Silicon Photonics 矽光子學 Fabrication of Silicon Waveguide[™]Devices (B)

> 課程編號:941 U0460 科目名稱:矽光子學 授課教師:黃鼎偉 時間地點:一678 明達館 303

Outline

- 5.3 OXIDATION
- 5.4 FORMATION OF SUB MICRON SILICON WAVEGUIDES
- 5.5 SILICON DOPING
- 5.6 METALLIZATION
- 5.7 SUMMARY

5.3 OXIDATION

OXIDATION

- The ability to easily form a high-quality, stable oxide is arguably the most important reason as to why silicon became the predominant material used in the microelectronics industry.
- SiO₂ has many applications in electrical IC fabrication, such as electrical isolation and implantation masking (see section 5.5).
- The deposition or growth of a SiO₂ layer on an SOI wafer also converts the slab waveguide to a symmetrical from an asymmetrical waveguide.
- In order to provide sufficient optical confinement, and for the SiO₂ refractive index to dominate the cladding properties of the waveguide, it is necessary to prevent penetration of the electric field beyond the cladding layer.

OXIDATION

- In the case of SiO₂ on silicon waveguides of several microns in dimension, the oxide thickness layer should be > 0.4 µ m.
- The most suitable method to achieve SiO₂ cladding layers on silicon waveguides is via thermal oxidation.
- This process not only guarantees a conformal, high-quality oxide suitable for integration with any electrical components on the same chip, but has the benefit of smoothing silicon surface roughness (either intrinsic or process-induced).

OXIDATION

- Thermal SiO₂ growth proceeds via a relatively simple chemical process:
- For dry oxidation where the process gas is oxygen

 $Si + O_2 \rightarrow SiO_2$

For wet oxidation where the process gas is steam.

 $Si + 2H_2O \rightarrow SiO_2 + H_2$

Silicon oxidation takes place in a quartz furnace in which the atmosphere and flow rate of the process gas are carefully controlled. Typical growth temperatures are in the range 750-1100°C.

OXIDATION

Thermal oxidation of silicon is well described by the general equation developed by Deal and Grove :

 $2d_0/A = \{1 + [(t+\tau)/(A^2/4B)]\}^{0.5} - 1$

where d_0 is the SiO₂ thickness, *t* is oxidation time, and *A*, *B* and τ are constants related to the reaction rate of oxygen at the silicon surface, the diffusion of the O₂ molecule in the oxide, and a shift in the time coordinate to account for the presence of an initial oxide, respectively.

OXIDATION

- These three constants are determined experimentally for specific growth conditions. As an indication, for a dry oxidation at 1000 °C and a pressure of 760 Torr, values of $A = 0.165 \,\mu$ m, $B = 0.012 \,\mu$ m²/h and $\tau = 0.37$ hours give reasonable approximations to real oxide growth.
- Directly from equation 5.3 one can infer a reaction-rate limited pro-cess during the initial stages of the oxidation process where (t + τ) << (A²/4B).
- This is related to the arrival of oxygen molecules to the wafer surface.
- For longer times, $t \gg \tau$ and $t \gg (A^2/4B)$, the process becomes limited by the diffusion of the oxygen or water molecules through the already present oxide layer before a surface reaction can take place.

OXIDATION



Figure 5.10 SiO₂ thickness versus oxidation time as derived from the Deal and Grove model. The process temperature is 1000° C.

OXIDATION

- Most striking is the difference between wet and dry oxidation. At 1000°C, the oxidation rate is approximately 20 times faster in a wet atmosphere.
- This is a direct result of the higher equilibrium concentration (by nearly three orders of magnitude) of H₂O compared to O₂ in SiO₂.
- Other factors affecting the oxidation rate are gas pressure, wafer crystal orientation, wafer dopant concentration, and wafer crystal damage (e.g. following ion implantation).

5.4 FORMATION OF SUB MICRON SILICON WAVEGUIDES (Silicon photonic wires)

FORMATION OF SUB MICRON SILICON WAVEGUIDES

- There is considerable interest in the fabrication of submicron devices which permit high device packing density, low-cost production, and high-yield.
- As well as more conventional use in the telecommunication network, such small waveguides also have applications in on-chip communication and device interconnects.
- Much of the process flow in fabricating low-dimensional waveguides is similar to that for the larger cross-section devices.
- However, the tolerances permitted in processing smaller devices scale with the device dimension, creating significant challenges for the process engineer.

Silicon Dioxide Thickness Considering Substrate Radiation

- Depends on the modes, polarization and wavelength
- For 1.3-1.6 μ m, Buried Oxide should > 0.4 μ m.
- Thinner core ⇒ thicker BOX layer



Figure 4.10 Buried oxide layer thickness vs planar SOI waveguide thickness for achieving \sim 0.001 dB/cm loss for the fundamental mode at the wavelength of 1550nm

Silicon Dioxide Thickness

- The maximum thickness for the buried SiO₂ layer in SOI material fabricated using the SIMOX process is approximately 0.5 μ m.
- Wafer-bonded SOI is the most viable option for the fabrication of submicron waveguides.
- SOI formed by wafer bonding with a buried SiO₂ thickness up to 3 μ m are commonly available.
- Of the two wafer-bonding techniques outlined in section 5.1, only UniBond® wafers possess the uniformity required for low-dimensional device fabrication.
- The range of silicon overlayer thickness that is available using the SmartCut® process (up to 1.5 µ m) provides the additional benefit of starting material with no requirement for additional epitaxial growth.

Silicon Dioxide Thickness

- Using thermal oxidation to produce a silicon surface oxide (upper cladding) of > 1 µ m is unreasonable, particularly in the case of low-dimensional silicon.
- It is likely that small waveguides will require CVDdeposited SiO₂ as an upper cladding layer.
- The principles of CVD were dealt with in 5.1.4 in relation to epitaxial silicon growth.
- High-quality conformal dielectric layers can be produced in an identical way with the use of a suitable processing precursor gas.
- Because CVD is a deposition process, there is no consumption of silicon and hence no reduction in silicon dimension.

Surface and Interface Roughness

As the thickness of the silicon overlayer is reduced, interface scattering loss increases for equivalent micro-roughness.

$$\alpha_{\rm s} = \frac{\cos^3\theta}{2\sin\theta} \left(\frac{4\pi n_1(\sigma_{\rm u}^2 + \sigma_{\ell}^2)^{\frac{1}{2}}}{\lambda_0}\right)^2 \left(\frac{1}{b + \frac{1}{k_{\rm yu}} + \frac{1}{k_{\rm y\ell}}}\right)$$

- *σ_u* and *σ_l*: surface roughness of upper and lower surfaces, respectively
- k_{yu} and k_{yl}: decay constant of upper and lower surface, respectively
- h: waveguide thickness

Surface and Interface Roughness

- The silicon roadmap indicates surface microroughness for starting material of <0.2 nm and this holds for SOI wafers formed via either the SIMOX or the SmartCut® process.
- The interface roughness between the silicon overlayer and the buried SiO₂ has been reported to be below 1 nm for SIMOX, while UniBond® wafers, in principle, can exhibit interface roughness comparable to that found between a thermally grown SiO₂ on a silicon substrate.
- The roughness of a silicon surface which has undergone a dry etch process may be increased by almost two orders of magnitude.

Surface and Interface Roughness



Figure 5.11 SOI rib formed via dry plasma etching. Etch-induced roughness is visible as the dark pitted regions each side of the rib.

Surface and Interface Roughness

- Thermal oxidation of silicon can be used as a method for the reduction of process-induced surface roughness.
- Specifically, it has been demonstrated that growing SiO₂ to a thickness > 100 nm will ensure an r.m.s. of approximately 0.3 nm.
- This has been demonstrated for surfaces with initial roughness values of 3 nm, and may be appropriate for smoothing surfaces of greater roughness.
- Of course, the use of thermal oxide for smoothing lowdimensional waveguides seems to oppose the deposition of thick CVD SiO₂ for waveguide upper cladding.
- However, the thermal oxide can be either sacrificial (i.e. removed by wet chemical etch before CVD deposition), or can form the initial layer of a thermal CVD SiO₂ stack.

Sidewall Roughness

- Lee et al. [22] recently investigated silicon waveguides with a height of 0.2 μ m and widths varying from 0.5 to 8 μ m.
- The ribs were formed using reactive ion etch and were bounded by a buried oxide of 1 μ m and a deposited upper cladding oxide of 0.2 μ m.
- A clear relationship between optical transmission loss and waveguide width was demonstrated and attributed to the increasing influence of sidewall roughness with reduction in waveguide dimension.
- The sidewall roughness of the etched waveguides was 9 nm needing to be reduced to 0.5 nm for the lowest dimensional waveguides to exhibit loss values of <0.1 dB/cm.</p>
- It is suggested the use of a wet chemical etch as opposed to dry etch to produce waveguides with low r.m.s. roughness.
- It is also possible that a thermally grown oxide would dramatically reduce the propagation loss by smoothing the sidewall roughness.

5.5 SILICON DOPING

SILICON DOPING

- Designs of *p-i-n* diodes rely on the fact that for low-loss waveguiding, the silicon overlayer must be practically free of dopant material.
- By selective introduction of doped regions close to the waveguide (but far enough away to ensure no optical absorption), a monolithically integrated diode allows the controlled injection of



carriers such that they interact with the optical signal under the rib.

Figure 5.12 Silicon waveguide with a monolithically integrated *p-i-n* diode

Ion Implantation

- For low-loss propagation of 1.55 μ m light in silicon the concentration of electrically active impurities such as phosphorus and boron must be <1 x 10¹⁶cm⁻³. This corresponds to a resistivity value> 10 Ω·cm.
- Wafers are readily available at resistivities of several tens of ohm·cm, including SOI fabricated via SIMOX and SmartCut®.
- Silicon epitaxial growth produces layers that are practically intrinsic (resistivity of hundreds of ohm·cm).

Ion Implantation

- To produce electrical and optical functionality in the same device and hence create a silicon photonic integrated circuit we are required to selectively reduce the resistivity of the silicon wafer.
- This is achieved by the introduction of group III atoms (e.g. boron for *p-type* behavior) and group V atoms (e.g. arsenic or phosphorus for *n-type* behavior) into the silicon crystal lattice, in a process known as *electrical doping*.
- As a consequence, the group III and V impurities are referred to as *dopants*. For successful electrical device fabrication the concentration and distribution of the dopants must be tightly controlled.
- Importantly, the introduction of the dopant must be restricted to volumes that do not form any significant part of the optical waveguide (i.e. there must be minimal unintentional, passive interaction of the dopant and the optical signal).

Ion Implantation

- During the first few decades of semiconductor device fabrication doping was performed via the deposition onto, and subsequent diffusion of impurity into, a silicon wafer in a raised temperature furnace rather like that used for thermal SiO₂ growth.
- Typical source gases might be arsine (AsH₃) or diborane (B₂H₆).
- The volume of silicon doped was controlled by the use of windows etched into overlayers of SiO₂, so called masking (the diffusion of dopants in oxide being several of orders of magnitude less than in silicon).
- To ensure a repeatable process the concentration of the dopant was determined in the deposition stage by its solid solubility.
- This severely restricted the concentration profiles and range in device performance that could be obtained.

Ion Implantation

- In the 1970s, ion implantation was developed as an alternative doping technology. Its main advantages are related to precise control of the delivered dopant purity and dose and flexibility in the concentration and profile that can be achieved.
- Further, because implantation is a low-temperature process it is compatible with masks with poor thermal stability, such as photoresist, removing the need to mask the doping process with a SiO₂ overlayer.
- A photoresist mask suitable for the production of a *p-i-n* diode integrated with a silicon rib is shown in Figure 5.13.
- In this example, a thermal SiO₂ upper cladding is in place and the ions are selectively implanted through the oxide and directly into the silicon wafer.

Ion Implantation



Figure 5.13 For a real selective implantation, doping windows are opened in an implantation mask allowing the passage of the ion beam into the silicon device

The Implantation System

- A significant drawback of ion implantation is the complexity and relative high cost of ownership of the implantation system, although these are outweighed by its ability to deliver precisely the vast range of doping profiles required for the fabrication of the most complex device structures.
- An ion implanter has four major subsystems: (1) ion source, (2) magnetic analyzer, (3) accelerator, and (4) process chamber, shown schematically in Figure 5.14.
- Doping by ion implantation, not surprisingly, requires the dopant impurities to be in the form of ions.
- Source material such as boron trifluoride gas or BF₃ is broken down into a variety of species including BF₂⁺, BF⁺, F⁺ and B⁺ in the ion source.
- This is achieved by the creation of localized plasma driven by the acceleration of electrons provided by a tungsten filament.

The Implantation System



Figure 5.14 Schematic of an ion implantation system

The Implantation System

- All positive ion species generated in the source are extracted through a small aperture, and this beam of charged particles is subjected to mass selection using an analyzing magnetic.
- The magnetic field (B) is positioned perpendicular to the direction of beam propagation. Under the influence of B, the ions that make up the beam are deflected to a degree dependent on their velocity, and mass.
- The heavier ions, such as BF₂⁺, are deflected less than the lighter ions, such as B⁺. With careful adjustment of B, ions of any particular mass can be selected from the analyzing magnetic subsystem through an aperture, and into the acceleration subsystem.
- Upon entering the accelerator, the desired ions are subjected to an electric field, the magnitude of which determines their final energy and hence ultimate depth in the silicon wafer.

The Implantation System

- Various machine designs exist which extend the range of ion energies from 200 keV to > 1 MeV, but the most common tools reliably implant ions with energies ranging from sub-keV to a maximum of 200 keV.
- The process chamber houses the silicon wafer to be implanted.
- Uniform exposure of the wafer to the ion beam, which may only be a few square centimeters in diameter, requires relative beam scanning.
- There are two approaches:
 - mechanical scanning of the wafer with a stationary beam, or
 - electrostatic scanning of the beam with the wafer remaining stationary.
 - The former is predominantly used for high-current, high-throughput machines, while the latter is commonly found in low- to medium-current, single-wafer tools.
- In-situ dose measurement is performed in real time by allowing the beam to be overscanned (i.e. beyond the diameter of the wafer) and impinge onto a Faraday cup where a representative beam current is measured.
- The beam dosimetery may be calibrated later by measuring the resistivity of the implanted wafer.

Implantation Parameters

The ion dose (Q) is the number of ions implanted per unit area (usually measured in cm⁻²) and is calculated using:

Q = I t/e n A

where *I* is the measured current, *t* is the time of implant, *e* is the electronic charge, *n* is the charge state of the ion (for B⁺, n = 1), and *A* is the area of the charge collector in square centimeters. The ability to monitor or measure *I*, *t* and *A* to an accuracy of <1 % leads directly to repeatable dose control for the implantation process.

Implantation Parameters

- Each implanted ion will impinge on the wafer with the same energy.
- However, the subsequent slowing down process is randomly depen-dent on ion interactions with the constituent free electrons and silicon atoms (discussed in more detail below).
- The average path length of the implanted ions is called the *range* (*R*).
- *R* has components in both the vertical and horizontal directions.
- Of significance to device designers is the average distance travelled in the direction perpendicular to the wafer surface, and this is called the *projected range* (*Rp*), with the thick-ness of the implanted distribution called the *ion straggle* or *LIRp*.
- The projected range is a function of the ion energy and mass.

Implantation Parameters

- The distribution of implanted ions can be approximated analytically by a gaussian, while more complex statistical distributions give better approximations.
- Numerical solutions for the ion distribution can be obtained also from Monte Carlo simulations and one of the most popular is the freely available TRansport of lons in Matter (TRIM) code.
- It is clear that as the implantation energy is increased both the projected range and *LIRp* increases.
- Also, the projected ranges associated with these typical energies are between 0.15 and 0.6 μm.
- The slowing down of the energetic ions in the silicon can be described by two general mechanisms: electronic stopping and nuclear stopping.

Implantation Parameters



Figure 5.15 Monte Carlo simulation of the distribution of boron ions following implantation into silicon for 50,100 and 150 keV implantation energy. The dose for each energy is 1×10^{15} cm⁻²

Implantation Parameters

- Electronic stopping occurs as the ions interact with the target electrons and creates no permanent damage in the silicon wafer.
- However, it will cause the ejection of electrons when they are situated within a few tens of nanometers of the wafer surface.
- This can result in severe wafer charging especially when the substrate is insulted from the outside world (as is the case for SOI wafers).
- Charging effects can be eliminated during the implantation process by providing a steady supply of low-energy electrons to the wafer surface extracted from a nearby remote gas plasma.

Implantation Parameters

- Nuclear stopping describes the collision of the implanted ions with the silicon atoms.
- These collisions cause wafer atoms to be displaced from their lattice site creating a lattice vacancy and a silicon interstitial (i.e. an atom which resides off the crystalline lattice).
- Although many of these defects 'repair' themselves at room temperature, approximately 5 % of the primary vacancies and interstitials form more complex defects that are stable at room temperature.

Dopant Activation and Drive-in

- A high-temperature anneal (> 900°C) will reconstruct the silicon lattice to its crystalline condition and simultaneously position the dopant ions onto lattice sites.
- This latter process is known as *electrical activation* and is the final step in the doping process.
- The anneal also causes the dopant atoms to diffuse.
- This can be useful if the dopant distribution is required to be broader, and the **Rp** deeper, than is available from the as-implanted profile.
- This post-implant redistribution of dopant is referred to as "drive-in" and typically requires a furnace anneal at temperatures in excess of 1000 °C for many minutes.
- Alternatively, there may be a requirement for the profile of the as-implanted ions to be retained as intact as possible. In this case, the electrical activation is achieved by a so-called "rapid thermal anneal (RTA)" (> 1000 °C for a few seconds).

Dopant Activation and Drive-in

- In the formation of the *p-i-n* structure shown in Figure 5.12 and the device structures described in section 6.1.3 of Chapter 6, the *p-type* and *n-type* regions are heavily doped in a similar manner to the source and drain contacts in CMOS technology (i.e. > 1 x 10¹⁵ cm⁻²).
- For doped regions on each side of the rib, it is necessary to prevent the attenuation of the optical mode in the device "off" state by opening the implant windows far enough away from the rib walls. In addition, consideration must be made for lateral thermal diffusion during the high-temperature anneal.
- For doped regions in the top of the rib (such as that shown in Figure 6.2), the electrical activation should proceed via a rapid thermal anneal to prevent any significant diffusion of the implanted dopant.

5.6 METALLIZATION

METALLIZATION

- The final step to the formation of an integrated silicon photonic device requires the pattern definition of metal on the silicon wafer for device interconnects and access to the device from the outside world.
- Similar caution is required with metallization as with doping, in that none of the metal and any part of the optical mode should interact, or else significant absorption will result.
- In practical terms this demands sufficient upper-cladding oxide thickness to prevent optical power decay beyond the oxide and into the metal layer.

Via Formation

- The resistivity of SiO₂ is extremely high (approximately 10¹⁵ Ω·cm) and hence the upper-cladding SiO₂ provides excellent electrical insulation between the metal/silicon and metal/metal in the case of multilevel metallization.
- If we again refer to our simple *p-i-n* structure in Figure 5.12, we observe the requirement for intimate contact only between the metal and the heavily doped *p* and *n* regions and hence the selective removal of SiO₂ from the doped silicon surface.
- Metal contact with undoped regions of the wafer surface is undesirable as it will result in current leakage and parasitic capacitance effects.

Via Formation

- Using photolithography, photoresist windows are formed inside the doped area.
- The SiO₂ is then etched away from the wafer surface using a hydrofluoric-acid (HF) based wet etch.
- The resulting SiO₂ windows are referred to as contact "vias".
- Metal film deposition is performed within quick succession of the removal of the surface oxide.
- This ensures the minimum regrowth of "native" SiO₂ following exposure to the atmosphere.

Thermal Evaporation

- A traditional method for depositing metal which is still used in small-scale research and development is thermal evaporation.
- The metal to be deposited is placed in a crucible inside a vacuum chamber.
- The chamber is evacuated to around 10⁻⁶ Torr and the metal source is heated by an electron beam until vaporization occurs.
- The wafer is placed far enough away from the source to ensure a uniform deposited thickness across the wafer area.
- Thermal evaporation provides a robust and inexpensive method for silicon metallization, but significant drawbacks are its inability to provide uniform step coverage of etched features and the requirement for each metal type to have a dedicated crucible.

Sputtering

- A more common method for metal deposition is *sputtering*.
- The sputtering technique users energetic, inert gas ions to strike a source-metal target and physically dislodge metallic atoms - rather like a cannon-ball striking a muddy field.
- The sputtered atoms are allowed to migrate through a vacuum until they are deposited on a target wafer.
- The inert gas (usually argon) is ionized inside an evacuated plasma chamber in a similar process to that described in section 5.2.2.

RF Sputtering

- The use of a 13.56-MHz power source has resulted in this technique being referred to as *RF sputtering*.
- The RF bias is applied to the back surface of the metal target.
- The formation of plasma and creation of the plasma potential attracts positive argon ions to the target where they dislodge target atoms.
- These atoms then migrate towards the wafer surface where they form the metal layer.
- RF sputtering suffers from a relatively poor sputtering yield and hence slow deposition rate.

RF Sputtering



Figure 5.16 Schematic of an RF sputtering system

Magnetron Sputtering

- Using magnetic confinement (magnetron sputtering), electrons close to the target are prevented from escaping from the discharge region.
- These electrons instead contribute to the ionization rate of the plasma, and subsequent bombardment of the target metal, significantly increasing the sputter yield and deposition rate.

Material Choice

- The choice of metal to be deposited depends on many parameters, such as available deposition tools, the specific application of the metallic layer and the reliability requirements for the device.
- In general though, semiconductor/metal contacts are required to have negligible resistance compared to the device itself, so-called "ohmic" contacts, while inter-connects between devices must have as lowa resistance as possible.
- The traditional "work-horse" metal of the microelectronics industry is aluminum. It has low room-temperature resistivity (2.5 μ Ω·cm), is rel-atively inexpensive, is easily deposited, and reacts with both silicon and SiO₂ ensuring excellent adhesion to device layers.

Material Choice

- In addition, aluminum can be patterned and etched in a straightforward manner to form the complex interconnects and contacts required in IC fabrication.
- Although aluminum is reaching the limits of use in deep submicron technology, it is worth mentioning that for prototype device development in academic or small-scale industrial operations, and devices where shallow doped regions are not required, aluminum is still an excellent choice for general-purpose metallization.

Material Choice

- In recent years copper has emerged as the replacement for aluminum in deep submicron interconnect fabrication.
- The main reason for this is its resistivity of only 1.7 μ Ω·cm increasing speed while reducing the power required to drive a device (alternatively increasing the packaging density for the same power).
- Processing with copper is not straightforward. It diffuses quickly in Si and SiO₂, is resistant to many wet and dry etch processes making pattern definition difficult, and unlike aluminum it oxidizes when exposed to atmospheric ambient.

Material Choice

- Refractory metals (such as titanium, cobalt and molybdenum) react with silicon to form silicide when heated to temperatures ranging from 900 to 1400°C, depending on the silicide material.
- The resulting com-pound has significantly greater thermal stability compared to aluminum.
- With a resistivity of tens of μΩ·cm, silicide is suitable for forming the contact between metal layers and the silicon surface in modern submicron devices

Sintering and Barrier Materials

- The formation of a robust ohmic contact between aluminum and silicon necessitates heating the interface to a temperature of around 500-550°C for several minutes in an inert or reducing ambient, in a process known as *sintering*.
- However, the use of pure aluminum can result in significant silicon diffusion into the aluminum layer during the sintering step.
- The resulting interface structure resembles the formation of aluminum spikes into the underlying silicon which may penetrate the doped junction (particularly if it is submicron in depth).
- This phenomenon is known as "junction spiking".

Sintering and Barrier Materials

- The most reliable method for the elimination of junction spiking is the incorporation of a diffusion barrier layer between the aluminum and the silicon.
- This thin layer must be able to prevent significant diffusion at the sinter temperature, have low electrical resistivity and possess good adhesion properties to both the Si and AI (or whatever metal is being used as a contact).
- Two of the most common barrier layers are titanium tungsten (TiW) and titanium nitride (TiN), although a thin layer of TiSi₂ is required below TiN to reduce the contact resistance.

5.7 SUMMARY

SUMMARY

- The development of silicon photonic fabrication benefits greatly from the vast library of knowledge that already exists in the silicon microelectronics industry.
- We concede that as with any fast-moving technology, silicon photonic processing will find a path of least resistance which may make some of the processes described here redundant and necessitate the use of methods not described (for instance in the inevitable development of sub micron devices).
- However, the overall process flow and much of the detail of this chapter will remain relevant for many years to come.