Continuous and Precise Work Function Adjustment for Integratable Dual Metal Gate CMOS Technology Using Hf–Mo Binary Alloys

Tzung-Lin Li, Chia-Hsin Hu, Wu-Lin Ho, Howard C.-H. Wang, and Chun-Yen Chang, *Fellow, IEEE*

***Abstract*—We demonstrate for the first time a continuous and almost linear work function adjustment between 3.93 and 4.93 eV using $Hf_xMo_{(1-x)}$ binary alloys deposited by co-sputtering. In view of the process integration, dual work function metal gate technology using Mo and $Hf_xMo_{(1-x)}$ formed by metal intermixing was proposed. Work function values were verified to be a function of the thickness ratio and accurate work function adjustment can be possible. Furthermore, one can be allowed to get around the thermal stability issue by choosing an appropriate total metal thickness corresponding to the thermal budget subsequent to gate deposition, since the thermal budget required for metal intermixing depends on the total metal thickness.**

***Index Terms*—Alloy, CMOS, hafnium (Hf), intermixing, metal gate, molybdenum (Mo), work function.**

I. Introduction

WITH the sustained scaling of CMOS technology for device performance improvement, the conventional polysilicon–SiO_2 structure keeps suffering a variety of challenges. Novel metal/high-κ gate stack has been extensively investigated as a potential solution. The introduction of high-κ gate dielectric can effectively reduce the tunneling leakage current due to its larger physical thickness under the same electrical thickness [1], [2]. On the other hand, the inherent drawback of a polysilicon gate such as a poly depletion effect and boron penetration which will lead to an undesired increase of effective oxide thickness (EOT) [3] and degrade device performance [4] can be eliminated. In addition, polysilicon gates are found to be thermodynamically unstable on many high-κ materials [5], [6] so that metals are expected to provide a lower gate resistance and a turning point in possessing better thermal stability.

The major superiority of the traditional polysilicon gate electrode is the ability to make a Fermi-level adjustment by either donor or acceptor implantation. In contrast, the adjustment of the metal work function is not easily achievable. For bulk devices, the required metal work functions for replacing the conventional n^+- and p^+-polysilicon gates are about 4 and 5 eV, respectively. On the other hand, for FinFET and/or ultrathin body silicon-on-insulator (UTB-SOI) devices, the gate over channel controllability is enhanced so that the required gate work function for n-channel (p-channel) devices have been increased (reduced) to 4.4–4.6 eV (4.8–5.0 eV) compared with bulk devices [7]. Since the required work function values for n- and p-channel devices are different in both cases, the dual metal gate technology with suitably chosen metal work function values has been proposed while the integration will lead to the unwanted gate dielectric integrity degradation [8] due to the removal of metal directly from the dielectric surface. It is also worth noting that, the adjustment of the substrate doping is no longer an effective way of threshold voltage control in FinFET and/or UTB-SOI devices, the importance of metal gate work function engineering will be more pronounced. Several metal work function modulation techniques have been investigated. The Ru-Ta alloys proposed by Zhong *et al.* can possess superior thermal stability and wide work function tuning range [9], [10]. However, the work function modulation seems not to be continuous so that the work function values with interest for advanced transistor structures (4.4–5 eV) would be unachievable. On the other hand, the Pt–Ta alloys proposed by Tsui *et al.* are demonstrated to possess wide and continuous work function modulation [11], but the issue of gate dielectric integrity degradation mentioned in [8] would be problematic due to the lack of suitable integration methods. Similarly, S. H. Bae *et al.* proposed that the laminated metal gate stacks HfN–Ti–TaN and Ti–Ta can possess p- and nMOS compatible work function values (5.1 and 4.35 eV) [12], however, the process integration of these two distinct laminated metal stacks into dual metal gate CMOS process are still problematic [8]. Otherwise, a novel work function modulation using nitrogen implanted Mo was proposed by Ranade *et al.* [13]. The major advantage of this method is the ease of process integration, while the Φ_m value strongly depends on the implant parameters and subsequent annealing conditions. Precise work function modulation would not be easily achievable.

In this paper, we first demonstrate the continuous and almost linear work function adjustment using $Hf_xMo_{(1-x)}$ binary alloys deposited by co-sputtering. Second, we propose a dual metal gate technology employing Mo and $Hf_xMo_{(1-x)}$ as gate electrodes for p- and n-channel devices, respectively. In view of the ease of process integration, the $Hf_xMo_{(1-x)}$ electrode in the proposed dual metal gate technology will be formed by metal intermixing. Metals need not to be removed from the dielectric interface and thus the uniformity and integrity of gate dielectric can be preserved. Furthermore, a parameter called the optimal annealing temperature, $T_{A,\mathrm{opt}}$, was expected to provide a prospective work function modulation without causing EOT

Manuscript received Septemeber 27, 2004; revised March 1, 2005. This work was supported by the Taiwan Semiconductor Manufacturing Company No. 94C001. The review of this paper was arranged by Editor V. R. Ramgopal.

T.-L. Li, C.-H. Hu, W.-L. Ho, and C.-Y. Chang are with the Institute of Electronics, National Chiao-Tung University, Hsinchu 30050, Taiwan, R.O.C. (e-mail: tlli.ee89g@nctu.edu.tw).

H. C.-H. Wang is with the Taiwan Semiconductor Manufacturing Company, Hsinchu 300, Taiwan, R.O.C.

Digital Object Identifier 10.1109/TED.2005.848108

0018-9383/$20.00 © 2005 IEEE

TABLE I
SAMPLE CONDITIONS AND EXTRACTED Φ_m OF CO-SPUTTERING EXPERIMENT. (ALLOY SAMPLES: 400 °C, 30 S; CONTROL SAMPLES: AS-DEPOSITED)

Sample	Hf power (W)	Mo power (W)	Hf power ratio (%)	Φ_m (eV)
ctrl.	0	150	0 (pure Mo)	4.932
1-1	22	150	12.8	4.727
1-2	50	150	25	4.597
1-3	90	150	37.5	4.492
1-4	150	150	50	4.350
1-5	150	90	62.5	4.213
1-6	150	50	75	4.150
1-7	150	22	87.5	4.062
ctrl.	150	0	100 (pure Hf)	3.930

variation. The value of $T_{A,\text{opt}}$ will strongly affect the application of the proposed technique. Since the phenomenon of metal intermixing is based on solid diffusion, we demonstrate that the optimal annealing temperature can be raised by increasing the metal thickness ($T_M = T_{\text{Hf}}+T_{\text{Mo}}$). Consequently, it is likely to overcome the thermal stability in conventional CMOS process by choosing an appropriate metal thickness.

In addition, the composition and work function of $\text{Hf}_x\text{Mo}_{(1-x)}$ were demonstrated to depend on the thickness combination of metal layers. The oxidation of Hf at the top surface leading to extra Hf consumption can be observed. A modified multilayer structure (TiN–Mo–Hf–Mo) was verified to improve the immunity to metal oxidation, and a quadratic equation relating work function (Φ_m) to composite metal thickness ratio ($T_R = T_{\text{Hf}}/T_{\text{Mo}}$) was also derived. Good consistency with experimental data assures the possibility of precise metal work function adjustment.

II. EXPERIMENTS

A. Co-sputtering

MOSCAP devices were fabricated on p-type (100) 6-in Si wafers and high frequency (1 MHz) capacitance–voltage (C–V) characteristics were measured using an Agilent 4284A precision LCR meter. After LOCOS isolation, SiO_2 with different thicknesses were thermally grown at 950 °C to serve as the gate dielectric. $\text{Hf}_x\text{Mo}_{(1-x)}$($\sim$ 50 nm) alloys were then deposited by co-sputtering in Ar ambient. The sputtering power of each target was varied as listed in Table I to modulate the composition of binary alloy. The $\text{Hf}_x\text{Mo}_{(1-x)}$ gate electrodes were then patterned by reactive ion etching (RIE) using Cl_2—based chemistry. All samples were then subjected to 400 °C annealing in N_2 ambient. The flatband voltage (V_{FB}) and EOT of each capacitor were extracted from the measured C–V curve using quantum mechanical C–V (QMCV) simulator so that one can avoid overestimating the EOT as well as extracted metal work function values [14].

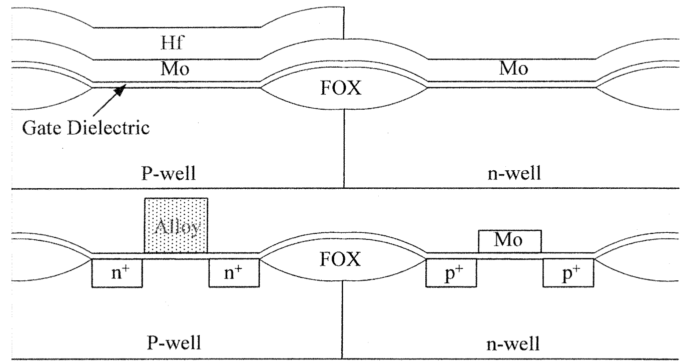


Fig. 1. Schematics of dual metal gate technology using metal and alloy formed by metal intermixing. Metals need not to be etched away from the dielectric surface so the uniformity and integrity of gate dielectric can be preserved.

B. Metal Intermixing

Fig. 1 shows the fabrication process of the proposed dual metal gate technology using Mo and $\text{Hf}_x\text{Mo}_{(1-x)}$ as gate electrodes. After LOCOS isolation and the gate dielectric deposition, the first layer metal Mo and the second layer metal Hf were then deposited over the entire wafer. A noncritical lithography step was performed and the second layer metal Hf was then selectively removed from the pMOS side. After gate electrodes patterning and S/D implantation, the thermal annealing was performed for dopant activation and metal intermixing at the nMOS side simultaneously.

To demonstrate this technique, MOSCAP devices were fabricated. The process flow was similar to that of co-sputtering experiment except that Mo and Hf were deposited in sequence. The Mo gate has been reported to possess high thermal stability ($>$ 1000 °C on SiO_2 gate dielectric [15]. Moreover, the work function value of Mo film varies with the bulk metal microstructure and consequently depends on deposition and annealing conditions [16]. The Mo film with (110) orientation can maintain possessing high work function value suitable for p-channel devices up to 900 °C [13]. By contrast, the thermal stability of pure Hf gate on SiO_2 was reported to be poor [17]. Therefore, Mo was chosen as the first layer metal in this work to retard the unwanted interaction between metal and SiO_2 during subsequent thermal treatment. The composition of binary alloy was modulated by varying thickness of each metal layer as listed in Table II. The Hf atomic fraction in each $\text{Hf}_x\text{Mo}_{(1-x)}$ alloy was approximately predicted by

$$\frac{x}{(1-x)} = \frac{9.38 T_{\text{Hf}}}{13.44 T_{\text{Mo}}} \tag{1}$$

where 13.44 and 9.38 are the molar volume of Hf and Mo, respectively. After gate electrode patterning, samples were then subjected to different rapid thermal annealing conditions in N_2 ambient to lead to metal intermixing for alloy formation and simulate the possible thermal cycle in the conventional CMOS process.

TABLE II
SAMPLE CONDITIONS AND EXTRACTED Φ_m OF METAL INTERMIXING EXPERIMENT. (700 °C, 30 s FOR SAMPLES 2–1 AND 2–2; 600 °C, 30 s FOR SAMPLES 2–3 AND 2–4)

Sample	T_{Mo} (Å)	T_{Hf} (Å)	Hf atomic fraction (%)	Φ_m (eV)
ctrl.	500	0	0 (pure Mo)	4.93
2-1	406	94	14	4.81
2-2	294	206	33	4.63
2-3	256	244	40	4.55
2-4	205	295	50	4.43

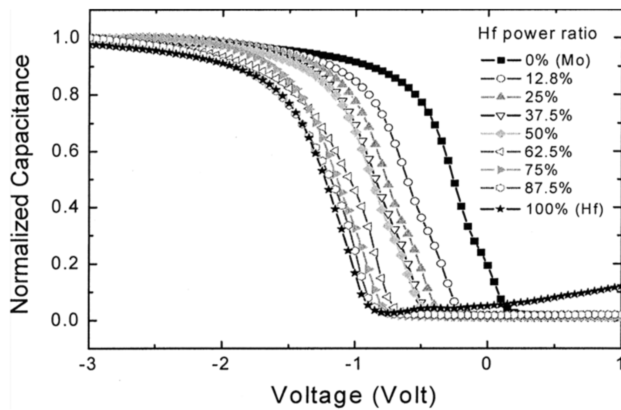


Fig. 2. C–V curves of as-deposited co-sputtering samples. Wide-ranging flat band voltage shift can be observed.

III. RESULTS AND DISCUSSION

A. Co-sputtering

The as-deposited high-frequency (1 MHz) C–V characteristics of capacitors gated by Hf–Mo binary alloys, pure Mo, and pure Hf films are shown in Fig. 2. The negative $V_{\rm FB}$ shift with the increase of Hf power ratio can be observed. To eliminate the contribution of oxide fixed charge, C–V measurements of MOSCAP with several oxide thicknesses were used to generate the $V_{\rm FB}$ versus EOT curve as shown in Fig. 3. All samples exhibited linear relationship from which the work function of each binary alloy can be extracted as listed in Table I. The extracted 4.93 eV work function for pure Mo film with (110) morphology demonstrated by the XRD spectra in Fig. 4 is close to other previous reports [13]. The as-deposited Mo with (110) orientation is different from the previous report [16], and this would be attributed to the different deposition condition. In addition, the process quality was also demonstrated according to the C–V measurement of the 50% Hf power ratio co-sputtering sample which exhibited a small hysteresis as shown in Fig. 5. The C–V curves of as-deposited and post-400 °C sintering pure Hf gated capacitor are shown in Fig. 6. Obvious EOT variation and V_{FB} shift after annealing illustrate the poor thermal stability of Hf on SiO_2 and exclude Hf from gate candidates even in the gate last process [18]. On the other hand, $Hf_xMo_{(1-x)}$ gated capacitors exhibit better thermal stability on SiO_2 as shown in Fig. 7. The dependence of Φ_m and EOT variation on annealing temperature of $Hf_xMo_{(1-x)}$ also indicates that the thermal stability of all alloy samples can be at least higher than 400 °C as shown in Fig. 8.

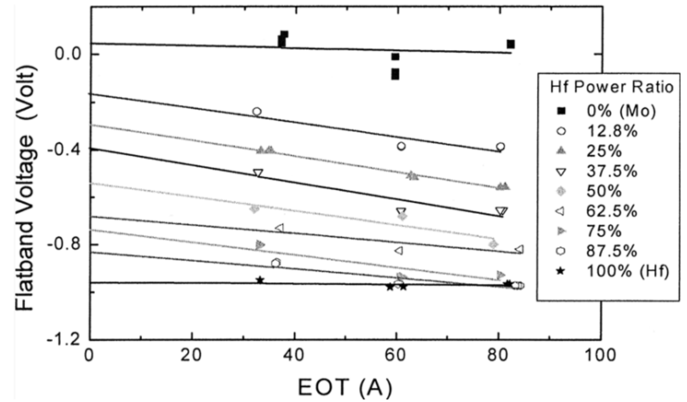


Fig. 3. All samples (Table I) exhibited linear behavior in $V_{\rm FB}$ versus EOT curves from which work function of each alloy could be extracted. (alloy samples: 400 °C, 30 s; control samples : as-deposited).

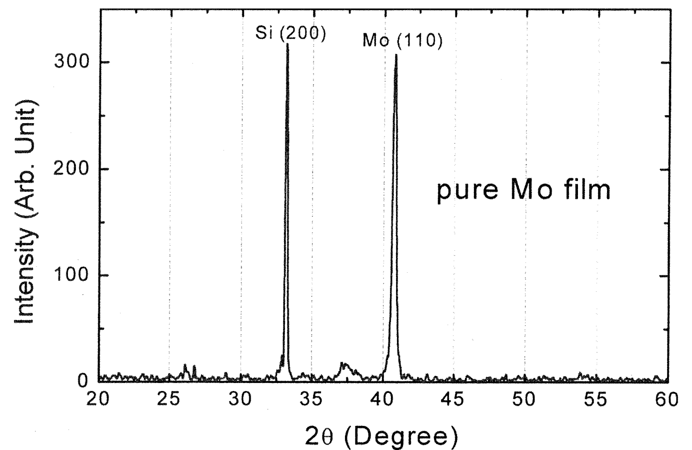


Fig. 4. As-deposited pure Mo film is found to have (110) orientation.

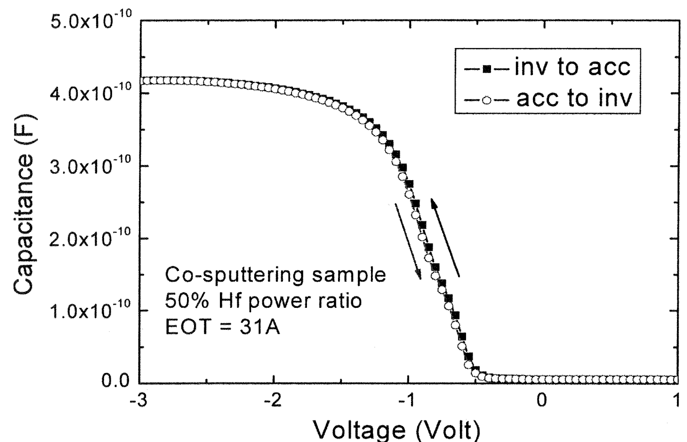


Fig. 5. Process quality was demonstrated, since the 50% Hf power ratio sample suffering maximum power summation during metal deposition still exhibited negligible hysteresis.

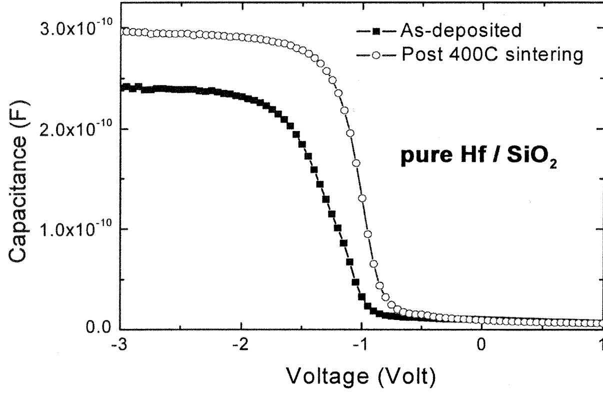


Fig. 6. C–V curve of post-400 °C annealing Hf gated capacitor shows noticeable EOT variation and flatband voltage shift.

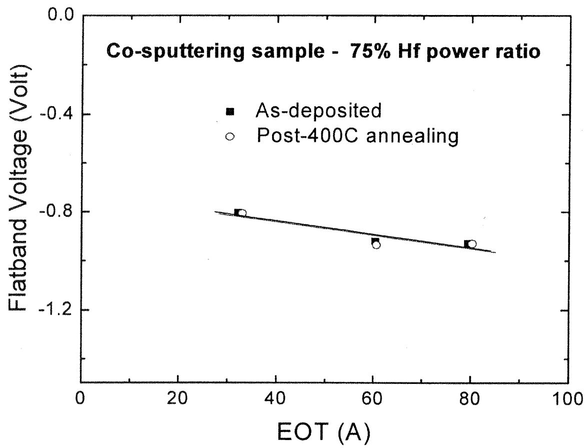


Fig. 7. Post-400 °C annealing co-sputtering sample shows negligible work function variation.

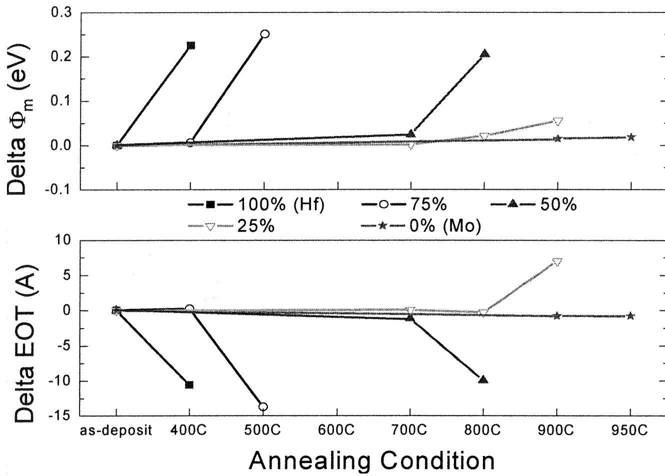


Fig. 8. Dependence of Φ_m and EOT variation on annealing temperature show that the thermal stability of alloy samples can be at least higher than 400 °C.

In Fig. 8, the decreases of EOT at certain temperature for alloy samples except for the one with 25% Hf power ratio were similar to the pure Hf sample. The noticeable EOT decrease might be attributed to that partial SiO_2 gate dielectric was transformed into high-κ materials such as HfO_2 or $HfSi_xO_y$ due to Hf–SiO_2 interaction and leading to corresponding noticeable work function variation. As for the 25% alloy sample, the abnormal EOT increase along with relative small work function variation can be observed make us speculate that the lower Hf concentration might make the effect of Hf–SiO_2 interaction be masked by the extra Si-substrate oxidation due to oxygen contamination. In the case of sample gated by pure Mo, the small amount of Φ_m increase (18 meV) along with negligible EOT variation (0.08 nm) after 950 °C RTA demonstrates the superior thermal stability of Mo on SiO_2 gate dielectric. Although the thermal stability was found to be degraded with the increasing of Hf atomic fraction, $Hf_xMo_{(1-x)}$ can still be adopted as gate material in a gate-last SiO_2 CMOS process. It is also worth to note that, alloy samples with Hf power ratio lower than 50% can possess work function value required for advanced transistor structures and exhibit thermal stability up to 700 °C. Hf–Mo alloys possess lower thermal stabilities than Ru–Ta alloys [9], but higher than Pt–Ta alloys [11].

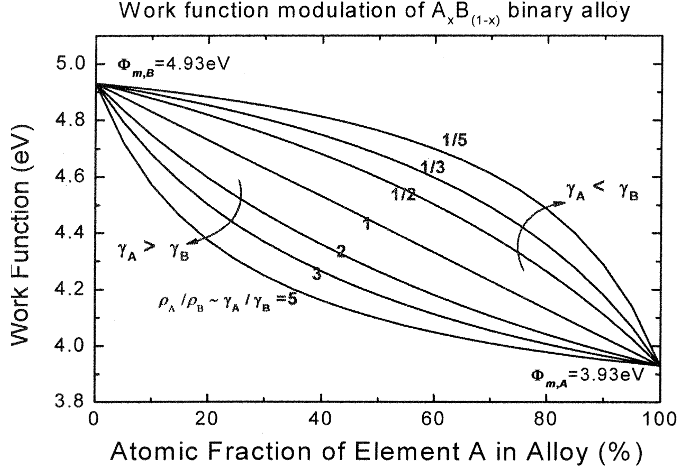


Fig. 9. Calculated work function value versus atomic fraction in binary alloy as a function of γ ratio. Metals with similar γ (Sommerfeld factor) will lead to a linear work function modulation which is a compromise between modulation efficiency and immunity to process variation. ($\Phi_{m,A}$ and $\Phi_{m,B}$ are set to be 3.93 and 4.93 eV for convenience).

In 1974, Gelatt and Ehrenreich proposed that the work function of an $A_xB_{(1-x)}$ alloy can be approximately expressed as [19]

$$\Phi_m(x) = x\Phi_{m,A} + (1-x)\Phi_{m,B} + x(1-x) \times \frac{(\Phi_{m,A} - \Phi_{m,B})\left(\frac{\rho_A}{\rho_B} - 1\right)}{x \cdot \frac{\rho_A}{\rho_B} + (1-x)} \quad (2)$$

where $\Phi_{m,A}$ and $\Phi_{m,B}$ are work function values of pure element A and B, respectively. ρ_A and ρ_B are effective density of states in Fermi level for pure element A and B, respectively. From (2), the first two terms represent that the work function of binary alloy is a linear combination of that of each pure element. On the other hand, the last term will lead to a deviation from the linear relationship. According to the theory of heat capacity of metal, the observable Sommerfeld factor γ of a metal is directly proportional to its density of state in Fermi level ρ [20]. The calculated results of (2) are shown in Fig. 9 where several γ ratios are used and values of $\Phi_{m,A}$ and $\Phi_{m,B}$ are set to be 3.93 and 4.93, respectively, for convenience. As expected, the work function modulation will deviate from the linear behavior with the

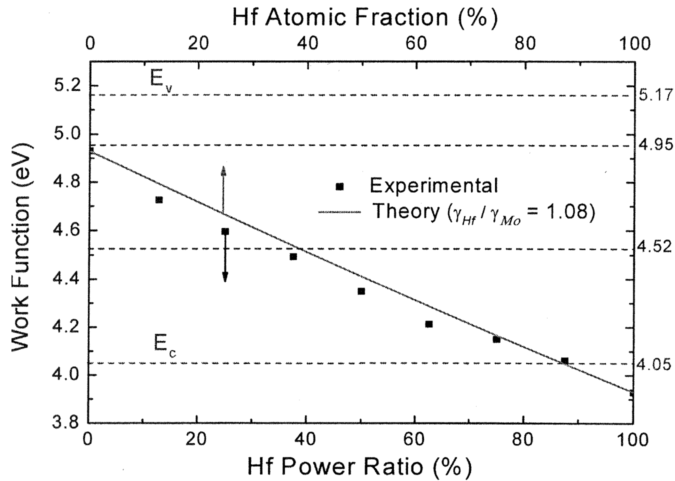


Fig. 10. Comparison between experimental (Table I) and theoretical work function values. A slightly deviation in lower Hf power ratio regime may be attributed to different sputtering rate between Hf and Mo. ($\gamma_{\rm Hf} = 2.16, \gamma_{\rm Mo} = 2.0$ are used for calculation).

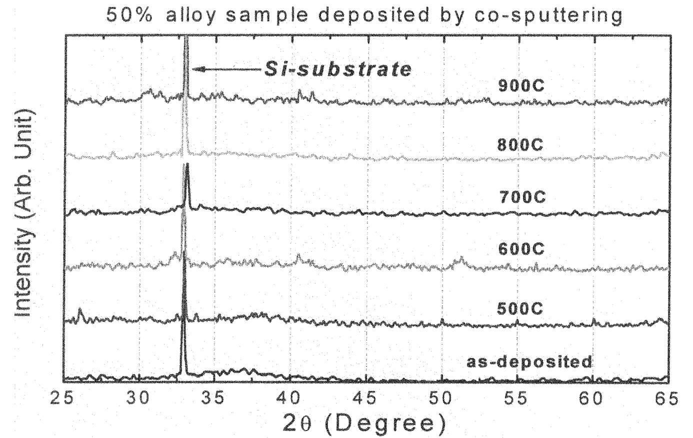


Fig. 11. XRD spectra of 50% Hf power ratio co-sputtering sample exhibited an amorphous film structure and only the c-Si was featuring.

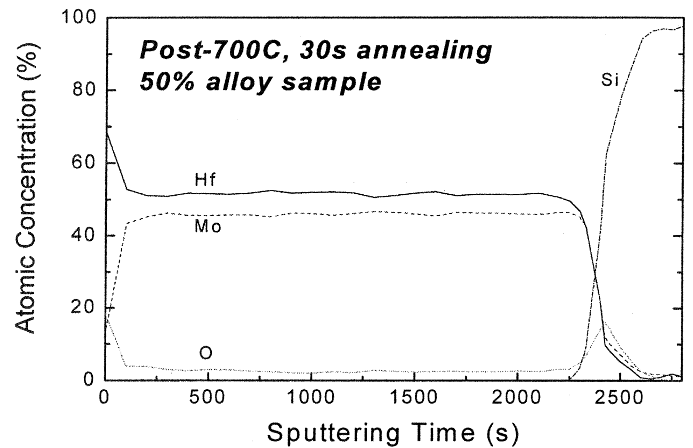


Fig. 12. AES profile of post-annealing 50% Hf power ratio co-sputtering sample. A uniform composition and abrupt interface can be observed.

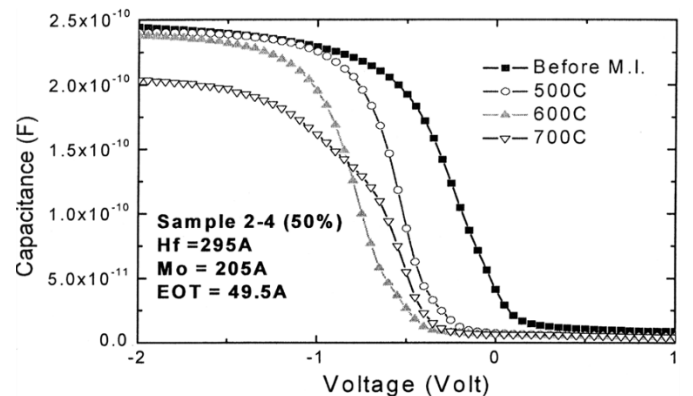


Fig. 13. C–V curves of Hf-295 Å/Mo-205 Å/ SiO_2 capacitor before and after thermal annealing. The optimal annealing temperature for this sample was found to be 600 °C.

difference in γ values between two metals. For a nonlinear behavior, work function modulation can be roughly divided into flat and sharp regime. In the flatter regime, the alloy system would be less susceptible to composition and process variation, but the Φ_m modulation efficiency will be lower. On the other hand, the alloy system will be more sensitive to process variation but possess better Φ_m modulation efficiency in the sharper regime. By contrast, the linear work function modulation can provide a compromise between modulation efficiency and immunity to process variation throughout the whole modulation range.

Abrupt work function modulation can be excluded since no specific compound will be formed under 1000 °C according to the binary alloy phase diagram of Hf-Mo system [21]. Moreover, the Sommerfeld factors for Hf and Mo are 2.16 and 2.0, respectively [20]. Thus a continuous and almost linear work function modulation using the $Hf_xMo_{(1-x)}$ solid solution can be expected. The calculated and experimental results of work functions of $Hf_xMo_{(1-x)}$ alloys are shown in Fig. 10. Compared with the experimental data, a good consistency with only a mild shift in the lower Hf power ratio regime can be observed. This deviation may be attributed to the difference between the Hf power ratio and the Hf atomic fraction due to a relatively lower deposition rate of Mo in our work. The XRD spectra and AES depth profile of the co-sputtering sample are shown in Figs. 11 and 12, respectively. An amorphous film with uniform composition and abrupt interface was obtained. It is worth noting that the relatively lower composition of Mo compared to that of Hf in the co-sputtering sample as shown in Fig. 12 also illustrates our speculation about the deviation observed in Fig. 10.

B. Metal Intermixing

Fig. 13 shows the C–V characteristics of an MOSCAP gated by Hf–Mo metal stack before and after rapid thermal annealing. The C–V results shows that the nearly optimal annealing temperature ($T_{A,\rm opt}$) for the 50% sample is 600 °C which can provide a sufficient and prospective metal work function modulation while keep EOT without variation. By contrast, smaller flat-band voltage shift for annealing temperature lower than 600 °C may be due to insufficient thermal budget needed for expected Hf concentration diffusion to the dielectric interface. As for annealing temperature higher than 600 °C, a noticeable capacitance decrease can be observed which is different with the observation in co-sputtering experiment. The novel increase of EOT might be attributed to the contribution

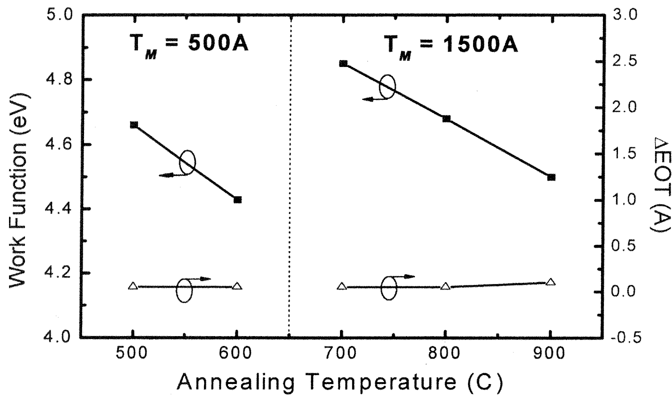


Fig. 14. Increase of total metal thickness under the same composite metal thickness ratio can effectively rise the optimal annealing temperature.

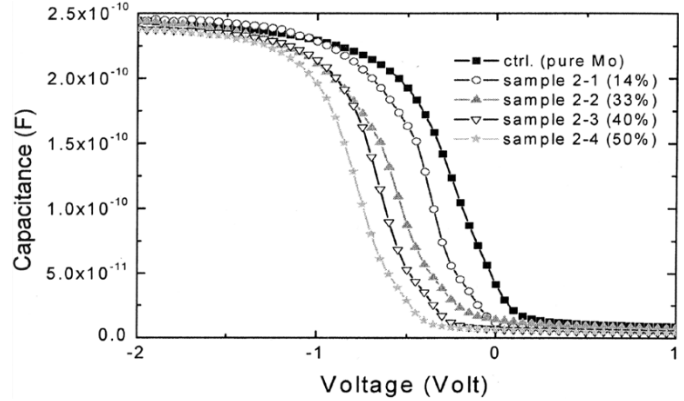


Fig. 15. C–V curves of post-annealing Hf–Mo–SiO_2 capacitors as a function of Hf atomic fraction.

of series capacitance due to metal oxidation and should be verified further. In view of process integration, lower $T_{A,opt}$ will restrict the proposed technique to the gate-last process where high temperature annealing for S/D dopant activation will be performed prior to the deposition of gate dielectric and gate electrode. It should be noted that the sum of metal thickness was kept constant (500 Å) for each sample in this work as listed in Table II. Since the phenomenon of metal intermixing is based on solid diffusion, the increase of first-layer metal thickness as well as the total metal thickness ($T_M = T_{\mathrm{Hf}} + T_{\mathrm{Mo}}$) under the same thickness ratio can be expected to raise the required thermal budget for prospective work function modulation. An investigation about the influence of total metal thickness T_M on the optimal annealing temperature $T_{A,opt}$ is shown in Fig. 14. Sample with 1500 Å total metal thickness was not optimized yet, but a positive correlation still can be observed. When the total metal thickness was increased from 500 Å to 1500 Å, the optimal annealing temperature for the 50% sample had been raised from 600 °C to 900 °C substantially. Consequently, one can possibly get around the thermal stability issue by using an appropriate total metal thickness corresponding to total thermal budget subsequent to the gate electrode deposition. According to the ITRS roadmap, however, the thickness of the gate electrode must be reduced with the miniaturization of MOSFET devices. To use thicker metal thickness for getting around the thermal stability issue in conventional CMOS process, the appropriate etch-back of gate electrode should be performed after the metal intermixing has almost finished. The extra gate electrode thin-down process without serious increase of the process complexity can be achievable. For instance, one can employ the interlevel dielectric (ILD) CMP for gate electrode etch-back, and only the polish-time control is needed to obtain the prospective gate electrode thickness. Compared to the gate-last process [18], the dummy gate removal, gate electrode re-deposition, and the gate electrode CMP will not be required.

The work function of $\mathrm{Hf}_x\mathrm{Mo}_{(1-x)}$ formed by metal intermixing was listed in Table II (optimal annealing temperature: 700 °C, 30 s for samples 2–1 and 2–2; 600 °C, 30 s for samples 2–3 and 2–4). The dependence of work function value on the thickness combination of the two metal layers can be observed. Fig. 15 shows the dependence of post-annealing capacitor-voltage characteristics on the Hf atomic fraction calculated

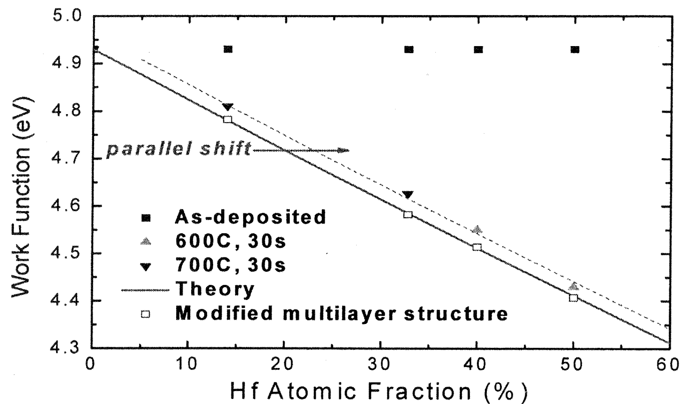


Fig. 16. Comparison between experimental (Table II) and theoretical results. A parallel shift may be attributed to the extra Hf consumption ($\sim 3\%$) due to surface oxidation. Also shown as open symbols are experimental results of multilayer (TiN–Mo–Hf–Mo) gated devices, good agreement on theoretical results can be achieved.

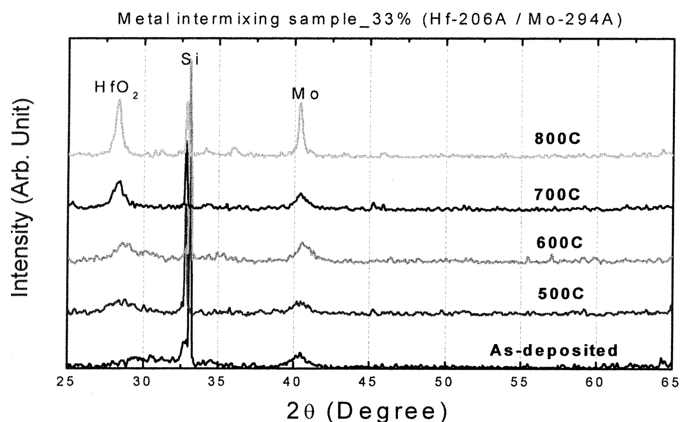


Fig. 17. XRD spectra of metal intermixing sample (Hf-206 Å/Mo-294 Å/SiO_2) exhibited HfO_2 peak as a result of oxidation of Hf at the top surface after thermal treatment.

by (1). Fig. 16 shows the dependence of extracted work function values on the Hf atomic fraction. A parallel shift of post-annealing work functions from the calculated values can be observed. This may be attributed to the extra consumption of Hf ($\sim 3\%$) due to surface oxidation. The XRD spectra in Fig. 17

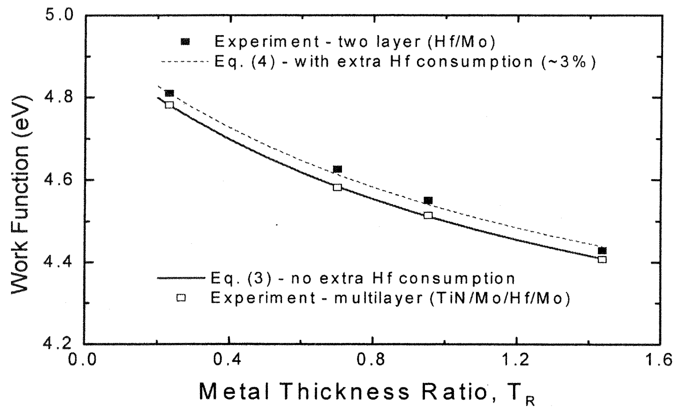


Fig. 18. Post-annealing work functions extracted from two-layer (Hf–Mo) and multilayer (TiN–Mo–Hf–Mo) gated MOSCAP versus the composite metal thickness ratio ($T_R = T_{\rm Hf}/T_{\rm Mo}$). Also shown are calculated results of derived quadratic equations with and without taking extra Hf consumption into account.

which exhibits HfO_2 peak also illustrate the speculation. With (1) and (2), one can obtain the following quadratic equation:

$$(1.053\Phi_m - 4.136)T_R^2 + (2.904\Phi_m - 12.808)T_R + (2\Phi_m - 9.86) = 0 \quad (3)$$

where T_R is the thickness ratio of Hf to Mo ($T_R = T_{\rm Hf}/T_{\rm Mo}$). If the extra consumption of Hf ($\sim 3\%$) due to surface oxidation is taken into account, (3) becomes

$$(1.05\Phi_m - 4.156)T_R^2 + (2.897\Phi_m - 12.864)T_R + (1.995\Phi_m - 9.89) = 0. \quad (4)$$

Fig. 18 shows the dependence of work function on the thickness ratio. Also shown are the calculated results of (3) and (4). A good consistency assures the possibility of precise work function tuning since the T_R value required for an expected Φ_m can be precisely determined. Furthermore, to overcome the surface oxidation of Hf, MOS capacitors with triple layer metal stack (Mo–Hf–Mo) capped by TiN (20 nm) were also fabricated and subjected to the same annealing conditions in the two-layer experiment. The extracted work function values are labeled in Figs. 16 and 18. In the multilayer experiments, the thickness of Mo in the two-layer experiment was split into 10 nm as first-layer metal and the remnant Mo was deposited atop Hf. Using the multilayer structure, the immunity against surface oxidation can be improved effectively. The extracted work function values will get emended and can be described accurately by (3).

IV. Conclusion

The continuous and almost linear work function adjustment using $Hf_xMo_{(1-x)}$ has been demonstrated for the first time. The work function value of Hf-Mo binary alloy deposited by co-sputtering ranges between 3.93 eV (Φ_m of pure Hf) and 4.93 eV (Φ_m of pure Mo). Although the thermal stability of $Hf_xMo_{(1-x)}$ on SiO_2 degrades with the increase of Hf atomic fraction, $Hf_xMo_{(1-x)}$ can still be used in a gate-last SiO_2 CMOS process. For the ease of process integration, $Hf_xMo_{(1-x)}$ formed by metal intermixing was also demonstrated. One can be allowed to get around the thermal stability issue by using an appropriate metal thickness T_M and possess precise controllability of metal work function by adjusting the composite metal thickness ratio T_R. This technique is not only attractive but is especially important for FinFET and/or UTB-SOI devices application.

Acknowledgment

The authors would like to thank the UC Berkeley Device Group for the use of the QMCV simulator.

References

[1] E. M. Vogel, K. Z. Ahmed, B. Hornung, W. K. Hensen, P. K. McLarty, G. Lucovsky, J. R. Hauser, and J. Wortman, "Modeled tunnel currents for high dielectric constant dielectrics," *IEEE Trans. Electron Devices*, vol. 45, no. 6, pp. 1350–1355, Jun. 1998.

[2] Q. Lu, D. Park, A. Kalnitsky, C. Chang, C. -C. Cheng, S. P. Tay, T. -J. King, and C. Hu, "Leakage current comparison between ultrathin Ta_2O_5 films and conventional gate dielectrics," *IEEE Electron Device Lett.*, vol. 19, no. 9, pp. 341–342, Sep. 1998.

[3] K. F. Schuegraf, C. C. King, and C. Hu, "Impact of polysilicon depletion in thin oxide MOS technology," in *Symp. VLSI Tech. Dig.*, 1993, pp. 86–90.

[4] J. R. Pfiester, F. K. Baker, T. C. Mele, H. H. Tseng, P. J. Tobin, J. D. Hayden, J. W. Miller, C. D. Gunderson, and L. C. Parrillo, "The effects of boron penetration on p^+ polisilicon gated pMOS devices," *IEEE Trans. Electron Devices*, vol. 37, no. 8, pp. 1842–1851, Aug. 1990.

[5] H. F. Luan, B. Z. Wu, L. G. Kang, R. Vrtis, D. Roberts, and D. L. Kwong, "Ultra thin high quality Ta_2O_5 gate dielectric prepared by *in situ* rapid thermal processing," in *IEDM Tech. Dig.*, 1998, pp. 609–612.

[6] C. H. Lee, H. F. Luan, W. P. Bai, S. J. Lee, T. S. Jeon, Y. Senzaki, D. Roberts, and D. L. Kwong, "MOS characteristics of ultra thin rapid thermal CVD ZrO_2 and Zr silicate gate dielectrics," in *IEDM Tech. Dig.*, 2000, pp. 27–30.

[7] L. Chang, S. Tang, T. J. King, J. Bokor, and C. Hu, "Gate length scaling and threshold voltage control of double-gate MOSFETs," in *IEDM Tech. Dig.*, 2000, pp. 719–722.

[8] Q. Lu, Y. C. Yeo, P. Ranade, H. Takeuchi, T. J. King, and C. Hu, "Dual-metal gate technology for deep-submicron CMOS transistors," in *Symp. VLSI Tech. Dig.*, 2000, pp. 72–73.

[9] H. Zhong, S.-N. Hong, Y.-S. Suh, H. Lazar, G. Heuss, and V. Misra, "Properties of Ru-Ta alloys as gate electrodes for nMOS and pMOS silicon devices," in *IEDM Tech. Dig.*, 2001, pp. 467–480.

[10] J. Lee, H. Zhong, Y.-S. Suh, G. Heuss, J. Gurganus, B. Chen, and V. Misra, "Tunable work function dual metal gate technology for bulk and nonbulk CMOS," in *IEDM Tech. Dig.*, 2002, pp. 359–362.

[11] B.-Y. Tsui and C.-F. Huang, "Wide range work function modulation of binary alloys for MOSFET application," *IEEE Electron Device Lett.*, vol. 24, no. 3, pp. 153–155, Mar. 2003.

[12] S. H. Bae, W. P. Bai, H. C. Wen, S. Mathew, L. K. Bera, N. Balasubramanian, N. Yamada, M. F. Li, and D. L. Kwong, "Laminated metal gate electrode with tunable work function for advanced CMOS," in *Symp. VLSI Tech. Dig.*, 2004, pp. 188–189.

[13] P. Ranade, Y.-K. Choi, D. Ha, A. Agarwal, M. Ameen, and T. -J. King, "Tunable work function molybdenum gate technology for FDSOI-CMOS," in *IEDM Tech. Dig.*, 2002, pp. 363–366.

[14] K. J. Yang, Y. -C. King, and C. Hu, "Quantum effect in oxide thickness determination from capacitance measurement," in *Symp. VLSI Tech. Dig.*, 1999, pp. 77–78.

[15] R. Beyers, "Thermodynamic considerations in refractory metal-silicon-oxygen systems," *J. Appl. Phys.*, vol. 56, no. 1, pp. 147–152, Jul. 1984.

[16] P. Ranade, Y. C. Yeo, Q. Lu, H. Takeuchi, T.-J. King, and C. Hu, "Molybdenum metal gate MOS technology for post-SiO_2 gate dielectrics," in *Mater. Res. Soc. Symp.*, vol. 611, 2000, pp. C3.2.1–C3.2.6.

[17] V. Misra, G. P. Heuss, and H. Zhong, "Use of metal-oxide- semiconductor capacitors to detect interactions of Hf and Zr gate electrodes with SiO_2 and ZrO_2," *Appl. Phys. Lett.*, vol. 78, no. 26, pp. 4166–4168, Jun. 2001.

[18] A. Yagishita, T. Saito, K. Nakajima, S. Inumiya, Y. Akasaka, Y. Ozawa, G. Minamihaba, H. Yano, K. Hieda, K. Suguro, T. Arikado, and K. Okumura, "High performance metal gate MOSFETs fabricated by CMP for 0.1 μm regime," in *IEDM Tech. Dig.*, 1998, pp. 785–788.
[19] C. D. Gelatt, Jr and H. Ehrenreich, "Charge transfer in alloys: AgAu," *Phys. Rev. B*, vol. 10, no. 2, pp. 398–415, Jul. 1974.
[20] C. Kittel, *Introduction to Solid State Physics*, 7th ed. New York: Wiley, 1996, pp. 151–157.
[21] B. M. Thaddeus, Ed., *Binary Alloy Phase Diagrams*. Metals Park, OH: American Society for Metals, 1990, pp. 2088–2089.

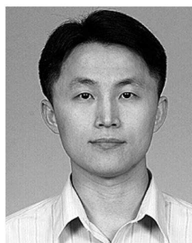
Tzung-Lin Li was born in Pingtung, Taiwan, R.O.C. on April 2, 1976. He received the B.S. degree in electrical engineering from the National Central University, Chungli, Taiwan, in 1998, and the M.S. degree from the Institutes of Electronics, National Chiao-Tung University, Hsinchu, Taiwan, in 2000, where he is currently pursuing the Ph.D. degree.

His current research interest is focused on the study and investigation of metal gate technology for nano devices.

Chia-Hsin Hu was born in Changhua, Taiwan, R.O.C., on July 29, 1980. He received the B.S. degree in electrical engineering from National Cheng-Kung University, Tainan, Taiwan, in 2002, and the M.S. degree from the Institute of Electronics, National Chiao-Tung University, Hsinchu, Taiwan, in 2004.

In 2004, he joined UMC, Hsinchu, Taiwan, where he has been engaged in research and development of RF devices technologies.

Wu-Lin Ho received the B.S. degree in industrial education from National Taiwan University, Taipei, Taiwan, R.O.C., in 1998. He is currently pursuing the M.S. degree with the Institute of Electronics, National Chiao-Tung University, Hsinchu, Taiwan.



Howard C.-H. Wang received the B.S. and Ph.D. degrees in electronic engineering from National Chiao-Tung University, Hsinchu, Taiwan, R.O.C., in 1992 and 2002, respectively.

He joined Taiwan Semiconductor Manufacturing Company (TSMC), Hsinchu in 1995, where he was initially responsible for process integration for 0.5- and 0.6-mm technology. Between 1997 and 2002, he was involved in the device design (front-end integration) for 0.18-, 0.15-, and 0.13-mm, and 90-nm technology in the Advanced Device Technology Department, TSMC. Since 2002 he has been a Manager in Exploratory Technology Development-1 Department under the supervision of Dr. C. Hu. He is currently working on the development of high-κ gate dielectric, metal gate, and strained-Si technologies for sub-65-nm node CMOS device technology. He holds 13 U.S. patents and has authored or coauthored 15 papers in technical journals and conferences.

Dr. Wang is a member of the CMOS Devices subcommittee for International Electron Devices Meeting (IEDM).

Chun-Yen Chang (F'88) received the B.S. degree in electrical engineering from Cheng-Kung University, Tainan, Taiwan, R.O.C., in 1960. He received the M.S. degree from the Institute of Electronics of National Chiao-Tung University (NCTU), Hsinchu, Taiwan, where his thesis was on tunneling in semiconductor/superconductor junctions. His Ph.D. dissertation, entitled "Carrier Transport across metal-semiconductor barrier," was completed in 1969 and published in 1970. It is cited as a pioneering paper in this field.

He has devoted himself to education and academic research more than 40 years. He has contributed to microelectronics and optoelectronics research and development, including the invention of low-pressure MOCVD using tri-ethylgallium method to fabricate LED, laser, and microwave transistors, Zn-incorporation of SiO2 for stabilization of power devices and nitridation of SiO2 for ULSIs, etc. From 1962 to 1963, he was in the military service at NCTU, establishing the first Taiwanese experimental TV transmitter. In 1963, he returned to NCTU to serve as an Instructor, and established the High Vacuum Laboratory. In 1964, he and his colleagues established the Semiconductor Research Center (SRC) at NCTU in April 1965, and the first IC in August 1966. In 1968, he published the first Taiwanese semiconductor paper in an international journal, *Solid State Electronics*. In the same year, he was invited by Prof. L. J. Chu, a Webster Chair Professor, Massachusetts Institute of Technology, to join the NCTU Ph.D. program. In 1969, he became a Full Professor, teaching solid state physics, quantum mechanics, semiconductor devices, and technologies. In 1987, he became Dean of Research (1987 to 1990), Dean of Engineering (1990 to 1994), and Dean of Electrical Engineering and Computer Science (1994 to 1995). Simultaneously, he was the founding President of National Nano Device Laboratories (NDL) from 1990 through 1997. Then he became Director of the Microelectronics and Information Systems Research Center (MIRC) of NCTU (1997 to 1998). In August 1, 1998, he was appointed President of NCTU by the NCTU and MOE search committees. He has supervised more than 300 M.S. and 50 Ph.D. students. They are now founders of most high technology enterprises in Taiwan, namely UMC, TSMC, Winbond, MOSEL, Acer, and Leo, among others.

Dr. Chang is a Member of Academia Sinica and the Foreign Associate of National Academy of Engineering.