SiGe/Si Heterostructures

C. W. Liu
National Taiwan University, Taipei, Taiwan, ROC

L. J. Chen
National Tsing Hua University, Hsinchu, Taiwan, ROC

CONTENTS
1. Introduction
2. Strain Effect
3. Device Performances
4. Fabrication Technologies
   Glossary
   Acknowledgments
   References

1. INTRODUCTION

The Si/SiGe:C heterostructures add a rich variety to the monotonic Si world. The misfit between Si and Ge, which is the initial obstacle to fabricating such material systems, becomes an advantage in tailoring the electronic properties of Si-based devices. Si/SiGe heterojunction bipolar transistors have a cutoff frequency of 210 GHz [1] and a maximum oscillation frequency of 285 GHz [2] with carbon incorporation in the base. The figures of merit of metal-oxide-silicon-field-effect transistors (MOSFETs) are also enhanced to some extent with the use of the strained Si on the relaxed SiGe buffers for both N and P channel devices [3, 4]. The reported peak hole mobility is 2700 cm²/Vs in the strained SiGe buffers for both N and P channel devices [3, 4]. The reported peak hole mobility is 2700 cm²/Vs in the strained SiGe buffers for both N and P channel devices [3, 4]. The reported peak hole mobility is 2700 cm²/Vs in the strained SiGe buffers for both N and P channel devices [3, 4]. The reported peak hole mobility is 2700 cm²/Vs in the strained SiGe buffers for both N and P channel devices [3, 4]. The reported peak hole mobility is 2700 cm²/Vs in the strained SiGe buffers for both N and P channel devices [3, 4].

2. STRAIN EFFECT

Si and Ge have the same diamond lattice structures and are completely miscible for a full range of composition, but have different lattice constants. The misfit is defined as the lattice constant difference between Si and Ge with respect to the Si lattice constant and is around 4%. For a linear assumption, the misfit between Si and Ge is ~0.04x. For example, the misfit between Si and Ge is ~0.04x. For example, the misfit between Si and Ge is ~0.04x.

2.1. Critical Thickness

Since the Si substrate is much thicker (~0.7 mm for 200-mm wafers), the original cubic Si−xGe−x lattice with a lattice constant larger than that of Si has to be distorted into the...
Table 1. The achieved performance of electronic and optoelectronic devices based on SiGe technology.

<table>
<thead>
<tr>
<th>Device</th>
<th>Figure of merit</th>
<th>Growth technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strained NMOSFET</td>
<td>Peak effective electron mobility 800 cm²/Vs [6]</td>
<td>Relaxed buffer</td>
</tr>
<tr>
<td>Strained PMOSFET</td>
<td>Peak effective hole mobility 2700 cm²/Vs [5]</td>
<td>Relaxed buffer</td>
</tr>
<tr>
<td>PNP-HBT</td>
<td>$f_T = 210$ GHz [1], $f_{max} = 285$ GHz [2]</td>
<td>Pseudomorphic</td>
</tr>
<tr>
<td>PNP-HBT</td>
<td>$f_T = 59$ GHz at RT, $f_T = 61$ GHz at 85 K [12]</td>
<td>Pseudomorphic</td>
</tr>
<tr>
<td>Heterojunction phototransistors</td>
<td>1.47 A/W at 850 nm, bandwidth = 1.25 GHz [13]</td>
<td>Pseudomorphic</td>
</tr>
<tr>
<td>NMOSFET</td>
<td>$G_m = 460$ ms/mm, $f_T = 76$ GHz, $f_{max} = 107$ GHz at RT [14], $G_m = 730$ ms/mm, $f_T = 105$ GHz, $f_{max} = 170$ GHz at 50 K</td>
<td>Relaxed buffer</td>
</tr>
<tr>
<td>PMOSFET</td>
<td>$G_m = 300$ ms/mm, $f_T = 70$ GHz, $f_{max} = 135$ GHz at RT [15]</td>
<td>Relaxed buffer</td>
</tr>
<tr>
<td>NKTD</td>
<td>$P/V \geq 7.6$ at RT [7], 2 at 4.2 K [8]</td>
<td>Relaxed buffer</td>
</tr>
<tr>
<td>PRTD</td>
<td>$P/V = 1.8$ at RT, 2.2 at 4.2 K [9]</td>
<td>Pseudomorphic</td>
</tr>
<tr>
<td>Detector (IR)</td>
<td>$1.3 \ \mu m, 1.5 \ \mu m$ [16]</td>
<td>Quantum dots</td>
</tr>
<tr>
<td>Detector (LWIR)</td>
<td>Schottky barrier: $\lambda_c = 10 \ \mu m$ [17]</td>
<td>Pseudomorphic</td>
</tr>
<tr>
<td>LED</td>
<td>$1.3 \ \mu m, 1.5 \ \mu m$ at RT [18, 19]</td>
<td>Quantum dots</td>
</tr>
</tbody>
</table>

The last column indicates the growth techniques used to obtain device-quality material.

The in-plane lattice constant of Si$_{1-x}$Ge$_x$ is reduced to that of Si, and the vertical lattice constant increases by 80% of the in-plane lattice constant reduction. Si$_{1-x}$Ge$_x$ under stress may yield to form dislocations if the compressive strain energy is high enough. Strained Si$_{1-x}$Ge$_x$ with dislocations grown on Si is shown schematically in Figure 1. The thickness of strained Si$_{1-x}$Ge$_x$ has to be thin enough to prevent the dislocation formation. The maximum thickness to have a defect-free Si$_{1-x}$Ge$_x$ epilayer is the so-called critical thickness. To minimize the sum of the strain energy and dislocation energy, the thermal equilibrium critical thickness ($h_c$) [20] can be derived as

$$x = 0.55 \ln(10h_c)/h_c$$

where $h_c$ is in nanometers. For example, the critical thickness is $\sim 13$ nm and $\sim 1.5$ nm for strained Si$_{0.8}$Ge$_{0.2}$ and Ge on Si, respectively. This value is too small to be used in many applications. The meta-stable growth away from thermal equilibrium is desired to increase the critical thickness. The selective epitaxial growth [21] and the low temperature epi-growth are common meta-stable growth methods. The enhancement by a factor of at least 4 can be obtained for $x$ less than $\sim 0.5$, but it really depends on the details of the growth conditions. Generally speaking, the smaller area of growth and the lower temperature of growth can have a larger meta-stable critical thickness. Most Si$_{1-x}$Ge$_x$ reactors have a growth temperature of 550–650 °C. However, metastable films will generate the defect if the thermal budget of the further process is too high. This is a challenge in the integration of Si$_{1-x}$Ge$_x$ into a CMOS process.

In ultra-large scale integrated circuits (ULSIs), the building devices are N-channel or P-channel MOSFETs. Strained channels can improve the device performance of MOSFET. Si$_{1-x}$Ge$_x$ grown on Si can be relaxed if it is much thicker than the critical thickness and/or is annealed at high temperature. Relaxed Si$_{1-x}$Ge$_x$ is not good enough for devices with vertical carrier transportation such as HBTs, since both the threading dislocations and misfit dislocations at the
Si/Si<sub>x−y</sub>Ge<sub>y</sub> interface can act as recombination centers. The Si or Ge grown on relaxed Si<sub>y</sub>Ge<sub>x</sub> is under tensile and compressive strain, respectively. The defects in the strained Si or Ge layer are the threading dislocation as shown in Figure 2, and most misfit network exists in the relaxed buffer layers. The threading dislocations in the strained Si and Ge channels have a small effect on in-plane carrier transport, especially for small feature devices, if the density is low enough (below 10<sup>5</sup> cm<sup>−2</sup>). The strained Si and Ge channels increase the carrier mobility, and thus the device speed is enhanced in the same MOSFET technology node. For the increase the carrier mobility, and thus the device speed is enhanced in the same MOSFET technology node. For the Si or Ge grown on relaxed Si<sub>x</sub>Ge<sub>y</sub> layers, the same critical thickness constraint is also applied. Note that the relaxed Si<sub>x−y</sub>Ge<sub>y</sub> can be obtained with graded relaxed buffers [22], silicon on insulators (SOI) [23], and regrowth after smart cut [24]. Equation (1) can be used to estimate the critical thickness of strained Si or Ge, but x is the Ge fraction difference between strained Si (Ge) and relaxed Si<sub>x</sub>Ge<sub>y</sub> buffers.

Because of the small lattice constant of diamond (made of carbon), the incorporation of carbon into SiGe on Si can compensate for the compressive strain from the large Ge atoms. A strain-free SiGeC film can be obtained if the ratio of the Ge fraction to the C fraction is ~11 (compensation ratio), obtained from the linear interpolation of the lattice constant between Ge and silicon carbide. Note that the silicon carbide is a compound with the same number of Si and C atoms in the face center cubic sublattices. However, only up to a few percent (~2.5%) of carbon can be incorporated into SiGe or Si [25, 26]. Note that Si and Ge can be mixed at any concentration.

Si<sub>x</sub>Ge<sub>y</sub>C<sub>z</sub> and Si<sub>x</sub>Ge<sub>y</sub>C<sub>z</sub>, alloys with tensile strain can be fabricated by incorporating carbon into Si and incorporating carbon content above the compensation ratio into Si<sub>x</sub>Ge<sub>y</sub>, respectively. In principle, the critical thickness concepts can still be applied to alloys with tensile strain. However, relaxation mechanisms in addition to the misfit dislocations, such as silicon carbide precipitation, and stacking fault formation, can also occur for thick and high-temperature annealed Si<sub>x</sub>Ge<sub>y</sub>C<sub>z</sub> and Si<sub>x</sub>Ge<sub>y</sub>C<sub>z</sub> films.

2.2. Growth Mode, Self-assembled Quantum Dots, and SiO<sub>2</sub> Dots

For Si<sub>x</sub>Ge<sub>y</sub> layers thicker than the critical thickness, the relaxation is mainly apportioned by misfit dislocation generation at low Ge fraction (< ~0.4), although the morphology roughening can also lower the elastic energy [27]. For the high Ge fraction, the roughening is the dominant mechanism for relaxation, since the roughening barrier is sufficiently small in energy for most growth conditions. The roughening can be represented by the Stranski-Krastanow growth model. After the two-dimensional layer-by-layer growth reaches some critical thickness, the growth becomes three-dimensional island growth. Note that the critical thickness here can differ from the thermal equilibrium critical thickness (Section 2.1). The two-dimensional layer is also called the “wetting layer,” and the islands are called “dots” or “quantum dots,” depending on the size. The three-dimensional islands remain dislocation-free, but the top layers of the islands start to relax. This relaxation can stop the build-up of strain energy in the Ge islands. Further growth can cause the islands to coalesce, and misfit dislocations can occur.

Because of the different strain conditions on Ge layers, that is, the strained wetting layer and relaxed islands, the Si layer grown on the Ge dots has a tensile strain, whereas the Si layer on the Ge wetting layer is strain-free. If this Si layer is thinner than the correlation length [28], the strain condition remains, and the further Ge deposited on the Si layer (spacer) can vertically align with the underside of the Ge layers. The Ge dots of individual Ge layers are vertically self-assembled in a row in the same position [29] (Fig. 3). Ge quantum dots can also be fabricated with fine structures such as Ge/Si/Ge composite dots (Fig. 4) to improve uniformity, dot area coverage (Fig. 5), and photoluminescence efficiency. The Ge quantum dots are used for photodetectors at wavelengths of 1.3, 1.5, and 10 μm (see Section 3.3.2).

To reduce the thermal budget of Si<sub>x</sub>Ge<sub>y</sub>-based devices, a low-temperature oxide is often used. Liquid-phase deposition can provide a low-temperature oxide on Si<sub>x</sub>Ge<sub>y</sub>. It is interesting that silicon dioxide dots can be deposited on the high-energy group of two valleys along the growth direction and a high-energy group of two valleys along the growth direction. Similarly, the originally sixfold degenerate light and heavy hole bands split into the high-energy heavy-hole band and the low-energy light-hole band. For tensile strain such as that of
strained Si on relaxed Si$_{1-x}$Ge$_x$, the results are similar, but the energy levels switch between the degenerate groups. The two conduction valleys along the growth direction are lower in energy than the other four valleys in the strained Si. So are the valence bands. The effects are sometimes referred to as “uniaxial splitting.”

The strained Si$_{1-x}$Ge$_x$ on Si has a smaller lattice constant in the two in-plane directions, and larger lattice constants in the growth direction as compared with originally relaxed Si$_{1-x}$Ge$_x$. The net volume decrease (hydrostatic term) of the tetragonal Si$_{1-x}$Ge$_x$ unit cell grown on Si leads to a shrinkage of the bandgap, that is, the valence band moves up and the conduction band moves down. Note that the strained Si on relaxed Si$_{1-x}$Ge$_x$ has the opposite effect, and the net volume increases. The bandgap of strained Si$_{1-x}$Ge$_x$ on Si is given in Figure 6 with the relaxed bandgap, and the optoelectronic devices for optical communication (1.3 and 1.5 µm) can be easily implemented by the strained Si$_{1-x}$Ge$_x$ bandgap. At low Ge concentration, the bandgap reduction is ~7.5 meV per % Ge for strained Si$_{1-x}$Ge$_x$ grown on Si. The conduction band discontinuity between strained Si$_{1-x}$Ge$_x$ and Si is often negligible. For strained Si grown on relaxed Si$_{1-x}$Ge$_x$ buffers, the bandgap reduction of Si is ~4 meV per % Ge. The smaller bandgap reduction of strained Si on relaxed Si$_{1-x}$Ge$_x$ as compared with the strained Si$_{1-x}$Ge$_x$ on Si is due to the fact that the tetragonal distorted lattice of strained Si has a net volume expansion to widen the conduction band edge and valence band edge. The conduction band edge of strained Si is lower than that of relaxed Si$_{1-x}$Ge$_x$ ($\Delta E_c \sim 6$ meV/%), and the electron confinement is possible in the relaxed Si$_{1-x}$Ge$_x$/strained-Si structures. The modulation doped devices can thus be made (Section 3.2.3). The bandgap of strained Si$_{1-x}$C$_y$ on Si is also reduced with respect to Si with the experimental value of 68 meV/% C [30].

There are two effects for carbon incorporation into Si$_{1-x}$Ge$_x$ on the bandgap:

1. Carbon can change the strain condition of the pseudomorphic film, and this effect on band structure can

---

**Figure 3.** (a) TEM micrograph of multilayer Ge dots. The spacer thickness is ~20 nm. The SiO$_2$ dots are also formed above the Ge dots. (b) Schematic diagram showing the strain field in Si layers. Adapted with permission from [29], C. W. Liu et al., Appl. Phys. Lett. (2003). © 2003, American Institute of Physics.

**Figure 4.** Composite dots with fine structures such as Ge/Si/Ge composite dots.

**Figure 5.** AFM picture of (a) composite dots and (b) conventional dots. The composite dots have a larger coverage area.

**Figure 6.** The bandgap of strained Si$_{1-x}$Ge$_x$ on a (001) Si substrate and relaxed bandgap of Si$_{1-x}$Ge$_x$ alloys. The shadow areas indicate the optical communication bandwidth of 1.3 and 1.5 µm. Reprinted with permission from [16], B.-C. Hsu et al., “International Electron Device Meeting” 2002, pp. 91–94. © 2002, IEEE.
be taken account of in a manner similar to that of Ge incorporation.

2. Carbon also changes the fundamental bandgap [25, 31], and the relaxed Si$_{1-x}$C$_x$ bandgap decreases with carbon concentration initially because of the local lattice disorder, despite the large bandgap of diamond.

2.4. Mobility

2.4.1. Mobility in Relaxed and Strained Si$_{1-x}$Ge$_x$ alloys

The theoretical calculation [32] and experimental [33] values of the majority hole mobility in relaxed Si$_{1-x}$Ge$_x$ alloys are plotted in Figure 7. The decrease in mobility with Ge concentration is due to the alloy scattering, and an alloy scattering potential of 0.7 eV is used in the calculation to fit the measurement data.

Because of the impurity scattering, the mobility depends on the doping concentration. The in-plane and out-of-plane majority hole mobilities in strained Si$_{1-x}$Ge$_x$ layers for different doping concentrations are shown in Figure 8 [34]. Note that the hole mobilities in n-type material and p-type material are different from the terms of majority hole mobility and minority hole mobility, respectively. So are electron mobilities.

In the calculation, the alloy scattering potential determines the mobility value [37]. The majority hole mobilities in the strained Si$_{1-x}$Ge$_x$ alloys based on Monte Carlo simulation decrease by a factor of 2–3 if an alloy potential of 1.4 eV is used, as compared with 0.7 eV up to the doping concentration of 10$^{16}$ cm$^{-3}$. The hole mobility of strained Si$_{1-x}$Ge$_x$ calculated by different groups [31, 35, 36] is qualitatively similar. For a doping concentration of ≥ 10$^{19}$ cm$^{-3}$, the mobility depends weakly on Ge content, and the alloy scattering is not important. The majority hole mobility is important for the p-channel strained Si$_{1-x}$Ge$_x$ FET, where the current is made of the majority hole current.

Both in-plane and out-of-plane minority electron mobilities at different doping concentrations are plotted in Figure 9 [37, 38]. The out-of-plane mobility is especially important for npn Si/SiGe heterojunction bipolar transistor applications, where the electron transit time in the base is shortened by the electrical field generated by the graded Ge profiles. In-plane electron mobility decreases with increasing Ge fraction, whereas out-of-plane electron mobility increases with Ge fraction at small Ge content. The out-of-plane mobility is larger than the in-plane mobility.

Klaassen model [39, 40] is commonly used for minority and majority mobility. Klaassen’s model takes into account four scattering mechanisms: lattice scattering, ionized impurity scattering, phonon scattering, and carrier-carrier scattering. The last mechanism makes the minority and majority carrier mobilities different.

2.4.2. Mobility in Strained Si

Figure 10 shows the in-plane electron and hole mobilities at 300 K in strained Si, based on [41]. Here the carrier-carrier scattering is neglected, and majority mobility and minority mobilities are the same. For tensile strain (positive strain value), the in-plane electron mobility increases rapidly, and electrons populate the two lower energy Δ valleys along the growth direction. The lower conductivity mass
SiGe/Si Heterostructures

is the main cause of increased mobility with the help of the small decrease of intervalley scattering (the strong g-type process remains effective). The electron mobility reaches \( \sim 2300 \text{ cm}^2/\text{Vs} \) at a tensile strain of about 0.8% (corresponding to a fully relaxed Si\(_{0.8}\)Ge\(_{0.2}\) buffer layer). On the other hand, the out-of-plane mobility in Si with tensile strain drops rapidly with strain, since the conductivity mass is now the large longitudinal mass. However, only the in-plane mobility matters in the field effect device applications.

For the compressively strained Si, the in-plane electron mobility initially decreases with the strain, since electrons mainly populate the four \( \Delta \) valleys perpendicular to the growth direction with a large effective conductive mass. With the further increase in compressive strain, the separation energy between the twofold degenerate valleys and fourfold degenerate valleys increases, and this reduces the intervalley scattering. The in-plane electron mobility increases again. Note that the out-of-plane electron mobility, controlled by the small transverse mass, remains high.

The hole mobility in strained Si increases for strain of either sign. The principal cause of the enhancement is the split of the degeneracy between the heavy-hole and light-hole bands at the \( \Gamma \) point in the reciprocal space and the reduction of the conductivity mass. For tensile strain, the light hole band has a much higher energy than the heavy-hole band, whereas for compressive strain the splitting of these two bands remains comparable to \( kT \). Therefore, interband scattering in Si with tensile strain is not as effective as in the compressively strained Si, and this results in larger hole mobility in Si with tensile strain.

Tensile strain improves both electron and hole mobilities by reducing the effective transport masses and suppressing intervalley phonon scattering. For the MOSFET applications, the carrier in the inversion layers also suffers the oxide roughness scattering, which decreases the mobility value in the Figure 10. The calculated electron and hole inversion layer mobility enhancements in strained Si are plotted in Figure 11 [42, 43]. The amount of strain in Si is proportional to the Ge content in the relaxed SiGe buffer layer. The electron and hole mobilities are expected to increase with increasing strain and saturate at Ge concentrations of 20% and 40%, respectively, in the buffer layers. The saturation of mobility enhancement is related to the diminishing benefit of strain-induced energy splitting. The suppression of the intervalley scattering by energy splitting increases the mobility, but other scattering mechanisms, such as oxide roughness scattering and other phonon scattering, begin to dominate the low field mobility. At low temperature, there are fewer phonons, and the smaller energy splitting is enough to suppress the intervalley phonon scattering. Therefore, the mobility enhancement saturation is expected at even lower Ge content. Please note that both NMOSFET and PMOSFET can use the strained Si channel to increase the device speed.

### 2.4.3. Mobility in Strained Ge

Figure 12 shows the in-plane electron and hole mobilities at 300 K in strained Ge [41]. The electron mobility shows a noticeably complicated behavior. For tensile strain, electrons begin to populate the conduction states at the \( \Gamma \) point at strains larger than 1%. Because of the small in-plane effective electron mass in the \( \Gamma \) valley, there is a dramatic enhancement of mobility for strains larger than 1%. For compressive strain, the electrons mainly populate the \( \Delta \) valleys at small strain, but the energy separation between

![Figure 11](image-url)
L valleys and the four-fold Δ valleys perpendicular to the growth direction shrinks. The resulting larger intervalley scattering causes a lower mobility. At large compressive strain, all electrons stay in the Δ valleys. The mobility now approximates a constant value, nearly equivalent to the Δ valleys’ electron mobility in relaxed Ge. In practice, only the compressively strained Ge can be grown. Therefore, there is no advantage to using the Ge channel in the NMOSFET, but it is interesting to note that the Ge with tensile strain can have a direct bandgap.

For in-plane hole mobility, the situation is qualitatively similar to that of strained Si. The compressive Ge has an in-plane hole mobility greater than 2000 cm²/Vs and is beneficial for the PMOSFET application.

2.5. Relaxed SiGe on an Insulator

To have fully relaxed Si₁₋ₓGeₓ grown on Si, the Si₁₋ₓGeₓ thickness has to be much larger (>>10x) than the critical thickness. The growth rate has to be high enough to ensure a reasonable growth time. The growth of films at low temperature takes a long time, but they might also have to be annealed at high temperature after growth to obtain sufficient relaxation.

The misfit and threading dislocations are generated in the relaxed films. All of the problems arise from the thick Si (500–800 μm) substrates. The thin epitaxial Si₁₋ₓGeₓ film has to fight with thick substrate to become relaxed with the pay-off of defects. It is unlikely that Si substrates can be thinned down far below 50 μm with mechanical stability. A smart way to effectively grow Si₁₋ₓGeₓ on thin Si is to use the silicon-on-insulator (SOI) substrates [44], and the Si can be as thin as 10 nm on the buried oxide (BOX, commonly used insulator). The SOI devices have the extra advantages of high speed, latch-up resistance, and radiation hardening. If the bonding force between Si and BOX can be small compared with Si and Si₁₋ₓGeₓ, a Si₁₋ₓGeₓ thicker than SOI can easily be relaxed without any defect generation, and the thin SOI becomes strained. Several variations have been invented based on the basic idea.

The BOX layer can be created by separation-by-implanted-oxygen (SIMOX) technology after the Si₁₋ₓGeₓ layer is grown on conventional bulk Si wafers [45]. The implantation depth can reach Si or Si₁₋ₓGeₓ. The relaxed films and BOX can be formed after high-temperature (>1100 °C) annealing. To obtain a relaxed Si₁₋ₓGeₓ layer with a large Ge fraction, the relaxed films can be further oxidized. The oxidation of Si₁₋ₓGeₓ can inject Ge into underlying Si₁₋ₓGeₓ films, and the underlying films can have a higher Ge fraction.

Bond-and-etch-back SOI (BESOI) wafers have the advantage of a cleaner silicon/oxide interface at the BOX than SIMOX. Oxidation of the seed and handle wafer, followed by bonding of the two wafers, generates the BESOI. The Si can be thinned down by the combination of CMP (chemical-mechanical polish) and etching to the desired film thickness. The relaxed Si₁₋ₓGeₓ can be grown on the BESOI wafers with thin Si (<100 nm) on the BOX.

The hydrogen ions implanted in Si at a sufficient dose (10¹⁶ to 10¹⁷ cm⁻²) can break Si into two pieces after annealing at 400 °C to 600 °C because of the formation of hydrogen molecules (bubbles) in the implanted region. The way to cut Si by hydrogen implantation is called “smart cut.” The splitting depth depends on the ion energy, and the splitting temperature depends on the ion dose. The combination of smart cut and wafer bonding can also yield relaxed Si₁₋ₓGeₓ with the use of one host wafer with oxide and the other handle wafer with hydrogen-implanted Si₁₋ₓGeₓ [46, 47]. Figure 13 shows one of the procedures for the smart-cut and layer transfer process to make strained Si on SiGe-on-insulator (SGOI) material, and Figure 14 is a transmission electron microscopy (TEM) picture of the resulting relaxed Si₁₋ₓGeₓ films with strained Si on the top. The as-split SGOI has a surface roughness of 5–12 nm (rms value); therefore a CMP process is needed to smooth the surface of the transferred layer. The relaxation of the Si₁₋ₓGeₓ layer is confirmed by X-ray diffraction measurement.

3. DEVICE PERFORMANCES

Advanced BiCMOS technology has fully utilized graded-base SiGe:C heterojunction bipolar transistors, which will dominate the radiofrequency (up to 10 GHz) and optical communication (up to 40 Gbit/s) markets. For mainstream Si technology, to increase the MOSFET speed with the same lithography feature, the strained Si MOSFET will be used in the 90-nm technology node and beyond. Si₁₋ₓGeₓ is one of
the enabling technologies that can make Si prosperous for at least two more decades.

In addition to conventional Si devices, because of the heterostructures of Si/Si$_{1-x}$Ge$_x$, some III–V device structures, such as modulation doped FET and resonant tunneling diodes, can be made with Si. It is possible that these Si based quantum devices can outperform the original III–V devices.

The bandgap of Si (∼1.1 eV) almost makes it impossible to have Si-based devices operating at 1.3-μm, 1.5-μm wavelengths and at the far-infrared region (∼10 μm) for optoelectronic and defense applications. However, Si$_{1-x}$Ge$_x$ with a small bandgap can absorb the light of 1.3 and 1.5 μm, and Ge/Si quantum structures with intraband transition can have far-infrared absorption. For light emission devices, Si$_{1-x}$Ge$_x$ p-i-n light-emitting diodes are also demonstrated.

### 3.1. Si/SiGe:C Heterojunction Bipolar Transistors

Since the first Si/SiGe heterojunction bipolar transistor (HBT) was reported in 1988 [48], tremendous progress of Si/SiGe:C HBTs has been achieved, reaching a maximum oscillation frequency of 210 GHz [1] and a cut-off frequency of 285 GHz [2], with circuit applications in wireless and optical communication.

In the NPN bipolar transistor, the collector current is controlled by application of a forward bias $V_{BE}$ to the emitter-base junction [49]. In the forward active region, the base-collector junction is reverse biased by $V_{CB}$, as shown in Figure 15. More electrons are injected into the base region in HBT than in bipolar junction transistor (BJT) because of the lowering of the potential barrier $\Delta V_n$. Note that the same base doping concentration is assumed in HBT and BJT, and therefore the separation between the Fermi level and the valence band edge is the same for HBT and BJT. This results in an increase in collector current as compared to Si BJT with the same base doping. In other words, to have the same barrier $\Delta V_n$ (same collector current), the HBT can have a heavy base doping, which moves the conduction band edge in the base upward. The narrow bandgap of the Si$_{1-x}$Ge$_x$ base can have a concentration of up to $2 \times 10^{20}$ cm$^{-3}$ [50], which significantly reduces the base resistance and hence leads to a high maximum oscillation frequency, high Early voltage, and a low noise figure, as compared with Si bipolar junction transistors.

A built-in electric field for electrons in the p-type base region can be obtained with graded Si$_{1-x}$Ge$_x$ (shown in Fig. 16). The electrical field accelerates the electron motion in the base and thus reduces the base transit time. The cut-off frequency $f_t$ can be enhanced. Many studies [51–56] are focused on the optimal Ge profile of SiGe base HBT to minimize the base transit time. The narrow base also reduces the base transit time. However, because of the boron out-diffusion, especially by transient-enhanced diffusion (TED) in the fabrication process, the base width is limited. An ultra-high-speed HBT requires a heavily doped and extremely narrow base. The carbon incorporation in

![Figure 15](https://example.com/figure15.png)
the Si$_{1-x}$Ge$_x$ base can reduce the boron out-diffusion to a narrow base with small base resistance and without the parasitic barrier in the collector. A low concentration of carbon atoms ($\sim 0.2\% = 10^{20}$ cm$^{-3}$) is introduced into the base region to effectively suppress the TED of boron [57]. The suppressed boron diffusion found in carbon-doped Si and SiGe is caused by an undersaturation of Si interstitials due to carbon out-diffusion from the base. However, it is reported that the addition of a carbon concentration higher than $\sim 1\%$ can degrade the electron mobility by a factor of 2 [58].

### 3.2. Field Effect Transistors

#### 3.2.1. N Channel Metal-Oxide-Semiconductor Field Effect Transistors

Very high electron mobilities in strained Si layers suggest a great potential for this material in high-performance N channel metal-oxide-semiconductor field effect transistors NMOSFETs. The low defect density buffers and high 14 quality oxide at low temperature are key technologies for ensuring a mobility advantage in device performance.

The enhancement of effective electron mobility has been reported for long-channel MOSFETs with both surface and buried channels [59]. Figure 17 shows schematic diagrams of these two possible configurations of strained Si channel NMOSFETs [60]. Both structures have relaxed SiGe buffer layers grown on either a graded-Ge layer or SOI. The surface channel device (Fig. 17a) has a single layer of thin strained Si grown on top of a relaxed buffer layer. This layer is oxidized to form gate oxide. The buried strained Si channel device (Fig. 17b) has a layer of strained Si buried beneath a thin layer of relaxed SiGe. An additional layer of strained Si is necessary to form gate oxide on top of the SiGe, and ideally this additional Si layer should be consumed during oxidation. If it is not, the residual sacrificial Si layer between the gate oxide and the SiGe barrier layer can act as a parallel conducting channel and strongly affects the device performance. Depending on the dopant type in the layers, these structures can also be used for PMOSFETs or NMOSFETs. The long channel ($W = 10 \mu m \times 168 \mu m$) surface and buried-channel NMOSFET devices fabricated on relaxed Si$_{1-x}$Ge$_{0.3}$ buffer layers showed well-behaved output characteristics. The effective low-field mobilities for these device structures [60] are shown in Figure 18. For the surface channel strained Si device, the effective mobility is enhanced with a similar dependence on the effective electric field as compared with a bulk-Si control device. The peak mobility value is 1000 cm$^2$/Vs, which shows 80% enhancement over the Si control device (550 cm$^2$/Vs). The peak mobility value of the buried channel device is over 1600 cm$^2$/Vs, which is almost 3 times that of the Si control device, but drops rapidly with increasing effective field. The enhancement of strained Si NMOSFET mobility becomes less significant, if the low Ge content buffer is used [60], since the strained Si mobility increases with increasing strain (more Ge content in the relaxed buffer layer). Note that buried channel devices with deep submicron length suffer from degraded threshold voltage roll-off and subthreshold characteristics.

Recently, IBM [6] has demonstrated current drive enhancements in the strained-Si NMOSFET with sub-100-nm physical gate lengths. Measured effective electron mobility characteristics are shown in Figure 19. For $\sim 1.2\%$ strain (i.e., 28% Ge content in buffer layer), effective electron
3.2.2. P Channel Metal-Oxide-Semiconductor Field Effect Transistors

Strained-Si P channel metal-oxide-semiconductor field effect transistors (PMOSFETs) have been studied [61–67] through the use of high hole mobility. The tensile strain in silicon grown on relaxed SiGe buffer raises the light-hole band and lowers the heavy-hole band. This reduces an intervalley scattering and increases the low field hole mobility significantly. Recently, IBM [6] has demonstrated current drive enhancements in the strained-Si PMOSFET with sub-100-nm physical gate lengths. Measured hole mobility characteristics are shown in Figure 20. For ~1.2% strain (i.e., 28% Ge content in the buffer layer), peak hole mobility was enhanced by 45% over that of the control Si. The hole mobility enhancement diminishes at high effective electric field. For lower strain in Si (13% Ge content in buffer layer), hole mobility is slightly degraded compared with the control Si. Since the strained SiGe also has high hole mobility (Fig. 8), it is desirable to use a SiGe channel in the PMOSFETs. With a strained SiGe buried channel below the strained Si in the buried channel device (Fig. 21), the hole mobility has an enhancement comparable to that of the Si surface channel structure with high strain (28% Ge content in the buffer layer). The same strained Si structure suitable for NMOSFETs and PMOSFETs makes it possible to apply strained Si technology in CMOS circuits.

Since the compressively strained Ge has a large hole mobility (Fig. 12), Ge channel PMOSFETs with an effective mobility of 2700 cm²/Vs have also been reported [5].

3.2.3. Modulation Doped Field Effect Transistors

Carriers can be separated from their parent donor or acceptor atoms for suitable band discontinuity. The impurity scattering can be further reduced if a spacer is inserted between the doped carrier-supply layers and undoped conduction channels. In principle, the modulation doped field effect transistors (MODFET) can have a much higher mobility than the MOSFET because of the lack of impurity scattering. With the band alignment requirement, NMODFETs have to use a strained Si channel on relaxed Si₁₋ₓGeₓ [68], whereas PMODFETs have to use strained Si₁₋ₓGeₓ on relaxed Si or strained Ge on relaxed Si₁₋ₓGeₓ [5, 69]. Figure 22 shows...
SiGe/Si Heterostructures

3.3. Optoelectronic Devices

In the last decade, interest in silicon-based optoelectronic devices has grown rapidly [76]. The trend of optoelectronics is to integrate both high-performance optoelectronic and electrical devices on the same chips with silicon processing technology. Among silicon-integrable materials, Ge and SiGe are of great importance to the fabrication of optoelectronic devices because of their optical absorption wavelengths.

Many SiGe/Si devices operated in the 1.3–1.5 μm wavelength range have been reported and demonstrated [77–78]. In this section, we discuss Si/SiGe photodetectors, light-emitting diodes, and heterojunction phototransistors.

3.3.1. Absorption of SiGe

Because of the large absorption length (≈16 μm) of Si at 820 nm [79] and the forbidden absorption at 1300 and 1550 nm, Si-based photodetectors have limited detection efficiency and wavelength range. Not only can the incorporation of Ge into Si increase the cutoff wavelength (Fig. 6); it can also enhance the absorption efficiency (small absorption length). The strained Ge could have an absorption length of 0.1 μm or less at a wavelength of 820 nm. Figure 24 shows the absorption length at 820, 1300, and 1550 nm versus Ge mole fraction. The absorption length decreases as the Ge mole fraction increases. For the large Ge fraction, the shadowed areas indicate the uncertainty of the estimation. The incorporation of strained Ge/SiGe into optoelectronic devices makes devices particularly useful over the important fiber optic communications wavelengths.

3.3.2. Photodetectors

The figures of merit for light detection are high efficiency, high responsivity, low noise, low dark currents, and fast response. Previous studies have focused on metal-semiconductor-metal (MSM) diodes and p-i-n diodes.

**Figure 22.** Typical MODFET structures. (a) NMODFETs with Si-channel on relaxed-SiGe buffer. (b) PMODFET with SiGe channel and (c) PMODFET with Ge channel on relaxed-SiGe buffer. Note that the doping layers and spacers can be located below the bottom of the channels (after [70]).

these three possible MODFET structures [70]. As in MOSFETs, the large defect density of these structures can seriously degrade the device performance.

An electron mobility up to 2700 cm²/Vs at 300 K for a sheet carrier density of 7 × 10¹² cm⁻² in a strained Si channel NMODFET is reported [71], and a hole mobility of 1880 cm²/Vs at 300 K for a sheet carrier density of 2.1 × 10¹² cm⁻² has been obtained with a pure Ge channel layer on a relaxed Si₀.₄Ge₀.₆ buffer [71]. All of the available data on effective electron and hole mobilities for MODFET are shown in Figure 23 [71–75]. From a comparison of Figure 19 and 23, it is evident the MODFETs have larger enhancement than MOSFETs.

**Figure 23.** Available experimental data [71–75] at 300 K for effective electron and hole mobility in MODFET. A denotes modulation doping above a strained Si channel, B denotes a doping supply layer below a strained silicon channel, x is the Ge fraction in the channel, and y is the Ge fraction in the buffer layers. DC, DaimlerChrysler Research Center; CNET, France Telecom.

**Figure 24.** The absorption length at 820, 1300, and 1550 nm versus Ge mole fraction. The absorption length decreases as the Ge mole fraction increases. For the large Ge fraction, the shadowed areas indicate the uncertainty of the estimation. Reprinted with permission from [16], B.-C. Hsu et al., “International Electron Device Meeting,” 2002, pp. 91–94. © 2002, IEEE.
In 1992, a metal-semiconductor-metal (MSM) diodes was reported by Splett et al. [80]. The 17 measured responsivity at 1.3 μm is 0.2 A/W over a 1-mm detector length with 500 pA/μm² dark current at 5-V bias. In 1996, Huang et al. reported an epitaxial SiGeC/Si photodetector with a response in the 1.3–1.55-μm wavelength range [77]. The active absorption layer of the p-i-n photodiode consists of a pseudomorphic SiGeC alloy grown on a Si substrate with a Ge content of 55% and a thickness of 80 nm. In 2001, a p-i-n Ge on Si photodetector was reported by Masini et al. [78]. The photodiodes exhibit short-circuit responsivities of 0.3 and 0.2 A/W at 1.3 and 1.55 μm, respectively, reverse dark currents of 20 mA/cm², and response times of 800 ps.

In 1999, photodetectors and light-emitting diodes for 1.1-μm wavelength have been demonstrated with metal-oxide-silicon (MOS) tunneling structures on both n-type and p-type silicon substrates [81–84]. The ultrathin thickness (<3 nm) of thermal oxide is required in these novel photodetectors to provide sufficiently large tunneling probability. As biased in inversion region, the tunneling diode works in the deep depletion region with the soft pinning of oxide voltage, instead of pinning of surface potential, which is very different from the conventional MOS diode with thick oxide. The Ge MOS detector can operate at 1.3 and 1.55 μm with high responsivity [85]. To avoid material degradation such as strain relaxation and Ge out-diffusion, low-temperature liquid phase deposition (LPD) oxide is introduced [86].

A MOS Ge quantum dot photodetector is demonstrated [16]. The photodetector has responsivities of 130, 0.16, and 0.08 mA/W at detection wavelengths of 820 nm, 1300 nm, and 1550 nm, respectively. The dark current is extremely low (0.06 mA/cm²). Si/Ge quantum dots are prepared by ultra-high-vacuum chemical vapor deposition (UHVCVD) on p-type Si (001) substrates. The structure is shown in Figure 25. With the careful design of intraband transition, the long-wavelength infrared (6–10 μm) can also be detected with the same devices.

### 3.3.3. Light-Emitting Diodes

Two different Si light emitters have been reported. A narrow-band infrared emitter at 1160 nm was implemented with a pn junction under forward bias with an external quantum efficiency of ~10⁻⁴ [87]. A broad-band (450–850 nm) visible-light emitter was also realized with an avalanche pn diode with an external quantum efficiency of ~10⁻⁸ [87] and ~10⁻⁶ [88]. Recently, the band-edge (1.1 μm) electroluminescence (EL) of a MOS tunneling light-emitting diode has also been reported [83]. Because of the indirect bandgap of Si, additional momentum is required for the light emission process. The phonon provides the additional momentum in bulk Si and bulk SiGe [89]. In the MOS structure, the interface/surface roughness can seriously affect the carrier transport in the inversion layer [90]. This indicates that the interface/surface roughness can scatter the carrier and can change the carrier momentum. Therefore, the rough oxide can enhance the emission efficiency of MOS light emitting diodes [91].

The addition of Si₁ₓGeₓ to p-i-n diodes can increase the emission wavelength. Both 1.3-μm [18] and 1.5-μm [19] emissions with Si₁ₓGeₓ/Si quantum wells and Ge/Si quantum dots, respectively, are demonstrated at room temperature, and erbium-doped Si₁ₓGeₓ can also have 1.5-μm emission [92]. However, MOS diodes with Si₁ₓGeₓ or Ge quantum dots are not reported to have long-wavelength emission so far.

### 3.3.4. Heterojunction Phototransistors

With high gain and low noise, the phototransistor can be used in the front end of an optical receiver. The Si₁ₓGeₓ/Si multiple quantum well heterojunction phototransistor with an ultrahigh responsivity of 1.47 A/W and a bandwidth of >1.25 GHz at a wavelength of 850 nm has been reported at an operation voltage as low as ≤1.5 V [13]. The responsivity at 1310 nm is 0.15 A/W. The multiple quantum wells in the collectors can enhance the photon absorption and increase the cutoff wavelength. The current gain of the transistors can further amplify the optical current. Si/Si₁ₓGeₓ quantum well phototransistors with far-infrared (6–20 μm) response with low leakage current and high gain were also proposed [93].

### 4. FABRICATION TECHNOLOGIES

#### 4.1. Growth of Si₁₋ₓGeₓ Films

Si₁₋ₓGeₓ can be epitaxially grown by chemical vapor deposition (CVD) and molecular beam epitaxy (MBE). Depending on the growth pressures, CVD has some variations: APCVD (atmospheric pressure), rapid thermal CVD (RTCD, several torr with lamp heating), low-pressure CVD (LPCVD, several tenths of a torr), and ultrahigh pressure CVD (UHV/CVD). UHV/CVD [94] has a growth pressure of 10⁻² to 10⁻⁵ torr and a base pressure of UHV (~10⁻⁹ torr). The designation UHV is sometimes confusing and only indicates the base pressure of the reactor chamber, not the growth pressure. The wafer temperature can be obtained by lamp heating or thermal equilibrium furnaces. For lamp heating, the wafer temperature can rise and fall very rapidly (on the order of 10 s), and is often called “rapid thermal CVD” [95]. The chamber used for RTCVD is often made of quartz to avoid contamination during growth. The quartz chamber does not absorb as much emission from the lamps.

![Figure 25. The MOS detector structure with five layers of Ge quantum dots.](Image)
as Si wafers. A cold-wall RTCVD can be obtained, whereas a quartz tube with hot walls is used in the furnace-type reactor, such as that used for UHV/CVD. Because of the configuration of heating lamps, RTCVD can process only a single wafer at a time, whereas a furnace-type reactor such as that used for UHV/CVD can process a batch of wafers (more than 20 wafers). Silane (SiH4) and dichlorosilane (DCS) are the common precursors of Si, and germane (GeH4) is the precursor of Ge. Phosphine (PH3) is used for n-type doping, and diborane (B2H6) is for p-type doping. Both RTCVD and UHV/CVD are used in semiconductor manufacturing laboratories as well as research laboratories.

MBE is a physical deposition system with a stainless chamber. The base pressure and growth pressure are about 10⁻¹¹ and 10⁻⁹ torr, respectively. The electron beams are used to heat source materials such as Si and Ge, and the Si and Ge atoms are evaporated on Si substrates. The chamber is cold and substrates are heated by a local heater. Because of the difficulty of using large-diameter wafers and some process issues, MBE is used in research laboratories, and HBTs with ideal base currents are grown by MBE [96] after CVD.

4.2. Metal Contacts

As the microelectronics industry moves into the era of deep submicron devices, metallization has become the main performance-limiting factor. Metal silicides have been widely used to reduce the contact resistance of the source/drain of microelectronic devices. The self-aligned silicidation technique has become a crucial part of recent ultra-high-speed silicon device technologies [97]. TiSi2 is currently the most commonly used silicide in the IC industry. In the deep submicron regime, a linewidth dependence of sheet resistance was observed for TiSi2 [98, 99]. CoSi2 has been introduced to replace TiSi2 in sub-quarter-micron technology. However, its use in sub-0.1-micron devices is in doubt unless a raised source/drain scheme becomes feasible. NiSi is the only silicide left with resistivity comparable to that of TiSi2 and CoSi2. On the other hand, intensive efforts have been made to extend IC devices to other substrate frames, such as Si-Ge. Metal germanosilicide/Ge_x/Ge_y/Si heterostructures are promising for use in devices such as the heterojunction bipolar transistor and infrared detectors with high cutoff wavelengths [100, 101].

Metal germanosilicides may be used for making metallic contacts with Si_{x-y}Ge_y alloys, and knowledge about the formation and stability of thin metal germanosilicide films is essential for such applications. The thermally induced metal/Si_{x-y}Ge_y reaction has previously been studied for many metals. Various degrees of germanium segregation and/or the formation of segregated layered structures were observed in the reactions of these metal/Si_{x-y}Ge_y systems. The difference in heats of formation of silicide and germanide offers the driving force for the segregation of Ge-rich Si-Ge alloy [102]. The phase formation paths were found to depend strongly on the composition of substrates. By extension from silicon to Si-Ge devices, the contact materials considered for Si-Ge devices have been focused on Ti, Co, and Ni contacts.

4.2.1. Co/Si-Ge System

Direct deposition of metal thin films on Si_{x-y}Ge_y in a self-aligned silicide process would be an efficient technique and would take advantage of established technologies. However, the major difficulties in using Co as a contact material for Si_{x-y}Ge_y appear to be the preferential reaction of Co with Si and the high consumption ratio of the Si to Co to form CoSi2. The undesirable characteristics lead to Ge segregation, deterioration of the Si_{x-y}Ge_y layer, degradation of junction integrity, and film agglomeration at low temperatures, resulting in poor thermal stability and high-resistivity contacts [103–105]. In addition, it has been demonstrated that CoSi2 is formed at much higher temperatures on a Si_{x-y}Ge_y layer than that on Si, owing to Ge expelled from the Co-Si-Ge compound, which blocks the Co diffusion paths, slowing down the reaction [106–108]. Successful formation of good quality CoSi2 on epitaxial Si_{x-y}Ge_y in the presence of a sacrificial a-Si layer has been achieved. CoSi2 was previously found to form at a lower temperature on a-Si with a smoother interface than that on single-crystal Si [109].

For convenience, the two sets of samples Co(15 nm)/Si_{x-y}Ge_y and Co(15 nm)/a-Si(50 nm)/Si_{x-y}Ge_y are designated as samples A and B, respectively. Five hundred-nanometer-thick Si_{x-y}Ge_y and 1-μm-thick strained layers of Si_{x-y}Ge_y (y varies from 1 to 0.7) were grown on (001)Si wafers at 550 °C by MBE. A 50-nm-thick sacrificial a-Si layer was deposited on a Si_{x-y}Ge_y substrate followed by the deposition of a 15-nm-thick Co thin film without breaking the vacuum at room temperature. Figure 26 shows the sheet resistance data of the two sets of samples after different heat treatments. For samples annealed at 600 °C, the sheet resistance of samples B is much lower than that of samples A. The low resistance is attributed to the early formation of low-resistivity CoSi2, which is consistent with the X-ray diffraction data. In contrast, CoSi2 was not detected in samples A annealed at a temperature as high as 800 °C. The further substantial increase in sheet resistance for Co(15 nm)/Si_{x-y}Ge_y samples after 800 °C annealing,
as shown in Figure 26, is correlated to the agglomeration of Co(Si$_{1-x}$Ge$_x$). Figure 27 is a plane-view TEM micrograph showing the formation of Co(Si$_{1-x}$Ge$_x$) islands in a Co(15 nm)/Si$_{0.7}$Ge$_{0.3}$ sample annealed at 800 °C. On the other hand, for the Co(15 nm)/a-Si(50 nm)/Si$_{0.7}$Ge$_{0.3}$ samples, the formation of CoSi$_2$ was completed after 700 °C annealing. The sheet resistance maintained the same low level for samples B after annealing at 700–900 °C. In 900 °C annealed samples, the silicide films were found to be continuous with the smooth interface with Si$_{0.7}$Ge$_{0.3}$, as revealed by plane-view TEM and XTEM images shown in Figure 28. No Ge segregation was detected from the analysis of TEM direct images and electron diffraction patterns. In contrast with samples B, rough interface and Ge segregation in samples A are evident. An example is shown in Figure 29.

Previous study showed that the retardation of the silicidation process on Si$_{1-x}$Ge$_x$ is related to a higher effective Ge concentration at the reaction front, therefore causing an increase in the interface energy. The expelled and segregated Ge tends to diffuse to the surface and interface of the silicide layer [104, 105]. These Ge atoms can decorate the grain boundaries, resulting in an increase in the grain boundary energy, which makes the silicide film more prone to agglomeration. On the other hand, the sacrificial a-Si layer is shown to improve the interfacial roughness and thermal stability of a CoSi$_2$ film grown on Si$_{0.7}$Ge$_{0.3}$. The increased uniformity of the silicide/Si$_{0.7}$Ge$_{0.3}$ interface is speculated to result from small-grained CoSi$_2$. The average grain size is smaller for silicides formed on a-Si than that on single-crystal Si and poly-Si in samples subjected to the same heat treatment [110]. The results indicated that the reaction between Co thin films and a-Si incurs the formation of small-grained CoSi$_2$, which in turn enhances the morphological stability of CoSi$_2$. In addition, the formation of thermally stable CoSi$_2$ leads to a low sheet resistance value.

### 4.2.2. Ti/Si-Ge System

Titanium disilicide has been widely used for contacts to Si in microelectronics devices. Its use is complicated by the existence of two crystalline phases. The high-resistivity C49-TiSi$_2$ forms first on heating of Ti on Si above 500 °C, and additional heating above 700 °C is needed to transform C49 into the low-resistivity C54-TiSi$_2$. However, in the Ti/Si$_{1-x}$Ge$_x$ system, Ge segregation was observed near the C54 phase formation temperature [111]. In addition, the resistivity of C54-Ti(Si$_{1-x}$Ge$_x$)$_2$ forms on Si$_{1-x}$Ge$_x$ layer is higher than that of C54-TiSi$_2$. The appearance and agglomeration temperature of low-resistivity C54-Ti(Si$_{1-x}$Ge$_x$)$_2$ were both found to decrease with the Ge concentration [111].

![Figure 27. Plane-view TEM image of a Co(15 nm)/Si$_{0.7}$Ge$_{0.3}$ sample after annealing at 800 °C for 30 s by RTA. Reprinted with permission from [109], W. W. Wu et al., Appl. Phys. Lett. 81, 820 (2002). © 2002, American Institute of Physics.](image)

![Figure 28. (a) Plane-view and (b) cross-sectional TEM images of Co(15 nm)/a-Si(50 nm)/Si$_{0.7}$Ge$_{0.3}$ samples after annealing at 900 °C for 30 s by RTA. Reprinted with permission from [109], W. W. Wu et al., Appl. Phys. Lett. 81, 820 (2002). © 2002, American Institute of Physics.](image)

![Figure 29. Cross-sectional TEM image of Co(15 nm)/Si$_{0.7}$Ge$_{0.3}$ samples after annealing at 900 °C for 30 s by RTA.](image)
An amorphous Si was also successfully used to decrease the formation temperature of low-resistivity C54-TiSi2 and alleviate the formation of islands as well as the segregation of Ge. For the Ti/a-Si/Si1−xGex system, the thicknesses of Ti and amorphous Si were selected to be 15 and 30 nm, respectively. Figure 30 shows an XTEM image of a Ti/Si0.7Ge0.3 sample annealed at 900 °C, revealing the formation of C54-Ti(Si1−xGex)2 islands and Ge segregation. An amorphous Si was used to alleviate the formation of islands as well as the segregation of Ge. The selection of the thickness ratio of Ti and a-Si was such that the a-Si was completely consumed in forming C54-TiSi2. If the a-Si was too thin, Ti would start to react with the Si1−xGex layer and accompanying problems in forming higher-resistivity Ti(Si1−xGex) compound became troublesome [102]. Ge segregation [111], a rough interface [112], and strain relaxation of the Si1−xGex layer are expected to occur [113].

GIXRD spectra of Ti(15 nm)/a-Si(30 nm)/Si0.7Ge0.3 samples revealed that low-resistivity C54-TiSi2 was the dominant silicide phase above 650 °C. The formation temperature of C54-TiSi2 was lowered by about 100 °C in samples with a sacrificial a-Si interlayer compared with that of the Ti/Si0.7Ge0.3 system [111]. Figure 31 shows the resistivity curves of the Ti(30 nm)/Si0.7Ge0.3 and Ti(15 nm)/a-Si(30 nm)/Si0.7Ge0.3 samples (samples A and B) after annealing at different temperatures. For samples annealed at 700 °C, the resistivity of samples B is 23 much lower than that of samples A. It resulted from the formation of low-resistivity C54-TiSi2 at low temperature. On the other hand, for samples A, as the annealing temperature was increased to 800 °C, a sharp drop in resistivity occurred. In samples B, the temperature range of low-resistivity C54-TiSi2 was extended to 700–900 °C. From both resistivity and TEM data, it was concluded that the resistivity of C54-TiSi2 in samples B is lower than that of C54-Ti(Si1−xGex) in samples A. The magnitude of such differences will be affected by the homogeneity of the microstructures and will depend specifically upon the grain size [110]. On the other hand, with higher solute concentration, the electronic structure or phonon spectrum of the alloy will start to suffer from perturbation, and then the resistivity is expected to increase. For Ti(15 nm)/a-Si(30 nm)/Si0.7Ge0.3 samples, the a-Si was completely consumed in forming C54-TiSi2, and the phase possesses lower resistivity than ternary C54-Ti(Si1−xGex)2 phase. In addition, no Ge segregation was detected. The silicide films were also found to be continuous with the smooth interface with Si0.7Ge0.3, as shown in Figure 32. As a result, a sacrificial a-Si layer was found to decrease the formation temperature of C54-TiSi2, prevent Ge segregation, lower the resistivity of silicide, and improve the morphological stability of C54-TiSi2 on the Si0.7Ge0.3 layer. If the a-Si was too thin, Ti would react with the Si0.7Ge0.3 layer significantly and result in a rough interface.

4.2.3. Ni/Si-Ge System

Low-resistivity NiSi is currently one of the most promising silicides to replace TiSi2 for the self-aligned technology of ULSI, owing to its favorable properties low resistivity, low silicon consumption, low processing temperature, and relative insensitivity to the linewidth of the silicide.

For the Ni/Si0.7Ge0.3 system, Ge tended to be less reactive with Ni than with of Si. Moreover, Ge was found to segregate to grain boundaries in 600 °C annealed samples, and Ni(Si0.7Ge0.3) islands were found to form in 700 °C annealed samples. In the XTEM image observation, the Ge segregation toward the grain boundary was evident. An example is shown in Figure 33. Ni(Si0.7Ge0.3) islands were observed to form in 700 °C annealed samples [114].

![Figure 30](image1.png) **Figure 30.** Cross-sectional TEM image of Ti(30 nm)/Si0.7Ge0.3 samples after annealing at 900 °C for 30 s by RTA.

![Figure 31](image2.png) **Figure 31.** Resistivity versus annealing temperature curves for samples A and B after annealing at 500–1000 °C for 30 s by RTA.

![Figure 32](image3.png) **Figure 32.** Cross-sectional TEM image of Ti(15 nm)/a-Si(30 nm)/Si0.7Ge0.3 samples after annealing at 900 °C for 30 s by RTA.
The scheme of an interposing a-Si layer was also attempted to overcome the detrimental effects of Ge segregation as well as island formation in Ni/Si-Ge systems. The thicknesses of Ni and amorphous Si were selected to be 15 and 27 nm, respectively. GIXRD spectra for Ni(15 nm)/a-Si(27 nm)/Si0\textsubscript{0.7}Ge0\textsubscript{0.3} samples annealed at 300–700 °C indicated that low-resistivity NiSi was the only phase present in all samples annealed above 500 °C.

Figure 34 shows the sheet resistance data for Ni(30 nm)/Si0\textsubscript{0.7}Ge0\textsubscript{0.3} and Ni(15 nm)/a-Si(27 nm)/Si0\textsubscript{0.7}Ge0\textsubscript{0.3} samples (samples C and D). After annealing at 400 °C, low-resistivity NiSi phase was formed in both sets of samples. The apparent increase in sheet resistance for Ni(30 nm)/Si0\textsubscript{0.7}Ge0\textsubscript{0.3} samples after annealing at 700 °C results from the formation of the island structure of Ni(Si0\textsubscript{0.7}Ge0\textsubscript{1}). On the other hand, even after annealing at 800 °C, the sheet resistance remained at the same low level for the Ni(15 nm)/a-Si(27 nm)/Si0\textsubscript{0.7}Ge0\textsubscript{0.3} samples. Furthermore, the sacrificial a-Si layer was also seen to improve the interface roughness and thermal stability of NiSi film grown on Si0\textsubscript{0.7}Ge0\textsubscript{0.3} from XTEM micrographs. In addition, no Ge segregation was observed. An example is shown in Figure 35. In the Ni thin films on the Si system, NiSi\textsubscript{2} is normally formed at 750 °C. The extraordinary stability of NiSi on Si0\textsubscript{0.7}Ge0\textsubscript{0.3} with a sacrificial a-Si layer is likely due to the high consumption ratio of the Si to Ni to form NiSi\textsubscript{2}, which requires the excess Ni atoms to react directly with the Si\textsubscript{1-x}Ge\textsubscript{x} layer. Ge was then expelled from the Ni-Si-Ge compound and slowed down the reaction. The blocking of Ni diffusion paths by Ge atoms may also slow down the reaction and result in a delay of the formation of NiSi\textsubscript{2}. The mechanism is similar to that of the formation of CoSi\textsubscript{2} on a Si1\textsubscript{−x}Ge\textsubscript{x} layer. In contrast, if the a-Si is too thin, as in Ni(15 nm)/a-Si(22 nm)/Si0\textsubscript{0.7}Ge0\textsubscript{0.3} samples, a rough interface is formed. The results indicate that a sacrificial a-Si layer with appropriate thickness works beneficially for Ni contacts on Si-Ge.

**GLOSSARY**

**CMP**  Chemical-mechanical polish. A chemical-mechanical method utilized to eliminate the surface roughness of a sample. This method is commonly used in ULSI processes.

**CVD**  Chemical vapor deposition. A method utilizing chemical gas to deposit film on substrates.

**Dislocation**  A kind of line defect due to the lattice misfit between SiGe and Si. There are two segments. The segment parallel to the Si/SiGe interface is called “misfit dislocation,” and the segment propagating to the surface is called “threading dislocation.”

**HBT**  Heterojunction bipolar transistors. The narrow bandgap base in the bipolar transistor to enhance the device performance.

**Metal germanosilicide**  A compound of metal, Si and Ge.

**SGOI**  Strained Si-on-SiGe-on-insulator. Introducing strained Si to the silicon-on-insulator (SOI) technology, which provides high-performance CMOS circuits due to combination of carrier mobility enhancement in strained Si with the advantage of SOI device/circuit.

**Silicidation process**  The process used to form silicide.

**Strained Si FET**  Strained Si field effect transistors. The tensile strained Si has higher carrier mobility than conventional Si. The new type of device can give larger current drive in VLSI circuits.
Universal mobility  The carrier mobility in the field effect transistor depending only on effective transverse electrical field regardless of the doping concentration of the channel.

ACKNOWLEDGMENTS

The help of S. T. Chang (Department of Electrical Engineering, National Taiwan University, Taiwan) and W. W. Wu (Department of Materials Science and Engineering, National Tsing Hua University, Taiwan) in manuscript preparation are gratefully appreciated. The work on the composite quantum dots by Dr. P. S. Chen, ERSO/TIRI, Taiwan, is highly appreciated.

REFERENCES
