

Oxidation (of Silicon)

Che-Hsin Lin*

Department of Mechanical and Electromechanical Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan

Synonyms

[Oxide deposition](#); [Oxide growth](#); [Silicon dioxide growth](#); [Thermal oxidation](#)

Definition

Oxidation of silicon to create a silicon dioxide layer on the silicon surface can be done either by the thermal growth methods or the undesired native growth process. The process is an adding procedure which adds oxygen to react with silicon to form silicon dioxide on the silicon surface. The process consumes silicon during the oxidation procedure. In general, the oxide layer can be grown either by wet or dry oxidation processes in a furnace at a temperature of 750–1,100 °C. The reaction can be expressed as $\text{Si}_{(\text{s})} + \text{O}_{2(\text{g})} \rightarrow \text{SiO}_{2(\text{s})}$.

Overview

Silicon dioxide is a form of pure glass with a melting temperature of 1,730 °C which is similar to fused silica. It exhibits excellent dielectric properties and has a good adhesion to silicon such that silicon dioxide is ideal to be used as a dielectric or barrier layer. Silicon dioxide is also very useful in the MEMS and semiconductor industry due to its excellent material properties. A grown oxide layer is conformal on the silicon surface, which makes it an ideal surface coating or passivation layer for MEMS structures. The electric resistivity is up to $1 \times 10^{20} \Omega \text{ cm}$ and the energy band gap is around 9 eV. A well-grown silicon oxide layer can sustain a high electric field of up to 10 V/m without breakdown [1]. Furthermore, the ease of formation and its excellent interface with the underlying silicon material make it the most popular thin film in VLSI technology. Essentially, silicon is reactive to oxygen at room temperature such that most silicon exists in the form of silicon dioxide like quartz sand in nature. In this regard, oxidation of silicon occurs immediately to form amorphous silicon dioxide film by exposing a silicon surface to oxygen in atmospheric environment. Silicon dioxide is a dense material that fully covers the silicon wafer to prevent the inner silicon from further oxidation. The thickness of this native oxide will reach to an upper limit of about 40 Å even over a long period of time. The native oxide layer grown from air is not uniform and is usually considered a contaminant in semiconductor processing. This native oxide can be simply removed in seconds using dilute HF solution.

Alternatively, a thermally grown oxide layer is thickness controllable and has various applications in the semiconductor industry. The thickness ranges for thermal oxide layers may vary, depending on the particular application. For example, the thickness for an oxide capacitor dielectric is as thin as several angstroms but the thickness of a field oxide layer can be up to several micrometers. Since the

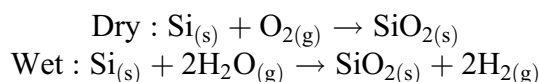
*Email: chehsin@mail.nsysu.edu.tw

grown oxide layer will fully cover the silicon surface to prevent it from further oxidation, it is common to elevate the temperature to drive the oxidant molecules to diffuse across the formed oxide layer to continue the oxidation process. In general, an oven or a furnace with high-purity oxygen supplied is usually adopted for the oxidation process. The oxidant can be either pure oxygen gas or water vapor. If the growth procedure occurs with dry oxygen, then the oxidation process is called *dry oxidation*. If the oxygen is introduced to the reaction chamber with water vapor, then the oxidation process is known as *wet oxidation*. The reaction mechanism of these two oxidation approaches is different such that the growth rate and the grown film quality of these two approaches are also different. In general, the wet oxidation process is faster than the dry oxidation process, but the density of the grown oxide layer obtained is lower when using the wet oxidation process. Higher temperature will significantly increase the growth rate of the oxide layer. The equipment for performing the oxidation processes, a mathematical model for evaluating the growth rate, and a well-known pattern of bird's beak effect in local oxidation of silicon (LOCOS) process are described in this entry.

Basic Methodology

Oxidation Modeling

Silicon dioxide is usually grown by consuming silicon during oxidation process at high temperature from 750 °C to 1,100 °C. Figure 1 shows the basic process for the oxidation of silicon. For successive oxidation, oxidants must diffuse across the formed oxide layer and then react with silicon at the SiO₂/Si interface. The chemical reactions for both dry and wet oxidation processes can be described as follows:



Apparently, silicon is consumed to form silicon dioxide in both dry and wet oxidation processes. The thickness of the silicon consumed is around 0.44 of the total oxide thickness. This value can be simply calculated using the following equation:

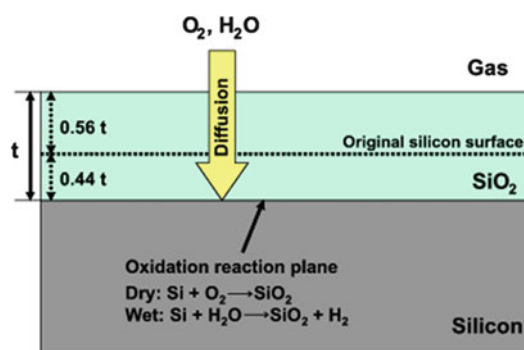


Fig. 1 The basic process for the oxidation of silicon

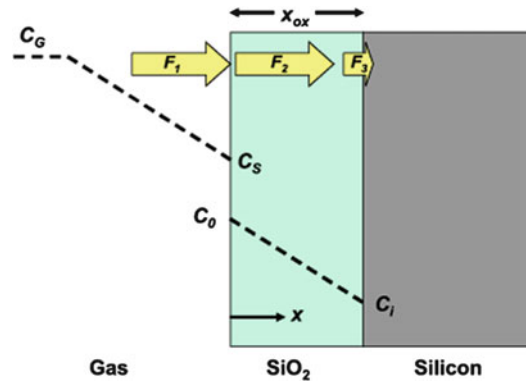


Fig. 2 Flux model for oxidant diffusion from the gas phase to silicon surface during thermal oxidation

$$\frac{t_{\text{Si}}}{t_{\text{SiO}_2}} = \frac{V_{\text{Si/mole}}}{V_{\text{SiO}_2/\text{mole}}} = \frac{M_{\text{Si}/\rho_{\text{Si}}}}{M_{\text{SiO}_2/\rho_{\text{SiO}_2}}} = \frac{2,809(\text{g/mole})/233(\text{g/cm}^3)}{6,008(\text{g/mole})/221(\text{g/cm}^3)} \approx 0.44$$

where t is the thickness, V is the volume, M is the molecular weight, and ρ is the material density. The growth rate of oxide depends on the parameters for the oxidation process such as temperature, gas pressure, oxidizing condition (wet or dry), silicon crystal orientation, and the dopant concentration in silicon. The model for oxide growth on silicon was first described by Deal and Grove in 1965 [2]. The Deal-Grove model describes two growth stages of the linear stage and the parabolic stage during oxidation. Figure 2 illustrates the one-dimensional model for the oxidant flux from the gas phase to the silicon surface during thermal oxidation.

Three first-order equations describing the three series parts of the sample fluxes, i.e., F_1 , F_2 , and F_3 , in the one-dimensional model are presented in Eqs. 1–3, where F_1 is the gas transport flux in number of molecules per square centimeter per second, F_2 is the diffusion flux through the oxide layer and can be expressed using Fick's first law, and F_3 is the reaction flux at the SiO_2/Si interface:

$$F_1 = h_G(C_G - C_S) \quad (1)$$

$$F_2 = -D \frac{\partial C}{\partial x} = D \left(\frac{C_0 - C_i}{x_0} \right) \quad (2)$$

$$F_3 = k_s C_i \quad (3)$$

where h_G is the mass transfer coefficient (molecules per cm^2s), C_G is the concentration of oxidant, C_S is the concentration at the surface, D is the oxidant diffusivity in the oxide (cm^2/s), C_0 and C_i are the concentrations at the two interfaces, x_0 is the oxide thickness, and k_s is the reaction rate constant (cm/s). The concentration of the oxidant just inside the oxide surface C_0 and C_S can be related using Henry's law:

$$C_0 = H P_S \quad (4)$$

where H is Henry's constant and P_S is the partial pressure of oxidant at the surface. From the ideal gas law, $PV = NkT$, Eq. 4 can be expressed as

$$C_0 = H \times kT \left(\frac{N_s}{V} \right) = H(kT C_s) \quad (5)$$

where the subscript s denotes at surface. Rearranging Eq. 5 will give

$$C_s = \frac{C_0}{HkT} \quad (6)$$

where k is the Boltzmann constant (1.38×10^{-23} J/K) and T is the temperature in K. C^* is defined to be the oxidant concentration in the oxide that would be in equilibrium with the bulk gas pressure P_G , which is known:

$$C^* = HP_G = H(kT C_G) \text{ then } C_G = \frac{C^*}{HkT} \quad (7)$$

Substitution of Eqs. 6 and 7 into Eq. 1 leads to the result that

$$\begin{aligned} F_1 &= h_G \left(\frac{C^*}{HkT} - \frac{C_0}{HkT} \right) \\ &\equiv \frac{h_G}{HkT} (C^* - C_0) \equiv h(C^* - C_0) \end{aligned} \quad (8)$$

where $h = h_G/HkT$, which is a large value because experimental results show that wide changes in gas flow rates in oxidation furnaces, changes in the spacing between wafers on the carrier in the furnace, and changes in the wafer orientation make little difference to oxidation rates. Under steady-state conditions, mass must balance such that the three flux equations representing the oxidation process must be equal. Thus, $F_1 = F_2 = F_3$. Solving the two unknowns C_i and C_0 from the two equations of $F_1 = F_2$ and $F_1 = F_3$, then we have

$$C_i = \frac{C^*}{1 + \frac{k_s}{h} + \frac{k_s x_{ox}}{D}} \quad (9)$$

$$C_0 = C^* \left(1 + \frac{K_s x_{ox}}{D} \right) \quad (10)$$

Substitution of Eq. 9 into Eq. 3 leads to

$$F(= F_1 = F_2 = F_3) = \frac{k_s \cdot C^*}{1 + \frac{k_s}{h} + \frac{k_s x_{ox}}{D}} \quad (11)$$

Let M be the number of molecules of the oxidant in a unit volume of the oxide. The rate of the thickness change of the oxide layer with time is then given by oxidizing flux divided by the number of molecules M of the oxidizing species (O_2 or H_2O) that are incorporated into a unit volume of the resulting oxide:

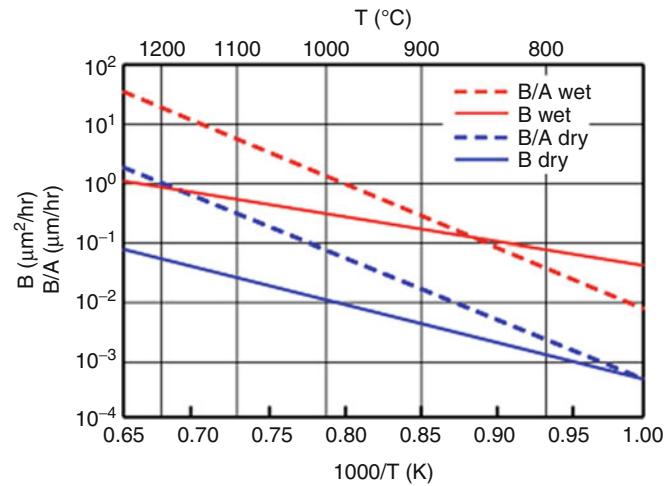


Fig. 3 Plot of B and B/A versus temperature for dry and wet thermal oxidation of $\langle 111 \rangle$ silicon

$$\frac{dx_{\text{ox}}}{dt} = \frac{F}{M} = \frac{k_x \cdot C^*}{M \left(1 + \frac{k_s}{h} + \frac{k_s x_{\text{ox}}}{D} \right)} \quad (12)$$

where $M = 2.2 \times 10^{22} \text{O}_2/\text{cm}^3$ for the dry oxidation process and $M = 4.4 \times 10^{22} \text{H}_2\text{O}/\text{cm}^3$ for the wet oxidation process. Note that $M = \text{density}/(\text{Avogadro's number} \times \text{molecular weight of SiO}_2)$.

Take the boundary condition of $x_{\text{ox}} = x_i$ at $t = 0$. The oxide growth kinetics can be obtained by integrating Eq. 12 from an initial oxide thickness x_i to a final thickness x_{ox} . The result is described as

$$\frac{x_{\text{ox}}^2}{B} + \frac{x_{\text{ox}}}{B/A} = (t + \tau) \quad (13)$$

where

$$A = 2D \left(\frac{1}{k_s} + \frac{1}{h} \right), \quad B = \frac{2DC_A}{M} \quad \text{and} \quad \tau = \frac{x_i^2 + Ax_i}{B}$$

The terms B and B/A can be referred to as the parabolic and linear rate constants for describing the oxide growth kinetics. The physical meanings of the rate constants B and B/A are the oxidant diffusion and interface reaction rate, respectively. The corresponding values for B and B/A for dry and wet oxidation of $\langle 111 \rangle$ silicon are plotted in Fig. 3. For $\langle 100 \rangle$ silicon, these values should be divided by 1.68. These expressions are better fit for the cases of a very thin oxide layer and in dry O_2 conditions. Solving Eq. 13 leads to the relationship between the oxide thickness and the growth time:

$$x_{\text{ox}} = \frac{A}{2} \left\{ \sqrt{1 + \left[\frac{t + \tau}{A^2/4B} \right]} - 1 \right\} \quad (14)$$

Considering two limiting cases of $t \rightarrow 0$ (small x_{ox}) and $t \rightarrow \infty$ (large x_{ox}), we have the following expressions corresponding to Eq. 14:

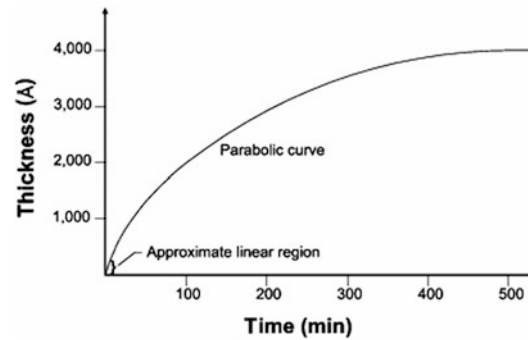


Fig. 4 Linear and parabolic stages for dry oxidation process at 1,100 °C

$$x_{\text{ox}} \rightarrow \frac{B}{A}t \text{ when } t \rightarrow 0$$

and

$$x_{\text{ox}} \rightarrow \sqrt{Bt} \text{ when } t \rightarrow 0$$

It is clear that the thickness of the grown oxide layer varies linearly with time at the initial stage of oxidation. In this stage, the oxide growth rate is controlled by the reaction occurring at the SiO_2/Si interface and is called the *reaction rate-controlled region*. The linear stage of oxide growth is valid up to about 150 Å of oxide thickness [3]. The growth of the oxide layer slows down after the linear stage because the oxidant must diffuse across a longer distance to reach the interface where the oxidation reaction occurs. The equation generates the shape of a parabola such that it is called the parabolic stage. In the parabolic stage, the growth rate of oxide is controlled by the mass transportation of the oxidant and is called the *diffusion-controlled region*. Figure 4 illustrates the general curve for both the linear and parabolic stages for dry oxidation at 1,100 °C [4]. It is noted that the linear stage is only at the very beginning of the thermal oxidation process.

In practical applications, it is not necessary to calculate the thickness of the grown oxide layer using the kinetic model described above. Users can simply obtain the approximate thickness from a chart which plots the oxide thickness for oxidation versus temperature and time, as an example shown in Fig. 5. Another simple way to estimate the thickness of the grown oxide layer is to use the

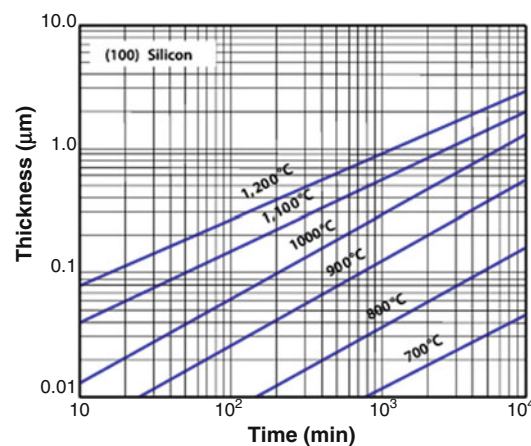


Fig. 5 The thickness of a grown oxide layer for different temperatures and times

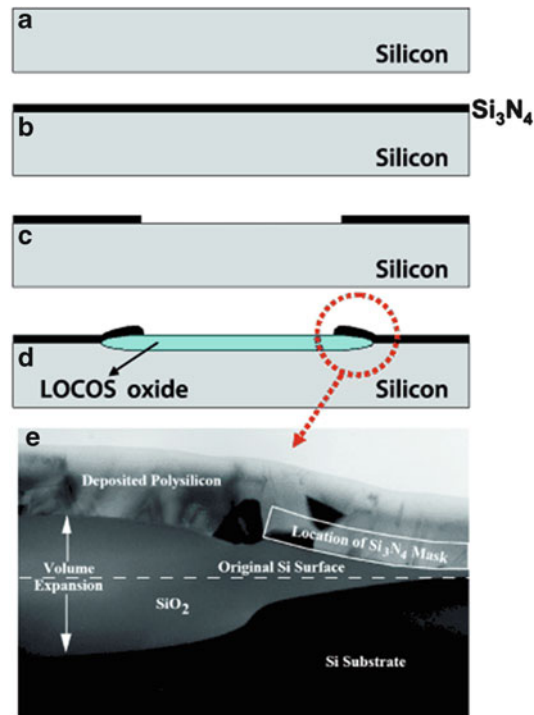


Fig. 6 (a–d) Schematic of a typical LOCOS formation process. (e) SEM image of a LOCOS structure. Note that the silicon nitride layer on the right-hand side was stripped and a polysilicon layer has been deposited on top of the structure to provide better contrast in the image (Image courtesy of J. Bravman, Stanford University)

well-developed simulator available on the internet [5]. In addition, the dopant concentration in silicon is also a factor that influences the growth rate of oxide. Here also a minor parameter which may affect the oxide growth rate is doping during the oxidation process. In general, heavily doped silicon oxidizes at a faster rate than a lightly doped substrate. In the parabolic stage, boron doping will oxidize faster than phosphorus. However, the growth rate shows little difference in the linear rate constant between boron and phosphorus doping [6].

In summary, the growth of thermal oxide follows a linear growth stage up to about 150 Å, followed by a slower parabolic growth stage. The chemical reaction for the thermal oxidation of silicon consumes silicon and oxidant, and the growth rate of the oxide layer is faster when using wet oxidation. Factors including dopant concentration, crystal orientation, pressure, and temperature may affect the oxide growth rate during oxidation.

LOCOS Process

The abbreviation LOCOS stands for *local oxidation of silicon*, in which selective areas on a wafer are oxidized for electric insulation purposes in semiconductor technology. This process is used for isolating two adjacent devices with a feature size of greater than 0.25 μm. Figure 6 presents a simplified schematic for the LOCOS process. A deposited silicon nitride layer is patterned and etched to expose the selective area for the oxidation process. The patterned Si_3N_4 layer serves as the oxidation barrier (Fig. 6a–c). There is a slight lateral growth of the oxide beneath the nitride oxidation barrier because the oxidants move in all directions when diffusing through the grown oxide. Since the oxidation is an adding procedure and the grown oxide is thicker than any silicon consumed, the oxide grown beneath the nitride layer serves to push up the nitride edge (Fig. 6d). The LOCOS process results in a non-planar Si/SiO₂ interface between the nitride oxidation barrier and

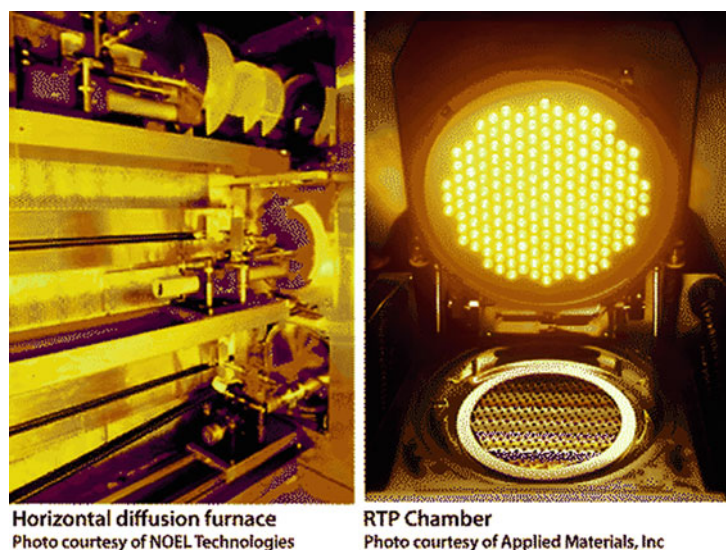


Fig. 7 Images of a conventional horizontal furnace (*left*) and a RTP chamber (*right*)

oxidized regions. This phenomenon is known as the *bird's beak effect* due to its shape. The bird's beak effect is an undesired by-product of LOCOS processes since it increases the area required for building a semiconductor device. Figure 6e illustrates a SEM image showing the cross section of a LOCOS process, the bird's beak shape being clearly observed in the image.

Furnace Equipment and Oxidation Processes

Furnace equipment serves as a workhorse in thermal processing and is among the simplest type of processing equipment in the semiconductor industry. The equipment is conceptually constituted of an oven capable of temperatures from 600 °C to 1,200 °C and a simple gas distribution system capable of introducing the oxygen or water molecule oxidants. Since the growth rate of the oxide layer is highly dependent on the temperature, precise temperature control is essential for oxidation furnaces. Modern furnaces are easy to handle up to several hundred 8 in. wafers, with a temperature uniformity of ± 0.5 °C [7]. There are three basic types of furnace, horizontal furnace, vertical furnace, and rapid thermal processor (RTP), used for thermal processing. Figure 7 shows two images of a conventional horizontal furnace and a RTP chamber. Simple construction and ease of maintenance are the major advantages of the horizontal furnaces such that they are popular in academic institutes and industries which require only small production volumes. The horizontal furnace is so named because of the horizontal position of the quartz tube where wafers are located and heated. This furnace was largely replaced in the early 1990s by the vertical furnace in the semiconductor industry, mainly because the vertical construction is easier to automate, has improved operator safety, and reduces particular contaminations [8]. Oxidation processes operated in the horizontal and vertical furnaces are batch processes, which are capable of processing large quantities of wafers (tens to 200 wafers) in one process cycle. The furnace process is considered a hot-wall procedure such that it requires more thermal budget for heating the furnace.

Figure 8 illustrates schematically the setup of a typical oxidation furnace which shows the conceptual configurations for both dry and wet oxidation processes. The main body of the system is a quartz process tube surrounded by a resistance heating element which can be controlled to produce multiple heating zones. It is common to form three to seven symmetric hot zones over the process tube. The temperature zones thus formed facilitate control of the furnace tube to attain a flat zone near the middle of the tube where the thermal oxidation takes place. The outer zones are

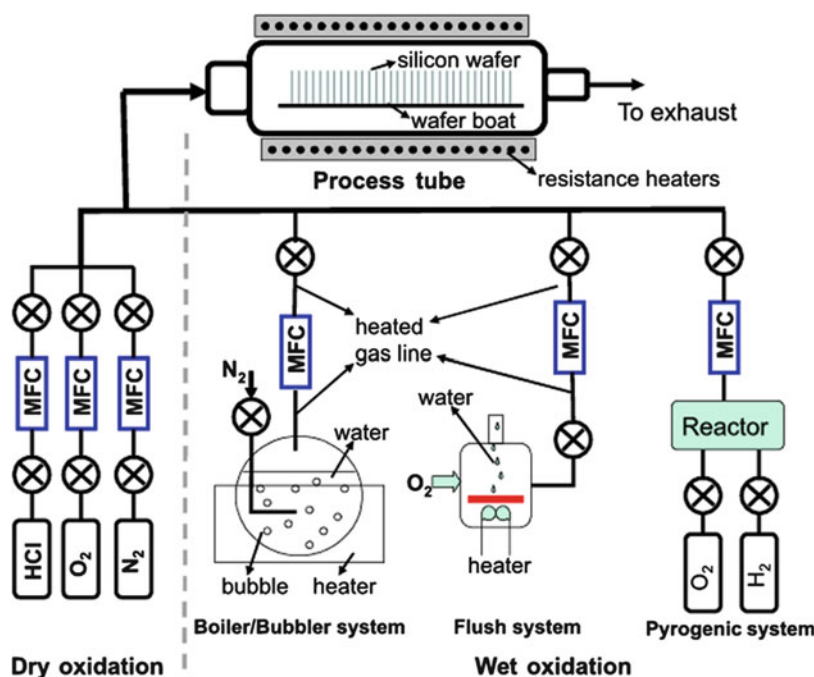


Fig. 8 Schematic for the conceptual construction of an oxidation furnace

designed to help compensate for heat losses out of the ends of the tube. The wafers are normally loaded into the furnace on quartz *boats* which hold 10–50 wafers in parallel. The boats are carried into the center section of the hot zones by the operator or the loading system. The wafers are normally loaded at a moderate temperature (600–800 °C) in order to prevent crystallographic defects caused by large thermal gradients across the wafers; thermal budget for this process can be saved as well. Nowadays, these furnaces are most commonly operated at atmospheric pressure. However, if necessary, one is able to increase the tube pressure to increase the growth rate of oxide film.

Dry Oxidation

The furnace structure for the dry oxidation process is simple and straightforward. Usually, high-purity oxygen gas is used to oxidize silicon. Nitrogen gas in the oxidation system is used as the process gas during system idle, temperature ramping, wafer loading steps, and chamber purging, because nitrogen does not react with silicon at the processing temperature. In dry oxidation, the process gas is usually introduced into the furnace with a trace amount of hydrogen chloride gas. The existence of chloride ions during oxidation is very important because chloride ions react with the mobile metallic ions, such as sodium ions, to form immobile chloride compounds. The undesired oxide charge occurring at the Si/SiO₂ interface can be neutralized by the use of HCl gas. In addition, the existence of hydrogen chloride will also allow some chlorine atoms to be integrated into the oxide film and bind with silicon at the silicon oxide interface. This process minimizes the number of dangling bonds and improves IC device reliability. Another important advantage of adding chloride ions to the thermal oxidation process is that they tend to increase the oxidation growth rate by about 10–15 % [9]. However, the concentration of chlorine ions should not be high (usually <3 %); otherwise the electric properties (e.g., bias voltage) of the electronic device would become unstable. Instead of using toxic and corrosive HCl gas, chlorine compounds such as dichloroethylene (DCE) or trichloroethylene (TCE) can be used to introduce the chlorine ions into the process tube [10]. Oxide films grown by dry oxidation have better quality than films grown by wet oxidation.

Therefore, high-quality thin oxide layers such as gate oxide, screen oxide, or pad oxide are usually grown using the dry oxidation process.

Wet Oxidation

The wet oxidation process uses water molecules instead of dry oxygen as the oxygen source to oxidize silicon. As a matter of fact, water molecules contact the silicon furnace in a normal wet oxidation process. Water molecules dissociate at high temperature and form hydroxide (HO) prior to reaching the silicon surface. Hydroxide has faster diffusion mobility in silicon dioxide than pure O₂, which explains why wet oxidation has a higher growth rate than dry oxidation. Wet oxidation is used to form thick oxides such as the LOCOS oxide, masking oxide, and field oxide. As shown in Fig. 7, several systems have been used to deliver water vapor into the process tube. The boiler system is the simplest setup which vaporizes ultrapure water and then drives the water vapor into the process tube via heated gas lines. However, it is difficult to have precise control of the vapor flow with this method. In addition, the water in the boiler is not *fresh* after a long boiling procedure and contaminants may be dissolved in the water. A modified setup is called a bubbler system which flows nitrogen gas into a hot water bath to generate wet nitrogen bubbles and carry vapor into the process tube. This scheme is the most commonly adopted in academic fields due to its simple structure and more precise flow control than a boiler system. The flush system is an alternative choice for introducing water vapor into the process tube. This system uses small water droplets on a quartz hot plate, which vaporizes the water. The vaporized steam is then carried into the process tube by flowing oxygen gas. However, this method has the problem that the steam generation is not a steady process. In order to have better control of the vapor flow rate, a pyrogenic system called *Dryox* was developed, which burns hydrogen gas to generate water vapor right at the inlet port of the process tube. The vapor flow can be precisely controlled since no liquid and vapor handling is required in this system. However, it needs extreme care to handle this process since hydrogen gas is flammable and explosive. The pyrogenic reaction is operated at a temperature of about 400 °C at which hydrogen automatically reacts with oxygen and forms water vapor. The typical mixing molar ratio for hydrogen and oxygen is 1.8:1 to 1.9:1, in order to completely consume the hydrogen gas [9]. Otherwise, hydrogen can accumulate inside the process tube and cause an explosion. A burn box and a scrubber system for exhaust gas treatment are necessary for operating a pyrogenic system.

Rapid Thermal Process (RTP)

The rapid thermal process (RTP) is a method which heats a single wafer to a temperature range of 400–1,300 °C in a very short time under atmospheric conditions or at low pressure. This process is usually referred to as rapid thermal oxidation (RTO). The RTP has a gas handling system, an optical pyrometer for fast temperature measurement, and a computer for controlling process operation. In order to achieve this fast heating process, a tungsten halogen lamp array located nearly at the top and bottom of the wafer is used as the radiant heating source. The temperature ramp up and down in a furnace is typically about 20 °C/min or less. Alternatively, the RTP can incorporate a fast-heating system which provides ramp-up rates up to 250 °C/min. In order to keep the temperature uniformity across the reaction surface, the RTP usually treats a single wafer at a time. The wafer-to-wafer uniformity of a RTP system is also better than that of a conventional furnace system. The silicon wafers are heated by selectively absorbing radiation from the lamps, which produce short-wavelength radiation. The energy from the lamps is mainly absorbed by the silicon but not the chamber wall such that the RTP is considered as a cold-wall process. In this manner, heat-generated contaminants from the chamber wall can be reduced. The RTP also uses less thermal budget than a furnace system but the throughput is an issue. More importantly, the oxidation process achieved

using the RTP can minimize dopant movement in wafers due to the shorter process time. This advantage makes the RTP ideal for processes like the barrier-layer formation and the oxide reflow process.

Future Directions for Research

Oxidation processes have been studied for more than 50 years and they still attract lots of interests from researchers who are working in this area. For example, in order to prevent gate leakage and dielectric breakdown, growth of thinner and denser dielectric layers with high dielectric constant for gate structures is an ongoing trend. The RTP will play a more important role in the semiconductor process since it offers better process control, less thermal budget, less processing time, and better wafer-to-wafer uniformity. The RTP incorporated with on-line film monitoring, rapid thermal annealing, and even rapid thermal chemical vapor deposition (RTCVD) processes has shown its potential for fabricating high-quality thin films with various material choices. Finally, an automation system to handle the multifunction RTP system is essential for improving the throughput of this system.

Cross-References

► [Thermal Oxidation](#)

References

1. DiMaria DJ, Arnold D, Cartier E (1992) Degradation and breakdown of silicon dioxide films on silicon. *Appl Phys Lett* 61:2329–2331
2. Deal BE, Grove AS (1965) General relationship for the thermal oxidation of silicon. *J Appl Phys* 36:3770–3778
3. Ganem JJ, Battistig G, Rigo S, Trimaille I (1993) A study of the initial stages of the oxidation of silicon using O_{18} and RTP. *Appl Surf Sci* 65(66):647–653
4. Quirk M, Serda J (2001) *Semiconductor manufacturing technology*. Prentice-Hall, New Jersey
5. <http://www.lelandstanfordjunior.com/oxcalcfaq.html>
6. Choi SS, Numan MZ, Chu WK, Irene EA (1987) Anomalous oxidation rate of silicon implanted with very high doses of arsenic. *Appl Phys Lett* 51:1001–1003
7. Plummer JD, Deal MD, Griffin PB (2000) *Silicon VLSI technology: fundamentals, practice and modeling*. Prentice-Hall, Upper Saddle River
8. Singer P (1997) Furnaces evolving to meet diverse thermal processing needs. *Semicond Int* 20:84–88
9. Xiao H (2001) *Introduction to semiconductor manufacturing technology*. Prentice-Hall, Upper Saddle River
10. Higashitani M, Hikazutani K (1997) Semiconductor device manufacturing method including dry oxidation. US Patent 5,637,528