ENHANCEMENT OF PROCESS CONTROL USING REAL-TIME SIMULATION

EWAN A. CAMERON

PhD
University of Edinburgh
1989

ACKNOWLEDGEMENTS

This work was sponsored by the Science and Engineering Research Council (SERC) and Plessey Research (Caswell) Ltd. under the SERC Co-operative Awards in Science and Engineering (CASE) scheme. I am grateful to both bodies for financial assistance. I would like to acknowledge the continued encouragement and support of my academic supervisors Professor J. M. Robertson and Dr. R. Holwill of the Edinburgh Microfabrication Facility (EMF) and my industrial supervisor Mr. H. Richards of Plessey Research. The assistance of the staff and research students at the EMF throughout the project is much appreciated. I would also like to acknowledge the co-operation of the people with whom I had contact at Plessey Research. *Vielen Dank* to Siemens AG for use of text processing facilities and to Dorothy Dickson for her hard work in preparing the text.

ABSTRACT

As semiconductor devices are reduced in size, it becomes increasingly more important to control the critical process parameters that determine device performance. A novel technique is proposed that enhances control by automatically taking into account all major variables encountered during processing. Real-time simulation, supplied with measured process variables, is the basis of the technique. The chosen application is thermal growth of thin silicon dioxide films. Gate oxide thickness is a critical parameter in MOS devices, appearing as a first order term in all principal device equations. Control using conventional means based on a fixed oxidation time becomes increasingly difficult as thickness is scaled to below 250Å for VLSI devices and growth times are correspondingly reduced.

A microprocessor-based system that uses simulated oxide growth to actively control oxide thickness has been developed. Process data available through the conventional furnace control system is utilised, with the advantage that little modification to the existing process equipment is necessary. The system provides closed-loop control and process data storage at the 'local' equipment level, and as such forms the foundation for a highly-automated processing module for oxidation that can be integrated into the automated production environment.

CONTENTS

CHAPTER 1

INTRODUCTION

Integrated Circuit (IC) manufacturing technology has advanced rapidly since the first ICs were produced in the late 1950's. This is most apparant in device geometries, which have shrunk steadily to allow increased levels of integration and improved circuit performance. VLSI (Very Large Scale Integration) processes are currently capable of producing circuit features with minimum geometries of around 1um, and minimum feature size will continue to decrease well into the sub-micron region in the future.

Associated with this trend towards smaller geometries is an overall increase in processing complexity, and advanced processes may typically involve over two hundred distinct process steps. For a stable, well-controlled process, it has therefore become increasingly critical that the individual steps have a high degree of reproducibility and uniformity. Future technologies will require much higher levels of control than can presently be achieved, if process quality and yield are to be maintained at acceptable levels. The means by which process control can be enhanced to meet future manufacturing needs is the topic addressed by this thesis.

Process control, in its broadest sense, involves the measurement and monitoring of important process parameters and analysis of the information so that the status of the process can be determined. Corrective action is taken when the status is found to be unacceptable (feedback). In IC fabrication, process control exists in many forms, and operates in various

timescales. For example, electrical testing is normally carried out at the end of processing on special test structures. This produces a valuable means of monitoring the overall process. However, it cannot be used very effectively to adjust the process because the feedback loop is too long. If there is a malfunction in the process, many batches will have been processed before it is detected and corrected.

By comparison, control at much shorter timescales operates on individual process steps or several steps grouped together. Thus layer thickness is measured after a deposition step; and sheet resistance after an implantation-anneal sequence, to verify that the respective processes have been carried out correctly. Through such in-process monitoring, it is possible for errors to be corrected at some later processing step or through repetition of the faulty step (rework). At worst, the waste of further resources on wafers that are already defective is prevented. Analysis methods such as process-control-charts can be used to detect trends in the measured data, and provide a basis for process feedback.

It is imperative that monitoring and analysis be carried out as quickly as possible so that problems can be readily identified and appropriate action taken to rectify faults. Only by doing so will the number of mis-processed batches (and therefore fabrication costs) be minimised. However, the data handling requirements associated with process control are large, and increase with increasing process complexity. Much of the current effort towards improving process control has therefore centred on the extensive information management capabilities of Computer Aided Manufacture (CAM) systems.

CAM [1,2] concentrates primarily on factory integration and management. Specific areas that it covers include tracking of all batches being processed (work-in-progress), maintenance of batch histories and materials management. Integration of process data into the CAM database will be a necessary part of the advanced manufacturing environment, allowing for organised, efficient use of process information.

Process equipment provides another level of control. In response to the tighter tolerances demanded by VLSI, equipment control systems are becoming increasingly sophisticated, and many incorporate microprocessor-based digital systems. It is useful here to distinguish between two types of control parameters : 'process inputs' and 'process outputs'. Process inputs are causal in nature. Obvious examples are temperatures, pressures and times. They can be thought of as characteristics of the process environment whose influence becomes apparent in one or more of the process outputs. These process outputs are the measured outcome of the process and include, for example, film thickness, sheet resistivities and dimensions of patterned structures. Equipment control systems generally have relatively fast control loops which confer control during the process run. However, they control the process environment and operate on process input parameters, when usually it is the related process output parameters that are of main concern. (Process outputs are of primary importance since they are of direct relevance to the characteristics of completed device structures.)

Process outputs can rarely be controlled directly in a feedback loop, because they are difficult to monitor in-situ. If the measured process outputs warrant an adjustment to the process, then this is normally carried

out by adjusting the set points of the appropriate equipment-controlled process inputs. Effective control depends on a well-understood relationship between outputs and inputs. The problem is that the relationship between the two is complex and sometimes ill-defined. Many separate process inputs can affect a single process output, and the effects are often interrelated. For example in the silicon oxidation process, the resultant oxide thickness will be influenced by a combination of process times, temperatures and gas flow rates. Conversely, a single process input may influence several outputs ; for example, a higher temperature in one process step will influence all dopants introduced into the substrate up to that point. Not surprisingly, true closed-loop control of the process step(s) is usually lacking. It is the experienced process engineer who provides the necessary understanding for instigating changes to equipment settings, based on information about the measured process outputs.

A means of closing the loop and enhancing control is via the factory CAM system [3]. In the advanced, highly-computerised manufacturing environment, automatic feedback (or feedforward) to appropriately interfaced equipment is possible. The CAM system would automatically adjust equipment settings by loading new process instructions (i.e. recipes) into the microprocessor-based equipment control systems. The process database, built up by the aforementioned input of measured process data into the CAM system, forms the basis for decision-making by the CAM software. Automatic data collection, through interfacing measurement equipment to the CAM system, would further enhance the process.

Although interface standards for equipment communications such as SECS[1] (SEMI Equipment Communications Standard) [4] now exist and are being increasingly implemented in the hardware and software of advanced process equipment, this approach is generally made difficult by the wide range of equipment that is available, each with different forms of internal control system and varying levels of capability for automatic control. Moreover, it would be a formidable task to develop an integrated system that applied to all areas of the production environment covered by the CAM system since, as already mentioned, the decision-making process necessary for feedback is seldom simple.

Even assuming that these difficulties could be overcome, the feedback loop would still be long and slow, and in all cases changes to equipment settings would arrive too late for the mis-processed or poorly processed wafers that had induced the change. The method also assumes that the process specification, in the form of the recipe delivered to the equipment, accurately describes the actual wafer environment. However, deviations from set point of parameters controlled by the equipment (i.e. process inputs) frequently occur, especially in the transient periods that occur during processing, and so nullify this assumption. This is particularly important for processes with short process times, where transient periods may form a significant part of the overall timescale.

In this work, a radically new approach to enhancing process control, that overcomes all of the above problems, is presented. This is possible by implementing process outputs in a feedback control loop at the equipment

[1] The SECS protocol is outlined in Appendix B

level. Real-time computer simulation of the process step forms the basis of the system, constituting the essential link between process inputs and process outputs. Housed in a microcomputer, the simulation is fed with real-time measurements of all important process inputs. If these are regularly updated, then a real-time estimate of process outputs can be calculated. The estimates can then be used to predict endpoints for the process. The microcomputer provides feedback to the process equipment, which is actively controlled during every run.

The advantages of this approach are:

i) In comparison to the control loop via the factory CAM system, this approach is implemented at the 'local' level of process equipment, and is more easily optimised to suit the individual process step(s) involved.

ii) Actual process conditions, as opposed to the idealised status of the process specification, are considered in the control loop.

iii) Any number of relevant output parameters can be considered, as long as the physical relationships to input parameters are known or an adequate empirical model exists.

iv) Process outputs, not process inputs are the basis of control. Desired values (i.e. set points) for the process outputs are entered at the microcomputer (e.g. the desired oxide thickness would be entered at an oxidation step).

v) Processing is optimised to give the desired process outputs for every run. Variation in any measured

input parameter during the process is automatically taken into account.

vi) The system is independent of deviation from set point of any parameter controlled by the equipment control system that is input into the real-time simulation. Thus allowance is made for transient periods, during which equipment parameters are generally poorly controlled.

The success of a real-time simulation-based control system obviously rests on the availability of an accurate process model. However, advanced models for many fabrication steps already exist, and are being constantly improved as the understanding of underlying physical processes increases. These models are most often applied in process simulation programs such as SUPREM III [5], where they have an important role as an aid to process development. The control system presented here introduces a new application of process simulation beyond this mainly passive role.

This work investigates whether the current simulation capability can be successfully utilised in an active control system. It focuses on a single aspect of silicon fabrication - thermal growth of thin silicon dioxide films. A real-time simulation-based control system for the oxidation process was developed. There were several reasons for the choice of silicon oxidation:

i) Oxidation is a critical process in Metal Oxide Semiconductor (MOS) device fabrication, currently the most commercially important fabrication technology. It is also of great importance in the fabrication of bipolar devices. The oxidation process directly determines the electrical properties of MOS devices ; hence the ability to

reproducibly form clean high-quality gate oxide is fundamental to successful MOS IC manufacture. The thickness of the oxide film is the primary process output of the oxidation process and must be carefully controlled. There is a particular need for enhanced control as oxide thicknesses are reduced to meet VLSI demands and process times are correspondingly decreased.

ii) Silicon oxidation has been extensively investigated in recent years. There is a relatively large amount of information available concerning the kinetics of oxide growth and the influence of processing conditions on oxide properties.

iii) Advanced oxidation equipment generally incorporates microprocessor-based controllers and interfacing possibilities to external microcomputers.

In order to provide the background for later chapters, Chapter 2 of the thesis presents a review of fabrication. Emphasis is given to aspects of fabrication which are of particular relevance to this project. Much of silicon fabrication is based on the iterative cycle of deposition-photolithography-etch. The principal individual operations in this cycle are discussed, including the applications and methods of thermal oxidation. A 5-micron NMOS process is described, a process which is representative of the fabrication technology used for LSI (Large Scale Integration). It illustrates the complexity of device fabrication and the importance of monitoring and controlling each of the individual steps if high-quality working devices are to be produced.

The advance to VLSI requires processing techniques in addition to those used for LSI fabrication and these

are discussed. Two of the most important topics for future VLSI, gate oxides for MOS devices and automation of the fabrication process, are especially relevant to this work, and are looked at in detail. Process simulation has an important role in VLSI process development. The major objectives of simulation and the principal computer packages currently in use are described.

In Chapter 3, the oxide growth process is discussed in detail. The discussion provides the background for the later implementation of a growth model in the simulation control system. Much of the current understanding of oxidation kinetics is based on the work of Deal and Grove carried out in the 1960's [6]. Their model of the growth kinetics is outlined, and the relation between this theory and the main process inputs which influence oxide growth is examined. The more detailed theories that have been proposed to explain the Si-SiO_2 interface reaction and the observed kinetics are also described. Even when first proposed, it was clear that the Deal-Grove model did not adequately describe the kinetics of thin oxides grown in oxygen. Such oxides are of increasing significance to VLSI processing. The additional models that have been developed to explain the anomalously high oxidation rate in this regime are compared. Of equal significance to VLSI is the use of chlorine-containing compounds in the oxidising ambient. The addition of a few percent HCl gas, for example, can considerably improve oxide quality. The additives used, and their influence on oxidation growth rates is discussed in the penultimate section of the chapter. The final section returns to process simulation, and looks at the implementation of oxide growth models in several packages. The models used are defined, along with their limitations. This forms an important

starting point for the real-time simulation of oxide growth.

Following the theory of oxidation in Chapter 3, Chapter 4 switches to the practical aspects of carrying out oxidation in the IC production environment. A generalised scheme for state-of-the-art oxidation equipment is suggested and the component parts are individually reviewed. Emphasis is placed on the control aspect, including the principal feedback loops used to control the oxidation process inputs. The furnace used for the experimental work of later chapters is described and related to the generalised outline. It is shown that current furnace systems are unable to confer full control on the oxidation conditions experienced by wafers. Experimental data specifically demonstrate problems with temperature control and control of the gases supplied to the oxidation equipment. It is established that significant transient periods exist during the oxidation process.

It is argued that pre-oxidation cleaning should be an integral part of practical oxidation procedures. As well as being crucial from the point of view of contamination, pre-oxidation cleaning influences the final oxide thickness, and must be well-regulated to obtain a stable, controlled process. The cleaning sequences in common use are reviewed.

Measurement and analysis of an oxide film is a necessary part of the investigations in Chapters 6 and 7. Many methods are available, and Chapter 5 is devoted to an investigation of the principal techniques. It is split into two broad sections covering measurement of oxide thickness and assessment of oxide quality. For thickness measurement, ellipsometry is generally considered to be one of the most accurate techniques.

However, it is shown by comparison of standard commercial equipment that optical interferometry can give comparable reproducibility. It is suggested that interferometry should be the preferred technique when large amounts of data are to be collected, since the time to complete a measurement on the available equipment is considerably shorter.

The discussion on oxide quality centres on MOS capacitor measurements. These are used to monitor both oxide charge concentrations and oxide dielectric integrity. It is important to distinguish the various types of oxide charge, and the standard terminology now in use is detailed. Capacitance and conductance measurements have potentially large benefits, but analysis methods are generally complex. An overview of the information that can be obtained is presented, along with important practical points concerning the measurement conditions. By comparison, dielectric breakdown measurements are relatively straightforward, but remain vitally important in the investigation of oxide defects and oxide reliability.

In Chapter 6 the uniformity in oxide thickness that is achievable with current furnace technology is investigated. Oxide thickness variability is measured for several different oxidation processes, and both between-batch and within-batch components are assessed. It is argued that significant between-batch variation is related to transient effects and the inadequacy of the current control methods. Specific patterns in the collected data are related to the control problems noted in Chapter 4. The furnace used in the investigation was first subjected to a rigorous verification procedure, the results of which are also presented in this chapter.

Chapter 7 describes the new control system, based on real-time simulation, that has been developed for furnace oxidation. The system implements the oxidation theory developed in Chapter 3. The accuracy of the model depends on suitably chosen rate constants. Simulation using literature values is compared with oxide growth rate data obtained from a series of oxidation runs. New values for the rate constants are extracted from the data set for additional comparison.

A bi-directional interface between oxidation model and furnace system was developed to realise active control of furnace operations. An extensive software system was also developed to execute all interface and modelling requirements, and provide a friendly user interface. The complete system was tested under a wide range of oxidation conditions, including processes involving substantial oxide growth under dynamic conditions. With this experimental study, the feasibility of the real-time simulation control technique is examined.

CHAPTER 2

REVIEW OF FABRICATION

2.1 Principles of Silicon Fabrication

The IC fabrication process is of a highly complex nature - typically involving hundreds of individual operations - and has many variations depending on the type of product. However, there is a basic sequence of operations common to all processing technologies (Fig. 2.1). Device structures are built up on a substrate material through many repetitions of the layer deposition-photolithography-etch sequence. The cycle may be repeated over 10 times in advanced VLSI processes and may include additional steps such as ion implantation and high-temperature "drive-in" treatments.

In silicon IC production the starting substrates are single crystal silicon wafers which are typically between 3 and 6 inches in diameter. These are processed in batches of up to 200 wafers. Coupled with the large number of individual chips that can be created on a wafer, this batch production is capable of giving very large numbers of ICs and is the main reason why the fabrication process is cost-effective.

The individual operations in the processing cycle are now discussed.

Layer deposition forms a controlled amount of new material on the wafer surface. Several materials, selected for their insulating, conducting or masking properties, are used and several deposition methods are involved. The following list describes some of the most

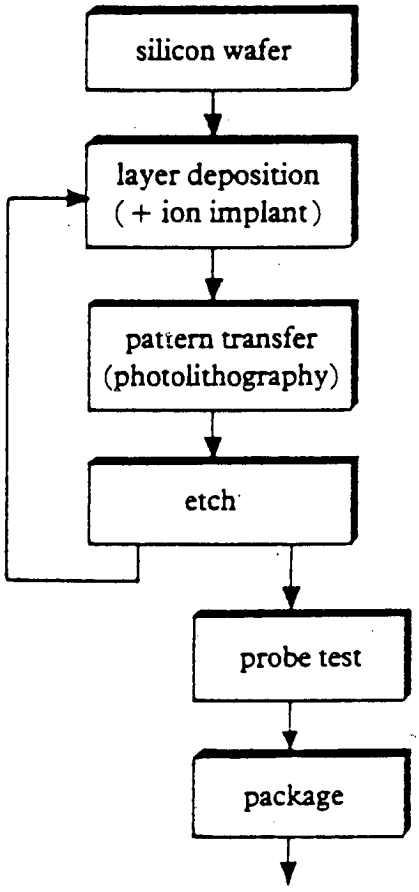


Fig. 2.1 Basic IC Fabrication Sequence.

important techniques. Polysilicon, silicon dioxide and silicon nitride can be deposited by Low Pressure Chemical Vapour Deposition (LPCVD) [7,8]. These materials have the advantage of being able to withstand high-temperature processing. Deposited oxide is often doped with phosphorous, using suitable additions to the reactant gases. The presence of phosphorus allows the oxide to flow during high-temperature treatment and is used to smooth sharp edges for improved step coverage. Aluminium is usually deposited by electron beam evaporation or by sputtering techniques [9,10]. In the latter case argon ions bombard an aluminium target, releasing aluminium which coats wafers placed in the sputter chamber.

Ion implantation [11,12] allows well-controlled adjustment to be made to the density of dopant atoms in the substrate material. The technique involves accelerating arsenic, boron or phosphorus dopant ions at energies in the range 20-200 keV. The ion beam is deflected over the silicon wafers, which are penetrated to depths of up to 2um. The implanted depth depends primarily on the ion energy. An alternative method of introducing dopant atoms into silicon is gas-phase diffusion [13], which is carried out as a furnace process at temperatures between 900 and 1200°C. Due to its poor controllability, diffusion has been mostly superseded by implantation, apart from instances where heavy doping levels are required.

Photolithography [14,15] provides the means of transferring a desired pattern into deposited material. The entire wafer is covered with a light sensitive film of photoresist which is then exposed to UV light through an appropriate mask. Exposure equipment must be able to accurately align successive layer patterns to each other

so that device structures can be built up. High demands are also made on the optical system which must be capable of resolving the minimum feature size of the technology. Wafer-steppers using masks 5x or 10x the final size offer several advantages over the older contact and proximity techniques and are being increasingly used in advanced processes [16], despite their much reduced throughput.

Of the two types of photoresist - positive and negative - positive is now generally preferred due to its better resolution. With positive resist, exposed areas are washed away upon development. A protective resist coat is left in unexposed areas which prevents the underlying material from being etched away in the subsequent etch step. Etching [17,18,19] transfers a resist pattern into deposited material and is carried out either with liquid chemicals (wet etching) or in a plasma process (dry etching). Plasma processes can be mainly anisotropic (as in reactive ion sputtering) or mainly isotropic (usually referred to as "plasma-etching") or a mixture of the two. Wet etching is isotropic. Careful choice of the etching method must be made since it can greatly effect feature size and edge profiles. After etching, remaining resist is removed by chemical or plasma treatments and the next cycle of the process is started.

When all iterations are complete, an electrical probe test is normally carried out on special test structures to check that all operations have been successfully carried out. Individual chips are then finally cut from the wafers using a diamond saw and packaged.

2.1.1 Thermally Grown SiO_2

Although not classified as a deposition process, the thermal oxidation of silicon to produce SiO_2 is also a means of forming new material on silicon. It is normally carried out in O_2 or H_2O ambients (corresponding to wet and dry oxidation respectively) at temperatures in the range 900-1150°C. Oxidation forms a passivating insulator on the silicon surface and is the main reason for the success of silicon as a semiconductor material in microelectronics. The oxide layer serves several functions in the fabrication of devices, and advanced processes may involve over 15 separate oxidation steps. The main functions are:

i) Isolation
In both MOS and bipolar technologies, thick field oxide regions (typically 5000-15000Å) provide insulation between active devices, and between conducting layers and the silicon substrate. In double polysilicon processes, isolation of the two poly layers involves thermal oxidation of polysilicon.

ii) Gate dielectric
The low charge densities associated with (dry) thermal oxidation are crucial to the stable operation of MOS devices.

iii) Implant/diffusion barrier
Patterned oxide is able to selectively mask parts of the silicon against introduction of dopant ions. Implantation through thin oxide is often used to adjust the dopant profile in the substrate.

iv) Stress relief

SiO_2 is used as a mechanical buffer between Si and Si_3N_4 in the LOCOS [20] process. This is necessary to relieve stress caused by the difference in thermal expansion between Si and Si_3N_4.

Oxidations are normally carried out in furnaces open to atmospheric pressure. Oxidation rate depends strongly on temperature which must therefore be carefully controlled[1]. The actual oxidation technique used depends on the particular application. Thicker oxides are grown under wet conditions since the oxidation rate is much faster than with dry oxidation. In modern furnaces water vapour is created by the pyrogenic technique, which involves direct combination of hydrogen with oxygen. Chlorinated dry oxidation has well-known beneficial effects on oxide quality[2] [21,22] and is frequently used to form gate oxide. Likewise, a high-temperature post-oxidation anneal in an inert ambient (such as N_2) can improve oxide quality [23]. Wafer warpage is a problem at higher oxidation temperatures, but is commonly minimised using a low loading/unloading temperature and ramping to the desired oxidation temperature.

When thicker oxides are required, high-pressure oxidation [24,25] (HIPOX) offers advantages. At high pressures, the growth rate is accelerated and allows oxidation to take place at lower temperatures whilst maintaining reasonable oxidation times. Thus the detrimental effects of long high-temperature oxidation, such as the growth of oxidation-induced stacking faults

[1] The features of advanced microprocessor-controlled furnaces are discussed fully in Chapter 4.

[2] This will be discussed in Section 3.4

and excessive dopant diffusion can be avoided. Plasma-enhanced oxidation [26] also considerably increases the oxidation rate and offers similar benefits.

2.2 NMOS Process Outline

MOS-based technologies have become dominant in the IC market, mostly due to their suitability for high levels of integration. The original commercial MOS process was PMOS and was introduced in the early 1960's. However, with the advent of ion implantation in the 70's, allowing accurate adjustment of transistor threshold voltage, NMOS processing became commercially viable. In NMOS technology, the basic circuit element is the n-channel MOSFET [27]. Because of the inherent higher switching speeds of n-channel devices over p-channel devices, NMOS soon overtook PMOS in importance. The combination of n-channel and p-channel in one chip - CMOS technology - offers lower static power consumption and is of increasing prominence, despite its added complexity. However, NMOS technology is still likely to be of significant importance in the future.

Table 2.1 shows a typical sequence of steps for NMOS fabrication. It has been based on the 5-micron NMOS process [28] which is run in the Edinburgh Microfabrication Facility (EMF) clean rooms at the University of Edinburgh. The major processing steps are now discussed:

i) The starting material is 3" diameter, <100> orientation, lightly-doped p-type silicon wafers.

ii) Active device areas are first isolated with thick field oxide (Fig. 2.2(a)) using the LOCOS process. An Si_3N_4 mask is used to inhibit oxide growth in

1. Substrate : P-type <100> 14-20 ohm.cm silicon
2. Pad Oxide : 500Å dry-wet-dry oxide (950°C)
3. Si_3N_4 Deposition : 500Å, LPCVD
4. 1st Photo : Definition of field areas
5. Field Implant : Boron 2 x 10^{13} cm^{-2} 130 keV
6. Si_3N_4 Etch : CF_4/O_2 plasma
7. Resist Strip : Fuming nitric acid
8. Oxide Etch : HF/NH_4F solution
9. Field Oxidation : 10 000Å wet oxidation (950°C, 16 hrs)
10. Si_3N_4 Etch : CF_4/O_2 plasma
11. 2nd Photo : Depletion threshold adjustment
12. Depletion Threshold Implant : Arsenic 1.5x10^{12} cm^{-2} 90 keV
13. Resist Strip : Fuming nitric acid
14. Gate Oxidation : 800Å dry-wet-dry oxide (950°C)
15. Enhancement Threshold Implant : Boron 11.3x10^{11} cm^{-2} 40 keV
16. Anneal : 950°C for 30 min in N_2
17. 3rd Photo : Buried contact
18. Oxide Etch : HF/NH_4F solution
19. Resist Strip : Fuming nitric acid
20. Polysilicon Deposition : 3500Å, LPCVD
21. 4th Photo : Polysilicon gate and interconnect
22. Polysilicon Etch : CF_4/O_2 plasma
23. Resist Strip : Fuming nitric acid
24. Phosphorus Diffusion : 1000°C in N_2 (solid Ph source)
25. Polysilicon Oxidation : 2000Å dry-wet-dry oxide (950°C)
26. Ph-doped Oxide Deposition : 7500Å phosphosilicate glass
27. Reflow 1 (densification) : 20 min at 1050°C in O_2
28. 5th Photo : Contact holes
29. Oxide Etch : HF/NH_4F solution
30. Resist Strip : Fuming nitric acid
31. Reflow 2 (step coverage) : 20 min at 1050°C in O_2
32. Aluminium Deposition : 15000Å evaporated aluminium
33. 6th Photo : Aluminium interconnect
34. Aluminium Etch : Aluminium etch sloution
35. Resist Strip : Fuming nitric acid
36. Sinter : Low temperature anneal (450°C)
37. Ph-doped Oxide Deposition : 7500Å phosphosilicate glass
38. 7th Photo : Pad areas
39. Oxide Etch : HF/NH_4F solution
40. Resist Strip : Fuming nitric acid

TABLE 2.1 Outline of the processing steps in the EMF 5um NMOS process.

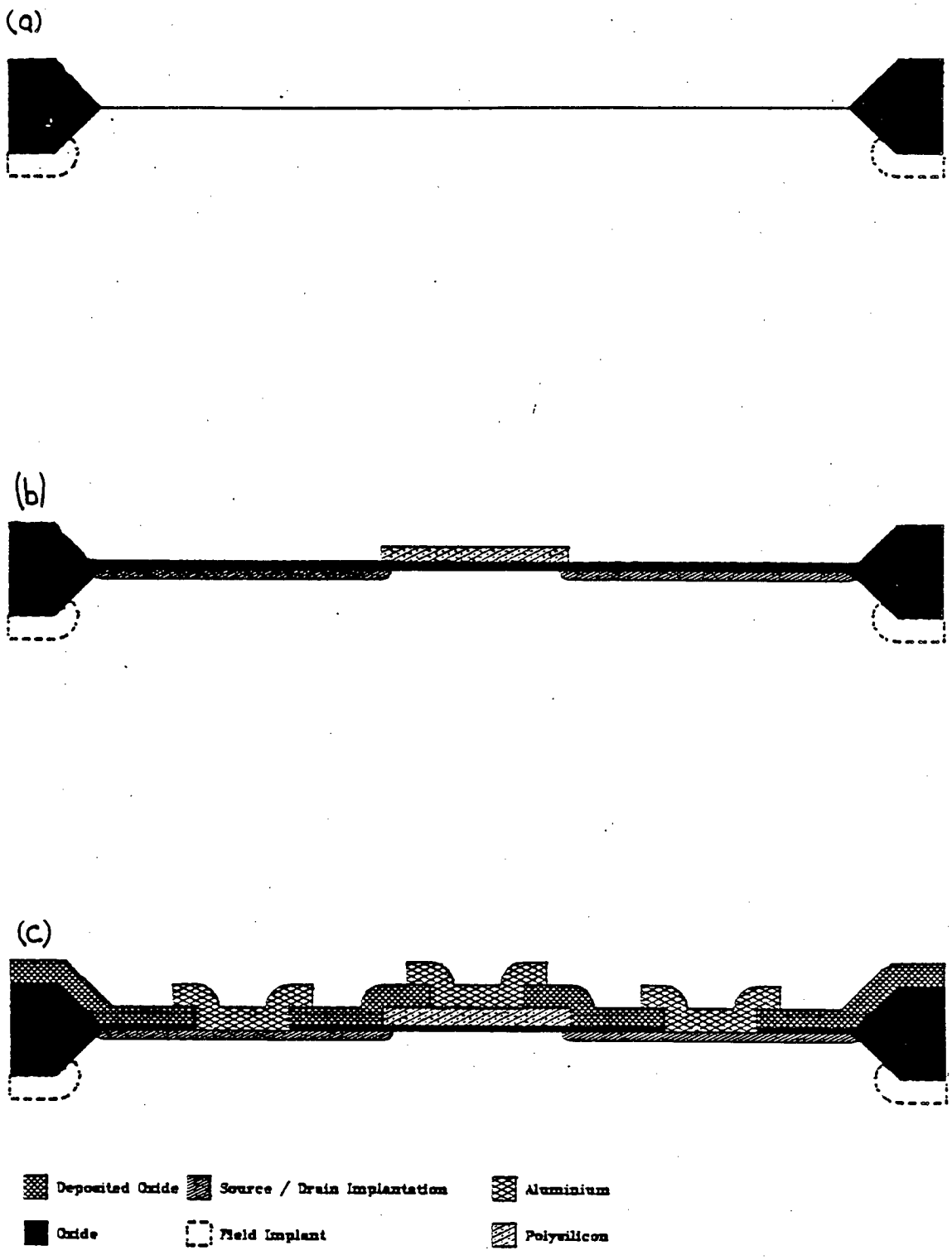


Fig. 2.2 Three stages in the production of an n-channel MOSFET (a) Isolation of active device areas (b) Fabrication of the MOSFET structure. (c) Interconnection with Aluminium.

active regions. A boron implant under the field areas provides the electrical isolation between active areas.

iii) The MOSFET is fabricated with a thin thermally-grown gate oxide and a phosphorous-doped polysilicon gate material (Fig. 2.2(b)). A self-aligning process [29] that aligns the start of the diffused source/drain regions to the edge of the gate is used to reduce parasitic overlap capacitances and give faster circuit operation. Implants of arsenic and boron are used to adjust the threshold voltages of the two types of transistors found in NMOS : enhancement (i.e. normally off) and depletion (i.e. normally on). An additional sequence not shown in Fig. 2.2(b) defines buried contacts, where polysilicon makes direct contact to the silicon substrate.

iv) The individual devices are interconnected with aluminium (Fig. 2.2(c)). Before its evaporation, a thick phosphorous-doped silicon dioxide layer is deposited over the whole wafer to isolate the underlying structures. Contact holes are etched in this layer where the aluminium is to contact polysilicon or diffused regions.

v) A final phosphorous-doped oxide layer provides mechanical protection and reduces the sensitivity of the devices to ionic contamination. Windows are cut in this passivation layer over the pad areas where external connection will be made to the chip.

2.3 Requirements for VLSI Processing

A process such as that just described produces devices with a minimum feature size of 5um. The shrinking of this minimum geometry to VLSI levels (i.e. <2um) requires many additional processing techniques and technological innovations. Gate oxidation and process automation are two particularly crucial areas of VLSI fabrication and will be discussed later in detail. Other VLSI techniques include:

i) replacement of the LOCOS process with alternative isolation techniques that reduce or eliminate the well-known bird's beak effect [30] and so enhance packing density. A variety of techniques exist including the SWAMI process [31], moat isolation [32] and trench isolation [33].

ii) the use of implantation to form source/drain regions. This allows shallower, better controlled junctions. Arsenic is used as the dopant, in preference to phosphorus, to reduce lateral dopant diffusion under the transistor gate.

iii) reducing unwanted dopant diffusion by using rapid thermal processing [34,35] instead of long, high-temperature furnace treatments.

iv) the use of refractory metal silicides in place of, or in conjunction with, polysilicon as gate and inter-connect material [36]. This reduces parasitic resistances and speeds up circuit operation.

v) drain-engineering techniques [37,38] such as lightly doped drain (LDD) that reduce the problems associated with high-fields encountered in small geometry MOSFETS.

vi) additional channel implants below the silicon surface. These raise the doping concentration and help prevent device punchthrough without increasing the bulk dopant concentration.

2.3.1 Gate Oxides for VLSI

The general downward scaling of device dimensions has necessitated a reduction in gate oxide thickness. Thicknesses of 150-250Å are now required for 1um technologies, and even thinner oxides will be needed in the future. Two problems are associated with thin oxides - thickness uniformity and oxide quality (i.e. dielectric properties and charge density levels).

Oxide thickness is one of the most important MOS parameters and appears as a first order term in expressions for threshold voltage, transconductance, gain and the body effect[3]. Control of oxide thickness is therefore crucial for stable device operation. As the oxide thickness is reduced, the tolerable thickness variation is also reduced (Table 2.2) and becomes increasingly difficult to achieve. The established method of growing gate oxide is furnace oxidation in a dry ambient at high temperature (950-1100°C). With thinner oxides, growth times become very short (e.g. about 30 minutes for 200Å at 950°C). This augments the problem since growth during transient periods becomes increasingly significant. The control capability of current furnace technology, with respect to transient effects, is investigated in Chapter 4.

[3] The principle MOSFET relationships are discussed in Appendix A. Comprehensive treatments on MOS device physics can be obtained in texts such as Sze [27] and Nicollian and Brews [39].

Minimum Device Feature (um)	Typical Gate Oxide Thickness (Å)	Required Tolerance (Å)	Approx. Dry Oxidation Time at 950°C (min)
5.0	800	40.0	245
2.5	400	20.0	90
1.25	250	12.5	40
0.5	100	5.0	9

TABLE 2.2 The effect of scaling on the tolerable variation in oxide thickness.

Lowering the oxidation temperature below 900°C will decrease the growth rate and allow longer oxidation times. Such low-temperature oxides have good breakdown characteristics, but the $Si-SiO_2$ interface properties are found to be poorer than those from high-temperature oxidations [40]. Since MOSFET characteristics are very sensitive to interface and oxide charges[4], the higher temperature oxides are still preferable.

The need to optimise dielectric properties and charge density levels has given rise to novel oxidation procedures for VLSI gate oxides. The beneficial effects of adding chlorine species to a dry oxidation ambient or a post-oxidation annealing in an inert ambient have been exploited in multi-step oxidation processes. These include dual-HCl [41], dual-TCE (trichloroethylene) [42] processes and oxidation with an intermediate anneal [43]. However, all involve very high-temperature steps (>=1050°C) and go against the general VLSI tendency towards lower temperature processing. Such complex procedures may also be expected to hinder control of oxide thickness.

Rapid Thermal Oxidation (RTO) has recently been suggested as an alternative method for the growth of high-quality thin oxides [44,45]. In this technique, single wafers are placed flat on support pins inside a small quartz tube. Heating is provided by tungsten-halogen lamps, which allow very fast temperature ramps. Without the thermal mass of a boat load of wafers to heat, as in conventional oxidation, high temperatures of up to 1200°C can be achieved in times of the order of a few seconds (even wafer heating prevents induced thermal stress and wafer warpage). Rapid purging of process

[4] Oxide charges are discussed in Chapter 5.

gases in the small reactor volume further reduces the impact of transient periods on the overall process. Oxidation rates are enhanced by this method [45,46] and very short high-temperature cycles (i.e less than 5 minutes) are possible, with the result that excessive silicon dopant redistribution is avoided.

Temperature in RTO systems is usually monitored with a pre-calibrated optical pyrometer and is computer-controlled with a closed feedback loop. The accuracy of this method seems to be adequate and wafer-to-wafer thickness uniformities of 1% to 6% have been reported for thicknesses in the range 50Å to 350Å [45]. However, problems at high temperatures with high lamp intensity interfering with the pyrometer measurements were also noted in [45]. Since very high oxidation temperatures are possible, the oxide films produced using RTO can be of high quality [47].

Despite the above advantages, RTO cannot match the throughput of conventional furnace systems and RTO films have not yet seen widespread use as a gate dielectric material. However, with future processes, where wafers may be larger than 6" diameter, RTO and other applications of rapid thermal processing (see previous section) are likely to see increased use in critical process steps. This is because wafers are heated (and cooled) uniformly, in comparison to conventional furnace systems where wafer heating is from the edge inwards. Additionally, in specific applications such as ASICs (which are becoming an increasingly important branch of IC production) batch size may be small and so the throughput by rapid thermal processing techniques becomes more comparable with that from conventional methods.

Nitrided SiO_2 and nitrided Si have also generated recent interest as alternatives to thermally grown gate

oxide [48]. The films are formed in pure ammonia gas at high temperature. Plasma enhancement is usually used since the thermal reaction is a very slow process. Although improved characteristics - including a higher dielectric constant - can be obtained, further investigation of such films is needed before they can fully replace the much better understood $Si-SiO_2$ system.

It is clear that there are significant advantages to be gained from conventional furnace oxidation including throughput, equipment reliability and general compatibility with existing fabrication procedures. But the ability to satisfy the thickness uniformity requirements of VLSI under the conditions necessary for high oxide quality is crucial to its continued use in gate oxidation. Advances in furnace technology over the past ten years have led to total automation of furnace operations. The significance of these developments with respect to oxide thickness control is investigated in Chapter 6.

2.3.2 Process Automation

Automation of the wafer fabrication process is becoming increasingly important to maintain VLSI productivity. It is driven by the need for high process yield in the face of increasing process complexity and reduced feature size. Automation takes many forms but can be most easily categorised with respect to the three main mechanisms that influence overall process yield. These are:

i) defects caused by particles and contamination
Even with advanced clean room capability, particulate contamination cannot be completely eliminated. Humans are amongst the largest

sources and considerable effort has been directed at keeping humans and wafers apart. Much automation has already been implemented at the individual equipment level, but only recently have the problems of wafer transport between stations and the equipment loading/unloading interface been addressed. Concepts such as the SMIF box [49] coupled with robotic handling systems [50] are now seeing increased use. Clean track systems suspended from a ceiling or wall, or automated guided vehicles are the main alternatives for wafer transportation [51].

ii) wafer breakage and misprocessing

Careful factory management is needed to concurrently steer many different batches successfully through the hundreds of individual process steps of advanced technologies, particularly when the factory may be running several different processes and process variations. Computer aided manufacturing (CAM) software [1,2], such as COMETS and PROMIS, can help eliminate wastage by keeping track of, and routing, work in progress.

iii) process control

As discussed in Chapter 1, many process steps run open-loop and process control would be enhanced by the provision of automatic feedback control. This thesis develops a specific means of implementing such control at the equipment level. In addition, automated data handling and movement, by means of equipment interfaced to a centralised CAM system, will reduce the scope for error and can also be expected to improve overall process yield.

Although CAM systems are now implemented in many major fabrication lines, and much new equipment incorporates standard interfaces like the SMIF concept or SECS, full process automation is still some way off. The transition will be facilitated by the development of individual highly-automated processing modules that can then be complemented by automatic wafer transport. These modules can gradually be introduced as they become available, with minimum disruption to the production environment. The active control system proposed here, being based at the equipment level, is in accordance with this approach to process automation.

2.4 Process Simulation

The driving force behind process simulation has been the need to predict the effect of process changes on device and circuit characteristics. Not only is experimentation thus reduced, but detailed information and added insight about the process can be obtained. A description of the process flow - including details of temperature and time - is used as input to the process simulation. The effect of each process step on dopant ions is modelled and the resultant impurity distribution forms the output (Fig. 2.3). This may then be coupled with a device simulator to estimate device parameters.

Of several process simulation packages available, SUPREM II, ICECREM, SUPREM III, SUPRA and SAMPLE are amongst the most well-known. SUPREM II [52], one of the earliest programs, incorporates one-dimensional models for oxidation, diffusion, deposition, implantation and etching. It is limited to two process layers and has been superseded by SUPREM III [5] which can model up to ten layers, including polysilicon, and includes some

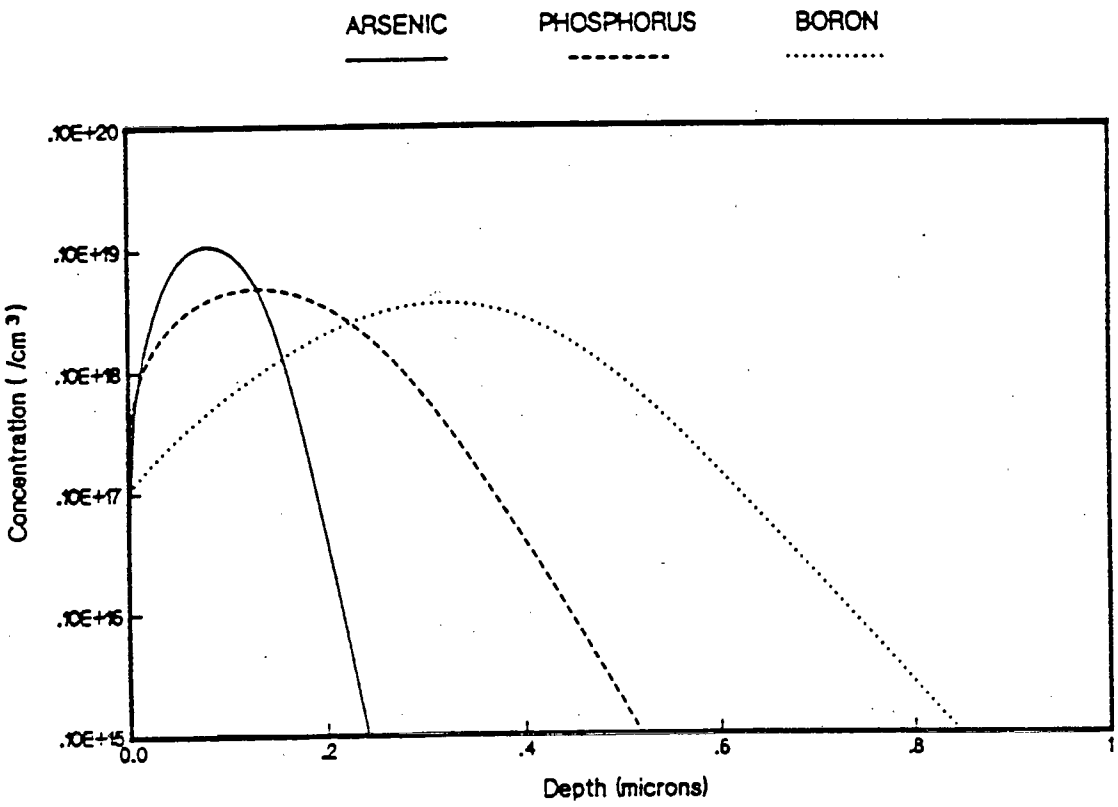


Fig. 2.3 Sample output form ICECREM simulation : Concentration profiles of implanted As, P and B after a high temperature diffusion.

enhancement to the individual models. ICECREM [53] has similar capabilities to SUPREM II.

As device geometries shrink, lateral effects must also be considered, and much effort has been directed at the development of 2-D simulators [54,55]. SUPRA models oxidation and diffusion in two dimensions and SAMPLE is a topographical simulator of the lithographic process. SUPREM IV [55] is a more advanced 2-D process simulator, that has recently been developed as the 2-D equivalent to SUPREM III.

The models used in simulators are based on a mixture of physical theory and empirical fits to experimental data. Numerical methods are used to solve the complex diffusion equations. Simulation is complicated by coupling between oxidation and diffusion, impurity clustering and impurity segregation. In two dimensions, the changing shape of the device structure (e.g. bird's beak in LOCOS) must also be considered and has prompted the use of new mathematical algorithms to avoid excessively long computing time.

Model parameters may sometimes be adjusted to suit individual processes and there remain gaps in the understanding of some physical processes (examples include phosphorus diffusion and thin oxide growth). Nevertheless, models have generally improved over the years and are applicable under an ever wider range of operating conditions. They are sufficiently refined that new applications of simulation, beyond the role as a tool in process development, can now be considered. One such application, simulation in real-time through direct coupling with process equipment, is the topic addressed by this thesis.

CHAPTER 3

OXIDATION OF SILICON

3.1 Kinetics of Oxide Growth

Silicon oxidises in the presence of oxygen or water vapour according to the following reactions, to give a film of SiO_2 on the silicon surface.

"Dry" oxidation : $Si + O_2 \text{ -----> } SiO_2$ (3.1)

"Wet" oxidation : $Si + H_2O \text{ ----> } SiO_2 + H_2$ (3.2)

Marker experiments using ^{18}O have shown that the mechanism of these reactions involves diffusion of the oxidising species through any existing SiO_2 film, and subsequent reaction with silicon at the $Si\text{-}SiO_2$ interface [56,57]. The interface therefore moves into the silicon as more and more Si is converted to SiO_2. The differences in the density and molecular weight of Si and SiO_2 indicate that a volume increase of approximately 100% accompanies the reaction, so that the SiO_2 is not coplanar with the original silicon surface (Fig. 3.1).

The following analysis of the kinetics of the oxidation reaction leads to the well-known linear-parabolic growth law and is based on the description proposed by Deal and Grove in 1965 [6].

For oxide to form at the $Si\text{-}SiO_2$ interface, the oxidising species must go through three distinct stages as shown in Fig. 3.2. It must first be adsorbed from

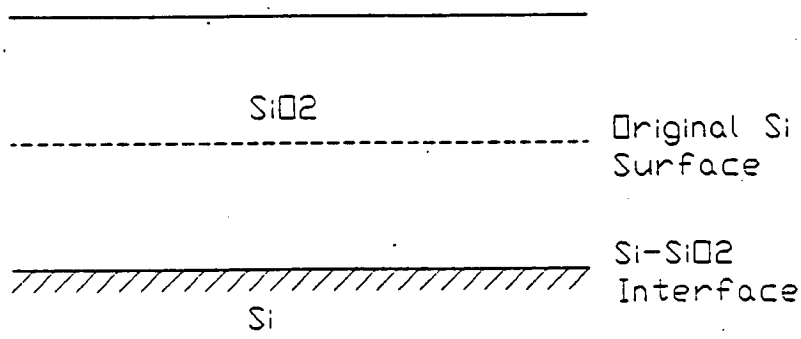


Fig. 3.1 The volume increase resulting from oxidation of silicon. Approximately 60% of the newly formed SiO_2 lies above the original Si surface.

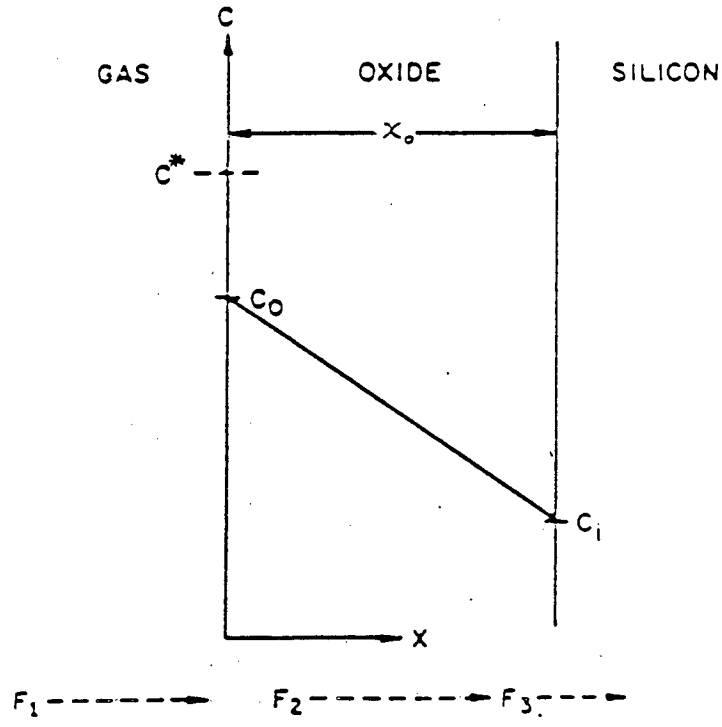


Fig. 3.2 Model for the thermal oxidation of silicon. The fluxes F_1 to F_3 correspond to the three stages of oxide growth: adsorption of the oxidising species from the gas phase, diffusion through the oxide, and reaction with silicon respectively.

the gas phase onto the outer surface of the oxide already formed. It then diffuses through the oxide to reach the interface where it finally reacts with the Si to form SiO_2.

Assuming that all these processes reach a steady-state, then the flux of oxidant in each are equal. (Flux is the number of molecules crossing a unit area in unit time.)

Flux F_1, as illustrated in Fig. 3.2, is assumed to be

$$F_1 = h \left(C^* - C_o \right) \qquad (3.3)$$

where C_o is the concentration of the oxidant in the outer surface of the oxide and C^* is the equilibrium concentration for the oxidant in the oxide. h is a gas-phase mass transport coefficient. Henry's Law states that the equilibrium solid concentration is proportional to the bulk gas partial pressure, so that

$$C^* = K P \qquad (3.4)$$

The use of Henry's Law assumes that there is no dissociation (or association) of the oxidant at the outer surface.

Flux F_2, the diffusion of oxidant across the oxide layer, is assumed to be given by Fick's Law, which states that flux is proportional to the concentration gradient.

$$F_2 = -D_{eff} \frac{dC}{dx} = D_{eff} \frac{(C_o - C_i)}{x_o} \qquad (3.5)$$

where D_{eff} is the effective diffusion coefficient, C_i is the oxidant concentration near the oxide-silicon interface and x_o is the oxide thickness.

Flux F_3, describing the oxidation reaction at the interface, is given by

$$F_3 = k_s C_i \qquad (3.6)$$

where k_s is the interface reaction rate constant for the oxidation.

If steady-state conditions exist, then $F_1 = F_2$ and $F_2 = F_3$, and it can be shown that

$$C_i = \frac{C^*}{1 + k_s/h + k_s x_o/D_{eff}} \qquad (3.7)$$

and

$$C_o = \frac{C^* \{ 1 + k_s x_o/D_{eff} \}}{1 + k_s/h + k_s x_o/D_{eff}} \qquad (3.8)$$

The flux of oxidant $F = F_1 = F_2 = F_3$ is given by

$$F = k_s C_i = \frac{k_s C^*}{1 + k_s/h + k_s x_o/D_{eff}} \qquad (3.9)$$

and the oxide growth rate by

$$\frac{dx_o}{dt} = \frac{F}{N_1} = \frac{k_s C^*/N_1}{1 + k_s/h + k_s x_o/D_{eff}} \qquad (3.10)$$

where N_1 is the number of oxidant molecules incorporated into a unit volume of the oxide.

This differential equation is solved by integration, assuming an initial condition $x_o = x_i$ at time t=0.

This yields the following generalised relationship:

$$\frac{x_o^2 - x_i^2}{B} + \frac{x_o - x_i}{B/A} = t \qquad (3.11)$$

or

$$x_o^2 + Ax_o = B (t + \tau) \qquad (3.12)$$

where

$$A = 2 D_{eff} (1/k_s + 1/h) \qquad (3.13)$$

$$B = 2 D_{eff} (C^*/N_1) \qquad (3.14)$$

$$\tau = (x_i^2 + Ax_i)/B \qquad (3.15)$$

τ corresponds to a shift in time which corrects for the presence of an initial oxide of thickness x_i.

The growth rate is given by

$$\frac{dx_o}{dt} = \frac{B}{2x_o + A} \qquad (3.16)$$

The solution to equation 3.12 gives two distinct limiting cases :

i) When $(t + \tau) >> A^2/4B$ then $x_o^2 = B(t + \tau)$ (3.17)

Parabolic kinetics : B is referred to as the parabolic rate constant.

ii) When $(t + \tau) << A^2/4B$ then $x_o = B/A(t + \tau)$ (3.18)

Linear kinetics: B/A is referred to as the linear rate constant.

$$B/A = \frac{C^*}{N_1 (1/k_s + 1/h)} \qquad (3.19)$$

The reaction rate in those two cases is determined by different mechanisms as shown by the different expressions for B and B/A (equations 3.14 and 3.19). In the linear regime, oxygen is in excess at the interface and the oxide growth rate is controlled only by the rate at which oxygen can react with silicon. This is

indicated by the dependence of B/A on k_s. In this respect, the gas-phase mass transport coefficient, h, is assumed to be large, so that the 1/h term in the expression for B/A can be neglected. Oxide growth is then said to be surface reaction-limited.

In the parabolic regime, oxidant reacts with silicon faster than it can be supplied to the interface and the oxide growth rate is controlled by the diffusion of oxidant across existing oxide, as indicated by the D_{eff} dependence of B. Here the oxidation is said to be diffusion-limited.

Whether the oxidation is reaction-limited or diffusion-limited depends upon temperature and oxide thickness. Generally, short oxidations are reaction-limited (linear kinetics) and long oxidations are diffusion-limited (parabolic kinetics).

3.2 Examples of Oxide Growth and Relation to Theory

Oxide growth is influenced by a number of experimental factors, notably oxidation ambient, temperature, substrate crystal orientation, pressure and substrate doping concentration. Although the Deal-Grove model provides a basis for all these observed effects, it does not give a complete understanding of the formation of SiO_2. Additional and more detailed theories have therefore been developed with the increased amount of experimental oxidation data that has become available. The above effects and their relation to Deal-Grove theory are now discussed.

3.2.1 Temperature Dependence of Rate Constants

The values of B and B/A can be evaluated from experimental data by rewriting equation 3.12 as

$$x_o = \{ (t + \tau)/x_o \} B - A \qquad (3.20)$$

Hence a plot of x_o versus $(t + \tau)/x_o$ gives a straight line of slope B and an intercept on the vertical axis equal to -A. Such an analysis was carried out by Deal and Grove on oxidation data of <111> orientated silicon [6]. The results are summarised in Table 3.1. It can be seen that the parabolic rate constant B is substantially greater for wet oxidation than for dry oxidation. This is reflected in the dependence of B on the equilibrium concentration of C^* (equation 3.14), which is three orders of magnitude larger for water than for oxygen.

These early experimental measurements of the rate constants showed them to exhibit Arrhenius behaviour. The temperature dependence of B (Fig. 3.3) is similar to that of D_{eff} as might be expected from equation 3.14, i.e. B increases exponentially with temperature. The activation energies are different for wet and dry oxidations, but the calculated values of 0.71 eV and 1.24 eV respectively (units converted from the original kcal/mole used in [6]) compare well with the diffusivity through silica of H_2O (0.80 eV) and O_2 (1.17 eV).

The temperature dependence of B/A (Fig. 3.4) shows almost identical activation energies for both wet and dry oxidation (1.96 eV and 2.00 eV respectively). Since the linear reaction regime is reaction-controlled, this indicates a similar reaction mechanism for the two

Ambient	Oxidation Temp (°C)	A (Å)	B (10^3 Å^2/min)	B/A (Å/min)	τ (min)
DRY	800	3700	1.8	0.5	540.0
DRY	920	2350	8.2	3.5	84.0
DRY	1000	1650	19.5	11.8	22.2
DRY	1100	900	45.0	50.0	4.5
DRY	1200	400	75.0	187	1.6
WET	920	5000	338	67.7	0
WET	1000	2260	478	212	0
WET	1100	1100	850	773	0
WET	1200	500	1200	2400	0

TABLE 3.1 Rate constants for oxidation of <111> silicon in wet and dry oxidation. (from [6])

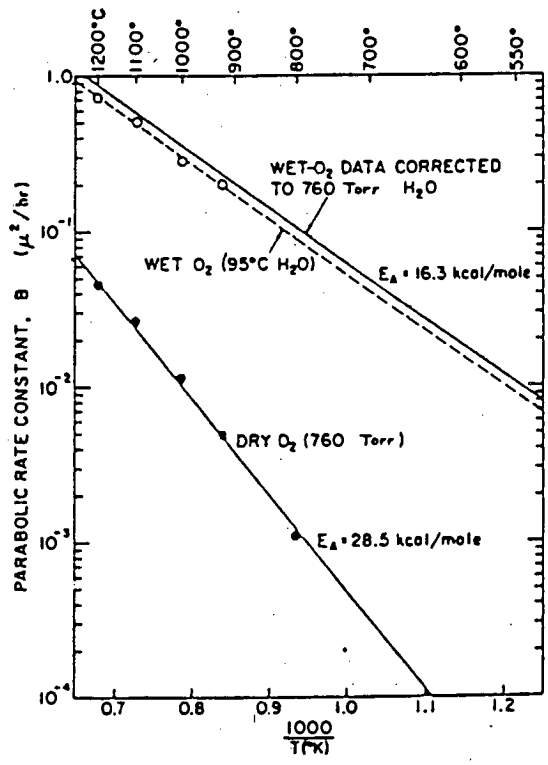


Fig. 3.3 Temperature dependence of the parabolic rate constant B (from [6]).

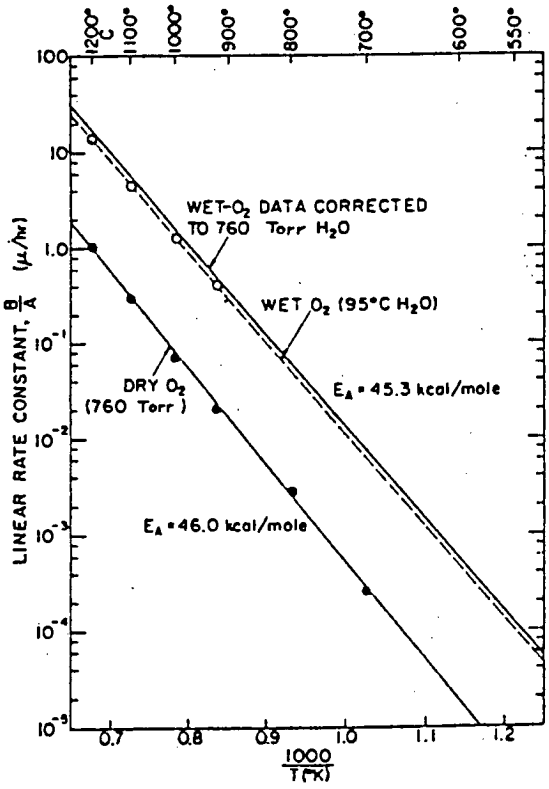


Fig. 3.4 Temperature dependence of the linear rate constant B/A (from [6]).

oxidants. The activation energies represent the temperature dependence of the interface reaction-rate constant k_s, and approximate to the Si-Si bond energy of 2.00 eV.

More recent investigations [58] were extended to include thinner oxides and lower oxidation temperatures than those previously studied. It was found that the Arrhenius plots had considerable curvature, indicating that the reaction mechanism may be more complex than that portrayed by the Deal-Grove model (this is discussed further in Section 3.2.5.). However, oxide growth data collected by this author (see Chapter 7), for the temperature range 900°C to 1025°C and thickness range 100Å to 4000Å, showed no clear curvature within the distribution of points in the Arrhenius plots, and a linear fit was used. The controversy here is due to propagation of error in the parameter extraction method and the general insensitivity of the Deal-Grove model to the parameter values used. In a recent paper [59] it was shown that a broad range of B and B/A combinations could give model fitting errors smaller than experimental error and that the existence of breakpoints in the Arrhenius plots could therefore be an artifact of the analysis method. In Chapter 7 two different parameter sets are shown to give comparable fitting errors, confirming the model insensitivity to the parameter values used.

3.2.2 Influence of Silicon Crystal Orientation

The data in Fig 3.5 show how oxidation kinetics depend upon the crystallographic orientation of the silicon surface. Orientation determines the silicon surface atom concentration. It is logical that this

will influence the interface reaction rate constant k_s and hence the linear rate constant B/A. In comparison, the parabolic rate constant would be expected to be independent of silicon orientation. These properties have been confirmed [60] for the wet oxidation of <111> and <100> orientation silicon (Table 3.2). The linear rate constants for <111> silicon are 1.68 times the values for <100> silicon at all corresponding temperatures, indicating a similar enhancement mechanism.

3.2.3 Pressure Dependence of Oxidation Rate

The expressions for both the linear and parabolic rate constants (equations 3.14 and 3.19) show a dependence on C^*, the equilibrium concentration of oxidant in the oxide. Since C^* is proportional to oxidant partial pressure (assuming Henry's Law - equation 3.4), both B and B/A are expected to have a linear partial pressure dependence.

Deal and Grove [6] presented confirmation of this for the parabolic rate constant based on the experimental data of Flint [61]. The argument has been extended to high pressure oxidation [62,63], where investigations have shown a linear pressure dependence of B up to 20 atmospheres for both wet and dry oxidation (Figs. 3.6 and 3.7). However, the same investigations have shown that the (high) pressure dependence of B/A is only linear for wet oxidation: in dry oxidation a sub-linear pressure dependence of B/A is observed (Fig. 3.8), with proportionality to the power of n where $0.7<n<0.8$. Similar results for pressures less than one atmosphere have also been found [64].

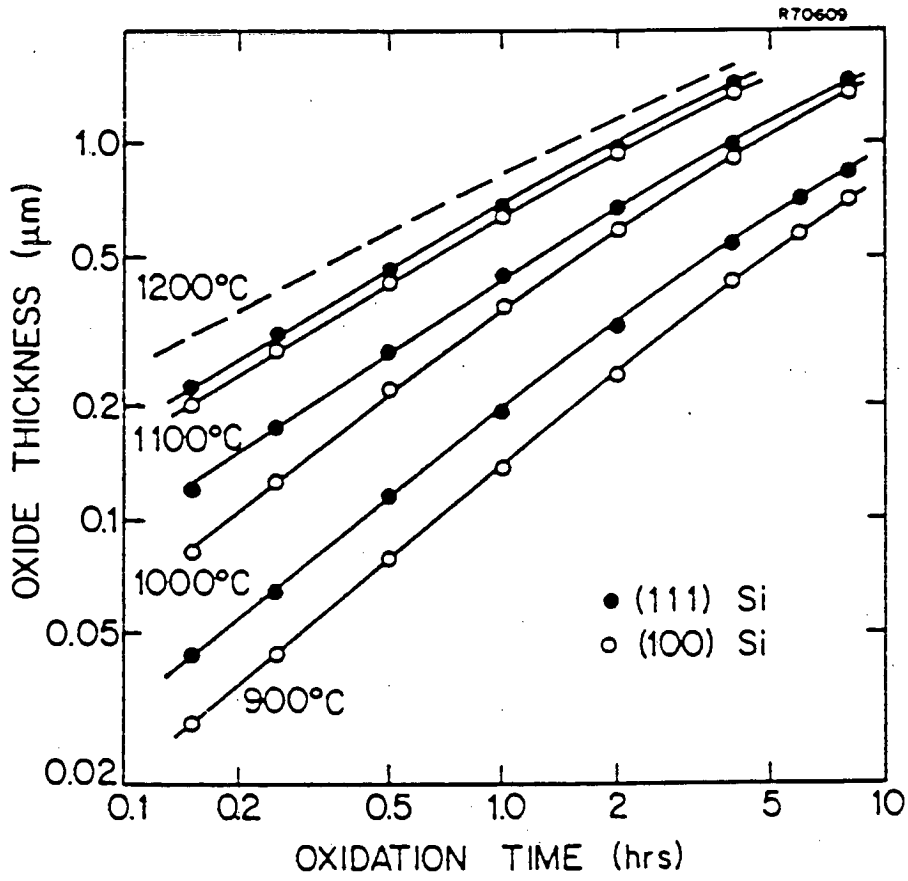


Fig. 3.5 The effect of crystal orientation on the oxidation rate of silicon. Data shown is for <111> and <100> silicon under wet oxidation conditions. (from [60])

Temp(°C)	Silicon Orientation	B (10^5 Å^2/min)	B/A (Å/min)	B/A ratio <111>/<100>
900	<100>	2.38	25.0	
	<111>	2.52	42.0	1.68
950	<100>	3.85	51.8	
	<111>	3.85	87.3	1.68
1000	<100>	5.23	110.7	
	<111>	5.23	193.8	1.75
1050	<100>	6.88	233.3	
	<111>	6.88	384.5	1.65
1100	<100>	8.68	496.2	
	<111>	8.62	821.0	1.65
			Average	1.68

TABLE 3.2 Rate constants for wet oxidation of <100> and <111> silicon. (from [60])

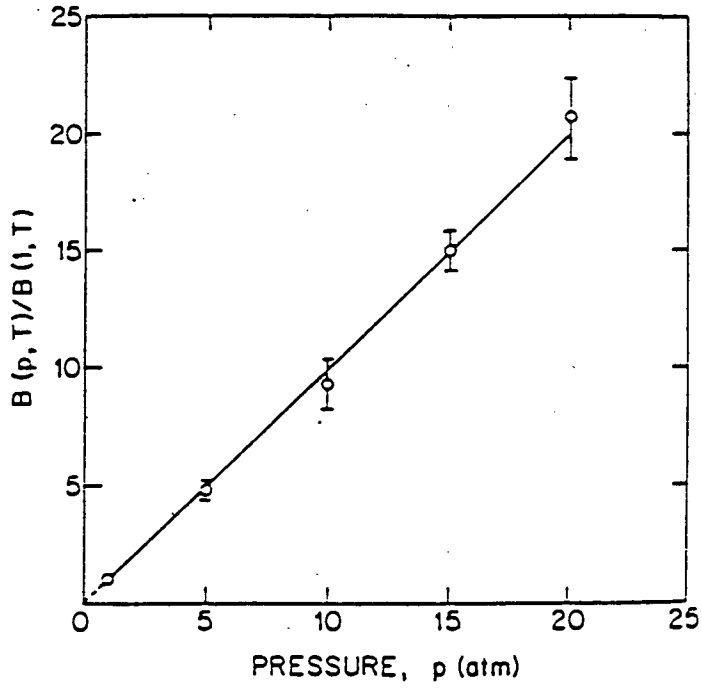


Fig. 3.6 Pressure dependence of B for wet oxidation. Data is normalised to P=1 atm. and is for the oxidation of <100> and <111> silicon in the temperature range 800-1000°C (from [62]).

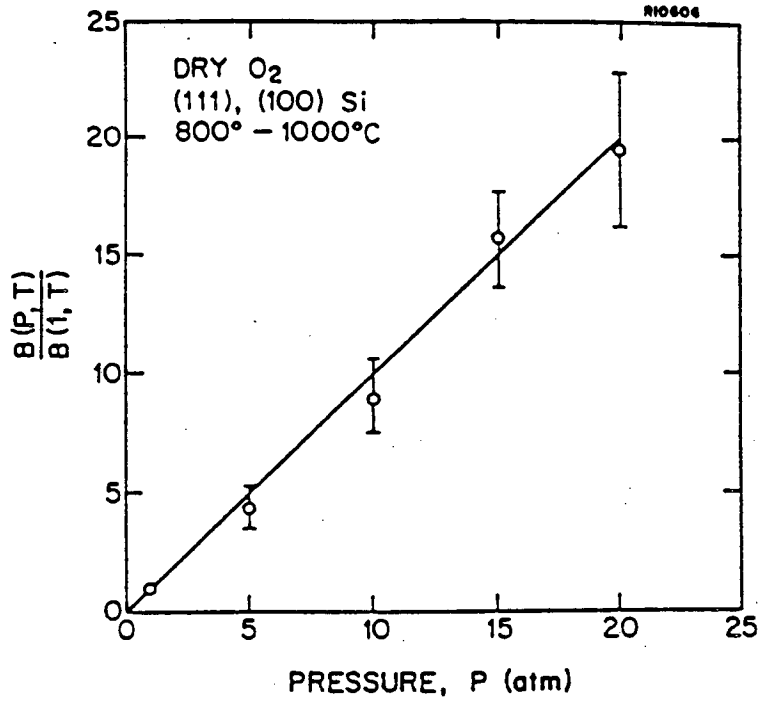


Fig. 3.7 Pressure dependence of B for dry oxidation. Data is normalised to P=1 atm. and is for the oxidation of <100> and <111> silicon in the temperature range 800-1000°C (from [63]).

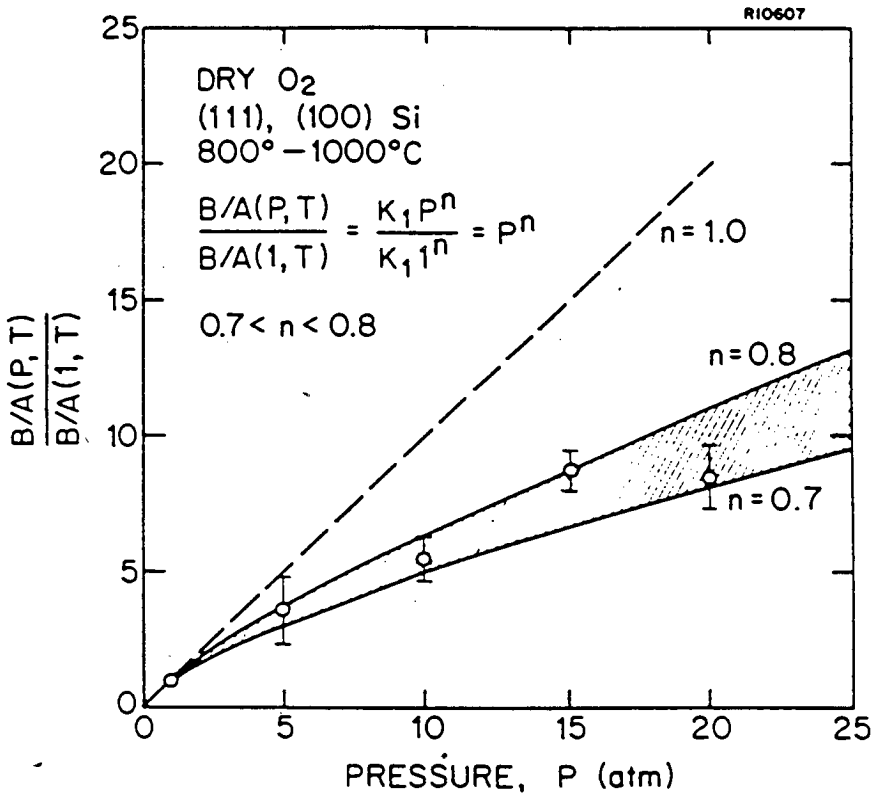


Fig. 3.8 Pressure dependence of B/A for dry oxidation. Data is normalised to P=1 atm. and is for the oxidation of <100> and <111> silicon in the temperature range 800-1000°C (from [63]).

The linear pressure dependence of both B and B/A in wet oxidation indicates the species diffusing in the oxide is molecular H_2O (i.e. no dissociation occurs at the gas-oxide interface). Similarly, the linear pressure dependence of B in dry oxidation indicates the diffusing species is molecular O_2. This is of course contradicted by the sub-linear pressure dependence of B/A in dry oxidation. Chez and Van der Meulen [65] have suggested that this discrepancy may be due to the dissociation of molecular oxygen at the $Si-SiO_2$ interface, yielding atomic oxygen, and that both species (molecular and atomic oxygen) then undergo reaction at the interface. The relative contribution of these species accounts for the observed pressure dependence. A similar explanation by Blanc [66] suggests the interface reaction occurs by atomic oxidation only, but this results in a square-root dependence on oxygen partial pressure (n=0.5) and is therefore inaccurate. That a full physical explanation of the B/A pressure dependence is still lacking, further reflects the complexity of the oxidation reaction mechanism, and suggests the Deal-Grove model is over-simplified.

3.2.4 Influence of Substrate Doping

During oxidation, substrate doping elements such as boron and phosphorous redistribute at the interface so that their chemical potential is equal on each side of the interface. This results in an abrupt shift in dopant concentration across the interface or, expressed differently, the dopant segregates either into the oxide or into the silicon substrate. It is defined quantitatively by the segregation coefficient, m :

$$m = \frac{\text{concentration of dopant in Si}}{\text{concentration of dopant in } SiO_2} \qquad (3.21)$$

Boron is observed to segregate into oxide ($m < 1$) as shown in Fig. 3.9(a). The segregation coefficient increases with increasing temperature and is greater for <100> orientation than for <111> orientation silicon [67]. It has also been found to be greater for dry oxidation than for wet oxidation [68]. In contrast to boron, phosphorus segregation is into the silicon ($m > 1$) as shown in Fig. 3.9(b), a phenomenon known as 'pile-up'.

The oxidation rate of silicon is found to be enhanced by high concentrations of substrate doping elements [70] such as might be found in the source and drain regions of MOS devices. Fig. 3.10 shows the effect of high phosphorus doping levels on the dry oxidation kinetics of <111> silicon. Phosphorus enhances the rate most effectively at lower temperatures. In contrast, boron gives a greater rate enhancement at high temperatures (>1050°C). This temperature dependence of the rate enhancement has been correlated with dopant segregation effects [58]. The enhancement for both boron and phosphorus is principally accounted for by an increase in the linear rate constant B/A [58,71] and the parabolic rate constant B remains relatively unchanged (Fig. 3.11). This relates to an influence on the Si-SiO_2 interface reaction rate constant k_s. Further explanation than this requires greater understanding of underlying physical mechanisms of the interface reaction, which are conveniently lumped together in the parameter k_s, and this is discussed in the next section.

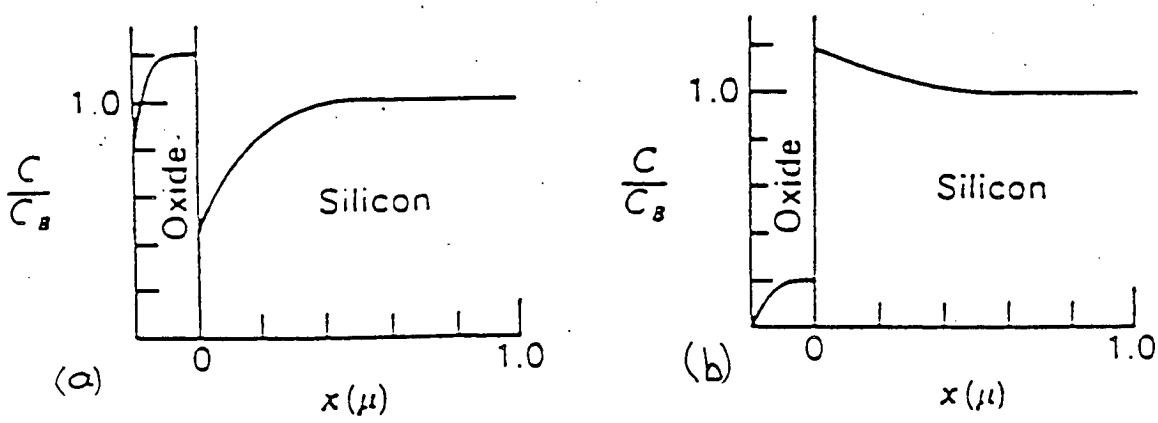


Fig. 3.9 Impurity concentration (after oxidation) against depth, for initially uniformly doped silicon : (a) m<1 (boron) - oxidation takes up impurity (b) m>1 (phosphorus) - oxide rejects impurity. (from [69])

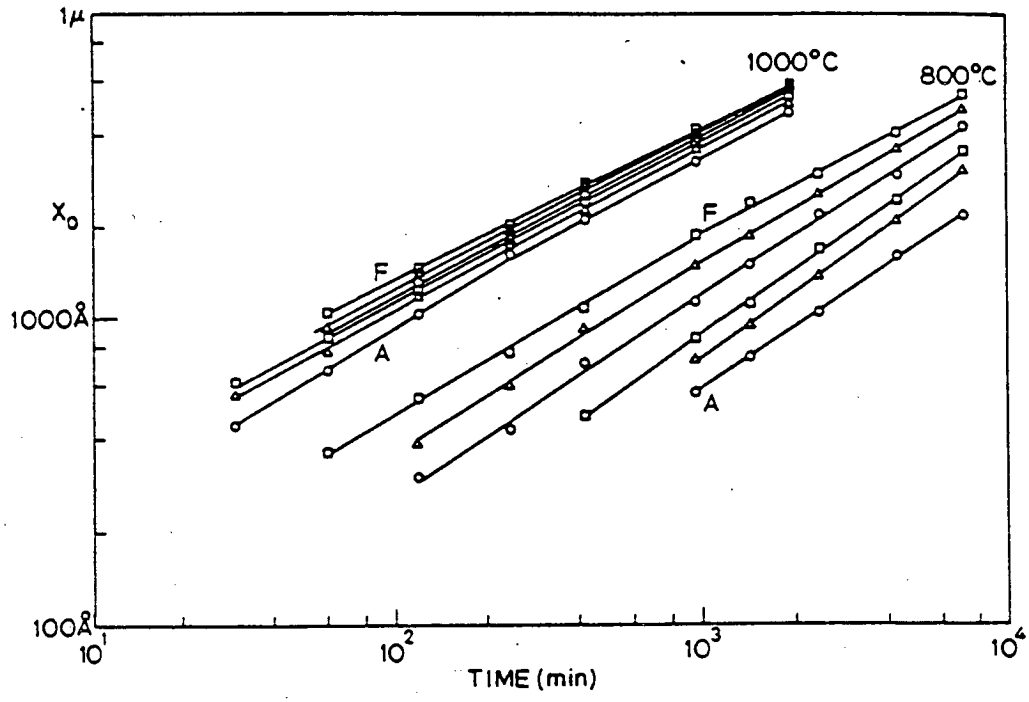


Fig. 3.10 Oxide thickness against time for <111> silicon under dry oxidation conditions. A to F indicate increasing phosphorus doping concentrations : A=$1x10^{15}$, B=$5.1x10^{19}$, C=$7.2x10^{19}$, D=$1.8x10^{20}$, E=$3.2x10^{20}$ (units in cm^{-3}). (from [71])

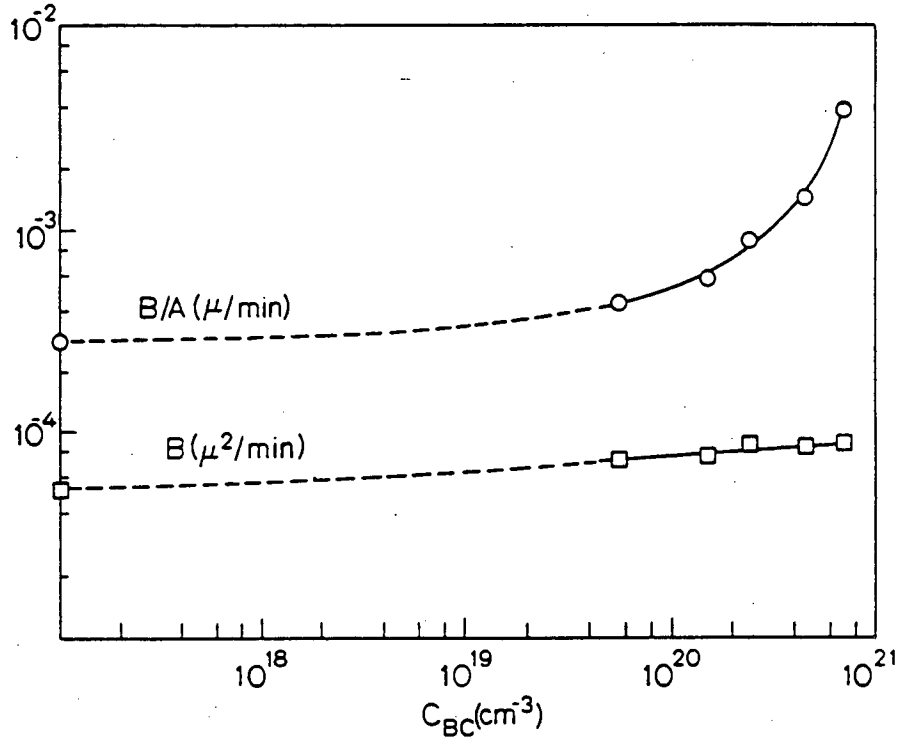


Fig. 3.11 Effect of phosphorous doping concentration on the rate constants B and B/A (data for dry oxidation of <111> silicon at 900°C). (from [71])

As is seen in Fig. 3.11, B is slightly enhanced with increasing dopant concentration, but the effect is much less distinct than that for B/A. It reflects an influence on the oxidant diffusivity and thus modification of the term D_{eff} in the expression for B (equation 3.14). Oxidant diffusivity is enhanced because incorporation of dopant in the oxide gives a loosened SiO_2 structure [58].

3.2.5 Mechanism of Oxide Growth

There is still uncertainty about the fundamental mechanisms of the Si-SiO_2 interface reaction. Certainly the linear-parabolic model offers no physical insight into details of the interface reaction mechanism, successful as it is in describing many of the observed kinetics. It is clear, however, that for oxidation to proceed, a quantity of 'free volume' must be supplied, as evidenced by the volume change that occurs when silicon becomes SiO_2. If insufficient volume is available, intrinsic stress develops in the oxide. This stress has been observed in oxides grown at low temperatures (<950°C) [72]. The absence of such stress in high temperature oxidations has been explained by Irene et al. [73] based on a visco-elastic model. In this model, high temperatures lower the oxide viscosity and allow oxide to flow in a direction normal to the wafer surface. This provides free volume and so relieves intrinsic stress. At low temperatures, the viscosity is too high to relieve all the stress, which can then accumulate. Note that intrinsic stress should not be confused with thermal expansion stress which is zero at oxidation temperature and only develops as the sample cools down due to mismatch in the thermal expansion coefficients between silicon and SiO_2.

The role of stress on oxide growth is further evidenced by densification of oxides grown at low temperatures [74]. Densification occurs as a mechanism to alleviate the build-up of stress and has been correlated with the increase in the refractive index of oxides grown at low temperatures [75,76]. The intrinsic stress would be expected to influence both the diffusion of oxidant through the oxide and the interface reaction mechanism, and therefore affect the parabolic and linear rate constants. Recent experiments [77,78] have confirmed this, and it may explain, at least in part, the non-Arrhenius behaviour of the rate constants (Section 3.2.1). Oxide stress can lead to structural damage in the silicon and adversely affect the performance of devices.

Fig. 3.12 shows a detailed model of the Si-SiO_2 interface reaction developed by a group at Stanford University. As oxidation proceeds and the interface progresses into the silicon, it leaves an excess of silicon atoms in each lattice plane [79,80]. This causes lattice mismatch at the interface which is stored as dislocation lines (dangling bonds) and strain in the SiO_2 film. It is represented by the top reaction in Fig. 3.12. Away from the interface, the strain is relieved by viscous flow that provides the required free volume, as discussed above.

Alternative methods of providing free volume at the interface region itself are shown by the bottom two reactions in Fig. 3.12. These are based on point defects, representing the presence of a silicon vacancy (Si_V) in the substrate and the removal of silicon atoms from lattice sites to create interstitials (Si_I) respectively. Point defects exist in several charge states [81] which have energy levels in the band gap, so

that their concentration depends upon the position of the Fermi level (Fig. 3.13). A shift in the position of the Fermi level will therefore change the number of point defects and so influence the oxidation. This can explain the observed rate enhancement with high phosphorous doping levels. The increased doping level shifts the Fermi level nearer the conduction band, resulting in an increased silicon vacancy concentration which affects the rate through the middle reaction of Fig. 3.12. A quantitative model based on this mechanism has been developed by Ho and Plummer [81,82] and has given a good fit to their data.

Although the model of Fig. 3.12 is not necessarily comprehensive in its treatment of interface mechanisms, it has explained several other oxidation-related phenomena. In particular, the creation of silicon interstitials has been linked to the growth of oxidation-induced stacking faults (OISF) [83] and the mechanism of oxidation-enhanced dopant diffusion [84]. All three mechanisms in Fig. 3.12 have experimental evidence to support them, as discussed above. It is reasonable to assume that all take part concurrently in the oxidation process, although any one may be dominant under specific conditions.

A final point concerns the nature of the oxidising species in dry oxidation. Diffusion through the oxide has been generally assumed to be by molecular neutral oxygen (O_2). However, early experiments showed that an applied electric field can influence the oxidation rate [56], indicating a charged oxygen ion may be involved. This is a matter of much controversy, with conflicting evidence for both neutral and charged diffusion species [85,86], and has not yet been fully resolved.

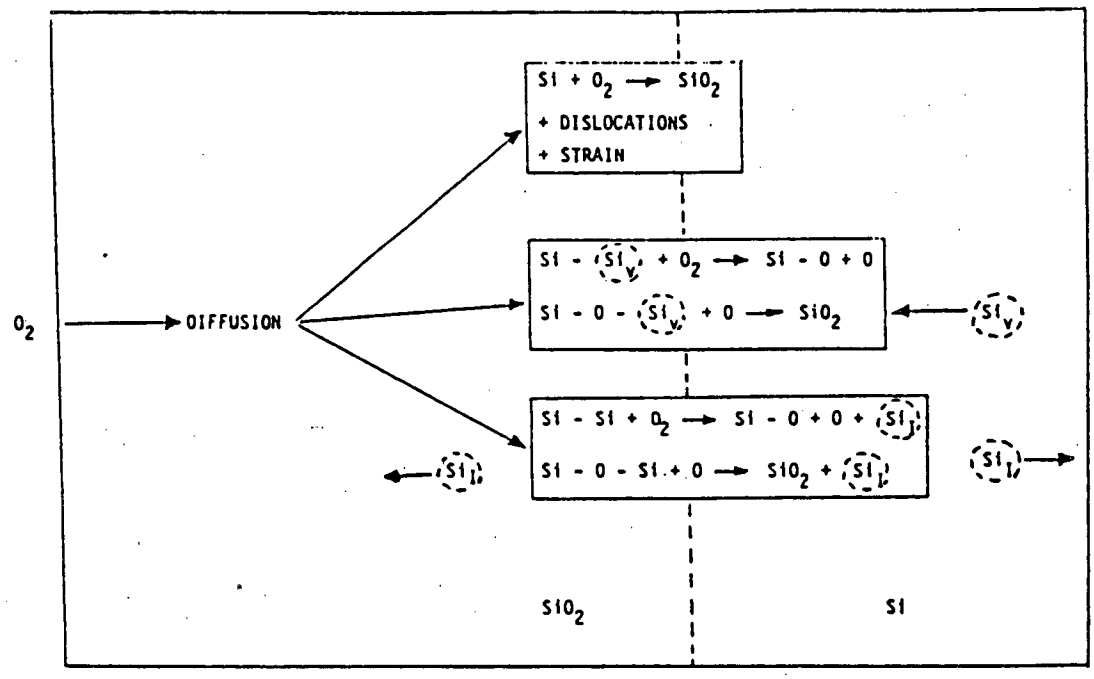


Fig. 3.12 Detailed model for silicon oxidation, showing the means of providing free volume: silicon interstitials (Si_I) and silicon vacancies (Si_V).

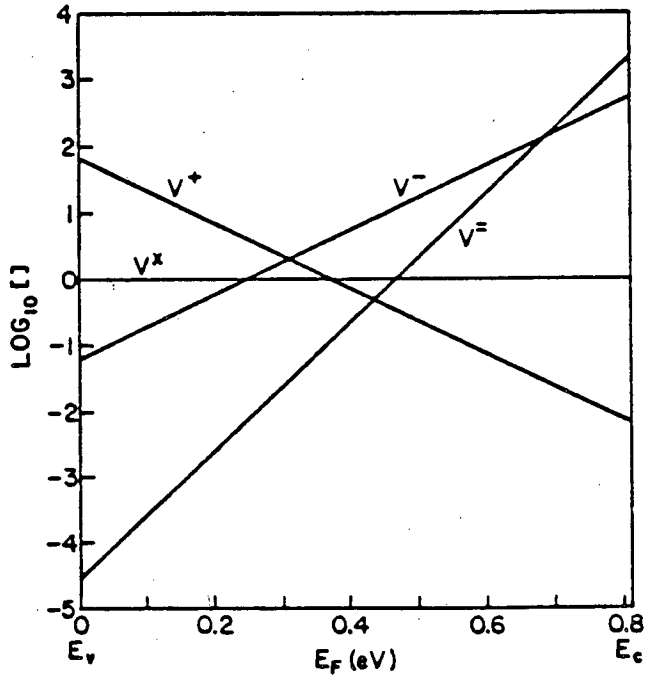


Fig. 3.13 Concentration of charged silicon vacancies as a function of the silicon Fermi level.

3.3 Kinetics of Thin Dry Oxidation

Table 3.1 shows the values of τ (equation 3.15) obtained by Deal and Grove from the extrapolation of oxide growth data to the horizontal axis. While wet oxidation can be seen to yield a value of τ equal to zero, corresponding to no initial oxide, dry oxidation yields non-zero values of τ. This implies the apparent existence of an oxide layer at zero time, the thickness of which is about 230Å [6]. In reality, the first few hundred Angstroms of oxide grow at an anomalously high rate and linear-parabolic kinetics are not fully established until after this growth. A different physical process must therefore operate during the initial regime. This is also manifested by different optical and physical properties in thin oxides [87,88]. Oxidation studies by Massoud et al. [89], using in-situ ellipsometry, clearly show the thin oxide rate enhancement over and above the expected result from Deal-Grove kinetics (Fig. 3.14).

Many workers have studied the initial oxidation regime and there are several models in the literature to explain its growth [6,89-94]. Deal and Grove [6] proposed an explanation based on the existence of an electric field across the oxide in the early stages of growth. An ionic oxidant species would then undergo enhanced diffusion, and they have suggested singly ionised molecular oxygen (O^-_2) as a possibility. Similarly Hu [90] has suggested that a flux of charged atomic oxygen (O^-) can exist in parallel to his proposed oxidation mechanism involving the chemisorption of oxygen. This flux is only important at thin oxide thicknesses, where electronic conduction in the oxide is possible by tunnelling and there is no overall charge flux. Counteracting these explanations is the

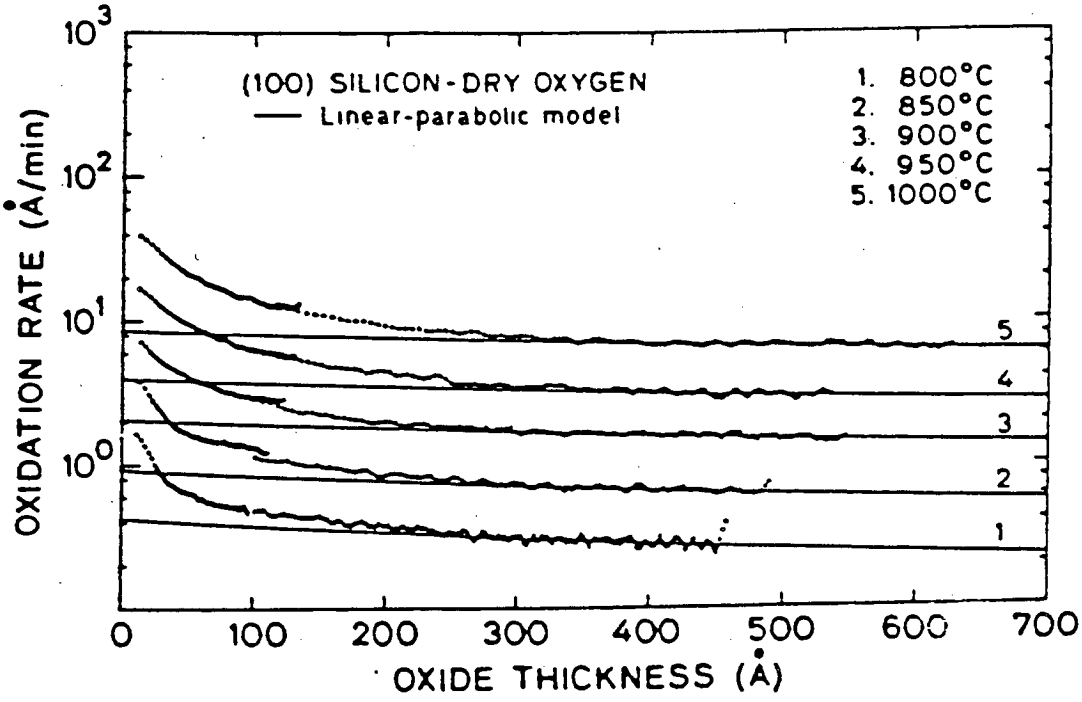


Fig. 3.14 Oxide growth rate under dry conditions of lightly doped <100> silicon, plotted against oxide thickness. The solid lines represent the rate expected from linear-parabolic kinetics. (from [89])

controversy over the existence of charged species and the fact that they both influence oxide diffusivity in a region of growth that is reaction-controlled.

Fargeix et al. [91,92] have proposed a model based on the influence of intrinsic stress on the diffusivity of the oxidising species. A decreased diffusivity nearer the interface is suggested, as a result of increasing stress, and this dominates oxidant diffusion. Again, this model is based on oxidant transport phenomena in a region where growth is reaction-controlled. The role of structural inhomogeneities (i.e. micropores) in thin oxides has also been investigated [93,94]. It is proposed that micropores of around 10Å diameter provide a short circuit path to the Si-SiO_2 interface for oxidant molecules. Lateral diffusion would then lead to a higher concentration of oxygen at the interface and an enhanced growth rate. The effect is not observed for wet oxidation because of greater interaction between H_2O and SiO_2. This forms SiOH species which clog the micropores. (This mechanism was also used to account for an improved dielectric breakdown in wet oxides [93].)

In a study involving oxidation in $^{18}O_2$ and Secondary Ion Mass Spectroscopy (SIMS) [95], it has also been suggested that a second oxidation mechanism dominates with thin oxides and so enhances the rate. The mechanism involves the diffusion of oxygen vacancies away from the Si-SiO_2 interface and their subsequent reaction with adsorbed oxygen at the SiO_2 surface. This mechanism was used to explain the presence of an ^{18}O concentration increase at the SiO_2 surface after a $^{16}O_2$ followed by an $^{18}O_2$ oxidation - an effect that would not be expected from the Deal-Grove theory. The mechanism of Ghez and Van der Meulen involving molecular and atomic oxygen species [65] (see Section 3.2.3) has also been

used to explain the thin oxide regime, as has the mechanism proposed by Blanc [66] involving only atomic oxygen (although, as already mentioned, this model is inaccurate in its O_2 partial pressure dependence and is therefore discredited).

An empirical model for thin oxide growth has been proposed by Massoud and Plummer [89] based on kinetic data obtained by in-situ ellipsometry. They were able to show that the excess oxidation rate has an exponential form (Fig. 3.15), and were able to model this by the addition of two exponential terms to the Deal-Grove expression for rate:

$$\frac{dx_o}{dt} = \frac{B}{2x_o + A} + C_1 \exp(-x_o/L_1) + C_2 \exp(-x_o/L_2) \qquad (3.22)$$

The decay length L_1 is around 10Å so the first exponential enhancement term vanishes after about 50Å. The second exponential term has a decay length L_2 of about 70Å and disappears at an oxide thickness of around 250Å, after which the growth kinetics become linear-parabolic. The first exponential term can be omitted to simplify equation 3.22, resulting in only 5-7% error. The parameters C_2 and L_2 were found to be approximately independent of substrate orientation, doping density and oxygen partial pressure. L_2 is independent of temperature and C_2 has a well-behaved temperature dependence of the form

$$C_2 = K \exp(-E_a/kT) \qquad (3.23)$$

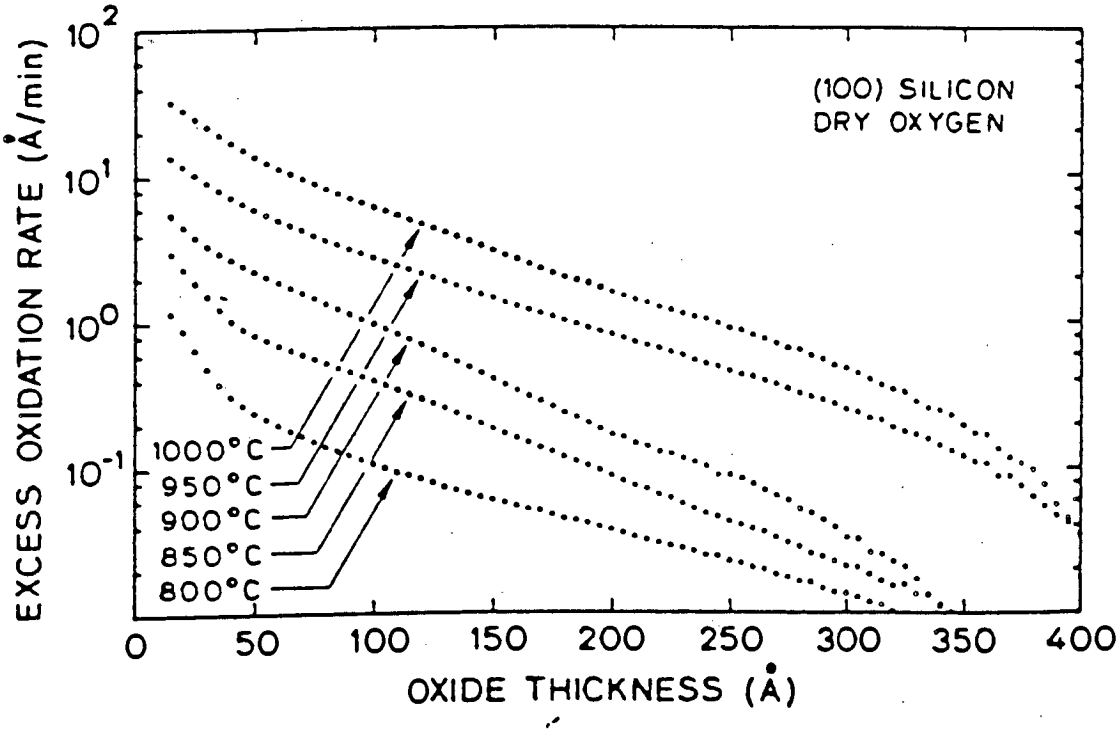


Fig. 3.15 Oxidation rate in excess of that expected from the linear-parabolic model, as a function of oxide thickness. (from [89])

A physical explanation of equation 3.22 has been proposed, based on the presence of a surface layer in the silicon substrate that contains additional sites for oxidation [96]. The nature of these sites may be weakened or broken Si-Si bonds. The possibility of oxygen diffusion into the silicon surface is also related. However, the physical origin of the parameters C_1, C_2, L_1, and L_2 was not established in [96]. The lack of enhanced growth in wet oxygen was suggested as being due to an elimination of the additional sites by hydrogen formed from the dissociation of water molecules.

The above list of models is not exhaustive but it does indicate the controversy surrounding the anomalously high initial rate of dry oxidation. The matter is not yet fully resolved. It is reasonable that the surface properties of silicon are different from its bulk properties and the steady-state assumption on which Deal-Grove kinetics are based is therefore not achieved until the (disordered) surface region has been oxidised. Hence the last explanation is favoured, at least as a general starting point for the interpretation of experimental results.

3.4 Chlorinated Oxidation

The introduction of a few percent of a chlorine-containing compound to a dry oxidising ambient has been found to improve the electrical properties of SiO_2 and has resulted in the widespread use of chlorinated oxides in semiconductor processing. Although HCl gas is the most common additive, it is a highly corrosive and dangerous gas and the use of trichloroethylene (TCE), trichloroethane (TCA) and other chlorine compounds has also been investigated.

The use of chlorinated oxides has a beneficial effect on the level of mobile ions. Passivation of sodium ions was reported by Kriegler et al. [21] who also observed a cleaning effect on the oxidation furnace tube by chlorine-containing ambients. The passivation effect virtually disappears at low temperatures (<1050°C) and is not present when HCl is added to an inert ambient [97]. TCE is observed to give similar levels of stability as HCl at eight times lower concentration [98]. A reduction in interface trapped charge density has also been observed through the use of HCl, TCE and TCA [99-101].

Higher and more uniform dielectric strengths have been reported for oxides grown in dry oxidation environments containing HCl, TCE, TCA and Cl_2 [101,102]. This is attributed to a reduction in oxide defect density. The extent of the effect depends on both the additive used and its gas concentration. The addition of HCl has been found to increase the minority carrier lifetime [103]. This is thought to be due to the removal of metal impurities through the formation of their volatile chlorides. A suppression of oxidation induced stacking faults (OISF) is observed in chlorine-containing ambients at high temperatures (>1050°C) and this may also be related to the gettering of metal impurities.

The oxidation kinetics of silicon are directly influenced by the addition of chlorine-containing compounds. Results for HCl/O_2 mixtures show a monotonic increase in growth rate as the HCl concentration is increased (Fig 3.16). This is observed at several temperatures and for both <100> and <111> silicon [104]. Both the linear and parabolic rate constants are found to increase with the addition of HCl (Figs. 3.17 and 3.18), though B/A saturates after the addition of 1% HCl.

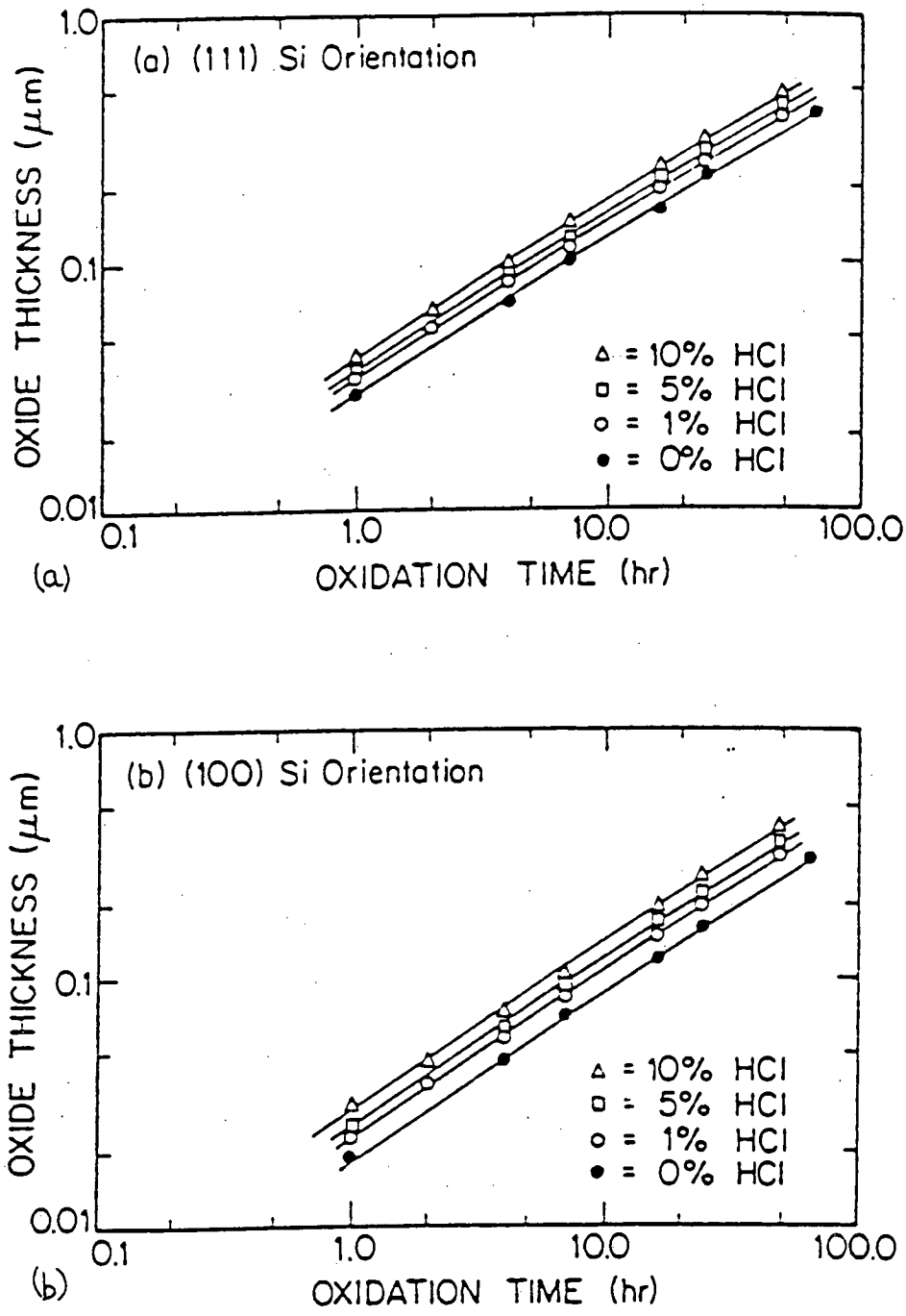


Fig. 3.16 Oxide growth in HCl/O_2 mixtures at 900°C for (a) <111> and (b) <100> silicon. (from [104])

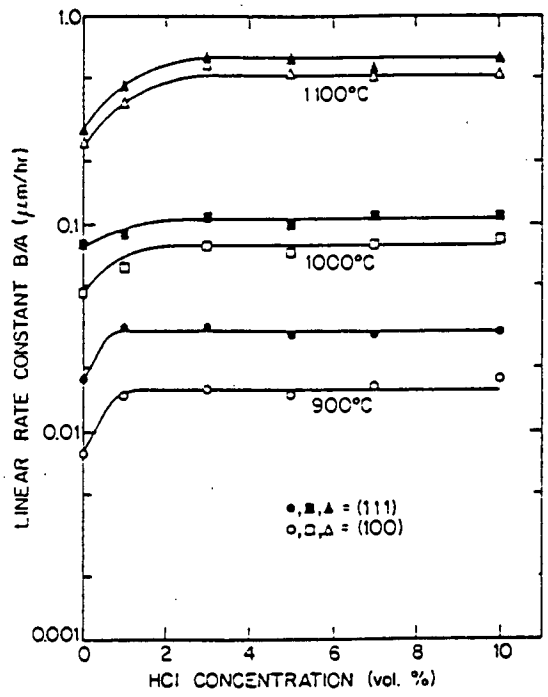


Fig. 3.17 The effect of HCl concentration on the linear rate constant B/A (data for dry oxidation of <111> and <100> silicon at 900°C, 1000°C and 1100°C). (from [104])

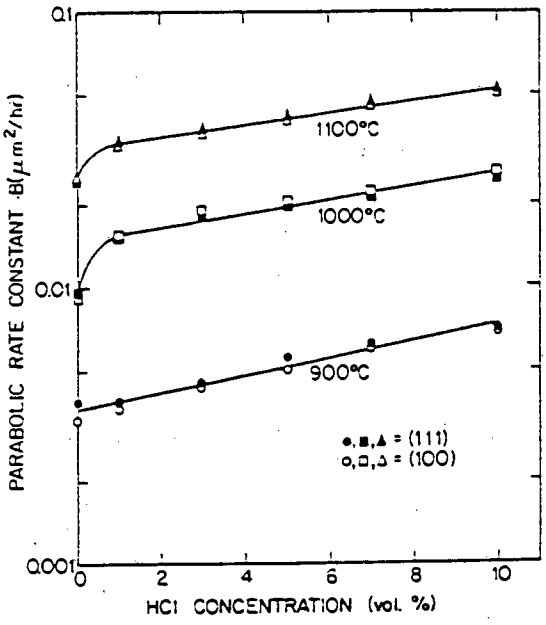


Fig. 3.18 The effect of HCl concentration on the parabolic rate constant B (data for dry oxidation of <111> and <100> silicon at 900°C, 1000°C and 1100°C). (from [104])

The mechanism of this enhanced growth is not fully understood. However it is known that H_2O and Cl_2 are formed in an HCl/O_2 atmosphere according to the reaction

$$4HCl + O_2 \longleftrightarrow 2H_2O + 2Cl_2 \qquad (3.24)$$

Although H_2O generation would be expected to increase the overall rate, gas-phase equilibria calculations [105] show that this is insufficient to account for the total observed enhancement. Of course the equilibria may be different in the solid SiO_2 phase. Addition of Cl_2 to O_2 increases the oxidation rate [106] so the Cl_2 produced by equation 3.24 must also influence the kinetics of HCl oxidation. It is known by SIMS analysis that chlorine piles up near the $Si-SiO_2$ interface [107,108]. Its role is therefore thought to involve bonding in Si-O-Cl complexes by replacing oxygen atoms at the interface. This would be expected to influence the interface reaction rate and therefore B/A, as well as straining the SiO_2 network so that oxidant diffusion is enhanced.

A similar rate enhancement is observed for the TCE/O_2 system [109], where an overall combustion reaction must be considered:

$$4C_2HCl_3 + 9O_2 \longleftrightarrow 2H_2O + 6Cl_2 + 8CO_2 \qquad (3.25)$$

$$2H_2O + 2Cl_2 \longleftrightarrow 4HCl + O_2$$

Under conditions of high temperature and long oxidation time the use of chlorine compounds can be detrimental due to the etching effect of Cl_2 on the

silicon surface. This is accompanied by the formation of gas bubbles or even flaking of the oxide from the silicon. The effect is worse for TCE than HCl or TCA since its behaviour is more chlorine-like.

In contrast to the behaviour under dry conditions, the addition of HCl to a wet oxidising ambient is found to decrease the oxidation rate [60]. This is due to a reduction of the H_2O partial pressure by dilution. The SIMS analysis of oxides prepared in H_2O/HCl ambients show that no appreciable amounts of chlorine are incorporated in the oxide and that there is no chloride peak at the interface (such as that obtained in O_2/HCl oxidising ambients) [107]. Although HCl additions to wet oxidation ambients do not getter, higher quality oxides are still obtained due to a reduction of impurity contamination in the furnace system. A decrease in oxidation rate has also been observed for the addition of TCE to a wet oxidation ambient [110], similar to that observed with HCl addition.

3.5 Simulation of Oxide Growth

A requirement of any simulation is the ability to model accurately over as wide a range of operating conditions as possible. This means taking account of all major parameters which may influence the process being simulated. In the case of oxide growth, the preceding sections have identified the most important parameters. These are:

i) Oxidation ambient (i.e. wet or dry)
ii) Temperature
iii) Oxidant pressure or partial pressure
iv) HCl concentration (dry oxidation only)

v) Substrate orientation

vi) Substrate doping level at the $Si-SiO_2$ interface

While the Deal-Grove model provides a good basis for simulating the effect of these parameters, it has many deficiencies. Although additional theories and models have evolved, there remain aspects of silicon oxidation where our understanding is not yet complete. The extent to which the oxidation models incorporated into process simulation programs have coped with this is now discussed.

The one-dimensional (1-D) simulation program SUPREM II [52], developed at Stanford University, utilises an incremental form of the Deal-Grove relationship (equation 3.12),

$$\Delta x_o = \frac{-(A + 2x_o) + ((A + 2x_o)^2 + 4B\,\Delta t)^{\frac{1}{2}}}{2} \qquad (3.26)$$

A and B are calculated from exponential expressions for the linear and parabolic rate constants.

$$B/A = \{C_L \exp(-E_{aL}/kT)\}.\{1 + \gamma\,(V_T - 1)\}.P \qquad (3.27)$$

$$B = \{C_P \exp(-E_{aP}/kT)\}.\{1 + \delta\,C_T^{0.22}\}.P \qquad (3.28)$$

The first terms in the equations 3.27 and 3.28 are intrinsic rate constants (i.e they apply to low doping concentrations). The values of C_L, E_{aL}, C_P and E_{aP} depend on the oxidation ambient and silicon crystal orientation, and SUPREM II incorporates default values

based on the experimental results of Deal [60] (Table 3.3).

The second term in equation 3.27 models the increase in B/A observed at high doping levels. It is the form proposed by Ho and Plummer [81], and is based on the total silicon vacancy concentration, V_T. C_T (equation 3.28) is the total n-type dopant concentration, and is used to model the dopant effect on B. γ and ζ are experimentally determined parameters. The form of equation 3.26 allows B and B/A to be time dependent and the influence of changes in the dopant concentration (due to diffusion and segregation) can be modelled. The simulation proceeds in time increments Δt, and from the calculated values of A and B, the corresponding increment of oxide thickness Δx_o is obtained.

The rate enhancement observed with thin dry oxides is rather crudely modelled in SUPREM II. Since the enhancement is in a regime where growth is reaction-controlled, the constant C_L is increased. Based on experimental evidence, it is increased by a factor of ten if the oxide thickness is less than 200Å, which results in a discontinuity. SUPREM II is thus clearly not suitable for accurate modelling of thin oxide growth.

The oxidation model in ICECREM [53] is similar to that in SUPREM II, at least in its validity. The main differences are:

i) The default values of C_L, E_{aL}, C_P and E_{aP} are slightly altered (Table 3.4).

ii) A phenomenological model is used to describe substrate doping enhancement. This takes the form of an oxidation enhancement factor, R.

Ambient	Silicon Orientation	C_L (10^9 Å/min)	E_{aL} (eV)	C_P (10^9 Å^2/min)	E_{aP} (eV)
DRY	<111>	1.04	2.0	1.29	1.23
DRY	<100>	0.62	2.0	1.29	1.23
WET	<111>	27.2	2.05	0.64	0.78
WET	<100>	16.2	2.05	0.64	0.78

TABLE 3.3 Default values for C_L, E_{aL}, C_P, E_{aP} used in SUPREM II. (from [60])

Ambient	Silicon Orientation	C_L (10^9 Å/min)	E_{aL} (eV)	C_P (10^9 Å^2/min)	E_{aP} (eV)
DRY	<111>	1.04	2.00	1.26	1.23
DRY	<100>	0.62	1.99	1.34	1.23
WET	<111>	27.2	2.05	0.64	0.78
WET	<100>	17.3	1.96	0.37	0.71

TABLE 3.4 Default values of C_L, E_{aL}, C_P and E_{aP} used in ICECREM.

$$R = 1 + \sum \frac{C_i}{C_{Si}} \exp\,(E_i/kT) \qquad (3.29)$$

where C_i are the concentrations of various dopants and C_{Si} is the atomic density of silicon. E_i is an empirically chosen constant.

iii) Although an incremental form of Deal-Grove's equation is used, the oxide thickness x_o is obtained directly without calculation of a thickness increment Δx_o.

iv) Thin dry oxides are not modelled and ICECREM assumes an initial thickness of x_i=230Å.

Both the above models are limited in their validity. Particular areas of concern in VLSI oxidation such as thin dry oxides and chlorinated oxidation are not covered. The more advanced 1-D simulator SUPREM III incorporates empirical models that covers these areas [5]. SUPREM III also incorporates modified terms that model non-Arrhenius behaviour of the rate constants and the sub-linear pressure dependence in dry oxidation. The expressions used in SUPREM III are:

$$B/A = lin_i . lin_p . lin_c . lin_{HCl} \qquad (3.30)$$

$$B = par_i . par_p . par_c . par_{HCl} \qquad (3.31)$$

lin_i and par_i are the intrinsic rate constants and have the same form as those used in SUPREM II:

$$lin_i = C_L \exp(-E_{aL}/kT) \quad (3.32)$$

$$par_i = C_P \exp(-E_{aP}/kT) \quad (3.33)$$

In wet oxidation both lin_i and par_i use two sets of coefficients, the change-over between them occurring at 900-950°C (Table 3.5). The data used are based on the results of Razouk et al. [62]. The coefficients for dry oxidation do not have a breakpoint and are the same as those used in SUPREM II.

lin_p and par_p model the pressure dependence and are of the form

$$lin_p = 1/2 \ (P_i^n + P_{i-1}^n) \quad (3.34)$$

$$par_p = 1/2 \ (P_i + P_{i-1}) \quad (3.35)$$

where n=0.75 for dry oxidation and n=1 for wet oxidation.

The influence of high doping concentrations is modelled by lin_c and par_c. These coefficients are similar to the expressions used in SUPREM II.

A quantitative physical model for the influence of HCl is lacking and SUPREM III uses a look-up table to model this. The values of lin_{HCl} and par_{HCl} are stored in the table as functions of %HCl and temperature.

To simulate thin dry oxides, a simplified version of the empirical fit by Massoud et al. [89] was used. This is equation 3.22 with the removal of the first exponential. Hence the form of Deal-Grove expression used in SUPREM III is :

Temperature (°C)	Silicon Orientation	C_L (10^9 Å/min)	E_{aL} (eV)
T > 900	<111>	29.5	2.05
	<100>	17.5	2.05
T < 900	<111>	0.345	1.60
	<100>	0.205	1.60

Temperature (°C)	Silicon Orientation	C_P (10^9 Å^2/min)	E_{aP} (eV)
T > 950	<111>	0.700	0.78
	<100>	0.700	0.78
T < 950	<111>	28.3	1.17
	<100>	28.3	1.17

TABLE 3.5 Default values for C_L, E_{aL}, C_P E_{aP} used in SUPREM III wet oxidation.

$$\frac{dx_o}{dt} = \frac{B}{2x_o + A} + C \exp(-x_o/L) \qquad (3.36)$$

Two dimensional models of oxidation are now being investigated [55] and although they have great importance for the simulation of VLSI structures, they are outside the scope of this project which is primarily concerned with planar oxide growth.

CHAPTER 4

PRACTICAL ASPECTS OF OXIDE GROWTH

4.1 Furnace Description

In IC manufacturing the oxidation of silicon is normally carried out in horizontal hot-wall furnaces. These are commonly referred to as 'diffusion' furnaces since their original use was the diffusion of impurities in silicon, but they are now mainly used in oxidation applications [111]. Diffusion furnaces allow large batches of silicon wafers to be processed, coupling high throughput with moderate costs and reliability. Despite alternative methods of producing layers of SiO_2 - such as the rapid thermal processing, high pressure oxidation and plasma enhanced oxidation techniques discussed in Chapter 2 - and the recent arrival of vertical furnace processing systems [112], it seems likely that the horizontal diffusion furnace will continue to have a major role.

Furnace systems are produced by several manufacturers, resulting in a variety of models on the market, but there are many common features. A typical state-of-the-art system is shown in Fig. 4.1. Furnace heating, gas flows and wafer loading/unloading are all controlled through a microprocessor-based process controller. The process controller also provides the necessary sequencing and co-ordination of furnace operations. The individual parts of the furnace system are now discussed.

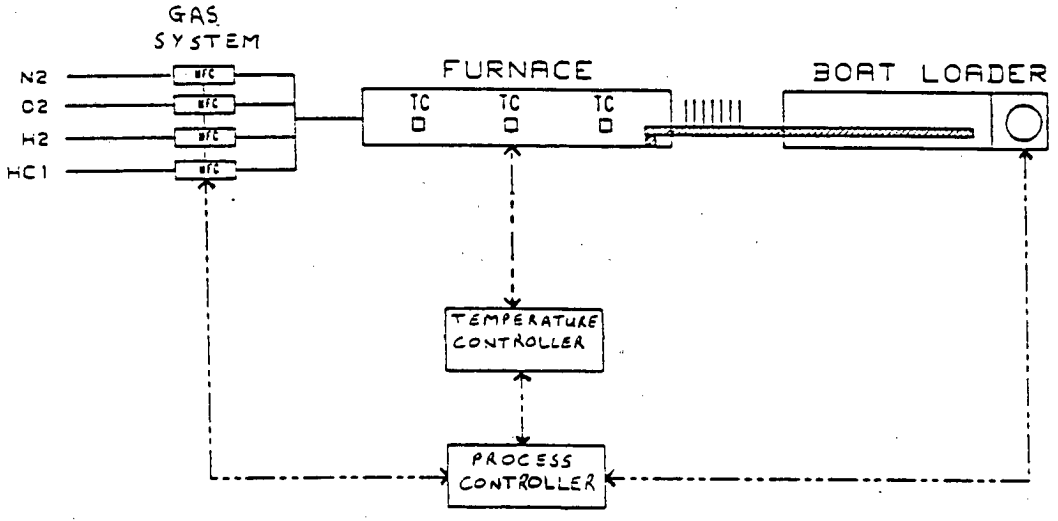


Fig. 4.1 Main features of an advanced digitally controlled furnace system, showing the gases normally supplied in oxidation applications and the position of 'spike' thermocouples (TCs) used to control temperature. In some systems the temperature controller and process controller are combined in a single unit.

4.1.1 The Furnace Tube

The furnace tube can be thought of as the reaction vessel in which oxidation takes place. It has an important role in preventing unwanted contamination from reaching the surface of silicon wafers during oxidation. The sensitivity of oxide quality to contamination will be discussed in Chapter 5. The contaminants of major concern are transition metals and alkali metals. Schmidt [113,114] has shown that the choice of material for the furnace tube can strongly influence contamination levels. The material must be effective as a barrier to diffusion due to the proximity of contaminant sources - furnace heating elements, supports and liners - and must also be of sufficient purity itself.

Most furnace processing takes place in fused quartz tubes. Fused quartz has very low impurity levels [114], can tolerate a wide temperature range and is easily manufactured. However, its transparency to impurities increases through use as a result of devitrivication and at very high temperatures (>1100°C) the tube is prone to sagging. These problems can be tolerated through good housekeeping procedures (i.e. cleaning, low idle temperatures etc.). The overall simplicity and low cost of quartz tubes has resulted in their widespread use in industry.

Silicon Carbide (SiC) has recently been considered as an alternative furnace tube material. Although originally produced with very high intrinsic impurity levels, causing much scepticism about its use in silicon processing, it has been shown that the present-day material is comparable in purity with fused quartz [115]. It offers a better barrier to impurity diffusion than fused quartz, as can be seen from the contamination levels measured in the surface regions of silicon wafers

oxidised using different tube materials (Table 4.1). The higher thermal conductivity of SiC compared to that of fused quartz also allows for more uniform temperature distributions. Although SiC has a higher initial cost than fused quartz it is offset by a longer service life, particularly at high temperatures.

Polysilicon has also been considered for the manufacture of tubes, again because of its high temperature durability. It has been found to have similar barrier properties to SiC (Table 4.1). Double-wall quartz tubes are another alternative [113,114]. A few per cent HCl gas in O_2 is passed between the inner and outer tubes to complex and/or sweep away incoming impurities. Devitrivication of the inner tube is reduced by removal of sodium (which is known to accelerate devitrivication) so tube life is increased.

4.1.2 The Boat Loader

Wafers are automatically moved in and out of the furnace tube by the boat loader. This is a motorised system that drives a 'paddle' on which the wafers are placed, usually in a quartz holder or 'boat'. The motor is interfaced to the process controller which supplies the movement direction and speed. Boat position may then be monitored by the process controller from a sensor on the loader (e.g. by using moving indicators on the drive shaft). Although control is not usually a true closed-loop, the loader allows accurate positioning of wafers in the furnace. Slow loading speeds (down to 10mm/min) are also possible, which reduces thermal shock effects, and heavy loads can be managed - particularly useful in the processing of large diameter wafers or for high throughput.

Element	SiC ($10^{15}/cm^3$)	Polysilicon ($10^{15}/cm^3$)	Fused Silica ($10^{15}/cm^3$)
Au	<0.6	<0.6	<0.6
Co	4.8	7.4	28
Cr	1.0	2.5	25
Cu	2.0	3.3	7.0
Fe	11	14	200
Mn	0.9	0.85	7.8
Ni	30	30	150

TABLE 4.1 Concentrations of impurities in the top 10 um of silicon wafers oxidised for 3 hrs at 1100°C in dry oxygen in SiC, polysilicon and fused silica furnace tubes, obtained by Atomic Absorption Spectroscopy. (from [115])

An important advancement due to the introduction of boat loaders has been the particulate reduction possible when non-contact paddles are used. Earlier contact systems, such as manual-loading or wheeled paddles that roll along the furnace floor, create friction and result in particulate contamination that may degrade yield. There are two main types of non-contact loaders : cantilever-suspended and soft-landing [116,117]. Both suspend the wafer boat during loading and unloading so that there is no contact with the furnace tube. Once the boat is in position the soft-landing system lowers it to rest on the tube floor and the paddle may withdraw ; the cantilever system continues to suspend the boat during the entire furnace process. Both methods require careful choice of construction material since large bending stresses are set up in the paddle. Quartz, alumina and SiC have all been considered, sometimes in combination [118]. Quartz is frequently used as a high purity sheath for less pure, but stiffer material. SiC in its recent high purity form (see Section 4.1.1) is particularly attractive since it has very high stiffness. This reduces deflection of the paddle tip, an important factor when there may be little clearance inside the furnace tube.

An end-cap is attached to the paddle to automatically seal the open end of the furnace tube when the loading operation is complete. This prevents the back diffusion of air and impurities into the tube. Process gases in the furnace escape through an exhaust port in the end-cap. Usually the exhaust feeds into a vented chamber so that the gases are safely removed from the furnace area.

4.1.3 The Furnace Gas System

The most common gases supplied to an oxidation furnace are O_2, H_2, HCl and N_2, which allow the furnace to be used for wet oxidation, dry oxidation or chlorinated oxidation. N_2 is used as an inert gas for annealing or as a carrier. The water vapour required for wet oxidation is produced by the pyrogenic technique, where H_2 and O_2 are directly combined inside the furnace tube to create H_2O. The two gases are mixed in a torch assembly (Fig. 4.2) placed at a point in the furnace where the temperature is high enough to ensure spontaneous ignition, producing a pale blue hydrogen flame. This technique is much cleaner than the original method of bubbling oxygen through water at 95-100°C.

The introduction of gases into the furnace is via a gas system (Fig. 4.3) controlled from the process controller. This is usually placed in a vented cabinet for safety in the event of leaks. Stainless steel piping and teflon tubing are generally used because of the corrosive nature of HCl. The key element is the mass flow controller (MFC), which confers localised closed-loop control on its input set point flow (which it receives from the process controller). One MFC is required for each gas. They are usually used in conjunction with solenoid or air-driven valves which ensure a gas line is sealed shut when no gas is required. Solenoid valves respond to a digital signal also provided by the process controller (0=closed, 1=open). Actual gas flows, as measured by the MFC, are returned to the processor controller for monitoring. The mass flow controller allows gas concentrations in the furnace to be accurately determined. This is of great importance since oxidant concentration can influence the oxide growth rate (Chapter 3). The precision of MFCs also

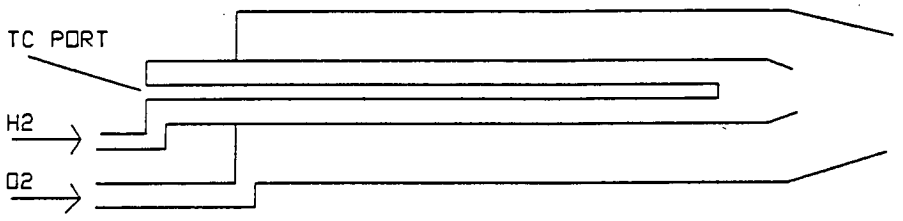


Fig. 4.2 Typical assembly used in the pyrogenic oxidation technique. A thermocouple is placed near the tip to warn of low temperatures and so ensure safe burning of hydrogen.

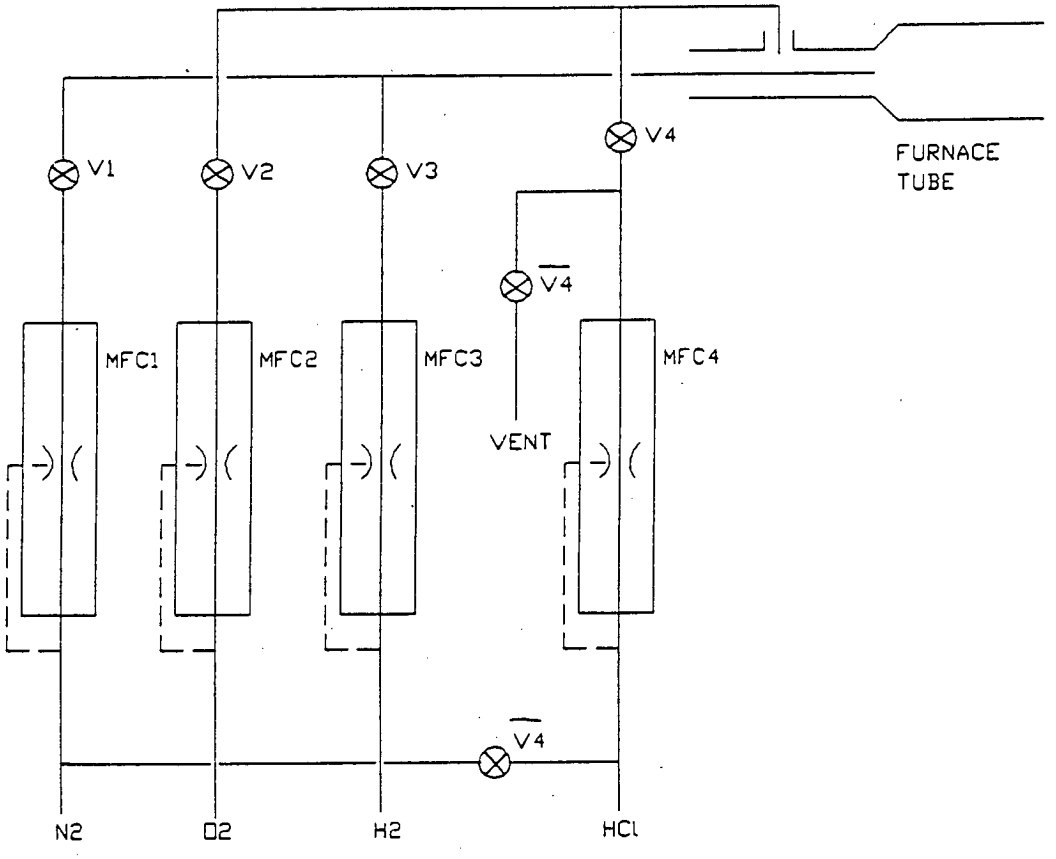


Fig. 4.3 Typical layout of a furnace gas system, showing valves (V1 to V4) and mass flow controllers (MFC1 to MFC4).

permits safe handling of dangerous gases such as HCl and H_2.

4.1.4 The Temperature Control System

The thermal oxidation of silicon is normally carried out at temperatures between 900°C and 1150°C. Heating to these temperatures is carried out by helically wound resistive heating elements, usually in a ceramic support, surrounding the furnace tube. These are controlled by the microprocessor-based temperature controller, which in some furnace systems is integrated into the process controller. The objective is to create a 'flat-zone' - a section of the furnace tube where there is uniform temperature and in which the wafers will be placed for oxidation. Flat-zones of 70cm to 100cm in length with an accuracy of better than +/- 0.5°C are frequently quoted by furnace manufacturers.

Control is complicated by higher heat losses at the end of a furnace tube than at the centre. The flat-zone is therefore split into 3 control zones which may be controlled either independently or in a master/slave configuration. Temperature is measured by 'spike' thermocouples (TCs) positioned close to the outside of the furnace tube, one for each zone. Platinum/platinum-rhodium TCs are frequently used due to their stability and wide operating range (0°C to 1400°C). Negative feedback control maintains the temperature in each zone at set point (Fig. 4.4). The software control algorithm incorporates separate proportional band, integral band, derivate band and gain coefficients to optimise temperature stability and prevent overshoot when an increase in temperature is required.

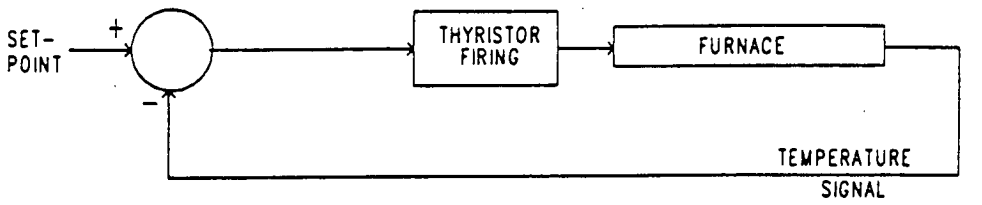


Fig. 4.4 Negative feedback loop in temperature control.

Controlled changes in temperature that maintain a flat-zone throughout the transition period are possible by ramping the closed-loop set point linearly with time. This controlled temperature ramping can be used in conjunction with low loading/unloading temperatures (i.e. lower than oxidation temperature) to reduce stress in the wafers caused by thermal expansion effects. It also means that controlled shifts to high annealing temperatures are possible.

Since the temperature at the spikes need not be the temperature inside the tube, the furnace must be profiled. Profiling involves placing a set of three standardised quartz-sheathed TCs inside the tube (one for each control zone) to determine the spike set points required for a specified tube temperature. Profiling data for several different tube temperatures is stored in the temperature controller as a profile-table that can be interpolated to determine spike set points at any temperature. The profile-table can normally be filled directly from the profile thermocouple data - a feature known as auto-profiling.

4.1.5 The Process Controller

The furnace process controller is the co-ordinator that communicates with other parts of the furnace and allows automated furnace operation. Controllers are now usually recipe-based. The recipe is a program that describes the complete oxidation process and it is formed as a set of steps. Each step contains definitions of time, gas flows, paddle position, paddle speed, and temperature. Upon initiation by an operator, these steps are executed sequentially by the controller and

appropriate signals are sent to the other furnace modules.

Several recipes can normally be stored in the controller, allowing a variety of oxidation processes to be easily accessed. Further versatility is usually available using a recipe editor to change the settings inside a recipe. There is no universal recipe language and the degree to which a recipe can be altered varies from manufacturer to manufacturer. For example, the Eurotherm FICS10 controller has a 'baseline' program stored in ROM that consists of the main program steps. Individual recipes can only be created within this framework. By contrast, in the Tempress DPC controller every step in every recipe is initially undefined and all commands and settings must be programmed as required.

As well as sending out control information, the process controller constantly receives measured parameters from various parts of the furnace. These allow the furnace operations to be monitored on a display. More significantly, alarm conditions can be set on these inputs, through programming in the recipe, to alert the operator of any problems. Recipe flow can be automatically changed, based on the input signals or the alarms generated from them. Thus by defining a limit alarm for O_2 flow (i.e. a maximum deviation between set point and actual flow), for example, an oxidation process can be automatically aborted if, say, the oxygen supply to the furnace failed. Safety in furnace operations is also enhanced through alarm generation by the process controller.

The process controller can normally communicate 'upstream' with a furnace supervisory system or remote computer. This feature, together with the continual monitoring of furnace parameters by the process

controller, is of particular interest to this project. It forms a key part of the new control technique described in Chapter 7.

4.1.6 EMF Research Furnace

In this section, the furnace used for all experimental work in this project, and for implementation of the control system introduced in Chapter 7, is described. The furnace used is of standard industrial design and follows the general outline given in Fig. 4.1. It is one of a four-furnace stack built by Tempress and is installed in the clean rooms of the Edinburgh Microfabrication Facility. It is controlled through the Tempress Digital Temperature Controller (DTC) [119] and Digital Process Controller (DPC) [120].

Tylan mass flow controllers are used in the gas system. Their preset ranges allow N_2 and O_2 flows from 0 to 10 standard litres per minute (SLM), H_2 flows from 0 to 5 SLM and HCl flows from 0 to 200 standard cubic centimetres per minute (SCCM). The furnace tube is manufactured from silicon carbide and has an internal diameter of 125mm (suitable for 3" diameter wafers). Three-zone temperature control gives a flat-zone of 100cm. Both spike and profiling thermocouples are of type 'R' (i.e. platinum/platinum-rhodium 13%).

The paddle is also manufactured from silicon carbide and is of a wheelbarrow-type design (i.e. a contact method). It is fitted with a quartz end-cap and has a hollow construction to allow the insertion of profiling thermocouples. The DTC has been designed as a stand-alone module and is recipe-based, a recipe consisting of temperatures and ramp rates for each zone.

Temperature changes are programmable at the DPC through selection of the appropriate temperature recipe.

4.2 Limitations of Furnace Control

In order to have a controlled oxidation process, stable, well-regulated conditions must be provided inside the furnace tube. Whilst advanced microprocessor-based systems such as those just described are certainly an improvement compared to their manual/analogue predecessors, full control over the furnace conditions is still not achieved. This is demonstrated by the following examination of furnace operating characteristics, carried out on the Tempress furnace.

4.2.1 Temperature Profiles

The flatness of the furnace flat-zone was investigated. Standard procedures for auto-profiling were carried out to build up the profile-table in the DTC. The profiled temperatures were 700, 800, 850, 900, 925, 950, 975, 1000, 1025 and 1050°C. Three profiling thermocouples were used, set 50cm apart to correspond to the position of the spike thermocouples. Auto-profiling was carried out with a nitrogen flow of 3.0 SLM.

On completion of auto-profiling, the set point for each control zone was set to 950°C. The temperature inside the tube was then measured every 5cm, using a thermocouple, allowing a stabilising time of three minutes each time the thermocouple was repositioned. The resultant temperature profile is shown in Fig. 4.5 (solid line). A rapid fall-off in temperature outside the region of the flat-zone is clearly seen. The measured temperature is within 0.5°C of 950°C around

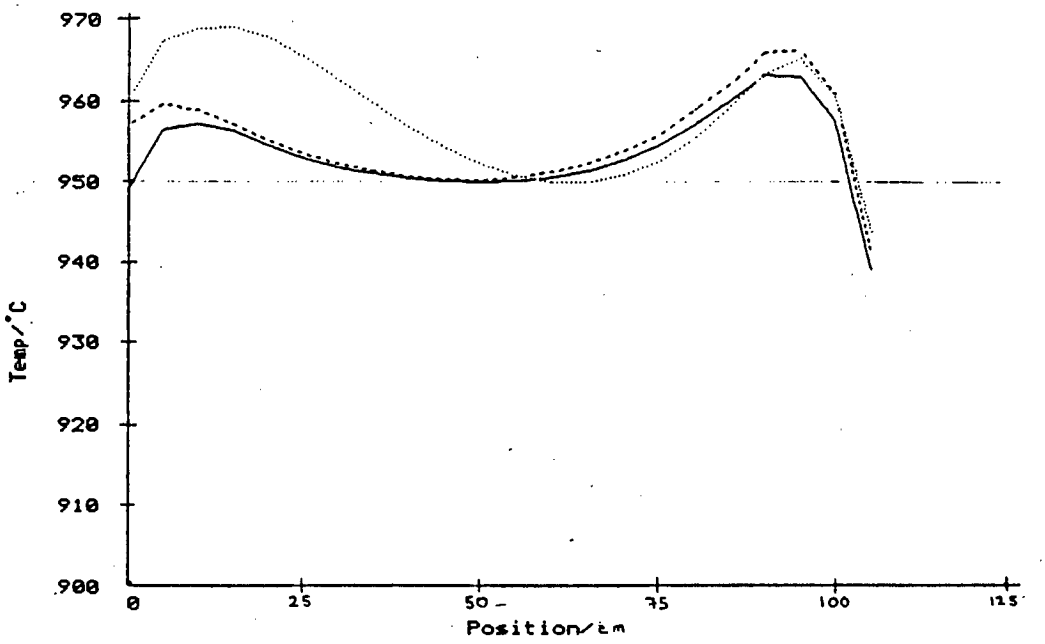


Fig. 4.5 Temperature profile within the furnace tube. The set point temperature is 950°C. Gas flows are : i) N_2=3.0 SLM (solid line); ii) H_2=3.0 SLM, O_2=2.5 SLM (dashed line); iii) H_2=6.0 SLM, O_2=5.0 SLM (dotted line). Position 50cm is the centre of the tube.

the sites of the profiling thermocouples (0cm, 50cm, and 100cm). However, much larger deviations (up to 13.2°C) are seen at other sites. Similar results were also obtained at set point temperatures of 975°C and 1025°C.

The measurement of radial temperature profiles (i.e. within a 'slice' of the furnace tube) is much more difficult. The data in Fig. 4.5 were obtained from a position near the bottom of the tube. An attempt was made to measure temperature at other positions using a thermocouple bent upwards near the junction to raise it to a higher level. Unfortunately, this required the quartz sheath to be bent as well and although possible, it was highly stressed. The sheath cracked when subjected to the high temperature inside the furnace and data were therefore not obtained. However, it can be surmised that there will be radial variation in temperature in addition to the axial variation described above.

Flat temperature profiles will be even more difficult to achieve as furnace diameters increase to accommodate larger wafers, due to greater heat losses at the ends. Two recent advances may help counteract this.

i) BTU/Bruce have introduced a five-zone heating system [121] which enhances the effectiveness of end-zone temperature adjustment.

ii) Tylan have introduced special insulation sections at each end of the furnace tube to reduce heat loss and so enhance temperature control [111].

4.2.2 The Hydrogen Flame

Hydrogen burning within the furnace tube during pyrogenic oxidation generates considerable amounts of heat. On the basis that one mole of hydrogen releases 284 kJ of energy on combustion, it can be calculated that a hydrogen flow of 1 SLM is equivalent to 211 watts. This heat is released directly into the furnace and therefore affects the temperature profile.

The temperature distribution inside the flat-zone was measured with a thermocouple after the hydrogen flame had been switched on and ample time allowed for the system to stabilise. Two different hydrogen flow rates were tried. For both an excess of oxygen was supplied to ensure complete combustion. The flows used were 3.0 SLM H_2 with 2.5 SLM O_2 and 6.0 SLM H_2 with 5.0 SLM O_2. The resultant profiles are also shown in Fig. 4.5, and should be compared with the profile obtained with the flame off (i.e. using a 3.0 SLM flow of N_2).

It is clear that the hydrogen flame has produced a localised heating effect that has not been sufficiently compensated for by the DTC. Not surprisingly, the effect is greater for the larger hydrogen flow rate, both in terms of temperature increase and extent into the furnace tube. It can be speculated that the radial temperature variations will also be greater as a result of the flame. These results highlight the problems of accurate temperature control. In particular the increase at position 0, a profiling thermocouple site, shows that the heating effect has not been detected in full by the spike thermocouple at that position. This indicates a limitation of the control that can be achieved based on temperature data from outside the furnace tube.

Care was taken to ensure that the above profiles were obtained under stable conditions, that is the furnace was effectively in a steady-state. However a transient period was also measured, in which furnace temperatures were changing. This is indicated in Fig. 4.6 which shows the temperature measured inside the furnace at position 0 before and after the hydrogen flame was switched on. Over 18 minutes are required for the temperature to stabilise after the flame is switched on. During this period, temperature is drifting to a new steady-state value and is not being actively controlled by the DTC.

4.2.3 Temperature Reduction on Loading

The effect of loading operations on furnace temperature was examined. After placing thirty 3" silicon wafers in a boat on the paddle, they were rapidly inserted into the furnace tube. The loading temperature was 950°C and a nitrogen flow of 3.0 SLM was used. The temperature measured by the central spike thermocouple (position 50cm) was recorded. The data had been adjusted by the DTC, which used the profile-table to shift the spike readings to the expected temperature inside the tube. (This is of no consequence to the examinations.) The results are plotted in Fig. 4.7.

Paddle, wafers and wafer boat constitute a considerable thermal mass that has caused a major temperature disturbance on loading into the furnace flat-zone. The DTC has responded, but has required a period of around 20 minutes to recover to its set point temperature. Its response is limited by the power that can be generated through the heating coils and by the control algorithm itself, which limits power output as

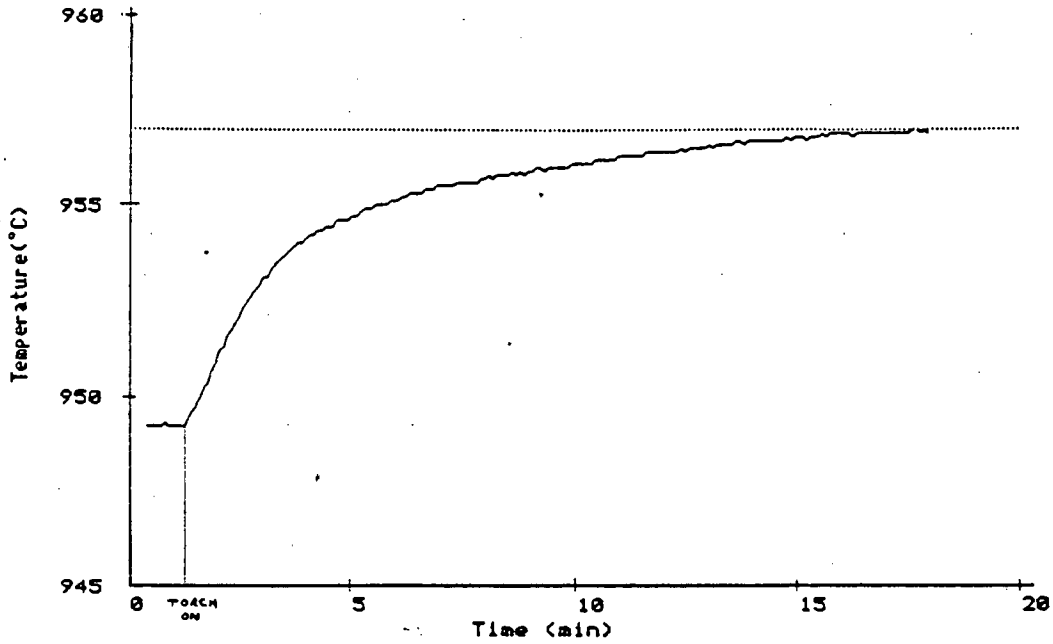


Fig. 4.6 Temperature against time measured at position 0 of Fig. 4.5. Gas flows were changed from N_2=3.0 SLM to H_2=3.0 SLM, O_2=2.5 SLM after 1 min.

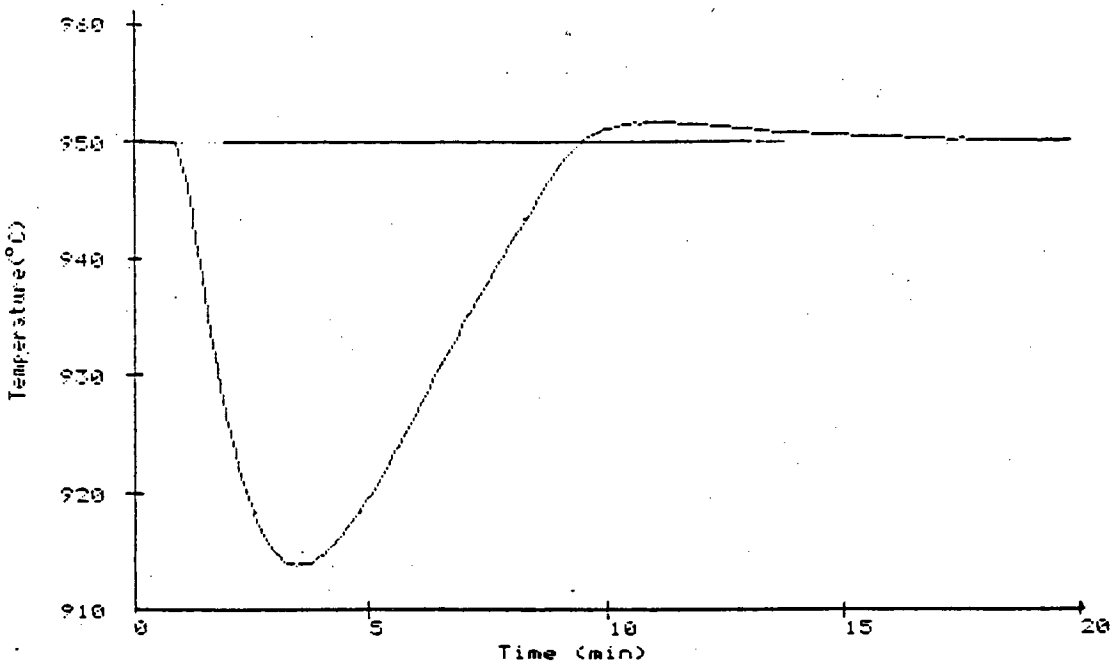


Fig. 4.7 Temperature measured by the central spike thermocouple upon insertion of wafers and paddle into the furnace. Set point temperature is 950°C.

the temperature approaches set point to minimise overshoot.

Further temperature measurements were taken from a thermocouple placed on the paddle as close as possible to the wafer boat (hereafter referred to as a paddle thermocouple). The paddle thermocouple recorded temperatures inside the tube during loading and the results are shown in Fig. 4.8. Closer examination (Fig. 4.9) confirms that around 20 minutes are required for temperature to stabilise inside the tube.

Wafer loading introduces another transient period where the furnace temperature is not being well-regulated by the DTC and non-steady-state conditions exist inside the tube. A relatively small load of thirty 3" wafers plus paddle has been used. The larger wafers and greater batch sizes which are now being introduced (e.g. two hundred 6" wafers) represent a much higher thermal load and it is likely that the effect will be worse under these conditions.

4.2.4 Mass Flow Controller Characteristics

When a gas is requested by the DPC, the gas system does not immediately provide that gas at the specified flow. This is demonstrated in Fig. 4.10, which shows the MFC output as a function of time. Mass flow controller output is a 0 to 5 volt dc signal that is directly proportional to mass flow (over the calibrated range). Fig. 4.10 was obtained by connecting a chart-recorder across this output on the H_2 MFC then programming an H_2 flow of 3.0 SLM through the DPC.

The MFC operates by directing a small sample of gas through a sensor tube fitted with two heated resistance

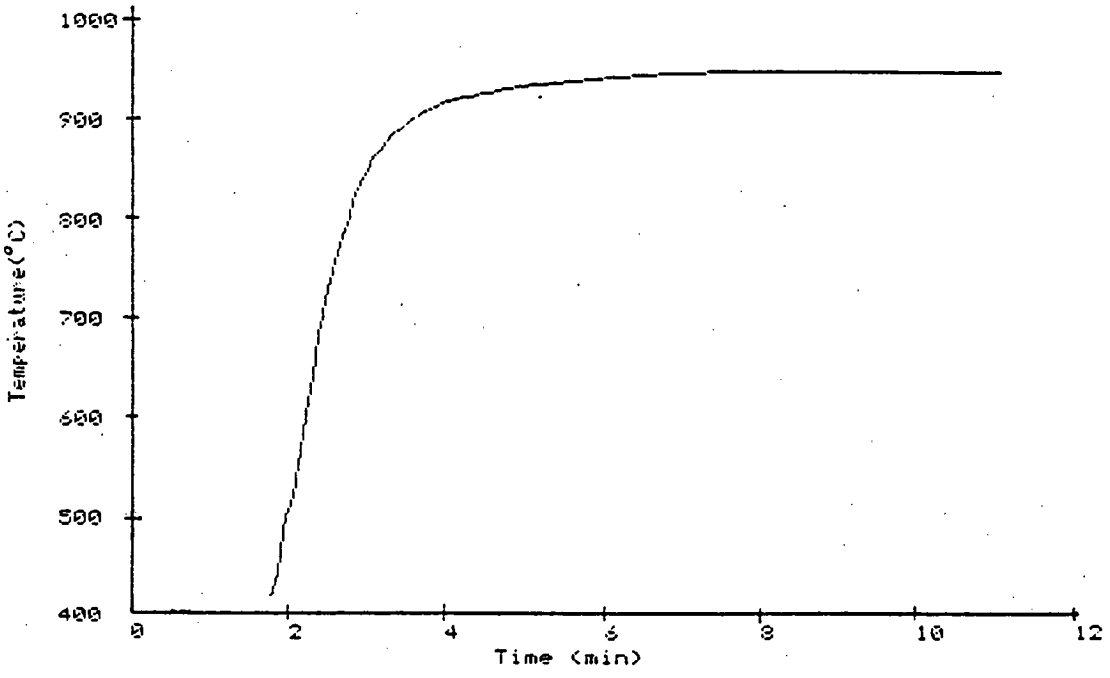


Fig. 4.8 Temperature measured by the paddle thermocouple upon insertion of wafers and paddle into the furnace.

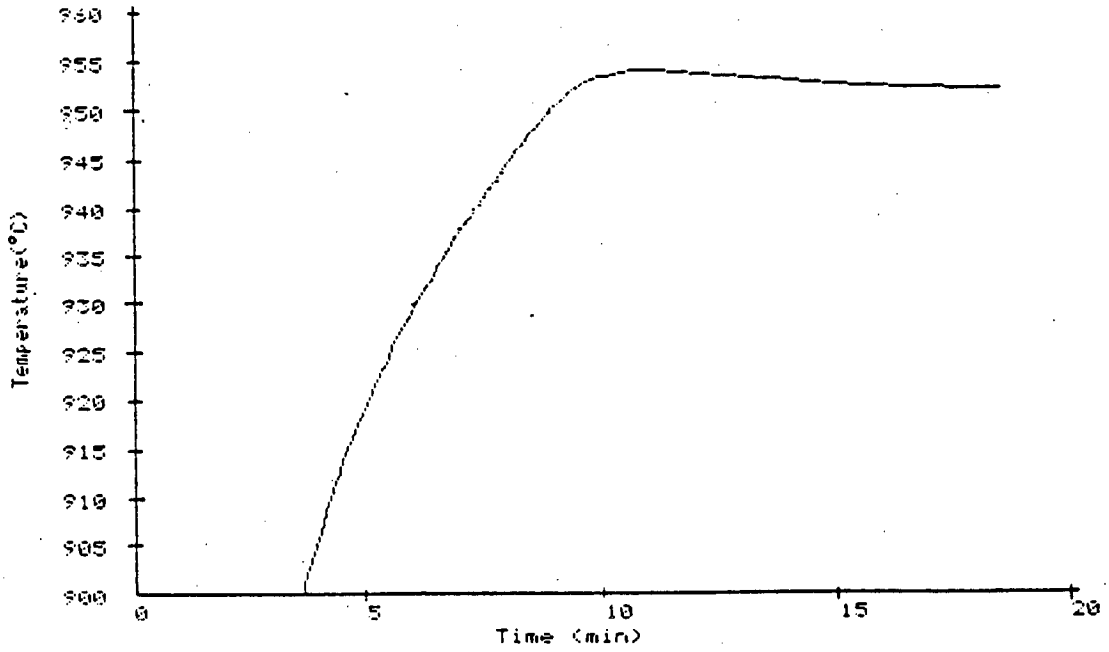


Fig. 4.9 Detailed view of the data in Fig. 4.8 as the measured temperature approaches the 950°C set point.

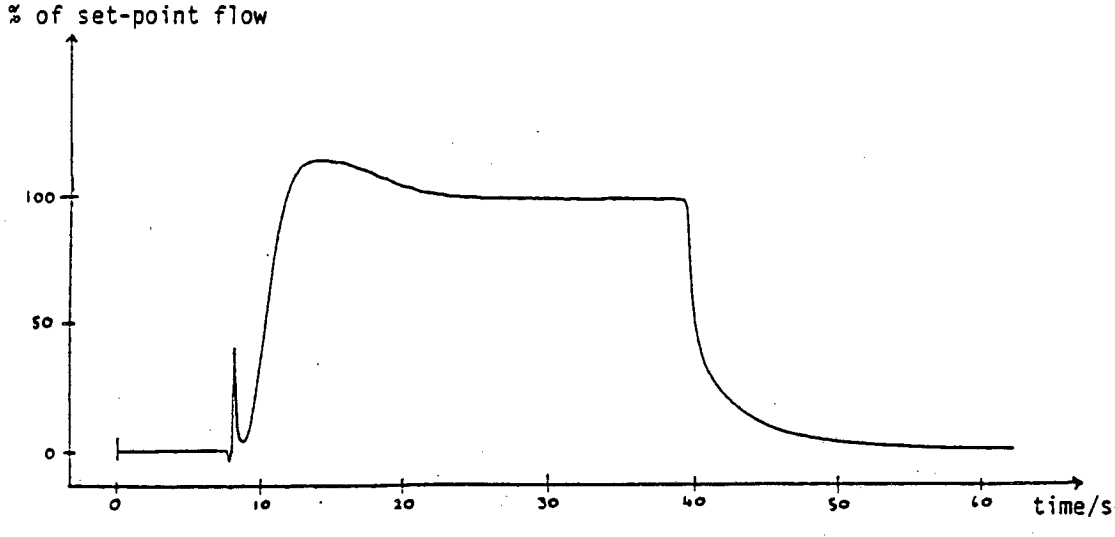


Fig. 4.10 Mass flow controller output for a programmed set point flow of 3.0 SLM.

thermometers that measure heat transfer in the gas and develop the flow-dependent output voltage. Closed-loop control is achieved by comparison of this signal with the set point provided by the DPC, the deviation determining the position of a control valve inside the MFC. The control valve can be one of several types, but the most responsive MFC has an 'expanding-rod' control valve that operates by the thermal expansion of the valve when power is applied to a heating coil. This is the type used in the EMF furnace. Because this type of valve is normally open, a burst of gas would occur at the beginning of a cycle. To avoid this, a 'soft-start' feature is usually incorporated into the process controller. Soft-start involves a delay period before the solenoid valve is opened to switch a gas on, during which time the MFC control valve is driven closed. This accounts for the initial 9 second delay in Fig. 4.10. After the solenoid valve is opened, the gas flow rises steadily from zero to the set point as the MFC closed-loop control opens the control valve. Overshoot occurs and a stable flow is not achieved until a total of 26 seconds after the DPC recipe has first requested the gas. Similarly, when the gas is switched off by the DPC, there is a 20 second delay before the gas flow rate is zero. All the MFCs in the gas system will have these on/off characteristics. Hence every time gases are changed there will be a period where the control of flow is poor.

A related effect is the purge time of the furnace tube. When a new gas is introduced into the furnace (e.g. a change over from nitrogen to oxygen) there will be a transition period during which the old gas is gradually replaced by the new gas. Steady-state conditions will only be achieved when enough new gas has entered the furnace to completely purge the old gas. The time required will depend on the gas flow dynamics

inside the tube. Since these are highly complex, a calculation is not attempted here. However an approximation can be obtained based simply on the tube volume and the flow rate of the gas entering the tube, accounting for thermal expansion. For example, a time of about 80 seconds is calculated to purge the EMF furnace tube (internal diameter 125mm, length 2.25m) at 950°C with a 5.0 SLM flow. During this period the gas composition in the tube will be constantly changing and the time will be extended by any gas mixing during the purging process.

4.2.5 Discussion

The main requirement of an oxidation process is the production of a thin layer of SiO_2 that is within thickness tolerances and is of acceptable quality. Concentrating on the former, oxide thickness is specified by the oxidation kinetics and time. Of the many factors that influence oxidation kinetics (Chapter 3), several are determined directly by the oxidation furnace: oxidant type, temperature, oxidant concentration (partial pressure) and HCl concentration. It is therefore important that these are accurately controlled to minimise oxide thickness variations, both within a single batch and from batch to batch.

However, the preceding sections show that even an elaborate digital control system does not realise this completely. The flat-zone is not flat. Transient periods exist where temperatures and gas flows are different from those defined in the process recipe. Dynamic gas concentrations occur whenever the furnace ambient is altered. All these effects will contribute to oxide thickness variation. As growth times are

reduced to produce the thinner oxides required for VLSI circuits, the transient effects will be increasingly more important. Improvement to the furnace control system is therefore necessary to maintain thickness tolerances.

4.3 Pre-oxidation Treatment

Wafers pick up both organic and inorganic contamination during handling and storage. This must be removed before furnace treatments, since contamination leads to poor device performance and reliability, as previously mentioned. Pre-oxidation cleaning of wafers is carried out immediately before they are loaded into the furnace to avoid subsequent recontamination. It should therefore be regarded as part of the oxidation process.

Wafers are cleaned chemically, the most widely used solutions consisting of various ratios of H_2SO_4-H_2O_2, HCl-H_2O_2-H_2O and NH_4OH-H_2O_2-H_2O. These are often used in conjunction with dilute HF acid which removes any oxide that may have formed on the wafer surface. De-ionised (DI) water is used for rinsing, and drying is carried out in hot nitrogen. The conventional procedure of dipping wafers in chemical immersion baths for cleaning has given way to chemical spray systems that continually flush away contaminants and allow all cleaning steps (including rinsing and drying) to be performed in the same chamber. Many pre-oxidation cleaning procedures are possible. The one described here is typical and is based on the well-known 'RCA-clean' [122]. It involves three cleaning steps separated with DI water rinses:

1. NH_4OH-H_2O_2-H_2O mixed in a ratio of 1-1-5 at 80°C. This solution removes organic contamination through the solvating action of ammonium hydroxide and the

oxidising action of the peroxide. It also complexes some metal ions from groups I and II of the periodic table.

2. $HF-H_2O$ mixed in a ratio of 1-9.
Native oxide and oxide formed as a result of Step 1 is removed from the silicon surface.

3. $HCl-H_2O_2-H_2O$ mixed in a ration of 1-1-6 at 80°C.
This solution removes heavy metal contaminants from the bare silicon surface, forming stable chloride complexes that prevents redeposition onto the silicon.

A final oxide removal by dilute HF is sometimes used, although this can lead to recontamination of the silicon [123].

If the final HF treatment is not carried out, a native oxide film is present on the silicon prior to oxidation. Since this is around 10-20Å thick, it cannot be measured accurately but its presence is observed by water retention on the wafer surface due to the hydrophilic nature of SiO_2. (By comparison, silicon is hydrophobic and bare wafers are observed to dewet.) The presence, or not, of native oxide could lead to an influence of pre-oxidation cleaning procedures on the control of oxide thickness. Little work has been carried out to investigate this effect, although it has been shown that the introduction of a final HF clean immediately before oxidation enhances the batch-to-batch reproducibility of 200Å oxides [124]. It has also been reported that the cleaning solutions $H_2SO_4-H_2O_2$ and $NH_4OH-H_2O_2-H_2O$ can effect oxidation rate [125]. The former enhances the rate and the latter is observed to decrease the rate, compared with oxidation of wafers that received no clean. The reason for this is unclear, but

it does indicate that pre-oxidation cleaning must be carefully controlled.

The pre-oxidation treatment of wafers does not necessarily concern just wafer cleaning. Some furnace oxidation processes may also involve heating up the wafers in a non-oxidising ambient such as nitrogen. Nitrogen is generally considered inert but it has been observed to react with silicon [126], although trace amounts of oxidant can effectively prevent this. A decrease in oxidation rate for thin oxides has also been reported when nitrogen is replaced by argon in the pre-oxidation warm-up of wafers [127]. Although only effective at high temperatures (>1100°C), these effects emphasise that control of oxide thickness depends on control of the pre-oxidation treatments as well as the oxidation itself.

CHAPTER 5

ASSESSMENT OF SILICON DIOXIDE

Various methods are available for evaluating the thickness and quality of oxide films. Each technique has its own merits and limitations. It is important that these are understood, since oxide assessment is a necessary part of the experimental work in Chapters 6 and 7 and the principal methods are now discussed.

5.1 Thickness Measurement

There are a number of methods for measuring the thickness of silicon dioxide on a silicon substrate. The methods reviewed here are optical interference, ellipsometry, surface profiling and capacitance. Film thicknesses are routinely monitored in IC manufacture and consequently there has been development in equipment that can perform measurements automatically. However, not all techniques are equally convenient, nor are they all non-destructive, and this must be taken into consideration when selecting a particular method of measurement.

5.1.1 Optical Interference Method

This technique is based on the interference between light reflected from the air-SiO_2 interface and light reflected from the SiO_2-Si interface. Light is directed at the SiO_2 film and the intensity of reflected light is measured, whilst either the angle of incidence is varied at constant wavelength [128], or wavelength is varied at constant angle of incidence [129]. Both produce maxima

and minima in intensity as a result of the difference in optical path lengths for the two reflecting interfaces. The film thickness can be calculated from the position of these maxima and minima using the equation

$$d = \frac{(M - N)\lambda}{2n^*} \qquad (5.1)$$

where d = film thickness
$n^* = (n^2 - \sin^2\theta)^{1/2}$
n = refractive index of SiO_2
θ = angle of incidence
M = order of interference
N = net phase shift = $\phi_f - \phi_s$
ϕ_f = phase shift at the air-SiO_2 interface
ϕ_s = phase shift at the SiO_2-Si interface
λ = wavelength

Calculation is simplified by assuming that the net phase shift and the refractive index are independent of wavelength, although this can lead to error [129].

The Nanometrics Nanospec is a commercially available microprocessor-based automatic film thickness measurement system that utilises the latter of the above methods i.e. variable wavelength at constant angle of incidence. A block diagram of the main components is shown in Fig. 5.1. It has a number of programs to measure various films and combinations of films, including SiO_2, Si_3N_4, polysilicon and photoresist. The program for SiO_2 scans between 480nm and 800nm and can measure oxide thickness from 400Å to 30000Å with a specified 2% accuracy. The Nanospec optical system has a choice of objective lens that allows the area of the

analysis region to be varied. The 10x objective gives a spot diameter of 35um and is the simplest to use. The higher power objectives that are available are useful for analysing patterned wafers.

An alternative program is available for oxides thinner than 500Å. This compares the reflectance of the oxide film with the reflectance of a bare silicon wafer at a single wavelength of 520nm. Reliable thickness measurements down to 50Å are specified for the program. However, it is a sensitive method that requires careful focusing of the Nanospec optics on the sample. The silicon reference must also be continually updated (every 10 minutes) for accurate results when using this program.

Interferometry is a non-destructive technique that requires no special treatment of the oxide samples. Through instrumentation such as the Nanospec, thickness measurements can be obtained quickly and simply, and the technique is widely used in the semiconductor industry. Incorporation of a programmed X-Y stepping table allows such equipment to automatically measure several sites per wafer and further simplify the collection of thickness data [130,131].

5.1.2 Ellipsometry

Ellipsometry [132] is a more complicated technique than interferometry and is based on the change in polarisation vector of a beam of monochromatic, polarised light when it is reflected from a surface. The components of the Applied Materials Ellipsometer II are shown diagrammatically in Fig. 5.2. A helium-neon laser emits monochromatic light which is plane polarised then elliptically polarised. This is directed onto the sample at a fixed angle of 70 degrees. The reflected

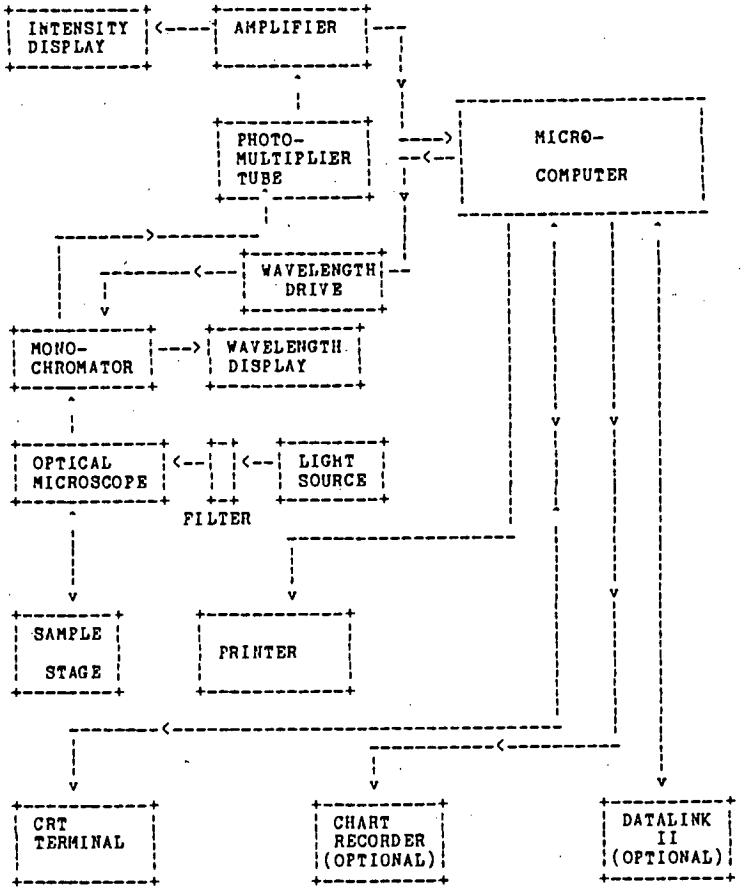


Fig. 5.1 Main components of the Nanometrics Nanospec.

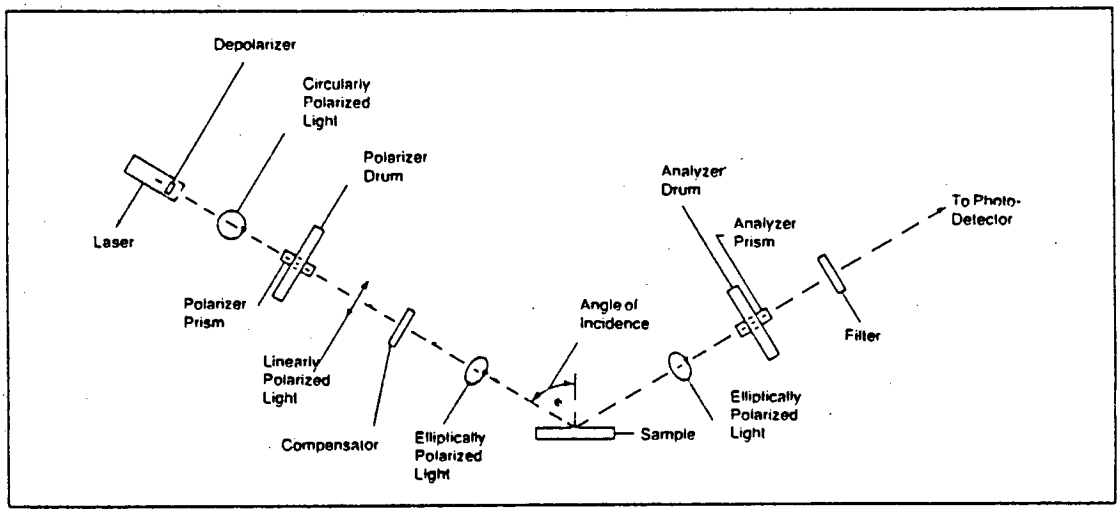


Fig. 5.2 Block diagram of the Ellipsometer.

beam passes through an analysing polariser and its intensity is measured. By varying the angle of polarisation of the incident beam and the angle of the analyser, two extinction points, where there is a minimum of light transmission, can be found. The positions for extinction depend on the SiO_2 thickness and its refractive index, both of which can be determined with the technique.

Two parameters, Ψ and Δ, are calculated from the polariser angles and the analyser angles at which extinction is obtained.

$$\Psi = \frac{(180 - A_2) + A_1}{2} \qquad (5.2)$$

$$\Delta = P_1 + P_2 \qquad (5.3)$$

where P_1 and A_1 are the polariser and analyser angles for the first extinction point, and P_2 and A_2 are the polariser and analyser angles for the second extinction point. The relationship between Ψ, Δ, refractive index and film thickness is complex, and it is normal to use a nomogram chart, tabulated data or a computer program to obtain values of refractive index and film thickness from the calculated values of Ψ and Δ.

The Ellipsometer can measure oxide thicknesses as low as 5Å. However, below about 400Å results become increasingly sensitive to the values of Ψ and Δ. In this region, lines on the nomogram converge rapidly and this limits accuracy. The use of a computer program is

therefore preferred for thin oxides, although this of course does not alter the basic problem of increased sensitivity.

There is no upper limit to the thickness that can be measured but the parameters Ψ and Δ are periodic functions of thickness and an approximate thickness (to within 2500Å) must be known. The measured thickness is then added to an appropriate thickness multiple to obtain the actual thickness. Oxides below about 2800Å have a thickness multiple of zero.

Ellipsometry is a non-destructive technique and is therefore suitable for the measurement of production wafers in the semiconductor industry. Only recently have automated Ellipsometers become commercially available and standard equipment such as the Ellipsometer II have a completely manual operation. This requires a degree of skill for efficient and accurate use. Extra calculation is also needed to obtain relevant information. This means oxide thickness data is obtained more slowly than with a Nanospec. However, the Ellipsometer provides a useful means of checking the oxide refractive index, a parameter that cannot be measured with the Nanospec.

5.1.3 Surface Profiling

Surface profiling requires a step to be etched into the oxide layer so that the silicon substrate is exposed. This could be carried out by defining a suitable pattern in photoresist on top of the oxide then etching in dilute hydrofluoric acid. The measurement technique involves tracking a stylus across the step. A transducer converts the stylus movements into an electrical signal which is displayed as a surface profile. From this it

is possible to estimate the step height and hence the oxide thickness.

The Sloane Dektak II is a commercially available instrument that utilises this technique. It is microprocessor-based and its operation is completely automated. Scan data is digitised and can be manipulated in a variety of ways to facilitate the measurement of structural features. The stylus tip has a radius of 12.5 microns and it can measure step heights in the range 200Å to 50000Å. Noise levels limit the resolution to around 10Å. A well-defined step is also required, otherwise the stylus will track up the wall, making it difficult to determine the exact step height.

Surface profiling requires some pre-processing of the oxide samples. It is a destructive technique and it is therefore not suitable for use on wafers that require further processing. Resolution is poorer than with the Nanospec and Ellipsometer, and these latter techniques are generally preferred to surface profiling, especially for measurement of thin films.

5.1.4 Capacitance Method

A MOS capacitor can be fabricated by the deposition of a thin metal or polysilicon electrode (gate) over the silicon dioxide film. An example using an evaporated aluminium gate is shown in Fig. 5.3. The gate is typically up to 10000Å thick and is defined by a photolithography-etch sequence. The MOS capacitor is frequently used to check oxide quality [133][1], but here it is shown how the oxide thickness can be obtained from capacitance measurements.

[1] See Section 5.2.

A simplified equivalent circuit for the MOS structure in Fig. 5.3 is two capacitors in series: a variable silicon capacitance (C_s) whose value depends on the applied bias, and the oxide capacitance (C_{ox}) which is independent of the applied bias. The total (measured) capacitance, C_m, is given by

$$C_m = \frac{C_{ox}\, C_s}{C_{ox} + C_s} \qquad (5.4)$$

Only a p-type substrate is now considered, but the treatment for n-type is similar with appropriate changes of physical constraints and sign changes.

The application of a dc bias to the gate (V_g) results in band bending close to the Si-SiO_2 interface (Fig. 5.4). For negative V_g the bands bend as in Fig. 5.4(a), known as the accumulation state. The Fermi level (E_f) and the valence band edge (E_v) are brought closer together, and majority carriers pile up at the Si-SiO_2 interface. This large concentration of majority carriers makes C_s very large and equation 5.4 becomes

$$C_m = C_{ox} \quad (C_s >> C_{ox}) \qquad (5.5)$$

As V_g is made more positive, the bands bend less, passing through the flatband condition (Fig. 5.4(b)) and into the depletion state (Fig. 5.4(c)). In depletion, majority carriers are repelled from the surface leaving a depletion layer, and C_s becomes significantly smaller than its value in accumulation. With further increase in V_g, the intrinsic level (E_i) crosses the Fermi level

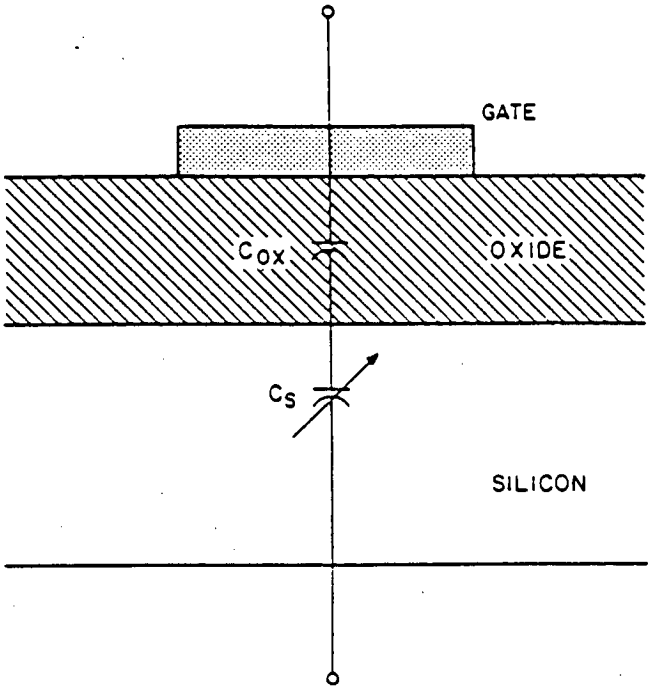


Fig. 5.3 MOS Capacitor structure and ideal equivalent circuit.

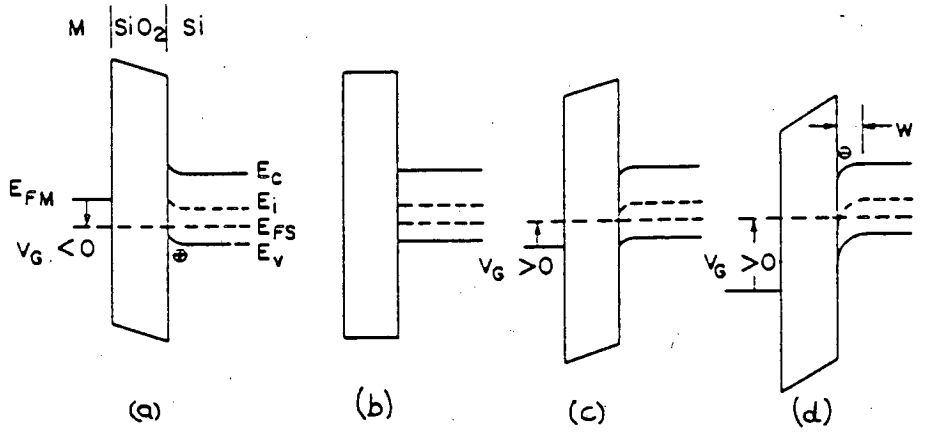


Fig. 5.4 Energy band diagram of the MOS capacitor structure : (a) accumulation mode (b) flatband condition (c) depletion mode (d) inversion mode. E_c and E_v are the energy levels of the conduction and valence bands, E_i is the intrinsic Fermi level, E_{FS} and E_{FM} are the Fermi levels in the silicon and the metal respectively.

and the surface layer becomes inverted (Fig. 5.4(d)). Although band bending is increasing, the depletion layer width does not change much and so C_s remains about constant.

Equation 5.5 shows that the oxide capacitance can be determined from the measured capacitance in accumulation. It is then possible to calculate the oxide thickness, x_o, from

$$x_o = \frac{A \cdot E_o \cdot E_r}{C_{ox}} \qquad (5.6)$$

where A = gate electrode area
E_o = permitivity of free space ($=8.854 \times 10^{-14} Fcm^{-1}$)
E_r= dielectric constant of SiO_2 ($=3.82$)[2]

The gate electrode area must be known accurately, although this is generally a simple task with a suitable microscope. The gate electrode dimensions must also be large compared to the oxide thickness to avoid significant fringing effects at the gate edge.

A capacitance bridge is the most sensitive method for measuring the high-frequency capacitance [39] and has been incorporated in instruments such as the Hewlett-Packard HP4275A LCR meter. The term high-frequency refers to the frequency of the small-signal ac voltage (<50mV) that is superimposed on the dc gate bias and is typically 100kHz to 1MHz. The bridge measures the differential capacitance defined as:

[2] E_r is generally insensitive to processing, oxide structure and capacitance measurement frequency. The value of 3.82 can therefore be applied under a wide range of conditions.

$$C = \frac{dQ_T}{dV_g} \qquad (5.7)$$

where Q_T is the total charge density in the capacitor. The capacitance-voltage (C-V) characteristic of a MOS structure shows the measured differential capacitance as a function of gate bias (Fig. 5.5). Band bending is directly responsible for the overall form of the C-V plot. For a p-type substrate, C_{ox} is the measured capacitance at negative values of V_g.

Computer-controlled semiconductor test systems such as the Hewlett-Packard HP4061 system, have greatly simplified C-V measurement [134]. Several test instruments are linked together on a standard IEEE-488 bus and are automatically co-ordinated under direction from a microcomputer. Appropriate software allows the automatic extraction of process parameters (including oxide thickness) from the measured data.

The capacitance method is not a rapid means of obtaining oxide thickness, nor is it generally as accurate as optical methods because of assumptions and difficulties associated with the measurement technique. It is a destructive method since material must be deposited over the oxide. The use of a mercury probe is a non-destructive alternative [135], where the oxide surface is contacted with a drop of mercury which then acts as a gate electrode, leaving the oxide underneath unaffected by the measurement. However, this is even less precise since the gate area is not known accurately.

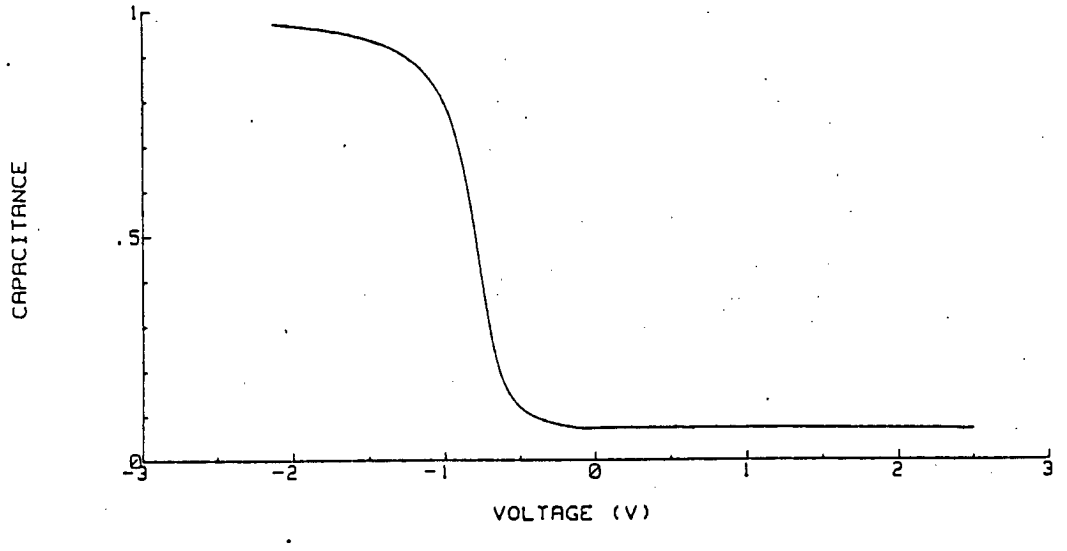


Fig. 5.5 High frequency C-V characteristic of a MOS structure (p-type substrate). Capacitance is normalised to C_{ox}, the value measured in accumulation.

5.1.5 Colour Charts

Different oxide thicknesses appear as different colours when observed under white light because of destructive interference. It is therefore possible to build up a thickness chart as shown in Table 5.1. Oxidised wafers can be compared with this chart to obtain the oxide thickness. Although this is not an accurate method (a resolution of about 200Å is possible), it provides a quick and easy estimation that can be combined with other techniques requiring approximate thickness to be known (e.g. ellipsometry). Thin oxides appear as the same colour as the silicon substrate and this method cannot be used below about 500Å. The colours also repeat apart from subtle variations, and some experience is required to use this technique effectively.

5.1.6 Comparison of the Nanospec and the Ellipsometer

Of the preceding thickness measurement techniques, optical interferometry and ellipsometry are the most convenient since no post-oxidation processing is required and since instrumentation is available to simplify the measurement operation. A comparison was made between two instruments utilising these techniques: the Nanometrics Nanospec and the Applied Materials Ellipsometer II. The repeatability of the instruments was studied at several oxide thicknesses. The same wafer sites were used with each instrument to allow a direct comparison of the measured thicknesses.

Eight dry oxidations at 950°C were carried out to give eight wafers with a range of oxide thicknesses from about 70Å to 1000Å. Two additional wafers with thicker oxides were obtained by wet oxidation at 950°C. Before oxide thickness measurements were made, the calibration

Film Thickness (Å)	Colour
500	Tan
700	Brown
1000	Dark violet to red-violet
1200	Royal Blue
1500	Light blue to metallic blue
1700	Metallic to very light yellow-green
2000	Light gold or yellow - slightly metallic
2200	Gold with slight yellow-orange
2500	Orange to melon
2700	Red-violet
3000	Blue to violet-blue
3100	Blue
3200	Blue to blue-green
3400	Light green
3500	Green to yellow-green
3600	Yellow-green
3700	Green-yellow
3900	Yellow
4100	Light orange
4200	Carnation pink
4400	Violet-red
4600	Red-violet
4700	Violet
4800	Blue-violet
4900	Blue
5000	Blue-green

TABLE 5.1 Colour chart for thermally grown SiO_2 films with thickness up to 5000Å. Colours apply to wafers observed perpendicularly under daylight fluorescent lighting.(from [128])

of each instrument was checked using procedures outlined in the respective operation manuals [136,137].

The oxide thickness at the centre of each wafer was first measured using the Ellipsometer. A program on the Edinburgh Microfabrication Facility VAX750 computer was used to calculate oxide thickness from the polariser and analyser settings. The measurement was repeated five times at exactly the same site. The average thickness and standard deviation obtained are shown in Table 5.2. Also shown is the calculated refractive index averaged over the five readings. For wafer 1, the VAX750 program calculated oxide thickness based on the assumption that the refractive index was 1.45 (all readings) and this may account for its particularly low standard deviation. Wafer 10 required the use of a non-zero thickness multiple and the extra calculation accounts for its large standard deviation in oxide thickness. The remaining wafers follow a pattern that can be expected from the nomogram chart: higher standard deviation in regions where lines converge rapidly (i.e. thickness of less than about 400Å and thickness between 2000Å and 2500Å), and improved repeatability in the less sensitive region between these.

Typical polariser and analyser settings obtained from the Ellipsometer are shown in Table 5.3 (the data shown are from wafer 5). The settings are obtained manually by reading vernier scales on the polariser and analyser. Changes of fractions of degrees result in the observed variation in measured oxide thickness. Readings must therefore be taken with great care to ensure accuracy when using this equipment.

Measurement of oxide thickness was repeated using the Nanospec. Again five measurements were taken at the centre site of each wafer. The results are shown in

Wafer No.	Oxide Thickness (Å)	s.d. (Å)	Refractive Index	s.d. (Å)
1	102.2	1.57	1.45	constant
2	201.7	8.12	1.42	0.037
3	298.8	8.10	1.46	0.028
4	378.2	4.76	1.45	0.016
5	452.6	4.78	1.46	0.009
6	585.6	5.45	1.47	0.009
7	794.6	6.44	1.47	0.005
8	1022.5	2.77	1.47	0.004
9	2495.5	58.50	1.43	0.019
10	8983.3	88.40	1.46	0.007

TABLE 5.2 Mean and standard deviation for five repeated measurements of oxide thickness and refractive index using the Ellipsometer.

Measurement Repetition	P_1	P_2	A_1	A_2	Oxide Thickness (Å)	Refractive Index
1	4.7	94.5	17.7	155.2	455.0	1.45
2	4.6	94.4	17.8	155.3	452.4	1.46
3	4.2	94.7	17.7	155.4	444.9	1.47
4	4.8	94.6	17.9	155.4	457.7	1.45
5	5.0	94.3	17.7	155.3	453.0	1.45
				Average	452.6	1.46

TABLE 5.3 Polariser (P_1, P_2) and Analyser (A_1, A_2) extinction positions (in degrees) for wafer 5, and derived values of oxide thickness and refractive index.

Table 5.4. The special thin oxide program (program 7) was used for thicknesses less than 500Å. The standard oxide program (program 1) was used for oxides thicker than this. Since program 1 is specified down to 400Å [136], the measurement of wafer 5 was also performed with this program. A value of 1.45 for the refractive index was assumed for all wafers. The Ellipsometer data indicates that all refractive indices are within 2% of 1.45, so this assumption is reasonable.

The repeatability of the Nanospec shows little dependence on oxide thickness. Apart from wafer 1, its value is lower than that for corresponding measurements with the Ellipsometer. Data obtained from wafer 1 show a much larger standard deviation, indicating poorer repeatability at very low oxide thicknesses (< 100Å). The thicknesses obtained using Nanospec program 1 compare well with the data from the Ellipsometer. Except for wafer 9, the two sets of data are within 1.5% of each other. It is not known exactly why wafer 9 exhibits such a large difference between the two techniques, although error in the Ellipsometer measurement is suggested by its particularly large standard deviation.

Measurements made using the Nanospec program 7 are consistently 15 - 25Å less than the corresponding Ellipsometer measurements. Wafer 5 also indicated a difference of about 15Å between the two Nanospec programs. This suggests that the Nanospec program 7 has a slightly different calibration from that of program 1.

The above analysis shows that both the Nanospec and Ellipsometer generally give reproducible results. Measurement of thinner oxides is less precise in both instruments. An important difference between the two, however, is the time taken to complete a measurement. The Nanospec allows oxide thickness data to be amassed

Wafer No.	Nanospec* Program	Oxide Thickness (Å)	s.d. (Å)
1	7	76.3	8.13
2	7	182.4	1.53
3	7	274.2	2.53
4	7	361.7	0.93
5	7	435.7	0.67
5	1	449.5	1.78
6	1	581.3	1.25
7	1	806.8	2.39
8	1	1031.8	0.42
9	1	2263.3	2.21
10	1	8981.5	0.53

* 7 = program for thin oxides, 1 = standard oxide program

TABLE 5.4 Mean and standard deviation for five repeated measurements of oxide thickness using the Nanospec.

much more rapidly than the Ellipsometer. Since this project requires a considerable amount of oxide thickness data to be collected, the Nanospec has therefore been the preferred instrument for thickness measurement.

5.2 Oxide Quality

Although thickness of the oxide film is the most important concern in this project, consideration must also be given to the oxide quality, since it directly affects device performance and reliability.

Thermally grown oxide can support locally unbalanced charges, both in the bulk film and in the oxide-silicon interface region. These charges influence the electrical properties of MOS devices, particularly the threshold voltage (V_T) which is one of the most important MOS parameters (see Appendix A). V_T for an n-channel device is given by

$$V_T = \emptyset_{ms} - Q_{eff}/C_{ox} + 2\emptyset_F - Q_B/C_{ox} \qquad (5.8)$$

where the terms have their usual meaning (also defined in Appendix A). Q_{eff} is the total effective oxide charge, and, as a first-order term in the expression for V_T, strongly determines device operating characteristics. Various types of charge contribute to Q_{eff} and these are discussed in Section 5.2.1.

The value of Q_{eff} can be altered by normal device operating conditions, and charge in the oxide film is closely related to device stability problems. It is therefore crucial that oxide charges are carefully controlled. Their magnitude becomes a direct measure of

the quality of the oxide film. The presence of charge is often brought about by the oxidation process itself, with important consequences for gate oxidation procedures.

The electrical properties of the MOS capacitor have been extensively studied and have provided the basis for much of the current understanding about oxide charges [39,138]. The MOS capacitor is widely used to monitor oxide quality, both in research applications and in routine process control. This interest in the MOS capacitor is due to its relative ease of fabrication (see Section 5.1.4) and to the relevance of its properties to more complex devices. In particular the MOS capacitor structure forms an integral part of MOSFETS and non-volatile memory cells. Methods of detecting and assessing oxide charges using MOS capacitor analysis are discussed in Section 5.2.2.

Dielectric breakdown measurements are also frequently used as an indicator of oxide quality. They relate to the presence of defects in the oxide film and its long-term reliability (constant electrical stress of an oxide film can ultimately lead to 'wear-out' of its dielectric properties). Oxide characterisation using breakdown measurements is discussed in Section 5.2.3.

5.2.1 Oxide Charges

This review of the various types of oxide charge is based on the classification produced by Deal [139] to remove the ambiguity associated with oxide charge terminology. Four types of charges have been identified and their spatial location in the MOS structure is shown in Fig. 5.6. The charge types are as follows.

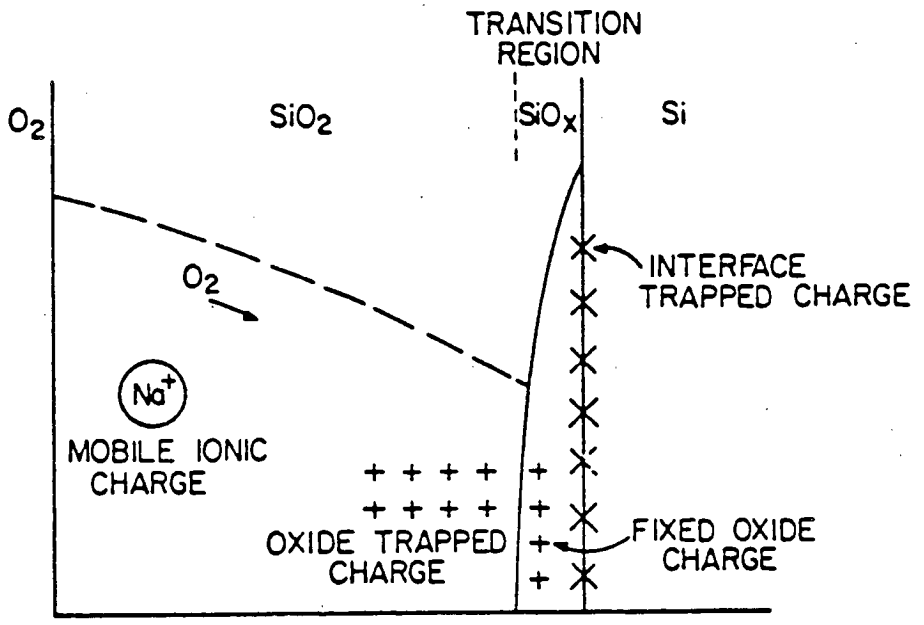


Fig. 5.6 The spatial location of charges in thermally oxidised silicon.

i) Fixed oxide charge (Q_f)

Q_f is located in the oxide within 30Å of the Si-SiO_2 interface. Its formation is directly associated with the oxidation process and is thought to be due to incompletely oxidised silicon. Q_f depends on both the oxidising ambient and temperature [140], and can be reduced by inert ambient annealing (N_2 or Ar). The value of Q_f does not change with applied gate bias (hence the term 'fixed' charge) It acts as a positive charge sheet near the Si-SiO_2 interface that induces a negative image charge at the surface of the underlying silicon. This affects MOSFET operating characteristics such as threshold voltage, drain junction breakdown and transconductance.

ii) Mobile ionic charge (Q_M)

Q_M is attributed mainly to alkali metal ions in the oxide [141], such as sodium and potassium. These are free to drift slowly through the oxide and, under the influence of an applied positive gate bias, will be driven towards the Si-SiO_2 interface. Here they pile up and, as for Q_f, affect device characteristics. This process causes a device stability problem since in-service operating conditions will cause variation in device characteristics through time. Alkali metal ions, sodium in particular, can be introduced into the oxide from several sources including processing equipment, chemical reagents and the human body. Clean room technology, chlorinated oxidation and wafer-cleaning procedures have all helped to substantially reduce this once serious contamination problem.

iii) Interface trapped charge (Q_{it})

Interface traps are defects in the Si-SiO_2 lattice associated with the thermal oxidation process. They are electronic states located at the Si-SiO_2 interface and can have energy levels which lie in the silicon bandgap. The occupancy of these traps changes in response to changes in band-bending at the silicon surface through electron or hole transitions from the silicon conduction and valence bands. The charge held in the traps (Q_{it}) is therefore not fixed, but rather depends on the applied gate bias. This implies a gate bias dependence of device parameters. The density of interface traps depends on the oxidation conditions used [142] and can be substantially reduced by a low temperature (450°C) anneal in hydrogen.

iv) Oxide trapped charge (Q_{ot})

Bulk charge traps are distributed throughout the oxide layer and, unlike interface traps, are not in electrical communication with the underlying silicon. They are associated with oxide defects such as impurities and broken bonds. Charge build-up occurs when these traps become filled with holes or electrons introduced into the oxide. This can occur as a result of electron-hole pair creation from ionising radiation (which can originate during processing from ion implantation or e-beam evaporation, for example) or from hot carriers injected into the oxide during operation. Although Q_{ot} can be reduced by low temperature annealing, the neutral traps may remain and become susceptible to subsequent charging. Q_{ot} causes MOSFET stability problems, but is used advantageously in electrically alterable memory devices.

5.2.2 Capacitance and Conductance Analysis

Much information on oxide charges can be obtained from MOS capacitor analysis, but unfortunately this is not a trivial task. Meaningful results require complex measurement procedures and careful attention to the assumptions and applicability of analysis techniques. Computer-controlled measurement systems, such as the HP4061 system mentioned in Section 5.1.4, provide a means of automating data collection and analysis. A MOS capacitor test system based on the HP4061 which has recently been developed at the EMF [143] allows oxide quality parameters to be easily and accurately determined.

The following discussion shows some of the methods that can be used in MOS capacitor analysis. It should be pointed out that alternative methods to the ones shown here often exist [39].

An ideal MOS capacitor has no locally unbalanced charges in the oxide, infinite oxide resistivity and zero metal-semiconductor work function difference. Based on these properties, a theoretical ideal high-frequency C-V characteristic can be calculated. The flatband condition (defined in Section 5.1.4) occurs when the silicon surface potential (ψ_s) is zero and the voltage at which this occurs is called the flatband voltage (V_{fb}). In the ideal case V_{fb} is zero and the corresponding capacitance is given by

$$C_{fb} = \frac{C_{ox} \cdot C_{sfb}}{C_{ox} + C_{sfb}} \qquad (5.9)$$

C_{sfb} is the silicon capacitance at the flatband voltage, and can be calculated using

$$C_{sfb} = \frac{E_s}{(2\ E_s kT/q^2\ N_{sub})^{\frac{1}{2}}} \qquad (5.10)$$

This assumes a uniformly doped substrate with doping density N_{sub}.

Oxide charges shift and distort the C-V characteristic from the ideal case (Fig. 5.7). Charges near the interface induce an opposite charge at the silicon surface, and in order to produce the flatband condition this must be compensated for by the applied bias, i.e. V_{fb} is shifted to non-zero values. The value of V_{fb} can be determined using equation 5.9 to calculate C_{fb} and then finding the voltage to which it corresponds from the C-V characteristic. V_{fb} is related to the charge concentration in the oxide by the following equation.

$$V_{fb} = \phi_{MS} - Q_{eff} / C_{ox} \qquad (5.11)$$

ϕ_{MS} accounts for non-zero metal-semiconductor work function differences. Q_{eff} is a combined charge density, with contributions from all types of oxide charges (Q_f, Q_M, Q_{ot}, Q_{it}), but it is mainly due to Q_f which, as already mentioned, acts as a large charge sheet at the interface. Since the distribution of Q_M and Q_{ot} in the oxide is not known, the individual contributions cannot be determined. The Q_{it} contribution corresponds to the net occupancy of interface traps at the flatband

voltage. Q_{eff} is generally positive although negative values can be obtained (this may be due to inaccurate metal-semiconductor work functions [143]).

Mobile charge density (Q_M) can be further assessed by bias-temperature (BT) stressing [39,133] of the MOS capacitor. BT stressing involves heating the sample to 200°C for 10 minutes whilst a negative bias of -10V is applied on the gate (values are approximate and will depend on the specific situation). After cooling, the high-frequency C-V characteristic is remeasured to determine the new shift in the flatband voltage V_{fb}^- (Fig. 5.8). The process is then repeated with positive applied bias, to determine the shift V_{fb}^+.

The effect of BT stress is to move mobile ions away from the $Si-SiO_2$ interface (-ve stress) or to pile them up in a charge at the interface (+ve stress). Both affect the image charge induced at the silicon surface and so change V_{fb}. Assuming that the BT stress has negligible effect on the other charge types, and that mobile ions do not discharge at the silicon and metal electrodes, then Q_M can be calculated using

$$Q_M = C_{ox} (V_{fb}^+ - V_{fb}^-) \qquad (5.12)$$

Small-signal admittance measurements on the MOS capacitor have been widely used in the investigation of interface trap properties. As mentioned in Section 5.2.1, interface trapped charge depends on gate bias. A small ac voltage superimposed on the gate bias results in changing occupancy of the interface traps. This can be detected in the measured capacitance and conductance of the device. Since interface traps have a response time

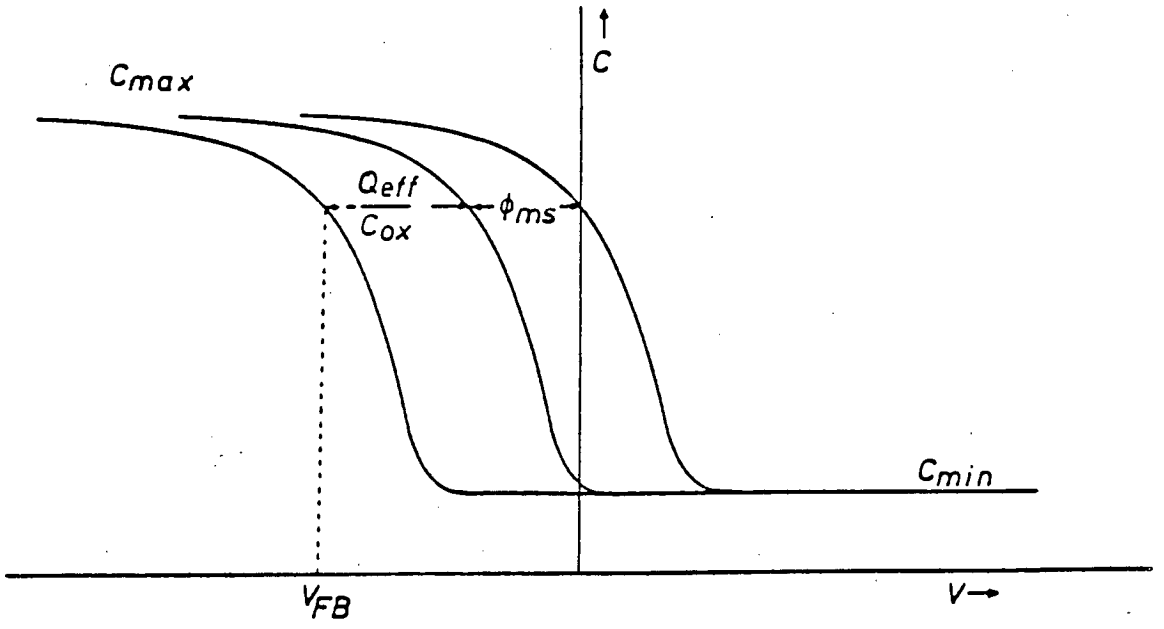


Fig. 5.7 Shift of the ideal high-frequency C-V characteristic due to the work function difference and effective oxide charge.

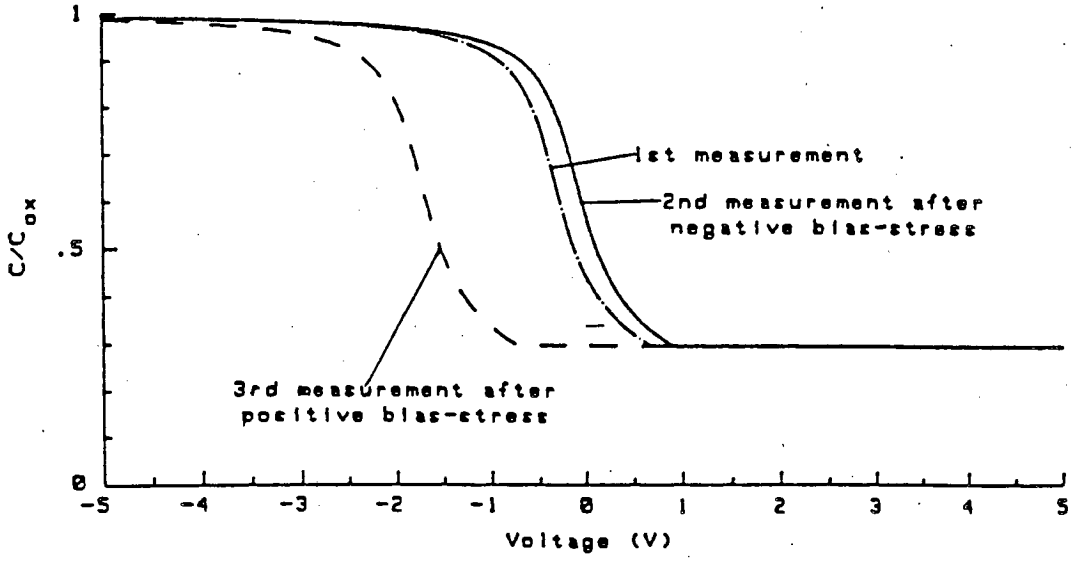


Fig. 5.8 C-V characteristics before and after BT stresses.

for carrier capture and emission, the measured effect is frequency dependent.

Of many procedures for analysis of interface traps, the conductance method is the most accurate and sensitive [138]. It involves measuring the device conductance (G) against gate bias at several different ac frequencies. Using a model that represents the MOS structure as equivalent parallel capacitance and conductance, the density of the interface traps (D_{it}) at different energy levels in the silicon bandgap can be calculated. The method is complex and is not dealt with here, but a detailed explanation can be found in Chapter 5 of Nicollian and Brews' book [39]. On a practical level, very small signal levels are involved, requiring instrumentation with high sensitivity and accuracy, and careful attention to measurement conditions (discussed later). Interpretation of measured data requires computing power for the method to be practically feasible. A system that allows fully automated measurement of D_{it} has recently been developed at the EMF [143], permitting this analysis of oxide quality to be carried out much more routinely.

Minority carrier generation lifetime (τ_g) provides another estimate of oxide quality since generation occurs at the silicon surface via interface traps (generation can also occur in the silicon bulk via bulk traps). τ_g can be measured by rapidly switching the sample from accumulation into deep depletion and measuring the changing capacitance as a function of time. Generation occurs to restore the minority carrier concentration to its equilibrium value, changing the depletion layer width and the measured capacitance. τ_g can be then determined by several analyses [39], but again this is not straightforward and the methods are not discussed here.

The complexity of MOS capacitor analysis has necessitated the development of advanced computer-based measurement systems. However, this alone will not guarantee reliable and accurate results. Careful attention must be paid to sample preparation and measurement conditions. The main points to consider are:

i) Gate material can affect the MOS capacitor behaviour through the metal-semiconductor work function difference and must be chosen accordingly.

ii) An ohmic back contact (low series resistance) is essential for accurate measurement. This is easily provided by the deposition of metal (e.g. aluminium or gold) on the stripped (i.e. no oxide) underside of the sample and the use of a vacuum chuck to hold the sample during testing.

iii) To minimise electrical interference, samples should be probed in a screened box. Coaxial connection between the sample and the measurement instrumentation should be used.

iv) The screened box should be light-tight to prevent light penetrating the MOS capacitor and affecting its properties.

Despite the recent equipment advances, MOS capacitor measurements are neither a straightforward nor immediate means of assessing oxide quality. The difficulties are confounded by the large range of analysis methods that can be used. This has meant that full analysis has not been widely introduced into production environments. Bearing in mind the crucial role of oxide charges in device operation, the risks inherent in such a policy are high.

The methods just discussed, as well as the dielectric breakdown measurements described in the following section, shall be used in Chapter 6 to examine the quality of oxides grown in the EMF research furnace.

5.2.3 Dielectric Breakdown

The MOS capacitor structure also provides a convenient means of investigating oxide integrity by dielectric breakdown measurements. Breakdown studies are inherently simpler than the capacitance and conductance methods just discussed, being less prone to measurement error and requiring no complex analysis (except perhaps on a statistical basis). No special physical assumptions are introduced and the measurements have a wide validity, allowing straightforward comparison of results. Breakdown data has consequently been widely used to report on oxide quality in a variety of studies [42,93,102,144-147].

Oxide breakdown voltage is normally measured by increasing the bias on the MOS capacitor gate until breakdown occurs. Breakdown is said to have occurred when the current through the oxide exceeds a certain critical level. Leakage currents through the insulating oxide are extremely small and breakdown levels are chosen typically around 10uA. To obtain thickness-independent data, results are usually quoted as breakdown field strengths. Good breakdown field strengths for silicon dioxide are 10-12 MV/cm.

In the absence of any defects in the oxide, the same breakdown field would always be obtained. This is referred to as intrinsic breakdown [148]. The presence of oxide defects - contaminants, dislocations, asperities and other inhomogeneities - causes localised regions of

decreased dielectric strength, where breakdown can occur at lower electric fields than for intrinsic breakdown. Because of the distribution of defects throughout an oxide film, the testing of many samples results in a range of breakdown fields with values up to the field strength at which intrinsic breakdown occurs. The probability of finding a defect increases with increasing gate electrode area, which must therefore be considered when making comparisons.

The variation in breakdown over many samples is often displayed as a histogram (Fig. 5.9) or as a probability of failure curve (Fig. 5.10). Both give a qualitative indication of the oxide quality. The defect density per unit area, D, can be calculated from the breakdown data using

$$D = \frac{1}{A} \ln (1 - F(E)) \qquad (5.13)$$

where A is the area of the test device and F(E) is the cumulative probability of failure at the electric field, E, corresponding to the onset of intrinsic breakdown (i.e. the breakpoint in the cumulative probability plot). Equation 5.13 then gives an upper limit estimate of D.

The defect density depends on the processing conditions used. Cleanliness is particularly important since the presence of impurities can lead to local defect formation. Thus inadequate pre-oxidation cleaning, incomplete drying and dirty or devitrified furnace tubes can all have an adverse effect on the defect density. Cleaning of the furnace tube with HCl/O_2 mixtures before oxidation or moderate use of HCl during oxidation have

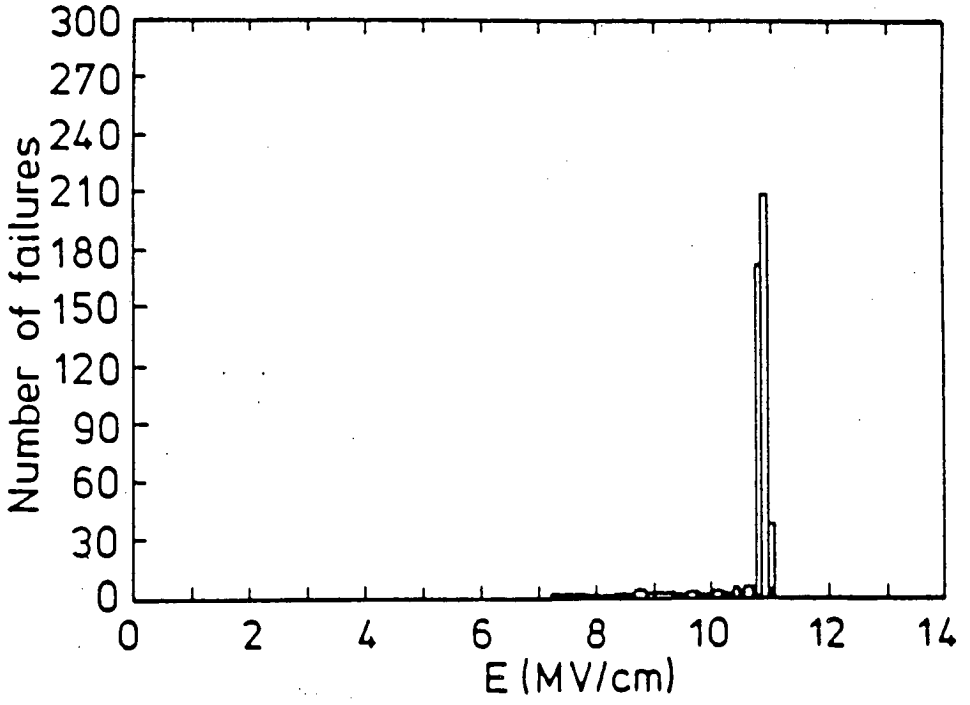


Fig. 5.9 Histogram showing the oxide breakdown field strength, E, of a number of samples.

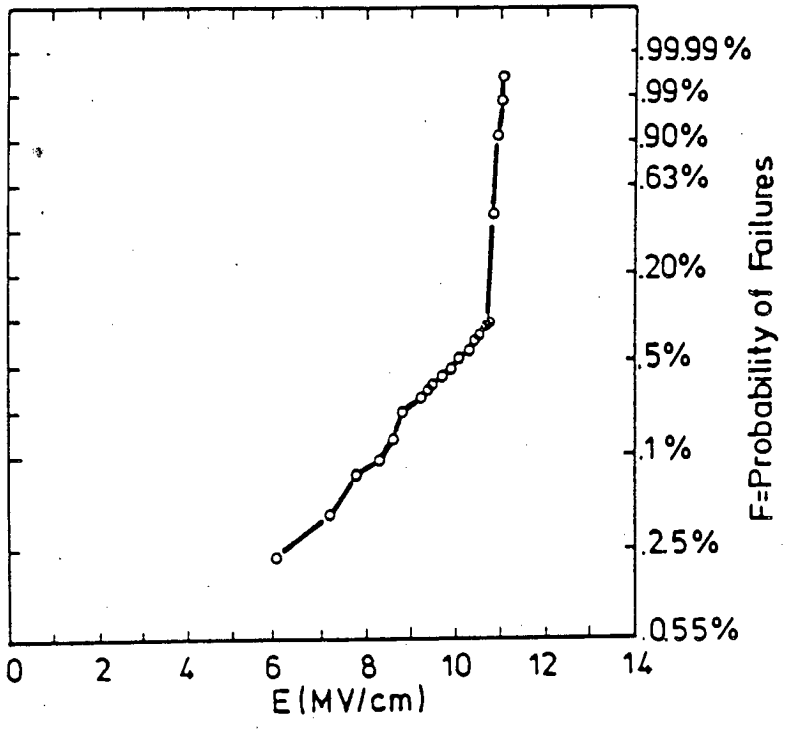


Fig. 5.10 The same breakdown data of Fig. 5.9 displayed as a probability plot.

been found to reduce defect densities [102]. The use of high oxidation temperatures and nitrogen annealing have been shown to increase defect density [147].

The mechanism of breakdown involves flow of leakage currents through the oxide [149]. These occur at high fields due to injection of charge into the oxide conduction band. It is thought that the oxide becomes permanently damaged when these charges give up their high potential energy at one of the electrodes or at charge traps (which may have been created by previous damage). This ultimately leads to a low resistance path between the electrodes, and to breakdown. A minimum level of charge (Q_{BD}) must be injected in order to evoke a breakdown. This mechanism also explains time-dependent breakdown, where oxides constantly stressed by application of high fields eventually breakdown. It is possible for Q_{BD} to be attained at fields below the breakdown value. Time-dependent dielectric breakdown presents a reliability problem for MOS ICs, since breakdown can occur as a result of prolonged usage under normal operating conditions.

CHAPTER 6

EVALUATION OF EXISTING FURNACE TECHNOLOGY

6.1 Introduction

In Chapter 4, the main features of an advanced oxidation furnace system were described. Using the EMF research furnace, the capabilities of a typical furnace control system were examined through measurement of selected furnace parameters such as temperature and gas flow rates. In this chapter, the performance of the furnace system is investigated by analysing the oxide films grown in the furnace. In particular, the extent to which the control problems discussed in Chapter 4 translate into variability in oxide thickness is investigated. Identification and understanding of the mechanisms of thickness variability is necessary if an improvement in control is to be achieved. Despite its importance, this is an aspect of oxidation that has received little attention in the literature.

Before the uniformity study was carried out, the general suitability of the EMF furnace for use in silicon oxidation was verified. The procedures and results are reported in the first section of the chapter. The second section then examines the uniformity of the oxide thickness in detail.

6.2 Furnace Verification

An initial test was carried out to check for oxidant leakage into the furnace tube. This is crucial, since leakage would influence oxidation rates by giving

'parasitic' oxide growth during supposedly inert periods of a furnace process, as well as suggesting a general contamination problem. Oxide quality was then checked, using a MOS capacitor structure, to ensure that there was no contamination problem which would render the furnace unsuitable for the growth of gate oxides.

6.2.1 Parasitic Oxide Growth

Unwanted oxidant can penetrate into the furnace from several sources - for example, through leaks in the furnace hardware or through back-diffusion of atmospheric oxygen via the furnace end-cap [150]. The presence of this oxidant leads to extra 'parasitic' oxide growth which may be localised to one part of the furnace tube and so seriously disrupt the study of oxide thickness uniformity. However, the extent of parasitic oxidation can be easily assessed by subjecting wafers to a high-temperature furnace treatment in an inert ambient, then checking for a change in oxide thickness.

Such a test was carried out using a nitrogen ambient at 1025°C for a period of 60 minutes. Table 6.1 gives details of the furnace recipe used. The nitrogen flow of 5.0 SLM corresponds to a displacement velocity down the tube of approximately 3 cms^{-1} (at temperature). Wafers were placed at three positions in the furnace hot-zone: one at each end and one in the centre. The wafers were held in 'transverse' quartz wafer boats (i.e. boats that positioned the wafers at 90° to the furnace tube axis).

The study into repeatability of the two instruments available for measuring the thickness of thin oxide layers - the Nanospec and the Ellipsometer - indicated poorer precision at very low thicknesses (Section 5.1.6).

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position (mm)	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	1025	-	'Standby' Position
1	5.0	0	0	0	250	1025	15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700	1025	15	Paddle replaced Temperature allowed to settle
3	5.0	0	0	0	1700	1025	60	Wait in N_2 ambient
4	5.0	0	0	0	250	1025	15	Paddle out of furnace to unload wafers

i) Paddle position is a number (in mm) used by the process controller. At 1700mm the paddle is fully inserted into the furnace tube, and the end-cap is flush against the end of the tube. At 250mm the furnace paddle is fully withdrawn for loading/unloading wafers.

TABLE 6.1 N_2 Process for furnace verification.

To maintain an acceptable level of accuracy, silicon wafers with an oxide layer of around 1200Å were therefore used instead of fresh wafers (whose native oxide could not be accurately measured). The wafers were given a standard RCA-clean (Section 4.3), then the oxide thickness was measured at five sites using the Nanospec. Fig. 6.1 shows the measurement sites used. Linear scales on the X-Y stage of the Nanospec allowed these positions to be determined to within $\pm$ 1mm. After the furnace treatment, the oxide thickness at the same five sites was remeasured and the increment calculated. The whole procedure was then repeated with the N_2 flow increased from 5.0 SLM to 10.0 SLM. The results are summarised in Table 6.2.

An increase in oxide thickness can be observed on all wafers but only to a relatively small extent. At the centre of the tube, the position where wafers will generally be placed during oxidation treatments, the parasitic growth is less than 10Å even at the low nitrogen flow rate of 5.0 SLM. Considering the extreme conditions of the test, it may be concluded that parasitic oxidation will have a negligible effect under normal processing conditions. A trend of increased parasitic growth at the end-cap end of the furnace is apparent in the data of Table 6.2. The effect is lower at the increased nitrogen flow of 10.0 SLM. This suggests that the cause of trace amounts of oxidant in the furnace tube may indeed be back-diffusion of air at the furnace end-cap.

6.2.2 Oxide Quality

The quality of SiO_2 films grown in the furnace was measured using aluminium-gated capacitor structures.

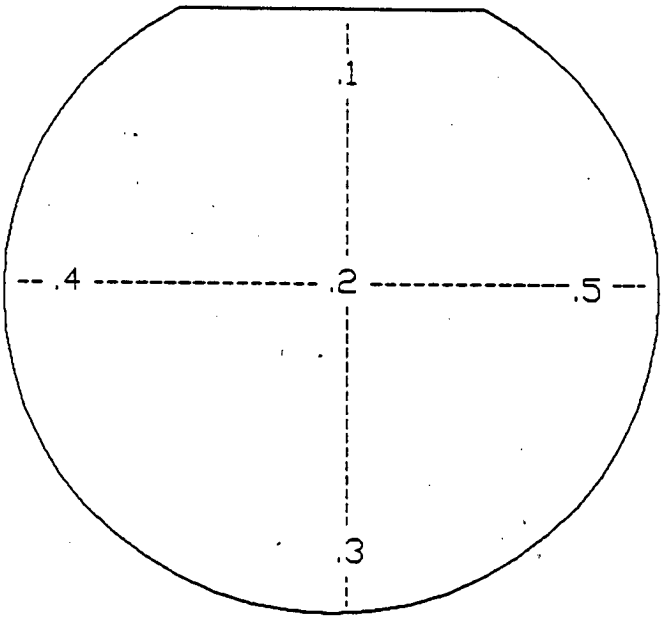


Fig. 6.1 Positions of the 5 sites measured on each wafer. The outer sites are 2.5cm from the centre site.

N_2 flow (SLM)	Wafer position	Δx_o(Å)	s.d.(Å)
5.0	1 'Gas inlet' End	1.0	1.6
	2 Centre of Tube	8.8	1.1
	3 'End-Cap' End	17.4	9.7
10.0	1 'Gas inlet' End	5.2	1.5
	2 Centre of Tube	6.4	1.1
	3 'End-Cap' End	8.0	1.2

TABLE 6.2 Mean change in oxide thickness (Δx_o) and standard deviation (s.d.) after high-temperature N_2 treatment (average of 5 sites on a single wafer).

These structures were fabricated using the processing sequence given in Table 6.3. The starting material was p-type <100> orientated silicon with a nominal resistivity of 10-14 ohm.cm.

A gate oxidation process which would give an oxide thickness of around 250Å, typical of the oxides required in VLSI processing, was used. It involved 20 minutes in a dry oxidation ambient with 5% added HCl, followed immediately by a 30 minute anneal in N_2. Both oxidation and anneal were carried out at a temperature of 950°C. The complete furnace procedure is given in Table 6.4.

Four separate batches were processed, each containing two test wafers. The oxide thickness was measured with a Nanospec directly after each oxidation. The average thickness over the eight wafers was 231Å. Patterning of the aluminium layer gave capacitor arrays over the entire wafer surface. Capacitor dots of area $10^{-3}cm^2$ were tested on wafers from each batch using a HP4061-based test system [143]. The methods have already been discussed in Section 5.2 and only the results are presented here.

Breakdown field strengths were generally good. The average value over 32 measurements was 11.1 $MVcm^{-1}$ with a standard deviation of 0.7 $MVcm^{-1}$. C-V measurements at 1MHz indicated a combined positive charge density (Q_{eff}) of 1 x 10^{11} cm^{-2} to 3 x 10^{11} cm^{-2} which is an acceptable level. Bias-temperature stressing at 200°C under both positive and negative bias at 2.5V (approximately equivalent to $10^6 Vcm^{-1}$) revealed virtually no shift in the C-V curves, indicating low levels of mobile charge. Capacitance-time analysis indicated minority carrier lifetimes of around 10usec, which again is an acceptable level. These results indicate that the

1. Pre-oxidation wafer clean
   The cleaning schedule given in Section 4.3 was used (RCA-clean). Wafers were rinsed in a four-cycle dump-rinser and spin-dried in nitrogen.

2. Oxidation
   Dry oxidation (5% HCl) at 950°C for 20 minutes.

3. Aluminium deposition
   Aluminium was deposited in an e-beam evaporator to give a thickness of 10000Å

4. Resist coat and pattern
   A mask that would give capacitor dots of various sizes was used.

5. Aluminium etch
   Wet chemical etch in phosphoric acid at 55°C.

6. Resist coat and oxide etch
   Removal of oxide from the back surface of wafers.

7. Resist removal
   Ten minutes in fuming nitric acid.

8. Aluminium deposition on backs of wafers
   As for step 3, 10000Å of evaporated aluminium to give an ohmic back-side contact.

9. Aluminium sinter
   Furnace process in H_2/N_2 mixture at 450°C for 30 minutes.

TABLE 6.3 Procedure for MOS capacitor fabrication.

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	950	-	'Standby' Position
1	5.0	0	0	0	250	950	15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700	950	15	Paddle replaced
3	0	5.0	0	0.25	1700	950	20	Oxidation
4	5.0	0	0	0	1700	950	30	Anneal
5	5.0	0	0	0	250	950	15	Paddle out of furnace to unload wafers

i) See Table 6.1

TABLE 6.4 Gate oxidation process.

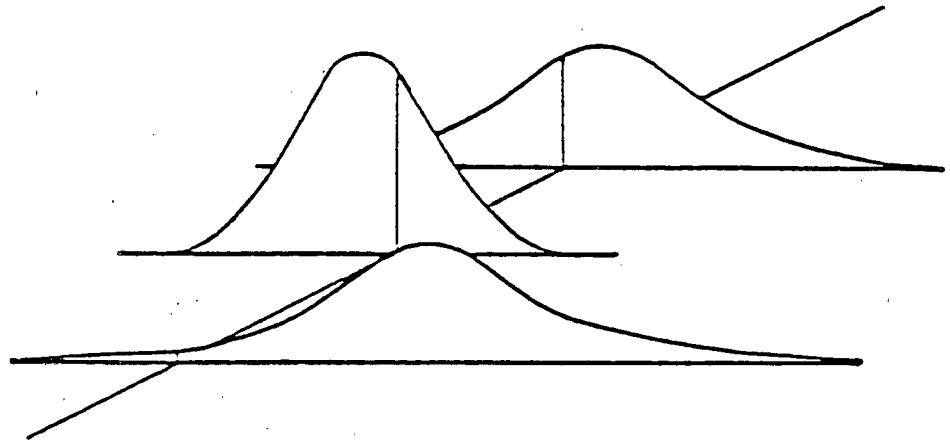
furnace is capable of growing gate oxides of a quality suitable for VLSI.

6.3 Oxide Thickness Variability

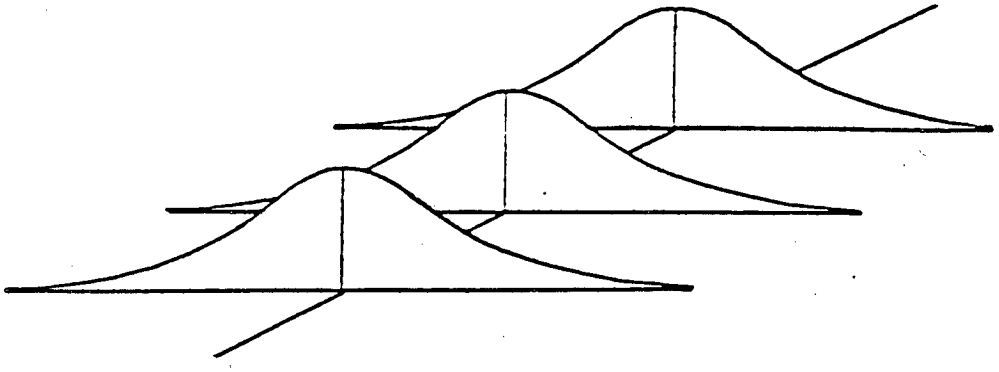
If control of oxide thickness is to be improved, then it is important to understand the causes of thickness variation. The oxidation process is subject to many sources of variation, such as those identified and described in Section 4.2, the cumulated effect of which becomes apparent in the measured oxide thickness. In a batch process like oxidation, oxide thickness forms a distribution within each batch. Additionally, the individual batch distributions will change from one batch to another. Splitting oxide thickness variation into the two components 'within-batch' and 'between-batch' greatly facilitates the task of process control, since each component can be examined separately. This is demonstrated schematically in Fig. 6.2. In Fig. 6.2(a), both within-batch and between-batch variation are present, representing a poorly controlled process. Fig. 6.2(b) shows the much more desirable situation where the between-batch variation has been eliminated. From this stable situation, within-batch variation can be targeted, to tighten the distributions and meet the specified tolerances.

In oxidation, the following independent mechanisms lead to between-batch and within-batch variation.

i) Between-batch variation is generally present because furnace treatments are not completely reproducible. The transient effects described in Chapter 4 result in unpredictable differences between furnace treatments. Degradation of furnace components (e.g. thermocouples) will also



(a)



(b)

Fig. 6.2 Distribution of oxide thickness. Each curve represents the thickness distribution from a single batch : (a) Both within-batch and between-batch variation is present. (b) The ideal case of a stable well-controlled process - only within-batch variation is present.

result in long-term drift in process parameters and hence in oxide thickness.

ii) Within-batch variation essentially represents the 'individuality' of the furnace itself. The non-uniform temperature profiles and hydrogen-torch heating effects already mentioned are just two examples of individual furnace features that cause variation within a single batch of wafers. Gas flow patterns will also play a role, causing a within-batch uniformity dependence on gas flow rates and furnace quartzware (e.g. baffles, type of wafer boat).

6.3.1 Quantitative Study of Contributions to Variability

An investigation to measure between-batch and within-batch variation and determine their relative importance was carried out. Three oxidation procedures were studied, involving both wet and dry ambients, in order to investigate the influence of oxide thickness and oxidation rate on the variability. All oxidations were carried out in the EMF research furnace at a temperature of 950°C. Two wet oxidation procedures with oxidation times of 5 minutes and 20 minutes and a 30 minute dry procedure were used (hereafter referred to as 5min-wet, 20min-wet and 30min-dry procedures respectively). The time/temperature combinations were chosen to include thicknesses typical of VLSI gate oxides. Details of the furnace recipes are given in Tables 6.5, 6.6 and 6.7. The basic formats are identical apart from oxidation time and ambient, to allow direct comparison.

Each oxidation procedure was repeated five times over a period of approximately 3 weeks. The substrate material was <100> lightly-doped p-type, 3" silicon

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	950	-	'Standby' Position
1	5.0	0	0	0	250	950	15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700	950	15	Paddle replaced
3	0	2.5	3.0	0	1700	950	5	Oxidation
4	5.0	0	0	0	1700	950	10	Anneal
5	5.0	0	0	0	250	950	15	Paddle out of furnace to unload wafers

i) See Table 6.1

TABLE 6.5 5 minute wet oxidation process.

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	950	-	'Standby' Position
1	5.0	0	0	0	250	950	15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700	950	15	Paddle replaced
3	0	2.5	3.0	0	1700	950	20	Oxidation
4	5.0	0	0	0	1700	950	10	Anneal
5	5.0	0	0	0	250	950	15	Paddle out of furnace to unload wafers

i) See Table 6.1

TABLE 6.6 20 minute wet oxidation process.

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	950	-	'Standby' Position
1	5.0	0	0	0	250	950	15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700	950	15	Paddle replaced
3	0	5.0	0	0	1700	950	30	Oxidation
4	5.0	0	0	0	1700	950	10	Anneal
5	5.0	0	0	0	250	950	15	Paddle out of furnace to unload wafers

i) See Table 6.1

TABLE 6.7 30 minute dry oxidation process.

wafers in all cases. These were given a standard RCA pre-oxidation clean with a final dip in dilute HF solution immediately before oxidation. The wafers for the wet procedures were placed in the furnace transverse to the furnace axis. The wafers in the dry oxidation were arranged in-line to provide further insight into the mechanisms of within-batch variation. For all oxidations, 30 wafers were used, held 'flats-up' in quartz boats and placed centrally in the furnace tube. After oxidation, wafers were selected for measurement from equally spaced positions in the boat, as shown in Figs. 6.3 and 6.4. The remaining wafers acted as dummies. Oxide thickness was measured at five sites on each of the selected wafers, again following the pattern given in Fig. 6.1. Because of the large amount of data involved, the Nanospec was the preferred instrument for oxide thickness measurement. As shown in Section 5.1, this has a fast, simple operation and good repeatability down to 150Å.

The mean thickness, standard deviation and coefficient of variation (CV[1]) were calculated for each batch. The results are presented in Tables 6.8, 6.9 and 6.10 for the 5-min wet, 20-min wet and 30-min dry procedures respectively. Average oxide thicknesses of 318Å and 224Å for the 5min-wet and 30min-dry processes were obtained, which are typical of VLSI gate oxide thicknesses. The 20min-wet process, with an average thickness of 906Å, provides a thick-oxide comparison. The CVs from each process generally show an increase as the oxide thickness is reduced. This is to be expected and illustrates that the absolute within-batch variability does not decrease linearly with decreasing thickness.

1 CV = $(s/\bar{x}) \times 100$, where s=sample standard deviation and $\bar{x}$=sample mean.

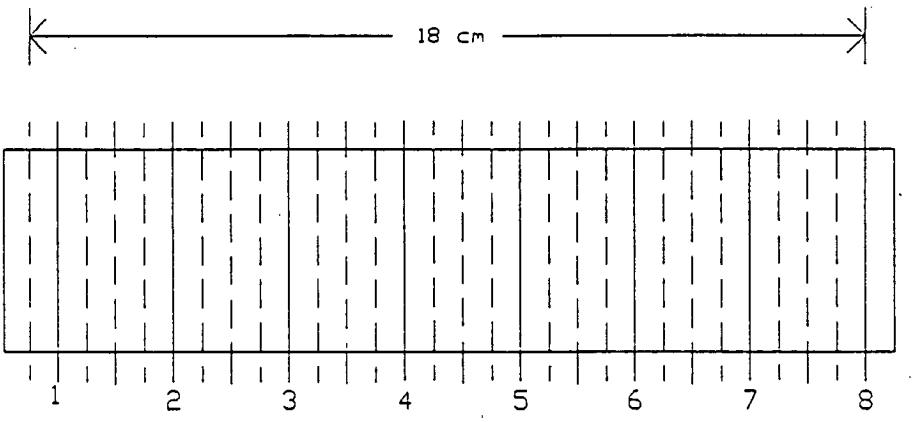


Fig. 6.3 Wafer boat used for the 5min-wet and 20min-wet oxidations. 8 wafers were selected for measurement from the positions shown in bold.

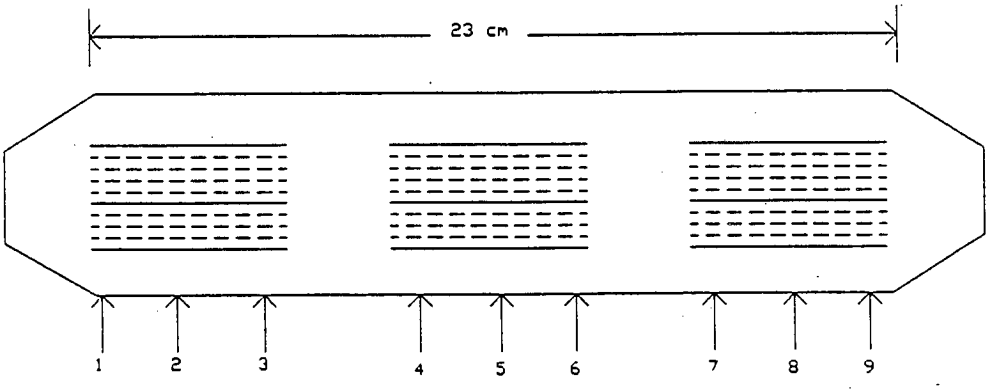


Fig. 6.4 Wafer boat used for the 30min-dry oxidation. 9 wafers were selected for measurement from the positions shown in bold.

Batch No.	N	Mean Thickness (Å)	s.d. (Å)	CV(%)
1	40	327.0	9.9	3.0
2	40	307.9	10.6	3.4
3	40	309.2	7.0	2.3
4	40	305.1	11.9	3.9
5	40	343.1	12.2	3.6
All Batches	200	318.5	17.9	5.6

TABLE 6.8 Summary data : 5 minute wet oxidation process.

Batch No.	N	Mean Thickness (Å)	s.d. (Å)	CV(%)
1	40	893.9	11.8	1.3
2	40	900.4	14.6	1.6
3	40	915.2	11.1	1.2
4	40	932.6	24.8	2.6
5	40	886.2	21.9	2.5
All Batches	200	905.7	24.1	2.7

TABLE 6.9 Summary data : 20 minute wet oxidation process.

Batch No.	N	Mean Thickness (Å)	s.d. (Å)	CV(%)
1	45	235.8	10.5	4.4
2	45	227.3	13.3	5.9
3	45	219.5	6.7	3.1
4	45	215.0	6.2	2.9
5	45	222.3	9.1	4.1
All Batches	225	223.6	11.9	5.3

TABLE 6.10 Summary data : 30 minute dry oxidation process.

The batch means are plotted in Figs 6.5, 6.6 and 6.7. No trends are seen in the plots, indicating that drift in furnace parameters has not been a problem over the three-week period of data collection. In all three cases, between-batch variation is clearly present and the situation is more analogous to Fig. 6.2(a) than to the stable situation of Fig. 6.2(b). In order to accurately quantify the contributions of between-batch and within-batch variation in the data, a detailed statistical analysis of each data set was carried out.

The general nature of this investigation makes it ideally suited to the standard statistical technique of analysis of variance (ANOVA). This tests whether the mean oxide thicknesses in the five batches differ with statistical significance. The method assumes that the thickness data are Normally distributed and that the batch variances are equal. The analyses are shown in Tables 6.11, 6.12 and 6.13. The given F-statistic tests the null hypothesis (H_0) that all batch means are equal. Departures from the null hypothesis give values of F greater than unity. Hence in all three cases there is very strong evidence against H_0 and it can be concluded that the batch means differ significantly.

Using the ANOVA table is it possible to partition the total variation directly into between-batch and within-batch components, as shown in Tables 6.14, 6.15 and 6.16.

The variance components represent the estimated variance (s^2) for each source of variability. The components can be summed to give an estimate of the total variance of the data, then expressed as a percentage of this total. The analysis shows that the proportion of variation attributable to differences between batches is 51% for the 20min-wet process, but that this proportion

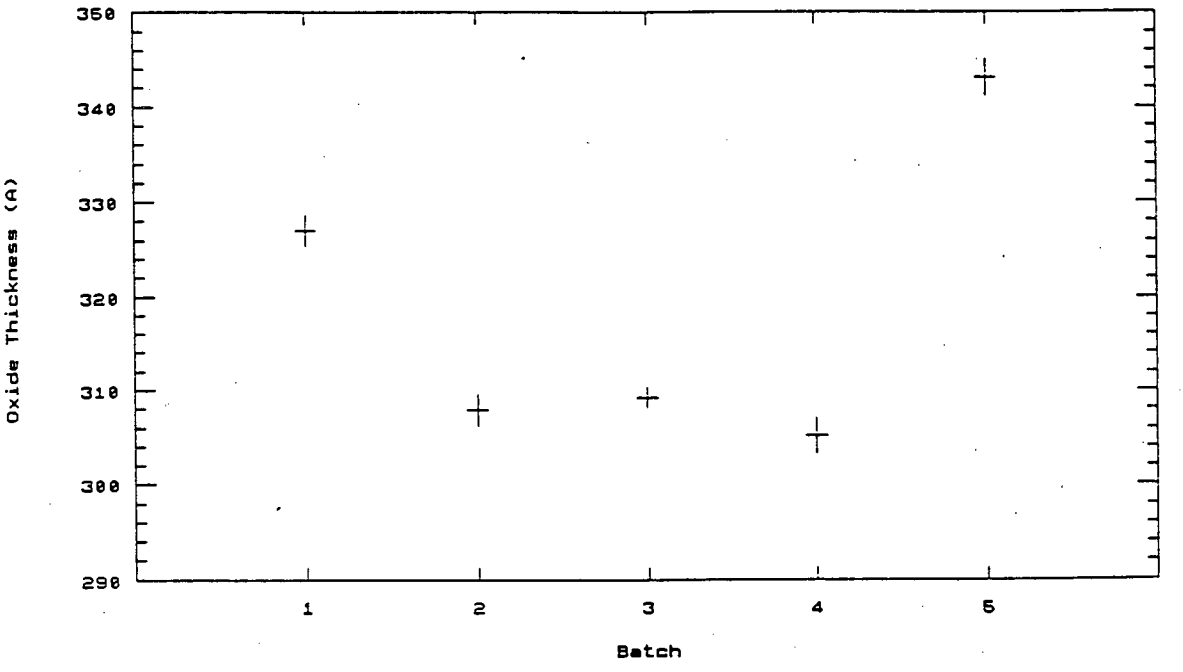


Fig. 6.5 Batch means with standard error bars for the 5min-wet process.

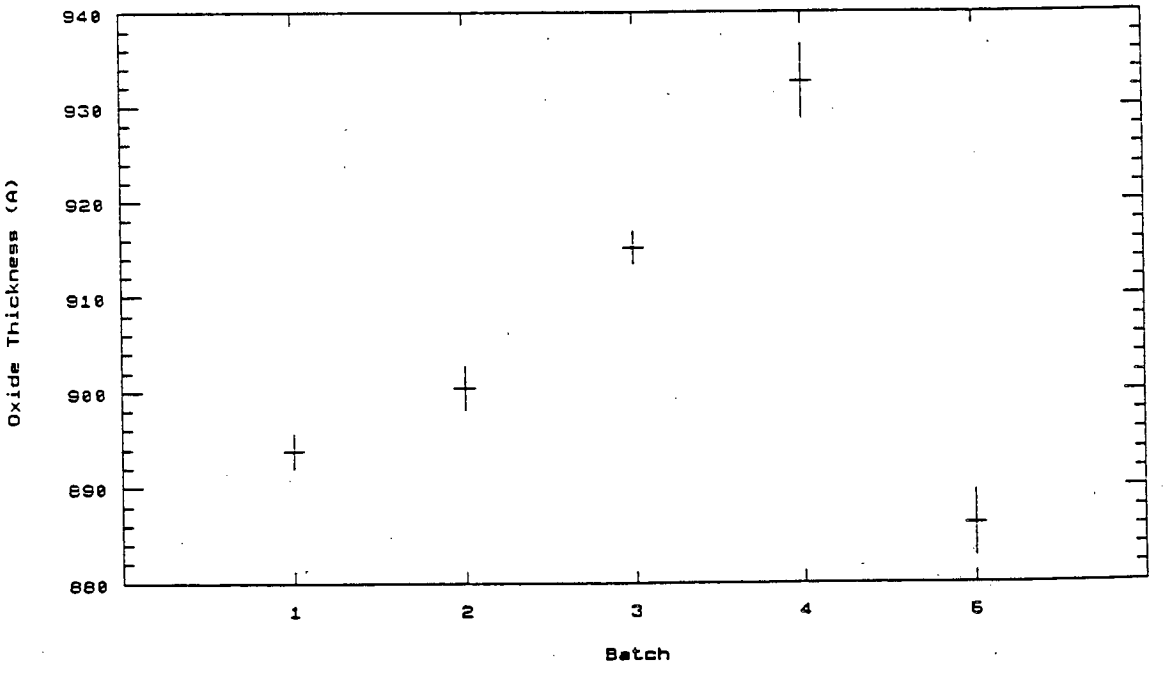


Fig. 6.6 Batch means with standard error bars for the 20min-wet process.

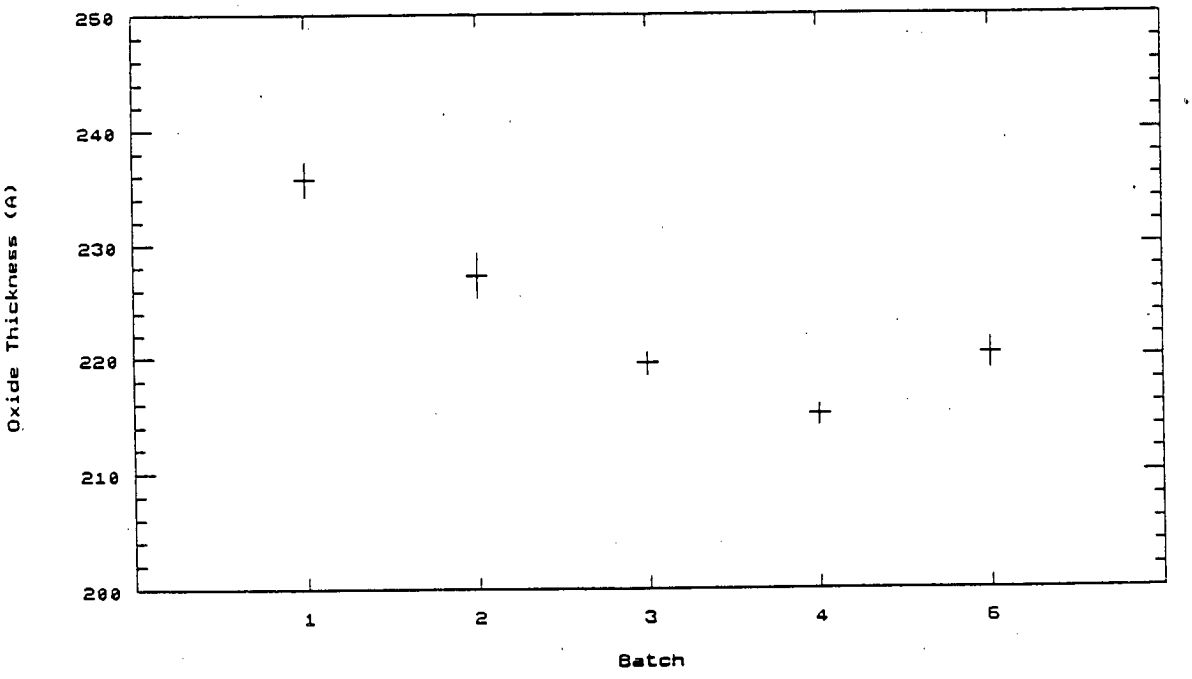


Fig. 6.7 Batch means with standard error bars for the 30min-dry process.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio
Between batches	42159.2	4	10539.8	95.4
Within batches	21534.5	195	110.4	
TOTAL	63693.7	199		

TABLE 6.11 Analysis of Variance : 5 minute wet oxidation process.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio
Between batches	54378.4	4	13594.6	43.3
Within batches	61259.8	195	314.2	
TOTAL	115638.2	199		

TABLE 6.12 Analysis of Variance : 20 minute wet oxidation process.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio
Between batches	11831.2	4	2957.8	32.7
Within batches	19881.7	220	90.4	
TOTAL	31712.9	224		

TABLE 6.13 Analysis of Variance : 30 minute dry oxidation process.

Source of Variation	Mean Square	Degrees of Freedom	Variance Component	%
Between batches	10539.8	4	260.7	70.2
Within batches	110.4	195	110.4	29.8

TABLE 6.14 Components of Variance : 5 minute wet oxidation process.

Source of Variation	Mean Square	Degrees of Freedom	Variance Component	%
Between batches	13594.6	4	332.0	51.4
Within batches	314.2	195	314.2	48.6

TABLE 6.15 Components of Variance : 20 minute wet oxidation process.

Source of Variation	Mean Square	Degrees of Freedom	Variance Component	%
Between batches	2957.8	4	63.7	41.4
Within batches	90.4	220	90.4	58.6

TABLE 6.16 Components of Variance : 30 minute dry oxidation process.

increases to 70% for the 5min-wet process. This reflects the increased relative importance of oxide growth during transient periods when the overall oxidation time is shorter. It is confirmed by the lower value of 41% obtained for the 30min-dry process. Here the proportion of between-batch variation decreases further, despite a reduction in the oxide thickness. It is worth noting that, even with the relatively long oxidation time of 30 minutes, a large proportion of variation can still be attributed to between-batch differences.

From Tables 6.14 and 6.15 it can be seen that the between-batch and within-batch variances for the 5min-wet process are both less than those for the 20min-wet process. However, the between-batch variance does not scale linearly with decreasing oxide thickness. It decreases from 332.0 to 260.7 (equivalent to a very slight change in the standard deviations from 18.2 to 16.1) for what is approximately a three-fold decrease in oxide thickness. This highlights the problem of moving to thinner oxides i.e. the between-batch variability becomes increasingly prominent relative to the oxide thickness. Although the between-batch and within-batch variances for the 30min-dry process are much lower than those obtained for the wet processes, they are still at an unacceptably high level relative to the mean thickness of 224Å (see Chapter 2).

Patterns in oxide growth within the data set are now investigated. This provides insight into the mechanisms causing within-batch variability. Figs. 6.8, 6.9 and 6.10 show the mean oxide thicknesses, calculated from all data points having the same horizontal distance along the furnace tube axis. For the wet oxidation procedures these are simply the wafer positions

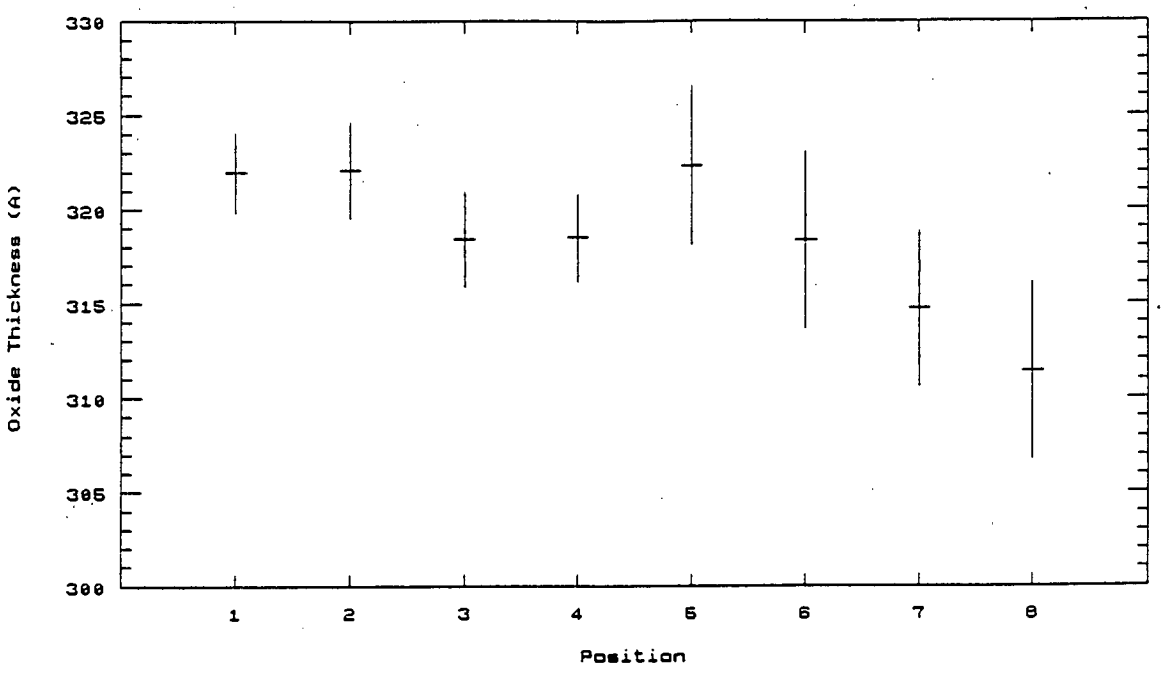


Fig. 6.8 Mean thickness and standard error bars plotted against horizontal position along the furnace tube for the 5min-wet process

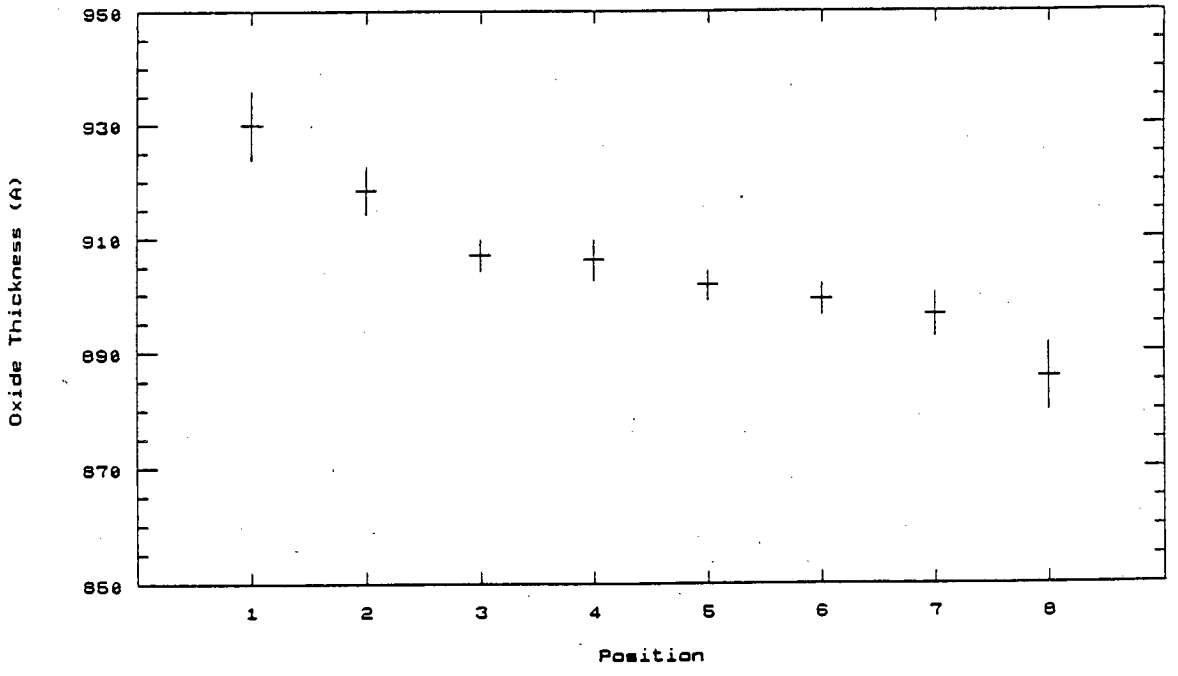


Fig. 6.9 Mean thickness and standard error bars plotted against horizontal position along the furnace tube for the 20min-wet process

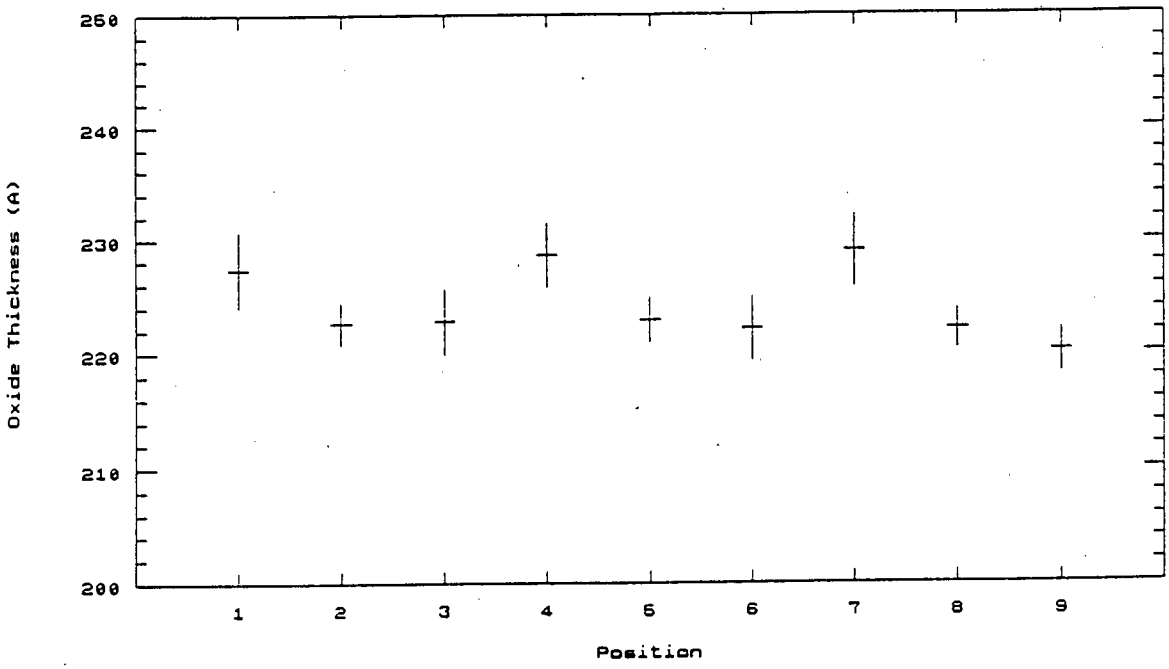
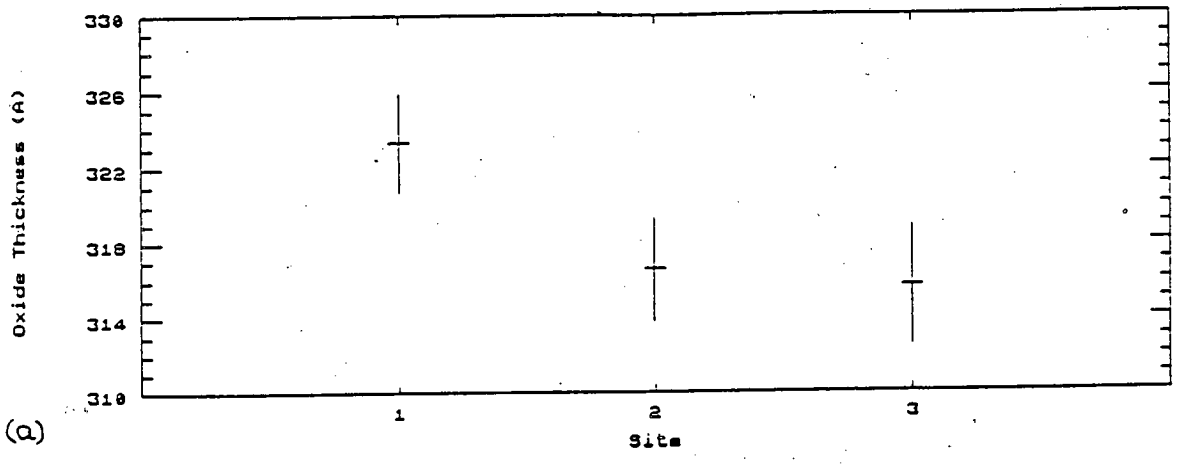


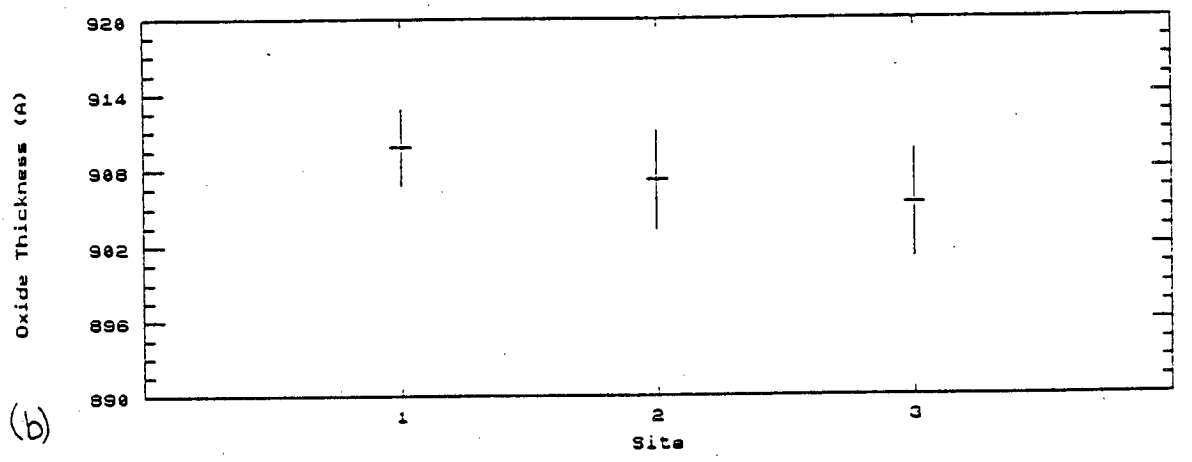
Fig. 6.10 Mean thickness and standard error bars plotted against horizontal position along the furnace tube for the 30min-dry process

which were given in Fig. 6.3, and for the dry procedure the positions were indicated in Fig 6.4. In both wet procedures, a general decrease in oxide thickness is observed in a direction away from the hydrogen flame. This indicates a non-uniform temperature profile within the tube and relates to the measurements made in Chapter 4, where the heating effect of the hydrogen torch was clearly seen. It is confirmed by the absence of such an effect from the dry procedure. Here another effect is observed due to the use of an in-line boat. The three positions corresponding to a wafer edge facing into the gas flow (positions 1, 4 and 7) all show a similar increase in oxide thickness compared to the other positions. This seems unlikely to be a temperature effect and explanation may lie in the complex gas flow patterns around the wafer edges. This would affect the rate at which oxygen can be supplied to the $Si-SiO_2$ interface and so influence the oxidation kinetics.

The growth patterns in a vertical direction, i.e. the averages of all points from the top, centre and bottom sites of each wafer, are shown in Fig. 6.11. Despite the large error bars - caused by the combined effect of 'horizontal' variation and the between-batch variation - the same pattern is clear for all three oxidation processes: thicker oxide at the top site, thinner oxide at the bottom. This is explained by a radial temperature distribution, as already mentioned in Chapter 4, whereby the tops of the wafers are subjected to higher temperature than the bottom of the wafers.

In summary, the above investigation has shown that significant batch-to-batch variation is still present in spite of elaborate digital control in furnace systems. Within-batch variability has also been assessed and related to specific influences of furnace hardware. The





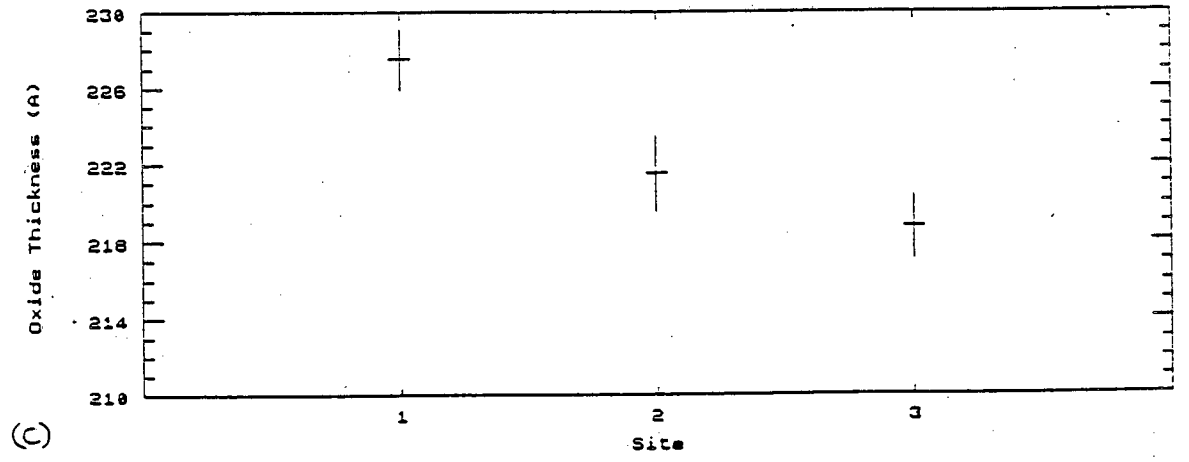


Fig. 6.11 Mean thickness against vertical position in the furnace tube for (a) the 5min-wet process (b) the 20min-wet process (c) the 30min-dry process. Position corresponds to the numbers given in Fig. 6.1. Standard error bars are also shown.

variability has been shown not to scale with decreasing oxide thickness, implying problems for thin gate oxides. It can be surmised that these problems will not go away or be improved in larger systems. Increased loads and larger wafer diameters will, on the contrary, be expected to compound the mechanisms of variability. In the following chapter, a means of alleviating the problem of batch-to-batch variation is described. The aim is a system that will permit the realisation of the stable situation which was illustrated in Fig. 6.2(b).

CHAPTER 7

A NEW FURNACE CONTROL SYSTEM

7.1 Control System Description

Current furnace systems control oxide thickness by carefully regulating the ambient and temperatures seen by the wafers. Furnace recipes contain all the necessary definitions and instructions to this effect. This approach fails however, because the assumption that there are no deviations (or insignificant deviations) from the recipe definitions does not hold. Because of transient effects, the actual oxidation conditions are not fully described by the recipe. Variability therefore exists between batches that have been processed using exactly the same recipe, as has been demonstrated in the preceding chapter. A solution can be seen in the following reasoning:

> If I knew exactly what the oxidation conditions throughout the furnace procedure were, and how these would influence oxide growth, then I could adjust the procedure to obtain the desired oxide thickness. By measuring the oxidation conditions during the furnace process itself, an active adjustment of the furnace procedure would be possible.

This reasoning details the concept of using real-time simulation as a means to enhance control of oxide thickness. The following requirements for realisation of this concept can be defined.

i) Accurate and continually updated measurement of all furnace variables (i.e. the process inputs discussed in Chapter 1) likely to affect the oxidation.

ii) A method of rapidly assessing the influence of these measurements on oxide growth, so that real-time estimates of thickness can be obtained. This estimate becomes the basis of control.

iii) The ability to change the furnace recipe during processing and so terminate the oxidation under conditions which will give the required oxide thickness.

Of the above points, ii) is the most critical. It signifies a new application of oxidation modelling beyond its predictive role in process simulation programs. The model used must be both highly accurate and very robust over the wide range of real processing conditions likely to be encountered. It must be run in real-time so that the model parameters can be immediately updated when the measured furnace variables indicate a change in conditions. For real-time simulation to be effective, the refresh period of the model parameters must be small compared with the overall system response. For furnace oxidation, update cycle times of around a few seconds are therefore preferable. This is not impractical - digital furnace controllers generally have sampling periods of less than a second.

A set of rules can be used to describe the relationship between model parameters and furnace variables. The most important model parameters were defined in Chapter 3 and are repeated here:

i) ambient (wet or dry)
ii) temperature
iii) oxidant pressure
iv) concentration of HCl
v) substrate orientation
vi) doping level at the interface

Ambient, oxidant pressure and HCl concentration can be determined by measuring the flow rates of all gases fed into the furnace. Temperature means the wafer temperature, and as a furnace variable should therefore be measured inside the tube at the positions of the wafers. The substrate orientation is simply a fixed parameter (it is unlikely to change during processing !). The doping level must only be considered for high concentrations ($> 10^{19}$ cm^{-3}), since at lower levels its influence on oxidation rate is negligible. Calculation requires modelling of the thermal redistribution of the dopant, which is not a simple task (the primary concern of the SUPREM and ICECREM simulators). It is complicated by the intimate relationship between oxidation and impurity redistribution and by segregation effects. Time-consuming iterative methods would have to be used, in which case a real-time model may not be practicable. However, since high doping levels in MOS technology apply to areas where the oxide thickness is generally not critically controlled (e.g. source/drain regions), the ability to model this effect can be conveniently omitted without seriously limiting the system capabilities.

A new control system based on real-time simulation which has been developed for use on the EMF furnace is now described.

7.2 Implementation of an Oxidation Model

The oxidation model from SUPREM III has been selected as a basis, since it adequately covers the range of oxide thicknesses and oxidation conditions likely to be encountered in VLSI processing. The model is described in detail in Section 3.5. As stated above, only lightly-doped silicon need be considered. The doping-dependent parameters can therefore be neglected, rendering the model suitable for real-time simulation.

The basic model equation (equation 3.36) in an incremental form is

$$\Delta x_o = \left[\frac{B}{(2x_o + A)} + R \right] . \Delta t \qquad (7.1)$$

where R models the rate enhancement observed with dry oxidation.

This form is ideal for real-time simulation since the values of B, A and R can be independently adjusted for each time interval Δt. The adjustment is determined by applying the appropriate rules to the most current furnace data. The change in oxide thickness for the time interval is then determined and added to the total. The total oxide thickness is used to calculate Δx_o in the following time interval as shown in Fig. 7.1.

The constants A and B are determined from exponential expressions for the linear and parabolic rate constants. The forms used are similar to those in SUPREM III (equations 3.30 and 3.31) with the omission of the factor that models high doping levels. The expressions are :

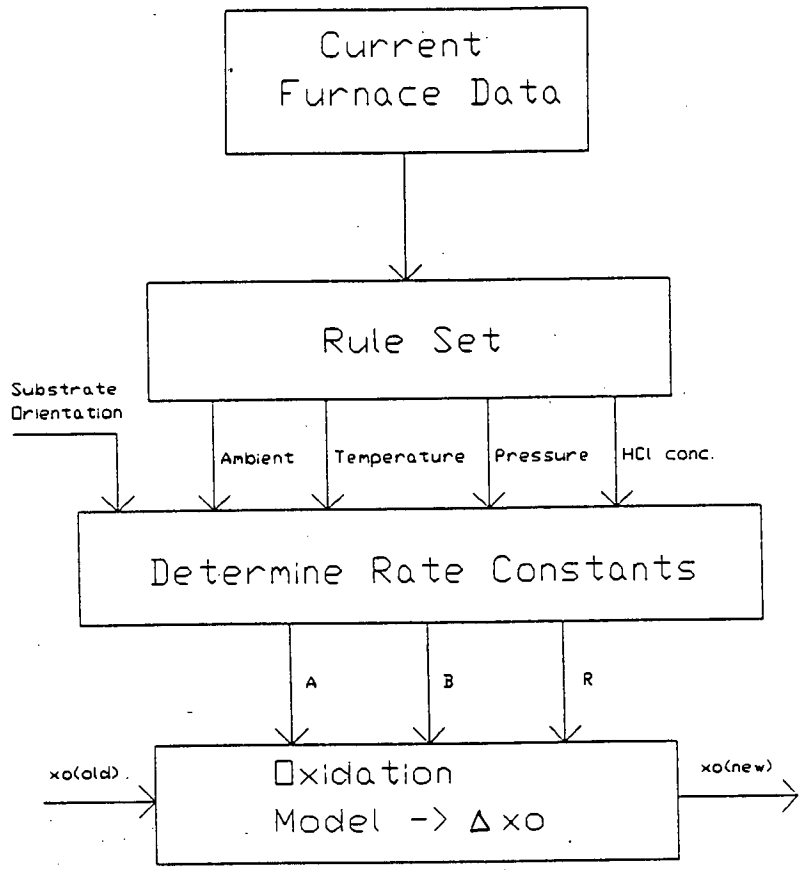


Fig. 7.1 Calculation of oxide thickness (x_o) from real-time furnace measurements.

$$B/A = C_L \exp(-E_{aL}/kT) . \gamma . H_L . P^n \qquad (7.2)$$

$$B = C_P \exp(-E_{aP}/kT) . H_P . P \qquad (7.3)$$

Here the values of C_L, E_{aL}, C_P and E_{aP} are for <100> silicon and the term γ is used to obtain the rate constants for <111> silicon ($\gamma_{<111>} = 1.68$, $\gamma_{<100>} = 1$ - see Section 3.2.2). The default values of C_L, E_{aL}, C_P and E_{aP} used in SUPREM III were given in Section 3.5. Because of the importance of these constants (they are essentially the basis of the whole model), the agreement of the SUPREM III values to oxide growth data from the EMF research furnace was examined. The experimental procedures and outcome is discussed in Section 7.3.

H_L and H_P are the rate enhancement factors which model the increase in oxidation rate upon addition of HCl to a dry oxidising ambient (Section 3.4). Since a quantitative physical model is lacking, either an empirical model or tabulated values for H_L and H_P must be used. SUPREM III uses values tabulated at various temperatures and HCl concentrations and this method has been adopted here (see Table 7.1).

The HCl concentration is calculated from the measured gas flow rates, F :

$$C_{HCl} = \frac{F_{HCl}}{F_{O2} + F_{N2} + F_{HCl}} \times 100\ \% \qquad (7.4)$$

The concentration must be rounded to the nearest 0.5% in order to use Table 7.1. Linear interpolation is then used for temperatures between the tabulated values.

	Temperature (°C)					
	900	1000	1100	900	1000	1100
HCl(%)	H_L			H_P		
0	1	1	1	1	1	1
0.5	1.48	1.32	1.32	1.05	1.34	1.21
1.0	1.89	1.49	1.61	1.11	1.41	1.34
1.5	2.04	1.64	1.89	1.16	1.48	1.38
2.0	2.04	1.74	1.96	1.21	1.54	1.42
2.5	2.04	1.74	2.09	1.27	1.61	1.46
3.0	2.04	1.74	2.21	1.32	1.68	1.49
3.5	2.04	1.74	2.21	1.37	1.74	1.53
4.0	2.04	1.74	2.21	1.42	1.81	1.57
4.5	2.04	1.74	2.21	1.48	1.88	1.60
5.0	2.04	1.74	2.21	1.53	1.95	1.64
5.5	2.04	1.74	2.21	1.58	2.02	1.68
6.0	2.04	1.74	2.21	1.64	2.08	1.71
6.5	2.04	1.74	2.21	1.69	2.15	1.75
7.0	2.04	1.74	2.21	1.74	2.22	1.79
7.5	2.04	1.74	2.21	1.80	2.28	1.82
8.0	2.04	1.74	2.21	1.85	2.35	1.86
8.5	2.04	1.74	2.21	1.90	2.42	1.90
9.0	2.04	1.74	2.21	1.95	2.49	1.94
9.5	2.04	1.74	2.21	2.01	2.56	1.97
10.0	2.04	1.74	2.21	2.06	2.62	2.01

TABLE 7.1 Values of H_L and H_P used in SUPREM III.

The flow rate of H_2 is not included in equation 7.4 since the presence of hydrogen indicates wet oxidation conditions. In this case $H_L = H_P = 1$. In wet oxidation, the dilution effect of added HCl gas is estimated in the calculation of the pressure, P.

P is also calculated from the measured gas flow rates. For dry oxidation, allowance must be made for the presence of added N_2 (which may be used as a dilutant).

$$P_{dry\ ox} = \frac{F_{O2}}{F_{O2} + F_{N2}} \times P_{Total} \qquad (7.5)$$

With pyrogenic oxidation the excess O_2, normally used to ensure complete combustion of the explosive H_2, must be taken into account. The partial pressure of H_2O is therefore given by

$$P_{wet\ ox} = \frac{2.F_{H2}}{F_{H2} + 2(F_{O2} + F_{N2} + F_{HCl})} \times P_{Total} \qquad (7.6)$$

Although the presence of both O_2 and H_2O strictly demands the use of a two-species model, wet oxidation is the dominant process and it is assumed that the Deal-Grove model still applies. The validity of the Deal-Grove model has been proven down to H_2O partial pressures of 0.25 [151].

The power dependence of P in the expression for B/A has been discussed in Section 3.2.3. Here the value of n is 1.0 for wet oxidation and 0.75 for dry oxidation.

The term R only applies to dry oxidation (for wet oxidation R=0). It is of the form

$$R = K \exp(-E_a/kT) \exp(-x_o/L) \qquad (7.7)$$

The first exponential term models the temperature behaviour of thin rate enhancement. SUPREM III uses equation 7.7 with the values of K, E_a, and L shown in Table 7.2. Because of the importance of these constants (they model oxide growth in the 'critical' thickness range up to 250Å), their applicability to oxide growth data from the EMF research furnace is also investigated in Section 7.3.

7.2.1 Simulation Dependence on Time Interval

The model represented by equation 7.1 calculates the increment in oxide thickness, $\triangle x_o$, for the time interval $\triangle t$. The calculated increment approaches the true value only when $\triangle t$ is sufficiently small. Too large a value of $\triangle t$ may therefore result in poor modelling. The value of $\triangle t$ can easily be reduced to a fraction of the parameter update period to avoid this problem. However, as $\triangle t$ is reduced, the number of time increments required for a given oxidation period increases. More calculations, and therefore a longer simulation time, result. For real-time simulation to be practicable, this calculation 'load' should be kept as small as possible. In the extreme case where $\triangle t$ is chosen so small that the calculation time actually exceeds the simulation period, then real-time simulation is of course no longer possible.

Silicon Orientation	K_o(Å/min)	E_a(eV)	L(Å)
<100>	7.48×10^{10}	2.38	69
<111>	6.58×10^{10}	2.33	78

TABLE 7.2 Values of K_o, E_a and L used in SUPREM III.

It is therefore important to know the sensitivity of the model to the value of Δt, in order that Δt can be optimally chosen. This was carried out by running computer simulations of oxide growth using equation 7.1 and several different values for Δt. Both dry and wet oxidations at 950°C were simulated, taking SUPREM III values for the model parameters. The simulations were repeated with Δt values of 0.5, 1.0, 5.0, 10.0 and 50.0 seconds. The simulated growth at fixed oxidation times of 20, 50 and 100 minutes for dry oxidation and 10, 25 and 50 minutes for wet oxidation are shown in Tables 7.3 and 7.4. For Δt values of up to 10 seconds, it can be seen that the model is very insensitive to the value used. Even at Δt=10 sec, the simulated growth is at most 1.2Å different from the more accurate result with Δt=0.5 sec (data for 20 min dry oxidation). This result means that Δt can be chosen as large as the parameter update period, providing this remains less than ten seconds. Moreover, because of this insensitivity, Δt need not be held at a fixed value throughout the simulation period.

7.3 Determination of Rate Constants

Values for C_L, E_{aL}, C_P and E_{aP} (equations 7.2 and 7.3) were determined from a total of 39 dry and 32 wet oxidations carried out in the EMF research furnace. Temperatures of 900, 950, 975, 1000, and 1025 °C were used. Oxidation times were selected to give oxide thicknesses up to 2500Å under dry oxidation conditions, and up to 4000Å for wet oxidation.

To ensure temperature accuracy, the furnace was first profiled using the standard procedures described in Section 4.1.4. Ten profile temperatures in the range

	Oxidation Time (min)		
△t(sec)	20	50	100
50	164.05	278.68	429.15
10	162.56	276.69	427.78
5	161.49	276.72	427.08
1	161.35	276.31	426.90
0.5	161.28	276.23	426.90

TABLE 7.3 Simulated oxide thickness (in Å) tabulated against the simulation time increment, △t, and total oxidation time. Dry oxidation at 950°C.

	Oxidation Time (min)		
△t(sec)	10	25	50
50	443.63	1025.09	1852.82
10	441.94	1021.92	1848.41
5	441.73	1021.58	1847.86
1	441.56	1021.21	1847.43
0.5	441.54	1021.17	1847.37

TABLE 7.4 Simulated oxide thickness (in Å) tabulated against the simulation time increment, △t, and total oxidation time. Wet oxidation at 950°C.

700°C to 1050°C were used. All oxidation temperatures to be studied were included in the profile data so that the exact set points for the control (spike) thermocouples were known, avoiding the need for interpolation at these temperatures. During oxidation, wafers were placed at the site of the centre thermocouple, thus minimising the effect of the linear temperature variation described in Section 4.2.1.

For all oxidations, the starting material was <100> 10-14ohm.cm 3" silicon wafers. A standard RCA pre-oxidation clean with a final dip in dilute HF to ensure an oxide-free surface was used. In order to obtain accurate oxidation times, the wafers were introduced into the furnace under an N_2 ambient, and ample time was allowed for the furnace temperature to settle before introducing the oxidising gases. The complete furnace procedures are described in Tables 7.5 and 7.6. The oxide thickness was measured at five sites on five wafers selected from each run. Because of the large number of thickness measurements involved, the Nanospec was again the preferred measurement instrument. The mean thickness and standard deviation over the 25 points from each run are presented in Tables 7.7 and 7.8. Although slightly erratic, the plots of standard deviation against mean thickness (Figs. 7.2 and 7.3) provide further confirmation of the observation in Chapter 6 that the magnitude of thickness variation does not scale linearly with oxide thickness.

The linear and parabolic rate constants at each temperature were evaluated using the method outlined in Section 3.2.1. The Deal-Grove equation is expressed in the form :

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	[ii]	-	'Standby' Position
1	5.0	0	0	0	250		15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700		15	Paddle replaced Temperature allowed to settle
3	0	2.5	3.0	0	1700		[iii]	Wet Oxidation
4	5.0	0	0	0	1700		10	Purge tube and anneal
5	5.0	0	0	0	250		15	Paddle out of furnace to unload wafers

i) Paddle position is a number (in mm) used by the process controller. At 1700mm the paddle is fully inserted into the furnace tube, and the end-cap is flush against the end of the tube. At 250mm the furnace paddle is fully withdrawn for loading/unloading wafers.

ii) The required temperature is set in the standby position and remains at this value throughout the rest of the process.

iii) Oxidation time is programmed as required for each run.

TABLE 7.5 Wet oxidation process for evaluation of rate constants.

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	ii)	-	'Standby' Position
1	5.0	0	0	0	250		15	Boat out of furnace to load wafers
2	5.0	0	0	0	1700		15	Boat replaced Temperature allowed to settle
3	0	5.0	0	0	1700		iii)	Dry Oxidation
4	5.0	0	0	0	1700		10	Purge oxygen from tube and anneal
5	5.0	0	0	0	250		15	Boat out of furnace to unload wafers

i) ii) iii) See Table 7.5.

TABLE 7.6 Dry oxidation process for calculation of rate constants.

Temperature (°C)	Time (Min)	Mean Thickness (Å)	s.d. (Å)	CV(%)
900	180	356.2	6.3	1.8
900	270	483.8	11.4	2.4
900	390	634.2	9.1	1.4
900	500	759.0	14.0	1.8
900	600	871.0	11.8	1.4
900	700	935.9	17.7	1.9
950	9	115.1	5.9	5.1
950	30	199.7	7.7	3.9
950	90	403.1	9.2	2.3
950	120	503.0	7.7	1.5
950	180	659.7	10.9	1.6
950	240	803.6	7.8	1.0
950	300	907.5	12.2	1.3
950	390	1084.3	9.7	0.9
950	510	1327.4	9.6	0.7
950	720	1655.3	8.5	0.5
975	90	511.6	8.8	1.7
975	140	705.8	5.1	0.7
975	160	779.6	13.3	1.7
975	200	915.6	8.9	1.0
975	330	1231.5	8.2	0.7
975	480	1539.5	21.9	1.4
975	720	2050.7	21.6	1.1
1000	30	309.4	5.8	1.9
1000	60	516.5	8.8	1.7
1000	80	615.9	9.1	1.5
1000	100	708.0	10.6	1.5
1000	120	830.5	15.7	1.9
1000	210	1176.7	10.8	0.9
1000	390	1770.3	20.7	1.2
1000	540	2138.6	22.6	1.1
1025	30	402.0	6.1	1.5
1025	45	530.0	8.2	1.6
1025	60	643.5	8.1	1.2
1025	75	751.4	10.3	1.4
1025	90	872.4	5.4	0.6
1025	120	1023.4	7.9	0.8
1025	180	1307.1	23.3	1.8
1025	300	1852.4	16.1	0.9

TABLE 7.7 Summary of dry oxidation data.

Temperature (°C)	Time (Min)	Mean Thickness (Å)	s.d. (Å)	CV(%)
900	20	457.5	12.0	2.6
900	40	816.3	47.2	5.8
900	80	1377.2	99.3	7.2
900	120	1995.8	14.3	0.7
900	160	2541.6	17.2	0.7
900	220	3155.9	122.6	3.9
950	9	449.4	5.9	1.3
950	22	921.3	7.7	0.8
950	32	1212.5	13.9	1.1
950	42	1540.0	53.1	3.4
950	62	2031.4	13.2	0.6
950	82	2535.4	39.0	1.5
950	105	3051.2	88.5	2.9
950	123	3462.9	74.7	2.2
975	10	601.3	35.3	5.9
975	15	851.4	48.1	5.6
975	20	1038.0	55.7	5.4
975	30	1418.7	92.3	6.5
975	40	1824.8	74.4	4.1
975	60	2383.8	142.3	6.0
1000	5	467.5	18.2	3.9
1000	10	776.1	59.8	7.7
1000	20	1384.0	62.5	4.5
1000	30	1846.9	58.6	3.2
1000	40	2258.5	77.9	3.4
1000	60	3071.4	98.3	3.2
1000	80	3711.2	106.5	2.9
1025	5	592.2	20.9	3.5
1025	10	1020.8	27.9	2.7
1025	15	1396.8	46.7	3.3
1025	20	1723.9	48.1	2.8
1025	25	1973.5	65.6	3.3

TABLE 7.8 Summary of wet oxidation data.

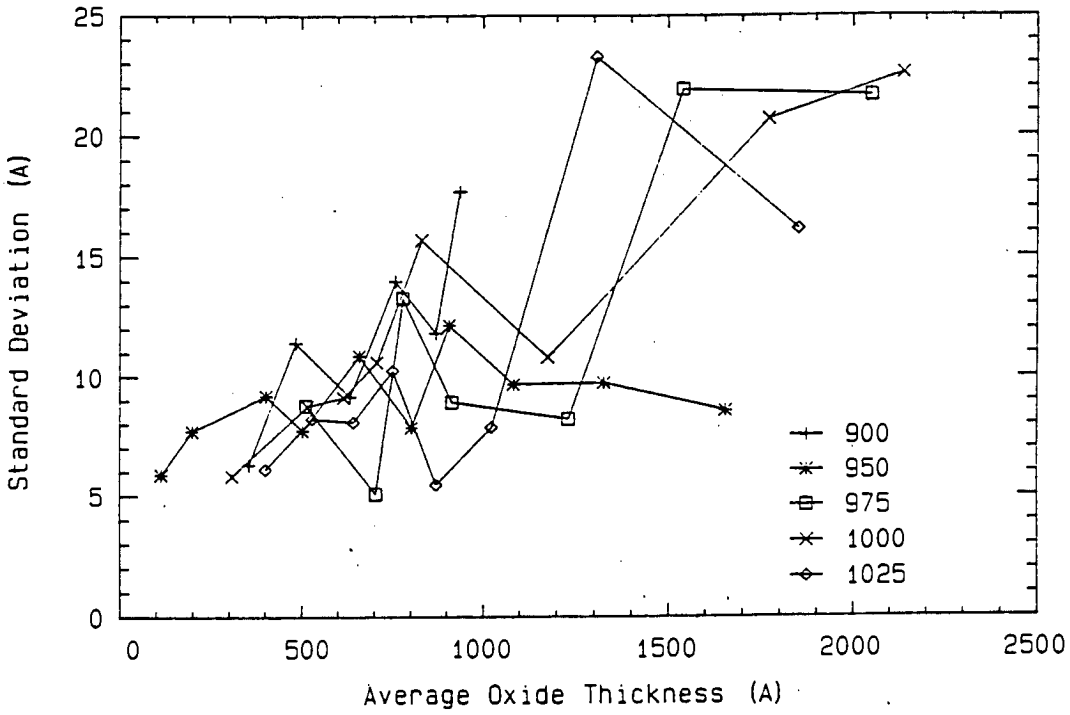


Fig. 7.2 Standard deviation against oxide thickness (dry oxidation).

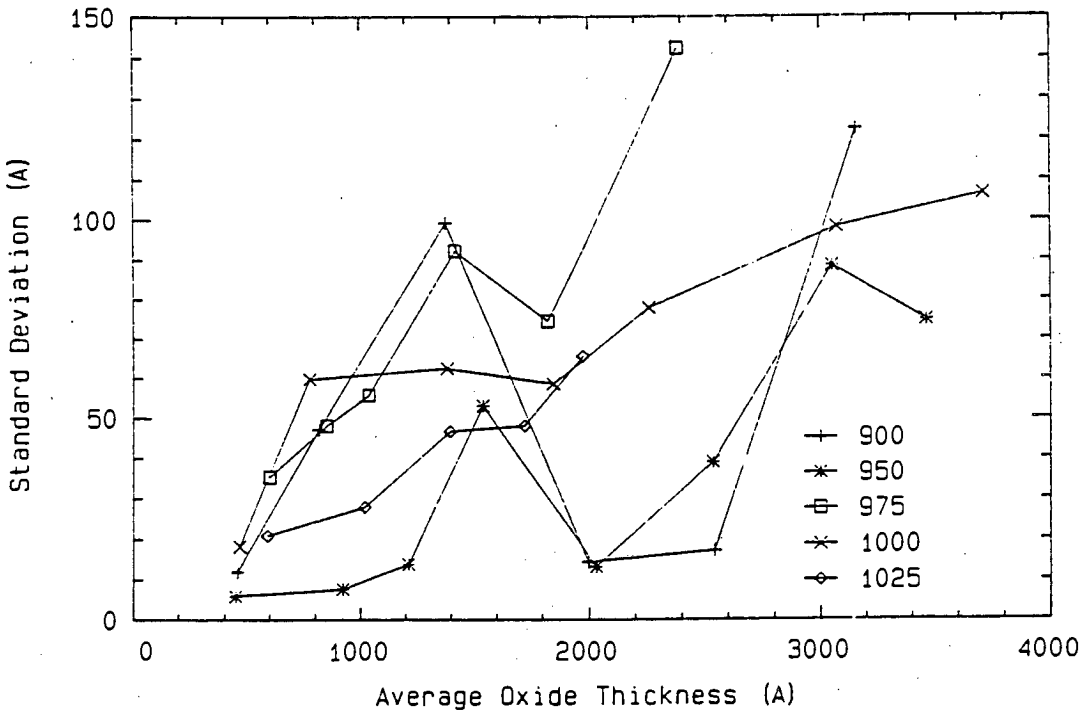


Fig. 7.3 Standard deviation against oxide thickness (wet oxidation).

$$x_o = \frac{(t + \tau).B}{x_o} - A \qquad (7.8)$$

from which it is clear that a plot of x_o against $(t + \tau)/x_o$ yields a straight line of slope B and vertical axis intercept equal to -A.

The parameter τ (the time correction factor equivalent to the presence of initial oxide on the silicon) was assumed to be zero for wet oxidation and for dry oxidation was calculated by extrapolation to the time axis of the linear portion of the x_o against t data. This is shown for the data at 950°C in Fig. 7.4. Values of τ were obtained at all temperatures by least squares analysis and are summarised in Table 7.9. Also shown are values of x_i, the y-axis intercept of the extrapolated line, which represents the oxide thickness required before the onset of linear-parabolic oxide growth kinetics. These data yield an average x_i of 178Å (s.d.= 13.9) which is in agreement with values quoted in the literature (these are generally in the range 100-350Å [6,89,104,152,153]).

The x_o against $(t + \tau)/x_o$ plots for the 950°C data are shown in Figs. 7.5 and 7.6. For dry oxidation, data outside the linear-parabolic region (i.e. in the initial oxidation regime, x_o<500Å) have been omitted since these points would not be expected to conform to the predicted linear relationship.

Values for the slope and intercept in each plot were obtained by least-squares analysis. The resultant rate constants together with similarly calculated values at other temperatures are presented in Table 7.10. For wet oxidation the constants have been corrected by a

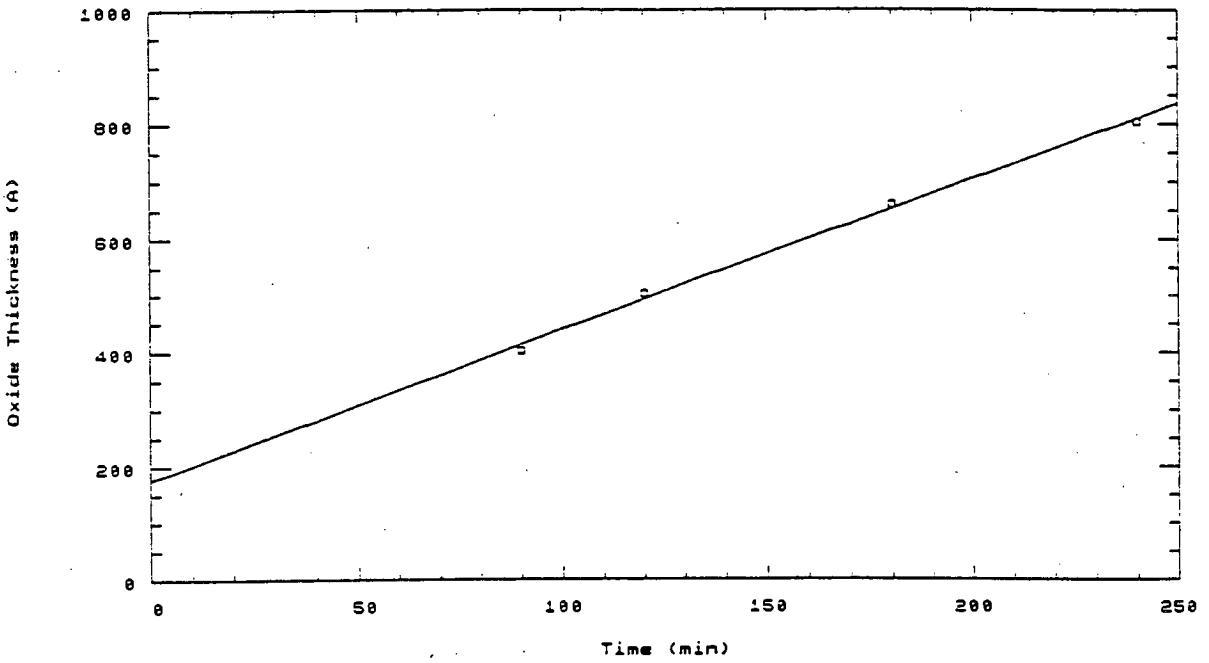


Fig. 7.4 Oxide thickness against time for dry oxidation at 950°C.

Temperature (°C)	τ (Min)	x_i(Å)
900	146.9	171.9
950	66.8	176.3
975	43.4	166.7
1000	39.2	202.4
1025	22.6	174.9
	Average	178.4

TABLE 7.9 τ and x_i extrapolated from oxide growth data for dry oxidation.

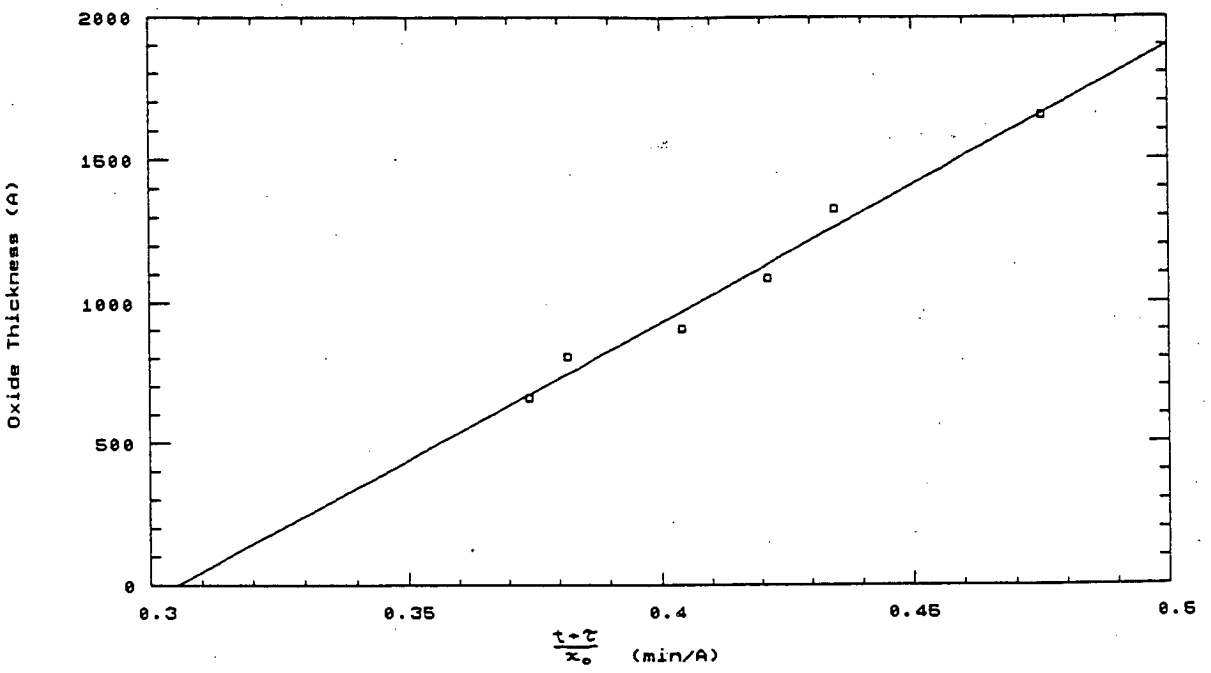


Fig. 7.5 Oxide thickness (x_o) against $(t+\tau)/x_o$ for dry oxidation at 950°C. The slope of the line is B and the intercept is -A.

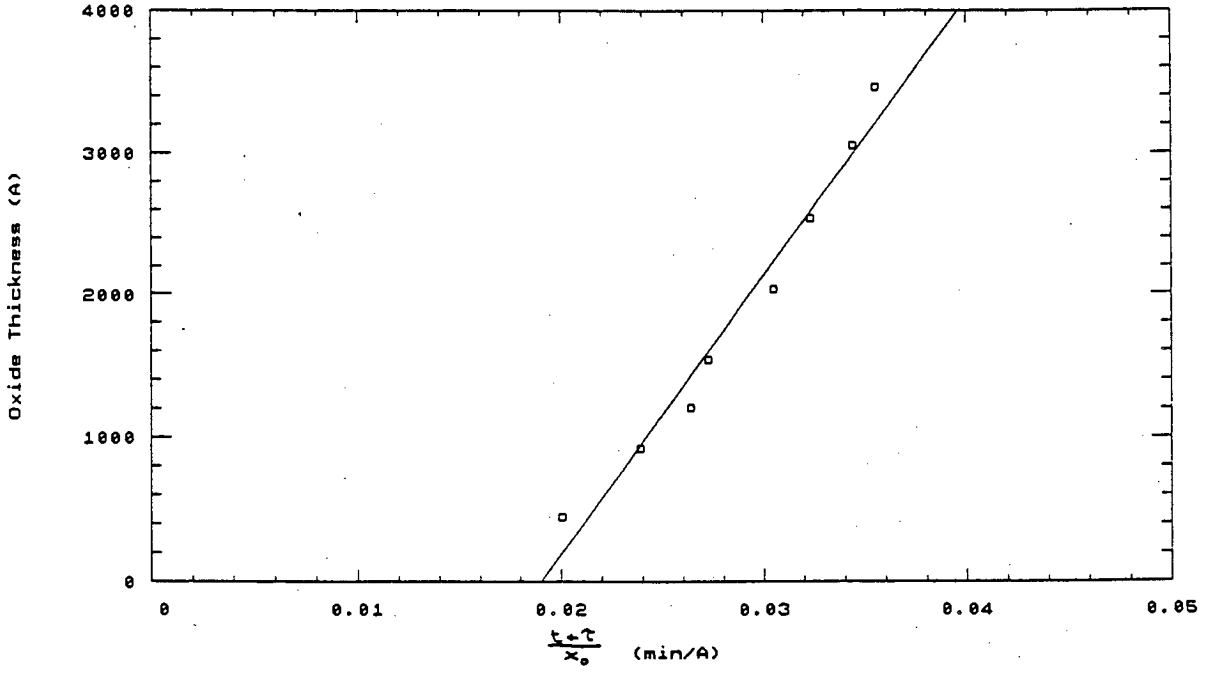


Fig. 7.6 Oxide thickness (x_o) against $(t+\tau)/x_o$ for wet oxidation at 950°C.

	Dry Oxidation		Wet Oxidation	
Temperature (°C)	B (10^3 Å^2/min)	B/A (Å/min)	B (10^3 Å^2/min)	B/A (Å/min)
900	3.83	1.56	106.32	24.36
950	9.75	3.28	194.49	52.57
975	11.59	5.02	210.52	72.12
1000	19.74	6.33	308.59	101.04
1025	25.01	9.94	337.36	149.85

TABLE 7.10 Calculated rate constants for dry and wet oxidation.

factor of 4/3 to correspond to an H_2O partial pressure of 1.

Arrhenius plots of the rate constants are shown in Figs. 7.7 and 7.8. The location of the data points from wet oxidation, in both the B and B/A plots, makes calculation of single activation energies for B and B/A more appropriate than the temperature breakpoint analysis on which the SUPREM III constants are based. This is therefore further evidence against the argument of curvature in the Arrhenius plots (see Section 3.2.1) although, as already mentioned, not too much should be interpreted from this observation due to the general insensitivity of the Deal-Grove model parameters.

C_L, E_{aL}, C_P and E_{aP} were calculated by a least-squares fit to the Arrhenius plots (Table 7.11). Direct comparison with SUPREM III values is possible for dry oxidation and gives a generally poor correlation. The original B and B/A rate constants for wet oxidation were also found to have a poor correlation with corresponding SUPREM III values. As for the above observations, the differences are caused by the influence of experimental error and model insensitivity. It is not possible to say which set of parameters most closely approach the true values. However, the preferable parameter set can be determined by examining how well each fits the oxidation data.

To provide a measure of goodness-of-fit, the fitting error, S, was calculated. S is defined as

$$S = \frac{\sum_i \{|x_{pi} - x_i|/x_{pi}\}}{N} \qquad (7.9)$$

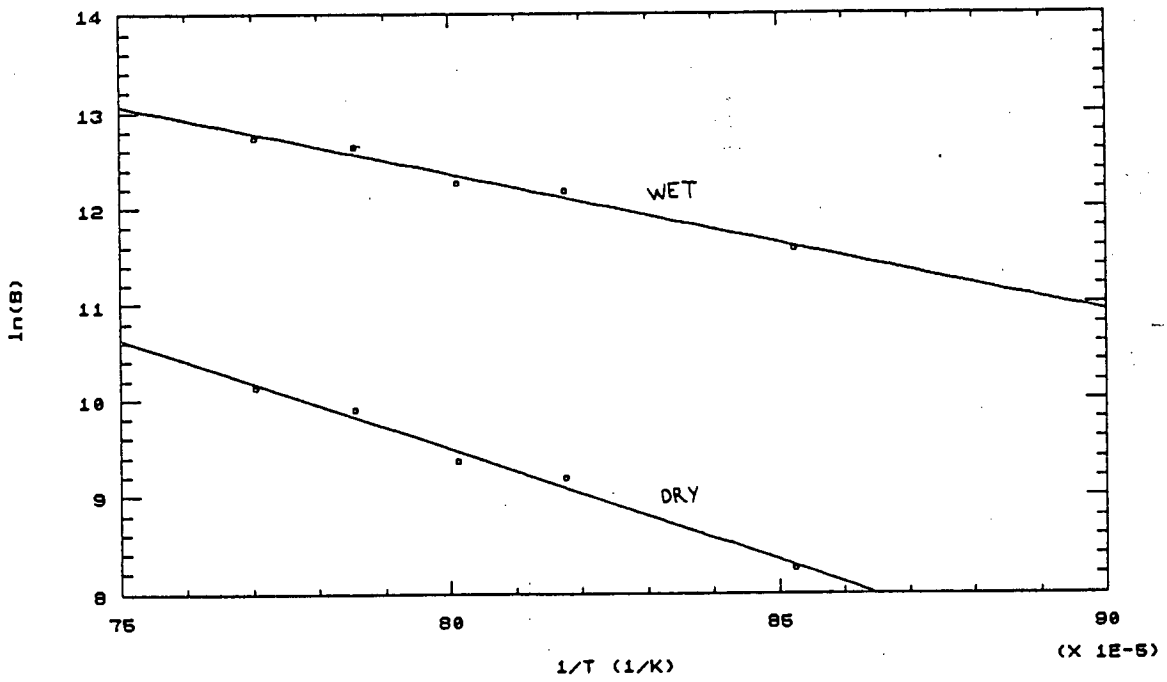


Fig. 7.7 Parabolic rate constant B against 1/T for dry and wet oxidation.

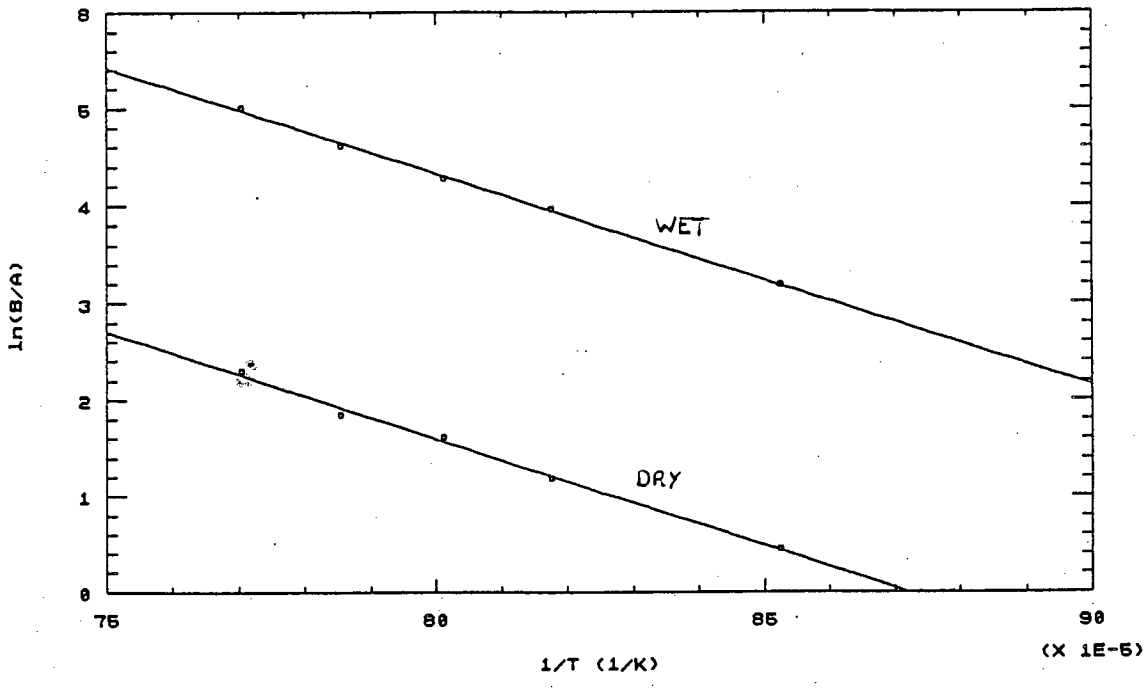


Fig. 7.8 Linear rate constant B/A against 1/T for dry and wet oxidation.

Ambient	C_L(Å/min)	E_{aL}(eV)	C_p(Å^2/min)	E_{aP}(eV)
DRY	2.51×10^8	1.91	1.21×10^{12}	1.98
WET	2.83×10^9	1.88	2.15×10^{10}	1.23

TABLE 7.11 Calculated pre-exponential constants and activation energies for wet and dry oxidation.

where x_i is the measured oxide thickness at a particular time and temperature and x_{pi} is the predicted oxide thickness for that point, determined from the model and the parameter set used. The relative errors (the bracketed term) are averaged over all i (i.e. all temperatures and times) so that an overall fitting error for the model is obtained.

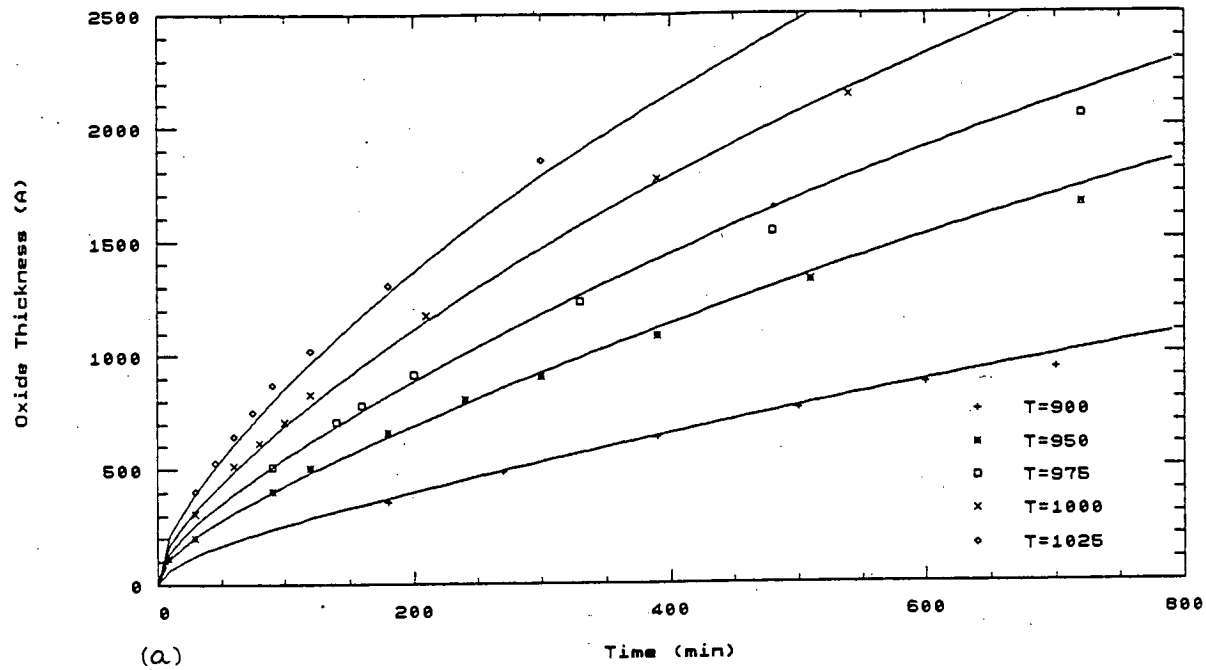
Calculated fitting errors for wet and dry oxidation, using the model given by equation 7.1 and the aforementioned parameters sets, are shown in Table 7.12. The simulations using the SUPREM III constants are shown in Fig. 7.9. These yield fitting errors of 3.2% for dry oxidation and 7.3% for wet oxidation. Simulation with the calculated constants (Fig. 7.10) gives values of 5.3% and 2.4% for dry and wet oxidation respectively. Despite the two very different parameter sets, fitting errors are of comparable order, demonstrating the model insensitivity discussed in Section 3.2.1.

The expected reduction in S from using the directly calculated parameter set is not obtained with dry oxidation because the rate enhancement, R, is modelled using the SUPREM III values of the constants K, E_a and L (equation 7.7). These have been calculated based on different B and B/A values [89] and they do not combine well with the values calculated here. This is illustrated by Fig. 7.11 which shows the thin oxide data at 950°C. Simulation with the SUPREM III values of B and B/A clearly lies closer to the data than simulation with the calculated values.

To avoid this problem, simulation of dry oxidation was repeated with the calculated parameter set, using the original form of model proposed by Deal and Grove (equation 3.11) and taking the value of τ from Table 7.9. This is shown in Fig. 7.12. With this form of the

Ambient	SUPREM III parameters	Calculated parameters
DRY	3.2%	5.3%
WET	7.3%	2.4%

TABLE 7.12 Fitting errors of the model to the oxide growth data of Tables 7.5 and 7.6. Simulations were run with either the SUPREM III or calculated values for C_L, E_{aL}, C_P and E_{aP}.



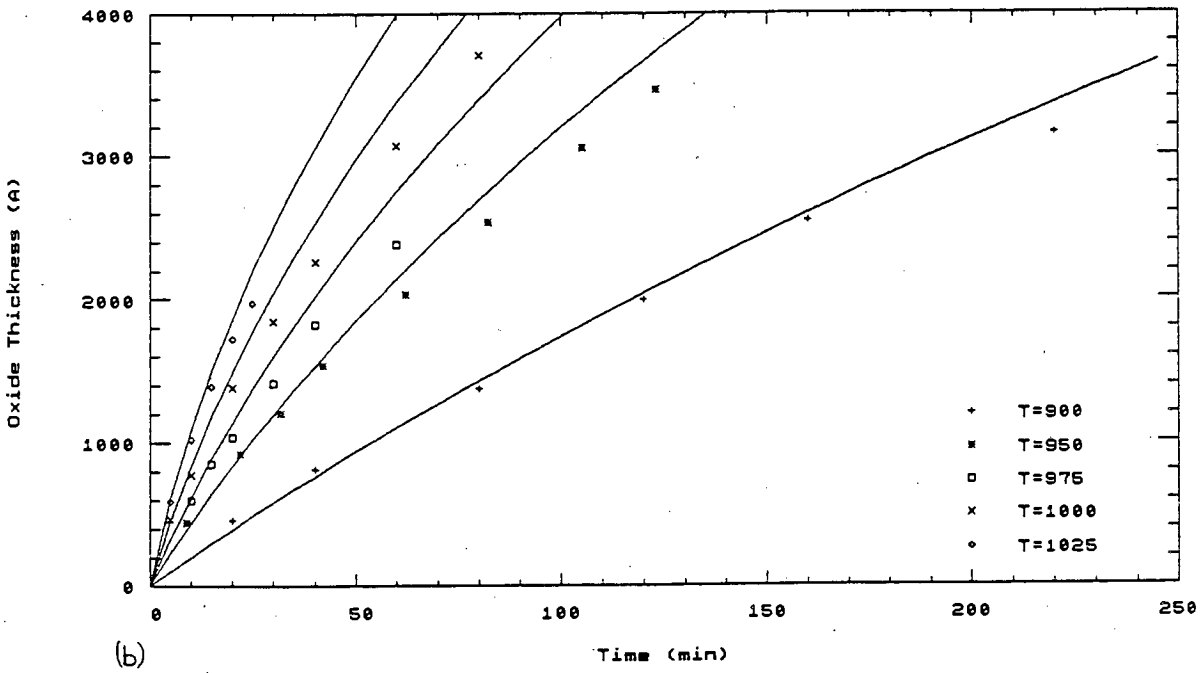
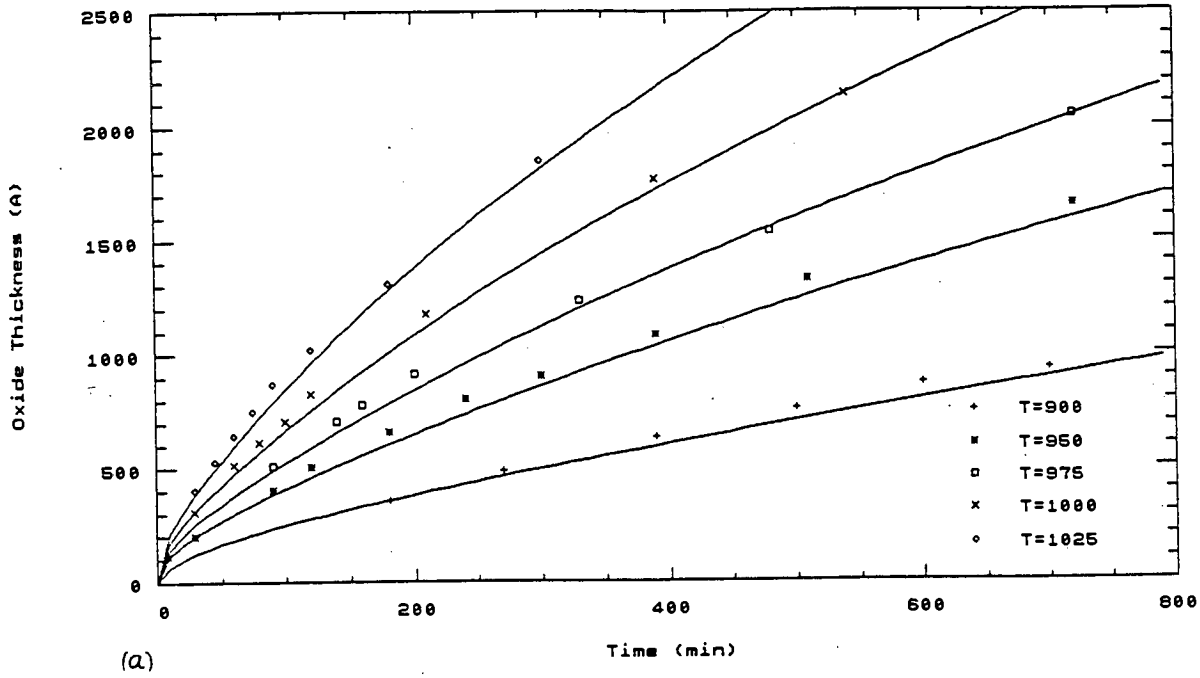


Fig. 7.9 Comparison between oxide growth data (points) and simulation (solid lines) using the model and constants from SUPREM III. Data shown is for (a) dry oxidation (b) wet oxidation at temperatures 900, 950, 975, 1000, 1025°C.



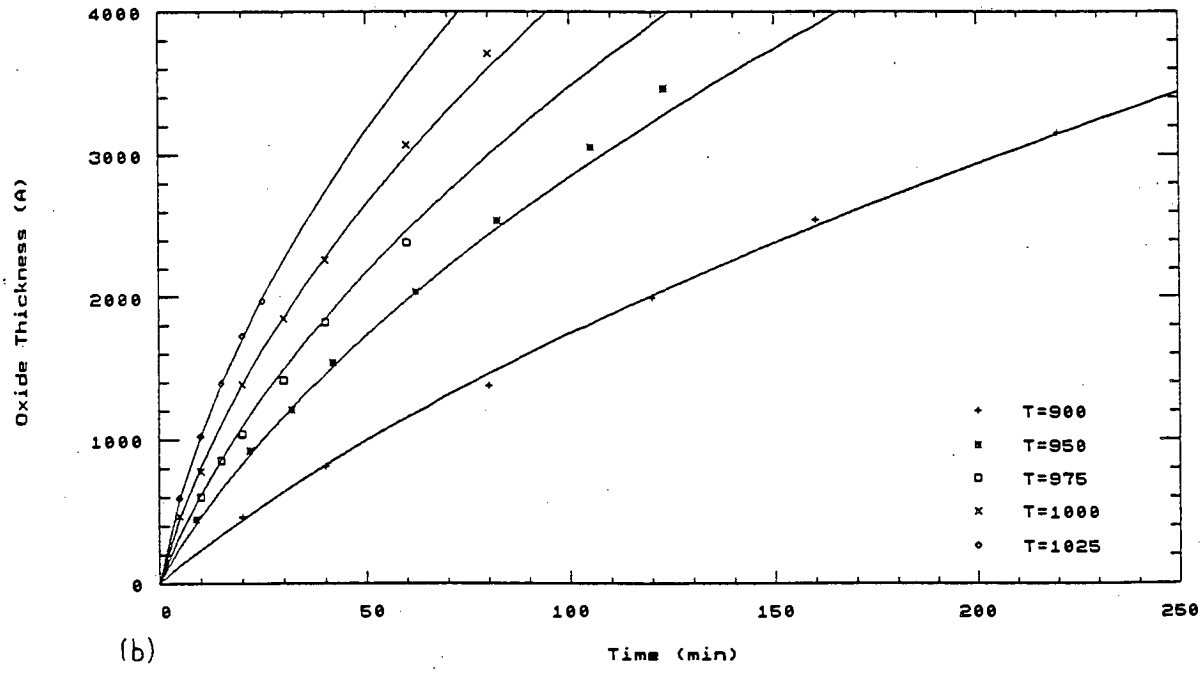


Fig. 7.10 Comparison between oxide growth data (points) and simulation (solid lines) using the model from SUPREM III, but with values of C_L, E_{aL}, C_P and E_{aP} calculated from the data set. Data shown is for (a) dry oxidation (b) wet oxidation at temperatures 900, 950, 975, 1000, 1025°C.

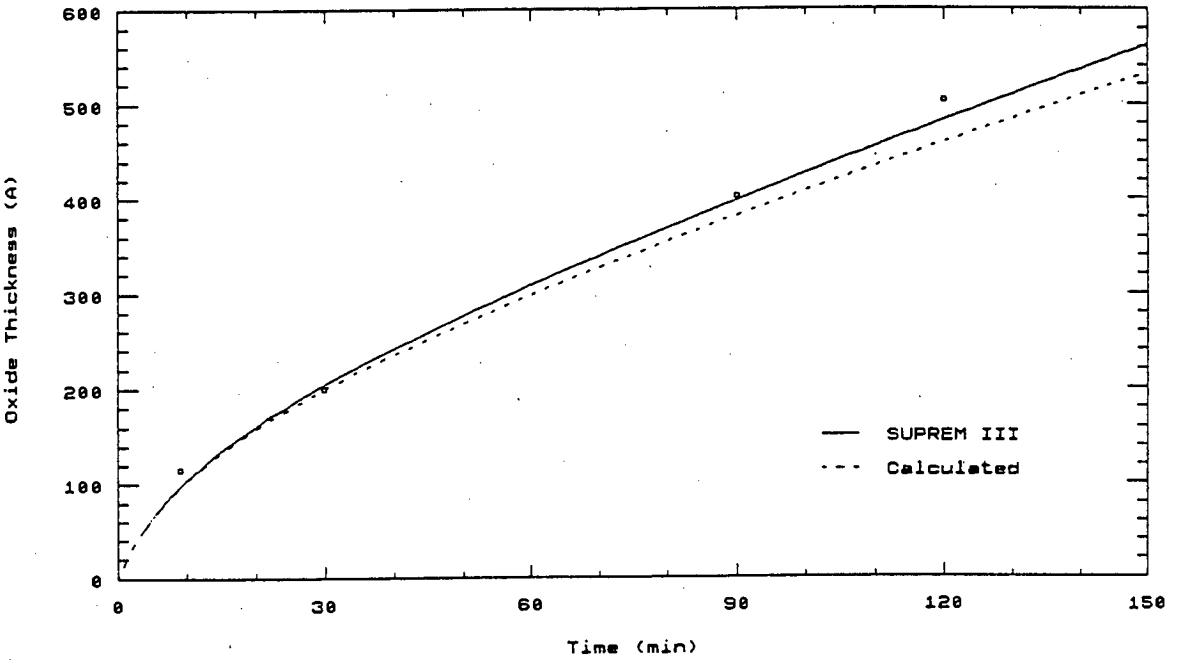


Fig. 7.11 Thin oxide growth data (points) for dry oxidation at 950°C and simulated growth curves using SUPREM III (solid line) and calculated (dashed line) values for C_L, E_{aL}, C_P and E_{aP}.

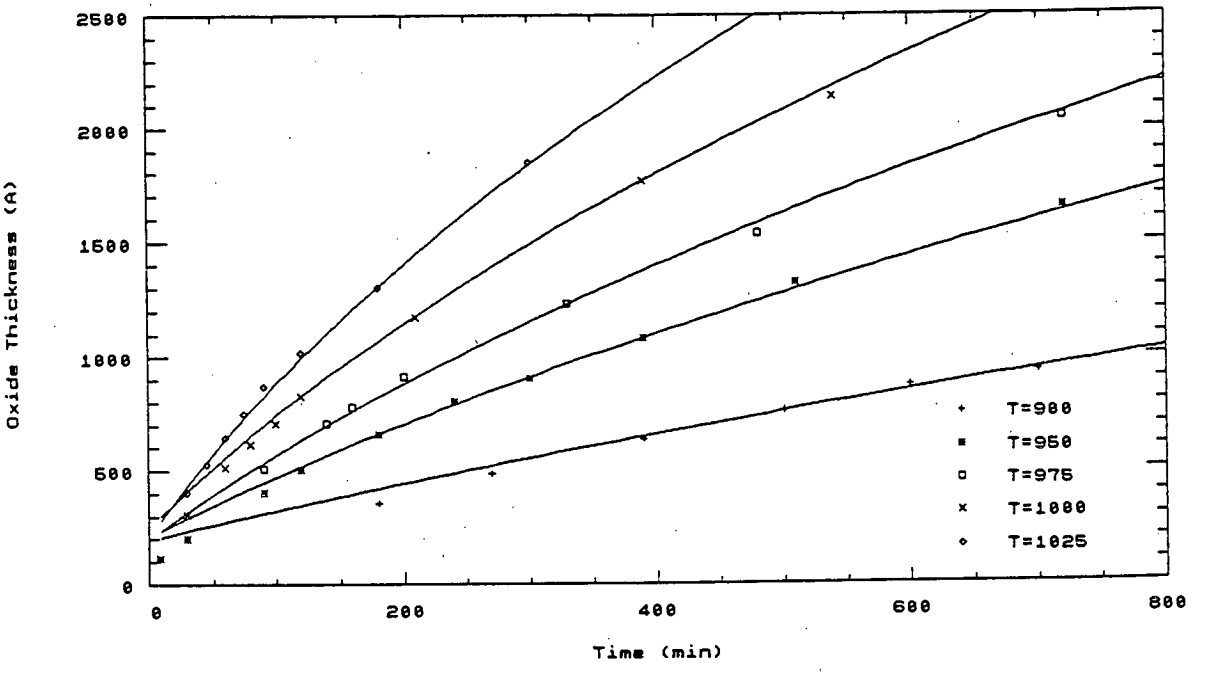


Fig. 7.12 Oxide growth data (points) and simulated growth curves (solid lines) using the original Deal-Grove model with calculated values for the model parameters C_L, E_{aL}, C_P, E_{aP} and τ. Data shown is for dry oxidation at temperatures 900, 950, 975, 1000, 1025°C.

model, the fitting error is reduced to 2.1% (excluding oxides below 500Å, where the model is not applicable) confirming that the problem lies in simulation of the thin oxide regime.

The SUPREM III constants yield a good fit to the data under dry conditions, even at thin oxides. It is therefore considered unnecessary at this point to redetermine by experiment 'optimal' values for K, E_a, and L in equation 7.7 for use with the calculated parameter set. The SUPREM III constants are implemented in the simulation-control system for the modelling of dry oxidation. In the case of wet oxidation, the SUPREM III constants give a relatively poor overall fit, and the calculated parameter set is favoured. The corresponding fitting errors of 3.2% and 2.4% are considered low enough to justify use of the model described in Section 7.2.

7.4 Interface between Model and Furnace

The interface between model and furnace was achieved by running the simulation as a BASIC program on an Apricot Personal Computer. The necessary inputs to the program are flow rates of all gases, temperature, and boat position (to establish whether the wafers are actually in the furnace or not). Knowledge of these variables is not a problem since they are already measured and collected by the furnace control system itself. They are available through the built-in ability of the furnace digital controller to communicate with an upstream remote computer (Section 4.1.5). Figure 7.13 shows schematically how data is passed to the simulation. A paddle thermocouple is positioned permanently at the site of the wafer boat to allow a more accurate simulation of the conditions seen by the wafers.

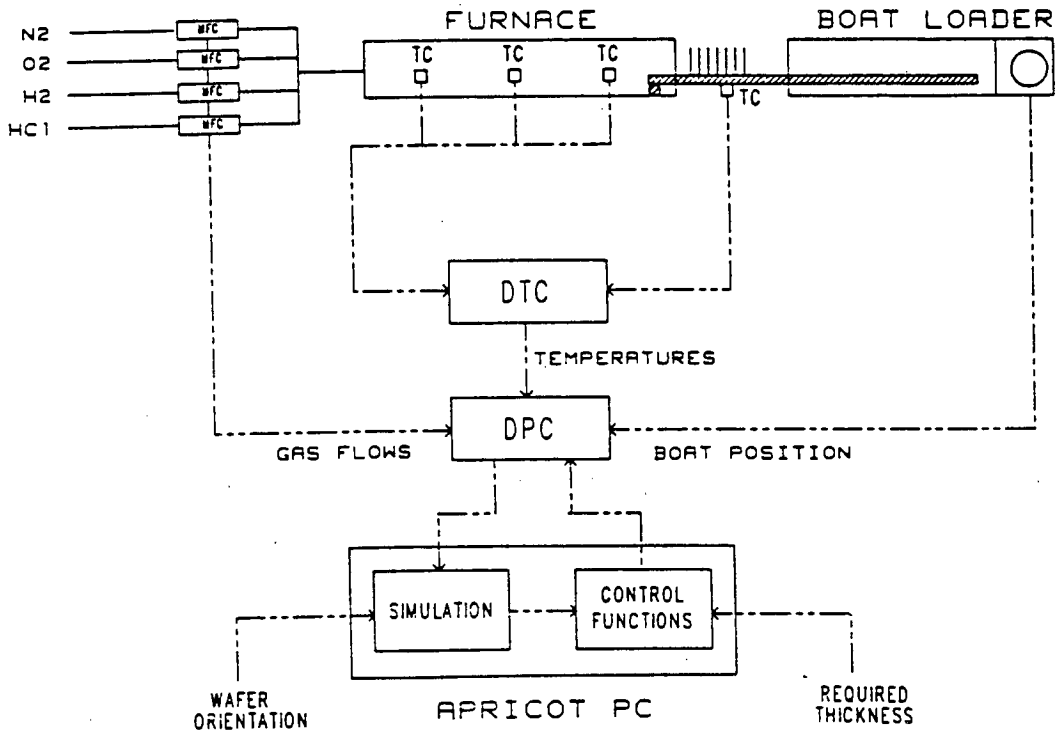


Fig. 7.13 Data collection for computer simulation of the oxide growth process.

Communication between the computer and the DPC uses the SECS protocol [4]. SECS is a communication standard for interconnected IC production equipment which utilises a standard RS232C serial interface. Messages are transferred over this interface using a defined handshaking procedure. This is described in detail in Appendix B. The protocol was implemented on the Apricot PC using the computer's normal serial interface port for the transmission and receipt of messages. The form of RS232C built into the DPC incorporates optical isolation, and a conversion circuit was required to enable coupling to the Apricot's serial interface. The circuit shown in Fig. 7.14 was built for this purpose. Its placement between computer and DPC has no effect on the operation of the SECS protocol.

A two-step process is needed to acquire furnace information over the communications link. Firstly, the computer must send a coded message to the DPC with a request for information. After the DPC has processed the request, it replies with a message containing the appropriate data which is collected and processed by the computer. Further details of this process are presented in Appendix B. The following are the main sources of delay in the process : (i) lag in response of the DPC after it is prompted for information (the DPC must ask the temperature controller (DTC) for relevant temperature-related information before responding to the computer) ; (ii) the transfer time of the relatively long message (197 bytes) between DPC and computer ; (iii) delays from the handshaking procedures ; (iv) processing delays in the computer. A minimum recycling time of around 3-4 seconds can be achieved. This is more than adequate for the implementation of real-time simulation, as previously discussed.

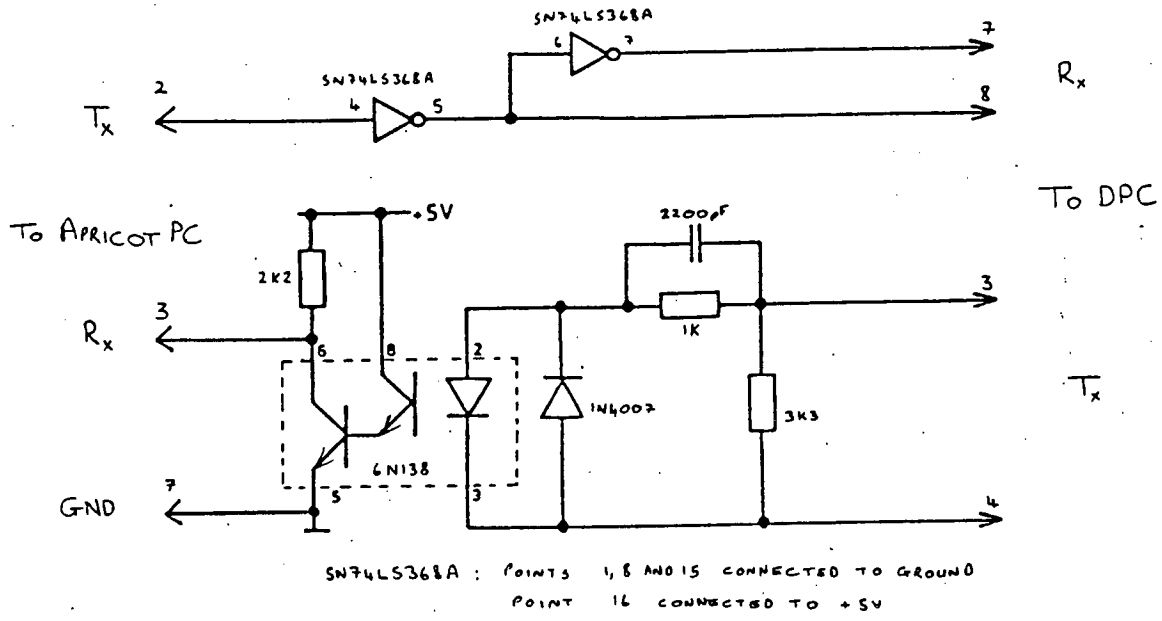


Fig. 7.14 Conversion circuit used to interface the Tempress DPC furnace controller to the Apricot PC for data collection.

A second interface between computer and furnace was built to realise active control of furnace operations. In contrast to the above interface - which is essentially furnace-system independent (most modern controllers implement the SECS protocol for remote computer communication) - the form of this control-interface was determined primarily by the configuration of the Tempress system. Here a recipe cannot be directly changed whilst it is being run. However, alternative options are available. The current recipe can be made to jump to a pre-specified 'Abort' recipe when a certain error condition is recognised by the DPC. Likewise the recipe can be programmed to remain at a particular step until the specified error condition exists, then advance to the next step.

These options provide a simple but effective means of steering furnace procedures. The triggering of available digital inputs to the DPC can be used to simulate an error condition. Oxidation recipes can then be programmed to stick at an oxidising step until receipt of a trigger, then advance to an appropriate Abort recipe or another step of the main recipe. Furnace conditions that would ultimately bring the oxidation to a controlled end must be pre-programmed into these jump-sequences.

Although several inputs and Abort recipes are possible, allowing the capability of complex steering functions, a relatively straightforward control philosophy was adopted for initial investigation. Oxidation was terminated by a jump to an anneal step in the same furnace recipe. In this step an N_2-only environment was defined. Only one digital input had to be triggered to initiate this jump.

The trigger was initiated from the Apricot PC when the simulated oxide thickness approached the control

value. An IEEE-488 standard interface was added to the computer for this purpose. The IEEE-488 bus was then used to send control characters to a digital I/O module [154]. An output of this module was linked to a DPC interconnection board. Although indirect, with this interface the furnace conditions are effectively under software control from the remote computer.

The delay between initiation of the trigger at the computer and the time when sufficient N_2 has reached the wafers to effectively stop the oxidation must be considered. This delay is comprised of (i) response time of the DPC, (ii) response time of the mass flow controllers (see Section 4.2.4) and (iii) the time taken by the N_2, introduced at the source end of the furnace tube, to expand and diffuse to the position of the wafers. Delays (i) and (ii) are more or less fixed time intervals and can be easily determined. However the N_2 diffusion delay cannot be known exactly. It is flow-dependent and may be roughly estimated as in Section 4.2.4.

In order to know when the trigger must be sent, the oxide grown during the total delay needs to be known. This of course depends on the oxidising conditions and the oxide thickness already grown. The problem is solved by predicting future growth, through simulation using the most current furnace data. Here it is assumed that the current furnace conditions will remain valid during the whole of the delay. Effectively, the control-point is adjusted to account for the predicted behaviour of the furnace system.

The delay period represents a fitting parameter for the furnace model and may need appropriate adjustment. A total delay of 61 seconds is estimated for a nitrogen flow rate of 5.0 SLM. This comprises a 7 second DPC

response delay, a 14 second MFC delay and a 40 second delay for the N_2 to diffuse and reach the position of the wafers. In processes involving an initial period in N_2, it can reasonably be assumed that the gas diffusion delays will be effectively cancelled by a corresponding delay at the start of oxidation (i.e. the time taken, after the change to an oxidising ambient, for oxidant gases to reach the wafers) providing the respective gas flows and temperatures are approximately equal. In such cases, the N_2 diffusion delay can be neglected and a total delay period of 21 seconds is more appropriate.

7.5 FACS Software Description

A new software system was developed which transforms the computer/furnace combination into an independent, actively-controlled processing module. The program encompasses all the aforementioned requirements for modelling and interfacing, and provides the capability of real-time control of furnace operations. Whilst the system was developed with this specific application in mind, additional features were included to make full use of the data-handling capability of the computer. The program has been given the acronym FACS (Furnace Active Control System).

With FACS, a regularly updated overview of the furnace system is permanently displayed on the computer screen. This provides at a glance all the important information about current furnace status, a feature not available at the local furnace level (i.e. DPC and DTC). Set point and actual values of temperature, gas flows and paddle conditions are all displayed, as are the recipe step and process times.

FACS also enables selected furnace data to be permanently stored on disk. ASCII files containing temperatures, gas flows or (simulated) oxide thickness, all as a function of time, are created. The stored data becomes part of a distributed process database, and can be used for later analysis or input into the main process database.

FACS operates from a baseline mode in which data is simply collected from the DPC and displayed on the screen. Data storage, simulation and active control are optional additions to this baseline procedure. A delay option allows execution of the additional routines to be automatically delayed until commencement of a specified step in the furnace recipe. This can be used to prevent unnecessary modelling and data storage during the furnace 'standby' period.

All options are selected individually from function keys on the computer keyboard. Function keys are also used for a reset function. Reset includes setting the oxide growth simulation to a starting thickness of 0Å, and initiation of a system timer. The timer is used to determine an elapsed time for data storage and simulation.

The main loop in the FACS program is shown in Fig. 7.15. In the baseline mode a cycle time of around 3 seconds is obtained, which increases to about 5 seconds when data storage, simulation and active control are added. Occasional errors in the data communication between the DPC and computer - which are automatically corrected through the 'Retry' philosophy of the SECS protocol - can increase the cycle time to 8 seconds. Even with no communication problems the response time of the DPC varies. The value of Δt in equation 7.1 is therefore not constant and must be calculated from the

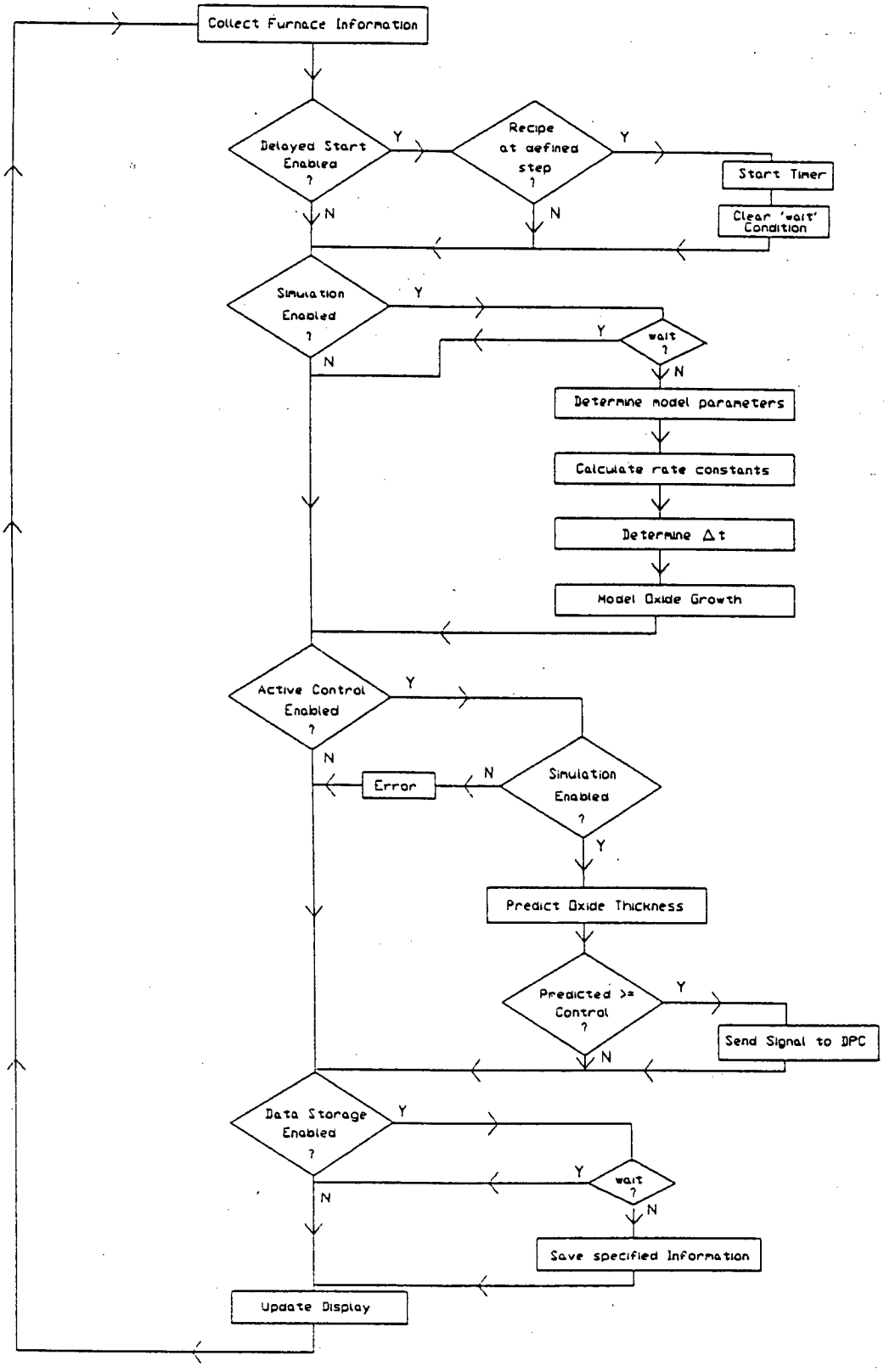


Fig. 7.15 Main loop of the FACS program.

system timer. It has already been shown that the choice of Δt in the range 0.5 to 10 seconds makes little difference to the simulation. For practical reasons, Δt is determined to the nearest second.

The complete program listing of FACS is presented in Appendix C and since this is extensively commented it is described here only briefly.

A series of routines initialise the system before the loop of Fig. 7.15 is entered. Initialisation includes keyboard and screen initialisation, configuration of communication ports, definition of functions, and assignment of constants (for use in later routines). Separate routines for data collection, the growth model, the control functions, data storage and visual presentation of furnace data are then executed as required. The routine to collect furnace data implements the SECS protocol in full. Subroutines which send and receive one character at a time on the serial interface are used. Further routines take care of error conditions and the function key options. In the latter case keyboard input may be requested for necessary information such as wafer orientation, required oxide thickness (control value) and filenames for stored data.

The routine for active control has been developed based on the control philosophy described in the preceding section. It is used in conjunction with a corresponding format for the recipe in the DPC. In this sense, the control routine in the program listing of Appendix C has been customised. However, all other routines are generally applicable and a switch to an alternative control philosophy requires modification only to the control routine.

7.6 Investigation of System Capabilities

The feasibility of the simulation control technique was examined by testing its ability to deliver a specified oxide thickness. Many trials were conducted, under a wide range of oxidation conditions and in this section the outcome of these trials is presented.

An initial set of tests examined the capability of the system under straightforward wet and dry conditions at several temperatures. The furnace recipes in Tables 7.5 and 7.6 were used as a basis, but were modified in accordance with the requirements of the FACS procedure. Hence the oxidation step (step 3) was programmed with 'infinite' duration, but configured to jump automatically to the anneal step when requested to do so by FACS. Gas flows were kept at the values in the tables. Oxidation temperature was set as before in the 'standby' step of the recipe and remained at this value throughout the process. Under these circumstances, a trigger time of 21 seconds is appropriate.

To replace the missing oxidation time, a control value for oxide thickness was entered at the computer. The system was tested with thicknesses of 250Å for dry oxidation and 250Å and 1000Å for wet oxidation. Temperature set points of 900°C, 950°C and 1025°C were used. Experimental procedures were the same as those described in Section 7.3, and after oxidation five wafers were measured at five sites using the Nanospec. The results are summarised in Table 7.13.

In the table, relative error (RE) is calculated using :

Ambient	Nominal Temp (°C)	Control Thickness (Å)	Measured Thickness (Å)	CV(%)	Diff* (Å)	Rel. Error (%)
DRY	900	250	247.6	1.9	-2.4	1.0
DRY	950	250	251.5	1.6	+1.5	0.6
DRY	1025	250	256.8	1.4	+6.8	2.7
WET	900	250	259.6	9.0	+9.6	3.8
WET	950	250	273.1	6.9	+23.1	9.2
WET	900	1000	952.2	4.5	-47.8	4.8
WET	950	1000	990.0	3.9	-10.0	1.0
WET	1025	1000	1050.0	3.6	+50.0	5.0

* Diff=Measured thickness - control thickness

TABLE 7.13 Results from the initial set of trials with FACS.

$$RE = \frac{|\text{measured thickness} - \text{control thickness}|}{(\text{control thickness})} \times 100\ \%$$

The relative errors generally agree with the previously calculated model fitting parameters. An exception is the wet 250Å oxidation at 950°C, although this may be explained by the relatively extreme conditions, in control terms, with an oxidation time of approximately 5 minutes. There is no obvious trend in the results to indicate that the trigger time of 21 seconds needs adjustment. The results are good, giving less than 5% error, and are initial proof that the concept of the control technique is practicable.

A second set of tests were carried out to prove the system under a wider range of conditions and to examine further the capabilities of the model. Again the same basic furnace procedures given in Tables 7.5 and 7.6 were used, but HCl-oxidation and oxidation under different oxidant partial pressures substituted the defined oxidation steps. The following conditions were programmed into step 3 of the recipe:

i) 2% HCl oxidation — Control thickness = 250Å
   O_2 flow = 5.0 SLM
   HCl flow = 0.1 SLM

ii) 5% HCl oxidation — Control thickness = 250Å
   O_2 flow = 5.0 SLM
   HCl flow = 0.25 SLM

iii) Wet oxidation P_{H2O} = 0.86 — Control thickness = 800Å
   H_2 flow = 3.0 SLM
   O_2 flow = 2.0 SLM

iv) Wet oxidation P_{H2O} = 0.75 — Control thickness = 800Å
   H_2 flow = 3.0 SLM
   O_2 flow = 2.5 SLM

v) Wet oxidation P_{H2O} = 0.46 Control thickness = 800Å
H_2 flow = 3.0 SLM
O_2 flow = 5.0 SLM

vi) Dry oxidation P_{O2} = 0.5 Control thickness = 250Å
O_2 flow = 2.5 SLM
N_2 flow = 2.5 SLM

vii) Dry oxidation P_{O2} = 0.2 Control thickness = 250Å
O_2 flow = 1.0 SLM
N_2 flow = 4.0 SLM

A temperature of 950°C was used in all cases.

The results from these trials are presented in Table 7.14 and are again seen to be acceptable, but on the whole poorer than those from the first trials. In the HCl oxidations, both results are below the control values indicating the oxidation time has been under-estimated by FACS. This suggests the HCl enhancement factors are too high, reflecting the inaccuracy of using interpolated values within a look-up table to determine these factors.

As the oxidant partial pressure decreases, the measured thickness tends to increase above the control value in both wet and dry cases. Assuming that the accuracy of the determined partial pressure is high (equations 7.5 and equation 7.6), this would suggest poorer modelling of the partial pressure dependence of rate at low values. An actual oxidation rate higher than that predicted by the simulation would explain the results.

To give a more demanding test of the system capability, a set of trials involving substantial oxide growth under dynamic conditions was carried out. In the first two of these trials, wet and dry processes that incorporated a temperature ramp during oxidation were

Ambient	Control Thickness (Å)	Measured Thickness (Å)	CV(%)	Diff* (Å)	Rel. Error (%)
i)	250	242.8	1.8	-7.2	2.9
ii)	250	238.1	3.0	-11.9	4.8
iii)	800	810.6	2.3	+10.6	1.3
iv)	800	821.8	2.2	+21.8	2.7
v)	800	851.4	1.8	+51.4	6.4
vi)	250	256.3	2.1	+6.3	2.5
vii)	250	263.1	1.9	+13.1	5.2

* Diff=Measured thickness - control thickness

TABLE 7.14 Results from the second set of trials using FACS to control oxide thickness. The ambients i) to vii) are described in the text.

used. For the wet process, wafers were loaded into the furnace at 700°C and the temperature was ramped to 950°C. In the dry process wafers were loaded at 850°C then ramped to 950°C. In both cases, oxidation was continued at 950°C until the control thickness had been achieved, then automatically terminated as before. The detailed furnace recipes are given in Tables 7.15 and 7.16. Oxide thicknesses of 1000Å for wet oxidation and 250Å for dry oxidation were programmed as control values.

These two processes would present difficulty if a conventional control system, based on fixed oxidation times, was used. Determination of the oxidation time in the fourth step requires knowledge of the oxide growth during temperature ramping. However the simulation-control system automatically and accurately takes this oxide growth into account, as indicated by the results given in Table 7.17.

With the data storage capability of the FACS system, it is possible to analyse the oxidation process in greater detail. Figs 7.16 and 7.17 respectively show the temperature measured at the wafer boat and the corresponding oxide growth, computed in real-time, for the wet oxidation. At the end of the temperature ramp an estimated 290Å of oxide has been grown. This is a significant proportion of the total, emphasising the need for careful assessment of oxide growth during the ramp. Oxidation has continued for a further 18 minutes before the change to an inert ambient. This time has been proven correct but it is based on furnace conditions during this particular run only and does not represent a 'standard' time for step 4 of this process. Simulation-based control will always adjust this time as necessary to ensure growth of the required oxide thickness. Similarly, for the dry oxidation process (Figs. 7.18 and

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	700	-	'Standby' Position
1	5.0	0	0	0	250	700	15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700	700	10	Paddle replaced
3	0	3.0	2.5	0	1700	950	25	Temperature ramped at 10°C/min
4	0	3.0	2.5	0	1700	950	ii)	Oxidation
5	5.0	0	0	0	1700	950	10	Purge tube and anneal
6	5.0	0	0	0	250	950	15	Paddle out of furnace to unload wafers

i) See Table 7.5

ii) Step 4 time is determined by FACS to give a total oxide thickness of 1000Å.

TABLE 7.15 Wet oxidation process with temperature ramp, used in conjunction with FACS.

Step	Gas Flows (SLM) N_2	O_2	H_2	HCl	Paddle[i] Position	Temp (°C)	Time (Min)	Comments
0	5.0	0	0	0	1700	850	-	'Standby' Position
1	5.0	0	0	0	250	850	15	Paddle out of furnace to load wafers
2	5.0	0	0	0	1700	850	15	Paddle replaced
3	0	5.0	0	0	1700	950	20	Temperature ramped at 5°C/min
4	0	5.0	0	0	1700	950	[ii]	Oxidation
5	5.0	0	0	0	1700	950	10	Purge tube and anneal
6	5.0	0	0	0	250	950	15	Paddle out of furnace to unload wafers

i) See Table 7.5

ii) Step 4 time is determined by FACS to give a total oxide thickness of 250Å.

TABLE 7.16 Dry oxidation process with temperature ramp, used in conjunction with FACS.

Ambient	Control Thickness (Å)	Measured Thickness (Å)	CV(%)	Diff* (Å)	Rel. Error (%)
WET	1000	1008.0	2.2	+8	0.8
DRY	250	252.1	1.6	+2	0.8

* Diff=Measured thickness - control thickness

TABLE 7.17 Results from trials incorporating temperature ramps during the oxidation period.

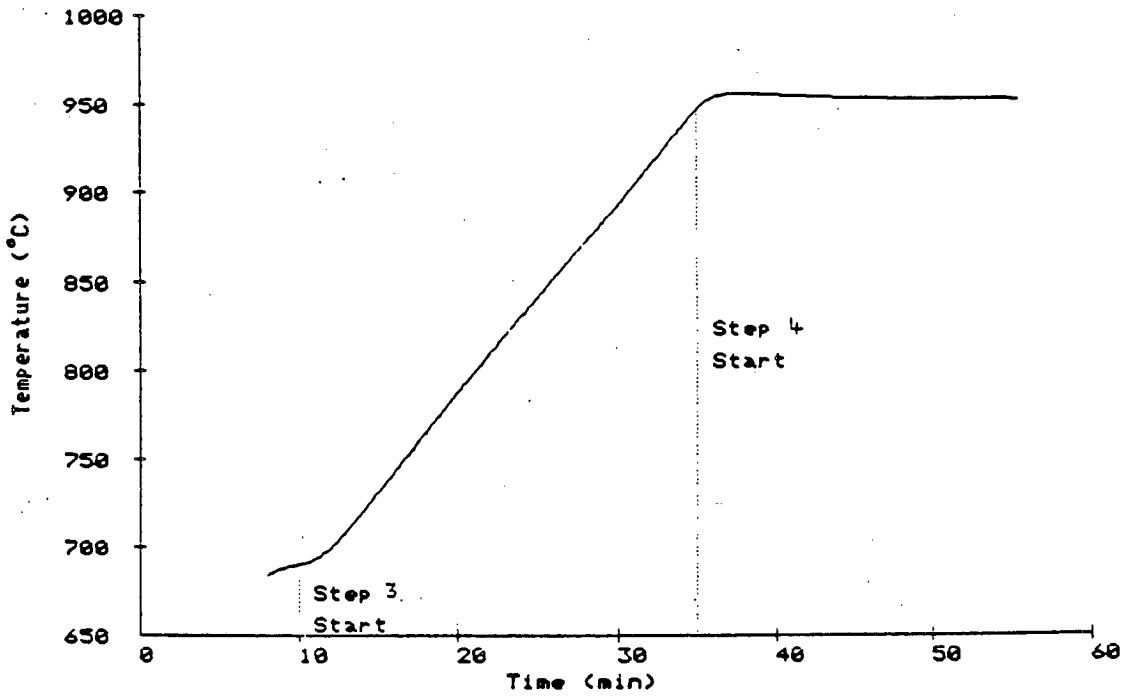


Fig. 7.16 Temperature measured at the wafers during the ramped wet oxidation process (Table 7.15). Time is measured from the start of step 2.

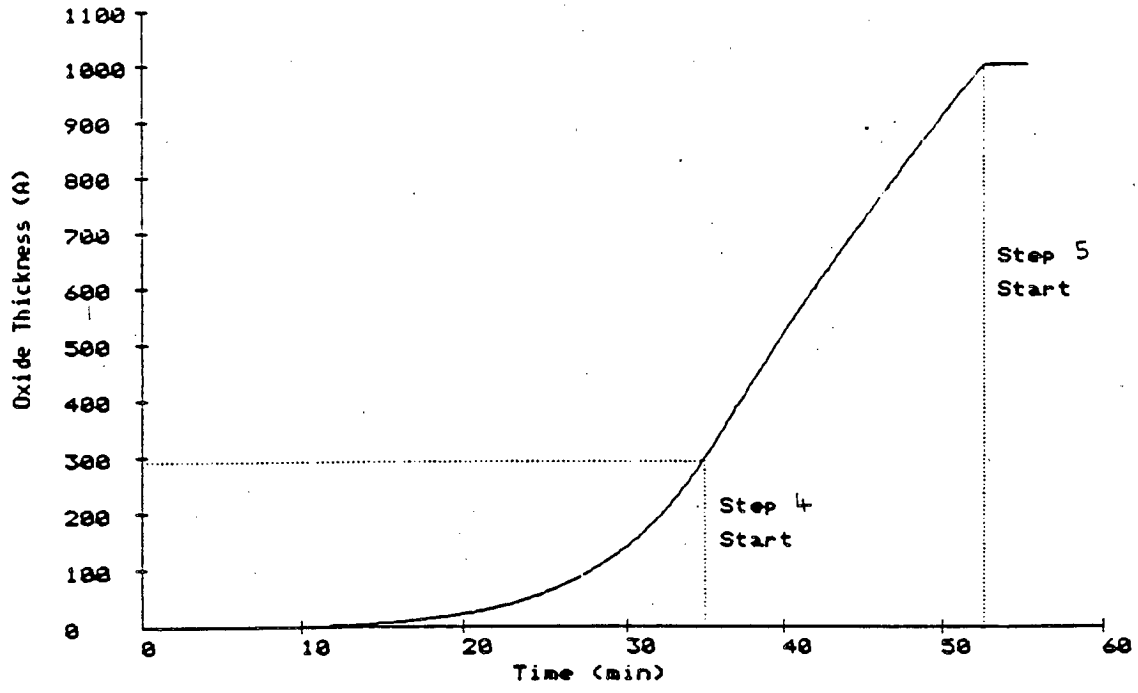


Fig. 7.17 Oxide growth simulated in real-time according to the temperature data of Fig. 7.16 (wet oxidation).

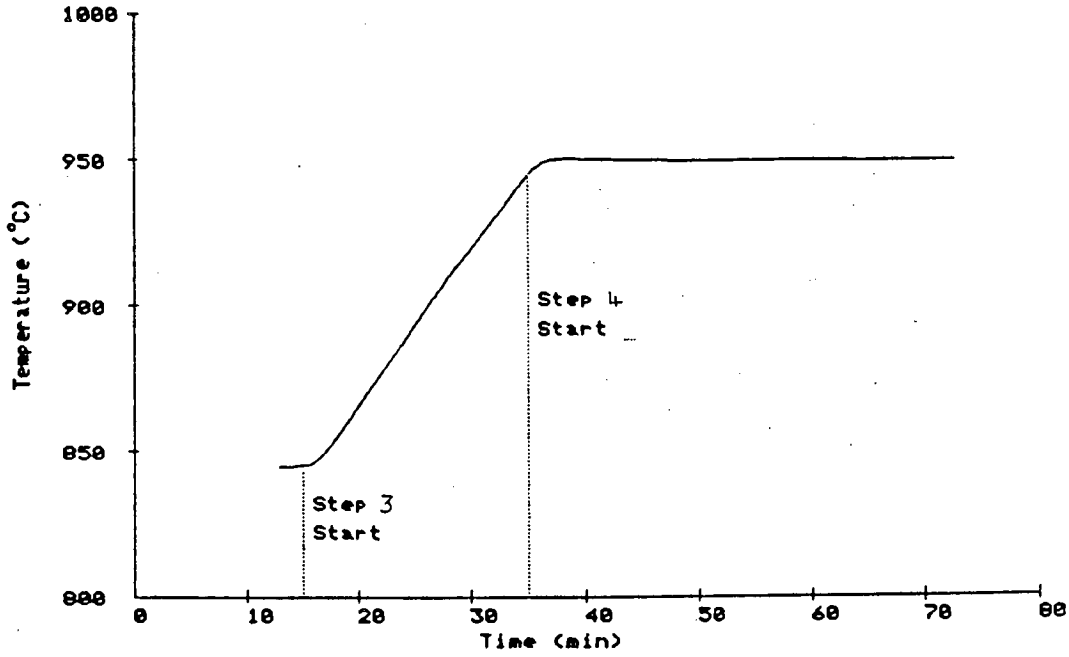


Fig. 7.18 Temperature measured at the wafers during the ramped dry oxidation process (Table 7.16). Time is measured from the start of step 2.

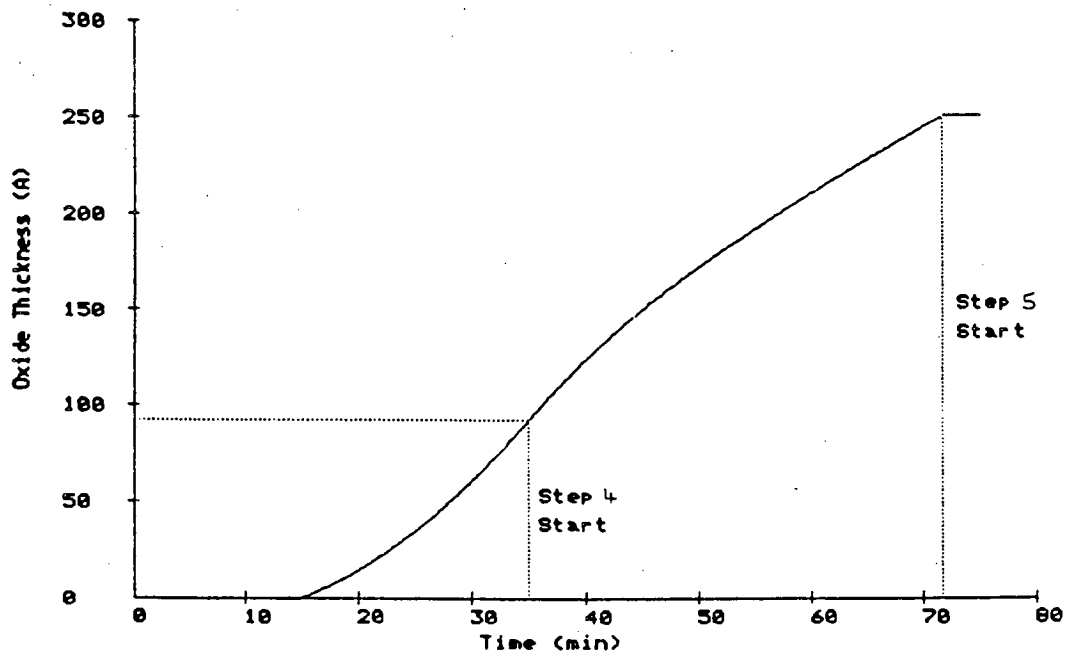


Fig. 7.19 Oxide growth simulated in real-time according to the temperature data of Fig. 7.18 (dry oxidation).

7.19) an estimated 90Å oxide has been grown by the end of temperature ramping (step 3) and a further 37 minutes oxidation time has been required to obtain the desired oxide thickness of 250Å.

The two temperature plots (Figs. 7.16 and 7.18) show that close to linear ramps at the specified rates (10°C/min and 5°C/min respectively) have been achieved at the position of the wafers. However, there is a lag of approximately one minute between the time-point at which the temperature set point starts to increase (step 3) and the point where a rise in temperature is seen at the wafers. This lag continues throughout the ramp and there is then a corresponding transient period in the supposedly 'stable' step 4 where the actual wafer temperature catches up to the set point. Whilst a temperature lag may be considered normal during ramping, only by collecting furnace data can it be clearly seen and analysed. The difference between the actual furnace conditions and the process description (as defined by the furnace recipe) is of course of no consequence to FACS, since the simulation is based entirely on the actual (measured) conditions.

A final trial was carried out, in which dynamic conditions were instigated by temporarily cutting power to the furnace heating coils whilst wafers were being oxidised. The furnace recipe of Table 7.6 was again used as a basis, and a control oxide thickness of 250Å was programmed into FACS. Twelve minutes after the oxidation step had started, power to the heating coils was switched off for ten minutes. As a result, the wafer temperature dropped to below 900°C, as shown in Fig. 7.20. Once power was reinstated, the temperature climbed back up to the 950°C set point in approximately

nine minutes, and remained at this value for the remainder of the oxidation step.

The oxidation rate during the period of reduced temperature shows a corresponding decrease (Fig. 7.21). Without real-time simulation it would be impossible to allow for this and still obtain the desired oxide thickness. FACS, however, automatically extends the oxidation time to compensate for the temperature reduction. Comparison with a simulation based directly on the furnace recipe (i.e. assuming a constant temperature of 950°C) shows that an extra 10 minutes have been needed (Fig. 7.22). The average oxide thickness measured was 249.8Å with a coefficient of variation of 2%, which shows that the real-time simulated oxide growth has corresponded well with the actual growth.

In summary, the correspondence between desired oxide thickness and the value obtained after processing, in the above trials, demonstrates the feasibility of the technique. Some limitations of the model and model parameters have been noted, but oxide thicknesses within 5% tolerance limits have still been achieved. The system has successfully accommodated dynamic process conditions in the form of deliberately-introduced temperature changes.

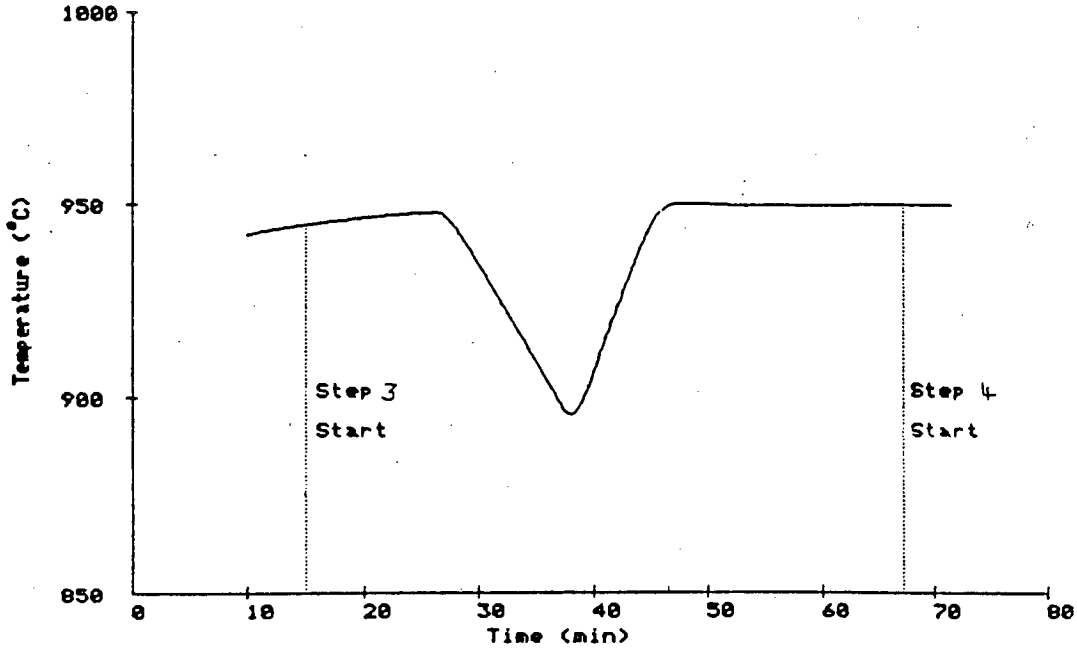


Fig. 7.20 Temperature measured at the wafers. Furnace heating was switched off at 27 mins and switched back on at 37 mins. Time is measured from the start of step 2 in Table 7.6.

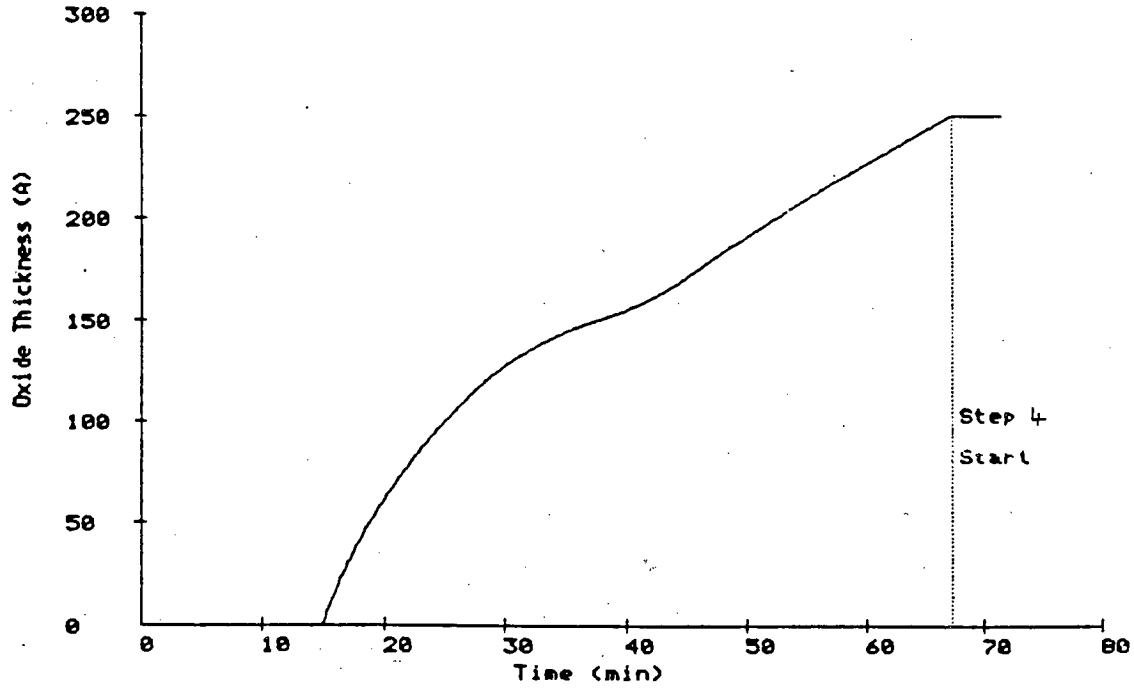


Fig. 7.21 Oxide growth simulated in real-time according to the temperature data of Fig. 7.20 (dry oxidation).

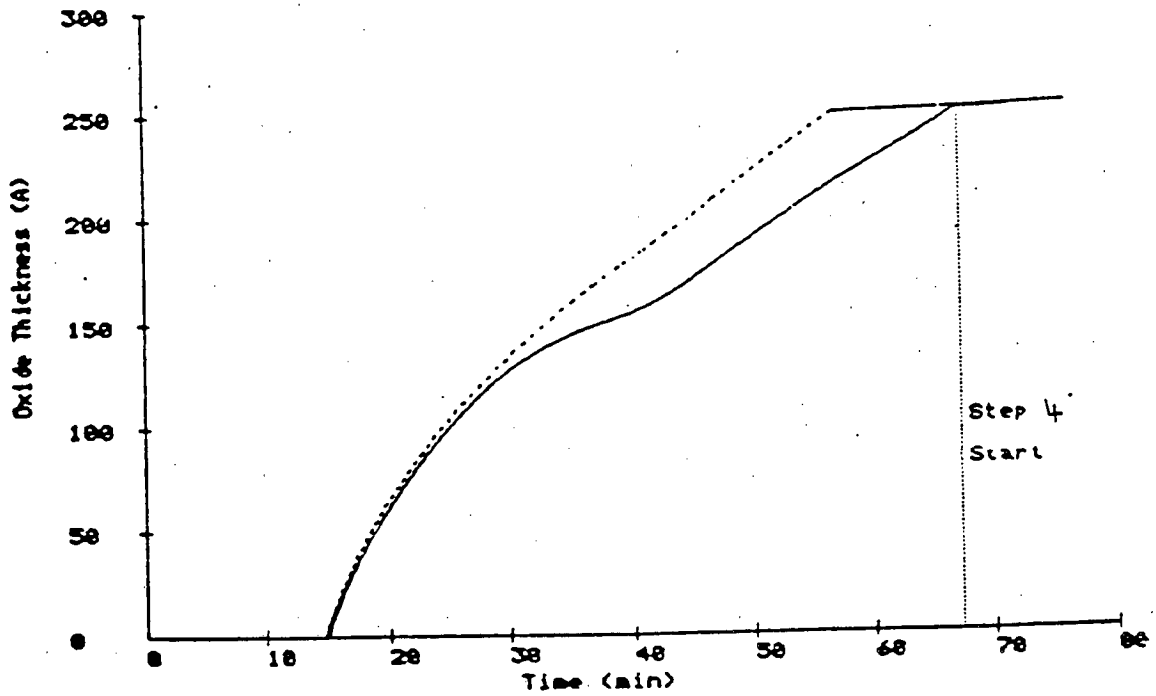


Fig. 7.22 Comparison between real-time simulation (solid line) and simulation assuming a constant temperature of 950°C (dashed line).

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

This thesis has addressed a topic that is crucial to the continued advancement of IC manufacturing, namely the enhancement of process control. Process output parameters have ever tighter tolerance requirements, which can only be met by improving the reproducibility of the individual steps involved. One such parameter was shown to be gate oxide thickness in MOS devices, and this work has concentrated on the growth of thin silicon dioxide films. A novel control technique for the oxidation process that overcomes the problems of conventional approaches has been developed and tested under a range of oxidation conditions.

The basis of the control technique is real-time simulation of the processes involved, and in Chapter 3 the current understanding of the silicon oxidation process was described. The Deal-Grove model for the growth kinetics has a central role in this understanding and was derived in detail. The principal factors that influence oxide growth were then discussed, namely oxidation ambient, temperature, oxidant pressure, HCl concentration, substrate orientation and doping level at the interface. It was shown to what extent oxidation theory would explain the observed effects.

Specific effects that are poorly understood or are not explained well by the theory were described. These are sub-linear pressure dependence of the linear rate constant (B/A) in dry oxidation ; the anomalously high oxidation rate observed with thin dry oxides ; enhanced oxidation rates with the addition of a few percent chlorine-containing compound to a dry ambient ; and

curvature in the Arrhenius plots of the rate constants. Difficulties with the accurate determination of rate constants was suggested as contributing to this lack of understanding. In the final section of Chapter 3, the oxide growth models implemented in process simulation packages were discussed. It was shown that SUPREM III used empirical additions to the theory to overcome some of the above difficulties and so extend its range of applicability.

Chapter 4 considered silicon oxidation from the practical point of view of wafer fabrication. A generalised scheme for an advanced digitally-controlled oxidation furnace was outlined. The extent to which such advanced equipment confers closed-loop control on the major furnace variables of temperatures, gas flows and boat movement was discussed. It was demonstrated that even with elaborate systems, transient periods are poorly controlled and that significant deviation from the control-loop set points exists. There is therefore a very important difference between the conditions expected (i.e. defined by the furnace recipe) and those actually experienced by the wafers.

The pre-oxidation treatment of wafers was shown to influence oxide growth and so it was concluded that wafer cleaning should be carefully regulated as an integral part of oxidation procedures.

Chapter 5 reviewed the principal techniques used to assess the thickness and quality of oxide films. With respect to thickness measurement, it was argued that the methods of interferometry and ellipsometry were preferable over alternatives such as capacitance measurements and surface profiling since they are non-destructive, accurate, and generally less time-consuming. It was shown that the Nanometrics Nanospec, a partially-

automated commercial instrument utilising the interferometry technique, has a sufficiently high reproducibility for application in the study of oxide thickness variability and oxidation growth rates.

In the review of quality assessment, the various oxide charge types, and their detection and measurement by MOS capacitor measurements, were first discussed. The complexity of the analysis methods has discouraged their widespread use in production environments, but it was pointed out that by automating the technique in a computer-based system and paying particular attention to the measurement conditions, accurate results can be obtained in a time-efficient manner. Commonly implemented methods for calculating the parameters Q_{eff}, Q_M, D_{it} and τ_g were presented. The importance of oxide breakdown measurements to obtain information about defect density and oxide reliability was then discussed.

MOS capacitor measurements were used in Chapter 6 as part of a verification procedure of the furnace used in this study. This concluded that there was no furnace contamination problem and that the furnace was capable of growing high-quality MOS gate oxides. It was then shown that such advanced furnace systems still give significant between-batch variation in oxide thickness and that this is related to inadequate control of the furnace variables. Oxide thickness variability was shown not to scale with decreasing oxide thickness. Growth patterns within a batch of wafers were related to the heating effect of the hydrogen torch and gas flows within the furnace tube.

Chapter 7 described in detail the hardware and software of the new control system. The software package developed was given the name FACS (Furnace Active Control System). This package provides all the

simulation and interfacing capabilities required by the control system, as well as a constantly-updated screen display of furnace conditions and the possibility of furnace data logging. It significantly extends the features offered by the furnace module. For active control, the system simply requires input of the desired oxide thickness and selection of an appropriate recipe at the furnace process controller (DPC).

A model based on that used in SUPREM III was implemented in the system. New rate constants for use with the model were evaluated. Although a poor comparison with SUPREM III values was obtained, both parameter sets gave comparable fitting errors to the oxide growth data, demonstrating general model insensitivity to the values used. No significant curvature in the Arrhenius plots was observed.

A series of trials involving oxidation in many different environments demonstrated the ability of the control technique to deliver a specified oxide thickness. Dynamic oxidation conditions were also successfully accommodated. Whilst all but one of the results were within 5% of the set point oxide thickness, those obtained under conditions where oxidation theory is weaker, notably chlorinated oxidation and oxidation under reduced oxidant partial pressures, were generally poorer.

By accurately accounting for the actual oxidation conditions and adjusting the oxidation times accordingly, the system enhances control of the oxidation process over the conventional approach where a fixed oxidation time is used. Batch-to-batch reproducibility benefits significantly, since every process run is optimised to the desired oxide thickness.

An immediate extension of the work presented here would be long-term examination of reproducibility using the system, preferably with several different oxidation processes. (This could most easily be carried out by implementing the system in a production environment.)

The proven ability of the system to impart control under dynamic growth conditions allows for greater flexibility in oxidation procedures without detriment to thickness reproducibility. Temperature ramping and wafer boat-loading can both be performed in oxidising ambients, for example. Advanced, more complex oxidation procedures can also be carried out. Thus sequential wet and dry oxidation or the multi-step oxidations discussed in Section 2.3.1 (e.g. dual-HCl) to improve oxide quality need not imply difficulties with thickness control.

Future development of the system could include enhancement to the oxide growth simulation. In particular, the modelling of HCl-related rate-enhancement could be improved, perhaps through a more detailed look-up table in the absence of an adequate empirical or theoretical model. Simulation at high doping densities may also be added to provide a more comprehensive system capability. Existing models (e.g. from SUPREM III) that estimate dopant redistribution could be implemented. Whether the extra calculation time would warrant software that runs more quickly than the BASIC program developed here, or even prohibits real-time simulation all together, would require investigation.

The furnace data collected by the system contains a detailed description of temperatures and gas flows encountered during every run. A simpler alternative development to simulation of dopant redistribution in real-time would be incorporation of this data into SUPREM III or other process simulators. This would provide a

basis for more accurate simulation of the overall process.

There is also some scope to extend the data acquisition part of the system so that the vast amount of collected data can be presented in a meaningful convenient format. Graphical display of furnace variables as a function of time is an obvious choice. Data handling at the local level is preferable to use of the overall CAM system for detailed process information.

All measured inputs to the control system are obtained from the local furnace controller. This makes full use of the data that is already present in the system and has the advantage of requiring little modification to the existing furnace hardware (apart form interfacing considerations, the major change was the addition of a permanent paddle thermocouple at the position of the wafer boat). This approach has been shown to be capable of thickness control to within 5% tolerance limits. However, the system need not be restricted to these inputs. The microcomputer could easily be used to co-ordinate data from additional sensors outside the domain of the local furnace controller. This could be used to obtain further information about the actual process conditions, or for a more rapid update of measured parameters and thus a more accurate real-time simulation. Developments here might include temperature measurement on the wafer surface using pyrometric techniques, such as those developed for RTO systems, or an enhanced method of determining the oxidant concentration at the wafers.

Real-time simulation of gas flow dynamics in the furnace tube may alternatively be considered as a basis for calculating oxidant concentration. The relatively simple approach utilised in FACS does not explicitly

determine gas composition at the wafers, but rather at the position of the mass flow controllers, and relies on a fitting parameter in the control routine to compensate for the difference.

Of more general significance, is the central role of the microcomputer within the framework of the complete oxidation operation. Active control could be coupled with the automatic generation of recipes in the computer and their subsequent download to the local furnace controller. This would ensure co-ordination of the real-time control system and the programmed furnace conditions. A highly-automated furnacing facility for oxidation that requires only an oxide thickness and a few essential details about growth conditions (e.g. ambient and temperature) as input, can be conceived. With the addition of automatic wafer loading and wafer transport, the facility could encompass all processing steps from wafer cleaning through to final measurement and verification of oxide thickness.

Although this thesis has focused on silicon oxidation, the implications of the results in Chapter 7 can be taken much further. Real-time simulation can be used to enhance control of critical parameters in many other areas of fabrication. The technique can be most easily extended to other furnace operations, such as deposition of polysilicon and silicon nitride, and dopant diffusion. Furnaces with common control systems are often used for these operations (within a given fabrication facility). Straightforward transfer of many routines between systems (e.g. interfacing, display and data acquisition routines) is therefore possible. Other areas where microprocessor-based equipment control systems are available, and where active simulation-based control is directly feasible, are ion-implantation and

plasma etching. The techniques developed in this work can be combined with understanding of the relevant processes involved.

The creation of independent modules which confer a high degree of control on critical process parameters, and their subsequent integration in a highly-automated manufacturing environment, will significantly contribute towards the goal of high process quality and yield in advanced processing.

APPENDIX A

PRINCIPAL MOSFET RELATIONSHIPS

The purpose of this appendix is to demonstrate the importance of gate oxide thickness as a MOSFET parameter, for which only first order equations need be considered. Second order effects, as well as derivation of the equations presented here, are treated in several advanced texts [27,39].

Gate oxide thickness determines the gate capacitance, C_{ox}, which has a central role in device characteristics.

$$C_{ox} = \frac{E_o E_r}{x_o} \quad \text{(F/unit area)} \qquad (1)$$

where E_o is the permitivity of free space, E_r is the dielectric constant for SiO_2 and x_o is the oxide thickness.

A key parameter in determining transistor operation is the threshold voltage, V_T, which defines the gate voltage needed to induce a conducting channel at the Si-SiO_2 interface. Threshold voltage for a large geometry n-channel device is given by

$$V_T = \phi_{ms} - Q_{eff}/C_{ox} + 2\phi_F - Q_B/C_{ox} \qquad (2)$$

where

ϕ_{ms} = metal-semiconductor work function difference.

Q_{eff} = effective insulator charge (this is the additive effect of all the various charge types associated with the $Si-SiO_2$ system - see Section 5.2.1).

$\emptyset_F$ = substrate Fermi level.

Q_B = bulk charge density at onset of strong inversion

$= - (2qE_sE_oN_A(2\emptyset_F))^{\frac{1}{2}}$

where E_S is the dielectric constant of silicon, q is the electric charge and N_A is the substrate doping concentration.

The MOS transistor characteristics can be split into two regions: the linear region when $V_{DS} < V_{GS}-V_T$ and the saturation region when $V_{DS} > V_{GS}-V_T$. The equations that define these two regions are:

$$\text{Linear:} \quad I_{DS} = u.C_{ox}.W/L.(V_{GS}-V_T-V_{DS}/2).V_{DS} \qquad (3)$$

$$\text{Saturation:} \quad I_{DS} = u.C_{ox}.W/L.(V_{GS}-V_T)^2/2 \qquad (4)$$

where u is the electron mobility in the device channel, W and L are respectively the width and length of the device channel, I_{DS} is the drain current, V_{GS} is the applied gate bias and V_{DS} is the applied drain bias.

Device transconductance is a frequently used MOSFET electrical parameter that is directly related to the device speed. It is defined by :

$$\left.\frac{dI_{DS}}{dV_{GS}}\right|_{V_{DS}=\text{constant}} \qquad (5)$$

In the two regions of device operation, g_m is given by

Linear : $g_m = C_{ox}.u.W/L.V_{DS}$ (6)

Saturation : $g_m = C_{ox}.u.W/L.(V_{GS}-V_T)$ (7)

Device swing, S, characterises the device sub-threshold region. It is defined by the variation in gate voltage needed to reduce the drain current by one decade. S can be shown to be

$$S = kT/q \log_{10}(1 + C_D/C_{ox}) \quad (8)$$

where C_D is the depletion capacitance.

The effect of substrate bias on the threshold voltage is normally characterised through the body effect parameter γ. Assuming uniform substrate doping,

$$\gamma = \frac{(2q.E_S.E_o.N_A)^{\frac{1}{2}}}{C_{ox}} \quad (9)$$

The actual device threshold V_T' is then given by

$$V_T' = V_T + \gamma \{(2\emptyset_F + V_{BS})^{\frac{1}{2}} - (2\emptyset_F)^{\frac{1}{2}}\} \quad (10)$$

where V_{BS} is the substrate bias and V_T is the threshold voltage without the body effect (i.e. when $V_{BS}=0$).

The presence of C_{ox} as a term in the above first-order MOSFET relationships, clearly shows that the gate

oxide thickness must be carefully controlled in order to attain stable device performance.

APPENDIX B

SEMI EQUIPMENT COMMUNICATIONS STANDARD (SECS)

SECS [4] is a software and hardware standard for the interconnection of semiconductor manufacturing equipment to a host computer. The communication link is based on the RS-232-C serial interface standard. Messages are sent as eight-bit characters in one direction at a time and data exchange rates up to 9600 baud are allowed. SECS provides a handshaking protocol which establishes the communication direction, resolves contention and allows for "retries" to correct communication errors. The four ASCII characters which are used in the data link protocol are given in Table B1.

Operation of the protocol is most easily understood by considering the example of communication between the Tempress DPC furnace controller and the Apricot PC. The computer requires information relating to the current furnace status. Fig. B1 shows the first part of the communication, in which the computer sends a message with the request for information. After receiving ENQ, the DPC responds with EOT signifying that it is not busy and ready to receive. The message is then sent : it consists of a length byte (=N), N bytes of data and a 2-byte "checksum" which is the total of all the N individual bytes. The DPC compares the checksum with its own estimate obtained by the addition of the N message bytes received. If they agree then the ACK signal is returned, otherwise NAK is sent to prompt a retry. Receipt of the appropriate message prompts the DPC to send out information. The second part of the communication then takes place, following exactly the same protocol as in Fig. B1 but with the positions of DPC

and computer reversed. In this case the message contains the relevant furnace status information.

Table B2 shows details of the messages used in the FACS program. Messages consist of a 10-byte header followed by (N-10) bytes of data. The function of the header is to help decode and organise transmitted data. It is made up of (i) a device ID to ensure correct message routing (every piece of equipment in the system should have a unique identification); (ii) a message ID which describes the type of message being sent so that the equipment software will know how to deal with the message format; (iii) block numbers which allow for the transmission of long messages in a series of blocks; (iv) system bytes which allow equipment to be linked to an unfamiliar host computer system.

As well as the checksum, SECS provides timeouts to detect communication errors. These determine how long data transactions should take before they are aborted and a retry is attempted. Table B3 summarises the timeout parameters and gives values used in the DPC - Apricot PC communication link. Their implementation is shown in the flow-chart of Fig. B2, which outlines the full operation of the SECS protocol. This flow-chart defines the logic flow that should be realised in software at both ends of the communication link (i.e. both in the processing equipment and in the host computer). Line contention - when both ends attempt to send at the same time - is resolved in SECS by designating the host computer as "slave" and forcing it to postpone its transmission and switch to receive.

Character Name	ASCII Code	Function
ENQ	5	Request to send
EOT	4	Ready to receive
ACK	6	Correct reception
NAK	21	Incorrect reception

TABLE B1 Data link protocol characters used in SECS.

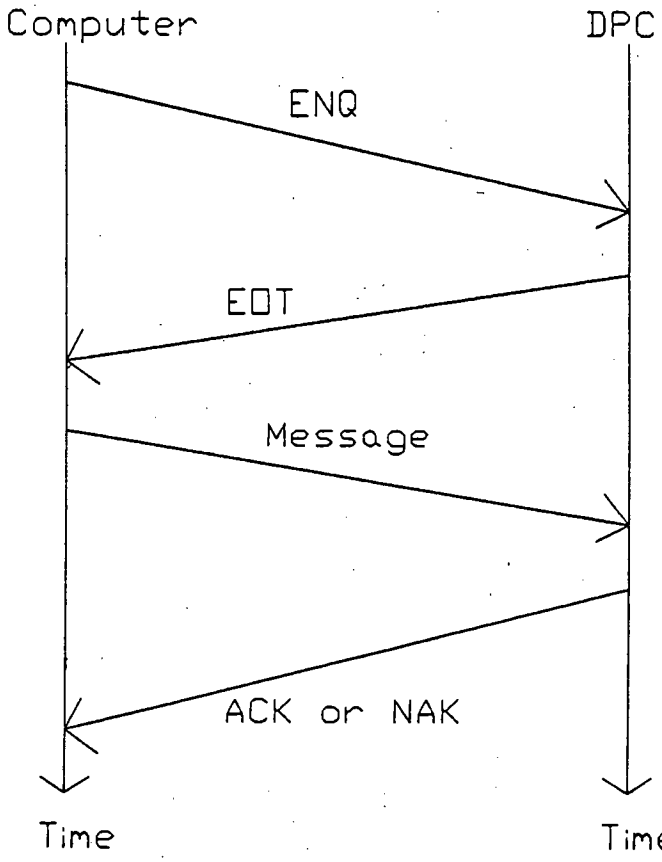


Fig. B1 Communication between host computer and DPC furnace controller using SECS.

Message sent from computer (10 bytes i.e.header only):

Byte	Value (Hex)	Definition
1 2	80 00	device ID
3 4	0B 00	message ID
5 6	00 00	block ID
7 8 9 10	00 00 00 00	system bytes

Reply from DPC (197 Bytes):

Byte	Definition
1-10	Header (as above)
11	Current Step No.
12-14	Remaining Step Time (hrs,min,sec)
15-18	Remaining Process Time (hrs,min,sec)
19-20	Boat Position (mm)
21-22	Boat Set Point (mm)
23-24	Boat Speed (mm/min)
25-26	Digital Inputs: bits 1-16 --> inputs 1-16
27-28	Digital Input Alarm Settings
29-30	Digital Outputs: bits 1-16 --> outputs 1-16
31-62	Gas Lines 1-16 actual flow (certified units)
63-94	Analog Inputs 1-16 (certified units)
95-126	Gas Lines 1-16 set point flow (cert. units)
127-132	DTC general Status
133-138	Paddle Thermocouples (TCs)-set points(0.1°C)
139-144	Paddle TCs - actual temperatures (0.1°C)
145-150	Spike TCs - set points (0.1°C)
151-156	Spike TCs - actual temperatures (0.1°C)
157-162	Temperature Deviation (set point - actual)
163-174	Ramp temperature and ramp rate
175-177	Power output (%)
178-181	Temperature Alarms
182-197	Alarm limit settings on gas lines 1-16 (%)

TABLE B2 Messages used in FACS to obtain detailed furnace status information.

Symbol	Parameter Name	Function	Parameter Range	Value used
T1	Receive Timeout	Detects interruption in received data	0.1-10s	0.5s
T2	Protocol Timeout	Detects lack of protocol response	0.2-25s	1s
T3	Reply Timeout	Detects lack of message reply	1-120s	2s
RTY	Retry Count	Maximum No. of send retries	0-31	4

TABLE B3 Data link protocol parameters.

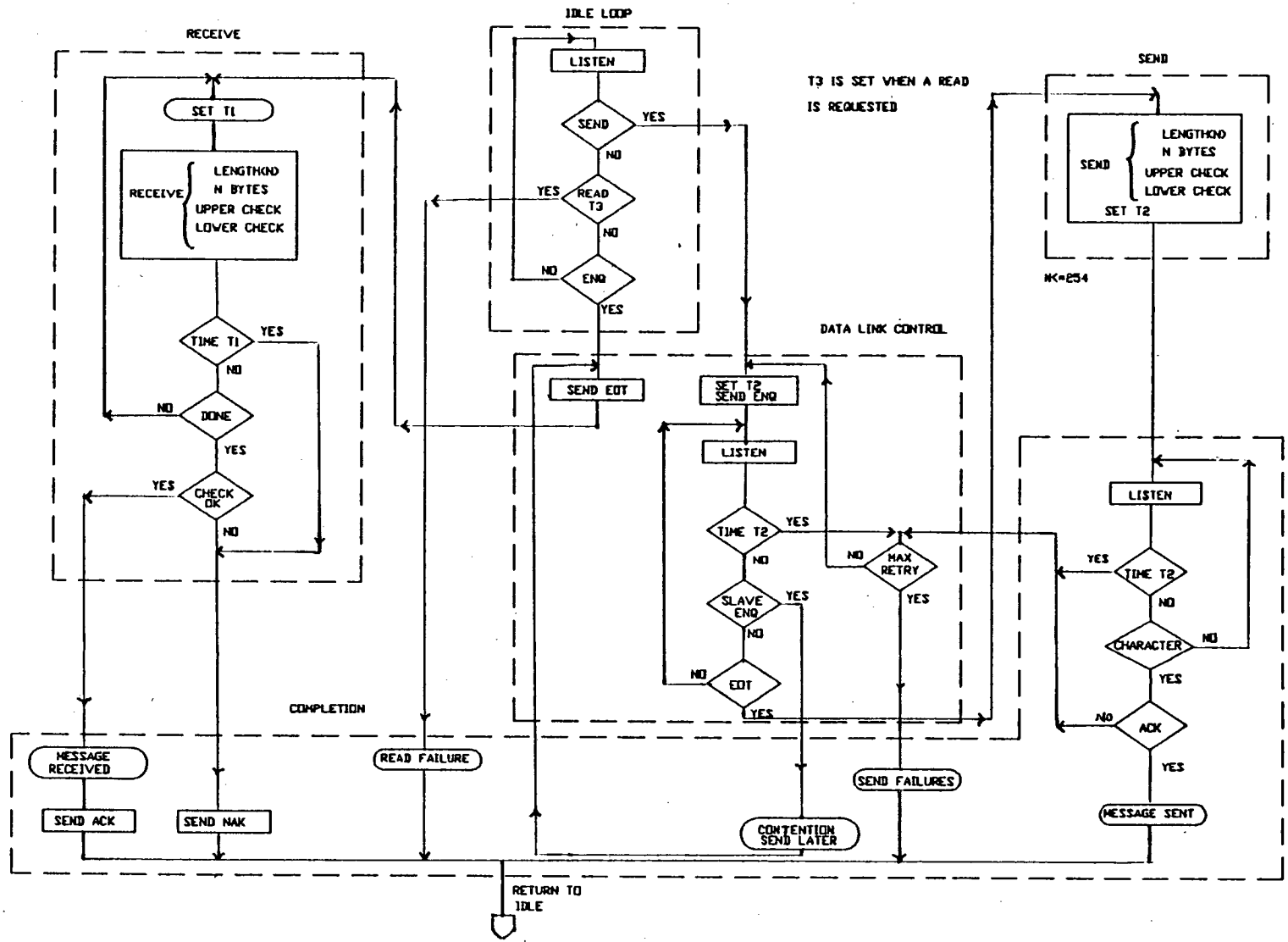


Fig. B2 SECS data link protocol.

APPENDIX C

FACS Program Listing

Definition of Selected Program Variables

```
General:
   CT      :Active Control Option                    0=Off, 1=On
   DI4     :Status of Digital Input 4 in DPC:0=low, 1=high
   D(250)  :SECS Message: DPC -> Computer (See Appendix B)
   GD      :Save Oxide Growth Data                   0=Off, 1=On
   GF      :Save Gas Flow Data                       0=Off, 1=On
   M(13)   :SECS Message: Computer -> DPC (See Appendix B)
   RTY     :SECS Retry Count
   SAV     :Data Storage Option                      0=Off, 1=On
   SC      :Delayed Start Option                     0=Off, 1=On
   SIM     :Real-time Simulation Option              0=Off, 1=On
   TP      :Save Temperature Data                    0=Off, 1=On
   WATE    :Delay Status (when SC=1)   0=proceed, 1=wait

Simulation:
 A,B                :Rate Constants
 DH,DM,DS           :Change in Time (hr,min,sec)
 DOX                :Change in Oxide Thickness
 DT                 :Calculated Time Interval
 EH,EM,ES           :Total Elapsed Time from Start
                     (hr,min,sec)
 G1$,G2$,G3$,G4$    :Furnace Gas Status (N2,O2,H2,HCL resp.)
 LLEN,PPEN          :HCl Enhancement Factors
                     (Linear,Parabolic)
 LUTL,LUTP          :Look-up Tables for HCl Enhancement
                     Factors (Linear,Parabolic)
 ORENT              :Silicon Crystal Orientation
                     0=<100>, 1=<111>
 OXSPA              :Required Oxide Thickness (Å)
 OXTHIA             :Simulated Total Oxide Thickness (Å)
 OXTYPE             :Oxidation Ambient
                     0=Dry,1=Wet,2=Non-oxidising
 PDT                :Time Interval for Future Oxide Growth
                     (min)
 PDOX               :Future Change in Oxide Thickness (Å)
 POXTHIA            :Future Total Oxide Thickness
 PRES               :Oxidant Partial Pressure
 SH,SM,SS           :Clock Start Time (hr,min,sec)
 TEMP               :Current Oxidation Temperature (°C)
 THINRATE           :Thin Rate Enhancement Factor
```

```
10 REM ************************************************************
20 REM ************************************************************
30 REM             FURNACE ACTIVE CONTROL SYSTEM (FACS)
40 REM                        Version 2.0
50 REM                  Ewan Cameron    17 June 1986
60 REM      Main Features : Display of Furnace info.
70 REM                      Data Storage
80 REM                      Real Time Oxide Growth Simulation
90 REM                      Active Control of Furnace Operations
100 REM
110 REM ************************************************************
120 REM ************************************************************
130 REM INITIALISATION
135 REM
140 DIM M(13), D(250)
150 GOSUB 280         'Define functions, constants etc.
160 GOSUB 1430        'Configure serial interface communications
170 GOSUB 4790        'Configure IEEE-488 interface communications
180 GOSUB 1600        'Create Display
190 GOSUB 3560        'Reset
200 REM ------------------------------------------------------------
205 REM MAIN PROGRAM LOOP
206 REM
210 GOSUB 590           'Collect furnace info
220 IF SC AND WATE THEN GOSUB 3500    'Delayed Start
230 IF SIM THEN GOSUB 3730            'Simulation
240 IF CT THEN GOSUB 5230             'Active Control
250 IF SAV THEN GOSUB 3640            'Data Storage
260 GOSUB 1870         'Update display
270 GOTO 210
280 REM ************************************************************
290 REM Define functions, constants etc.
300 REM
310 ON ERROR GOTO 3140
320 REM
330 REM Configure function keys
340 ON KEY(1) GOSUB 2330
350 ON KEY(2) GOSUB 2520
360 ON KEY(3) GOSUB 2700
370 ON KEY(4) GOSUB 2780
380 ON KEY(5) GOSUB 2890
390 REM
395 REM Functions to convert Hex. data to decimal
400 DEF FNBD(I)=(D(I)*256)+D(I+1)
410 DEF FNTD(I)=(D(I)*256+D(I+1))/10
420 DEF FNGD(I)=(D(I)*256+D(I+1))/100
430 DEF FNHD(I)=(D(I)*256+D(I+1))/1000
435 REM
436 REM Function to Separate TIME$ into hrs,min,secs
440 DEF FNTM(STP$,I)=VAL(MID$(STP$,I,2))    'I=1,4,7 ==> hr,min,sec
490 REM
495 REM Start up Configuration
500 RC=22 : CC=2
```

```
510 PAUSE=1500
515 REM
520 SIM$="OFF"  :SIM=0              'Simulation off
530 SAV=0  :TP=0  :GF=0  :GD=0      'Data Storage off
540 CT$="OFF"  :CT=0  :DI4=0        'Active Control off
550 SC$="IMM  "  :SC=0  :WATE=0    'Delayed Start Option off
560 GOSUB 1350                      'Initialise SECS message
570 LABEL$="1.SIMULATION  2.DATA STORAGE  3.CONTROL  4.TIME
5.RESTART"
580 RETURN
590 REM ************************************************************
600 REM COMMUNICATATION WITH DPC USING SECS
610 REM
615 REM ------------------------------------------------------------
620 REM IDLE
630 REM
635 FOR I=1 TO 5  :KEY(I) STOP  :NEXT I     'Disable function keys
640 RTY=0  :CHANGE=0  :DEF SEG=SER
650 GOSUB 690                       'Send request for info to DPC
660 FOR I= 1 TO 150  :NEXT I        'Pause
670 GOSUB 1310 :IF A4%=5 THEN GOSUB 920   'Receive info from DPC
675 FOR I=1 TO 5  :KEY(I) ON  :NEXT I     'Re-enable function keys
680 RETURN
685 REM ------------------------------------------------------------
690 REM SEND MESSAGE
700 REM
710 'Data Link Control (Send)
720 A3%=5  :GOSUB 1260
730 FOR I=1 TO 10
740 GOSUB 1310
750 IF A4%=4 THEN GOTO 780
760 NEXT I
770 GOSUB 1200
780 'Send Message
790 A2%=1
800 FOR B=1 TO 13
810 A3%=M(B)
820 CALL IO%(A1%,A2%,A3%,A4%)
830 NEXT B
840 'Complete Send
850 FOR J=1 TO 10
860 GOSUB 1310
870 IF A4%=6 THEN RETURN
880 IF A4%=21 THEN GOSUB 1200
890 NEXT J
900 GOSUB 1200
910 RETURN
915 REM ------------------------------------------------------------
920 REM RECEIVE MESSAGE
930 REM
940 'Data Link Control (Receive)
950 A3%=4  :GOSUB 1260
960 'Receive Message
```

```
970 SUM=0
980 A2%=2
990 CALL IO%(A1%,A2%,A3%,A4%)
1000 IF A4%=5 THEN GOTO 990
1010 IF A4%=21 THEN RETURN
1020 LENGTH=A4%
1030 FOR L=1 TO LENGTH+2
1040 CALL IO%(A1%,A2%,A3%,A4%)
1050 SUM=SUM+A4%
1060 D(L)=A4%
1070 NEXT L
1080 SUM=SUM-D(LENGTH+1)-D(LENGTH+2)
1090 CHK=(D(LENGTH+1)*256)+D(LENGTH+2)
1100 IF SUM <>CHK THEN GOTO 1150
1110 'Complete Receive (Correct)
1120 A3%=6  :GOSUB 1260
1130 CHANGE=1
1140 RETURN
1150 'Complete Receive (Incorrect)
1160 A3%=21  :GOSUB 1260
1170 CHANGE=0
1180 RETURN
1185 REM ------------------------------------------------------
1190 REM SEND RETRY
1200 REM
1210 RTY=RTY+1
1220 IF RTY<=4 THEN GOTO 710
1230 GOTO 640
1240 RETURN
1245 REM ------------------------------------------------------
1250 REM ROUTINE TO SEND CHARACTER ON SERIAL INTERFACE
1260 REM
1270 A2%=1
1280 CALL IO%(A1%,A2%,A3%,A4%)
1290 RETURN
1295 REM ------------------------------------------------------
1300 REM ROUTINE TO RECEIVE CHARACTER FROM SERIAL INTERFACE
1310 REM
1320 A2%=2
1330 CALL IO%(A1%,A2%,A3%,A4%)
1340 RETURN
1350 REM ******************************************************
1360 REM SECS MESSAGE FOR FULL STATUS INFORMATION FROM DPC
1370 REM
1380 RESTORE 1400
1390 FOR I=1 TO 13  :READ M(I)  :NEXT I
1400 DATA &HA,&HB0,0,&HB,0,0,0,0,0,0,0,0,&H8B
1410 RETURN
1420 REM ******************************************************
1430 REM CONFIGURE SERIAL INTERFACE
1435 REM
1440 SER=&H60  :IO%=0
1450 DEF SEG=SER
```

```
1460 A1%=52  :A4=0
1470 A2%=4   :A3%=14  :GOSUB 1570    'Baud rate=9600
1480 A2%=5   :A3%=14  :GOSUB 1570
1490 A2%=6   :A3%=8   :GOSUB 1570    '8 Bits/Character
1500 A2%=7   :A3%=8   :GOSUB 1570
1510 A2%=8   :A3%=1   :GOSUB 1570    'One stop bit
1520 A2%=9   :A3%=0   :GOSUB 1570    'No parity
1530 A2%=3            :GOSUB 1570    'Update the SIO
1540 DEF SEG
1550 A1%=52  :A2%=0  :A3%=0  :A4%=0
1560 RETURN
1570 CALL IO%(A1%,A2%,A3%,A4%)
1580 RETURN
1590 REM ******************************************************
1600 REM CREATE SCREEN DISPLAY FORMAT
1610 REM
1620 CLS :DRAW "S10BM5,10D125R315U125L315"
1630 LINE (5,39)-(790,39)  :LINE (5,165)-(790,165)
1640 LINE (5,295)-(790,295) :LINE (380,165)-(380,295)
1650 LOCATE 2,3,0  :PRINT  "FURNACE 10 :"
1660 LOCATE 2,35   :PRINT  "Step No.    Step Time (h/m/s):"
1670 LOCATE 4,30   :PRINT  "Temperature Status"
1680 LOCATE 6,32   :PRINT  "Rear        Centre        Load"
1690 LOCATE 7,3    :PRINT  "Temp"  :LOCATE 7,20  :PRINT  "SP(°C)"
1700 LOCATE 8,3 :PRINT "Recipe" :LOCATE 8,16 :PRINT  "Actual (°C)"
1710 LOCATE 9,17   :PRINT  "Spike (°C)"
1720 LOCATE 10,18  :PRINT  "Power (%)"
1730 LOCATE 12,14  :PRINT  "Gas Status"
1740 LOCATE 12,54  :PRINT  "Boat Status"
1750 LOCATE 14,13  :PRINT  "SP(SLM)  Actual (SLM)"
1760 LOCATE 14,45  :PRINT  "SP(mm)    Posn(mm)    Speed(mm/min)"
1770 LOCATE 15,3   :PRINT  "N2"
1780 LOCATE 16,3   :PRINT  "O2"
1790 LOCATE 17,3   :PRINT  "H2"
1800 LOCATE 18,3   :PRINT  "HCl"
1810 LOCATE 20,3   :PRINT  "Elapsed Time (h/m/s)"
1820 LOCATE 20,40  :PRINT  "Oxide Thickness (A)"
1830 LOCATE 24,10  :PRINT LABEL$
1840 LOCATE 25,16  :PRINT SIM$;  :LOCATE 25,28  :PRINT TP;GF;GD;
1850 LOCATE 25,43  :PRINT CT$;   :LOCATE 25,53  :PRINT SC$
1860 RETURN
1870 REM ******************************************************
1880 REM UPDATE DISPLAY
1890 REM Use Furnace Information coded in received DPC message
1895 REM
1900 LOCATE 2,43  :PRINT USING "##";D(11);
1910 LOCATE 2,65  :PRINT USING "## ";D(12);D(13);D(14)
1920 LOCATE 7,26  :PRINT USING "########.#" ;FNTD(133) ;FNTD(135)
;FNTD(137)
1930 LOCATE 8,26  :PRINT USING "########.#" ;FNTD(139) ;FNTD(141)
;FNTD(143)
1940 IF D(128)=0 THEN RT$="N " ELSE RT$="P "
1950 LOCATE 9,3  :PRINT RT$;D(127)
```

```
1960 LOCATE 9,26  :PRINT USING "#########.#" ;FNTD(151) ;FNTD(153)
;FNTD(155)
1970 LOCATE 10,26 :PRINT USING "#########  ";D(175);D(176);D(177)
1980 IF D(30)<>DPREV THEN GOSUB 2120  'Get gas status (if changed)
1990 LOCATE 15,8  :PRINT G1$;
2000 LOCATE 15,11  :PRINT USING "######.#";FNGD(63);FNGD(31)
2010 LOCATE 15,39  :PRINT USING "###########" ;FNBD(21) ;FNBD(19)
;FNBD(23)
2020 LOCATE 16,8  :PRINT G2$;
2030 LOCATE 16,11  :PRINT USING "######.#";FNGD(65);FNGD(33)
2040 LOCATE 17,8  :PRINT G3$;
2050 LOCATE 17,11  :PRINT USING "######.#";FNGD(67);FNGD(35)
2060 LOCATE 18,8  :PRINT G4$;
2070 LOCATE 18,11  :PRINT USING "######.#";FNGD(69);FNGD(37)
2080 LOCATE 20,26  :PRINT USING "## ";EH;EM;ES;
2090 LOCATE 20,26  :PRINT USING "#####";OXTHIA;
2100 RETURN
2110 REM ******************************************************
2120 REM DETERMINE GAS SYSTEM STATUS (FROM BYTE 30)
2130 REM
2140 DPREV=D(30)
2150 G1$="OFF"  :G2$="OFF"  :G3$="OFF"  G4$="OFF"
2160 IF D(30)<16 THEN GOTO 2180
2170 D(30)=D(30)-16
2180 IF D(30)<8 THEN GOTO 2210
2190 G4$="ON "
2200 D(30)=D(30)-8
2210 IF D(30)<4 THEN GOTO 2240
2220 G3$="ON "
2230 D(30)=D(30)-4
2240 IF D(30)<2 THEN GOTO 2270
2250 G2$="ON "
2260 D(30)=D(30)-2
2270 IF D(30)<1 THEN GOTO 2290
2280 G1$="ON "
2290 RETURN
2300 REM ******************************************************
2310 REM ROUTINES TO DISABLE/ENABLE FUNCTION KEY OPTIONS
2320 REM
2330 REM ------------------------------------------------------
2335 REM OPTION 1 : OXIDE GROWTH SIMULATION
2340 GOSUB 3080
2350 LOCATE RC,CC,1 :INPUT "Oxide Growth Simulation (0=OFF,1=ON)
:",IN$
2360 IF IN$="0" AND IN$<>"1" THEN 2490 ELSE SIM=VAL(IN$)
2370 IF SIM=0  AND SIM$="OFF" THEN GOTO 2490
2380 IF SIM=1  AND SIM$=" ON" THEN GOTO 2490
2390 IF SIM THEN SIM$=" ON" :GOTO 2440
2400 SIM$="OFF"
2410 ERASE LUTP,LUTL
2420 IF SAV=0 THEN EH=0:EM=0:ES=0 :EHP=0:EMP=0:ESP=0
2430 GOTO 2490
2440 LOCATE RC,CC :INPUT "    Si orientation - 0=<100> 1=<111>
```

```
:",ORENT
2450 IF ORENT <>0 AND ORENT <>1 THEN 2440
2460 IF ORENT THEN ALPHA=1.68 ELSE ALPHA=1
2470 GOSUB 4180                    'HCl enhancement coefficients
2480 IF SC=0 AND SAV=0 THEN STRT$=TIME$ :SH=FNTM(STRT$,1)
:SM=FNTM(STRT$,4) : SS=FNTM(STRT$,7)
2490 GOTO 3020
2500 RETURN
2510 REM ------------------------------------------------------------
2515 REM OPTION 2 : STORAGE OF FURNACE DATA
2520 GOSUB 3080
2530 LOCATE RC,CC,1 :INPUT "  Filename for temp data (0 to close)
:",IN$
2540 IF IN$="" THEN GOTO 2570
2550 IF IN$="0" THEN CLOSE #1 :TP=0 :GOTO 2570
2560 IF TP=0 THEN OPEN IN$ FOR OUTPUT AS #1 :TP=1 :ELSE ERROR 203
2570 LOCATE RC,CC :INPUT "Filename for gas flow data (0 to close)
:",IN$
2580 IF IN$="" THEN GOTO 2610
2590 IF IN$="0" THEN CLOSE #2 :GF=0 :GOTO 2610
2600 IF GF=0 THEN OPEN IN$ FOR OUTPUT AS #2 :GF=1 :ELSE ERROR 203
2610 LOCATE RC,CC :INPUT " Filename for growth data (0 to close)
:",IN$
2620 IF IN$="" THEN GOTO 2650
2630 IF IN$="0" THEN CLOSE #3 :GD=0 :GOTO 2650
2640 IF GD=0 THEN OPEN IN$ FOR OUTPUT AS #3 :GD=1 :ELSE ERROR 203
2650 IF SC=0 AND SIM=0 AND SAV=0 THEN STRT$=TIME$ :SH=FNTM(STRT$,1)
:SM=FNTM(STRT,4) :SS=FNTM(STRT$,7)
2660 IF TP=0 AND GF=0 AND GD=0 THEN SAV=0 ELSE SAV=1
2670 IF SAV=0 AND SIM=0 THEN EH=0:EM=0:ES=0: EHP=0:EMP=0:ESP=0
2680 GOTO 3020
2690 RETURN
2700 REM ------------------------------------------------------------
2705 REM OPTION 3 : CONTROL BASED ON REAL TIME SIMULATION
2710 GOSUB 3080
2720 LOCATE RC,CC,1 :INPUT "Required oxide thickness (A): ",IN$
2730 IF IN$="" THEN GOTO 2760
2740 OXSPA=VAL(IN$) :OXSP=OXSPA/10000
2745 PDT=21/60   'Time interval for predicting future oxide growth
2750 IF OXSP=0 THEN CT$="  OFF  " :CT=0 :ELSE CT$=IN$+"  " :CT=1
2760 GOTO 3020
2770 RETURN
2780 REM ------------------------------------------------------------
2785 REM OPTION 4 : DELAY START UNTIL SPECIFIED DPC STEP
2790 GOSUB 3080
2800 LOCATE RC,CC,1 :INPUT "Clock start - 0=Immediate 1=DPC :",IN$
2810 IF IN$="" THEN GOTO 2870
2820 IF IN$<>"0" AND IN$<>"1" THEN 2870 ELSE SC=VAL(IN$)
2830 IF SC=0 THEN SC$="IMM  " :GOTO 2870
2840 LOCATE RC,CC :INPUT "                Step change at DPC :",STP
2850 SC$="DPC"+STR$(STP)
2860 WATE=1
```

```
2870 GOTO 3020
2880 RETURN
2890 REM ------------------------------------------------------------
2895 REM RESET SYSTEM VARIABLES AND TIMER
2900 GOSUB 3080
2910 LOCATE RC,CC,1 :INPUT "Restart - Are you sure ? (Y/N) :",IN$
2920 IF IN$="N" OR IN$="n" THEN GOTO 2990
2930 OXTHI=0 :OXTHIA=0
2940 EHP=0 : EMP=0 :ESP=0
2950 EH=0 :EM=0 :ES=0
2960 IF SC THEN WATE=1 :GOTO 2980
2970 STRT$=TIME$ :SH=FNTM(STRT$,1) :SM=FNTM(STRT$,4)
:SS=FNTM(STRT$,7)
2975 WATE=0
2980 IF DI4 THEN DI4=0 :A$="IO" :GOSUB 5160
2990 GOSUB 3020
3000 REM GOTO 200
3010 RETURN
3020 REM ------------------------------------------------------------
3025 REM RETURN TO SCREEN POSITION AND RE-ENABLE FUNCTION KEYS
3030 LOCATE 25,16,0 :PRINT SIM$; :LOCATE 25,28 :PRINT TP;GF;GD;
3035 LOCATE 25,43 :PRINT CT$; :LOCATE 25,53 :PRINT SC$;
3040 LOCATE RC,1 :PRINT STRING$(79,32);
3050 FOR I=1 TO 5 :KEY(I) ON :NEXT I
3060 LOCATE SAVER,SAVEC
3070 RETURN
3080 REM ------------------------------------------------------------
3085 REM SAVE SCREEN POSITION AND DISABLE FUNCTION KEYS
3090 FOR I=1 TO 5 :KEY(I) STOP :NEXT I
3100 SAVER=CSRLIN  :IF SAVER=0 THEN SAVER=1
3110 SAVEC=POS(I)  :IF SAVEC=0 THEN SAVEC=1
3120 LOCATE RC,1 :PRINT STRING$(79,32)
3130 RETURN
3140 REM ************************************************************
3150 REM ERROR HANDLING ROUTINES
3160 REM
3170 LOCATE RC,CC :PRINT "** ERROR";ERR;"**";
3180 BEEP
3190 IF ERR=7 THEN PRINT ERL; :RESUME 200
3195 REM Attempt to re-open file
3200 IF ERR<>203 AND ERR<>55 THEN 3280
3210 PRINT " File already open and storing data";
3220 FOR I=1 TO PAUSE : NEXT I
3230 LOCATE RC,1 :PRINT STRING$(79,32)
3240 'RESUME NEXT
3250 IF ERL=2560 THEN RESUME 2570
3260 IF ERL=2600 THEN RESUME 2610
3270 IF ERL=2640 THEN RESUME 2650
3275 REM Disk full
3280 IF ERR<>61 THEN GOTO 3340
3290 PRINT " Disk full - Data storage stopped
3300 SAV=0 :TP=0 :GF=0 :GD=0
3310 CLOSE
```

```
3320 LOCATE 25,28 :PRINT TP;GF;GD;
3330 RESUME 260
3335 REM Time interval too long
3340 IF ERR<>202 THEN 3410
3350 PRINT " Time interval too great - ";
3360 IF SAV THEN PRINT "Data storage stopped"; :SAV=0 :TP=0 :GF=0
:GD=0
3370 IF SIM THEN PRINT " Simulation stopped"; :SIM$="OFF" :SIM=0
3380 CLOSE
3390 LOCATE 25,26 :PRINT SIM$; :LOCATE 25,28 :PRINT TP;GF;GD
3400 RESUME 200
3405 REM Other errors
3410 CLS
3420 PRINT :PRINT
3430 PRINT "**** FATAL ERROR ****"
3440 PRINT "Error at line :";ERL;"  Error code :";ERR
3450 PRINT "Program stopped"
3460 CLOSE
3470 LOCATE 7,1,1 :ON ERROR GOTO
3480 END
3490 RETURN
3500 REM ************************************************************
3510 REM DELAYED START
3515 REM Check current DPC step and start timer if necessary
3517 REM
3520 IF D(11)<>STP THEN GOTO 3550
3530 STRT$=TIME$ :SH=FNTM(STRT$,1) :SM=FNTM(STRT$,4)
:SS=FNTM(STRT$,7)
3540 WATE=0
3550 RETURN
3560 REM ************************************************************
3570 REM RESET
3575 REM
3580 OXTHI=0 :OXTHIA=0
3590 EHP=0 :EMP=0 :ESP=0
3600 EH=0 :EM=0 :ES=0
3610 IF SC THEN WATE=1 :GOTO 3625
3620 STRT$=TIME$  :SH=FNTM(STRT$,1)  :SM=FNTM(STRT$,4)
:SS=FNTM(STRT$,7)
3625 FOR I=1 TO 5 :KEY(I) ON :NEXT I
3630 RETURN
3640 REM ************************************************************
3650 REM WRITE DATA TO DISK
3660 REM
3670 IF WATE THEN GOTO 3720
3680 IF SIM=0 THEN GOSUB 4630
3690 IF TP THEN PRINT #1, USING "##  "; EH,EM,ES;  :PRINT #1, USING
"####.#   ";FNTD(153),FNTD(141),FNTD(143)
3700 IF GF THEN PRINT #2, USING "##  "; EH,EM,ES;  :PRINT #2, USING
"####.#   ";FNGD(31),FNGD(33),FNGD(35),FNHD(37)
3710 IF GD THEN PRINT #3, USING "##  "; EH,EM,ES;  :PRINT #3, USING
"#####.#";OXTHIA
3720 RETURN
```

```
3730 REM ************************************************************
3740 REM OXIDE GROWTH SIMULATION
3750 REM
3760 IF WATE THEN RETURN
3765 REM
3770 REM Determine Oxidation Ambient
3775 IF D(30)<>DPREV THEN GOSUB 2120  'Get gas status (if changed)
3780 OXTYPE=2        'Non-oxidising ambient
3790 IF G2$="ON " AND FNGD(33)>.2 THEN OXTYPE=0      'Dry Oxidation
3800 IF G3$="ON " AND FNGD(35)>.2 THEN OXTYPE=1      'Wet Oxidation
3804 REM
3805 REM If non-oxidising ambient or wafers not in furnace then no
oxide growth
3810 IF OXTYPE=2 OR FNBD(19)<800 THEN GOSUB 4630  :RETURN
3815 REM
3816 REM Determine Oxidant Partial Pressure
3820 IF OXTYPE THEN PRES=2*FNGD(35)/(FNGD(35) + 2*(FNGD(33) +
FNGD(31) + FNHD(37))) :ELSE PRES=FNGD(33)/(FNGD(33)+FNGD(31))
3830 REM
3835 REM Determine Oxidation Temperature (=Paddle Temperature)
3840 TEMP=FNTD(141)
3845 REM
3846 REM Determine HCl Enhancement Factors
3850 IF OXTYPE OR G4$="OFF" THEN HCLPC=0 :LLEN=1 :PPEN=1
 :GOTO 3900 'No Enhancement
3855 HCLPC=FNHD(37)*100/(FNGD(33)+FNGD(31)+FNHD(37))
3860 IF HCLPC<=0 THEN HCLPC=0 :LLEN=1 :PPEN=1 :GOTO 3900
3861 IF HCLPC>10 THEN HCLPC=10
3865 HC=CINT(HCLPC*2)  :TC=INT(TEMP/100)-8
3870 IF TC<1 THEN LLEN=LUTL(HC,1) :PPEN=LUTP(HC,1)  :GOTO 3900
3875 IF TC>=3 THEN LLEN=LUTL(HC,3) :PPEN=LUTP(HC,3)  :GOTO 3900
3880 LLEN=LUTL(HC,TC)+(TEMP/100-TC-8)*(LUTL(HC,TC+1)-LUTL(HC,TC))
3890 PPEN=LUTP(HC,TC)+(TEMP/100-TC-8)*(LUTP(HC,TC+1)-LUTL(HC,TC))
3895 REM
3900 REM Determine Time Step
3905 GOSUB 4630
3910 DT=DM+DS/60
3915 REM
3916 REM Oxidation Model
3920 GOSUB 3980
3930 XX=2*OXTHIA+A
3940 DOX=((B/XX)+THINRATE)*DT
3950 OXTHIA=OXTHIA+DOX
3960 OXTHI=OXTHIA/10000    'Oxide Thickness in um
3970 RETURN
3980 REM ************************************************************
3990 REM DETERMINATION OF GROWTH PARAMETERS
4000 REM
4010 REM Data for <100> Si
4015 REM EAP,EAL in eV ;CP in A^2/min ;CL in A/min
4020 REM Experimental Data for Wet Oxidation
4030 REM SUPREM III Data for Dry Oxidation
4040 IF OXTYPE THEN CL=2.83EXP09 :EAL=1.88 :CP=2.15EXP10 :EAP=1.23
```

```
:ELSE CL=0.62EXP09 :EAL=2.0 :CP=1.29EXP09 :EAP=1.23
4050 REM
4060 REM Calculation of Linear and Parabolic Rate Constants
4070 TEMPK=TEMP+273.16
4080 KT=8.61708E-05*TEMPK
4090 RATEL=CL*EXP(-EAL/KT)
4100 RATEP=CP*EXP(-EAP/KT)
4105 REM
4110 IF OXTYPE THEN RATEL=RATEL*PRES*ALPHA*LLEN :THINRATE=0
 :GOTO 4150
4120 RATEL=RATEL*(PRES^.75)*ALPHA*LLEN
4125 REM
4126 REM Calculate Thin Rate Enhancement (Dry Oxidation)
4130 IF ORENT THEN KK=6.58EXP10*EXP(-2.33/KT) :L=78
4135 ELSE KK=7.48EXP10*EXP(-2.38/KT) :L=69
4140 THINRATE=KK*EXP(-OXTHIA/L)
4145 REM
4146 REM Calculate A and B
4150 B=PRES*RATEP*PPEN
4160 A=B/RATEL
4170 RETURN
4180 REM ************************************************************
4190 REM LOOK-UP TABLE FOR HCL ENHANCEMENT OF RATE CONSTANTS
4200 REM
4210 DIM LUTL(21,3),LUTP(21,3)
4220 RESTORE 4400
4230 FOR I=0 TO 20
4240 FOR J=1 TO 3
4250 READ LUTL(I,J)
4260 NEXT J
4270 FOR J=1 TO 3
4280 READ LUTP(I,J)
4290 NEXT J
4300 NEXT I
4310 REM
4340 REM Enhancement factors for HCl addition to dry ambient
4350 REM Data only available for Temp=900,1000,1100°C
4360 REM Data for %HCL = 0,½,1,1½,.....10
4370 REM
4380 REM              LINEAR                    PARABOLIC
4390 REM      900      1000     1100     900      1000     1100
4400 DATA   1,       1,       1,       1,       1,       1
4410 DATA   1.48,    1.32,    1.32,    1.05,    1.34,    1.21
4420 DATA   1.89,    1.49,    1.61,    1.11,    1.41,    1.34
4430 DATA   2.04,    1.64,    1.89,    1.16,    1.48,    1.38
4440 DATA   2.04,    1.74,    1.96,    1.21,    1.54,    1.42
4450 DATA   2.04,    1.74,    2.09,    1.27,    1.61,    1.46
4460 DATA   2.04,    1.74,    2.21,    1.32,    1.68,    1.49
4470 DATA   2.04,    1.74,    2.21,    1.37,    1.74,    1.53
4480 DATA   2.04,    1.74,    2.21,    1.42,    1.81,    1.57
4490 DATA   2.04,    1.74,    2.21,    1.48,    1.88,    1.60
4500 DATA   2.04,    1.74,    2.21,    1.53,    1.95,    1.64
4510 DATA   2.04,    1.74,    2.21,    1.58,    2.02,    1.68
```

```
4520 DATA   2.04,    1.74,    2.21,    1.64,    2.08,    1.71
4530 DATA   2.04,    1.74,    2.21,    1.69,    2.15,    1.75
4540 DATA   2.04,    1.74,    2.21,    1.74,    2.22,    1.79
4550 DATA   2.04,    1.74,    2.21,    1.80,    2.28,    1.82
4560 DATA   2.04,    1.74,    2.21,    1.85,    2.35,    1.86
4570 DATA   2.04,    1.74,    2.21,    1.90,    2.42,    1.90
4580 DATA   2.04,    1.74,    2.21,    1.95,    2.49,    1.94
4590 DATA   2.04,    1.74,    2.21,    2.01,    2.56,    1.97
4600 DATA   2.04,    1.74,    2.21,    2.06,    2.62,    2.01
4610 REM
4620 RETURN
4630 ************************************************************
4640 REM CALCULATE TOTAL ELAPSED TIME AND INCREMENTAL TIME STEP
4650 REM
4660 FT$=TIME$ :ES=FNTM(FT$,7)-SS :EM=FNTM(FT$,4)-SM
:EH=FNTM(FT$,1)-SH
4670 IF ES<0 THEN EM=EM-1 :ES=ES+60
4680 IF EM<0 THEN EH=EH-1 :EM=EM+60
4690 IF EH<0 THEN EH=EH+24
4700 DH=EH-EHP
4710 DM=EM-EMP
4720 DS=ES-ESP
4730 IF DS<0 THEN DM=DM-1 :DS=DS+60
4740 IF DM<0 THEN DH=DH-1 :DM=DM+60
4750 IF DH<0 THEN DH=DH+24
4760 IF DH<>0 THEN ERROR 202
4770 EHP=EH :EMP=EM :ESP=ES
4780 RETURN
4790 ************************************************************
4800 REM CONFIGURE IEEE-488 INTERFACE
4810 REM
4830 GOSUB 4900
4840 DI4=0
4850 DEF SEG=IEEE
4860 CALL IFC
4870 A$="QOP"+"" :GOSUB 5160
4880 A$="00000000" :GOSUB 5160
4890 RETURN
4900 REM ------------------------------------------------------
4910 REM DEFINE VARIABLES REQUIRED BY MSDOS IEEE ROUTINES
4920 REM
4930 DEF SEG=0
4940 REM Get starting address
4950 IEEE=PEEK(641)*256+PEEK(640)
4960 REM Offset addresses of subroutines
4970 NUI=&H103
4980 NUO=&H106
4990 CHI=&H109
5000 CHO=&H10C
5010 FAS=&H10F
5020 SRQ=&H112
5030 CCOM=&H115
5040 AATN=&H118
```

```
5050 AFL=&H11B
5060 PPL=&H11E
5070 IFC=&H121
5080 REN=&H191
5090 RFL=&H194
5100 PA%=13
5110 SA%=0
5120 TC%=10     'Terminating character = LF
5130 TOUT%=10000     'Timeout
5140 FL%=4     'Flags: Add TC and set EOI after last character
5150 RETURN
5160 ************************************************************
5170 REM Routine to send characters on IEEE-488 bus
5180 REM
5190 DEF SEG=IEEE
5200 CALL CHO (TOUT%,ST%,FL%,TC%,A$,SA%,PA%)
5210 DEF SEG
5220 RETURN
5230 ************************************************************
5240 REM CONTROL DPC
5250 REM
5260 IF SIM=0 THEN RETURN
5270 IF DI4 THEN RETURN
5274 PXX=2*OXTHIA+A
5275 PDOX=((B/PXX)+THINRATE)*PDT
5276 POXTHIA=OXTHIA+PDOX
5280 IF POXTHIA<OXSPA THEN RETURN
5290 DI4=1
5300 A$="I2" :GOSUB 5160
5310 RETURN
```

REFERENCES

1. P. Burggraaf, "CAM Software - Part 1 : Choices and Capabilities," Semiconductor International, (June 1987) p.56.

2. P. Burggraaf, "CAM Software - Part 2 : Implementation and Expectations," Semiconductor International, (July 1987) p.63.

3. B. Johnston, "Process Control in Wafer Fabrication," Solid State Technology, (May 1986) p.195.

4. Book of Semiconductor Standards, Vol. 2, Equipment Division, SEMI (1983).

5. C.P. Ho, J.D. Plummer, S.E. Hansen and R.W. Dutton, "VLSI Process Modeling - SUPREM III," IEEE Trans. Electron Devices, ED-30, (1983) p.1438.

6. B.E. Deal and A.S. Grove, "General Relationship for the Thermal Oxidation of Silicon," J. Appl. Phys., 36, (1965) p.3770.

7. W. Kern and G.L. Schnable, "Low-Pressure Chemical Vapour Deposition for VLSI Processing - A Review," IEEE Trans. Electron Devices, ED-26, (1977) p.647.

8. W.A. Brown and T.I. Kamins, "An Analysis of LPCVD System Parameters for Polysilicon, Silicon Nitride and Silicon Dioxide Deposition," Solid State Technology, (July 1979) p.51.

9. P.S. Burggraaf, "Advances in Metallisation," Semiconductor International, (November 1985) p.72.

10. D.B. Fraser, "Metallisation," in VLSI Technology, Ed. S.M. Sze, McGraw Hill, Singapore (1983) p.347.

11. T.E. Siedel, "Ion Implantation," in VLSI Technology, Ed. S.M. Sze, McGraw Hill, Singapore (1983) p.219.

12. J. Sainsbury, "Application of Ion Implantation in Semiconductor Processing," Solid State Technology, (November 1986) p.31.

13. J.C.C. Tsai, "Diffusion," in VLSI Technology, Ed. S.M. Sze, McGraw Hill, Singapore (1983) p.169.

14. J.K. Hassan and H.G. Sarkary, "Lithography for VLSI : An Overview ," Solid State Technology, (May 1982) p.49.

15. M.C. King, "Principles of Optical Lithography," in VLSI Electronics Microstructure and Science, (Vol. 1), Ed. N.G. Einspruch, Academic Press, London (1981) p.42.

16. P.S. Burggraaf, "Advances Keep Optical Aligners Ahead of Industry Needs," Semiconductor International, (February 1984) p.88.

17. C.J. Mogab, "Dry Etching," in VLSI Technology, Ed. S.M. Sze, McGraw Hill, Singapore (1983) p.303.

18. P.S. Burggraaf, "Wet Etching Today,"Semiconductor International, (February 1983) p.48.

19. D.L. Tolliver, "Plasma Processing in Micro-electronics - Past, Present and Future," Solid State Technology, (November 1980) p.99.

20. J.A. Appels, E. Kooi, M.M. Paffen, J.J. Schlorj and W.H.C.G. Verkuylen, "Local Oxidation of Silicon and its Application in Semiconductor Technology," Philips Res. Repts., 25, (1970) p.118.

21. R.J. Kriegler, Y.C. Cheng and D.R. Colton, "The Effect of HCl and Cl_2 on the Thermal Oxidation of Silicon," J. Electrochem. Soc., 119, (1972) p.388.

22. T. Hattor, "Chlorine Oxidation and Annealing in the Fabrication of High Performance LSI Devices," Solid State Technology, (July 1982) p.83.

23. B.E. Deal, M. Sklar, A.S. Grove and E.H. Snow, "Characteristics of the Surface State Charge (Q_{ss}) of Thermally Oxidised Silicon," J. Electrochem. Soc., 124, (1967) p.266.

24. N. Tsubouchi, "Oxidation Technology - The Application of High Pressure Oxidation to VLSI," in Handbook of Advanced Semiconductor Technology and Computer Systems, Ed. G. Rabbat, Van Nostrand Reihhold Company Inc., New York (1988) p.73.

25. E. Bussman, "High Pressure Oxidation in n-channel MOS Technology," Semiconductor International, (April 1983) p.162.

26. A.K. Ray and A. Reisman, "The Formation of SiO_2 in an RF Generated Oxygen Plasma," J. Electrochem. Soc., 128, (1981) p.2466.

27. Physics of Semiconductor Devices (2nd Edition), Ed. S.M. Sze, Wiley, New York (1981) Chapter 8, p.431.

28. J.M. Robertson, "Microelectronic Circuit Fabrication," in Impact of Microelectronics Technology, Ed. M. Jack, Edinburgh University Press, Edinburgh (1982).

29. D.M. Brown, "The Self Registered MOSFET : A Brief Review," Solid State Technology, (April 1972) p.33.

30. T.C. Wu, W.T. Stacy and K.N. Ritz, "The Influence of the LOCOS Processing Parameters on the Slope of the Birds Beak Structure," J. Electrochem. Soc., 130, (1983) p.1563.

31. R.Y. Chiu, J.L. Moll, K.T. Cham, J. Lin, C. Lage, S. Angelos and R.L. Tillman, "The Sloped Wall SWAMI - A Defect-free Zero Birds Beak Local Oxidation Process for Scaled VLSI Technology," IEEE Trans. Electron Devices, ED-30, (1983) p.1506.

32. K.L. Wang, S.A. Saller, W.L. Hunter, P.K. Chatterjee and P. Yang, "Direct Moat Isolation for VLSI," IEEE Trans. Electron Devices, ED-29, (1982) p.541.

33. S.Y. Chiang, K.M. Chan, D.W. Wenocur, A. Hiu and R.D. Rung, "Trench Isolation Technology for MOS Applications," Proceedings of the First International Symposium on VLSI Science and Technology, Eds. K.E. Bean and G.A. Rozyonyi ,Vol. 82-7, (1982) p.339.

34. R.T. Fulks, C.J. Russo, P.R. Hanley and T.I. Kamins, "Rapid Isothermal Annealing of Ion Implant Damage using a Thermal Radiation Source," Appl. Phys. Lett., 39, (1981) p.604.

35. S.R. Wilson, W.M. Paubon and R.B. Gregory, "Rapid Annealing Technology for Future VLSI," Solid State Technology, (June 1985) p.185.

36. H.J. Geipel Jr., N. Hsieh, M.H. Ishaq, C.W. Koburger and F.R. White, "Composite Silicide Gate Electrodes - Interconnections for VLSI Device Technologies," IEEE Trans. Electron Devices, ED-27, (1980) p.1409.

37. M. Chan, "CMOS Carrier Protection with LDD," Semiconductor International, (April 1988) p.78.

38. S. Oruga, P.J. Tsang, W.W. Walker, D.L. Critchlow and J.F. Shepard, "Design and Characteristics of the Lightly Doped Drain-Source (LDD) Insulated Gate Field-effect Transistor," IEEE Trans. Electron Devices, ED-27, (1980) p.1359.

39. E.A. Nicollian and J.R. Brews, MOS Physics and Technology, Wiley (1982).

40. J. Kassabov, D. Dimitrov and J. Kaprinova, "Inversion Electron Mobility in Si-SiO_2 Structures Oxidised at Low and High Temperatures," Microelectronics Journal, Vol. 17, No. 5, (1986) p.15.

41. C. Hashimoto, S. Muramota, N. Shiono and O. Nakajima, "A Method of Forming Thin Highly Reliable Gate Oxides - Two Step HCl Oxidation," J. Electrochem. Soc., 127, (1980) p.127.

42. Y.C. Cheng and B.Y. Liu, "Oxidation Characteristics and Electrical Properties of Low Pressure Dual TCE Oxides," J. Electrochem. Soc., 131, (1984) p.354.

43. A. Bhattacharyya, C. Vorst and A.H. Carim, "A Two Step Oxidation Process to Improve the Electrical Breakdown Properties of Thin Oxides," J. Electrochem. Soc., 132, (1985) p.1900.

44. J. Nulman, J.P. Krusius and A. Gat, "Rapid Thermal Processing of Thin Gate Dielectrics. Oxidation of Silicon," IEEE Electron Device Letters, EDL-6, (1985) p.205.

45. M.M. Moslehi, S.C. Shatas and K.C. Saraswat, "Rapid Thermal Oxidation of Silicon," Extended Abstracts of the 169th Meeting of the Electrochemistry Society, Abstract No. 155, Vol. 86-1, (1986) p.217.

46. R. Rajsuman, R. Singh and J. Nulman, " Theory of Thermal and Rapid Isothermal Oxidation of Silicon," Extended Abstracts of the 169th Meeting of the Electrochemistry Society, Abstract No. 265, Vol. 86-1, (1986) p.389.

47. J. Nulman, "Thin High Quality Silicon Dioxide Films Grown by Rapid Thermal Processing," Extended Abstracts of the 169th Meeting of the Electrochemistry Society, Abstract No. 259, Vol. 86-1, (1986) p.378.

48. T. Ito, T. Nakamura and H. Ishikawa, "Advantages of Thermal Nitride and Nitroxide Gate Films in VLSI Processing," IEEE Journal of Solid State Circuits, SC-17, (1982) p.128.

49. M. Parikh and U. Kaempf, "SMIF A Technology for Wafer Cassette Transfer," Solid State Technology, (July 1984) p.111.

50. A. Chestnut, P.H. Singer and K. Kearney, "Robotic Systems Enhance Manufacturing Efficiency," Semiconductor International, (October 1988) p.58.

51. C. Murray, "Planning for Inter-equipment Transport Automation," Semiconductor International, (September 1985) p.74.

52. D.A. Antoniadis and R.W. Dutton, "Models for Computer Simulation of Complete IC Fabrication Processes," IEEE Trans. Electron Devices, ED-26, (1979) p.490.

53. H. Ryssel, K. Habsberger, K. Hoffman, G. Prinke, R. Dumcko and A. Sachs, "Simulation of Doping Processes," IEEE Solid State Circuits, SC-15, (1980) p.549.

54. D. Chin, M.R. Kump, H.G. Lee and R.W. Dutton, "Process Design using Two-Dimensional Process and Device Simulators," IEEE Trans. Electron Devices, ED-29, (1982) p.336.

55. C.S. Rafferty, M.E. Law and R.W. Dutton, "Two-Dimensional Process Modelling and SUPREM IV," Paper presented at the 2nd International Conference on Simulation of Semiconductor Devices and Processes, University College of Swansea, Swansea, July 21-23 1986.

56. P.J. Jorgenson, "Effect of an Electric Field on Silicon Oxidation," J. Chem. Phys., 37, (1962) p.874.

57. J.R. Ligenza and W.G. Spitzer, "The Mechanisms for Silicon Oxidation in Steam and Oxygen," J. Phys. Chem. Solids, 14, (1980) p.131.

58. E.A. Irene and D.W. Dong, "Silicon Oxidation Studies: The Oxidation of Heavily B- and P- Doped Single Crystal Silicon," J. Electrochem. Soc., 125, (1978) p.1146.

59. Z.M. Ling, L.H. Dupas and K.M. DeMayer, "Sensitivity Analysis of the Deal-Grove Model Parameters. Problem Definition and Strategy Considerations," J. Electrochem. Soc., 135, (1988) p.1245.

60. B.E. Deal, "Thermal Oxidation Kinetics of Silicon in Pyrogenic H_2O and 5% HCl/H_2O Mixtures," J. Electrochem. Soc., 125, (1978) p.576.

61. P.S. Flint, "The Rates of Oxidation of Silicon," Extended Abstracts of The Electrochemistry Society, Abstract No. 94, Spring Meeting, Los Angeles, May 6-10 1962.

62. R.R. Razouk, L.N. Lie and B.E. Deal, "Kinetics of High Pressure Oxidation of Silicon in Pyrogenic Steam," J. Electrochem.Soc., 128, (1981) p.2214.

63. L.N. Lie, R.R. Razouk and B.E. Deal, "High Pressure Oxidation of Silicon in Dry Oxygen," J. Electrochem. Soc., 129, (1982) p.2828.

64. Y.J. Van der Meulen, "Kinetics of Thermal Growth of Ultra-Thin Layers of SiO_2 on Silicon. Part 1 Experiment," J. Electrochem. Soc., 119, (1972) p.530.

65. R. Ghez and Y.J. Van der Meulen, "Kinetics of Thermal Growth of Ultra-Thin layers of SiO_2 on Silicon. Part 2 Theory," J. Electrochem. Soc., 119, (1972) p.1100.

66. J. Blanc, "A Revised Model for the Oxidation of Si by Oxygen," Appl. Phys. Lett., 33, (1978) p.424.

67. J.W. Colby and L.E. Katz, "Boron Segregation at the $Si-SiO_2$ Interface as a Function of Temperature and Orientation," J. Electrochem. Soc., 123, (1976) p.409.

68. R.B. Fair and J.C.C. Tsai, "Theory and Direct Measurement of Boron Segregation in SiO_2 during Dry, Near Dry and Wet O_2 Oxidation," J. Electrochem. Soc., 125, (1978) p.2050.

69. A.S. Grove, O. Leistiko and C.T. Sah, "Redistribution of Acceptor and Donor Impurities during Thermal Oxidation of Silicon," J. Appl. Phys., 35, (1964) p.2695.

70. B.E. Deal and M. Sklar, "Thermal Oxidation of Heavily Doped Silicon," J. Electrochem. Soc., 112, (1965) p.430.

71. C.P. Ho, J.D. Plummer, J.D. Meindl and B.E. Deal, "Thermal Oxidation of Heavily Phosphorous Doped Silicon," J. Electrochem. Soc., 125, (1978) p.665.

72. E.P. Eer Nisse, "Stress in Thermal SiO_2 during Growth," Appl. Phys. Lett., 35, (1979) p.8.

73. E.A. Irene, E. Tierney and J. Angilello, "A Viscous Flow Model to Explain the Appearance of High Density SiO_2 at Low Oxidation Temperatures," J. Electrochem. Soc., 129, (1982) p.2549.

74. E.A. Irene, D.W. Dong and R.J. Zeto, "Residual Stress, Chemical Etch Rate, Refractive Index and Density Measurements on SiO_2 Films prepared using High Pressure Oxygen," J. Electrochem. Soc., 127, (1980) p.396.

75. E.A. Taft, "The Optical Constants of Silicon and Dry Oxygen Oxides of Silicon at 5461Å," J. Electrochem. Soc., 125, (1978) p.968.

76. E.A. Taft, "Index of Refraction of Steam Grown Oxides on Silicon," J. Electrochem. Soc., 127, (1980) p.994.

77. J.K. Srivastava and E.A. Irene, "A Measurement of the Effect of Intrinsic Film Stress on the Overall Rate of Thermal Oxidation of Silicon," J. Electrochem. Soc., 132, (1985) p.2815.

78. L.M. Landsberger and W.A. Tiller, "Stress Relaxation Technique for Thermally Grown SiO_2," Appl. Phys. Lett., 49, (1986) p.143.

79. W.A. Tiller, "On the Kinetics of the Thermal Oxidation of Silicon. I. A Theoretical Perspective," J. Electrochem. Soc., 127, (1980) p.619.

80. W.A. Tiller, "On the Kinetics of the Thermal Oxidation of Silicon. II. Some Theoretical Evaluations," J. Electrochem. Soc., 127, (1980) p.625.

81. C.P. Ho and J.D. Plummer, "Si/SiO_2 Interface Oxidation Kinetics: A Physical Model for the Influence of High Substrate Doping Levels. I. Theory," J. Electrochem. Soc., 126, (1979) p.1516.

82. C.P. Ho and J.D. Plummer, "Si/SiO_2 Interface Oxidation Kinetics: A Physical Model for the Influence of High Substrate Doping Levels. II. Comparison with Experiment and Discussion," J. Electrochem. Soc., 126, (1979) p.1523.

83. S.M. Hu, "Formation of Stacking Faults and Enhanced Diffusion in the Oxidation of Silicon," J. Appl. Phys., 45, (1974) p.1567.

84. D.A. Antoniadis, A.G. Gonzalez and R.W. Dutton, "Boron in Near Intrinsic <100> and <111> Silicon under Inert and Oxidising Ambients - Diffusion and Segregation," J. Electrochem. Soc., 125, (1978) p.813.

85. D.O. Rayleigh, "Transport Processes in the Thermal Oxidation of Silicon," J. Electrochem. Soc., 113, (1966) p.782.

86. T.G. Mills and F.A. Kroger, "Electrical Conduction at Elevated Temperatures in Thermally Grown SiO_2 Films," J. Electrochem. Soc., 120, (1972) p.1582.

87. E. Harari, "Dielectric Breakdown in Electrically Stressed Thin Films of Thermal SiO_2," J. Appl. Phys., 49, (1978) p.2478.

88. E. Taft and L. Cordes, "Optical Evidence for a Silicon - Silicon Oxide Interlayer," J. Electrochem. Soc., 126, (1979) p.131.

89. H.Z. Massoud, J.D. Plummer and E.A. Irene, "Thermal Oxidation of Silicon in Dry Oxygen: Growth Rate Enhancement in the Thin Regime. I. Experimental Results," J. Electrochem. Soc., 132, (1985) p.2685.

90. S.M. Hu, "New Oxide Growth Law and the Thermal Oxidation of Silicon," Appl. Phys. Lett., 42, (1983) p.872.

91. A. Fargeix, G. Ghibaudo and G. Kamarinos, "A Revised Analysis of Dry Oxidation," J. Appl. Phys., 54, (1983) p.2878.

92. A. Fargeix and G. Ghibaudo, "A New Model of Growth including Relaxation of Stress by Viscous Flow," J. Appl. Phys., 54, (1983) p.7153.

93. E.A. Irene, "Silicon Oxidation Studies: Some Aspects of the Initial Oxidation Regime," J. Electrochem. Soc., 125, (1978) p.1708.

94. V.K. Samalam, "Theoretical Model for the Oxidation of Silicon," Appl. Phys. Lett., 47, (1985) p.736.

95. C.J. Han and C.R. Helms, "$^{18}O_2$ Tracer Study of Si Oxidation in Dry O_2 using SIMS," J. Electrochem. Soc., 135, (1988) p.1824.

96. H.Z. Massoud, J.D. Plummer and E.A. Irene, "Thermal Oxidation of Silicon in Dry Oxygen: Growth Rate Enhancement in the Thin Regime. II. Physical Mechanisms," J. Electrochem. Soc., 132, (1985) p.2693.

97. C.L. Claeys, "The Use of Chlorinated Oxides and Intrinsic Gettering Techniques for VLSI Processing," NATO Advanced Study Institute on Process and Device Simulation for MOS-VLSI Circuits, SOGESTA, Urbino, Italy. July 12-23 1982.

98. A. Rohatgi, S.R. Butler and F.J. Feigl, "Mobile Sodium Ion Passivation in HCl Oxides," J. Electrochem. Soc., 126, (1979) p.149.

99. G. Baccarini, M. Severi and G. Soncini, "Electrical Characteristics of MOS Structures on <111> and <100> Orientated N-Type Silicon Influenced by Use of Hydrogen Chloride during Thermal Oxidation," J. Electrochem. Soc., 120, (1973) p.1436.

100. G.J. Declerck, T. Hattori, G.A. May, J. Bedouin and J.D. Meindl, "Some Effects of 'Trichloroethylene Oxidation' on the Characteristics of MOS Devices," J. Electrochem. Soc., 122, (1975) p.436.

101. E.J. Janssens and G.J. Declerck, "The Use of 1,1,1-Trichloroethane as an Optimized Additive to Improve the Silicon Thermal Oxidation Technology," J. Electrochem. Soc., 125, (1978) p.1696.

102. C.M. Osburn, "Dielectric Breakdown Properties of SiO_2 Film Growth in Halogen- and Hydrogen-containing Environments," J. Electrochem. Soc., 121, (1974) p.809.

103. R.S. Ronen and P.H. Robinson, "Hydrogen Chloride and Chlorine Gettering: An Effective Technique for Improving Performance of Silicon Dioxide," J. Electrochem. Soc., 119, (1972) p.747.

104. D.W. Hess and B.E. Deal, "Kinetics of the Thermal Oxidation of Silicon in O_2/HCl Mixtures," J. Electrochem. Soc., 124, (1977) p.7359.

105. R.E. Tressler, J. Stach and D.M. Metz, "Gas Phase Composition Considerations in the Thermal Oxidation of Silicon in Cl-H-O Ambients," J. Electrochem. Soc., 124, (1977) p.607.

106. B.E. Deal, D.W. Hess, J.D. Plummer and C.P. Ho, "Kinetics of the Thermal Oxidation of Silicon in O_2/H_2O and O_2/Cl_2 Mixtures," J. Electrochem. Soc., 125, (1978) p.339.

107. B.E. Deal, A. Hurrle and M.J. Schulz, "Chlorine Concentration Profiles in O_2/HCl and H_2O/HCl Thermal Silicon Oxides using SIMS Measurements," J. Electrochem. Soc., 125, (1978) p.2024.

108. J.R. Monkowski, M.D. Monkowski, I.S.T. Tsong and J. Stach, "Hydrogen/Chlorine Distributions in Silicon Dioxide: Detection and Model," in Silicon Processing, Ed. D.C. Gupton, American Society for Testing and Materials STP804 (1983) p.245.

109. B.R. Singh and P. Balk, "Thermal Oxidation of Silicon in O_2 - Trichloroethylene," J. Electrochem. Soc., 126, (1979) p.1288.

110. T. Hattori, "Thermal Oxidation of Silicon in Wet O_2/TCE Mixtures at 1200°C," J. Electrochem. Soc., 126, (1979) p.1789.

111. P.H. Singer, "Advances in Diffusion Furnace Technology," Semiconductor International, (May 1985) p.276.

112. P.H. Singer, "Trends in Vertical Diffusion Furnaces," Semiconductor International, (April 1986) p.56.

113. P.F. Schmidt, "Furnace Contamination and its Remedies," Solid State Technology, (June 1983) p.147.

114. P.F. Schmidt, "Contamination-Free High Temperature Treatment of Silicon and Other Materials," J. Electrochem. Soc., 130, (1983) p.196.

115. B.D. Foster and R.E. Tressler, "Silicon Processing with Silicon Carbide Furnace Components," Solid State Technology, (October 1984) p.143.

116. R. Iscoff, "Furnace Loading Systems Review," Semiconductor International, (August 1984) p.82.

117. C. Murray, "'Quartzware' for Diffusion Furnaces," Semiconductor International, (April 1986) p.64.

118. J.L. Lambert and C.J. Bayne, "A Suspended Boat Loader based on the Cantilever Principle," Semiconductor International, (April 1983) p.150.

119. Instruction Manual for the Tempress Digital Temperature Controller.

120. Instruction Manual for the Tempress Digital Process Controller.

121. J. Maliakal, D. Fisher Jr. and A. Waugh, "Trends in Automated Diffusion Furnace Systems for Large Wafers," Solid State Technology, (December 1984) p.105.

122. W. Kern and D.A. Puotinen, "Cleaning Solutions Based on Hydrogen Peroxide for Use in Silicon Semiconductor Technology," R.C.A. Review, 31, (1970) p.187.

123. I.K. Bansal, "Autodoping and Particulate Contaminations during Pre-diffusion Chemical Cleaning of Silicon Wafers," Solid State Technology, (July 1986) p.75.

124. A. Butler, Private Communication. Result of an investigation into oxide thickness variation carried out at Plessey Research (Caswell) Ltd., 1985.

125. F.N. Schwettmann, K.L. Chiang and W.A. Brown, "Variation of Silicon Dioxide Growth Rate with Pre-oxidation Clean," Electrochem. Soc. Extended Abstracts, 78-1, No. 276, (1978) p.688.

126. S.I. Raider, R.A. Gdula and J.R. Petrak, "Nitrogen Reaction at a Silicon-Silicon Dioxide Interface," Appl. Phys. Lett., 27, (1975) p.150.

127. J. Ruzyllo, "Effects of Pre-oxidation Ambient in Very Thin Thermal Oxide on Silicon," J. Electrochem. Soc., 133, (1986) p.1677.

128. W.A. Pliskin and E.E. Conrad, "Non-destructive Determination of Thickness and Refractive Index of Transparent Films," IBM J. Res.& Dev., 8, (1964) p.43.

129. F. Reizman and W. Von Gelder, "Optical Thickness Measurement of SiO_2-Si_3N_4 Films on Silicon," Solid State Electronics, 10, (1967) p.625.

130. D.S. Perloff, F.E. Wahl and J.D. Reimer, "Contour Maps Reveal Non-uniformity in Semiconductor Processing," Solid State Technology, (February 1977) p.31.

131. D.S. Perloff, T.S. Hasan and E.R. Blome, "Real-Time Monitoring of Semiconductor Process Uniformity," Solid State Technology, (February 1980) p.81.

132. E. Passaglia, R.R. Stranberg and J. Kruger (Eds), Ellipsometry in the Measurement of Surfaces and Thin Films, NBS Misc. Publ. Washington DC, (1984) p.256.

133. K.H. Zaininger and P. F. Heiman, "The C-V Technique as an Analytical Tool. Part 2," Solid State Technology, (June 1970) p.46.

134. C.G. Shirley, "A Computer Controlled CV Characterization System," Semiconductor International, (July 1982) p.81.

135. A. Lederman, "Vacuum Operated Mercury Probe for CV Plotting and Profiling," Solid State Technology, (August 1981) p.121.

136. Nanometrics Nanospec AFT Model No. 010-0180 Instruction Manual.

137. Applied Materials Instruction Manual for Ellipsometer II.

138. E.H. Nicollian, "Electrically Active Defects in the Si-SiO_2 System - Present Understanding," Proceedings of the Electrochemistry Society Meeting on Semiconductor Silicon, Vol 86-4, (1986) p.437.

139 B.E. Deal, "Standardised Terminology for Oxide Charges Associated with Thermally Oxidised Silicon," J. Electrochem. Soc., 124, (1980) p.979.

140. B.E. Deal, "The Current Understanding of Charges in the Thermally Oxidised Silicon Structure," J. Electrochem. Soc., 121, (1974) p.189C.

141. E.H. Snow, A.S. Grove, B.E. Deal and C.T. Sah, "Ion Transport Phenomena in Insulating Films," J. Appl. Phys., 36, (1965) p.1664.

142. R.R. Razouk and B.E. Deal, "Dependence of Interface State Density on Silicon Thermal Oxidation Process Variables," J. Electrochem. Soc., 126, (1979) p.1467.

143. I.G. McGillivray, "The Measurement of Electrical Parameters and True Impurity Effects in MOS Capacitors," PhD Thesis (1987), University of Edinburgh.

144. A. Bhattacharyya, T. Bril, C. Vorst, B. Westlund and F. Van Roosmalen, "Effect of Dry Etching of a Thermal Oxide on Subsequent Growth and Properties of Thin Oxides," J. Electrochem. Soc., 133, (1986) p.1670.

145. A.H. Carim and A. Bhattacharyyu, "Si/SiO_2 Interface Roughness : Structural Observations and Electrical Consequences," Appl. Phys. Lett., 46, (1985) p.872.

146. N. Lifshitz, "Study of Breakdown Fields of Oxides Grown on Reactive Ion Etched Silicon Surface : Improvement of Breakdown Limits by Oxidation of the Surface," J. Electrochem. Soc., 130, (1983) p.1549.

147. C.M. Osburn and D.W. Ormond, "Dielectric Breakdown in Silicon Dioxide Films on Silicon. II Influence of Processing and Materials," J. Electrochem. Soc., 119, (1972) p.597.

148. D. R. Wolters and J. J. Van der Schoot, "Dielectric Breakdown in MOS Devices Part 1: Defect-Related and Intrinsic Breakdown," Philips J. Res., 40, (1985) p.115.

149. D.R. Wolters and J.J. Van der Schoot, "Dielectric Breakdown in MOS Devices Part 3 : The Damage Leading to Breakdown," Philips J. Res., 40, (1985) p.164.

150. E. Cameron, "Investigation of Parasitic Oxidation," Edinburgh Microfabrication Facility / Plessey Research (Caswell) Ltd., Internal CASE Project Industrial Training Report, (August 1985).

151. A.N. Crandorkar, V.T. Karulker and K.V. Ramanathan, "Dependence of Partial Pressure of H_2O on Pyrogenic Growth of Silicon Dioxide," J. Electrochem. Soc., 132, (1985) p.415.

152. H.Z. Massoud, J.D. Plummer and E.A. Irene, "Thermal Oxidation of Silicon in Dry Oxygen. Accurate Determination of the Kinetic Rate Constants," J. Electrochem. Soc., 132, (1985) p.1745.

153. E.A. Irene, and Y.J. van der Meulen, "Silicon Oxidation Studies: Analysis of SiO_2 Film Growth Data," J. Electrochem. Soc., 123, (1976) p.1380.

154. FIOBUS SC032 IEEE-488 Compatible Digital I/O Module Application Notes.