Chapter 2

Advanced Transistor Process Technology from 22- to 14-nm Node

Huaxiang Yin and Jiaxin Yao

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.78655

Abstract

Transistor performance meets great technical challenges as the critical dimension (CD) shrinking beyond 32/28-nm nodes. A series of innovated process technologies such as high-k/metal gate, strain engineering, and 3D FinFET to overcome these challenges are reviewed in this chapter. The principle, developing route, and main prosperities of these technologies are systematically described with theoretical analysis and experimental results. Especially, the material choice, film stack design, and process flow integration approach with high-k/metal gate for sub-22-nm node is introduced; the film growth technique, process optimization, and flow integration method with advanced strain engineering are investigated; the architecture design, critical process definition, and integration scheme matching with traditional planar 2D transistor for 14-nm 3D FinFET are summarized.

Keywords: CMOS, high-k/metal gate, strain, FinFET, process

1. Introduction

The metal-oxide-semiconductor field effect transistors (MOSFETs) are core switch devices of current large-scale complementary-metal-oxide-semiconductor integrated circuits (CMOS ICs). The performance of the transistor has a critical effect on the performance of IC. As the continuous scaling of the transistor CD for a higher IC performance and integration density, the fabrication process technologies and methods of the transistor are fast changing and becoming relatively complicated. To suppress the short-channel effect as well as the performance degradation of devices, three main new technologies, including strain engineering, high-k/metal gate (HKMG), and FinFET of MOSFETs, are implemented into state-of-art IC manufacture technology. The three technologies are quite important and firstly applied in the

CMOS IC manufacturing process by Intel Corporation in the years 2003, 2007, and 2011 at 90-, 45-, and 22-nm node, respectively, which is developed into the industrial standards and widely adopted by other IC manufacturing corporations including TSMC, and Samsung. While the process node of CMOS IC scaling from 22- into 14-nm node, advanced technologies such as film growth, structure design, process optimization, and integration flow of them become more complicated, which often need elaborated process development with diversified knowledge and techniques from different fields.

2. Strain engineering

The effective carrier mobility in the channel of the transistor is crucial to the device's performance. With the gate length aggressively shrinking down, the electric field magnitude in the channel is strengthened with channel-doping concentration rising, resulting in obvious degradation in effective carrier mobility due to ionized impurity scattering. Mobility both for electron and for hole can be enhanced by changing the silicon atom arrangement of crystal lattice in the channel through the external stress. It is investigated that the tensile and compressive strain for silicon can improve the electron and hole mobility, respectively.

Mobility is closely dependent on the mean free time and the effective mass of the carrier. As we consider the simplified band structure of silicon, there are six equivalent minima at k = (x, 0, 0), (−x, 0, 0), (0, x, 0), (0, −x, 0), (0, 0, x), (0, 0, −x) with $x = 5\ nm^{-1}$ for the conduction band in Ref. [1]. There is one maximum containing two sub-bands at k = 0 for the valence band. These two sub-bands are referred to as the light and heavy hole bands with a light hole effective mass and a heavy hole effective mass. Therefore, the effective mass of these anisotropic minima is characterized by a longitudinal mass along the corresponding equivalent (1, 0, 0) direction and two transverse masses in the plane perpendicular to the longitudinal direction. For the electron in conduction band, the external stress can cause tensile or compressive strain in the silicon lattice. The longitudinal band valley will change. Thus, the corresponding longitudinal mass is changed leading to the mean free time increasing or decreasing for the carriers. For hole in the valence band, the strain effect on the light hole band and heavy hole band is familiar with electron.

For the device, several of strain types for the mobility enhancement are listed in **Table 1**. Different axial tensile and compressive strain can introduce different mobility enhancement or

Direction	NMOS	PMOS
Channel length	Tensile	Compressive
Channel width	Tensile	Tensile
Perpendicular to channel plane	Compressive	Tensile

Table 1. Strain type for carrier mobility enhancement.

http://dx.doi.org/10.5772/intechopen.78655

degradation for electron and hole. Therefore, strain technique is very practical and important for device's performance promotion. In the point of IC process, this technique is called as strain engineering, which is divided into global strain stress and local strain stress. Global strain stress is less employed in the manufacturing process due to the application regime. Local strain stress can be targeted to enhance carrier's mobility in the specified region and widely used in the modern IC process flow. Local strain stress can be induced and achieved by IC processes, such as selective epitaxy growth (SEG) of silicon-germanium (SiGe) source/drain, dielectric etch-stop layer (ESL), metal gate, and contact. For PMOS, the most important strain engineering technology is to use selective source/drain epitaxy of SiGe with large lattice constant in order to provide channel with axial compressive stress for hole mobility enhancement as shown in **Figure 1**.

The embedded SiGe in source/drain (S/D) region has been widely used to induce uniaxial strain in the channel, especially for the sigma SD epitaxy in Ref. [2]. The strain engineering in 22-nm planar transistor is becoming more complicated. The film growth technique often requires relatively low temperature and the decreasing of pattern dependency. In 22-nm node, the SiGe layers need to be grown at 650°C in a reduced-pressure chemical vapor deposition (RPCVD) reactor and with a series of complicated steps. First, in situ cleaning is performed by annealing in the range of 740–825°C for 3–7 min. Then, dichlorosilane (SiH_2Cl_2), 10% germane (GeH_4) in H_2, and 1% diborane (B_2H_6) in H_2 are used as Si, Ge, and B precursors, respectively. Moreover, HCl is utilized as Si etchant to obtain selectivity during the epitaxy. The SiGe growth rate can be denoted by an empirical model (Eq. (1)) in Refs. [3, 4], which considers the contribution of a variety of molecule fluxes coming from different directions toward Si planes during epitaxy in Ref. [5]:

$$R_{total} = R_{Si}^{V} + R_{Si}^{LG} + R_{Si}^{SO} + R_{Si}^{CO} + R_{Si}^{IP} + R_{Ge}^{V} + R_{Ge}^{LG} + R_{Ge}^{SO} + R_{Ge}^{CO} + R_{Ge}^{IP} - R_{HCl}^{V} - R_{HCl}^{LG} - R_{HCl}^{SO} - R_{HCl}^{CO} + R_{HCl}^{IP}, \quad (1)$$

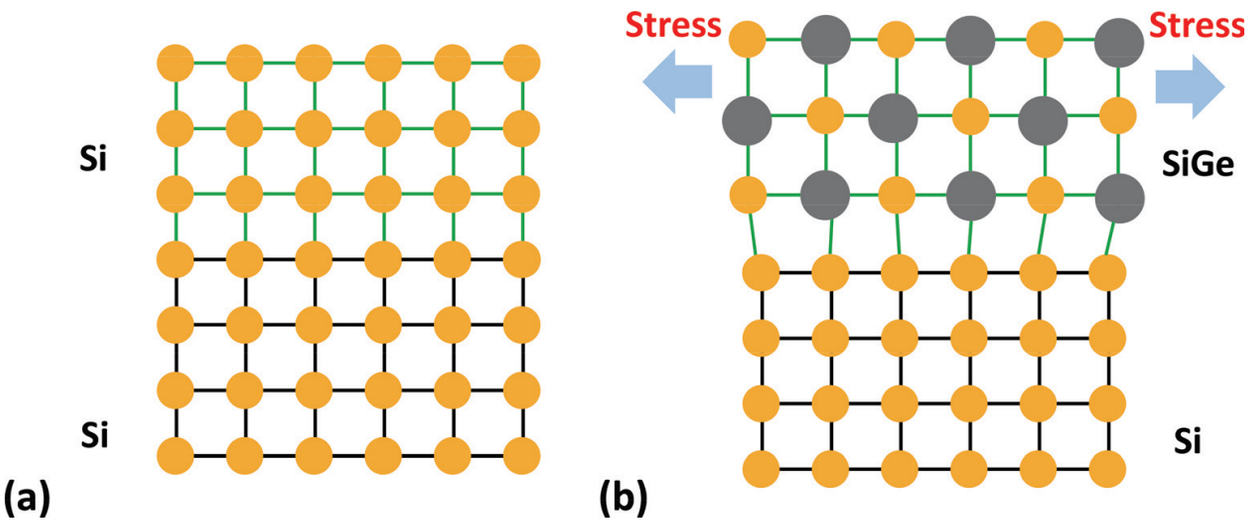

Figure 1. (a) Homoepitaxy and (b) heteroepitaxy: SiGe with high stress growth on the Si substrate.

where R^V and R^{LG} for Si, Ge, and HCl are the contribution of gas molecules in vertical and lateral directions; R^{SO} and R^{CO} are the mobile reactant molecules on the oxide surface surrounding or within a chip; R^{IP} is the contribution from atoms which diffuse from the edges toward the Si plane. After considering the reaction species and atom activation energy, gas partial pressure, and growth temperature, the final expression for the total growth rate is given by (Eq. (2))

$$\begin{aligned} R_{Total} = {} & \beta \frac{\left(1-\theta_{H(Si)}-\theta_{Cl(Si)}\right)}{\mathrm{N}_0} \times \frac{P_{SiH_2Cl_2}}{(2\pi m_{SiH_2Cl_2} k_b T)^{\frac{1}{2}}} \left(\frac{E_{SiH_2Cl_2\ on\ Si}}{k_b T}+1\right) \exp\left(-\frac{E_{SiH_2Cl_2\ on\ Si}}{k_b T}\right) \\ & + \chi \frac{(1+m_r)\left(1-\theta_{H(Si)}-\theta_{Cl(Si)}\right)}{N_0} \times \frac{P_{GeH_4}}{(2\pi m_{GeH_4} k_b T)^{\frac{1}{2}}} \left(\frac{E_{GeH_4\ on\ Si}}{k_b T}+1\right) \exp\left(-\frac{E_{GeH_4 on\ Si}}{k_b T}\right) \\ & + \chi \frac{(1+m_r)\left(1-\theta_{H(Si)}-\theta_{Cl(Si)}\right)}{N_0} \times \frac{\left(BP_{GeH_4} \ln\left(\frac{1}{c}\right)\right)}{(2\pi m_{GeH_4} k_b T)^{\frac{1}{2}}} \left(\frac{E_{GeH_4\ on\ Si}+0.1eV}{k_b T}+1\right) \\ & \times \exp\left(-\frac{E_{GeH_4 on\ Si}+0.1eV}{k_b T}\right) - \frac{\gamma}{N_0} \frac{P_{HCl}^{0.596}}{(2\pi m_{HCl} k_b T)^{\frac{1}{2}}} \left(\frac{E_{Etching}}{k_b T}+1\right) \exp\left(-\frac{E_{Etching}}{k_b T}\right) \end{aligned} \tag{2}$$

where θ_{Cl} and θ_H parameters stand for the occupied dangling bonds by hydrogen and chlorine atoms on Si; N_0 is the number of atoms per unit volume for Si; E and P are activation energy and partial pressure for different reactant molecules, respectively. The variable c is the exposed Si coverage of Si chip where B is a unit-less constant which is dependent on the architecture of the mask. The equation constants, β, χ, and γ, are tooling factors which depend on the temperature distribution and gas kinetic over the susceptor in the CVD reactor. Therefore, a series of process parameters are affecting the growth of SEG SiGe in 22-nm PMOS transistor, resulting in different growth rates, Ge content, film quality as well as the compressive stress to the channel. The stress distribution has also a strong relationship to the growth area, pattern intensity, and locations around the wafer.

For integrating SEG SiGe into 22-nm PMOSFET, a sacrificial epitaxy-block Si3N4 layer is deposited on whole wafer after the formation of dummy polysilicon gate and spacers. In the next step, the block layer is selectively opened and low-temperature epitaxy SiGe with high stress is performed at the source/drain region of PMOS.

The amount of strain induced by SiGe is dependent on the initial recess shape, interfacial quality of SiGe/Si, and defect density in the epilayers. Sigma-shaped recesses with (100) and {111} planes are very suitable shape for embedded SiGe in source/drain regions with the highest stress. Moreover, in such transistors, shorter distance between sigma-shaped recesses and channel region can generate a higher stress to the channel region. By applying dry etch together with wet etch in Si substrate, the sigma-shaped recesses turn more large and induce a stronger stress of embedded SiGe with a closer distance to the channel in Ref. [5].

For NMOS, PMD (pre-metal dielectric) layer is deposited as ESL, which can offer axial tensile stress to the channel for electron mobility enhancement [6]. In the modern IC manufacturing

integrated process, more and more strain process is employed for the devices, which is of great significance to suppress the device's performance degradation. However, the film thickness of ESL is limited due to the scaling of gate pitch between transistors. New techniques, such as metal gate and contact electrode stress of NMOS, are necessary. The TiN metal gate and the W plug often bring effective tensile stress into the channel of NMOSFET, resulting in the enhancement of motilities for electrons.

3. High-k/metal gate

High-k/metal gate (HKMG) is a very important technique for modern CMOS IC manufacturing process. While the transistor CD scaling down, conventional oxide dielectric/polysilicon gate was formally replaced by high-k dielectric/metal gate, in order to suppress the unbearable leakage in the ultra-thin oxide dielectric film in Ref. [7]. HKMG technique has found a new effective path for equivalent oxide thickness (EOT) scaling tendency, which is of deep significance to continuous scaling of MOS transistors. However, HKMG brings about a series of challenges, including new high-k dielectric and metal gate materials, threshold voltage modulation, and process integration scheme. To some extent, the scaling of MOS transistor relies on the scaling of EOT of gate dielectric. When the conventional oxide film thickness shrinks to about 11–12 A, the transistor shrinking cannot be continued due to the extremely large leakage current from the gate to the substrate by the electron direct tunneling through the ultra-thin oxide film. EOT is defined as (Eq. (3)).

$$\text{EOT} = T_{HK} \cdot \frac{\varepsilon_{OX}}{\varepsilon_{HK}} \tag{3}$$

where T_{HK} is the physical thickness of high-k dielectric, ε_{OX} is the SiO_2 dielectric constant, and ε_{HK} is the high-K dielectric constant. When it comes to high-k dielectric materials, the physical thickness of the gate dielectric is increased because of the high value of dielectric constant parameter. Hence, the gate leakage current induced by direct tunneling is reduced dramatically to continue the scaling of EOT.

Many high-k materials have been investigated for CMOS devices including metal oxide (HfO_2, ZrO_2, Al_2O_3, etc.) as shown in **Table 2**. Among these metal-oxide materials, HfO_2 has the advantages of the moderate relative permittivity value, the basically symmetrical energy band offset to silicon conduction band and valence band, and the uniformly amorphous structure. Therefore, HfO_2 material is applied in the IC manufacturing production.

In the early 1990, it is reported that the integration of polysilicon gate with HfO_2 dielectric results in serious Fermi Level Pinning (FLP) phenomenon, where the Fermi level of polysilicon gate is fixed at the poly/HfO_2 interfacial energy level. Although some theories including oxygen vacancy model, and dipole formation, are put forward to explain the pinning effect, the process cannot successfully release the effect of FLP, which causes huge difficulties on the device's threshold voltage modulation. Therefore, different metal gates with the high-k material are corresponded to different threshold voltage modulation regimes for PMOS and NMOS.

Material	Dielectric constant	Material	Dielectric constant
Al_2O_3	8–11.5	$NdAlO_3$	22.5
(Ba, Sr)TiO_3	200–300	$PrAlO_3$	25
$BeAl_2O_4$	8.3–9.43	Si_3N_4	7
CeO_2	16.6–26	$SmAlO_3$	19
HfO_2	26–30	$SrTiO_3$	150–250
Hf silicate	11	Ta_2O_5	25–45
La_2O_3	20.8	TiO_2	86–95
$LaAlO_3$	23.8–27	Y_2O_3	8–11.6
$LaScO_3$	30	ZrO_2	22.2–28

Table 2. High-k dielectric constant.

3.1. HKMG film stack

The introduction of high-k/metal gate provides great potential of transistor's scaling down under 45-nm node. Metal gate can reduce oxide thickness by eliminating polysilicon gate depletion effect. Metal gate has a low gate resistance and can suppress boron penetration to the substrate in Refs. [8–10].

In 22-nm node, the main challenge and research hotpot for HKMG stack lie in the effective work function (EWF) modulation of metal gate. The EWF can be defined as the value between Fermi level of metal gate and vacuum level in the metal-oxide-semiconductor system. As shown in **Figure 2**, EWF is defined as (Eq. (4))

$$\mathrm{EWF} = E_0 - E_{FM}, \tag{4}$$

where E_0 is the vacuum energy level, and E_{FM} is the Fermi level of metal gate. When E_{FM} is close to the conduction band edge E_C or valence band edge E_V of Si substrate, EWF will get the minima or the maximum value for the MOS device. Generally, the mid-gap EWF is around 4.6 eV, and the band edge EWF is less than 4.4 for conduct band in NMOS or is high than 4.8 eV for valence band in PMOS.

Fermi level of metal gate can be shifted both upwards and downwards. The Fermi level (E_F) of metal gate is set to be in the position of mid-gap in the substrate. When Fermi level shifts to the conduction band of Si substrate, the effective work function of metal gate decreases. On the other hand, when Fermi level shifts to the valence band of Si substrate, the effective work function of metal gate increases. The EWF movement behaviors can directly drive the threshold voltage (Vt) modulation for the MOS devices.

The most effective method to manipulate EWF is the selection of metal gate. For PMOS, a large EWF is preferred to achieve high Vt for low-power (LP) IC performance. Hence, Fermi level of PMOS metal gate is ideally near to the valence band maximum of silicon substrate, where the position in the valence band minima of silicon is the best choice for Fermi level of metal gate.

http://dx.doi.org/10.5772/intechopen.78655

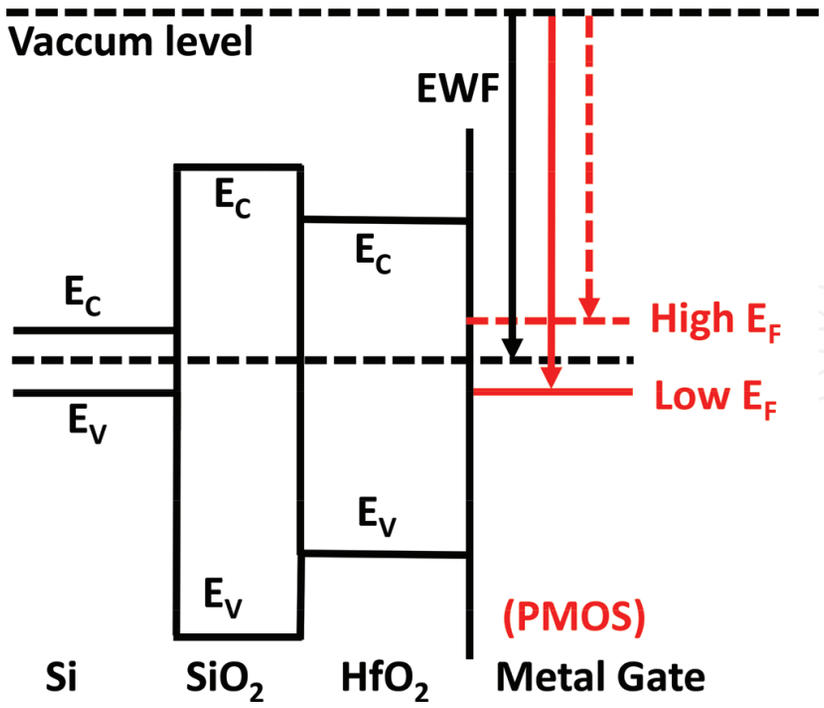

Figure 2. Illustration of band edge EWF of metal gate for PMOS.

For NMOS, a small EWF is preferred to achieve high Vt for LP IC performance, where the Fermi level of NMOS metal gate is ideally near to or at the conduction band minima of silicon substrate. Therefore, the demand of the large EWF for PMOS and small EWF for NMOS is selecting TiN with a high work function and TiAl with a low work function metals.

For the multi-Vt modulation of PMOS and NMOS, the most general method is to tune the gate-stack thickness control in Refs. [11, 12], in order to realize the regular, low, and high Vt levels. As shown in **Figure 3**, the metal gate stack can be divided into three layers: the first is the bottom-capping layer for the high-k dielectric, the second is exactly the work function layer, and the last is the top-capping layer for the contacted metal. Moreover, the etch-stop layer should be considered for the dual work function metal integration of PMOS and NMOS. Although the gate stack contains three parts, the effective work function of the entire gate electrode is dominated by the work function layer metal, where EWF sensitivity is strictly limited by the bottom-capping layer thickness, and the top-capping layer acts as the barrier layer for the contacted metal (tungsten). Therefore, the thickness control of metal gate-stack design is exactly of precision and significance.

3.2. Gate-first and gate-last integration scheme

The novel gate-stack structure of HKMG has been implemented for MOSFETs to promise conventional scaling of the high-performance CMOS process down to the 45/32-nm node. Two completely different integration schemes were proposed [13]. With the HKMG in the IC process flow, a big question arises that the module of HKMG structure formation is ahead of or after the module of source/drain process. Gate-first process integration scheme is familiar with

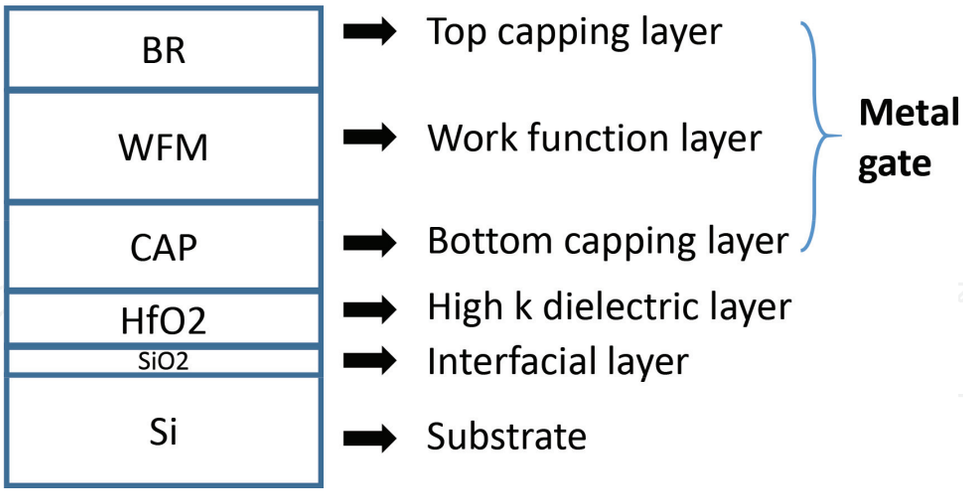

Figure 3. HKMG gate stack.

poly-Si/SiO2 process flow. HKMG module is firstly deposited after the active-region formation module, and then source/drain module formation module is following until the end. However, with the source/drain formation later than HKMG formation module, the high annealing temperature for the S/D doping profile has a serious impact on HKMG characteristics and its reliability.

To overcome shortcoming caused by gate-first integration scheme, gate-last integration scheme is put forward. In the gate-last process, conventional poly-Si/SiO2 is still formed on the wafer substrate firstly. After poly-Si/SiO2 formation module, it is followed by the S/D impurity doping and its activation with annealing process at high temperature ambient. Then, PMD layer is deposited on the poly-Si dummy gate, where PMD is also called an inter-layer-dielectric zero layer (ILD0). With poly-Si-open planarization (POP) chemical mechanical polishing (CMP), poly-Si gate is exposed for the following removal of poly-Si/SiO_2 process. Finally, HKMG is deposited in the position where poly-Si dummy gate previously existed, which is called gate-last process due to HKMG module later than the middle end of line (MEOL) process. The implement of gate-last integration scheme avoids the damage to devices by the high annealing temperature of S/D process. Therefore, gate-last integration scheme has obvious performance advantages for HKMG devices and becomes popular technique applied beyond 28 nodes.

In gate-last technique, it is divided into two integration schemes: high-k first/metal-gate last and high-k last/metal-gate last. In the first approach, the high-k layer is deposited together with the formation of dummy gate and before the annealing of source/drain, where only metal gate stack is formed with gate-last scheme. In the second approach, both high-k and metal gate are formed after the annealing of source/drain, which is also called all gate-last integration scheme. It has better film quality and process adjustment window than the former and is widely adopted for CMOS IC fabrication process in 22 nm and beyond node. In this integration scheme, multilayer HKMG stacks are IL/HfO_2/TiN/TaN/TiN/W and IL/HfO_2/TiN/TiAlC/TiN/W for PMOS and NMOS, respectively. IL layer is an interfacial layer between HK and substrate and is normally SiO_2 forming by chemical oxidation method. All HKMG depositions

are finished by atomic layer deposition (ALD) approach with a high conformality and a precise thickness control ability.

4. FinFET technology

While process node scaling from 22 to 14 nm, the basic architecture of the transistor is changing from 2D planar device to 3D volume inversion device for a better control of SCE in channel with less leakage. The device design as well as the process techniques turns more complicated and needs a more elaborated technologies.

4.1. FinFET transistors

With feature size of CMOS IC shrinking to 20-nm node and beyond, the structure of the conventional planar MOSFET consisting of single-gate electrode to control channel potential distribution and the flow of current in the channel region is faced with the undesirable parasitic effects called short-channel effect (SCE) and drain-induced barrier lowering (DIBL) effect. Via voltage-doping transformation (VDT) model [14], the device's structure and material parameters can be translated into electrical parameters with electrostatic integrity (EI) (Eq. (5)). SCE and DIBL can be derived as (Eqs. (6) and (7))

$$\mathrm{EI} = \left[1 + \frac{{x_j}^2}{{L_{ch}}^2}\right] \frac{t_{ox}}{L_{ch}} \frac{t_{dep}}{L_{ch}} \tag{5}$$

$$\mathrm{SCE} = 0.64 \frac{\varepsilon_{Si}}{\varepsilon_{ox}} EI\, V_{bi}, \tag{6}$$

$$\mathrm{DIBL} = 0.8 \frac{\varepsilon_{Si}}{\varepsilon_{ox}} EI\, V_{ds} \tag{7}$$

where L_{ch} is the effective channel length, V_{bi} is the source or drain built-in potential, t_{ox} is the gate oxide thickness, x_j is the source/drain junction depth, and t_{dep} is the penetration depth of the gate electric field in the channel region. The parameter EI is denoted as electrostatic integrity factor.

The threshold voltage of MOSFET can be denoted as (Eq. (8))

$$V_T = V_{T\_long} - SCE - DIBL \tag{8}$$

According to the above expression, SCE can be minimized by reducing the junction depth, gate oxide thickness, and depletion depth via increasing the doping concentration in the channel region. However, the limits on the reducing junction depth and gate oxide thickness have become very toughly serious in the practical device. Hence, SCE and DIBL values of the planar MOSFET are not controlled well in the ultra-short-channel length.

The most efficient and direct way to suppress SCE is to strengthen the gate electric field control capability by double-gate (DG) or multi-gate (MG) structure. DG or MG structures on thin Si channel improve the electrostatic integrity of MOSFET (Eq. (9)) with the transistor working in a volume inversion mode due to a reduced device structure parameter, which decreases the SCE and DIBL effects on the device electric parameters, such as threshold voltage, sub-threshold slope (SS), and DIBL voltage. In the equation, since the thickness of Si is much smaller than that of depletion region in planar transistor, EI is obviously improved. The whole new structures of MOSFET extend the shrinking boundary of the ultra-short gate length

$$\mathrm{EI} = \frac{1}{2}\left[1 + \frac{t_{Si}^2/4}{{L_{ch}}^2}\right]\frac{t_{ox}}{L_{ch}}\frac{t_{Si}/2}{L_{ch}} \tag{9}$$

FinFET is a typical double-gate or multi-gate device with a three-dimensional channel structure, as shown in **Figure 4**. The FinFET is made of a tall and narrow silicon island. The 3D channel is standing above the silicon substrate, where the ultra-thin silicon body is familiar with the fin of the fish. The fin channel under the gate can be fully depleted by electrostatic potential, providing a strong ability of controlling the carriers' behaviors in the channel. FinFET can really expand the limit of the shrinking size and is widely adopted for the 16/14-nm technology node and beyond. FinFET can effectively suppress the leakage of the sub-surface channel, which can obviously reduce the off-state current for the device's current-voltage transfer characteristic. In the meantime, the fully depleted channel can obtain benefit of carriers' mobility with less scattering. For the 3D fin structure, the transistor's width can be doubled compared to the planar one in the projected plane, which can improve the driving current at on-state in the saturation regime. With the same drive current, the supply voltage of FinFET can be significantly reduced regardless of the planar transistor's power limit, where the suppression of power consumption in modern integrated circuits emphasizes energy efficiency ratio.

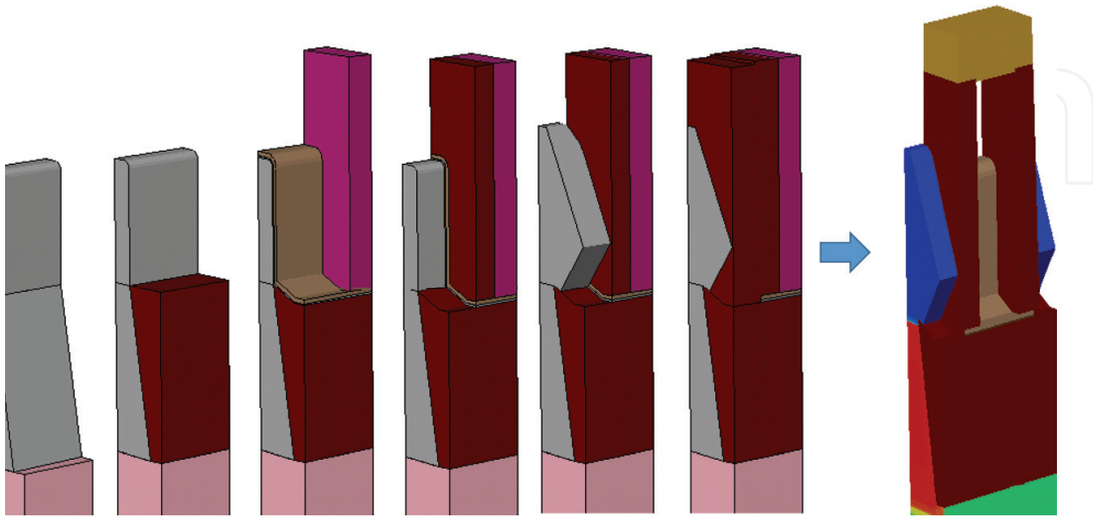

Figure 4. FinFET from fin to whole device.

4.2. FinFET integration process

Since 22-nm technology node, FinFET has been utilized for several process nodes [15–17]. It is firstly introduced by Intel in 22-nm node and widely adopted by different companies in 16- or 14-nm process node. The process integration scheme of FinFET is compatible with that of the planar transistor. In a general way, the critical fabrication steps of FinFET transistor include silicon fin formation on the substrate by the spacer-transfer lithography (STL), shallow trench isolation (STI) formation and recess, 3D dummy gate formation and planarization, 3D spacer formation, source/drain with 3D selective SEG, 3D HKMG formation, and back-end-of-line (BEOL) metallization and contact techniques. It added a little extra process steps than those of planar transistor fabrication. It is very meaningful to understand the integration process of FinFET. In future, the next-generation devices, such as gate-all-around nanowire transistor or nanosheet FET, are still dependent on current FinFET integration flow [18, 19].

4.2.1. Spacer-transfer lithography for bulk fin formation

Oxide by plasma-enhanced CVD (PECVD), poly-Si by low-pressure CVD (LPCVD), and SiNx by PECVD are sequentially deposited in the substrate for the formation of etch-hard-mask (EHM). After etching EHM with pattern, another SiNx is deposited as the spacer of the core layer of oxide/poly-Si/SiNx structure. After spacer and Si dry etch, the 3D Si fin is formed and the Si fin width depends on the SiN spacer thickness, as shown in **Figure 5**. The fin width may be beyond the lithography resolution limit and often smaller than 10 nm.

4.2.2. STI formation and recess

For adjacent fins isolation, high-aspect-ratio-process (HARP) oxide deposition is widely used with a good step coverage on 3D fins. The oxide for HARP STI is deposited by sub-atmospheric CVD (SACVD) with the reaction by tetraethoxysilane (TEOS) precursor and O3. After the isolation oxide annealing, chemical mechanical polishing is utilized for the planarization of deposited dielectric on 3D fins. In following steps, the oxide is precisely etched back and making the fin final formation with shallow trench isolation structures as shown in **Figure 6**.

4.2.3. 3D dummy gate formation

On 3D fins with STI, thin oxide is firstly formed on the surface. Then, amorphous-Si (α-Si) is deposited as dummy gate on the fin. However, the dummy gate etch is the most challenging, for which the top dummy gate needs to be protected during the etching and sidewall and the foot of the dummy gate needs strong etching capability to prevent the residue of Si and no process damage on the exposed fin tip (**Figure 7**).

4.2.4. Source/drain 3D SEG

On 3D fin, it often needs SEG on source/drain regions for less contact resistance. Source/drain selective epitaxy growth normally employs SiH2Cl2, GeH4, and HCl gases. Especially, for PMOS source/drain, B2H6 is mixed into the carrier gas of the reaction. The selectivity of SiGe

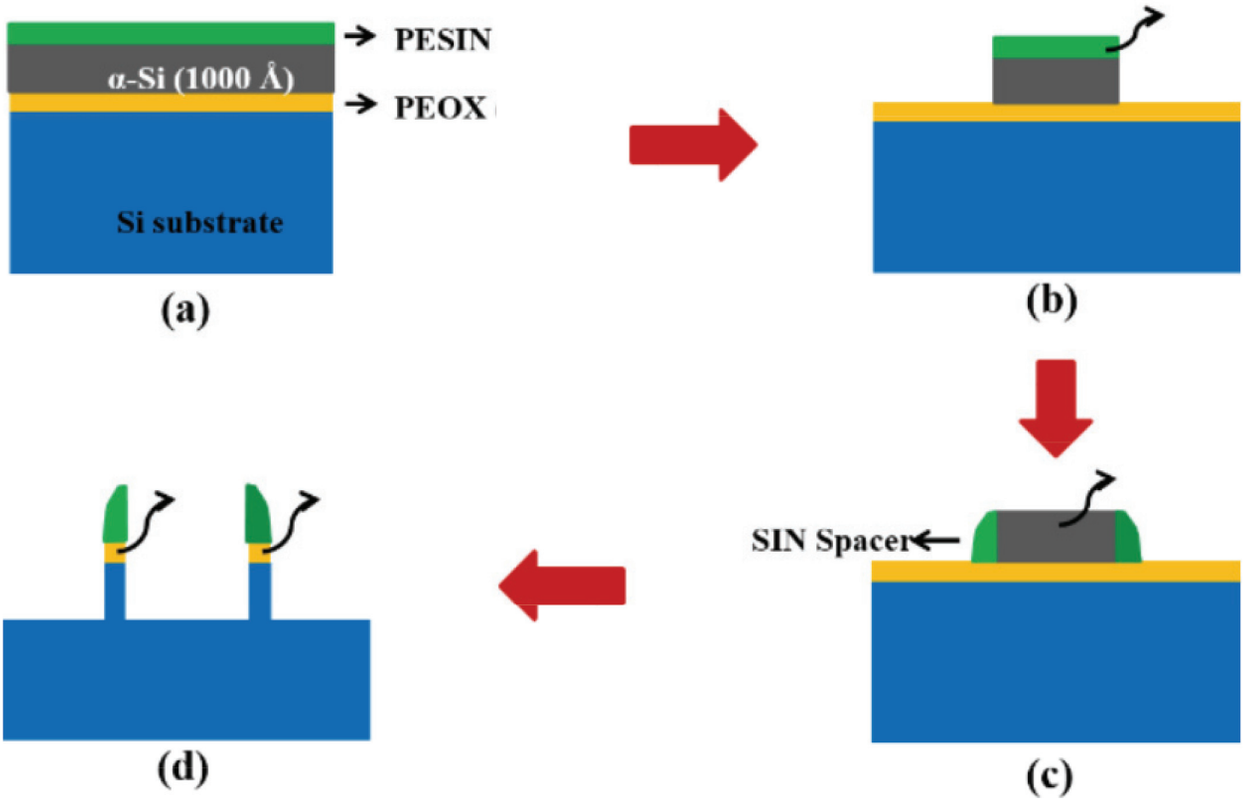

Figure 5. STL for bulk fin formation (a) Hard mask deposition; (b) Hard mark etch; (c) SiN spacer deposition and etch; (d) Fin structure etch.

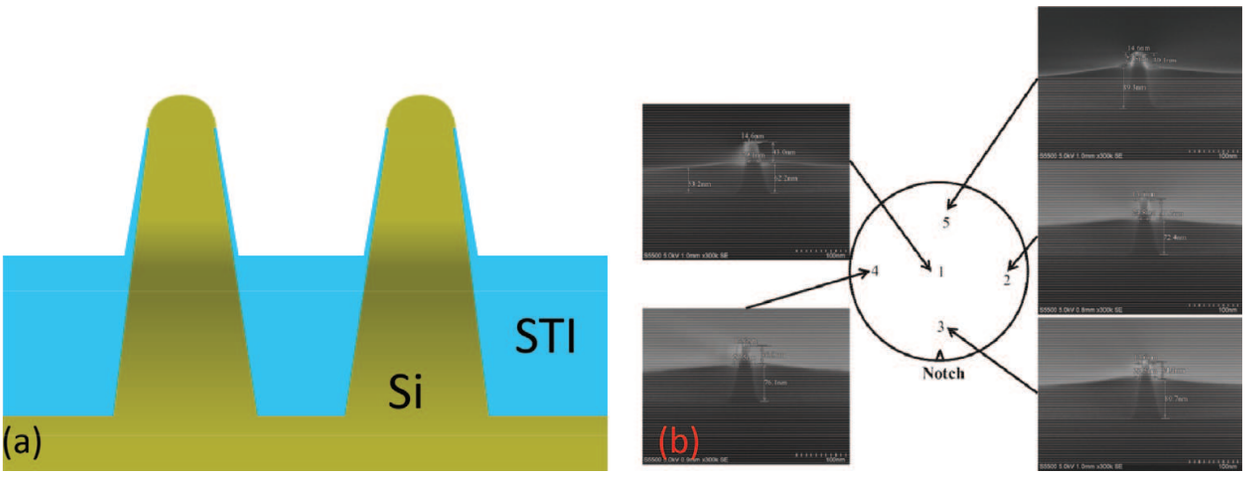

Figure 6. STI formation and recess on 3D fins (a) Fin and STI structure after recessng; (b) SEM images for Fin and STI after recessing.

epitaxy is mainly due to the function of HCl gas, where the etch rate of polycrystalline SiGe is higher than that of single crystalline of SiGe by HCl. In the whole process, the dilution protective gas contains N_2 or H_2 all the time. Due to the slowest growth rate on Si (111) lattice plane, as shown in **Figure 8**, the final formed SiGe shape on 3D fin is more like a diamond. The film stress not only depends on the process conditions but also is strongly affected by the surface quality of fins.

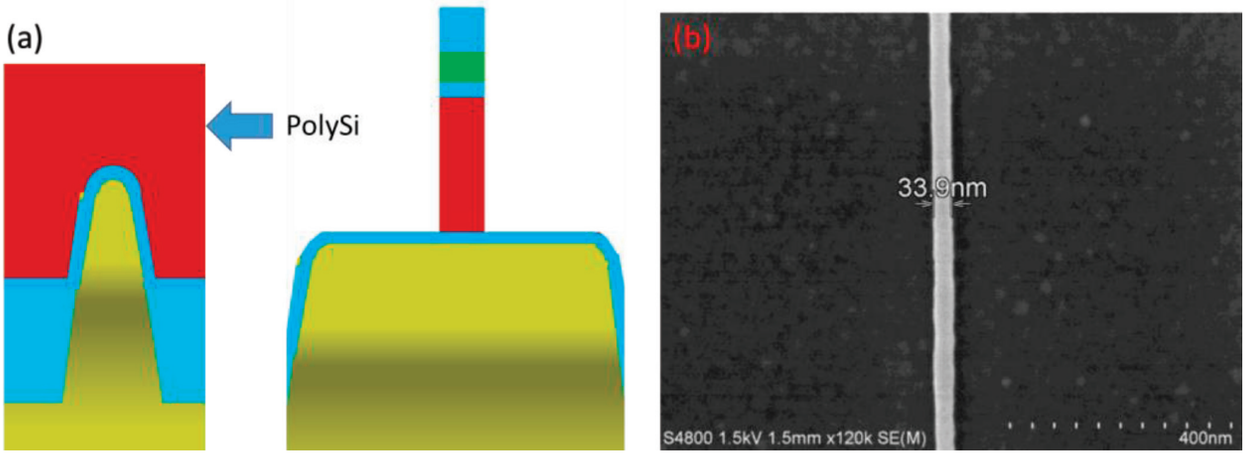

Figure 7. 3D dummy gate formation (a) PolySi deposition, planarization and dummy gate etch; (b) SEM image after dummy gate formation.

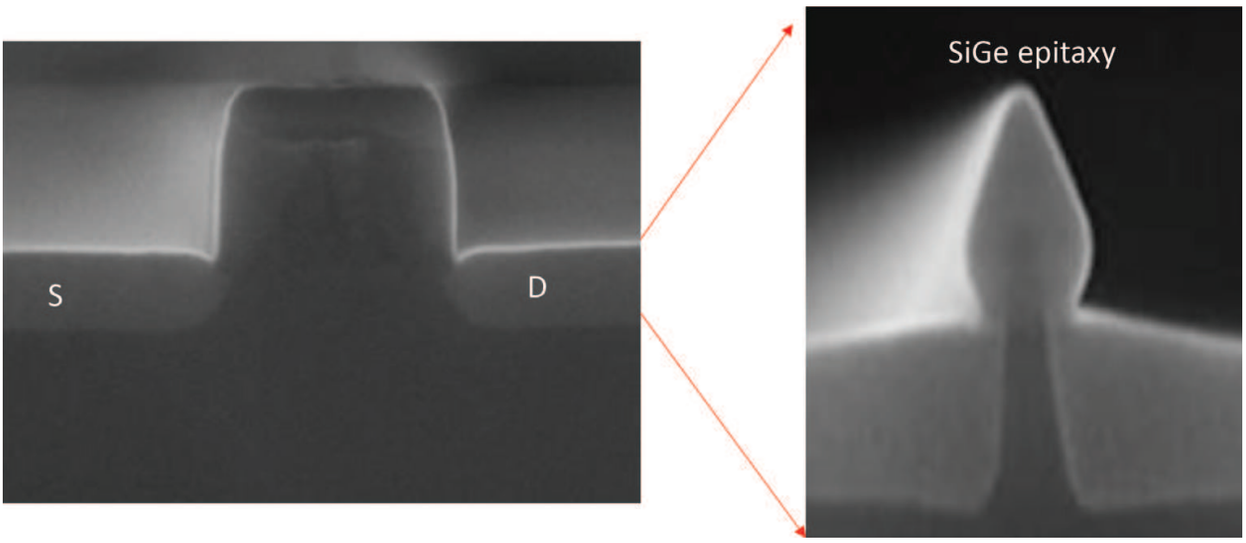

Figure 8. SEG SiGe on 3D fin.

5. Conclusions

Advanced transistor technologies were extensively implemented into the CMOS IC manufacture with the process node scaling from 22 to 14 nm. They require new materials and novel structures as well as complicated process techniques and different device integration flow. This chapter presented a summary on the three important techniques, strain engineering, high-k/metal gate, and FinFET. Both the process theory related to the suppress on SCE for device's shrinking and the detailed illustration on material choice, film growth method, architecture design, critical process definition, and integration are presented in a comprehensive and systematic manner. The process condition optimizations for suppressing stress release are key technologies of strain engineering. The high-k/metal gate needs multilayer structure for

modulating Vt in a different manner for PMOS and NMOS, respectively. The integration scheme is also changed from gate-first to all-last integration. FinFET requires a sophisticated device integration structure and a flow design with less extra process cost. It also has some new fabrication techniques, such as ultra-thin fin formation with STL and improved process methods, including HKMG and SiGe SEG in 3D approach.

Acknowledgements

The reported work was supported by "16/14nm Basic Technology Research" of national 02 IC R&D projects in China (No. 2013ZX02303). The authors would like to thank all the colleagues in Integrated Circuit Advanced Process Center, IMECAS, for their kind and great support.

Conflict of interest

The authors declare that they have no competing interests.

Author details

Huaxiang Yin* and Jiaxin Yao

*Address all correspondence to: yinhuaxiang@ime.ac.cn

Key Laboratory of Microelectronics Devices and Integrated Technology, Institute of Microelectronics of Chinese Academy of Science, Beijing, China

References

[1] Vasileska D, Goodnick SM. Computational electronics. Synthesis Lectures on Computational Electromagnetics. 1st ed. San Rafael, CA, USA: Morgan & Claypool Publishers; 2006. 20 p. 1-216. DOI: 10.2200/s00026ed1v01y200605cem006

[2] Qin C, Yin H, Wang G. Study of sigma-shaped source/drain recesses for embedded-SiGe pMOSFETs. Microelectronic Engineering. 2017;**181**:22-28. DOI: 10.1016/j.mee.2017.07.001

[3] Wang G, Abedin A, Moeen M. Integration of highly-strained SiGe materials in 14nm and beyond nodes FinFET technology. Solid-State Electronics. 2015;**103**:222-228. DOI: 10.1016/j.sse.2014.07.008

[4] Radamson HH, Kolahdouz M. Selective epitaxy growth of Si1−xGex layers for MOSFETs and FinFETs. Journal of Materials Science: Materials in Electronics. 2015;**26**(7):4584-4603. DOI: 10.1007/s10854-015-3123-z

[5] Qin C, Wang G, Kolahdouz M. Impact of pattern dependency of SiGe layers grown selectively in source/drain on the performance of 14nm node FinFETs. Solid-State Electronics. 2016;**124**:10-15. DOI: 10.1016/j.sse.2016.07.024

[6] Yin H, Meng L, Yang T. CMP-less planarization technology with SOG/LTO etchback for low-cost high-k/metal gate-last integration. ECS Journal of Solid State Science and Technology. 2013;**2**(6):268-270. DOI: 10.1149/2.011306jss

[7] Auth C, Cappellani A, Chun J-S. 45nm high-k + metal gate strain-enhanced transistors. In: Symposium on VLSI Technology (VLSI '08); 17-19 June 2008; Honolulu. New York: IEEE; 2008. pp. 128-129

[8] Robertson J. High dielectric constant gate oxides for metal oxide Si transistors. Reports on Progress in Physics. 2006;**69**(2):327-396. DOI: 10.1088/0034-4885/69/2/r02

[9] Wilk GD, Wallace RM, Anthony JM. High-κ gate dielectrics: Current status and materials properties considerations. Journal of Applied Physics. 2001;**89**(10):5243-5275. DOI: 10.1063/1.1361065

[10] Choi J, Mao Y, Chang J. Development of hafnium based high-k materials—A review. Materials Science and Engineering: R: Reports. 2011;**72**(6):97-136. DOI: 10.1016/j.mser.2010.12.001

[11] Ma X, Yang H, Wang W. An effective work-function tuning method of nMOSCAP with high-k/metal gate by TiN/TaN double-layer stack thickness. Journal of Semiconductors. 2014;**35**(9):096001-096004

[12] Xu J, Wang A, He J. 14nm metal gate film stack development and challenges. In: China Semiconductor Technology International Conference (CSTIC '2017); 12-13 March 2017; Shanghai. New York: IEEE; 2017. p. 1-3

[13] Veloso A, Ragnarsson L-A, Cho M-J. Gate-last vs. gate-first technology for aggressively scaled EOT logic/RF CMOS. In: Symposium on VLSI Technology (VLSI '11); 14-16 June 2011; Honolulu. New York: IEEE; 2011. pp. 34-35

[14] Skotnicki T, Merckel G, Pedron T. The voltage-doping transformation: A new approach to the modeling of MOSFET short-channel effects. IEEE Electron Device Letters. 1988;**9**(3):109-112. DOI: 10.1109/55.2058

[15] Auth C, Allen C, Blatter A. A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors. In: Symposium on VLSI Technology (VLSI '12); 12–14 June 2012; Honolulu. New York: IEEE; 2012. pp. 131-132

[16] Natarajan S, Agostinelli M, Bost M. A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588um2 SRAM cell size. In: IEEE International Electron Devices Meeting (IEDM '14); 15-17 Dec. 2014; San Francisco. New York: IEEE; 2014. pp. 3.7.1-3.7.3

[17] Auth C, Aliyarukunju A, Asoro M. A 10nm high performance and low-power CMOS technology featuring 3rd generation FinFET transistors, Self-Aligned Quad Patterning, contact over active gate and cobalt local interconnects. In: IEEE International Electron Devices Meeting (IEDM '17); 2-6 Dec. 2017; San Francisco. New York: IEEE; 2017. pp. 29.1.1-29.1.4

[18] Loubet N, Hook T, Montanini P. Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET. In: Symposium on VLSI Technology (VLSI '17); 5-8 June 2017; Kyoto, Japan. New York: IEEE; 2017. pp. T230-T231

[19] Zhang Q, Yin H, Meng L. Novel GAA Si nanowire p-MOSFETs with excellent short-channel effect immunity via an advanced forming process. IEEE Electron Device Letters. 2018;**39**(4):464-467. DOI: 10.1109/LED.2018.2807389