Metal-dielectric band alignment and its implications for metal gate complementary metal-oxide-semiconductor technology

Yee-Chia Yeo, a) Tsu-Jae King, and Chenming Hu
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California 94720

(Received 31 July 2002; accepted 20 September 2002)

The dependence of the metal gate work function on the underlying gate dielectric in advanced metal-oxide-semiconductor (MOS) gate stacks was explored. Metal work functions on high-κ dielectrics are observed to differ appreciably from their values on SiO₂ or in vacuum. We applied the interface dipole theory to the interface between the gate and the gate dielectric of a MOS transistor and obtained excellent agreement with experimental data. Important parameters such as the slope parameters for gate dielectrics like SiO₂, Al₂O₃, Si₃N₄, ZrO₂, and HfO₂ were extracted. In addition, we also explain the weaker dependence of n⁺ and p⁺ polysilicon gate work functions on the gate dielectric material. Challenges for gate work function engineering are highlighted. This work provides additional guidelines on the choice of gate materials for future MOS technology incorporating high-κ gate dielectrics. © 2002 American Institute of Physics.

I. INTRODUCTION

Aggressive scaling of gate length L_G and gate oxide thickness in complementary metal-oxide-semiconductor (CMOS) transistors for higher performance and circuit density aggravates the problems of polysilicon (poly-Si) gate depletion, high gate resistance, high gate tunneling leakage current, and boron penetration from the p⁺-doped polysilicon gate into the channel region. 1 To reduce the high gate resistance and boron penetration from the depletion, high gate resistance, high gate tunneling leakage current, and boron penetration from the p⁺-doped poly-Si gate into the channel region. 1 To reduce the high gate resistance and gate depletion problems, the active dopant density in the poly-Si gate material must be increased. In fact, the active dopant density must be greater than 1.87 × 10²⁰ cm⁻³ at L_G = 25 nm CMOS technology generation for the poly-Si gate depletion layer to be less than 25% of the equivalent SiO₂ thickness t_{ox,eq}. 1 This presents a great difficulty since the active poly-Si dopant density at the gate-dielectric interface saturates at 6 × 10¹⁹ cm⁻³ and 1 × 10²⁰ cm⁻³ for p⁺- and n⁺-doped poly-Si, respectively. 2 Insufficient active dopant density in the gate results in a significant voltage drop across the gate depletion layer, and increases the equivalent gate oxide thickness, as illustrated in Fig. 1. In effect, it reduces the gate capacitance in the inversion regime and hence the inversion charge density, or leads to a lower effective gate voltage, thus compromising device performance. As a result, there is immense interest in metal gate technology. 3 A metal gate material not only eliminates the gate depletion and boron penetration problems, but also greatly reduces the gate sheet resistance. One approach is to employ a metal with midgap work function for both N- and P-MOS field-effect transistors (FET’s). 4 The drawback is that the threshold voltages |V_TH| would be too large for a reasonable channel doping concentration; counterdoping the channel to reduce V_TH degrades transistor short-channel and turn-off characteristics. 2,4 Hence, an approach analogous to the established dual-doped poly-Si gate technology, i.e., a dual-work-function metal gate technology, would be preferred. 3

The major challenge is to find two metals with suitable work functions and a way to integrate them into a CMOS process. Metal gates with work functions corresponding to the conduction and valence band edges of Si are desired for the optimal design of bulk-Si N- and P-MOSFET’s, respectively. 5 The energy band alignment between the metal gate and the gate dielectric is very important, because it determines the effective metal gate work function Φ_m,eff, which affects the transistor threshold voltage and optimum current drive. 5 An accurate understanding of the metal-dielectric interface is essential for the design of transistors with metal gates. However, there has been very little work to study this important interface. 6 In the selection of metal gate materials, a common assumption is that the effective work function of a metal on a gate dielectric is the same as that in vacuum. This has been experimentally observed to be incorrect. 6 In this paper, the physics of the interface between the gate electrode and the gate dielectric is investigated, and the dependence of the gate work functions on the gate dielectric material is explained. Section II briefly reviews the physics of the metal-semiconductor or metal-dielectric interface. Section III applies the interface dipole theory 3 to the metal-dielectric interface to explain the experimental observations and extract important material parameters that characterize the gate dielectrics. Fundamental differences between the metal-dielectric and the poly-Si-dielectric interfaces are also noted. Section IV summarizes this work and provides guidelines for the selection of metal gate materials for integration in a CMOS process.

II. THEORETICAL BACKGROUND: INTERFACE DIPOLE THEORY

According to the Schottky model, 9 when a metal and a semiconductor or a dielectric form an interface, there is no charge transfer across the interface and the barrier height for
the electrons is given by the difference between the work function of the metal in vacuum $\Phi_{m,vac}$ and the electron affinity of the semiconductor $\chi$. However, it was observed experimentally that the Schottky model is not generally obeyed. Bardeen then proposed the well-known surface state model to explain this observation. Bardeen’s model postulates a high density of surface states of the order of 1 per surface atom at a well-defined energy relative to the conduction band edge of the semiconductor, and these states act to pin the metal Fermi level. However, arguments presented by Heine showed that semiconductor surface states do not exist in the fundamental gap for most metal-semiconductor interfaces, except possibly those in Bardeen’s time where metal-semiconductor contacts were made by pressing the metal against the semiconductor, each coated with an oxide film. This interfacial oxide film sustains a potential difference to account for the difference between the theoretical Schottky barrier height and experimental observation, and is an essential feature of Bardeen’s model. Modern metal-semiconductor interfaces are comprised of materials in intimate atomic contact, and therefore render Bardeen’s surface state model inapplicable.

Heine was the first to point out that the wave functions of electrons in the metal tail or decay into the semiconductor in the energy range where the conduction band of the metal overlaps the semiconductor band gap. These resulting states in the forbidden gap are known as metal-induced gap states or simply intrinsic states. The existence of metal-induced gap states at the metal-semiconductor or metal-dielectric interface was also predicted by Louie and Cohen using self-consistent pseudopotential calculations of the electronic structure of material interfaces. Recently, metal-induced gap states were experimentally observed by Muller et al. using electron energy loss spectroscopy. These states are predominantly donorlike close to $E_v$, and mostly acceptorlike near $E_c$ (Fig. 2). The energy level in the band gap at which the dominant character of the interface states changes from donorlike to acceptorlike is called the charge neutrality level $E_{CNL}$. Charge transfer generally occurs across the interface due to the presence of intrinsic interface states. Charging of these interface states creates a dipole that tends to drive the band lineup toward a position that would give zero dipole charge. Figure 2 illustrates the case where the metal Fermi level $E_{F,m}$ is above the charge neutrality level in the dielectric $E_{CNL,d}$, creating a dipole that is charged negatively on the dielectric side. This interface dipole drives the band alignment so that $E_{F,m}$ goes toward $E_{CNL,d}$ and the effective metal work function, $\Phi_{m,eff}$, therefore differs from the vacuum metal work function $\Phi_{m,vac}$. This work function change is proportional to the difference between $E_{CNL,d}$ and $E_{F,m}$, or equivalently, the difference between $\Phi_{m,vac}$ and $\Phi_{CNL,d} = (E_{vac} - E_{CNL,d})/q$. Thus, $\Phi_{m,eff}$ is given by

$$\Phi_{m,eff} = \Phi_{CNL,d} + S(\Phi_{m,vac} - \Phi_{CNL,d}),$$

where $S$ is a slope parameter that accounts for dielectric screening and depends on the electronic component of the dielectric constant $\varepsilon_e$. With larger dielectric screening, the slope parameter $S$ becomes smaller, as shown in Figs. 3(a) where the slope parameters for a variety of materials are documented. The slope parameter $S$ obeys an empirical relationship given by
Materials with a smaller $S$ tend to pin the metal Fermi level more effectively to $E_{\text{CNL}}$, as illustrated in Fig. 3b. The maximum value for $S$ is unity, which corresponds to no pinning of the metal Fermi level.

### III. EXPERIMENTAL DATA AND EXTRACTION OF MATERIAL PARAMETERS

Table I summarizes the experimental work function data for metals in vacuum and metals on dielectrics such as SiO$_2$, Al$_2$O$_3$, ZrO$_2$, HfO$_2$, and Si$_3$N$_4$. The work function of a metal is a measure of the minimum energy required to extract an electron from the surface of the metal. It is commonly measured in vacuum by making use of phenomena such as the photoelectric effect or thermionic emission. In general, it differs for each facet of a monocrystalline sample.\textsuperscript{16,19,22} Partial crystalline orientation of film specimens can account for observed differences in some cases, since anisotropies of up to \~0.5 eV exist for the different crystal facets. Given that the work function measurement is also dependent on the cleanliness of the metal surface, data reported in the literature often cover a considerable range, and Table I documents the typical values for a reasonably clean surface. For the measurement of a metal work function in a metal-dielectric system, the most common method is an extraction of the flat-band voltage from the capacitance-voltage characteristics of a metal-dielectric-semiconductor capacitor.\textsuperscript{23,29} Internal photoemission may also be used to extract the barrier height at the interface of a metal-dielectric system,\textsuperscript{7} and the metal work function may be deduced from the barrier height and the electron affinity of the dielectric. The electron affinities of SiO$_3$, Al$_2$O$_3$, and ZrO$_2$ are 0.95, 1.0, and 2.5 eV, respectively.\textsuperscript{30}

The experimental data (symbols) are plotted in Fig. 4 to illustrate the varying degrees of pinning of metal work functions toward $E_{\text{CNL,d}}$ of gate dielectrics such as SiO$_2$, Al$_2$O$_3$, ZrO$_2$, HfO$_2$, and Si$_3$N$_4$. A fit of Eq. (1) to the experimental data not only reveals the good agreement between the interface dipole theory and measured data, but also yields the first extraction of $E_{\text{CNL,d}}$ and $S$ for dielectrics like ZrO$_2$, HfO$_2$, and Si$_3$N$_4$. The extracted $(E_{\text{CNL,d}} - E_p)$ values obtained in this work are compared with theoretical values\textsuperscript{30} in Table II, showing reasonable agreement. It should be noted that theoretical values of $E_{\text{CNL,d}}$ are derived from electronic band structure calculations, the local density approximation,\textsuperscript{32}

\begin{equation}
S = \frac{1}{1 + 0.1(\epsilon_n - 1)^2}.
\end{equation}

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|}
\hline
Metal/dielectric & Work function (eV) & Barrier height (eV) & Measurement method \tabularnewline
\hline
\hline
Mg & 3.66 & & Photoelectric effect \tabularnewline
Mg/Al$_2$O$_3$ & 3.6 & 2.6 & Internal photoemission \tabularnewline
Mg/SiO$_2$ & 3.45 & 2.5 & Internal photoemission \tabularnewline
Mg/ZrO$_2$ & 4.15 & 2.6 & Internal photoemission \tabularnewline
Ta (polycrystalline) & 4.25 & & Thermionic emission \tabularnewline
Ta/SiO$_2$ & 4.2 & & MOS capacitor $V_{FB}$ \tabularnewline
Al & 4.28 & & \tabularnewline
Al/Al$_2$O$_3$ & 3.9 & 2.9 & Internal photoemission \tabularnewline
Al/SiO$_2$ & 4.14 & & MOS capacitor $V_{FB}$ \tabularnewline
Al/Si$_3$N$_4$ & 4.06 & & MOS capacitor $V_{FB}$ \tabularnewline
Al/ZrO$_2$ & 4.25 & 2.7 & Internal photoemission \tabularnewline
W & 4.63 & & Field emission \tabularnewline
W/SiO$_2$ & 4.6–4.7 & 3.65–3.75 & Fowler-Nordheim tunneling \tabularnewline
\hline
Mo(110) & 4.95 & & Photoelectric effect \tabularnewline
Mo/SiO$_2$ & 5.05 & & MOS capacitor $V_{FB}$ \tabularnewline
Mo/Si$_3$N$_4$ & 4.76 & & MOS capacitor $V_{FB}$ \tabularnewline
Mo/HfO$_2$ & 4.95 & & MOS capacitor $V_{FB}$ \tabularnewline
Pt & 5.65 & & Photoelectric effect \tabularnewline
Pt/SiO$_2$ & 5.59 & & MOS capacitor $V_{FB}$ \tabularnewline
Pt/HfO$_2$ & 5.23 & & MOS capacitor $V_{FB}$ \tabularnewline
Pt/ZrO$_2$ & 5.05 & & MOS capacitor $V_{FB}$ \tabularnewline
Ni(110) & 5.04 & & Photoelectric effect \tabularnewline
Ni/Al$_2$O$_3$ & 4.5 & 3.5 & Internal photoemission \tabularnewline
Ni/ZrO$_2$ & 4.75 & 3.25 & Internal photoemission \tabularnewline
Au & 5.31–5.47 & & Photoelectric effect \tabularnewline
Au/Al$_2$O$_3$ & 5.1 & 4.1 & Internal photoemission \tabularnewline
Au/ZrO$_2$ & 5.05 & 3.5 & Internal photoemission \tabularnewline
Hf (polycrystalline) & 3.95 & & Photoelectric effect \tabularnewline
Hf/SiO$_2$ & 4.0 & & MOS capacitor $V_{FB}$ \tabularnewline
\hline
\end{tabular}
\caption{Experimental data on metal work functions and conduction band barrier heights at metal-dielectric interfaces.}
\end{table}

\begin{equation}
S = \frac{1}{1 + 0.1(\epsilon_n - 1)^2}.
\end{equation}
which results in an underestimation of the band-gap energy, is often employed. Thus, theoretically derived values of $E_{\text{CNL},d}$ with reference to the conduction or valence bands should be used with caution. For this reason, our experimental extraction of $(E_{\text{CNL},d} - E_v)$ should be more accurate and useful.

Since metal gates with work functions near the conduction and valence band edges of Si are ideal for the optimum performance of bulk N- and P-MOSFET’s, respectively, the desired effective metal work function should be $\sim 4.05$ V for the N-MOSFET and $\sim 5.17$ V for the P-MOSFET. Consider ZrO$_2$ for example. Figure 4(c) suggests that in order to obtain $\Phi_{\text{m,eff}}$ of 4.05 V (or 5.17 V) for a N-MOS (or P-MOS) device, a metal with an even smaller (or larger) $\Phi_{m,\text{vac}}$ has to be used. In fact, the selection of metal gate materials to achieve the desired effective work functions depends on the choice of the gate dielectric materials. As a result, research work on metal gate technology should be performed in close conjunction with work on alternative gate dielectrics. This dependency is examined in Fig. 5 where the required $\Phi_{m,\text{vac}}$

![Graph](image)

FIG. 4. Effective work functions of metals on various dielectrics such as (a) SiO$_2$, (b) Al$_2$O$_3$, (c) ZrO$_2$, (d) HfO$_2$, and (e) Si$_3$N$_4$ versus their work functions in vacuum. Good agreement is observed between experimental data (symbols) and the theoretical fit (lines) based on the interface dipole theory.

<table>
<thead>
<tr>
<th>Material</th>
<th>$E_{\text{G}}$ (eV)</th>
<th>$\varepsilon_s$</th>
<th>$E_{\text{CNL}} - E_v$ (eV)</th>
<th>$S$</th>
<th>$E_{\text{CNL}} - E_v$ (eV)</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>1.12</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ge</td>
<td>0.66</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SiO$_2$</td>
<td>9</td>
<td>2.25</td>
<td>4.5</td>
<td>0.86</td>
<td>5.04</td>
<td>0.95</td>
</tr>
<tr>
<td>Al$_2$O$_3$</td>
<td>8.8</td>
<td>3.4</td>
<td>5.5$^b$</td>
<td>0.63</td>
<td>6.62</td>
<td>0.69</td>
</tr>
<tr>
<td>Si$_3$N$_4$</td>
<td>5.3</td>
<td>3.8</td>
<td>2.6$^c$</td>
<td>0.56</td>
<td>2.79</td>
<td>0.59</td>
</tr>
<tr>
<td>HfO$_2$</td>
<td>6.0</td>
<td>4.0</td>
<td>3.7$^b$</td>
<td>0.53</td>
<td>3.64</td>
<td>0.52</td>
</tr>
<tr>
<td>ZrO$_2$</td>
<td>5.8</td>
<td>4.8</td>
<td>3.6$^b$</td>
<td>0.41</td>
<td>3.82</td>
<td>0.52</td>
</tr>
</tbody>
</table>

$^a$Taken from Ref. 8.

$^b$Taken from Ref. 30; electron band structures obtained using tight-binding method.

$^c$Calculated from band structures obtained using the self-consistent orthogonalized linear combination of atomic orbitals method (Ref. 32).

$^d$Determined from empirical model: $S = [1 + 0.1(\varepsilon_s - 1)^2]^{-1}$. 

TABLE II. Comparison of theoretical and experimental slope parameters and charge neutrality levels for several gate dielectrics.
to achieve $\Phi_{m,\text{eff}} = 4.05$ V for a N-MOS (solid circles) and $\Phi_{m,\text{eff}} = 5.17$ V for a P-MOS (solid squares) device is plotted as a function of $\varepsilon_\infty$ for five different dielectrics. The data points in Fig. 5 are obtained using the extracted values of $E_{\text{CNL},d}$ and $S$ (Table II). For a high-$\kappa$ gate dielectric, a high-work-function inert metal must be used as the P-MOS gate electrode. High-work-function metals such as platinum are resistant to chemical or plasma etching, so gate patterning would be particularly challenging. On the other hand, a low-work-function metal must be used as the N-MOS gate electrode. Low-work-function metals such as magnesium and hafnium are extremely reactive, and this might introduce extrinsic interface states due to defects arising from an interfacial reaction. Defect-related extrinsic interface states should be distinguished from intrinsic interface states. Any large deviation of the measured effective work functions from the linear trend in Fig. 4 may be attributed to the existence of extrinsic interface states. We have carefully eliminated data that show evidence of an interfacial reaction. To avoid the need for either an inert or reactive metal gate, a possible solution is to introduce at least a monolayer of SiO$_2$ at the metal-dielectric interface to achieve a large $S$. Nevertheless, this approach leads to a lower effective dielectric permittivity for the gate dielectric stack and compromises gate leakage current and power consumption.

It is worthwhile to compare the metal gate technology with the well-established dual-doped poly-Si technology on the issue of electrode work function dependence on the choice of gate dielectric materials. It has recently been shown that poly-Si on HfO$_2$ and SiO$_2$ has about the same work function. The important characteristic that differentiates the poly-Si-dielectric interface from the metal-dielectric interface is the absence of interface states at energies where the band gaps of both materials overlap (Fig. 6). The band lineup at the poly-Si-dielectric interface is thus driven by a smaller dipole charge. Mathematically, the effective work function of $n^+$ or $p^+$ poly-Si, $\Phi_{\text{Si,eff}}$, is determined by the alignment of $E_{\text{CNL,Si}}$ and $E_{\text{CNL,d}}$.

$$\Phi_{\text{Si,eff}} = \Phi_{\text{Si,vac}} + (S-1)(E_{\text{CNL,Si}} - E_{\text{CNL,d}}).$$

In Eq. (3), $\Phi_{\text{Si,vac}}$ is the work function of $n^+$ or $p^+$ poly-Si in vacuum, and $\Phi_{\text{CNL,Si}}$ is the charge neutrality level in Si. Using Eqs. (1) and (3), the theoretical effective work functions of $n^+$ poly-Si ($\Phi_{\text{Si}} = 4.05$ V) and a metal with $\Phi_{m,\text{vac}} = 4.05$ V on different high-$\kappa$ dielectrics are compared in Fig. 7. It shows that the $n^+$ poly-Si gate is less vulnerable to Fermi pinning compared to the metal gate with $\Phi_{m,\text{vac}} = 4.05$ V. This is also true for $p^+$ poly-Si ($\Phi_{\text{Si}} = 5.17$ V) and a metal with $\Phi_{m,\text{vac}} = 5.17$ V.

**IV. CONCLUSION**

The metal-dielectric interface in advanced MOS gate stacks was examined. Metal work functions on high-$\kappa$ dielectric materials have been explored in detail. In particular, the effective work functions of $n^+$ or $p^+$ poly-Si have been calculated using the charge neutrality level and the band lineup at the poly-Si-dielectric interface. The results show that the $n^+$ poly-Si gate is less vulnerable to Fermi pinning compared to the metal gate with the same work function. This is also true for $p^+$ poly-Si. Poly-Si or poly-SiGe gates do not suffer appreciably from this effect.
ZrO$_2$, and HfO$_2$ were extracted. To achieve the desired dual-vacuum work functions smaller than 5.17 V for SiO$_2$ at the interface could relax this requirement. Additional guidelines were provided for the choice of gate materials for future CMOS technology incorporating high-$\kappa$ gate dielectrics. The effective work functions of $n^+$ and $p^+$ poly-Si gates show less dependence on the gate dielectric material. This weaker dependence on the gate dielectric material was explained.

**ACKNOWLEDGMENTS**

Y.-C. Y. acknowledges financial support from a NUS OGS Award, Singapore, and an IEEE Electron Device Society Graduate Student Fellowship Award.


$^9$W. Schottky, Phys. Z. 41, 570 (1940).


$^{29}$E. H. Nicollian and J. R. Brew, Metal Oxide Semiconductor Physics and Technology, (Wiley, New York, 1987); see Chap. 10, pp. 423–491, for work functions of metals on SiO$_2$.


