Chapter 5
Metal Gate Electrodes

Jamie K. Schaeffer

Abstract Metal gate electrodes are a vital enabler for the use of high-k gate dielectrics in advanced complimentary metal oxide semiconductor (CMOS) technology. This chapter will detail how metal gate electrodes were selected over poly-Si electrodes to be used in conjunction with high-k gate dielectrics, how metal gate electrodes are an important component of device scaling, and how the of the metal gate can be tuned to optimize the device performance. In this chapter you will see how the properties of the metal gate are inseparable from those of the high-k gate dielectric and how the metal gate can be manipulated to intentionally influence the properties of the dielectric material to produce a desired device response. This chapter will discuss the different ways in which metal gates can be used to change the effective work function of a MOS transistor, including the vacuum work function, dielectric/metal capping layers, and oxygen vacancy manipulation. In each of these cases there is a clear interplay between the metal gate and the gate dielectric and it is important to understand both materials systems in order to understand the device response. Finally, this chapter will detail the common approaches of integrating metal gate electrodes in a complimentary metal oxide semiconductor. Because there are various methods of modulating the effective work functions by using metal gate electrodes, there are multiple integration schemes that have been devised in order to produce optimized nMOS and pMOS transistors.

5.1 Reasons for Using Metal Gate Electrodes

In almost all instances metal gate electrodes are being introduced simultaneously with high-k gate dielectrics. There are two primary reasons for this. The first is that metal gate electrodes eliminate poly-silicon depletion and help with MOSFET

J. K. Schaeffer (✉)
GLOBALFOUNDRIES, Dresden 01109, Germany
e-mail: jamie.schaeffer@globalfoundries.com

S. Kar (ed.), *High Permittivity Gate Dielectric Materials*,
Springer Series in Advanced Microelectronics 43, DOI: 10.1007/978-3-642-36535-5_5,

© Springer-Verlag Berlin Heidelberg 2013

(Metal Oxide Semiconductor Field Effect Transistor) scaling. The second became apparent when researches first attempted to introduce high-k gate dielectrics with poly-Si electrodes, and noticed an incompatibility between high-k based gate dielectrics and poly-silicon electrodes that results in the pinning of the work functions of n+ and p+ poly-silicon to undesirable values.

5.1.1 Elimination of Gate Depletion

Since increasing the gate capacitance is required for increasing MOSFET drive current, device scaling mandates a consideration of factors other than the gate oxide that contribute to the inversion capacitance. When a MOSFET is operated in inversion there are actually two additional capacitances in series with the oxide capacitance. Taking these factors into consideration the total capacitance in inversion, C_{inv}, is given by

$$\frac{1}{C_{inv}} = \frac{1}{C_{sub}} + \frac{1}{C_i} + \frac{1}{C_{poly}}$$

where C_i is the aforementioned oxide capacitance, C_{sub} is a capacitance due to quantum mechanical effects which force the centroid of inversion charge in the substrate to be a few Ångstroms away from the Si/SiO_2 interface, and C_{poly} is due to a gradual potential drop, or band bending, [1] in the poly-Si gate electrode (Fig. 5.1).

The C_{poly} contribution, which is frequently referred to as poly-silicon depletion, is a result of the sum of inversion and depletion charges in the substrate being greater than the impurity density (N_{poly}) near the poly-Si—oxide interface. Since charge neutrality requires that the field lines for every carrier in the substrate are screened by the ionized impurities in the electrode, the field lines from the substrate penetrate a finite distance into the poly-Si electrode before encountering enough charges to cancel the sum of their electric fields. The screening of charge over a finite distance into the poly-Si is the origin of the gradual potential drop and band bending in the poly-Si electrode.

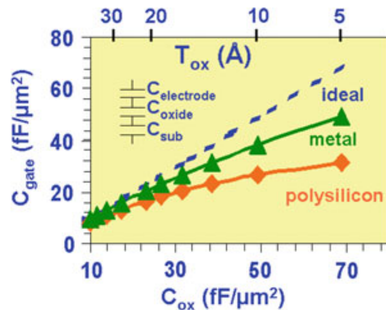


Fig. 5.1 The total gate inversion capacitance (C_{gate}) versus the oxide capacitance (C_{ox}) for metal gate electrodes and for poly-silicon electrodes. Higher C_{gate} with metal gate electrodes is due to the absence of gate-depletion effects

To overcome the additional inversion capacitance, poly-Si electrodes can be replaced with metal electrodes. Since metals have a shorter Debye length than poly-silicon, meaning a metal has a higher carrier concentration and is more effective at screening charge, no band bending occurs with metal electrodes. This eliminates the depletion capacitance and assists device performance.

5.1.2 Incompatibility of High-k Materials with Poly-Si Gates

Although there were many attempts to first introduce high-k with poly-Si electrodes, this encountered multiple challenges. These challenges include nFET and pFET threshold voltages that were higher than were expected based on n+ and p+ doped poly with SiO_2 electrodes and the observance of some physical reactions that occurred at the high-k/Si interface.

An experiment by Hobbs et al. [2], was one of the first experiments that investigated the high threshold voltages observed in poly-Si gated HfO_2 gate dielectrics. In this experiment, Hobbs et al. deposited sub-monolayers of HfO_2 on the surface of a 23 Å thermally grown SiO_2 gate dielectric (Fig. 5.2).

Atomic layer deposition (ALD) was used to deposit the sub-monolayers of increasing surface coverage of HfO_2. When analyzing poly-Si/HfO_2/n-type Si capacitors Hobbs et al. reported monotonic shifts in the flat band voltage electrodes that encouraged more researchers to evaluate flat band voltage of poly-Si with increasing HfO_2 surface coverage (Fig. 5.3). Compared to a SiO_2/Poly-Si system, flat band voltage shifts are observed with only one cycle of ALD HfO_2, and the shifts saturate at approximately 20 cycles, or when complete surface coverage is obtained.

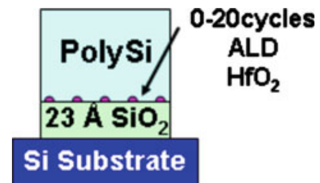


Fig. 5.2 Schematic showing the experimental technique used by Hobbs et al. to study the Fermi level pinning of poly-silicon electrodes on HfO_2

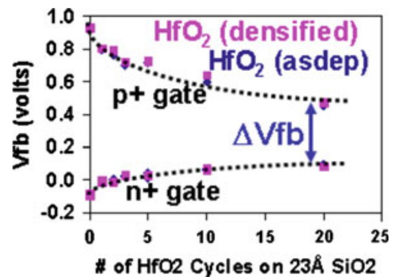


Fig. 5.3 V_{fb} shifts as a function of ALD precursor cycles for HfO_2 growth. The shifts tend to saturate when complete HfO_2 surface is obtained [2]

High-k dielectrics also had many undesirable physical interactions with poly-Si electrodes that encouraged more researchers to evaluate metal gate electrodes. The problems were first noticed when ZrO_2 was studied in conjunction with poly-Si electrodes. Large $ZrSi_x$ nodules would form at the ZrO_2/poly-Si interface demonstrating an incompatibility at the ZrO_2/poly-Si interface [3]. The gross $ZrSi_x$ formation at this interface is the primary reason HfO_2 became the preferred gate dielectric material even though ZrO_2 and HfO_2 are chemically very similar materials.

5.2 Work Function Considerations for Metal Gate Electrodes

This section focuses on the desired gate work function characteristics for metal gate electrodes and various methods of achieving the desired effective work function.

5.2.1 Work Function Requirements of Devices

For bulk CMOS devices, simulations indicate that the desired work function for NMOS (PMOS) electrodes is near the conduction (valence) band edge of silicon [4, 5]. These simulation studies imposed the constraint of a constant off current (I_{off}) for a sub-nominal length device. This constraint means that if the metal work function is shifted towards the middle of the Si band gap, the substrate doping has to be reduced to maintain a constant I_{off}. Reducing the substrate doping can benefit performance by improving carrier mobility. However, reducing the substrate doping too far also degrades the short channel performance of a MOSFET by lowering the potential barrier and increasing the depletion width between the source and the channel regions. As the gate length is reduced electric fields from the drain contribute to lowering this potential barrier. This phenomenon is referred to as drain induced barrier lowering (DIBL). DIBL manifests itself as a degraded sub-threshold swing, a measure of how rapidly the drain current responds to the gate bias. A degraded sub-threshold swing results in a lower inversion charge density for the same gate voltage which degrades the current drive of the MOSFET in saturation ($I_{d,sat}$). This suggests that work functions closer to $\sim$4.1 eV ($\sim$5.2 eV) near the conduction (valence) band edge of silicon are required to optimize NMOS (PMOS) device performance for short channel MOSFET devices.

It should also be mentioned that work function engineering provides an alternative means of achieving multiple threshold voltages on a semiconductor chip. Traditionally with poly-Si, the effective work function could not be engineered and therefore the only way to achieve multiple threshold voltages V_t was to change the

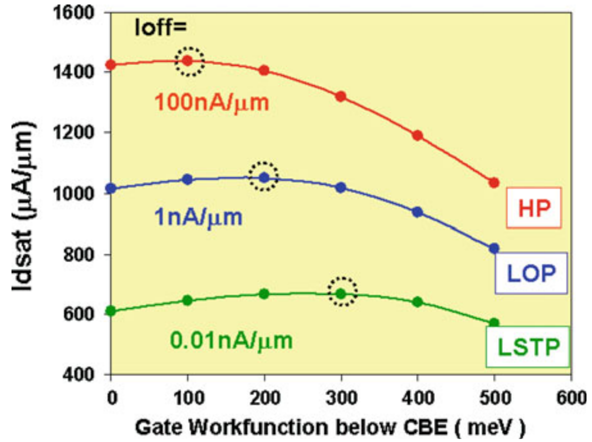


Fig. 5.4 Simulation of I_{dsat} as a function of nMOS gate work function relative to the silicon conduction band for different I_{off} targets. A maximum drive-current is achieved by using an off-band-edge metal gate to increase device V_t, and then lowering the substrate doping to increase mobility [5]

gate length or to increase the substrate well doping. One benefit of using metal gate electrodes is that a high-V_t device can be achieved by moving the metal work function towards mid-gap without increasing the substrate doping. This results in an inherent mobility and ultimately performance advantage for the devices in the circuit that utilize higher threshold voltages (Fig. 5.4).

5.2.2 *Methods of Achieving Desired Effective Work Function*

To achieve the appropriate threshold voltages in the devices, researchers are exploiting various methods broadly termed "work function engineering". Along with manipulating the metal work function to shift threshold voltages, there has also been a significant amount of work dedicated to modulating threshold voltages via interface dipoles and dielectric fixed charges. This section will review the factors that impact effective work function and also discuss the various methods that are currently being utilized to achieve this.

Accurate extraction of effective work function and dielectric fixed charge is the subject of multiple papers [6, 7]. For conventional SiO_2 gate dielectrics, the effective work function can be extracted from the y-intercept of plots of flat band voltage (V_{fb}) versus effective oxide thickness (EOT). The basic equation governing the electrostatics in a MOS capacitor is:

$$V_{fb} = \phi_{m,eff} - \phi_s - \frac{\int_0^{EOT} x\rho dx}{\varepsilon_{SiO_2}}$$

for a typical SiO_2-based MOSFET with negligible contributions from the bulk charges this is equivalent to:

$$V_{fb} = \phi_{m,eff} - \phi_s - \frac{Q_2 EOT}{\varepsilon_{SiO_2}} \tag{5.1}$$

where $\phi_{m,eff}$ = the effective metal work function, ϕ_s is the Fermi level of the silicon, $\rho(x)$ is the bulk charge distribution in the dielectric, Q_2 is the fixed charge at the Si/dielectric interface, and ε_{siO_2} is the permittivity of SiO_2.

For HfO_2 films deposited on silicon a ~ 1 nm SiO_2-like interface layer typically exists between the silicon substrate and the HfO_2. Therefore, (5.1) does not adequately account for the contributions of the fixed charges at the SiO_2/HfO_2 interface, nor does it capture the fact that HfO_2 can have a rather significant bulk fixed charge value, something that is often neglected for SiO_2 dielectrics. A more rigorous equation for the extraction of fixed charges in a high-k gate stack is given by

$$V_{fb} = \phi_{m,eff} - \phi_s - \frac{Q_1 EOT_1}{\varepsilon_{SiO_2}} - \frac{Q_2 EOT}{\varepsilon_{SiO_2}} - {}^1\!/_2 \frac{\varepsilon_{HfO_2}}{\varepsilon_{SiO_2}} \frac{\rho_i (EOT_1)^2}{\varepsilon_{SiO_2}} \tag{5.2}$$

where Q_1 is the fixed charge at the SiO_2/HfO_2 interface, EOT_1 is the effective oxide thickness of the HfO_2 layer, EOT_2 is the effective oxide thickness of the SiO_2 layer, ρ_1 is the HfO_2 bulk charge density, ε_{HfO_2} is the permittivity of HfO_2, and $EOT = EOT_1 + EOT_2$ (Fig. 5.5).

This equation now illustrates some of the various layers in the gate stack that can be manipulated in order to shift threshold voltages to the desired device requirement. The next sub-sections will provide more detail into ways the flat-band voltage, or threshold voltage, can be controlled via the effective work function of the gate material or the fixed charge at various layers in the dielectric gate stack.

5.2.2.1 Vacuum Work Function

The discussion of effective work function begins with a quick review of the work functions of the elements [8]. Vacuum work functions, $\phi_{m,vac}$, show a periodicity between work function and atomic number. Work functions increase from left to right across a row on the periodic table (Fig. 5.6). This indicates that metals with work functions above the conduction band edge of Si (<4.1 eV) are typically in the first three columns of the periodic table. Alternatively, metals with work functions below the valence band edge of Si (>5.2 eV) tend to be late transition

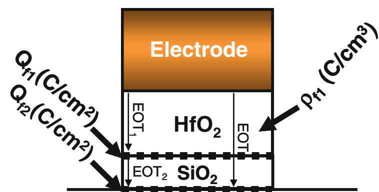


Fig. 5.5 Image showing interface fixed charges at the Si/SiO_2 and SiO_2/HfO_2 interfaces and bulk fixed charges inside the HfO_2 gate dielectric

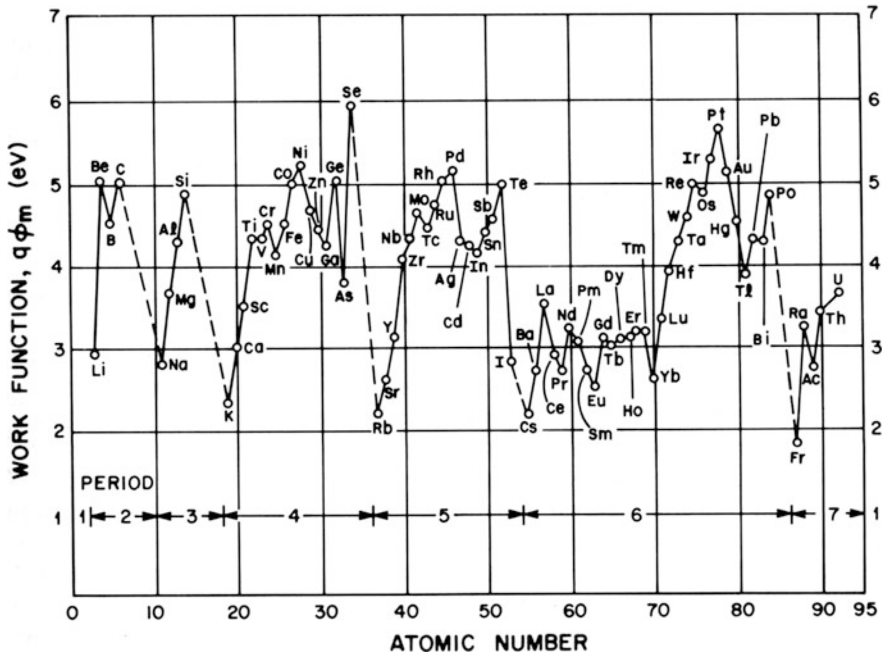


Fig. 5.6 Plot of work function versus atomic number. Work functions are generally observed to increase moving across a row of transition metal elements [20]

metals. In particular, the platinum group metals such as Pt, Ir, Os, Au, Ni, Ru, Pd, and Rh have high work functions. However, the vast majority of the transition metals have work functions that exist within the band edges of silicon (4.1 eV < ϕ_m < 5.2 eV) making it difficult to find metals that can easily replace p+ and n+ poly-silicon gates.

The limited number of metal gate electrode candidates with vacuum work functions between the conduction band edges of silicon is restricted even further by Fermi level pinning. Fermi level pinning is a consequence of forming an interface between a metal and a dielectric (or semiconductor). When an interface is formed, the effective metal work function becomes "pinned" at a different energy than its vacuum work function. This results from interfacial charge exchange between the metal Fermi level and gap states at the metal-dielectric interface causing it to shift with respect to its unpinned location. A recent comprehensive review of Schottky barrier concepts has been published [9] and Fermi level pinning models have been extensively tested on a broad class of interfaces [10–13], including for metal gate electrodes on HfO_2 [14, 15].

The most widely accepted of the Fermi level pinning models is the metal induced gap states (MIGS) model [10]. The origin of the gap states in the MIGS model is from the dangling bonds of under-coordinated surface atoms. These dangling bonds produce surface states that are dispersed in continuum at energies

throughout the band gap of the dielectric. The electron wave function for surface states is given by $\psi = u(r)exp(ikr)exp(-ik\perp z)$[16] (k is the wave vector, r, z are position vectors, and u is a periodic function in r). The two components of this equation are related to electron propagation in the plane of the surface and electron propagation normal to the surface. For electrons traveling parallel to the surface the wave vector $k_{\|}$ is real, and $\psi = u(r)exp\left(ik_{\|}r\right)$. However, for electron propagation normal to the surface $k_{\perp}$ is complex. This is because the $E - V(z)$ term under the square root in the expression $k_{\perp} = \left(2m\left[E - V(z)/\hbar^2\right)^{1/2}$ is negative. This term is negative due to the potential energy $V(z)$ being greater than the electron energy E when an electron tunnels outside the crystal or when an electron tunnels into the crystal bulk. Therefore, the $exp(-ik\perp z)$ term results in an exponential decay of the electron wave function for directions normal to the surface. These are often referred to as evanescent states. Similar to a surface state on a dielectric, a cleaved metal surface also has a surface wave function that decays exponentially into vacuum. In the MIGS model, when a metal is placed in contact with a dielectric, the metal surface states induce the gap states in the dielectric. This results in interfacial charge exchange between the metal and the dielectric gap states causing the metal Fermi level to shift with respect to its unpinned location towards a characteristic energy level in the semiconductor. In the MIGS model this characteristic energy is referred to as the charge neutrality level ($\phi_{CNL,d}$). The charge neutrality level is the location of the highest occupied surface state in the dielectric band gap. The charge neutrality level has been described as the position where the surface states change from acceptor-like to donor-like character. Therefore, when the two surfaces are brought into contact, charge is exchanged between the metal and the dielectric surface states resulting in the formation of an interfacial dipole. The magnitude of the interfacial dipole depends on the pinning strength of the semiconductor which is defined by the value of the Schottky pinning parameter (S). The barrier height between a metal Fermi level and the dielectric conduction band depends on the pinning parameter and is given by $\Phi_b = S\left(\phi_{m,vac} - \Phi_{CNL,d}\right) + \left(\Phi_{CNL,d} - \chi_d\right)$ [17] (χ_d is the electron affinity of the dielectric). In the Schottky, or weak pinning, limit $S = 1$, while in the Bardeen, or strong pinning, limit $S = 0$ (Fig. 5.7). The magnitude of Fermi level pinning in a MOS capacitor can also be measured in terms of an effective metal work function $\left(\phi_{m,eff}\right)$, as opposed to barrier height shifts. The effective work function is related to the vacuum work function $\left(\phi_{m,vac}\right)$ by $\phi_{m,eff} = \phi_{CNL,d} + S\left(\phi_{m,vac} - \phi_{CNL,d}\right)$ [14]. It should be noted that the gap states in the MIGS model are intrinsic to any metal-dielectric interface.

The pinning parameter (S) is the slope obtained on plots of barrier height (ϕ_b) versus the vacuum work function $(\phi_{m,vac})$ for a given dielectric or semiconductor, $S = d\phi_b/d\phi_{m,vac}$. This relationship was developed for Schottky contacts, but in MOS capacitor structures effective work function $(\phi_{m,eff})$ should be substituted for ϕ_b since the metal Fermi level is being referenced with respect to the silicon Fermi level instead of the semiconductor (or dielectric) conduction band as in studies of

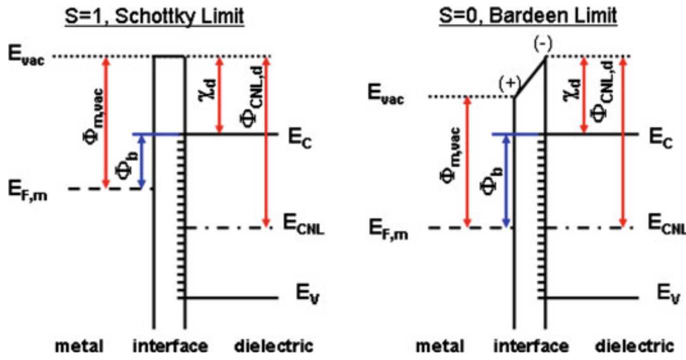


Fig. 5.7 Schematic of Fermi level pinning in the Schottky or weak-pinning limit and in the Bardeen or strong-pinning limit

Schottky contacts. The pinning parameter $S{=}d\phi_{m,eff}/d\phi_{m,vac}$ is identical to $S{=}d\phi_b/d\phi_{m,vac}$ if it is assumed that any internal dipoles that might exists at the SiO_2/HfO_2 or the Si/SiO_2 interfaces as well as any internal fields in the dielectric are independent of the electrode work function. Plots of $\phi_{m,eff}$ and ϕ_b versus $\phi_{m,vac}$ will therefore only differ by a parallel shift of the data, but the slope, or pinning parameter, remains unchanged.

An alternative approach is to express the slope parameter in terms of metal electronegativity (X) instead of vacuum work function, where $S' = d\phi_{m,eff}/dX$ [18, 19]. This variation on the pinning parameter arises because the Pauling electronegativity has been correlated to the vacuum work function of elemental metals by the expression $\phi_{m,vac} = 2.27X + 0.34$ [20, 21]. When this relationship is applied to elemental metals, the Schottky limit occurs on plots of X versus $\phi_{m,eff}$ when $S' \sim 2.27S$ [22, 23].

The relationship between group electronegativity and barrier height or effective work function has been extended to multi-element metal gate electrodes in contact with HfO_2. A plot of effective work function versus the geometric mean of electronegativity for a diverse set of electrode materials on HfO_2 produces a reasonable linear fit (see Fig. 5.8). This relationship provides a useful means of predicting effective work functions of either elemental or alloyed metals.

It should be noted that the sub-lattice elements in certain metal gate compounds can be a large factor in the calculation of the mean electronegativity. A rather sizeable electronegativity difference of $\Delta X_{\text{B to O}} = 1.5$ exists across four common sub-lattice elements (B, C, N, and O) that appear sequentially in the periodic table. For comparison, the electronegativity range across 30 transition metals is only $\Delta X_{\text{La to Au}} = 1.44$. Because of this fact, it is not surprising that low effective work functions have been obtained with LaB_6 and TaC, that interstitial carbide materials have consistently lower work functions than their analogous interstitial nitrides, and that conductive oxides, such as iridium oxide (IrO_2) and ruthenium dioxide

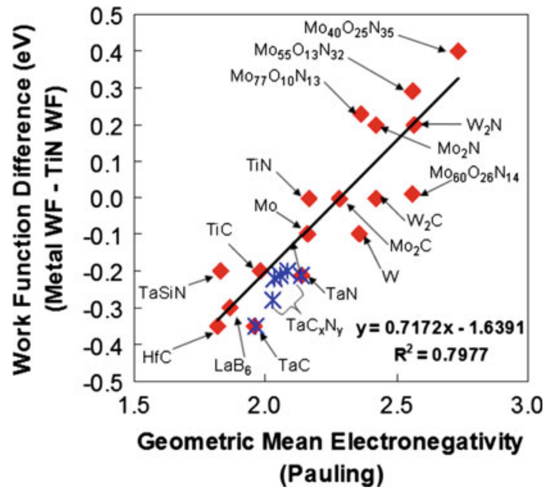


Fig. 5.8 Plot of the mean electronegativity of the constituent elements in metal gate electrodes versus the effective work function relative to TiN on HfO_2 gate dielectrics. All films were exposed to >900 °C anneal temperatures [65]

(RuO_2) have higher effective work functions than Ir or Ru. It should be noted that the correlation is not perfect, indicating that factors other than the geometric mean electronegativity likely contribute to the effective work function of metals on HfO_2. Some of these factors include the fact that the volume fraction of elements in the electrode may not be representative of the actual interface bonding, and differences in crystal phase and preferred orientation may contribute to some of the variability as well.

5.2.2.2 Manipulation of Fixed Charges/Interface Dipoles

Dielectric Capping Layers

Some promising techniques for engineering the metal work function towards band-edge involve the use of dielectric capping layers such as MgO or La_2O_3 for nFETs [24] and Al_2O_3 for pFETs [25], or by incorporating the La, Mg or Al into the metal electrode, instead of as an oxide capping layer. Both approaches work well for modulating the effective work function or threshold voltages (V_t) of MOS devices, but alloying the metal directly into the gate electrode can be advantageous in terms of effective oxide thickness scaling.

Capping Layers for NFET Electrodes

There has been a significant amount of work discussing the use of column IIA, IIIB, and rare earth elements for "tuning" the nFET threshold voltages [24, 26, 27]. The typical approach is to deposit oxides of these materials 1–10 Å thick between the high-k gate dielectric and the metal gate electrode in a gate first integration scheme. These capping layers produce negative V_t shifts, and reduce the threshold voltage of nFET device by up to 400 mV (Fig. 5.9). In some instances dielectric capping layers can be used for V_t shifts with slight reduction or

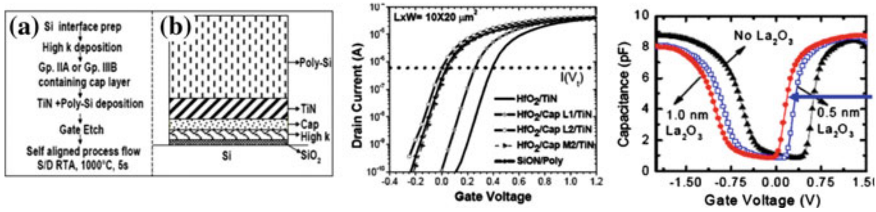


Fig. 5.9 Dielectric capping layer approach showing well behaved $I_d - V_g$ [24] and $C_g - V_g$ [75] curves

minimal increases in the effective oxide thickness of the gate dielectric. However, increasing the dielectric cap thickness too much will eventually result in increases in the effective oxide thickness. Well behaved $I_d - V_g$ (drain current–gate voltage) and $C_g - V_g$ (gate capacitance–gate voltage) curves showing monotonic shifts in the threshold voltage and no degradation in the sub-threshold slope are observed with this approach (Fig. 5.9).

To avoid the effective oxide thickness penalty associated with using an oxide capping layer, the V_t shift element can be incorporated directly into the electrode instead. One example of this is an experiment where Mg was incorporated directly into a TaMgC inter-layer between the gate dielectric and a TaC capping layer. To simultaneously decrease the NMOS threshold voltage and the effective oxide thickness, the thickness of the TaMgC inter-layer can be gradually increased or the Mg concentration in the TaMgC inter-layer can be gradually increased (Fig 5.10). Both approaches show V_{fb} shifts comparable to those achieved with MgO capping layers [28].

As will be explained in more detail later, the V_t shift mechanism for both the oxide capping and the alloyed electrode approach depends on the controlled diffusion of the capping elements to the SiO_2/HfO_2 interface where they react with SiO_2 at this interface to produce an interface dipole or positive fixed charge which results in the favorable reduction of the nFET threshold voltage. One potential negative side-effect of achieving threshold voltage reductions in this manner is that the creation of dipoles or fixed charges at this interface can result in the degradation of the mobility, sub-threshold slope, or gate leakage if the dielectric capping layer is too thick [24]. For device performance optimization, careful consideration needs to be given to the trade-offs between mobility, inversion capacitance equivalent thickness T_{inv}, and short channel control [29].

Capping Layers for PFET Electrodes

For pFETs, additional V_t shifts to help satisfy the requirements of high performance devices can be achieved by using Al_2O_3 capping layers or by alloying Al into the metal gate electrode. The impact of Al_2O_3 capping layers has been well documented [30–33] and shown to produce positive V_t shifts of up to ~200 mV that are desirable for reducing the pFET threshold voltage (Fig. 5.11). The PMOS work function tuning situation is analogous to the NMOS one in that V_t lowering

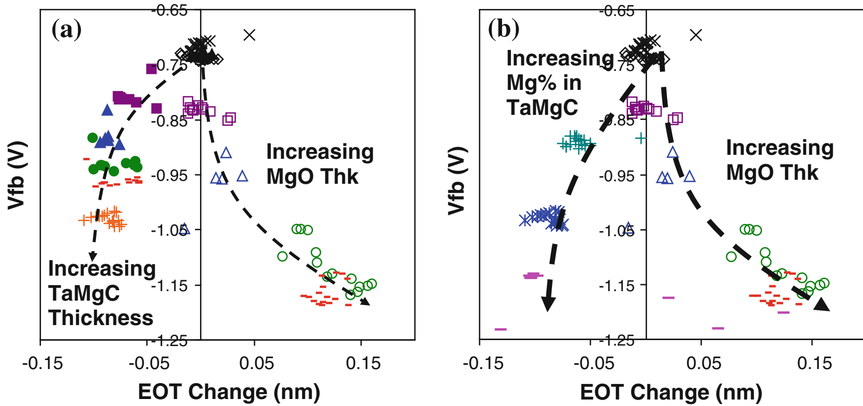
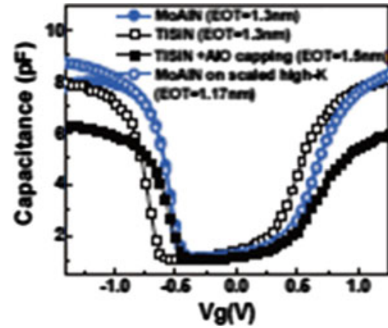


Fig. 5.10 V_{fb}/EOT shifts for increasing TaMgC thickness (**a**) and for increasing at percentage Mg in TaMgC (**b**) with V_{fb}/EOT for increasing MgO cap layer thickness as a reference

Fig. 5.11 Capacitance–Gate Voltage curves showing *positive* V_t shifts compared to reference devices by using MoAlN. MoAlN does not have any T_{inv} increase which is an added benefit over AlO capping layers [34]



can be achieved by either using Al_2O_3 capping or by alloying Al into a metal gate electrode.

For the alloyed metal electrode approach, MoAlN has been regularly evaluated [28]. Similar to the results achieved with nFET dopant species, comparisons of the Mo_2N/Al_2O_3 and MoAlN approaches indicate there is less effective oxide thickness increase with the alloyed electrode approach [34]. Unfortunately, increasing the total dose of Al in the MoAlN layer, either by increasing the MoAlN layer thickness or by increasing the Al concentration can result in a degradation of the device mobility. Tuning the threshold voltage of MOSFET devices with a MoAlN electrode requires careful engineering control to avoid some of the device degradations that occur when the Al dose is too large.

Mechanism for Threshold Voltage Shifts

Although there remains some debate as to whether the mechanism for the V_t, EOT, mobility, sub-threshold slope, and leakage changes is caused by internal dipoles, fixed charges, or modulation of dielectric oxygen vacancy concentrations the experimental results do suggest that the V_t shift elements must be located at the

SiO_2/high-k interface in order to produce the desired threshold voltage shifts. Some of the most compelling data elucidating this phenomenon was from the work that placed an Al_2O_3 inter-layer either at the SiO_2/HfO_2 interface or at the HfO_2/electrode interface [25]. This work indicated that large V_t shifts are achieved even with no thermal budget applied to the gate stack when the Al_2O_3 layer is placed at the SiO_2/high-k interface. On the other hand, when the Al_2O_3 is placed at the high-k/electrode interface, high temperature annealing is required in order to diffuse the Al towards the SiO_2/high-k interface where it produces the desired threshold voltage shift. This is a very critical result that suggests that without a significant thermal budget applied to the gate stack, the capping oxides must be placed at the SiO_2/HfO_2 interface in order to produce the desired V_t shifts. However, V_t shifts can also be produced if a thermal budget (>900 °C) is applied to “drive-in” the dopant species through the high-k to the bottom interface.

This has been extensively studied with various depth profiling techniques such as EELS (Electron Energy Loss Spectroscopy) and SIMS (Secondary Ion Mass Spectroscopy) to verify that the V_t shift elements are diffusing down to the SiO_2/high-k interface after annealing. In one experiment on MoAlN metal gate electrodes, back-side SIMS confirmed that there is indeed an inward diffusion of the Al from the electrode to the SiO_2-based interface layer with thermal budget and that this coincides with increases in the effective work function of the MoAlN metal electrode with increasing anneal temperature.

SIMS has also been performed on gate stacks with MgO caps or TaMgC electrodes before and after annealing. This analysis clearly shows that after a high temperature anneal there is a pile-up, or segregation, of Mg into the SiO_2-like interface layer. This is consistent with the fact that many of these materials are strong silicate forming elements [35]. It should be noted that the dopant species diffuse to the HfO_2/SiO_2 interface regardless of whether it is incorporated in the metal gate electrode or applied as an oxide dielectric cap layer.

There are multiple mechanisms that have been proposed to explain the observed V_t shifts. The first possible mechanism is that the effective work function of the electrode is actually changed by the electropositive (nFET) or electronegative (pFET) elements used as capping layers. This result is consistent with the effective work function and electronegativity correlation. However, it is not consistent with the electrical results in Fig. 5.12 that suggest the dopant species needs to be located at the SiO_2/HfO_2 interface to produce the desired V_t shifts. It is also possible that as the elements diffuse through the gate dielectric the HfO_2 defect chemistry and equilibrium oxygen vacancy concentration is being modulated [36]. In this argument one would expect Al to produce similar V_t shifts to La since they both have a valence of +3 and would favor the creation of positively charged oxygen vacancies in HfO_2. Since the opposite is observed this is also not the likely V_t shift mechanism. Finally, as the elements diffuse down further toward the SiO_2/high-k interface, they can react to form a silicate and create fixed charges or a dipole layer at the SiO_2/high-k interface layer. There is much supporting evidence for this model. First, electrical results suggest that for capping layers a high thermal budget needs to be applied to produce V_t shifts that are similar to those

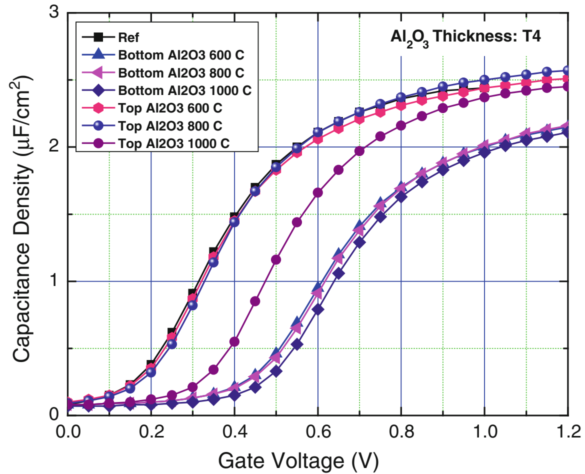


Fig. 5.12 Capacitance–Gate Voltage curves showing V_t shifts for *top*- and *bottom*-Al_2O_3 interface layers with different anneal temperatures. V_t shifts are achieved when the Al_2O_3 is at the bottom interface, or when the thermal budget diffuses the Al_2O_3 from the *top* to the *bottom* interface [25]

achieved at low thermal budgets when the cap layer is placed directly at the SiO_2/HfO_2 interface. SIMS and EELS depth profiling indicates that the capping elements diffuse down to the HfO_2/SiO_2 interface with anneal. The pile-up of dopant species at the SiO_2/HfO_2 interface is consistent with the fact that many of these dopant species are known to be strong silicate forming elements [35], and it has been theoretically demonstrated that these elements energetically prefer to segregate to the HfO_2/SiO_2 interface or the SiO_2 intermediate layer [37, 38]. Theoretical studies have evaluated a set of dopant species with different valences and work-functions/electro-negativities. Finally, the direction of the V_t shifts is consistent with the group electro-negativities of the doping oxide relative to HfO_2. The majority of evidence so far supports a model that the V_t shifts are caused by interface dipoles at the SiO_2/HfO_2 interface (Table 5.1, Figs. 5.13, 5.14).

Mechanism for Effective Oxide Thickness Reduction

The mechanism for EOT scaling that is achieved when the La, Mg or Al is incorporated into the gate electrode instead of as a dielectric capping layer is related to the thermodynamic competition for oxygen between the layers in the

Table 5.1 Comparison of the mean electro-negativities of some dopant oxides commonly employed to shift the threshold voltages in Hf-based gate dielectrics

Capping dielectrics	Geometric mean electronegativity (Pauling)
MgO	2.11
La_2O_3	2.18
Gd_2O_3	2.26
Y_2O_3	2.27
HfO_2	2.49
Al_2O_3	2.54
TiO_2	2.63
Ta_2O_5	2.71
SiO_2	2.82

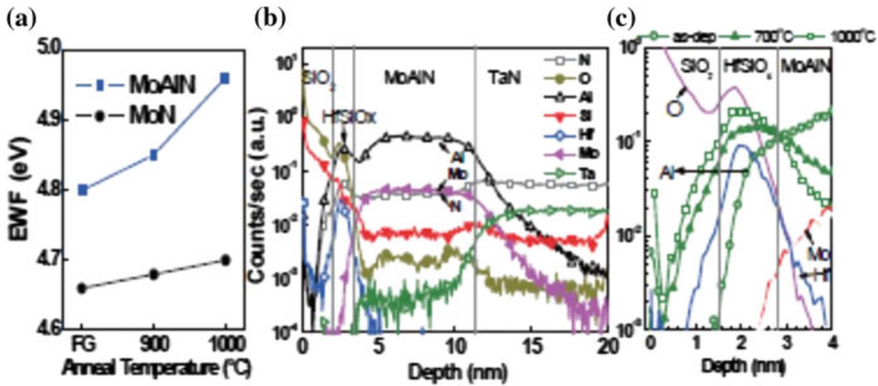


Fig. 5.13 Effective work function shift and aluminum diffusion into $HfSiO_x$ gate dielectric as a function of applied thermal budget. Backside SIMs profile of the10-nm MoAlN/2 nm $HfSiO_x$/2 nm terraced oxide stack showing **a** MAIN EWF as a function of annealing temperature **b** uniform MoAlN film even at 1,000 °C **c** local diffusion of Al through $HfSiO_x$ to the SiO_2 interface after annealing (a possible cause for EWF modulation) [88]

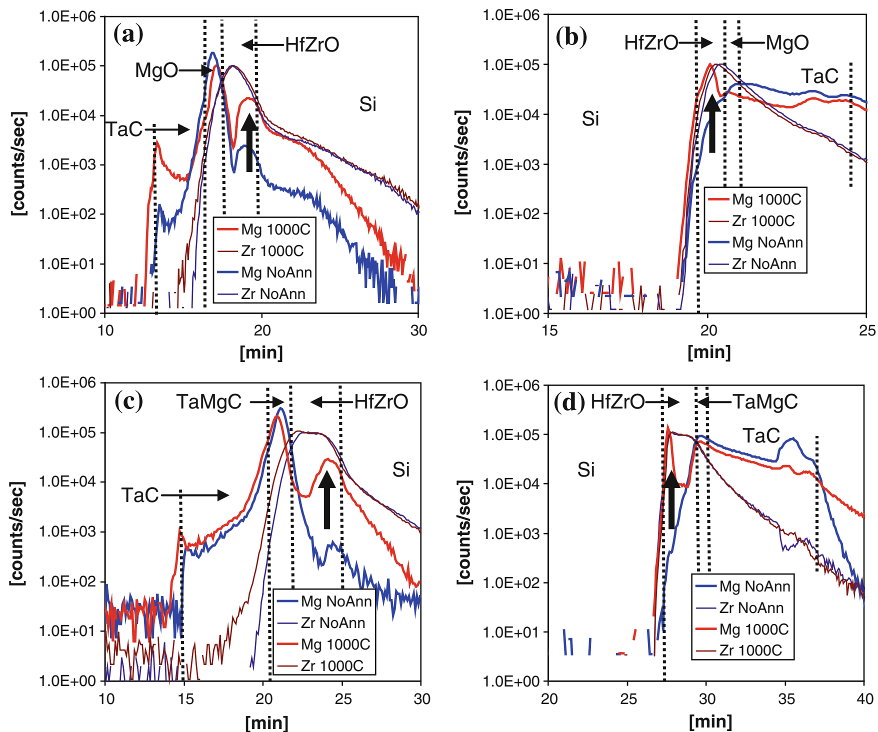
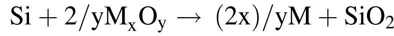


Fig. 5.14 *Front*- (**a**) and *back*- (**b**) side SIMS with MgO capping. *Front*- (**c**) and *back*- (**d**) side SIMS for TaMgC. The dielectric is $Hf_xZr_{1-x}O_2$ and only the Zr and Mg profiles are shown for clarity [28]

gate stack. This is summarized by the following reaction which illustrates how SiO_2 is converted into Si and a metal oxide with a higher dielectric constant if there is a large positive Gibbs free energy change at 1,000 K ($\Delta G°_{1000K}$) [39].



$$Si + 2/yM_xO_y \rightarrow (2x)/yM + SiO_2$$

While using metallic capping layers to scavenge oxygen away from the bottom interface for EOT scaling, a commensurate degradation of mobility is also observed. A careful study comparing the use of La capping layers versus a remote interface scavenging technique (interface scavenging achieved without diffusing dopants into the high-k that would cause V_t shifts) suggests that the diffusion of La to the bottom interface does not result in additional mobility degradation beyond what is expected due to the thinning of the bottom SiO_2 interface layer. However, one should use caution because an extrinsic mobility degradation and device scatter can be observed depending on whether the dopant species or even with oxide cap layers if the cap layer is too thick. It should be noted that excellent process control is required for the deposition of these thin capping layers since the total dose of these layers is closely related to important device parameters such as mobility, EOT, and V_t.

Oxygen Vacancies

The threshold voltage shift mechanisms so far have focused on effective work function and charges at the SiO_2/high-k interface, but many experiments have shown that the bulk charge term in (5.2): $\Delta V_{fb} = -1/2\left(\varepsilon_{HfO_2}/\varepsilon_{SiO_2}\right)\rho_1 EOT_1^2/\varepsilon_{SiO_2}$ can be manipulated by controlling the oxygen vacancy concentration in the high-k dielectric. By annealing high-k devices in oxygen at an appropriate temperature and pressure positive shifts of the threshold voltage of up to 450 mV can be achieved by annihilating positively charged oxygen vacancies that exist in the high-k material.

Most of the studies on this topic focused on the top-down diffusion of oxygen through a metal gate electrode such as Pt [40], Re [41, 42], and TiN [43] as well as the lateral diffusion of oxygen from the sides of the gate [44]. The key results indicate that there are three key temperature ranges. At low temperatures no threshold voltage shifts are observed when the samples are annealed in oxygen. This is a diffusion limited temperature range where oxygen can not penetrate the metal gate electrode in the top–down experiments or diffusion through the high-k is sufficiently low such that only the shortest gate lengths show threshold voltage shifts in the lateral diffusion experiments. Above the critical temperature, a thermodynamic equilibrium exists between the oxygen partial pressure and the oxygen vacancy concentration in the HfO_2 film. As the oxygen annihilates positively charged oxygen vacancies large positive shifts of the threshold voltage can be produced. As the temperature continues to increase the maximum threshold voltage shift produced by the oxygen anneal reduces with increasing temperature.

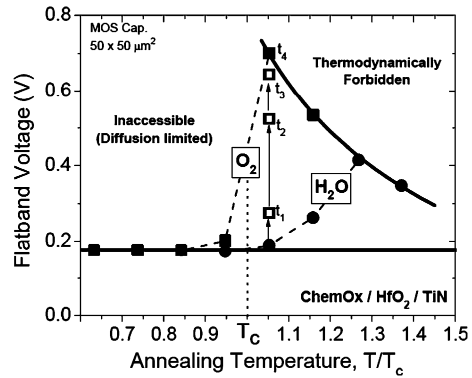


Fig. 5.15 Flatband voltage shifts in HfO_2 gate dielectrics as a function of oxygen annealing temperature [43]

This is because the equilibrium oxygen vacancy of the high-k film increases as temperature increases and because the equilibrium oxygen vacancy of the high-k film is balanced by the reaction of silicon and oxygen to form SiO_2 at the silicon substrate. When this occurs, an effective oxide thickness increase or gate leakage decrease accompanies the positive threshold voltage shifts due to the growth of additional SiO_2 at the silicon substrate (Fig. 5.15).

The studies of lateral oxidation are consistent with the top-down oxidation studies [44]. In general, at low temperatures only the devices with the shortest gate lengths begin to show threshold voltage shifts. With either increasing oxygen anneal time or increasing oxygen temperature, the threshold voltage shifts of the short devices become larger and the longer devices begin to show threshold voltage shifts also. This result indicates that a concentration gradient exists for the oxygen vacancy concentration with the highest concentration of vacancies in the center of the gate and the lowest concentration at the edge of the gate.

The positive threshold voltage shifts caused by the oxygenation of a high-k gate stack can be reversed by annealing in vacuum [41, 44], reducing ambients [40], or as previously discussed, at elevated temperatures where there is a higher equilibrium oxygen vacancy in the high-k material. This is the fundamental reason why it is difficult to achieve a low threshold voltage pFET device in a gate first integration scheme. This has been discussed in the context of a threshold voltage “roll-off” where low pFET threshold voltage devices can be achieved if the gate stack is never exposed to elevated temperatures, but at elevated temperatures there is an increase in the pFET threshold voltage when the interfacial layer separating the Si substrate from the high-k film is thin (Fig 5.16). It should be noted that the proximity of the high-k to the substrate appears to facilitate the formation of oxygen vacancies in the high-k film at elevated temperatures. When the high-k is near the substrate, oxygen exchange between the high-k and the silicon substrate readily occurs and the oxygen chemical potential is shifted. This results in a dipole offset in the gate stack that increases the pFET threshold voltage after high temperature processing [45]. The impact of this high temperature processing on the

pFET threshold voltage is an important consideration for integrating high-k materials into a CMOS-compatible process flow.

5.3 CMOS Metal Gate Integrations

This section deals with the integration of metal gate electrodes with different work functions to build both NMOS and PMOS devices. CMOS integration is easily done with poly-silicon electrodes by ion implanting n-type or p-type dopants into NMOS or PMOS electrodes to get appropriate work functions for both device polarities. For metal gate electrodes the integration scheme is more complex. Both conventional gate-first integrations [46, 47] and replacement gate [48–51] integrations have been proposed. Gate-first integrations have the advantage of being structurally similar to a poly/SiO_2-based CMOS process. In replacement gate integrations, the highest processing temperatures are performed prior to depositing the gate materials. Replacement gate integrations can therefore avoid the increase in the pV_t due to the generation of positively charged oxygen vacancies in the high-k.

5.3.1 Materials Considerations for Metal Gate Electrodes

When used in a gate-first integration scheme, metal gate electrode candidates must possess thermal stability up to the dopant activation temperatures of approximately 1,000 °C. Thermal stability includes the absence of gross reactions between the dielectric and other surrounding materials, no inter-diffusions with surrounding materials, sufficient bulk phase stability, and smooth interfaces with the dielectric [52]. Due to these stringent thermal requirements refractory metals such as W, Re, Ta, and Mo have attracted interest.

Metal nitrides and carbides have also gathered attention as metal gate electrodes. These materials are renowned for their use as diffusion barriers in the semiconductor industry. Transition metal nitrides and carbides are comprised of a

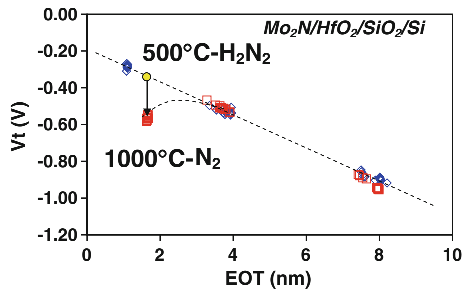


Fig. 5.16 Threshold voltage versus EOT for a constant thickness of HfO_2 on varying thicknesses of underlying SiO_2. Note that with high temperature processing the threshold voltage only shifts when the interface is thin

face-centered-cubic metal structure with nitrogen or carbon atoms occupying the octahedral interstices. This results in the rock-salt structure (NaCl). Since N and C occupy interstices the lattice parameter is typically only $\sim 5\ \%$ larger than that of the pure metal compound. Stuffing the interstices with N or C helps give these materials their excellent diffusion barrier properties. These materials exhibit metal-like conduction, but at the same time are highly refractory compared to the pure metal constituent. It is believed that a mixture of metallic, covalent, and ionic bonding character is responsible for this behavior [53]. There have been numerous studies on the work function of interstitial nitrides for electrode and emitter applications. This includes MoN [54], WN [52], NbN [55], TaN [52, 56], TiN [52, 57–59], ZrN [55], and HfN [60], among others. TiAlN has also been widely studied as a metal gate electrode [61]. Although it is a ternary metal nitride, it is more closely related to the binary nitrides than the amorphous ternary metals because this compound exists in the rock salt structure for Al:Ti ratios <0.40 and in a mixture of rock salt and wurtzite structures for >0.40 Al:Ti ratio [62]. TiAlN loses its conductive properties with increasing presence of the AlN wutzite phase. Metal carbide materials have also shown promise as thermally stable metal gate electrodes. [27, 63, 64]. Solid-solutions of metal carbides and nitrides have been compared and it has been shown that carbide materials have lower effective work functions than their analogous nitride making them good nFET electrode candidates [65, 66].

Ternary alloys of a transition metal (TM), silicon, and nitrogen have also been widely investigated for their properties as diffusion barriers. These films are found to exist in a highly metastable amorphous structure [67]. This property makes them excellent diffusion barriers to elevated temperatures. Ta–Si–N has been studied as a stable NMOS electrode candidate [52, 68–70].

Even if a thermally stable metal gate electrode is selected there are more subtle effects detected with electrical measurements that are related to the thermo-chemistry of the gate stack. As previously discussed in the section on alloyed metal gate electrodes for work function tuning, if a metal/high-k/SiO_2/Silicon stack is annealed at high temperatures there is a thermo-chemical competition for oxygen in the gate stack that takes place between the silicon substrate and the metal gate material where oxygen transport across the HfO_2 layer is too rapid to limit the reactions [71]. Thermo-chemical calculations of a simple TiN_x/HfO_2/SiO_2/Si system were performed to illustrate this effect. If the composition is such that $x < 1$, then the oxygen in the SiO_2 interface layer can be dissolved into metal gate electrode and the Si returns to the substrate. However, if $x > 1$ then the calculations indicate that the system drives towards the creation of extra interfacial SiO_2. As was observed in the examples of alloyed metal gate electrodes these thermo-chemical interactions can be manipulated to achieve a desired electrical response (Fig. 5.17).

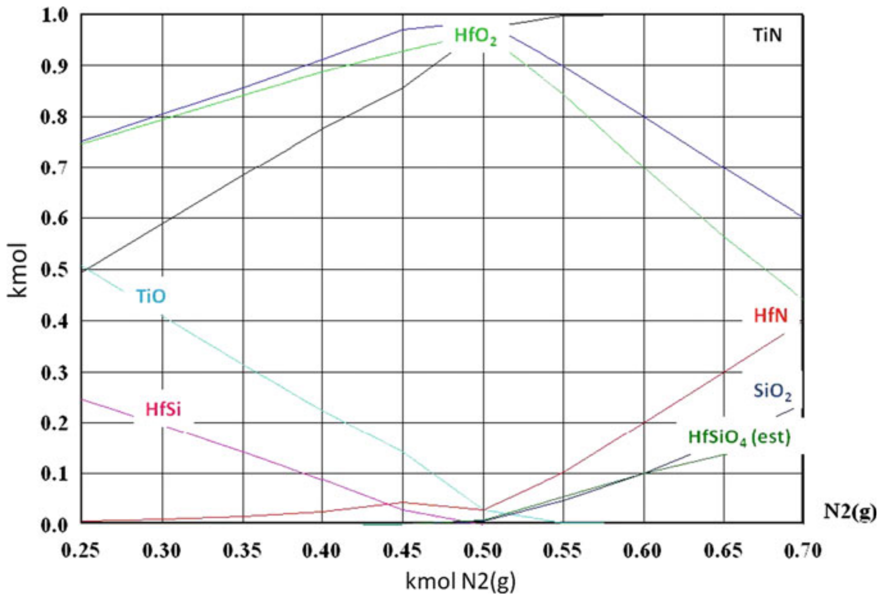


Fig. 5.17 Illustrates the fraction of each species present (Si, SiO_2, HfO_2, HfSi, TiN, TiO, HfN, and $HfSiO_4$) by varying the composition of a TiN_x metal gate electrode. For N-deficient TiN, there is a tendency to form TiO and HfSi. For N-rich TiN there is a tendency to form HfN and extra SiO_2 interface layer

5.3.2 Gate First Integration

A gate first integration is defined as one in which the high-k material and the metal gate electrodes are deposited prior to implanting the extension, halo, and source and drain dopants into the transistor. The gate first integration is desirable because it is the most similar to the conventional poly/SiON transistor integration and the least disruptive to the other modules that need to be integrated with the gate stack. Typically, in a gate first integration, thin metal gate layers are capped with poly-Si allowing for gate silicidation to be performed in the same fashion as it was in a poly/SiON integration. In a gate-first integration, the high-k metal gate materials must withstand significant process temperatures, typically in excess of 1,000 °C, which are required to activate the implanted dopants in the silicon substrate. Exposing the gate stack to such temperatures can present some challenges. It limits the number of gate stack metals to those that are thermally stable at these temperatures. In addition, as detailed in the section on the role of oxygen on the effective work function of the transistor, high processing temperature can result in higher than desired pFET threshold voltages.

5.3.2.1 Dual Metal Gate Electrodes

There are two main approaches to dual-metal gate electrode integration. The first is to use two separate metal gate electrodes with different effective work functions. To get different work functions, two separate metals can be used for nFET and pFET electrodes or one of the metals can be alloyed in a manner that shifts its effective work function. A second approach is to combine the dual-metal gate electrodes with some of the other threshold voltage engineering techniques that manipulate fixed charges or dipoles inside the gate dielectric layer. In this situation, the final structure is one in which the gate dielectric and the metal gates are different between nFET and pFET electrodes.

Dual-Metal Gate, Single-Gate Dielectric

The integration approach that accommodates vastly different metals for NMOS and PMOS electrodes is the stacked dual metal gate integration (Fig. 5.18). This integration involves depositing one metal over both the NMOS and the PMOS regions, then patterning one of the two regions and removing the electrode using a wet etch chemistry that is selective to the underlying gate dielectric. A second metal can then deposited over the first metal. The most difficult aspect of this integration is the plasma etching of two different gate stacks of different heights. This integration has been demonstrated on Si_3N_4 gate dielectric [72] and on HfO_2 gate dielectrics [73, 74].

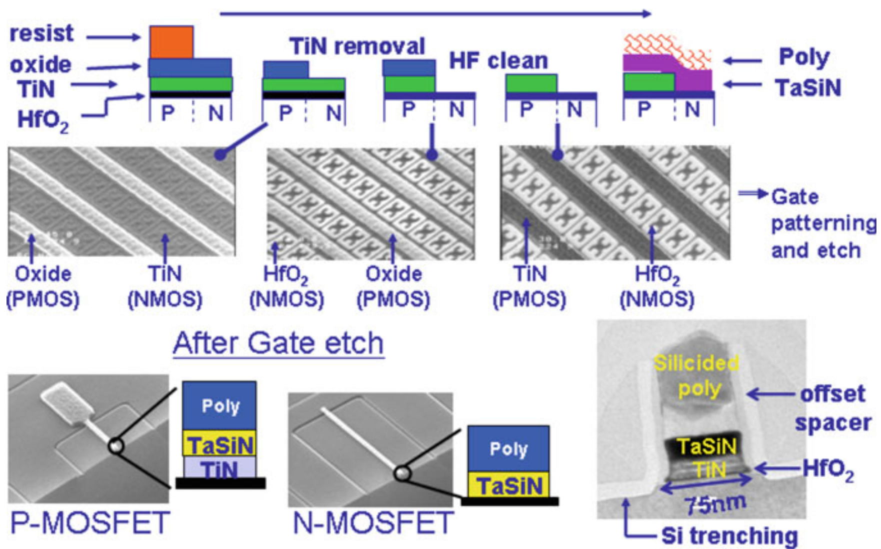


Fig. 5.18 Process flow for the stacked dual metal gate integration. The difficulty in this approach is that two gate stacks of different heights need to be etched simultaneously

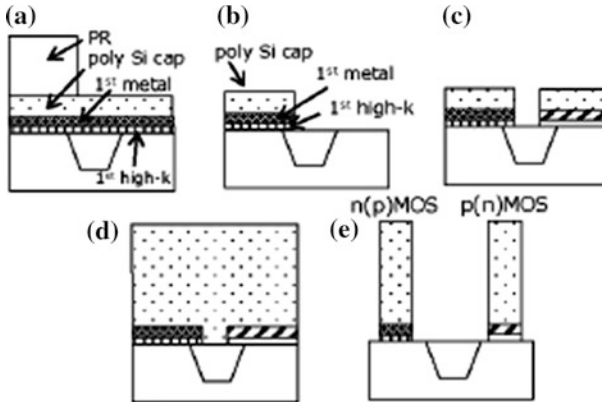


Fig. 5.19 Dual Metal Dual Dielectric (DMDD) CMOS integration scheme. [75]

The basic dual-metal gate integration scheme has been expanded to include additional work function tuning elements into the gate stack. One of these integration approaches uses two separate gate dielectrics, referred to as a dual-metal, dual-dielectric integration scheme. This allows for the use of different capping oxides on both the nMOS and pMOS devices as well as two separate work function metal gate electrodes to provide additional flexibility in threshold voltage tuning [75, 76] (Fig. 5.19).

An even more complicated approach employed dual-dielectrics, dual-metals, and dual-channels [77]. In this integration scheme the valance band-offset between SiGe and Si results in an additional reduction of the pFET threshold voltage.

Alloying for Dual-Metal Gate Work Functions

There have been numerous attempts to simplify the dual-metal gate integration to avoid having to selectively remove one metal from the gate dielectric without having any deleterious impact on the device. This can be achieved by alloying or tuning the work function of the metal in either the nMOS or pMOS region without actually depositing a second metal.

In the metal alloy approach, two metals are initially deposited on top of each other. Either the NMOS or PMOS region is then patterned so that the top metal can be selectively removed. A high temperature annealing process is then used to inter-diffuse the two metals. This leaves an inter-diffused metal for one of the electrodes and an elemental metal for the other electrode. This approach has been performed using Ru-Ta [78, 79], Ti-Ni [80], and Pt–Ta [81] alloys. The alloyed electrode approach has two foreseeable problems. First, it may be difficult to get one alloy system to span the entire 1.1 eV range to meet the NMOS and PMOS work function requirements. Second, some of the elements used in the alloy may not be thermally stable in contact with a high permittivity dielectric.

The nitrogen modulation integration has been reported in two forms. The first method is to ion implant nitrogen into a metal to shift its work function. This has been demonstrated for Mo-based electrodes [82, 83], and for TiN-based electrodes where the N concentration is modulated via implantation of N [84]. Adjusting the nitrogen concentration has also been accomplished by solid-state diffusion of nitrogen from a nitrogen-rich film into a nitrogen-deficient film [85]. Like the alloy approach, it may be difficult to span the entire work function range to meet the NMOS and PMOS electrode requirements by implanting nitrogen.

One material system that has been well suited for the nitrogen modulation technique is TaC. TaC has been shown to have a lower effective work function than its analogous nitride and it has been demonstrated that the work function of TaC can be controllably increased by alloying in nitrogen [63]. To exploit this property of the TaC_xN_y system in a metal gate integration, there has been some work to nitride TaC either by annealing or with plasma nitridation (Fig. 5.20).

5.3.2.2 Single Metal Gate, Dual-Dielectric

One of the primary manufacturing challenges for a dual-metal gate integration is the patterning and removal of one metal from the surface of the high-k film and the need to perform the simultaneous etch of two gate stacks with varying heights and/or materials. A simplification for manufacturing would be to integrate a gate stack that has a single metal gate electrode, but still achieves threshold voltages ideal for NMOS and PMOS devices by using separate dielectric capping layers for nMOS and pMOS. In this situation the challenge becomes the selective patterning of the dielectric capping layers. Depending on the integration approach, the dielectric capping layers can be placed at the SiO_2/high-k interface or on top of the high-k film and selective patterning with good resist adhesion is achieved by using a developable bottom anti-reflective coating (BARC) layer and wet chemistry to pattern the dielectric capping layers [86] (Fig. 5.21).

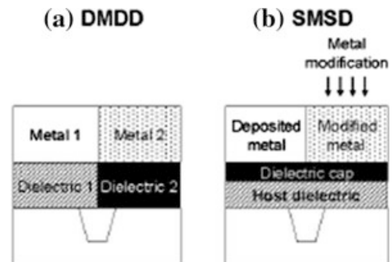


Fig. 5.20 Schematics of high-k/metal gate CMOS with deposited **a** dual metal-dual-dielectric (DMDD) and **b** single-metal-single-dielectric (SMSD) gate stacks [27]

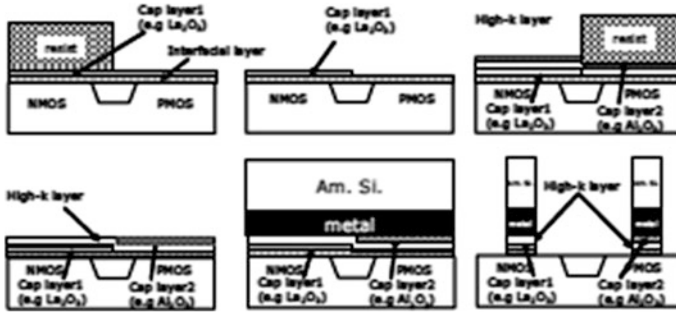


Fig. 5.21 Single-metal, dual-dielectric integration scheme (SMDD). CMOS process flow, with the nMOS cap located below and pMOS cap located above the bulk high-k dielectric [86]

5.3.3 Gate Last or Replacement Gate Integration

An alternative approach is the gate-last or replacement-gate integration. The main concept in the gate-last integration is that there is a "dummy gate" in place during the high temperature processing steps that are required to activate the dopants. This "dummy gate" can then be removed and replaced with the work function setting high-k/metal gate materials which is then never exposed to extreme subsequent processing temperatures. The replacement gate integration was the first high-k metal gate integration to be successfully implemented into manufacturing.

There are variations of the replacement metal gate integration scheme. One such integration that has been demonstrated is a hybrid gate-first/gate-last integration. In this integration the high-k layer, a thin portion of metal gate, and poly-Si capping layer are deposited. After encapsulation with spacers and implantation of the extensions/halos and source/drain, the poly-Si is eventually removed from both the nFET and pFET devices. At this stage the poly-Si voids are filled back in with a first work function setting metal. This will need to be patterned and removed in a manner that is similar to those used in the gate-first dual-metal gate integration. A second gate metal is then deposited for the opposite polarity transistor. The remainder of the gate is then filled back in with a low resistivity metal [48].

Another approach is a full replacement gate integration. In this approach, a poly/SiON gate is used for all the implantation and anneal steps. At a later stage of processing, the poly-Si is then removed and replaced with the high-k gate dielectric and the work function setting metal gate electrodes. In this scheme even the high-k film is not exposed to the source-drain activation temperatures [49].

There are a couple of benefits to the replacement gate integration that should be discussed. First, the replacement gate integration avoids many of the complications associated with the exposure of high-k/metal gate materials to elevated temperatures. For replacement gate integration, the gate stack materials are likely not exposed to temperatures greater than ~500 °C. In this temperature range, there

are more options for engineering a pFET transistor by modulating the oxygen vacancy concentration. Another widely touted benefit of the replacement gate integration is that the removal of the "dummy gate" can enhance the strain applied to the channel region from some of the other commonly used stress enhancement techniques like embedded SiGe source-drain regions and stress liners. The stress of the work function setting metal gates and the fill metal can also be tailored to directly couple to the channel strain [87].

5.4 Conclusions

In conclusion, metal gate electrodes are now being used in production by multiple companies. Some companies have opted for gate-last integration approaches and others have opted for gate-first integration approaches. In both cases, a significant number of materials, device, and integration challenges had to be overcome. As CMOS technologies continue to scale to the next technology nodes, metal gate electrodes will be used together with entirely new device architectures, channel materials, and gate dielectrics. Much of the learning from the initial development of metal gate electrodes will continue to apply in these situations, but there will also be many exciting new challenges and opportunities to build upon the years of research that were required to ultimately implement metal gate electrodes into advanced planar CMOS technology nodes.

References

1. T. Hori, *Gate Dielectrics and MOS ULSIs* (Springer, Berlin, 24 1997)
2. C. Hobbs, L. Fonseca, V. Dhandapani, S. Samavedam, B. Taylor, J. Grant, L. Dip, D. Triyoso, R. Hegde, D. Gilmer, R. Garcia, D. Roan, L. Lovejoy, R. Rai, L. Hebert, H. Tseng, B. White and P. Tobin. in *Symposium on VLSI Tech. Digest*, pp. 9–10, 2003
3. M. Gutowski, J.E. Jaffe, C.L. Liu, M. Stoker, R.I. Hegde, R.S. Rai, P.J. Tobin, Appl. Phys. Lett. **80**(11), 1897–1899 (2002)
4. I. De, D. Johri, A. Srivastava, C.M. Osburn, Solid State Electron. **44**, 1077 (2000)
5. B. Cheng, B. Maiti, S. Samavedam, J. Grant, B. Taylor, P. Tobin, J. Mogab, 2001 IEEE International SOI Conference, pp. 91–92, 2001
6. H.-C. Wen, et al, IEEE Elec. Dev. Lett. **27**(7), 598–601 (2006)
7. R. Jha, J. Gurganos, Y.H. Kim, R. Choi, J. Lee, V. Misra, IEEE Elec. Dev. Lett. **25**(6), 420 (2004)
8. S.M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981)
9. R.T. Tung, Mat. Sci. Engr. R, **35,** 1–138 (2001)
10. J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984)
11. W.E. Spicer, P.W. Chye, P.R. Skeath, C.Y. Su, I. Lindau, J. Vac. Sci. Technol. **16**, 1422 (1979)
12. H. Hasegawa, H. Ohno, J. Vac. Sci. Technol. B **4**, 1130 (1986)
13. J.L. Freeouf, J.M. Woodall, Appl. Phys. Lett. **39**, 727 (1981)
14. Y. C. Yeo, P. Ranade, T.-J. King, C. Hu, IEEE Elec. Dev. Lett. **23**(6), 342 (2002)

15. Y.C. Yeo, T.-J. King, C. Hu, J. Appl. Phys. **92**(12), 7266 (2002)
16. J.A. Venables, *Introduction to Surface and Thin Film Processes* (Cambridge University Press, Cambridge, 2000), p. 31
17. J. Robertson, Appl. Surf. Sci. **190**, 2 (2002)
18. S. Kurtin, T.C. McGill, C.A. Mead, Phys. Rev. Lett. **22**(26), 1433–1436 (1969)
19. S.G. Louie, J.R. Chelikowsky, M.L. Cohen, J. Vac. Sci. Technol. **13**, 790 (1976)
20. H.B. Michaelson, IBM J. Res. Develop. **22**(1), 72–80 (1978)
21. W. Gordy, W.J.O. Thomas, Phs. Rev. **24**, 439 (1956)
22. S.G. Louie, J.R. Chelikowsky, M.L. Cohen, J. Vac. Sci. Technol. **13**(4), 790–797 (1976)
23. M. Schluter, J. Vac. Sci. Technol. **15**(4), 1347–1376 (1978)
24. V. Narayanan, et al., in *Digest of Technical Papers of Symposium on VLSI Technology*, pp. 178–179 (2006)
25. H. Jagannathan, V. Narayanan, S. Brown, Abstract ECS Transactions, Meet, **16**(5), 19–26 (2008)
26. H.R. Harris, et al., IEEE International Electron Devices Meeting IEDM, pp. 1–4 (2006)
27. V.S. Chang, et al., IEEE International Electron Devices Meeting IEDM 2007, **10**–**12**, 535–538 (2007)
28. J. Schaeffer, M. Raymond, D. Gilmer, R. Gregory, B. Taylor, J. Jiang, D. Triyoso, R. Hegde, S. Samavedam, ECS Trans. **16**(5), 3–10 (2008)
29. M. Goto, S. Kawanaka, S. Inumiya, N. Kusunoki, M. Saitoh, K. Tatsumura, A. Kinoshita, S. Inaba, Y. Toyoshima, in *Symposium on VLSI Technology*, pp. 214–215 (2009)
30. H.N. Alshareef, H.F. Luan, K. Choi, H.R. Harris, H.C. Wen, M.A. Quevedo-Lopez, P. Majhi, B.H. Lee, Appl. Phys. Lett. **88**, 112–114 (2006)
31. B. P. Linder, V. Narayanan, V. K. Paruchuri, E. Cartier, S. Kanakasabapathy, Technical Digest Solid State Devices and Materials Conference Japanese Applied Physical Society, Tokyo, 16 (2007)
32. C.C. Hobbs, L.R.C. Fonseca, A. Knizhnik, V. Dhandapani, S.B. Samavedam, W.J. Taylor, J.M. Grant, L. Dip, D.H. Triyoso, R.I. Hegde, D.C. Gilmer, H.H. Tseng, B.E. White, P.J. Tobin, IEEE Trans. Electron. Devices **51**, 971 (2004)
33. S.B. Samavedam, et al., IEDM Tech. Dig., 307 (2003)
34. H.-H. Tseng, P. Kirsch, C.S. Park, G. Bersuker, P. Majhi, M. Hussain, R. Jammy, Microelectron. Eng. **86**(7–9), 1722–1727 (2009)
35. D.G. Schlom, J.H. Haeni, MRS Bull. **27**(03), 198–204 (2002)
36. S. Guha, V.K. Paruchuri, M. Copel, V. Narayanan, Y.Y. Wang, P.E. Batson, N.A. Bojarczuk, B. Linder, B. Doris, Appl. Phys. Lett. **90**, 092902 (2007)
37. M. Copel, S. Guha, N. Bojarczuk, E. Cartier, V. Narayanan, V. Paruchuri, Appl. Phys. Lett. **95**(21), 212903 (2009)
38. J. Robertson, L. Lin, Appl. Phys. Lett. **95**, 012906 (2009)
39. T. Ando, M.M. Frank, K. Choi, C. Choi, J. Bruley, M. Hopstaken, M. Copel, E. Cartier, A. Kerber, A. Callegari, D. Lacey, S. Brown, Q. Yang, V. Narayanan, IEEE International Electron Devices Meeting (IEDM), pp. 1–4 (2009)
40. J.K. Schaeffer, Appl. Phys. Lett., **85**(10), 1826 (2004)
41. E. Cartier, et al., in *Proceedings of VLSI Symposium*, p. 230 (2005)
42. S. Guha, V. Narayanan, Phys. Rev. Lett. **98**, 196101 (2007)
43. E. Cartier, M. Hopstaken, M. Copel, Appl. Phys. Lett. **95**, 042901 (2009)
44. E. Cartier, M. Steen, B.P. Linder, T. Ando, R. Iijima, M. Frank, J.S. Newbury, Y.H. Kim, F.R. McFeely, M. Copel, R. Haight, C. Choi, A. Callegari, V.K. Paruchuri, V. Narayanan, in *Symposium on VLSI Technology*, pp. 42–43 (2009)
45. K. Shiraishi, K. Yamada, K. Torri, Y. Akasak, K. Nakajima, M. Konno, T. Chikyow, H. Kitajima, T. Arikado, Jap. J. Appl. Phys. **52**, L1413 (2004)
46. M. Chudzik, et al., in *IEEE Symposium on VLSI Technology*, pp. 194–195 (2007)
47. K. Henson, H. Bu, M.H. Na, Y. Liang, U. Kwon, S. Krishnan, J. Schaeffer, R. Jha, N. Moumen, R. Carter, C. DeWan, R. Donaton, D. Guo, M. Hargrove, W. He, R. Mo, R. Ramachandran, K. Ramani, K. Schonenberg, Y. Tsang, X. Wang, M. Gribelyuk, W. Yan,

J. Shepard, E. Cartier, M. Frank, E. Harley, R. Arndt, R. Knarr, T. Bailey, B. Zhang, K. Wong, T. Graves-Abe, E. Luckowski, D.-G. Park, V. Narayanan, M. Chudzik, M. Khare, IEEE International Electron Devices Meeting 2008 (IEDM 2008), pp. 1–4 (2008)
48. K. Mistry et al., in *IEEE International Electron Devices Meeting 2007 (IEDM 2007)*, pp. 247–250, 2007
49. S. Natarajan, M. Armstrong, M. Bost, R. Brain, M. Brazier, C.-H. Chang, V. Chikarmane, M. Childs, H. Deshpande, K. Dev, G. Ding, T. Ghani, O. Golonzka, W. Han, J. He, R. Heussner, R. James, I. Jin, C. Kenyon, S. Klopcic, S.-H. Lee, M. Liu, S. Lodha, B. McFadden, A. Murthy, L. Neiberg, J. Neirynck, P. Packan, S. Pae, C. Parker, C. Pelto, L. Pipes, J. Sebastian, J. Seiple, B. Sell, S. Sivakumar, B. Song, K. Tone, T. Troeger, C. Weber, M. Yang, A. Yeoh, K. Zhang, in *IEEE International Electron Devices Meeting 2008 (IEDM 2008)*, pp. 1–3, 2008
50. F. Ducroquet, H. Achard, F. Coudert, B. Previtali, J-F., Lugand, L. Ulmer, T. Farjot, Y. Gobil, M. Heitzmann, S. Tedesco, M-E. Nier, S. Deleonibus, Trans. Electron Devices **48**, 8, (2001)
51. R. Li, Q. Xu, Trans. Electron Devices **49**, 11 (2002)
52. J. Schaeffer et al., J. Vac. Sci. Tech. B **21**(1), 11 (2003)
53. H. Pierson, Handbook of Refractory Carbides and Nitrides (Noyes, 1996)
54. P. Ranade, H. Takeuchi, T-J. King, C. Hu, Electrochem. Sol. State Lett., 4 G85 (2001)
55. Y. Saito, S. Kawata, H. Nakane, H. Adachi, Appl. Surf. Sci. **146**, 177 (1999)
56. C.S. Kang, H.-J. Cho, Y.H. Kim, R. Choi, K. Onishi, A. Shahriar, J.C. Lee, J. Vac. Sci. Technol. **B 21**, 5 (2003)
57. M. Moriwaki, T. Yamada, Jpn. J. Appl. Phys., 40 Part 1, **4**B 2679 (2001)
58. D-G. Park, H-J. Cho, K-Y. Lim, T-H. Cha, I-S. Yeo, J-W. Park, J. Electrochem. Soc., **148**(9), F189 (2001)
59. D-G. Park, K-Y. Lim, H-J. Cho, T-H. Cha, J–J. Kim, J-K. Ko, I-S. Yeo, J-W. Park, in *Symposium VLSI Technology Digest*, pp. 65–66 (2001)
60. H.Y. Yu, H.F. Lim, J.H. Chen, M.F. Li, C. Zhu, C.H. Tung, A.Y. Du, W.D. Wang, D.Z. Chi, D.-L. Kwong, IEEE Elec. Dev. Lett **24**, 4 (2003)
61. T-H. Cha, D-G. Park, T-K. Kim, S-A Jang, I-S. Yeo, Appl. Phys. Lett., **81**(22), 4192 (2002)
62. F. Esaka, K. Furuya, H. Shimada, M. Imamura, N. Matsubayashi, T. Kikuchi, H. Ichimura, A. Kawana, Surf. Interface Anal. **27**, 1098 (1999)
63. J.K. Schaeffer, C. Capasso, L.R.C. Fonseca, S. Samavedam, D.C. Gilmer, Y. Liang, S. Kalpat, B. Adetutu, H.-H. Tseng, Y. Shiho, A. Demkov, R. Hegde, W.J. Taylor, R. Gregory, J. Jiang, E. Luckowski, M.V. Raymond, K. Moore, D. Triyoso, D. Roan, B.E. White Jr, P.J. Tobin, IEDM Tech. Dig., p. 287 (2004)
64. H.-H. Tseng, C.C. Capasso, J.K. Schaeffer, E.A. Hebert, P.J. Tobin, D.C. Gilmer, D. Triyoso, M.E. Ramon, S. Kalpat, E. Luckowski, W.J. Taylor, Y. Jeon, O. Adetutu, R.I. Hegde, R. Noble, M. Jahanbani, C. El Chemali, B.E. White, IEEE International Electron Devices Meeting (IEDM) Technical Digest, pp. 821–824 (2004)
65. J.K. Schaeffer, D.C. Gilmer, C. Capasso, S. Kalpat, B. Taylor, M.V. Raymond, D. Triyoso, R. Hegde, S.B. Samavedam, B.E. White Jr, Microelectron. Eng. **84**(9–10), 2196–2200 (2007)
66. J.K. Schaeffer, C. Capasso, R. Gregory, D. Gilmer, L.R.C. Fonseca, M. Raymond, C. Happ, M. Kottke, S.B. Samavedam, P.J. Tobin, B.E. White Jr, J. Appl. Phys. **101**, 014503 (2007)
67. M.-A. Nicolet, P.H. Giauque, Microelectron. Eng. **55**, 357 (2001)
68. Y-S. Suh, G. Heuss, V. Misra, J. Electrochem. Soc. **150**(5), F79–F82 (2003)
69. Y.-S. Suh, G.P. Heuss, J.-H. Lee, V. Misra, Electron. Dev. Lett **24**(7), 439–441 (2003)
70. D-G. Park, T-H. Cha, K-Y. Lim, H-J. Cho, T-K. Kim, S-A. Jang, Y-S. Suh, V. Misra, I-S. Yeo, J-S. Roh, J-W. Park, H-K. Yoon, IEDM Technical Digest, 671 (2001)
71. M.J. Hopstaken, J. Bruley, D. Pfeiffer, M. Copel, M.M. Frank, E. Cartier, T. Ando, V. Narayanan, ECS Trans. **28**(1), 105–113 (2010)
72. Y.-C. Yeo, Q. Lu, P. Ranade, H. Takeuchi, K.J. Yang, I. Polishchuk, T.-J. King, C. Hu, S.C. Song, H.F. Luan, D.-L. Kwong, IEEE Electron. Dev. Lett **22**(5), 227–229 (2001)

73. S. B. Samavedam, L. B. La, J. Smith, S. Daksha-Murthy, E. Luckowski, J. Schaeffer, M. Zavala, R. Martin, V. Dhandapani, D. Triyoso, H. H. Tseng, P. J. Tobin, D. C. Gilmer, C. Hobbs, W. J. Taylor, J. M. Grant, R. Hegde, J. Mogab, C. Thomas, P. Abramowitz, M. Moosa, J. Conner, J. Jiang, V. Aruachalam, M. Sadd, B-Y. Nguyen, B. White, IEDM Technical Digest, pp. 433–436 (2002)
74. S.C. Song, Z.B. Zhang, M.M. Hussain, C. Huffman, J. Barnett, S.H. Bae, H.J. Li, P. Majhi, C.S. Park, B.S. Ju, H.K. Park, C.Y. Kang, R. Choi, P. Zeitzoff, H.H. Tseng, B.H. Lee, R. Jammy, in *Symposium on VLSI Technology Digest of Technical Papers*, p. 13–14 (2006)
75. H.N. Alshareef, et al., in *Symposium on Digest of Technical Papers VLSI Technology*, 7–8 (2006)
76. S. Kubicek, T. Schram, E. Rohr, V. Paraschiv, R. Vos, M. Demand, C. Adelmann, T. Witters, L. Nyns, A. Delabie, L.-Å. Ragnarsson, T. Chiarella, C. Kerner, A. Mercha, B. Parvais, M. Aoulaiche†, C. Ortolland, H. Yu, A. Veloso, L. Witters, R. Singanamalla†, T. Kauerauf†, S. Brus, C. Vrancken, V.S. Chang1, S-Z. Chang1, R. Mitsuhashi2, Y. Okuno2, A. Akheyar3, H.-J. Cho4, J. Hooker5, B. J. O'Sullivan, S. Van Elshocht, K. De Meyer†, M. Jurczak, P. Absil, S. Biesemans, T. Hoffmann, in *Symposium on VLSI Technology Digest of Technical Papers*, pp. 130–131 (2008)
77. H.R. Harris, P. Kalra, P. Majhi, M. Hussain, D. Kelly, O. Jungwoo, D. He, C. Smith, J. Barnett, P.D. Kirsch, G. Gebara, J. Jur, D. Lichtenwalner, A. Lubow, T.P. Ma, S. Guangyu, S. Thompson, B.H. Lee, H.-H. Tseng, R. Jammy, in *IEEE Symposium on VLSI Technology*, pp. 154–155 (2007)
78. J-H. Lee, H. Zhong, Y.-S. Suh, G. Heuss, J. Gurganus, B. Chen, V. Misra, IEDM Technical Digest, pp. 359–362 (2002)
79. V. Misra, H. Zhong, H. Lazar, IEEE Electron. Dev. Lett. **23**(6), 354–356 (2002)
80. I. Polishchuk, P. Ranade, T-J. King, C. Hu, IEEE Electron Dev. Lett., **22**(9), 444–446 (2001)
81. B.-Y. Tsui, C.-F. Huang, IEEE Electron. Dev. Lett. **24**(3), 153–155 (2003)
82. Q. Lu, R. Lin, P. Ranade, T-J. King, C. Hu, in *Symposium on IEEE VLSI*, pp. 45–46 (2001)
83. P. Ranade, H. Takeuchi, T.-J. King, C. Hu, Electrochem. Sol. State Lett. **4**, G85 (2001)
84. H. Wakbayashi, Y. Saito, K. Takeuchi, T. Mogami, T. Kunio, IEEE Trans. Electron Dev. **48** 2363 (2001)
85. R.J.P. Lander, J.C. Hooker, J.P. van Zijl, F. Roozeboom, M.P.M. Maas, Y. Tamminga, R.A.M. Wolters, Mater. Res. Soc. Proc. V **716**, 253 (2002)
86. T. Schram, S. Kubicek, E. Rohr, S. Brus, C. Vrancken, S.-Z. Chang, V.S. Chang, R. Mitsuhashi, Y. Okuno, A. Akheyar, H.-J. Cho, J.C. Hooker, V. Paraschiv, R. Vos, F. Sebai, M. Ercken, P. Kelkar, A. Delabie, C. Adelmann, T. Witters, L.-A. Ragnarsson, C. Kerner, T. Chiarella, M. Aoulaiche, M.-J. Cho, T. Kauerauf, K. De Meyer, A. Lauwers, T. Hoffmann, P.P. Absil, S. Biesemans, in *Symposium on VLSI Technology*, pp. 44–45 (2008)
87. C. Auth et al., in *Symposium on VLSI Technology*, 128 (2008)
88. H.-C. Wen, S.C. Song, C.S. Park, C. Burham, G. Bersuker, K. Choi, M.A. Quevedo-Lopez, B.S. Ju, H.N. Alshareef, H. Niimi, H.B. Park, P.S. Lysaght, P. Majhi, B.H. Lee, in *IEEE Symposium on VLSI Technology*, pp. 160–161 (2007)