};

Semiconductor surface potential or surface band bending ϕ_s;

Band bending due to the depletion charge ϕ_{depl};

These parameters are interrelated by the following equation due to the law of electrostatics:

$$N_{inv} + N_{depl} = Q(\phi_s) \tag{4-1}$$

Equation (4-1) states that the total charge per unit area Q below the MOS gate is the sum of the inversion electron charge N_{inv} and the depletion charge N_{depl}:

$$\phi_{depl} = \phi_s - \frac{qN_s\bar{z}_{QM}}{\varepsilon_{Si}\varepsilon_0} - \frac{kT}{q} \tag{4-2}$$

From (4-2), the surface potential ϕ_s is made up of the following components: the second term of the right side of (4-2) stems from the influence of inversion charge N_{inv}, the terminal term kT/q from the gradual transition of the space charge region

into the substrate, and the term ϕ_{depl} due to the space charge N_{depl}.

$$N_{depl} = \sqrt{\frac{2\varepsilon_{Si}\varepsilon_0\phi_{depl}\left(N_D\right)}{q}} \tag{4-3}$$

The equivalent thickness of the quantum confined electron gas can be expressed as

$$\overline{z}_{QM} = \sum_{i,j} Z_{i,j} N_{i,j}\left(E_{i,j}, E_F\right) \Big/ N_{inv} \tag{4-4}$$

(4-4) reveals that the average of the subband widths $z_{i,j}$ is weighted with the corresponding subband occupation factors $N_{i,j}$, then constituting the mean quantum mechanical channel width $\overline{z}_{QM}$. In this case, the wave functions are given by Airy functions, and therefore the mean subband width $z_{i,j}$ is

$$z_{i,j} = \int \left|\Psi i, j(z)\right|^2 z dz = \frac{2E_{i,j}\varepsilon_{Si}}{3q\varepsilon_{ox}E_{ox}} \tag{4-5}$$

The energy eigenvalues $E_{i,j}$ can be written by

$$E_{i,j} = \left(\frac{\hbar^2}{2m_i}\right)^{\frac{1}{3}} \left[\frac{3\pi q\varepsilon_{ox}F_{ox}\left(j-0.25\right)}{2\varepsilon_{Si}}\right]^{2/3} \tag{4-6}$$

where F_{ox} is the oxide electric field and the index i determines the value of the normal mass m_i for two or four fold valleys. In addition, $N_{i,j}$ can be derived from Fermi-Dirac statistics

$$N_{i,j} = \int_{E_{i,j}}^{\infty} D_i(E) f(E) dE = \frac{m_{di}}{\pi\hbar^2} \ln\left(1+\exp\left(\frac{E_F - E_{i,j}}{kT}\right)\right) \tag{4-7}$$

where $D_i(E)$ is the density of states of the subband for two-dimensional gas and $f(E)$ is the Fermi-Dirac occupation factor.

It can be confirmed that the lowest subband occupies most of all inversion electrons.

Although higher energy states share far less carriers, the other factor such as the transmission probability can be much larger than the ground state due to effective lower barrier heights for tunneling. Thus, the resulting tunneling current contains a substantial component from the secondary subbands and beyond. Besides the electron direct tunneling from the secondary subbands and beyond can be comparable to that from first subbands. Consequently, we can not ignore the contribution of the secondary and beyond to the electron tunneling current.

Chapter 5 Tunneling Model for High-K Gate Stacks

5.1 MOS Electrostatics

MOS electrostatics law is utilized to deal with the potential drops across the high-K gate stacks. The flat-band voltage for the poly gate and metal gate case can be written as

$$V_{FB} = \frac{\varphi_{metal}}{q} - \left[\frac{\varphi_{sub}}{q} + \frac{E_G}{q} + \frac{KT}{q} \mathbf{log}\left(\frac{N_{sub}}{N_v} \right) \right] \quad ... \quad for\ metal\ gate \tag{5-1}$$

$$V_{FB} = \frac{E_G}{q} + \frac{KT}{q} \mathbf{log}\left(\frac{N_{sub}}{N_v} \right) \quad ... \quad for\ \text{poly}\ gate \tag{5-2}$$

where φ_{metal} is the workfunction of metal, and φ_{sub} means electron affinity of p-substrate. Doping concentration of p-substrate is named N_{sub}. E_G (=1.12 eV) is band gap energy of silicon. Because there is high doping concentration in poly-silicon, the quasi-Fermi level approaches the conduction band edge for poly-silicon in energy band diagram.

By Gauss' law we can get V_{high-K} (potential drop across high-K dielectric) easily,

$$F_{IL} = \frac{V_{IL}}{t_{IL}} \quad , \quad F_{high-K} = F_{IL} \frac{\varepsilon_{IL}}{\varepsilon_{high-K}}$$
$$V_{high-K} = F_{high-K} \times t_{high-K} \tag{5-3}$$

The thickness of interfacial layer (IL) or high-K layer can be estimated by C-V method. Moreover, F_{IL} and F_{high-K} mean the electric field of IL and high-K dielectric, respectively. The potential drop on IL (V_{IL}) is obtained by a self-consistent calculation.

The band bending of substrate surface can be computed:

$$\phi_s = |V_G - V_{FB}| - t_{IL}|F_{IL}| - t_{high-K}|F_{high-K}| - \phi_{poly} \tag{5-4}$$

where $t_{high-K}|F_{high-K}|$ and $t_{IL}|F_{IL}|$ are the potential drop across the high-K dielectric and interfacial layer, respectively. The n^+-poly gate potential drop (ϕ_{poly}) can be expressed as

$$\phi_{poly} = \frac{\varepsilon_{IL}^2 F_{IL}^2}{2q\varepsilon_{Si} N_{poly}}$$

In the metal gate case, it is also simple for modulating ϕ_s. The formula is also suitable for metal gate case when deleting ϕ_{poly}.

5.2 WKB Probability for High-K Stacks

The energy band diagram as shown in Fig. 8 is constructed for a two-layer high-K stack system. To calculate the tunneling probability, we have selected a modified WKB method with $P_t = T_{WKB} \cdot T_R$, where T_{WKB} is the usual WKB tunneling probability obtained for smooth varying potentials and T_R is the correction factor considering reflections from potential discontinuities. They can be treated as the most important part of electron tunneling model. The transmission probability $P_t(E)$ not only goes across the interfacial layer (IL), but also the high-K dielectric [5], [6]. Therefore, T_{WKB} has two portions needed to be treated, one is with regard to interfacial layer, and the other is on high-K dielectric:

$$\begin{aligned} T_{WKB}(E) &= \exp\left(-2\int_{x_R}^{x_L} \kappa(x)\,dx\right) \\ &= \exp\left(-2\int_{x_R}^{x'} \kappa_{IL}(x)\,dx + \left(-2\int_{x'}^{x_L} \kappa(x)_{high-K}\,dx\right)\right) \\ &= \exp\left(\frac{4\sqrt{2m_{IL}}\left(\phi_{IL,an}^{\frac{3}{2}} - \phi_{IL,cath}^{\frac{3}{2}}\right)}{3q\hbar|F_{IL}|}\right) \exp\left(\frac{4\sqrt{2m_{high-K}}\left(\phi_{high-K,an}^{\frac{3}{2}} - \phi_{high-K,cath,}^{\frac{3}{2}}\right)}{3q\hbar|F_{high-K}|}\right) \end{aligned} \tag{5-6}$$

The index IL corresponds to the interfacial layer. The index high-K applies to the high-K dielectric. Now, this new equation can apply to high-K stacks, but it is still problematic in some conditions. The band diagram essentially varies with the change of the gate bias voltage. Consequently, Fig. 9 shows three different cases. Case 1 is normal situation that have been discussed above. Comparing Fig. 9, if a subband energy is lying between the two edges of high-K layer, namely case 2. When gate voltage gets larger, the subband energy of electrons is even higher than high-K edge which is located at interface of high-K dielectric and IL. In other words, the classification can be represented in the form of the following equations:

$$\begin{aligned} &E < \phi_{high-K,an} \;\; \& \;\; E < \phi_{high-K,cath} \quad ... \quad case\ 1 \\ &E \geq \phi_{high-K,an} \;\; \& \;\; E < \phi_{high-K,cath} \quad ... \quad case\ 2 \\ &E \geq \phi_{high-K,an} \;\; \& \;\; E \geq \phi_{high-K,cath} \quad ... \quad case\ 3 \end{aligned} \tag{5-7}$$

where, E means the total energy including the transversal and longitudinal energies. In case 2 and 3, they correspond to the following WKB approximation methods:

$$T_{WKB}(E) = \exp\left(\frac{4\sqrt{2m_{IL}}\left(\phi_{IL,an}^{\frac{3}{2}} - \phi_{IL,cath}^{\frac{3}{2}}\right)}{3q\hbar\left|F_{IL}\right|}\right)\exp\left(\frac{4\sqrt{2m_{high-K}}\left(0 - \phi_{high-K,cath}^{\frac{3}{2}}\right)}{3q\hbar\left|F_{high-K}\right|}\right) \quad ... \quad case\ 2$$

$$T_{WKB}(E) = \exp\left(\frac{4\sqrt{2m_{IL}}\left(\phi_{IL,an}^{\frac{3}{2}} - \phi_{IL,cath}^{\frac{3}{2}}\right)}{3q\hbar\left|F_{IL}\right|}\right) \quad ... \quad case\ 3 \tag{5-8}$$

Further, the potential energy for each edge of two layers could be calculated with the following equation,

$$\begin{aligned} &\phi_{IL,cath} = \varphi_{IL} - E \\ &\phi_{IL,an} = \varphi_{IL} - E - V_{IL} \\ &\phi_{high-K,cath} = \phi_{IL,an} - \left(\varphi_{IL} - \varphi_{high-K}\right) \\ &\phi_{high-K,an} = \phi_{IL,an} - \left(\varphi_{IL} - \varphi_{high-K}\right) - V_{high-K} \end{aligned} \tag{5-9}$$

In case 2, electrons pass through IL by direct tunneling, but with F-N tunneling for high-K part. The behavior of electron F-N tunneling in high-K can appear in I-V characteristic. For example, Fig. 10 shows such a turning point when gate voltage increases to 1.6 V. As the gate current increases and exceeds the turning point different situation appears. Similarly, case 3 has only one kind of tunneling mechanism. Because the electrons have higher energy in case 3, they can go through IL by direct tunneling and surpass high-K layer immediately. The expression used for case 3 is like the only dielectric layer as discussed in Chapter 3. In case 3, the gate current increases with gate voltage.

5.3 Correction factor for High-K Stacks

In principle, there is a special T_R which is in connection with high-K dielectric. But here, we only consider the reflection at the Si/IL interface in our model. The correction factor of interface reflection typically approaches unity for metal/high-K dielectric interface and for the interface of high-K gate stacks. The correction factors at the interface mentioned above can be neglected in our calculations. Therefore, T_R still maintains the original shape. The correcting T_R for high-K stacks is as follows:

$$T_R(\mathrm{E})=\frac{4v_{si\perp}(\mathrm{E})\cdot v_{\mathrm{IL}}(\phi_{IL,cath})}{v_{si\perp}^2(E)+v_{\mathrm{IL}}^2(\phi_{IL,cath})} \qquad (5\text{-}10)$$

Chapter 6 Application and Discussion

We calculate the first and second subbands. The electron tunneling current from such two lowest subbands is almost the same as the actual tunneling current. Thus, for the computation time saving, the first and second subbands are considered only in this work. Although we have done many simplifications in the tunneling situation, the results are reasonable relation to the experimental measurement.

The tunneling model is constructed using an accurate description of electron quantization and a modified WKB method for tunneling through the barriers of high-K stacks. But the high-K stacks electron tunneling current simulator has a drawback: it could not simulate gate tunneling current correctly for the gate bias below threshold voltage. Our simulated model is based on direct tunneling mechanism; unfortunately, the direct tunneling is not dominant in subthreshold region. Although simulator can not work in subthreshold region, WKB approximation would show good agreements with the experiments once the gate bias is large enough. For this reason, the following discussion will be focused on the range of gate bias larger than the threshold voltage.

Finally, the new model for high-K stacks can be used to fit the results in literature about high-K tunneling. In addition, we also measure the I-V and C-V characteristics of devices with HfSiON and HfO_2, and experimental reproduction has been consistently achieved.

The influence of each model parameters on the gate tunneling current has been calculated as shown in Fig. 11-15. Here m_{IL} and m_{high-K} are the effective mass for interfacial layer and high-K, respectively. The band offsets of IL and high-K relative to the conduction band of p-substrate are expressed as φ_{IL} and φ_{high-K}, respectively.

Similarly, ε_{IL} and ε_{high-K} mean the permittivity of IL and high-K, respectively.

From Fig. 11-15, φ_{high-K} determines the onset of case 2. The range that m_{high-K} can have a significant effect concentrates on case 1 and 2, where the tunneling current is almost the same at high gate voltage. The influence of t_{high-K} can also be seen, especially its ability to determine the rage of case 2. As for ε_{high-K}, its influence is not only on case 1 and 2 but also case 3, because ε_{high-K} will influence the electric field of IL indirectly. The parameters about IL impact the tunneling significantly in all cases. Finally, φ_{metal} and N_{sub} can control the threshold voltage, which can help the calibration step in using our model. Briefly, In Fig. 11-15, different tunneling parameters have been varied with our simulator. The change of the tunneling current is seen clearly. Each parameter has different influences, and the results can help us fit data efficiently.

Moreover, we test the model by fitting experimental results in some open literature articles [4], [7]-[12]. This can verify the validity of the model for tunneling current simulation; for example, an excellent agreement between our simulated and the measured tunneling currents in the literature [7] has been obtained as shown in Fig. 16. The tunneling parameters used are those published [7], except several parameters not available such as doping concentrations.

We have also measured the gate tunneling current for two kinds of high-K stacks: TaC/HfSiON/IL and TiN/HfO_2/IL. There is one thing needed to explain specially. IL means interfacial layer, but it is not always formed by SiO_2. In fact, the IL of our device under test is SiON and SiO_2 for HfSiON and HfO_2, respectively. The measured capacitance versus gate voltage for two different gate stacks is exhibited in Fig. 17. From C-V fitting method, the metal work function φ_{metal} and p type substrate doping

concentration N_{sub} can both be obtained. Meanwhile, the effective oxide thickness (EOT) can also be determined. These parameters which are obtained from C-V fitting are the inputs to the gate tunneling current model. Moreover, from the drain currents versus gate voltage shown in Fig. 18, we can obtain threshold voltage V_{th} by the maximum trans-conductance method. The gate current fitting results for TaC/HfSiON/IL and TiN/HfO_2/IL are shown in Fig. 19. It can be seen that for two level high-K stacks, the experimental gate current can be fitted well. The high-K stacks tunneling simulator can apply to either metal gate or poly gate. The tunneling parameters and energy band diagram can be established via the data fitting. Furthermore, a better understanding of the gate tunneling mechanisms of electrons in high-K stacks can therefore be captured clearly.

Chapter 7 Conclusion

A physically based tunneling model is constructed using an accurate description of the electron quantization in the confined inversion-layer. This model contains a modified Wentzel–Kramers–Brillouin (WKB) method for tunneling through the barrier, including the effects of an ultrathin interfacial layer. The physical model of the electron tunneling through high-K stacks has been built up and experimental reproduction has been consistently achieved.

Our electron tunneling model is applied to high-K stacks. This model is easy to realize and computationally efficient. The data from the other tunneling models published in the open literature is also presented to confirm the accuracy of our model. From the model, the effect of high-K film will appear when the F-N tunneling begins to dominate in high gate voltage. As the gate voltage goes higher, because of the reduced barrier height of the high-K film, electrons will tunnel within the Fowler Nordheim regime instead of the direct tunneling regime. So it is important to find a high-K dielectric with a sufficient barrier height to reduce tunneling current in a wide range of gate voltages.

This model has evidenced its potential applications in enabling in-depth understanding of the different subbands in the confined inversion layer in affecting electron tunneling conduction. The high-K stacks tunneling simulator can apply to either metal gate or poly gate. The tunneling parameters and energy band diagram can be established by the data fitting. We can therefore capture the underlying gate tunneling mechanisms of electrons in high-K gate stacks.

But there are efforts needed to investigate for further, such the hole tunneling across the high-K gate stacks. Even the temperature effect may be significant and must be

taken into account. In addition, the electron tunneling model which was built in this thesis is focused on the inversion condition. Its extension to the accumulation may be needed.

References

1. L. F. Register, E. Rosenbaum, and K. Yang, "Analytic model for direct tunneling current in polycrystalline silicon-gate metal-oxide-semiconductor devices," *Appl. Phys. Lett.*, vol. 74, no. 3, pp. 457-459, Jan. 1999.
2. K. N. Yang, H. T. Huang, M. C. Chang, C. M. Chu, Y. S. Chen, M. J. Chen, Y. M. Lin, M. C. Yu, S. M. Jang, D. C. H. Yu, and M. S. Liang, "A physical model for hole direct tunneling current in p^+ poly-gate PMOSFETs with ultrathin gate oxides," *IEEE Trans. Electron Devices*, vol. 47, no. 11, pp. 2161-2166, Nov. 2000.
3. N. Yang, W. K. Henson, J. R. Hauser, and J. Wortman, "Estimation of the effects of remote charge scattering on electron mobility of n-MOSFET's with ultrathin gate oxides," *IEEE Trans. Electron Devices*, vol. 47, no. 2, pp. 440-447, Feb. 2000.
4. F. Li, S. P. Mudanai, Y. Y. Fan, L. F. Register, and S. K. Banerjee, "Physically based quantum-mechanical compact model of MOS devices substrate-injected tunneling current through ultrathin (EOT ~ 1 nm) SiO_2 and high-K gate stacks," *IEEE Trans. Electron Devices*, vol. 53, no. 5, pp. 1096-1106, May. 2006.
5. F. Li, Y. Y. Fan, L. F. Register, and S. K. Banerjee, "Compact model of MOSFET electron tunneling current through ultra-thin SiO_2 and high-k gate stacks," in *Device Res. Conf.*, Jun. 2003, pp. 47-48.
6. Y. Zhao and M. H. White, "Modeling of direct tunneling current through interfacial SiO_2 and high-k gate stacks," in *Semicond. Device Res. Symp.*, Dec. 2003, pp. 468-469.
7. Y. T. Hou, M. F. Li, H. Y. Yu, and D. L. Kwong, "Modeling of tunneling current through HfO_2 and $(HfO_2)_x(Al_2O_3)_{1-x}$ gate stacks," *IEEE Electron Device Lett.*, vol.

24, no. 2, pp. 96-98, Feb. 2003.

8. M. A. Quevedo-Lopez, S. A. Krishnan, P. D. Kirsch, H. J. Li, J. H. Sim, C. Huffman, J. J. Peterson, B. H. Lee, G. Pant, B. E. Gnade, M. J. Kim, R. M. Wallace, D. Guo, H. Bu, and T. P. Ma, "High performance gate first HfSiON dielectric satisfying 45nm node requirements," in *IEDM Tech. Dig.*, Dec. 2005, pp. 425-428.
9. R. G. Southwick, J. Reed, C. Buu, H. Bui, R. Butler, G. Bersuker, and W. B. Knowlton, "Temperature (5.6-300K) dependence comparison of carrier transport mechanisms in HfO_2/SiO_2 and SiO_2 MOS gate stacks," in *IRW*, Oct. 2008, pp. 48-54.
10. J. Coignus, C. Leroux, R. Clerc, G. Ghibaudo, G. Reimbold, and F. Boulanger, "Experimental investigation of transport mechanisms through HfO_2 gate stacks in nMOS transistors," in *Solid State Device Res. Conf. ESSDERC*, Sept. 2009, pp. 169-172.
11. J. Coignus, R. Clerc, C. Leroux, G. Reimbold, G. Ghibaudo, and F. Boulanger, "Analytical modeling of tunneling current through SiO_2-HfO_2 stacks in metal oxide semiconductor structures," *J. Vac. Sci. Technol. B*, vol. 27, no. 1, pp. 338-345, Feb. 2009.
12. A. Campera, G. Iannaccone, and F. Crupi, "Modeling of tunneling currents in Hf-based gate stacks as a function of temperature and extraction of material parameters," *IEEE Trans. Electron Devices*, vol. 54, no. 1, pp. 83-89, Jan. 2007.

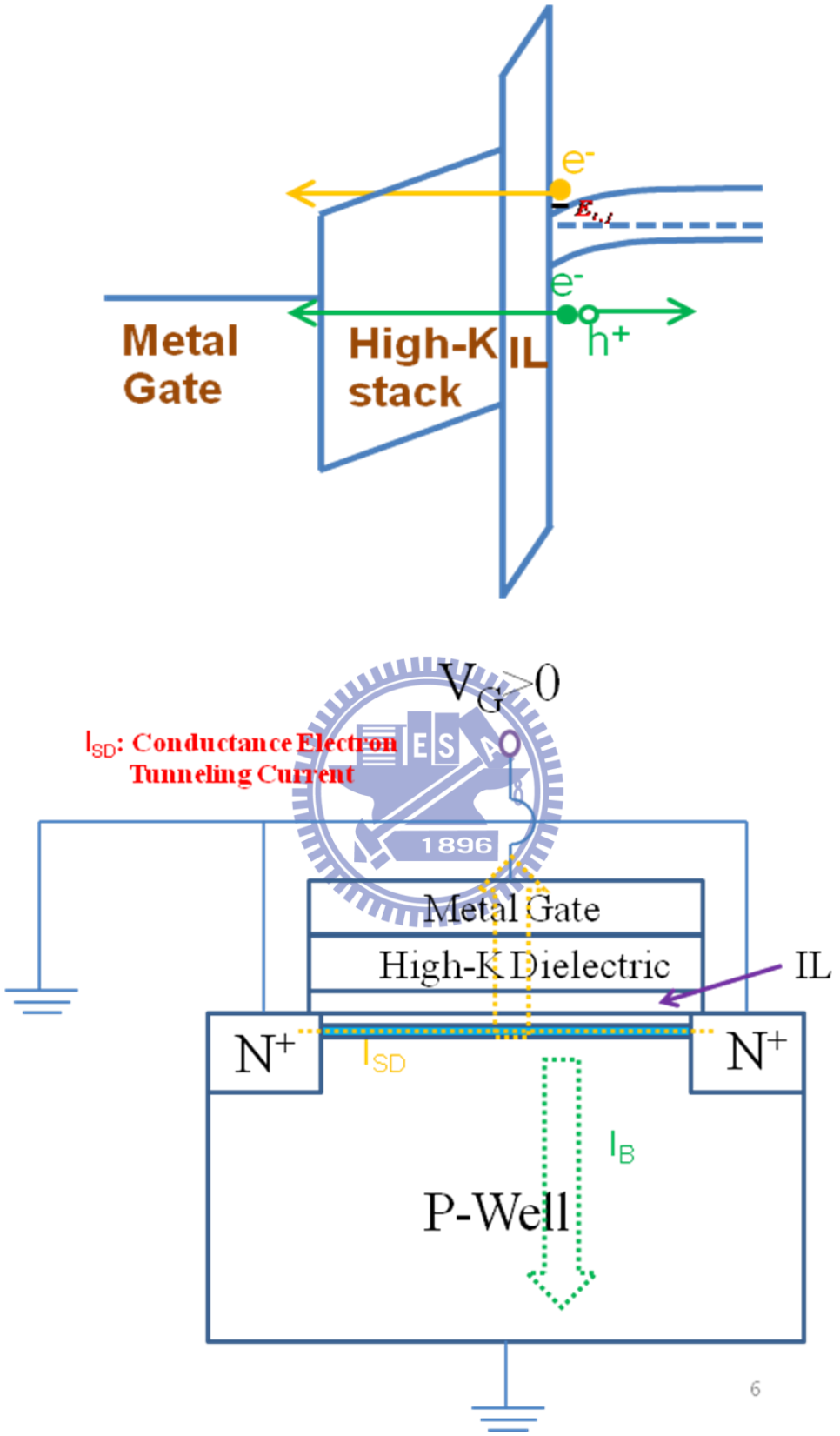


Fig. 1. The gate leakage current is composed of conductance electron tunneling current (I_{SD}) and valence hole current (I_B).

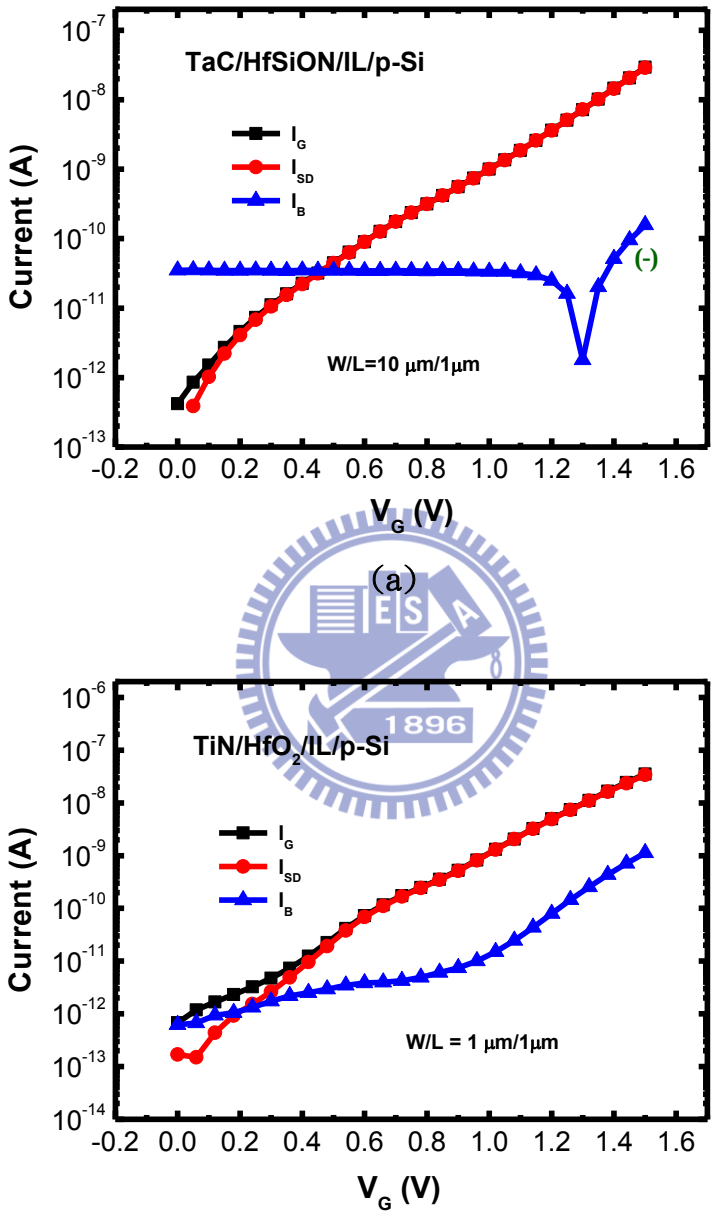


Fig. 2. Gate current separation shows the conductance electron tunneling current is dominant the gate leakage current in inversion region.

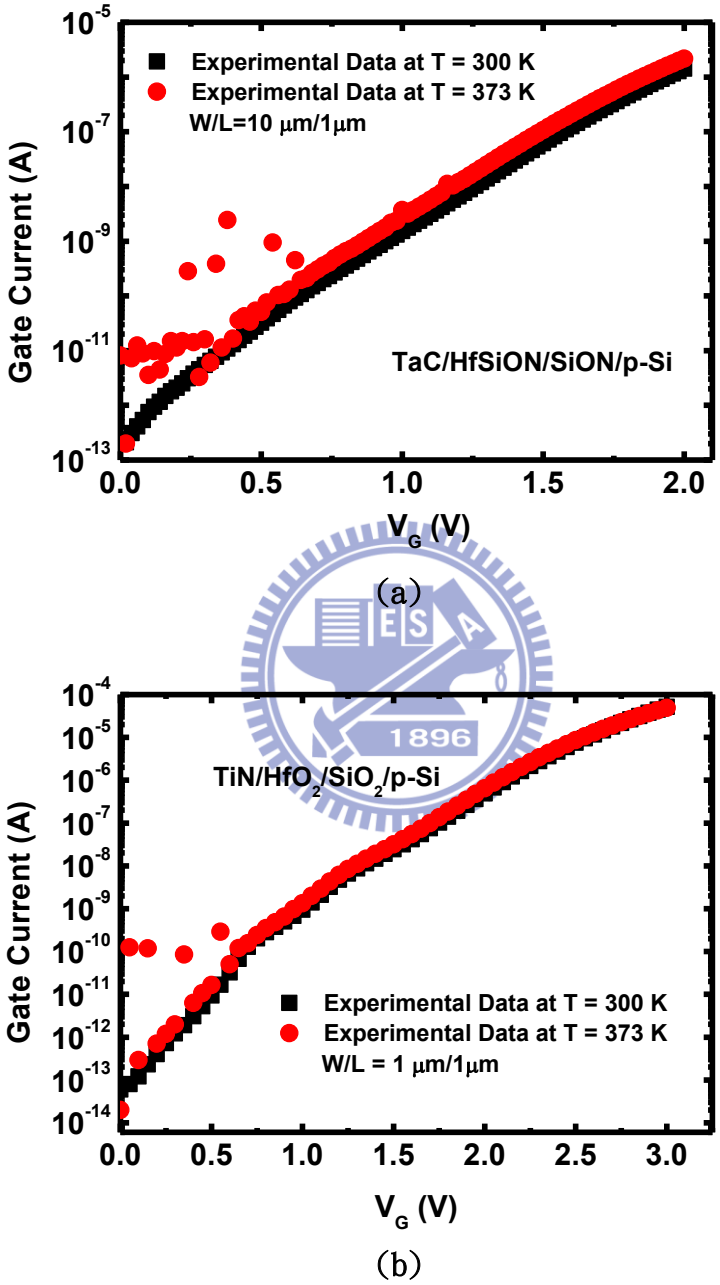


Fig. 3. Measured gate current versus gate voltage for T = 373K and T=300K of (a) TaC/HfSiON/IL gate stacks and (b) TiN/HfO_2/IL gate stacks of nMOSFETs.

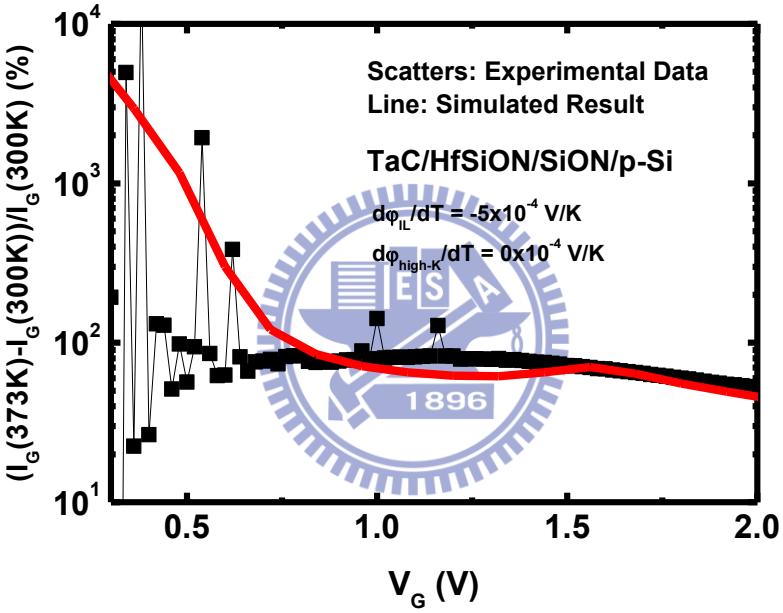


Fig. 4. Comparison of simulated (line) gate current change of T = 373K with respect to T = 300K versus gate voltage with our measured current (scatters).

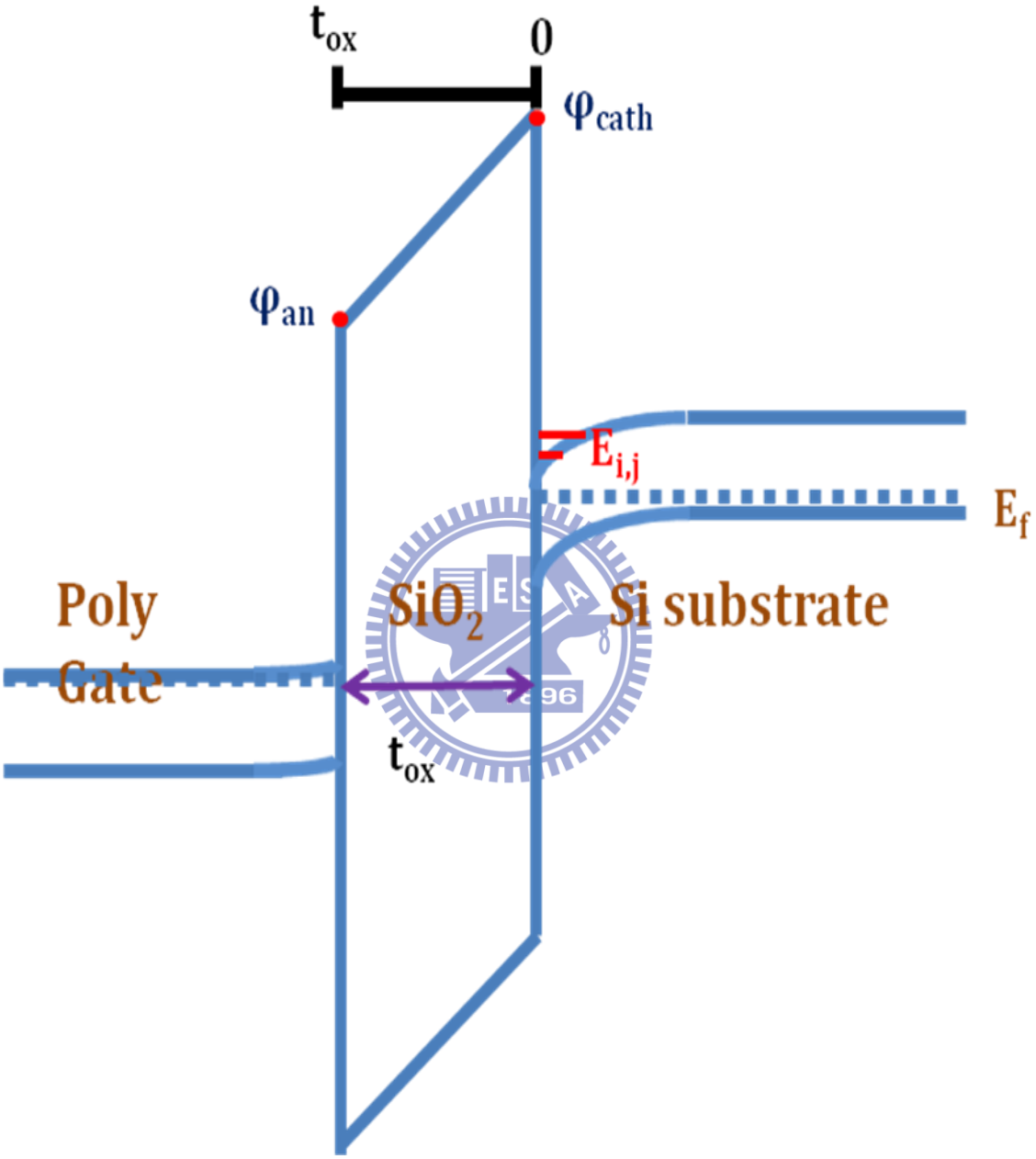


Fig. 5. The energy band diagram schematically shown for the n^+poly/SiO_2/p-substrate system.

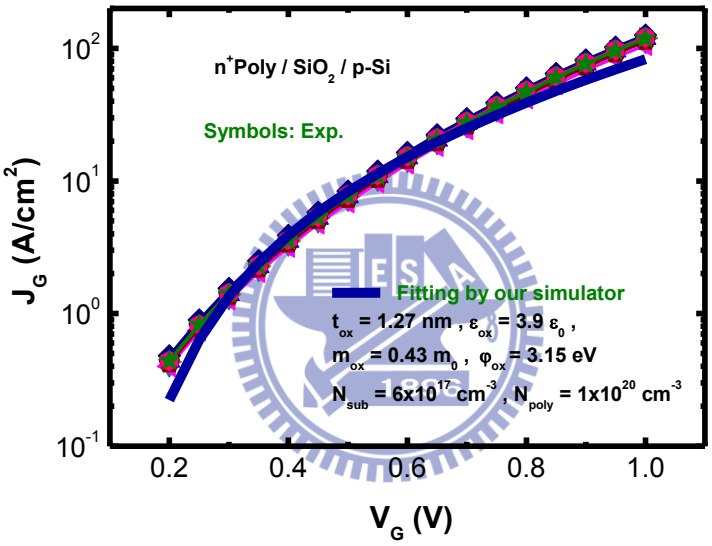


Fig. 6. The measured gate current versus gate voltage and its fitting for n^+poly gate MOSFETs.

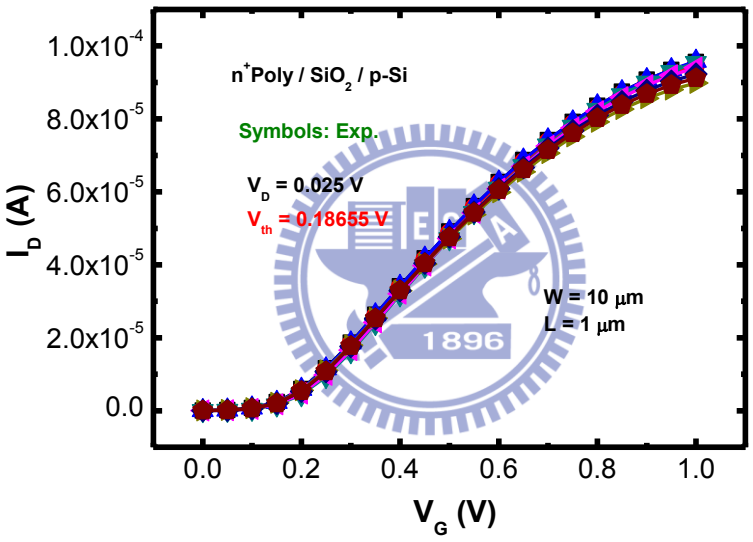


Fig. 7. The measured drain currents versus gate voltage for n^+poly gate MOSFETs.

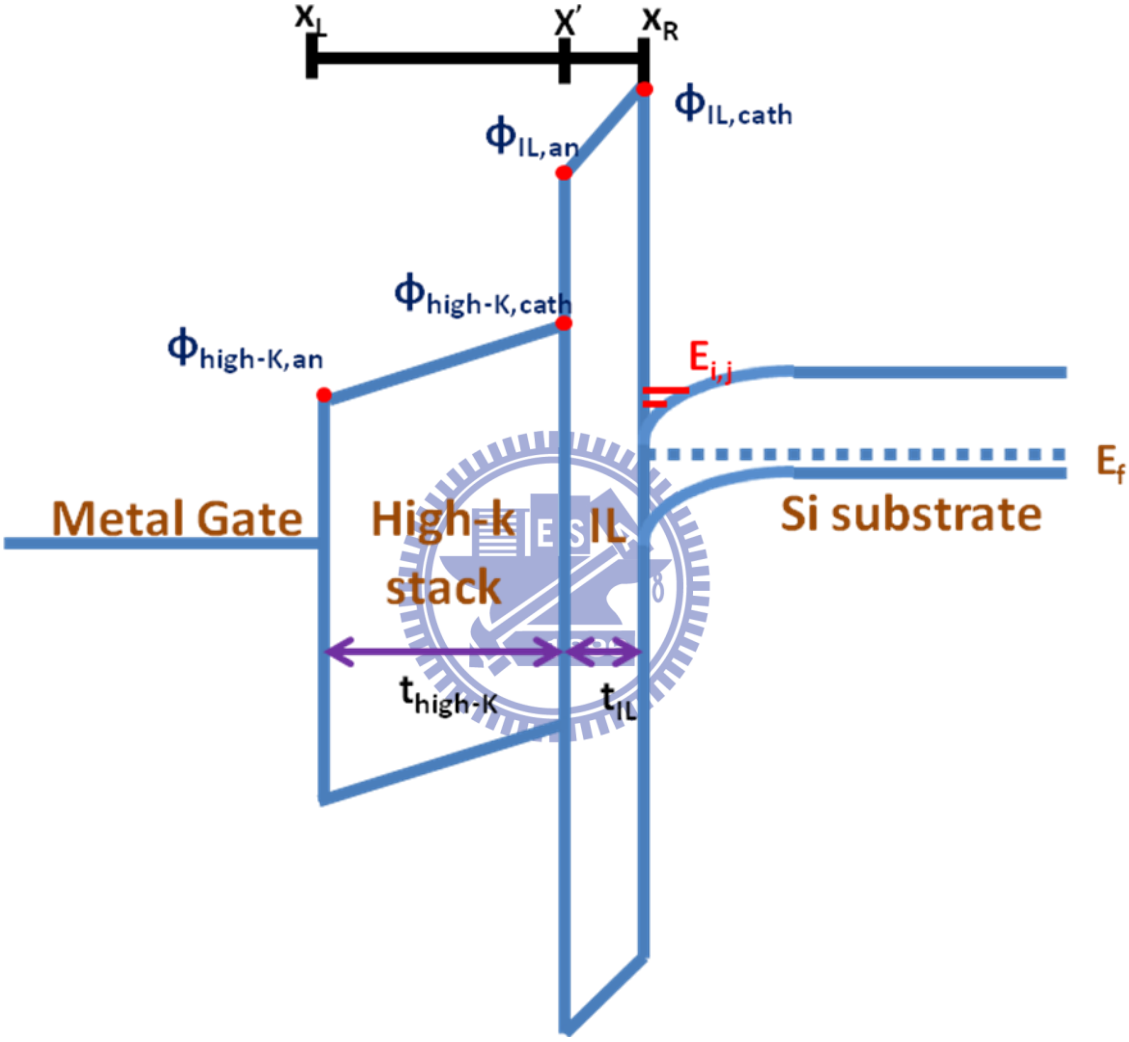


Fig. 8. The energy band diagram for the metal gate/high-K/interface layer/Si substrate system.

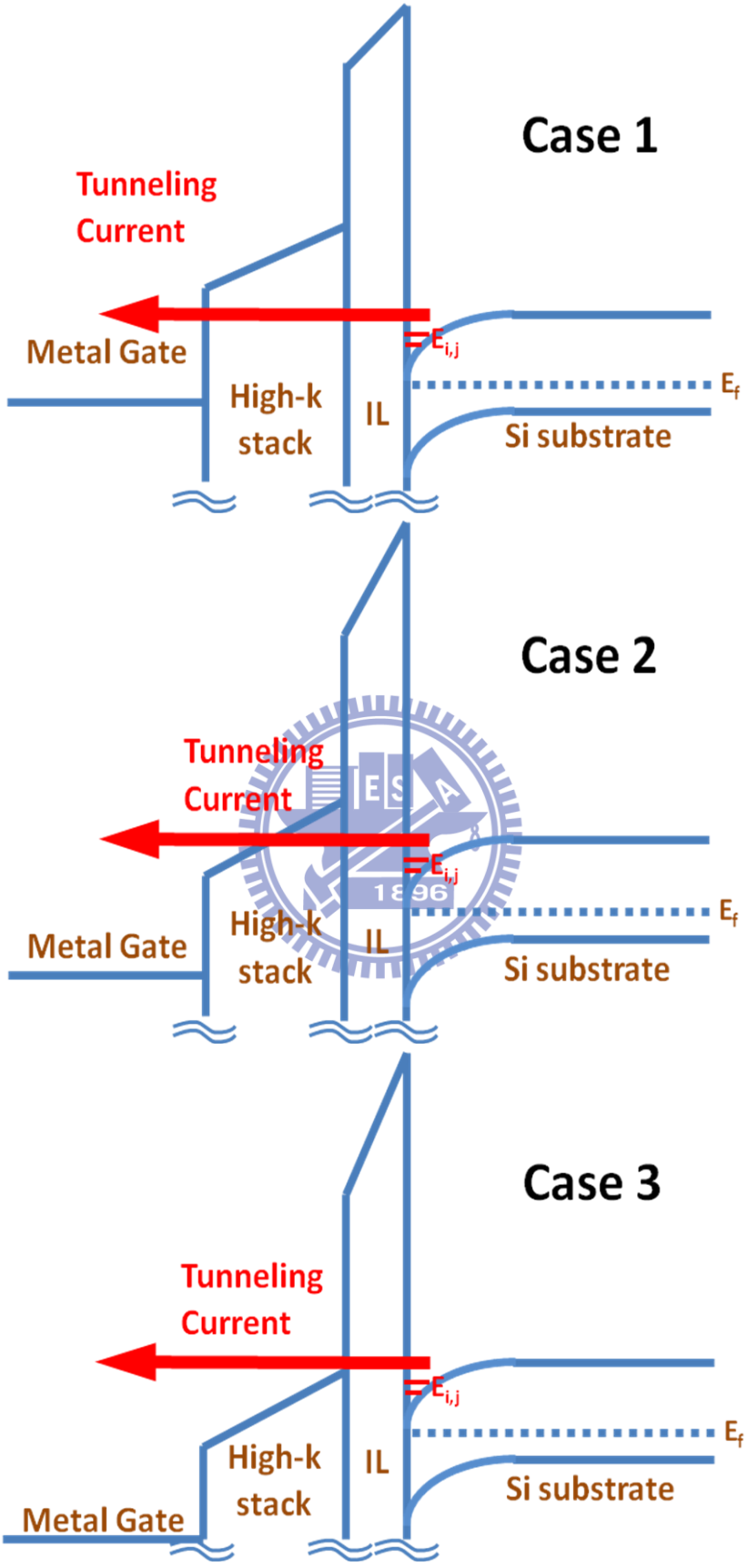


Fig. 9. Three different cases for electron tunneling in varying gate voltage.

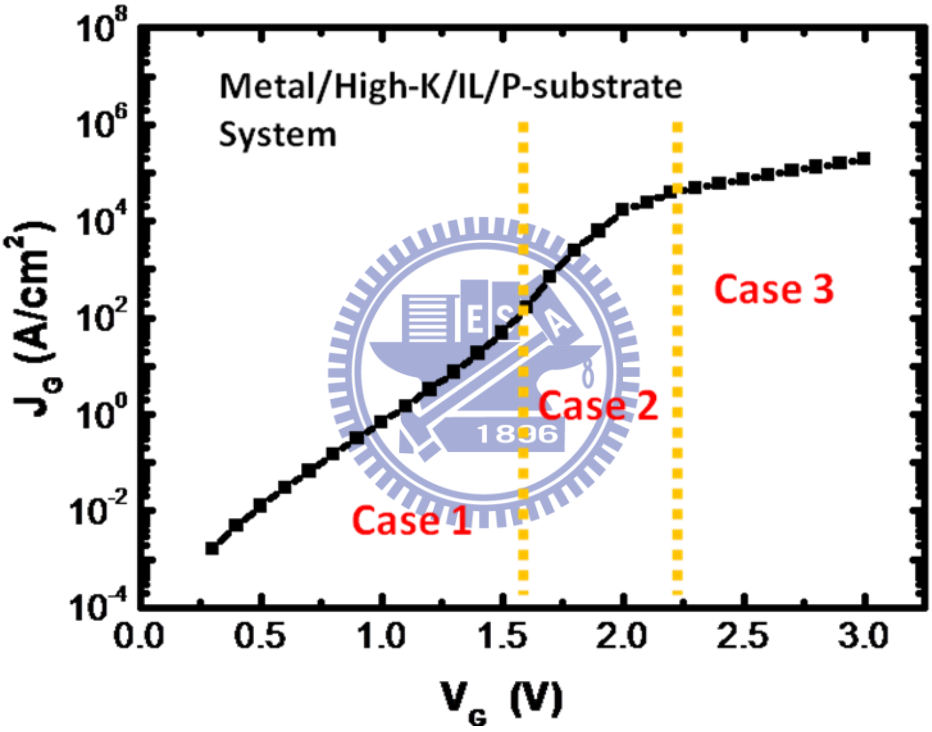
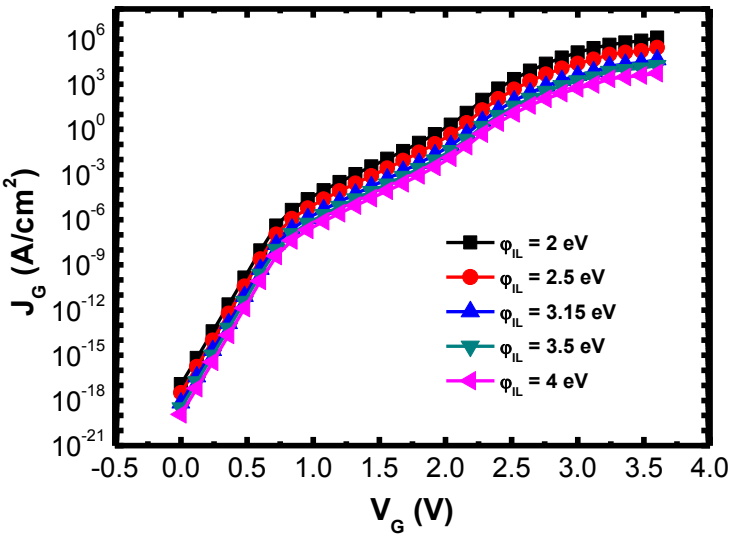
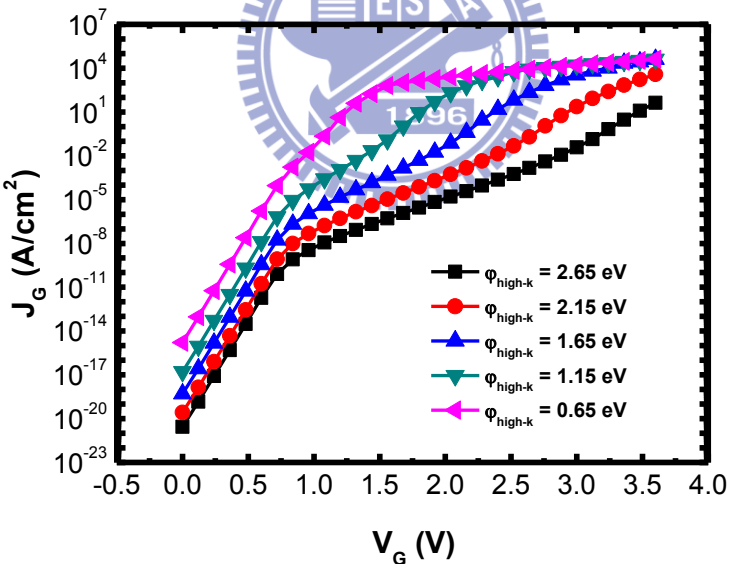


Fig. 10. The demonstration of three different tunneling regions.

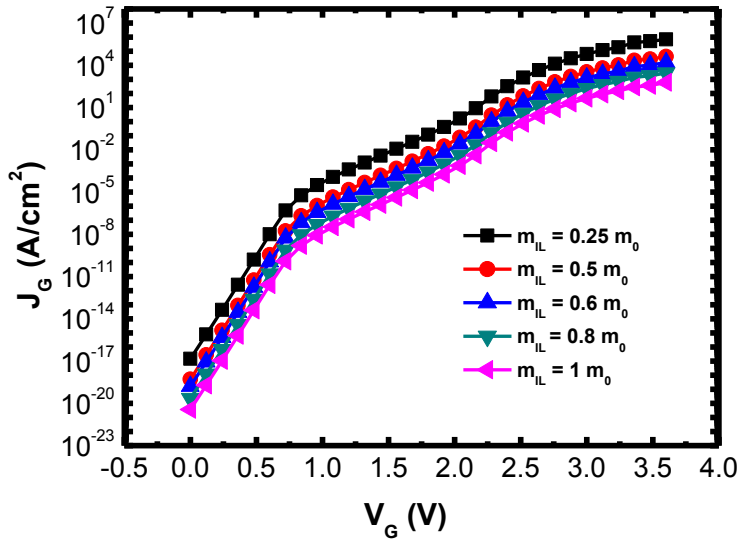


(a)

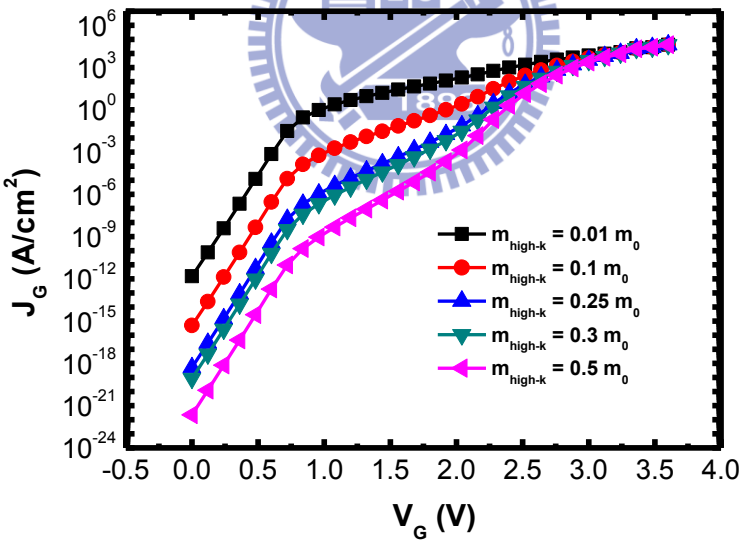


(b)

Fig. 11. The influence of (a) φ_{IL} and (b) φ_{high-K} on the calculated gate tunneling current.

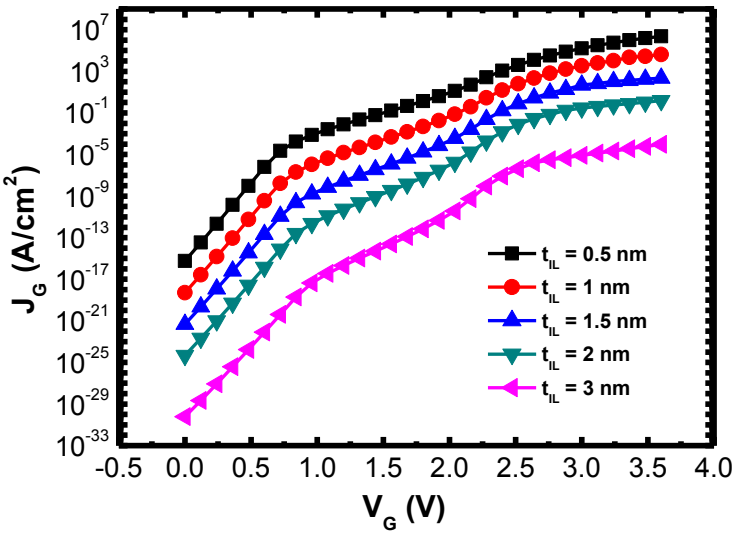


(a)

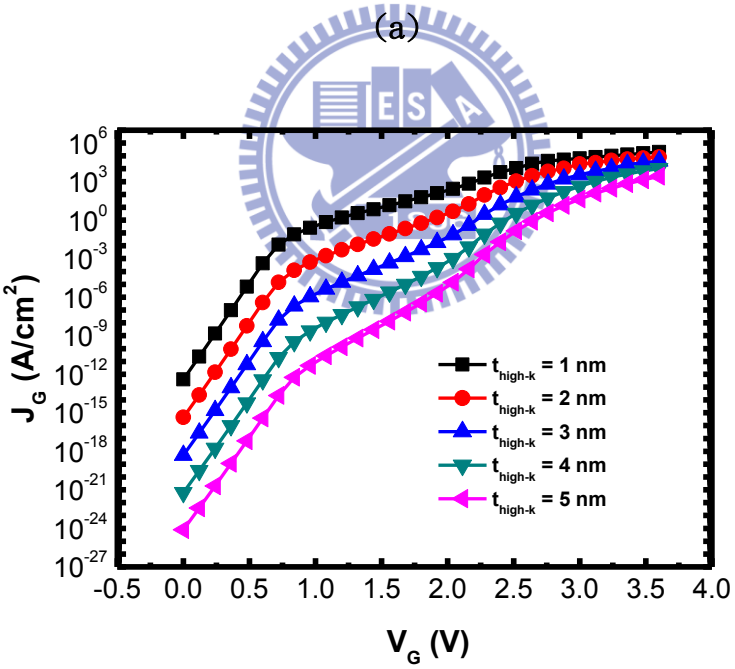


(b)

Fig. 12. The influence of (a) m_{IL} and (b) m_{high-K} on the calculated gate tunneling current.

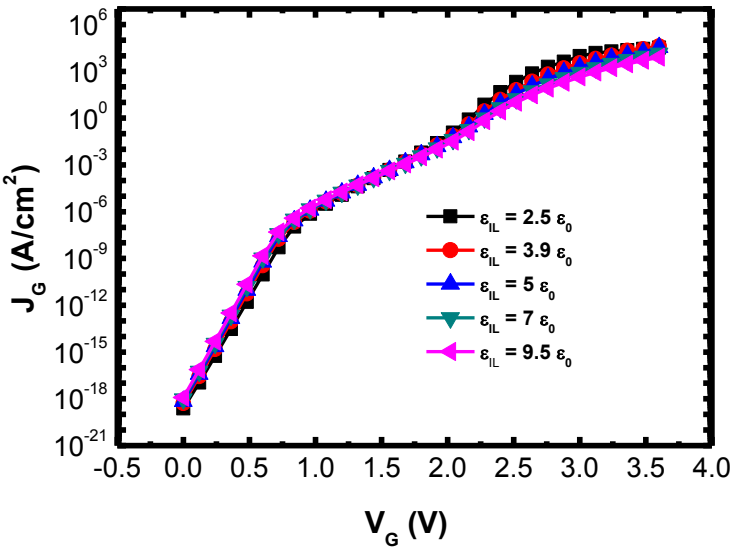


(a)

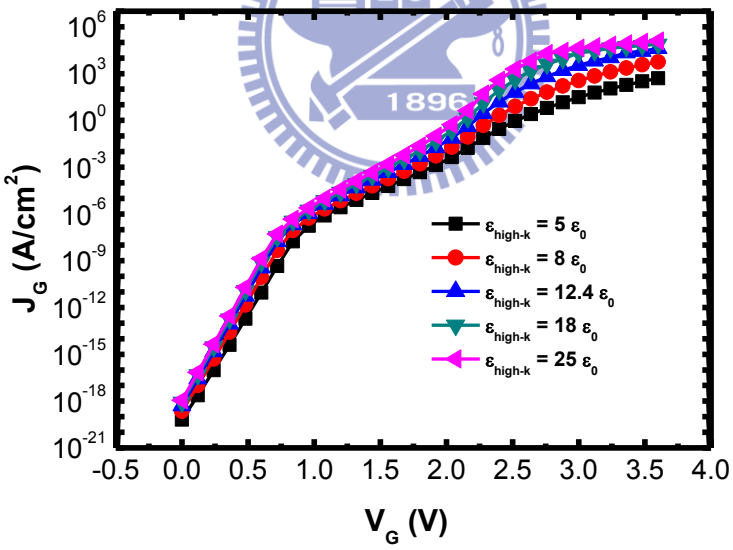


(b)

Fig. 13. The influence of (a) t_{IL} and (b) t_{high-K} on the calculated gate tunneling current.



(a)



(b)

Fig. 14. The influence of (a) ε_{IL} and (b) ε_{high-K} on the calculated gate tunneling current.

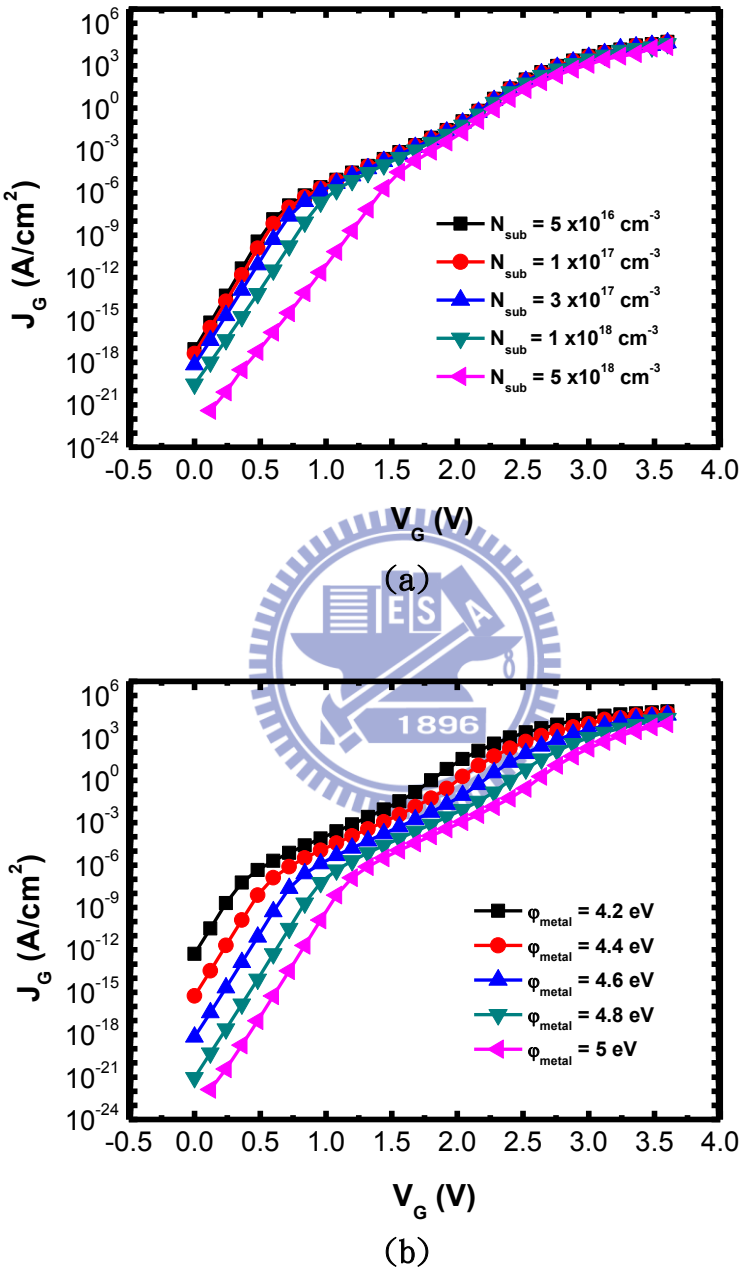


Fig. 15. The influence of (a) N_{sub} and (b) φ_{metal} on the calculated gate tunneling current.

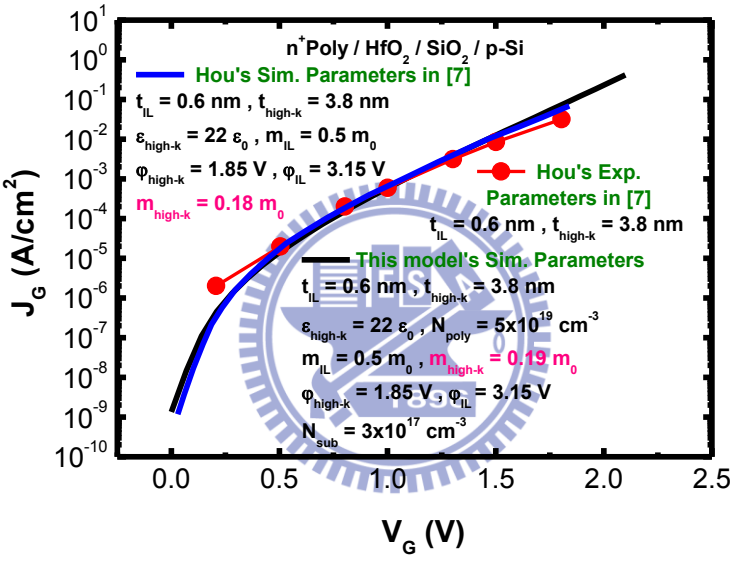


Fig. 16. The fitting results by our model in comparison with the measured tunneling current in the literature [7]. The tunneling parameters are the same as those in [7].

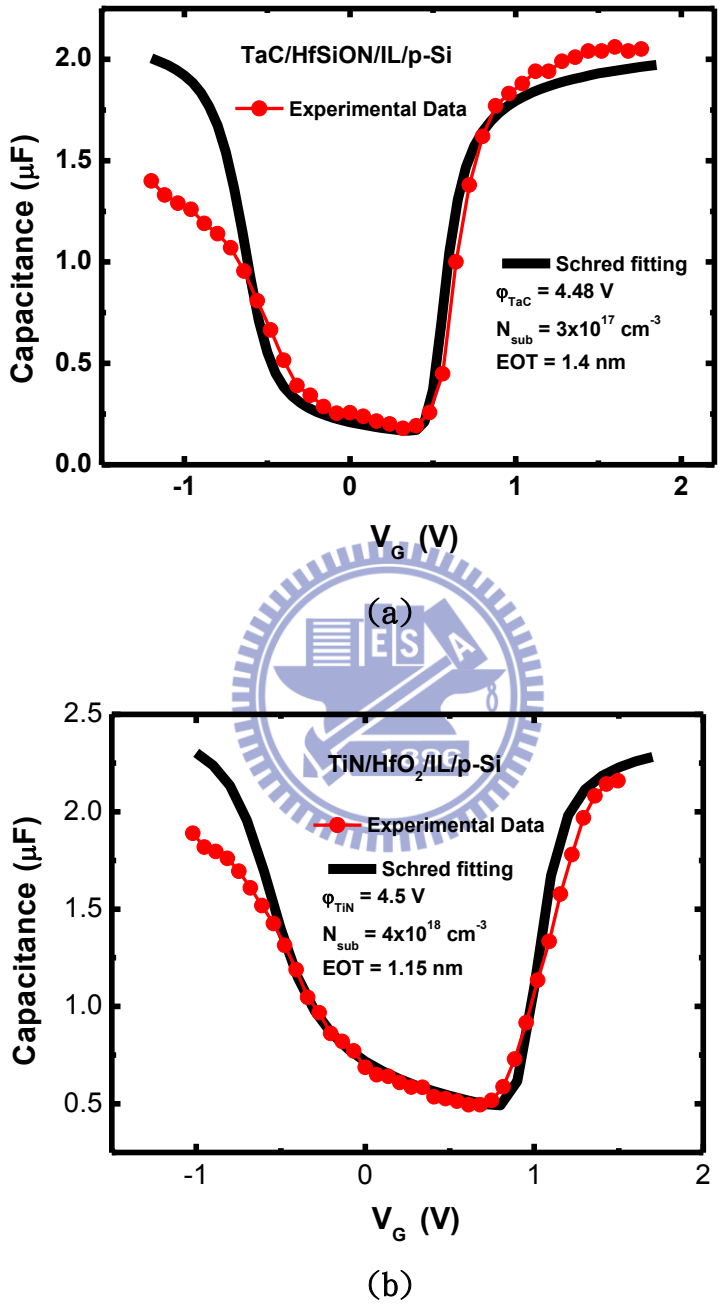


Fig. 17. The measured capacitance versus gate voltage for two different gate stacks. And it is also shown the fitting lines.

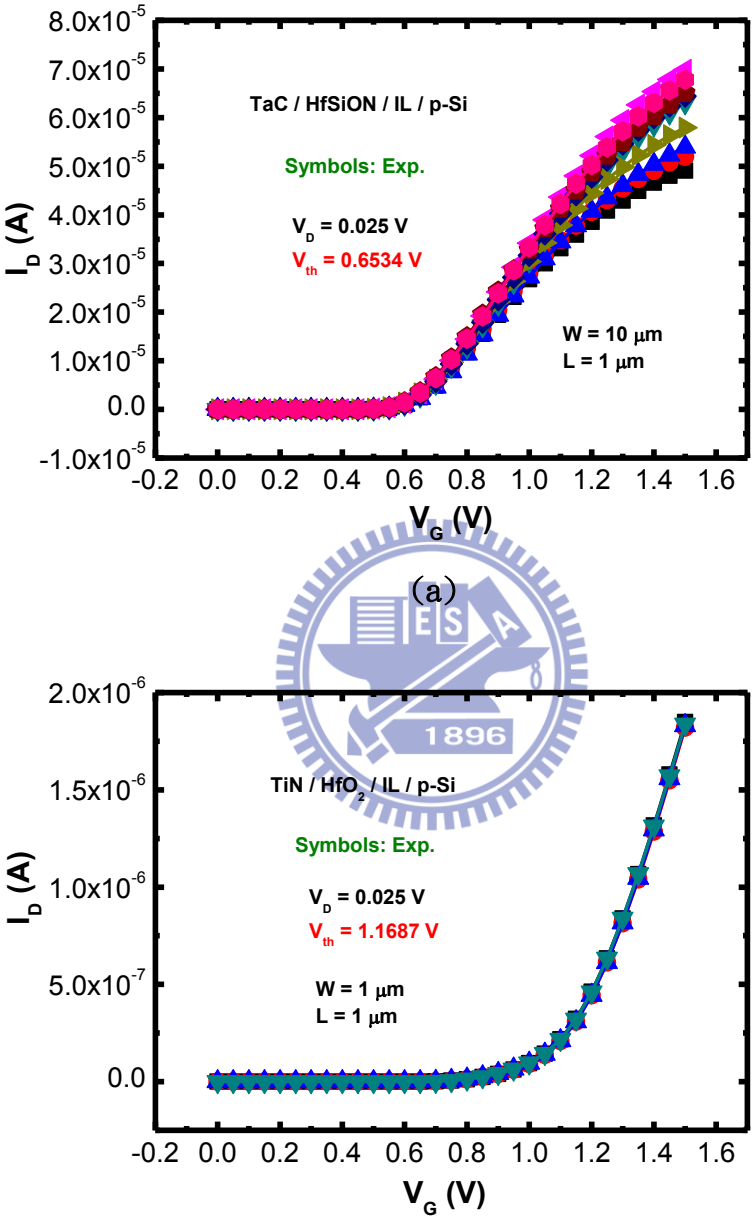


Fig. 18. The measured drain currents versus gate voltage for (a) TaC/HfSiON/IL and (b) TiN/HfO_2/IL gate stacks.

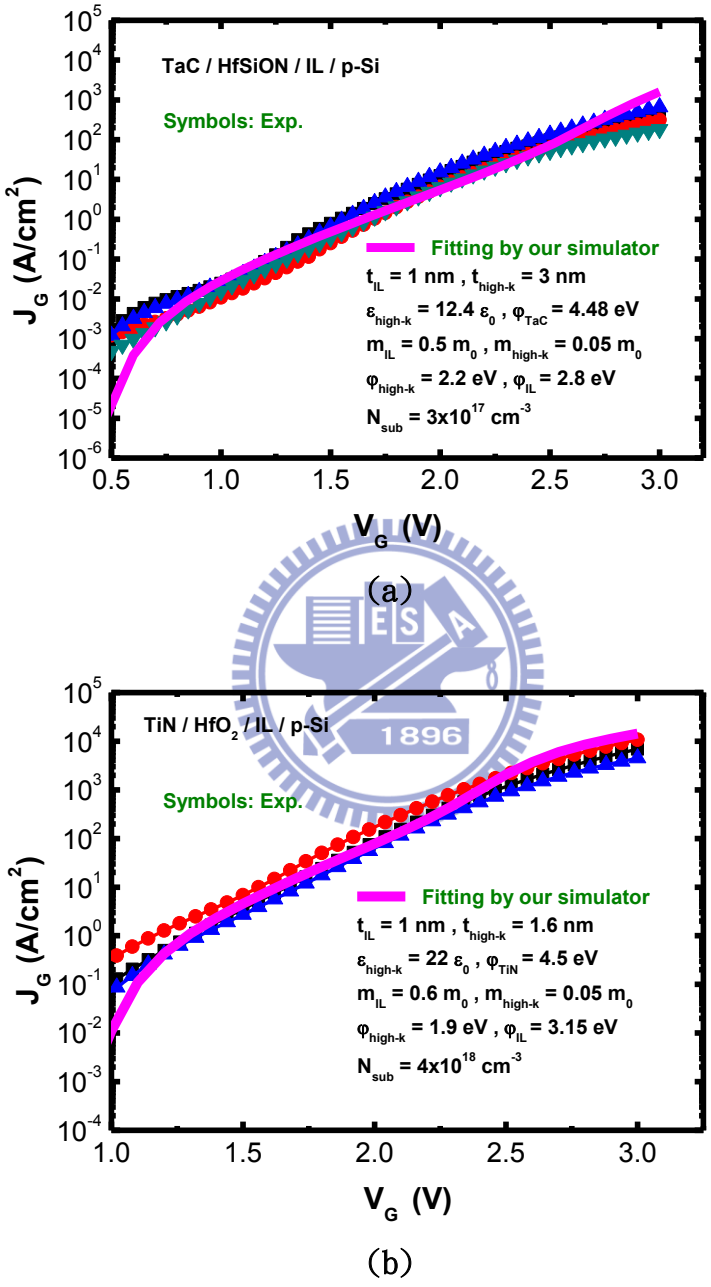


Fig. 19. Comparison of the measured and calculated gate currents versus gate voltage for (a) TaC/HfSiON/IL and (b) TiN/HfO_2/IL gate stacks.