2-1-1989

Three-Dimensional MOS Process Development

J. A. Friedrich
Purdue University

G. W. Neudeck
Purdue University

Follow this and additional works at: https://docs.lib.purdue.edu/ecetr

https://docs.lib.purdue.edu/ecetr/639

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Three-Dimensional MOS Process Development

J. A. Friedrich
G. W. Neudeck

TR-EE 89-7
February, 1989

School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>v</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>vi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ABSTRACT</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>xi</td>
</tr>
</tbody>
</table>

## CHAPTER 1 - INTRODUCTION

1.1 Background ............................................................. 1  
1.2 Statement of Purpose ........................................... 2  
1.3 Proposed Structure ............................................. 3  
1.4 Thesis Overview ................................................ 5  

## CHAPTER 2 - LITERATURE REVIEW

2.1 Introduction ........................................................... 7  
2.2 Background ........................................................... 7  
2.3 CMOS Circuit Design ............................................. 9  
  2.3.1 Fully Complementary CMOS .................................. 10  
  2.3.2 Clocked CMOS .................................................. 10  
  2.3.3 Complex CMOS Circuits ...................................... 12  
2.4 CMOS Technology .................................................. 15  
2.5 SOI Technologies .................................................. 25  
  2.5.1 Full Insulation by Porous Oxidized Silicon (FIPOS) ..... 25  
  2.5.2 Polysilicon Recrystallization .............................. 27  
  2.5.3 Separation by Implanted Oxygen (SIMOX) ................ 28  
  2.5.4 Lateral Solid-Phase Epitaxial Growth ................... 29  
  2.5.5 Heteroepitaxial SOI ......................................... 29  
  2.5.6 Wafer Bonding and Wafer Thinning ....................... 30
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6 Epitaxial Lateral Overgrowth</td>
<td>30</td>
</tr>
<tr>
<td>2.6.1 Basic Terminology</td>
<td>31</td>
</tr>
<tr>
<td>2.6.2 Transport and Fluid Dynamics</td>
<td>35</td>
</tr>
<tr>
<td>2.6.3 Chemical Reactions in Si-Epitaxy</td>
<td>38</td>
</tr>
<tr>
<td>2.6.4 Applied SEG and ELO</td>
<td>41</td>
</tr>
<tr>
<td>2.7 Summary</td>
<td>46</td>
</tr>
<tr>
<td>CHAPTER 3 - PROCESS DEVELOPMENT I: TEST STRUCTURES</td>
<td>46</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>47</td>
</tr>
<tr>
<td>3.2 3-D CMOS Process Flow</td>
<td>47</td>
</tr>
<tr>
<td>3.3 Test Device Structures</td>
<td>53</td>
</tr>
<tr>
<td>3.3.1 ELO Growth and Planarization</td>
<td>53</td>
</tr>
<tr>
<td>3.3.2 Electronic Device Structures</td>
<td>55</td>
</tr>
<tr>
<td>3.4 Process Simulation</td>
<td>63</td>
</tr>
<tr>
<td>3.4.1 3-D Capacitor Simulation</td>
<td>63</td>
</tr>
<tr>
<td>3.4.2 3-D CMOS Inverter Simulation</td>
<td>68</td>
</tr>
<tr>
<td>CHAPTER 4 - PROCESS DEVELOPMENT II: EXPERIMENTAL</td>
<td>72</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>72</td>
</tr>
<tr>
<td>4.2 Polysilicon</td>
<td>73</td>
</tr>
<tr>
<td>4.3 Oxide Degradation During SEG</td>
<td>76</td>
</tr>
<tr>
<td>4.4 Silicon Epitaxy</td>
<td>88</td>
</tr>
<tr>
<td>4.4.1 Crystalline Quality</td>
<td>88</td>
</tr>
<tr>
<td>4.4.2 Growth Rate Characteristics</td>
<td>96</td>
</tr>
<tr>
<td>4.4.2.1 Reactor Description</td>
<td>96</td>
</tr>
<tr>
<td>4.4.2.2 Reactor Feed Composition Dependence</td>
<td>99</td>
</tr>
<tr>
<td>4.4.2.3 Masking Oxide Dependence</td>
<td>102</td>
</tr>
<tr>
<td>4.4.2.4 Temperature Dependence</td>
<td>109</td>
</tr>
<tr>
<td>4.4.2.5 Total Flow Dependence</td>
<td>112</td>
</tr>
<tr>
<td>4.4.2.6 Area Dependence</td>
<td>113</td>
</tr>
<tr>
<td>4.4.3 Growth Mechanism</td>
<td>115</td>
</tr>
<tr>
<td>4.5 ELO Planarization</td>
<td>119</td>
</tr>
</tbody>
</table>
CHAPTER 5 - ELECTRONIC DEVICES .......................................................123

5.1 Introduction...........................................................................123
5.2 Interface States....................................................................123
5.3 Quasistatic D_{it} Measurement ...........................................126
  5.3.1 Data Acquisition.............................................................127
  5.3.2 Data Analysis.................................................................130
5.4 SOI Interface Characteristics ..............................................136
5.5 3-D PMOS Transistors.........................................................144

CHAPTER 6 - SUMMARY AND PROPOSED FUTURE RESEARCH ......153

LIST OF REFERENCES......................................................................158

APPENDICES..................................................................................178

Appendix A....................................................................................178
Appendix B....................................................................................184
Appendix C....................................................................................192
Appendix D....................................................................................193
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SOI technologies</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>Summary of device and material properties</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>Polyoxide leakage at 5V Bias after 40 min. exposure to SEG ambient at 950°C and 150 torr</td>
<td>77</td>
</tr>
<tr>
<td>4</td>
<td>Average breakdown field $\overline{E}$, defect density $D$, and standard deviation of breakdown field $\sigma$ for the four different groups of 80 nm oxides</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>Preclean experiment matrix</td>
<td>94</td>
</tr>
<tr>
<td>6</td>
<td>Growth experiment matrix</td>
<td>95</td>
</tr>
<tr>
<td>7</td>
<td>Defects delineated by Secco etching</td>
<td>95</td>
</tr>
<tr>
<td>8</td>
<td>Temperature dependence of growth from 0.36 vol % DCS in H₂ at 150 Torr</td>
<td>109</td>
</tr>
<tr>
<td>9</td>
<td>Effect of total flow on growth rate</td>
<td>113</td>
</tr>
<tr>
<td>10</td>
<td>Effect of wafer oxide coverage on growth rate and uniformity</td>
<td>115</td>
</tr>
<tr>
<td>11</td>
<td>Parameters of capacitors fabricated under different processing conditions</td>
<td>140</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 3-D CMOS Inverter</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Compact 3-D CMOS Inverter</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Comparison of operational amplifier circuits in NMOS and CMOS</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Basic CMOS inverter circuit diagram</td>
<td>11</td>
</tr>
<tr>
<td>2.3 CMOS EXOR circuit diagram</td>
<td>11</td>
</tr>
<tr>
<td>2.4 Two-input CMOS NAND gate</td>
<td>13</td>
</tr>
<tr>
<td>2.5 Three-input clocked CMOS NAND gate</td>
<td>13</td>
</tr>
<tr>
<td>2.6 CMOS master-slave flip-flop</td>
<td>14</td>
</tr>
<tr>
<td>2.7 CMOS static RAM</td>
<td>16</td>
</tr>
<tr>
<td>2.8 CMOS content addressable memory</td>
<td>16</td>
</tr>
<tr>
<td>2.9 Twin-tub CMOS inverter cross section</td>
<td>17</td>
</tr>
<tr>
<td>2.10 Recessed LOCOS process flow</td>
<td>19</td>
</tr>
<tr>
<td>2.11 Sealed Interface Local Oxidation (SILO) structure</td>
<td>20</td>
</tr>
<tr>
<td>2.12 BOX fabrication sequence</td>
<td>21</td>
</tr>
<tr>
<td>2.13 Direct moat isolation process sequence</td>
<td>22</td>
</tr>
<tr>
<td>2.14 Trench isolation process flow</td>
<td>23</td>
</tr>
<tr>
<td>2.15 Dielectric isolation by Selective Epitaxial Growth</td>
<td>24</td>
</tr>
<tr>
<td>2.16 Selective epitaxy and epitaxial lateral overgrowth</td>
<td>32</td>
</tr>
<tr>
<td>2.17 Basic model of the epitaxial growth process</td>
<td>33</td>
</tr>
<tr>
<td>2.18 Growth rates of silicon films as a function of temperature for various silicon gas sources</td>
<td>35</td>
</tr>
<tr>
<td>2.19 Development of a boundary layer in gas flowing over a flat plate and expanded view of boundary layer</td>
<td>36</td>
</tr>
<tr>
<td>2.20 Undercutting of the SiO$_2$/Si interface at 1150 °C during a conventional in situ preclean in H$_2$ at 50 Torr</td>
<td>43</td>
</tr>
<tr>
<td>2.21 Arrhenius plot of undercutting length for preclean in H$_2$ at 50 Torr</td>
<td>43</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>3.1 Proposed 3-D CMOS inverter cross section</td>
<td>48</td>
</tr>
<tr>
<td>3.2 3-D CMOS process flow</td>
<td>50</td>
</tr>
<tr>
<td>3.3 Epitaxy and planarization test mask</td>
<td>54</td>
</tr>
<tr>
<td>3.4 3-D capacitor cross section</td>
<td>56</td>
</tr>
<tr>
<td>3.5 3-D capacitor electrical equivalent</td>
<td>56</td>
</tr>
<tr>
<td>3.6 Threshold voltage as a function of doping density</td>
<td>57</td>
</tr>
<tr>
<td>3.7 3-D capacitor mask layout</td>
<td>59</td>
</tr>
<tr>
<td>3.8 PMOS stacked transistor layout</td>
<td>60</td>
</tr>
<tr>
<td>3.9 3-D CMOS inverter layout</td>
<td>60</td>
</tr>
<tr>
<td>3.10 CAM circuit diagram</td>
<td>61</td>
</tr>
<tr>
<td>3.11 Dynamic CAM mask layout</td>
<td>61</td>
</tr>
<tr>
<td>3.12 Electronic device characterization mask</td>
<td>62</td>
</tr>
<tr>
<td>3.13 Simulation of 3-D capacitor, polysilicon after P-implant</td>
<td>64</td>
</tr>
<tr>
<td>3.14 Simulation of 3-D capacitor, polysilicon after oxidation</td>
<td>64</td>
</tr>
<tr>
<td>3.15 Simulation of 3-D capacitor, 860 °C ELO before contact anneal</td>
<td>65</td>
</tr>
<tr>
<td>3.16 Simulation of 3-D capacitor, 860 °C ELO after 920 °C contact anneal</td>
<td>65</td>
</tr>
<tr>
<td>3.17 Simulation of 3-D capacitor, 900 °C ELO before contact anneal</td>
<td>66</td>
</tr>
<tr>
<td>3.18 Simulation of 3-D capacitor, 950 °C ELO before contact anneal</td>
<td>66</td>
</tr>
<tr>
<td>3.19 Simulation of 3-D capacitor, ELO after 1000 °C contact anneal</td>
<td>67</td>
</tr>
<tr>
<td>3.20 Simulation of 3-D CMOS inverter, field oxide region</td>
<td>70</td>
</tr>
<tr>
<td>3.21 Simulation of 3-D CMOS inverter, source/drain region</td>
<td>70</td>
</tr>
<tr>
<td>3.22 Simulation of 3-D CMOS inverter, common drain (SEG) region</td>
<td>71</td>
</tr>
<tr>
<td>3.23 Simulation of 3-D CMOS inverter, gate/ELO region</td>
<td>71</td>
</tr>
<tr>
<td>4.1 Polysilicon growth rate as a function of temperature</td>
<td>74</td>
</tr>
<tr>
<td>4.2 Polysilicon growth rate as a function of doping</td>
<td>74</td>
</tr>
<tr>
<td>4.3 Yield by processing group for oxides of different thickness</td>
<td>79</td>
</tr>
<tr>
<td>4.4 Breakdown field distribution for 80-nm control oxides (group I)</td>
<td>81</td>
</tr>
<tr>
<td>4.5 Breakdown field distribution for 80-nm H₂ bake oxides (group II)</td>
<td>81</td>
</tr>
<tr>
<td>4.6 Breakdown field distribution for 80-nm H₂ + HCl oxides (group III)</td>
<td>82</td>
</tr>
<tr>
<td>4.7 Breakdown field distribution for 80-nm H₂ + HCl + DCS oxides (group IV)</td>
<td>82</td>
</tr>
<tr>
<td>4.8 Regions of silicon dioxide growth as a function of oxygen partial pressure and temperature</td>
<td>83</td>
</tr>
<tr>
<td>4.9 Regions of silicon dioxide growth as a function of water vapor partial pressure and temperature</td>
<td>83</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>4.10</td>
<td>Critical conditions for oxide decomposition by SiO formation to take place</td>
</tr>
<tr>
<td>4.11</td>
<td>Yield for 3-D capacitors with epitaxy grown at 860 °C, 900 °C, and 950 °C</td>
</tr>
<tr>
<td>4.12</td>
<td>Critical temperature vs carrier gas moisture content, with reactor chamber pressure as a parameter</td>
</tr>
<tr>
<td>4.13</td>
<td>SEM of a seed precleaned at 150 Torr, above the required minimum temperature (point 2 in Fig. 4.12)</td>
</tr>
<tr>
<td>4.14</td>
<td>SEM of a seed precleaned at 760 Torr, just below the required minimum temperature (point 1 in Fig. 4.12)</td>
</tr>
<tr>
<td>4.15</td>
<td>SEM of growth at 150 Torr, above the required minimum temperature (point 4 in Fig. 4.12)</td>
</tr>
<tr>
<td>4.16</td>
<td>SEM of growth at 150 Torr, below the required minimum temperature (point 3 in Fig. 4.12)</td>
</tr>
<tr>
<td>4.17</td>
<td>Cross section of the Purdue Gemini I reactor</td>
</tr>
<tr>
<td>4.18</td>
<td>Growth rate vs. HCl/DCS ratio of feed</td>
</tr>
<tr>
<td>4.19</td>
<td>Growth rate vs. HCl^2/DCS ratio of feed</td>
</tr>
<tr>
<td>4.20</td>
<td>Growth rate profile for various feed gas compositions</td>
</tr>
<tr>
<td>4.21</td>
<td>Non-uniformity as a function of feed gas composition</td>
</tr>
<tr>
<td>4.22</td>
<td>Etch rate profile at 950 °C and 150 Torr</td>
</tr>
<tr>
<td>4.23</td>
<td>Oxide thickness characterization test pattern cross section</td>
</tr>
<tr>
<td>4.24</td>
<td>Oxide thickness characterization inverted test pattern</td>
</tr>
<tr>
<td>4.25</td>
<td>Growth rate as a function of masking oxide thickness</td>
</tr>
<tr>
<td>4.26</td>
<td>Growth rate as a function of masking oxide thickness compared to a model where only thermal conduction is considered</td>
</tr>
<tr>
<td>4.27</td>
<td>Bulk (unmasked) silicon growth rate profiles at 920 °C, 950 °C, and 1000 °C</td>
</tr>
<tr>
<td>4.28</td>
<td>Selective epitaxy growth rate profile on unpatterned wafers at 950 °C and 1000 °C</td>
</tr>
<tr>
<td>4.29</td>
<td>Selective epitaxy growth rate profile on patterned wafers at 860 °C, 920 °C, 950 °C, and 1000 °C</td>
</tr>
<tr>
<td>4.30</td>
<td>Effect of total gas flow on selective epitaxial growth rate profile</td>
</tr>
<tr>
<td>4.31</td>
<td>Effect of masking oxide coverage on selective epitaxial growth rate profile</td>
</tr>
<tr>
<td>4.32</td>
<td>Effect of reactor loading on selective epitaxial growth rate profile</td>
</tr>
<tr>
<td>4.33</td>
<td>Growth rate as a function of HCl/DCS ratio at 850 °C and 30 Torr</td>
</tr>
</tbody>
</table>
4.34 Effect of HCl flow on the thickness ratio between a 5 \( \mu \)m and a 50 \( \mu \)m aperture on a patterned wafer ................................................................. 117
4.35 Planarization procedure by resist reflow and plasma etching ................ 120
4.36 Planarization by preferential polishing .......................................................... 122

5.1 Quasi-static C-V measurement instrumentation ............................................. 128
5.2 Quasi-static C-V measurement timing diagram .......................................... 128
5.3 HF-LF C-V measurement instrumentation set-up ....................................... 129
5.4 MOS capacitor small signal equivalent circuit ............................................... 131
5.5 Overetched polysilicon finger structure ....................................................... 134
5.6 Total high-frequency capacitance exhibiting lateral effects ...................... 134
5.7 I-V curve for ELO capacitor with epitaxy grown at 900 ° C .......................... 137
5.8 Capacitance \( C_{ELO} \) for epitaxy grown at 860 ° C ........................................ 139
5.9 Capacitance \( C_{ELO} \) for epitaxy grown at 900 ° C ........................................ 139
5.10 Interface state density \( D_{it} \) as a function of band gap energy
    for epitaxy done at 860 ° C and 900 ° C .................................................. 141
5.11 Capacitance \( C_{ELO} \) for epitaxy grown at 860 ° C,
    after 10 min. anneal in \( H_2 \) at 1000 ° C ................................................. 141
5.12 Capacitance \( C_{ELO} \) for epitaxy grown at 900 ° C
    after 10 min. anneal in \( H_2 \) at 1000 ° C .................................................. 143
5.13 Interface state density \( D_{it} \) as a function of band gap energy
    for epitaxy done at 860 ° C and 900 ° C, after 10 min. anneal
    in \( H_2 \) at 1000 ° C ............................................................................. 143
5.14 Comparison of experimental and ideal LF C-V curve for epitaxy
    grown at 860 ° C ................................................................................. 145
5.15 Comparison of experimental and ideal HF C-V curve for epitaxy
    grown at 860 ° C ................................................................................. 145
5.16 Comparison of experimental and ideal LF and HF C-V curves
    for epitaxy grown at 900 ° C .................................................................. 146
5.17 Scanning electron beam micrograph of a 3-D PMOS stacked
    transistor with 1.4 \( \mu \)m thick overgrowth ................................................. 146
5.18 I-V characteristics of bad 3-D stacked PMOS transistor
    with source-drain short circuit ............................................................... 149
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.19 Scanning electron beam micrograph of a 3-D PMOS stacked transistor with 0.8 μm thick overgrowth</td>
<td>149</td>
</tr>
<tr>
<td>5.20 $I_D$ vs $V_{DS}$ characteristics of field effect modulated 3-D stacked PMOS transistor</td>
<td>151</td>
</tr>
<tr>
<td>5.21 $I_D$ vs $V_G$ characteristics of field effect modulated 3-D stacked PMOS transistor</td>
<td>151</td>
</tr>
<tr>
<td>5.22 I-V characteristics of PMOS-drain/substrate diode</td>
<td>152</td>
</tr>
<tr>
<td>5.23 SPICE simulation output of extreme short channel MOSFET</td>
<td>152</td>
</tr>
</tbody>
</table>
ABSTRACT


A novel MOS technology for three-dimensional integration of electronic circuits on silicon substrates was developed. Selective epitaxial growth and epitaxial lateral overgrowth of monocrystalline silicon over oxidized silicon were employed to create locally restricted silicon-on-insulator device islands.

Thin gate oxides were discovered to deteriorate in ambients typically used for selective epitaxial growth. Conditions of general applicability to silicon epitaxy systems were determined under which this deterioration was greatly reduced.

Selective epitaxial growth needed to be carried out at low temperatures. However, crystalline defects increase as deposition temperatures are decreased. An exact dependence between the residual moisture content in epitaxial growth ambients, deposition pressure, and deposition temperature was determined which is also generally applicable to silicon epitaxy systems.

The dependences of growth rates and growth rate uniformity on loading, temperature, flow rates, gas composition, and masking oxide thickness were investigated for a pancake type epitaxy reactor. A conceptual model was discussed attempting to describe the effects peculiar to selective epitaxial growth.

The newly developed processing steps were assembled to fabricate three-dimensional silicon-on-insulator capacitors. These capacitors were electrically evaluated. Surface state densities were in the order of $10^{11} \text{cm}^{-2}\text{eV}^{-1}$ and
therefore within the range of applicability for a practical CMOS process.

Oxidized polysilicon gates were overgrown with silicon by epitaxial lateral overgrowth. The epitaxial silicon was planarized and source and drain regions were formed above the polysilicon gates in silicon-on-insulator material. The modulation of the source-drain current by bias changes of the buried gate was demonstrated.
CHAPTER 1

INTRODUCTION

1.1 Background

Initial attempts to build active solid-state transistors that operated on a field effect principle were undertaken by J.E. Lilienfeld in the early 1930s. Patents on a metal-oxide-semiconductor (MOS) structure were filed in 1935 by O. Heil and later by D. Kahng and M.M. Atalla who proposed a Si-SiO₂ based field effect transistor. Not until the early 1960s however was the fabrication of MOSFETs successful because of the lack of a controllable and stable silicon surface. With the introduction of the silicon planar process it became possible to create stable insulating layers on top of silicon by means of thermal oxidation. In 1964 the first simple integrated circuits (ICs) with MOS transistors were fabricated.

The first complex circuits were built using p-channel, and later n-channel transistors. In 1962, C.T. Sah and F. Wanlass proposed the pairing of complementary n-channel and p-channel transistors to form low-power integrated circuits. For over a decade after its invention, this technology, which in fact is as much a circuit technology as a processing technology, called "Complementary MOS" or CMOS, had its primary use only in specialty applications where the benefits of low power consumption and high noise immunity outweighed the higher production costs.

With the need for increased device density and shrinking dimensions, NMOS technologies became more and more complex. In the late 1970's it became apparent that power dissipation would set a limit on the level of integration of NMOS circuits. At that time, the complexities of advanced NMOS and CMOS process technologies were comparable, and CMOS became a viable alternative to NMOS which had been dominating the field.
Today CMOS has developed into the technology of choice for Very Large Scale Integration (VLSI) and Ultra Large Scale Integration (ULSI). Currently it is possible to integrate more than 500000 transistors in non-repetitive circuit designs such as microprocessors, and more than 4 million transistors in regular circuit designs such as random access memories.

1.2 Statement of Purpose

The development of process technologies is driven by the demand for more powerful systems, such as computers and signal processing equipment. Current two-dimensional planar CMOS technologies will have reached their performance limits soon, making it necessary to develop more sophisticated processes that eliminate or greatly reduce the key problems found in conventional technologies. Three-dimensional integration, first at the interconnect level, and later at the active device level, has potential to significantly increase the complexity of monolithically integrated systems.

In addition to lithographic restrictions, scaling of CMOS technologies is limited by the required large spacing between transistors of different type, in order to prevent latch-up which can render a circuit temporarily or permanently inoperable. Furthermore, susceptibility to radiation increases as device dimensions shrink. Studies on VLSI interconnects suggest that as the lateral spacing between adjacent metal lines approaches the 0.5 \( \mu \text{m} \) level, the increased mutual capacitance between neighbouring lines and not the intrinsic switching speed of the active devices themselves will become the limiting factor in circuit performance, density and power dissipation for integrated circuits fabricated using two-dimensional planar silicon technology.\(^1\) For these reasons, there has long been interest in silicon-on-insulator (SOI) and three-dimensional technologies, which would imply higher operating speed, lower dynamic power consumption, greater packing density, increased radiation hardness, more layout flexibility, and immunity to latch-up.

However, single crystalline silicon is not easily formed on top of amorphous insulating material such as SiO\(_2\). Crystalline insulators have been devised, such as sapphire and calcium fluorides, which match the lattice spacing of silicon reasonably close. Other methods of creating SOI include zone-melting recrystallization of polysilicon deposited by standard chemical vapor deposition techniques, separation by implanted oxygen, lateral solid-phase epitaxy, and full insulation by porous oxidized silicon. All these methods have in common that they are either prohibitively expensive, lead to poor quality SOI
layers, or are incompatible with common clean room processing technologies.

The primary goal in SOI is to realize perfect silicon on top of an insulator. There exist stringent requirements that SOI technologies have to meet. For example, all SOI processing steps have to be IC compatible which would exclude the use of extreme high temperature anneal steps. All SOI processing has to be clean room compatible, which may exclude technologies such as SOI by wafer bonding and thinning by polishing. Also, an SOI technology should be economically feasible. For that reason batch processing capability would be preferred over single wafer processing. Lastly, the SOI technology should allow for three-dimensional integration of active device layers.

To date, no single SOI technology has been able to meet all these demands. The purpose of this work was to develop a new CMOS SOI technology, by epitaxial lateral overgrowth of silicon over oxidized silicon to form a second active device layer. This new SOI technology was designed to meet most of the demands stated above.

1.3 Proposed Structure

There is a need to have a CMOS technology that allows for the three-dimensional integration of active devices, where stringent boundary conditions limit the methods by which the required SOI layers can be obtained. In 1982, J.F. Gibbons and K.F. Lee proposed a CMOS inverter with stacked transistors to form silicon-on-insulator device islands. Their method of 3-D integration included the recrystallization of the polysilicon top substrate. Although the design was simple and basically compatible with well established IC fabrication methods, this approach lead to high defect densities in the recrystallized material itself and to damage of the underlying layers.

Also in 1982, K. Tanno et al. reported on selective epitaxial growth (SEG) of silicon in a reduced temperature, reduced pressure epitaxial reactor. Soon thereafter methods were proposed to create SOI structures by extending the SEG of silicon towards epitaxial lateral overgrowth (ELO). In 1984, B. Hoefflinger and S.T. Liu proposed to build CMOS inverters with stacked transistors, structurally similar to the design by Gibbons and Lee, but where the second active device layer was to be formed by epitaxial lateral overgrowth (Fig. 1.1). A more planar structure by S.T. Liu following a similar approach is shown in Fig. 1.2. Here the SEG is not extended over the polysilicon gate, and the PMOS transistor is formed along the gate sidewall.
Figure 1.1: 3-D CMOS Inverter

Figure 1.2: Compact 3-D CMOS Inverter
The stacking of transistors can greatly reduce the area required to build circuits. The method by which this stacking is accomplished inherently leads to a monocrystalline second active device layer. For the structure as shown in Fig. 1.2, the sidewall transistor can have effective gate lengths that are much smaller than the lithographical resolution. The key steps in this new process technology are selective epitaxial growth and epitaxial lateral overgrowth of monocrystalline silicon. These methods are economical batch processes, and integrate well into clean room environments. Epitaxy can be carried out at low temperatures, so that significant movement of junctions already incorporated in the substrate can be avoided.

1.4 Thesis Overview

This thesis describes the development and investigation of a process technology that allows for the fabrication of three-dimensional silicon-on-insulator MOS structures created by the epitaxial lateral overgrowth or selective epitaxial growth of single crystalline silicon over thermally oxidized polysilicon gates, as shown in Fig. 1.1 and Fig. 1.2. The electronic interface of the stacked PMOS transistor, created by the epitaxial overgrowth of thermally oxidized polysilicon rather than by conventional thermal oxidation of monocrystalline silicon is investigated in depth. This novel ELO-Si/polyoxide/polysilicon interface is not just important for active devices in SOI but also for dielectric device isolation techniques by selective epitaxial growth that have already been proposed and realized elsewhere.

Chapter 2 contains a review of the relevant literature on conventional CMOS technologies and advanced device isolation concepts, including currently known SOI technologies, selective epitaxial growth, and epitaxial lateral overgrowth.

Chapters 3 and 4 describe the process sequence development for the shared gate silicon-on-insulator structure by epitaxial lateral overgrowth. In Chapter 3 test devices and structures are discussed and process simulations are contemplated. Also, the test pattern layout used throughout the experimental work is presented.

Chapter 4 reports on the various steps of the new technology and their development. Focus is placed on an in-depth investigation of epitaxy at reduced pressures and temperatures and gate oxide quality.
The electrical characterization of the fabricated test structures, namely 3-D stacked capacitors and 3-D PMOS transistors, is reported in Chapter 5. The measurement apparatus is described, and functionality of the test devices is demonstrated. In Chapter 6, the investigation is summarized and recommendations for future work are given.
CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Over the last several years, integrated circuit technology, driven by demanding new applications, has advanced at a rapid pace. During the last 20 years, the functional density of integrated circuits has increased 200-fold, and speed has increased 20-fold. VLSI has been employed in solving computationally intensive problems in the fields of artificial intelligence and signal processing. In many cases particularly in the field of information processing it is now more cost effective to design application specific integrated circuits (ASIC) than to employ a high-performance general purpose computer.

From all the available processing technologies, CMOS has emerged as the single most important one for VLSI and ULSI. The SOI technology developed in this thesis work is essentially part of a CMOS technology. For this reason, two areas of research are reviewed. The first part of this chapter summarizes issues regarding conventional CMOS circuit and process design methodologies. The second part gives a brief overview over different SOI technologies, and then focuses on SOI by ELO and on epitaxy itself, as they constitute key processing steps in the newly developed SOI technology. General processing subjects such as oxidation, chemical vapor deposition (with the exception of epitaxy), and lithography are not within the scope of this review.

2.2 Background

CMOS devices were first fabricated employing metal gate electrodes. The principal merit of CMOS technology at that time was its low power consumption as compared to the less complex, more cost effective, and widely used unipolar MOS technologies. For that reason it found its principal applications in
battery powered consumer appliances such as wrist watches and pocket calculators, and in aeronautical equipment. During the 1970's, NMOS became the dominant technology for Large Scale Integration (LSI), with refinements such as polysilicon gates, self aligned source and drain regions, and highly resistive polysilicon loads. As device dimensions shrunk and the complexity of systems that could be integrated monolithically rapidly increased it became apparent that power dissipation due to the required steady current flow through the transistors would soon limit the level of integration possible in NMOS technologies. Because of its near-zero quiescent current, CMOS became the technology of choice for VLSI circuit integration.

Beside its low power consumption, CMOS has several other important advantages over bipolar and unipolar MOS technologies. It accepts a wide tolerance in parameter variations such as power supply voltage and temperature. CMOS requires only small drive currents and is less sensitive to backgate biasing than NMOS or PMOS. No bootstrapping is required to feed a signal through a chain of clocked inverters, since no threshold voltage loss occurs. It has excellent noise immunity and soft error protection. The ratio-less design allows for flexibility in circuit layout, which facilitates the use of cell libraries and automated design environments. The high layout symmetry of CMOS is of advantage in analog circuits, thereby making it easy to integrate analog and digital function on the same chip. Operational amplifiers can be integrated in CMOS occupying considerable less area than bipolar or NMOS amplifiers, as illustrated in Fig. 2.1. CMOS furthermore allows for the symmetric operation about the ground potential, and CMOS switches have no offset voltage.

CMOS process technologies are challenging with regard to obtaining high-quality n- and p-channel devices on the same wafer. CMOS also introduces a unique problem named "latchup", where a parasitic bipolar structure can short-circuit transistors of opposite polarity and render the circuit temporarily or permanently inoperable. CMOS usually requires more transistors per function than unipolar technologies, although circuit design methods exist where this is circumvented. Compared with bipolar technologies, CMOS has smaller current drive capabilities, is sensitive to electrostatic discharge, has less gain, and exhibits higher 1/f noise, mostly due to surface states. These disadvantages are the main reason that considerable development work is put into combined bipolar-CMOS (BiCMOS) technologies, since there the strength of each technology can be exploited without having to deal with the restrictions.
Today, MOSFETs with good characteristics have been fabricated with effective gate lengths down to 0.1 μm. The estimated propagation delay time for gates in these technologies is in the order of 20 ps.

2.3 CMOS Circuit Design

As has been mentioned previously, CMOS is as much a processing technology as it is a circuit technology. The link between circuit and process design has become increasingly more important, as circuit and system performance ultimately set the requirements for device and process technology performance. Generally one likes to have systems operate at great speed, and because of cost reasons the minimization of the wafer area occupied by a given circuit is desirable. This can be accomplished by scaling of an existing technology, by the development of a new technology, as pursued in this thesis work, or by new circuit design concepts. Oftentimes it is found that new circuit design concepts require new processing technologies, and vice versa. Dynamic random access memories are a good example for circuits that strongly depend on a particular fabrication technology.
2.3.1 Fully Complementary CMOS

The most basic circuit in CMOS is the inverter (Fig. 2.2). The inverter is composed of a PMOS transistor, whose source is connected to the positive supply voltage, and an NMOS transistor with its source connected to ground. The gates and drains of both transistors are connected together, respectively. When a positive voltage is supplied to the input, the NMOS transistor will open a conducting path from the output to ground. Respectively, if ground potential is supplied to the input, the NMOS transistor will not conduct, and the PMOS transistor will open a path to the positive supply voltage. Therefore, at any one time only one of the transistors will be conducting.

The mobility of holes in silicon is smaller than that of electrons. Since the gain of a transistor is proportional to the carrier mobility, the PMOS transistor in an inverter has usually twice the width of the complementary NMOS transistor to compensate for its smaller gain and to achieve the same current for a given gate and drain voltage. If two inverters are connected as shown in Fig. 2.3, an EXOR gate is obtained. As inputs the logic levels and their complements have to be supplied. A two-input CMOS NAND is given in Fig. 2.4. This also demonstrates one of the drawbacks found in conventional CMOS circuit design: four transistors are needed for the two-input NAND, imposing four capacitive gate loads on the driving circuitry. This problem of fully complementary CMOS designs, resulting in lower speed and greater area compared to equivalent NMOS circuits, can be circumvented by a different approach to CMOS circuit design.

2.3.2 Clocked CMOS

Because of the good electrical isolation of MOS transistor gates with regard to drain, source, and bulk it is possible to employ the capacitor composed of the gate-oxide-bulk structure to store small charges on the gate nodes. Now, for each logic gate its respective output node is precharged to the positive supply voltage (i.e. logical "1"). During this precharge phase the logical function evaluation network is disabled. When the precharge is completed the charging current is switched off and the logical network is enabled. If the input signals satisfy the condition for a logical "0" the precharged output node is discharged over the network. If the input configuration does not satisfy the "0" condition nothing will happen and the precharged output node keeps its logic level. The output node has a valid logic level only after the logical function network has been enabled. It therefore has to be made certain
Figure 2.2: Basic CMOS inverter circuit diagram

Figure 2.3: CMOS EXOR circuit diagram
that successive logic processes the output signal of the previous stage only at the end of the conditional discharge phase.\textsuperscript{18-20}

To illustrate the dynamic CMOS circuit design consider Fig. 2.5. During the precharge time the CLOCK ($\phi$) signal is at logic level "0" and thus enables the PMOS precharge transistor, charging the gate node of the output buffer to the positive supply voltage such that the output node is at logic level "0". The precharge phase is terminated when the CLOCK signal goes to logic level "1", thereby disabling the PMOS precharge transistor and enabling the logical function network via the NMOS transistor T1. If the input signals are such that a conducting path is opened between the precharged inverter input node and ground, the inverter output will go from logic level "0" to one. If the input signal configuration does not meet the requirements of the logic network to evaluate to "1", the precharged node will stay charged and the inverter output will remain at logic level "0".

In this technique, only two PMOS transistors per gate are needed, independent of the complexity of the logic function network. There never is a dc path to ground opened and therefore the low power dissipation of fully complementary CMOS is preserved. There are some drawbacks to this approach, however.

First, the load is always driven by a PMOS transistor, since the transition of the output node is always from ground potential to the supply voltage. This transition is caused by the flow of current through the PMOS transistor of the output buffer, which has a smaller current driving capability than an NMOS transistor. Secondly, because of the non-inverting nature of this design technique, not every logic function can be implemented. However, because clocked CMOS and fully complementary CMOS are compatible, in such cases a conventional design could be used to implement such gates. The third problem stems from the dynamic operation of the circuit. This can lead to charge sharing between the precharge node capacitance and parasitic capacitances of the logic function network, thus jeopardizing proper gate function.

\textbf{2.3.3 Complex CMOS Circuits}

The basic CMOS circuits such as inverter, NAND, and EXOR gates have been reviewed in the last two subchapters. This was to demonstrate that there is a class of circuits that contain the CMOS inverter structure as their basic building block. Another class of circuits, such as multiple input AND and NAND gates, are not easily built from inverters, but most often contain
Figure 2.4: Two-input CMOS NAND gate

Figure 2.5: Three-input clocked CMOS NAND gate
A third important class of circuit modules is that of bistable storage elements, e.g. master-slave flip flops, random access memories, D-flip flops, and J-K flip flops. Also mentioned should be content addressable memory cells.

Figure 2.6 shows an MS flip flop, consisting of four inverters and four transfer gate transistors.

![CMOS master-slave flip-flop diagram](image)

Figure 2.6: CMOS master-slave flip-flop

The static random access memory cell (SRAM) in Fig. 2.7 employs two inverters and two transfer gate transistors that connect the data bit lines to the bistable memory element of two cross coupled inverters. By connecting an EXOR to the SRAM cell such that one input probes into one of the two storage nodes comprised of a cross coupled inverter pair and the other senses the bit line, the output of the EXOR will indicate if the bit line data will match the stored data. This is the function of a content addressable memory cell (CAM) as shown in Fig. 2.8.

It should be noted that a full CMOS SRAM cell requires 6 transistors. Adding the EXOR as shown in Fig. 2.3 would require an additional 4 transistors, bringing the total transistor count of the CAM cell to 10. Making use of the specific properties of the proposed 3-D technology we have proposed a CAM design that only needs three transistors, two of which form an inverter. In conventional CMOS technologies the PMOS transistors of the storage cell are oftentimes substituted by highly resistive polysilicon loads in order to
make the cells smaller. This can be done because the current driving capability of the PMOS transistors is not critical in terms of circuit function. The relaxation on the requirements for the PMOS transistor could be exploited in designing a technology tailored towards the fabrication of SRAMs or CAMs. It demonstrates the need to understand the relationship between circuit design requirements and requirements imposed on a specific fabrication technology by the desired circuit function.

2.4 CMOS Technology

In CMOS, n-type and p-type transistors are combined on the same substrate. Special wells or tubs must be created for one of the transistor types, where the semiconductor type is opposite to the substrate type. For example, an NMOS transistor could be defined in a p-type substrate, and the complementary PMOS transistor would have to be defined in an n-type well. Modern CMOS processes however oftentimes employ a twin-well process, where both, n-type and p-type regions are formed in a lightly doped substrate. The benefits of such an arrangement include the close placement of the NMOS and PMOS devices, the need for only a single lithographic mask step to form both well types, the independent doping control of the bodies of each of the CMOS device types, and the choice of either substrate type for different circuit applications with basically no change in process flow. Generally, a p-type substrate is preferred to an n-type because it is less sensitive to process induced material defects. In addition, the p-type material provides higher conductivity under the NMOS device compared to the conductivity of a p-well in an n-type substrate. For certain applications it is beneficial to use n-type substrates. In static random access memories for example, the p-well to n-type substrate junction provides protection against radiation induced charge loss at the storage nodes.

As device dimensions shrink the problem of device degradation by hot electron injection into the gate oxide increases. This problem is related to the long term stability of the devices. For gate lengths of less than about 2 μm lightly doped drain and source regions are required to minimize threshold voltage shift and oxide deterioration.

To increase operating speeds it is important to keep interconnect resistances as small as possible. The sheet resistance of the polysilicon interconnect level with typical values between 40 and 100 Ω-cm can be reduced to less than 10 Ω-cm by deploying a silicide on top of it.
Figure 2.7: CMOS static RAM

Figure 2.8: CMOS content addressable memory
A cross section of an advanced twin-tub CMOS inverter is given in Fig. 2.9.

Figure 2.9: Twin-tub CMOS inverter cross section

Here devices are isolated from each other by a field oxide region that was created by local oxidation of silicon (LOCOS). In LOCOS, the active device areas are masked by silicon nitride. Prior to the field oxidation a channel stop implant is performed which is selfaligned to the active areas. The nitride mask prevents oxygen from diffusing to the active areas on the substrate, thus permitting to selectively oxidize the wafer. However, because some lateral oxidant diffusion underneath the nitride layer takes place, a "bird's beak" develops, leading to lateral encroachment of the active area. The resulting structure is more planar than that resulting from oxidation and following active area definition by etching of the oxide. The masking nitride can damage the underlying silicon and can lead to problems in subsequent gate oxide formation. For that reason a stress relief oxide is usually grown prior to nitride deposition, or a sacrificial oxide is grown after nitride removal. The outdiffusion of the channel stop dopant into the active area can contribute further to lateral encroachment by reducing effective channel width.

Because of the shortcomings of regular LOCOS other ways of electrically isolating active device areas have been devised. In recessed LOCOS (LOCOS) the wafer substrate is anisotropically etched after the silicon nitride deposition and active device area definition. The local oxidation is carried out following
the silicon etch (Fig. 2.10). This leads to desirable improved surface planarity compared to the standard LOCOS process. Surface planarity becomes more important as minimum line widths become smaller. Steps in the surface can lead to local line width loss during lithography.

Sidewall Masked Isolation (SWAMI) was first proposed by K.Y. Chiu et al. It takes full advantage of LOCOS processing without suffering the drawbacks such as bird's beak formation and lateral encroachment. The SWAMI process sequence is similar to that of recessed LOCOS, except that after the anisotropic silicon etch the developed sidewalls are masked by a stress relief nitride. This can be accomplished without additional masking steps and leads to a defect free local oxidation isolation technique.

The undesirable bird's beak formation in conventional LOCOS has been found to be due to lateral oxidation underneath the stress relief oxide. To eliminate this effect a technique called Sealed Interface Local Oxidation (SILO) was developed. In SILO technology, a sandwich layer consisting of plasma enhanced silicon nitride, LPCVD oxide, and LPCVD nitride is used to mask portions of the substrate from oxidation (Fig. 2.11). The LPCVD oxide film serves as a stress relief between the rigid LPCVD nitride on the top of the sandwich and the substrate, while the bottom nitride film seals off the silicon/silicon dioxide interface from oxidant diffusion. The plasma enhanced bottom nitride is quite thin and less rigid, so that intrinsic nitride stress and volume expansion induced stress are kept small.

Yet another way of circumventing the problems associated with LOCOS was suggested by K. Kurosawa et al., called Buried Oxide (BOX). They masked active device geometries by aluminum and etched the unmasked silicon field regions by reactive ion etching. Then they plasma-deposited SiO2 and etched it in a buffered HF solution. This led to preferential etching of the oxide at the side walls forming V-grooves there. After removal of the aluminum mask and another oxide deposition, the wafer was planarized using a conventional resist reflow and etch back process (Fig. 2.12).

In Direct Moat Isolation (Fig. 2.13) the wafer is completely oxidized. After the oxidation the top region of the oxide film is damaged by ion implantation. Successively the field regions are masked by a thin nitride layer, and an isotropic wet etch is performed. Because of the damaged top oxide layer, the resulting profiles are tapered off. Straight wet etching of thick field oxides would lead to sharp profiles that would impose serious limitations on succeeding fine line lithography and high yield metallization.
Figure 2.10: Recessed LOCOS process flow\textsuperscript{25}
Figure 2.11: Sealed Interface Local Oxidation (SILO) structure
Figure 2.12: BOX fabrication sequence\textsuperscript{28}
Figure 2.13: Direct moat isolation process sequence
Figure 2.14: Trench isolation process flow\textsuperscript{30}
An almost latch-up free isolation technique allowing very densely packed devices was presented with the Trench Isolation concept as shown in Fig. 2.14.\textsuperscript{30} The devices are separated by a deep (> 5 μm) trench which is anisotropically etched into the silicon substrate and then refilled with insulating and non-insulating material. This procedure has found recent application in proposed 4 and 16 MBit DRAM technologies and advanced BiCMOS processes.\textsuperscript{15,31,32} One of the key problems with this approach are the parasitic channels that form where active devices touch the vertical trench sidewall. This leads to increased subthreshold leakage currents.\textsuperscript{33} One way around this is to fill the trenches with heavily doped polysilicon. To avoid outdiffusion of dopants out of the trench during subsequent processing, the trench has to be lined with a thin nitride prior to its filling with polysilicon.

Novel isolation techniques that are not yet established in production lines include device isolation by selective epitaxial growth of monocrystalline silicon (Fig. 2.15).\textsuperscript{34-43}

![Figure 2.15: Dielectric isolation by Selective Epitaxial Growth\textsuperscript{34}](image)

Here problems are similar to that of trench isolation, namely enhanced leakage along the sidewall silicon/silicon dioxide interface. In addition, the problem of non-planarity introduced by faceting during the epitaxial growth has not yet been completely solved. Since selective epitaxial growth is one of the key processing steps in the newly developed 3-dimensional CMOS technology we will review principles and research results in epitaxial depositions of silicon on
silicon in one of the following subchapters.

2.5 SOI Technologies

As device dimension shrink into the submicron range with increasingly better lithographical methods and improved process control, circuit performance is less and less limited by the characteristics of the active devices but by parasitics inherent in todays planar technologies.\textsuperscript{44} New device isolation techniques will be required to prevent parasitic coupling between adjacent devices through the Si substrate. To further increase the complexity of systems that can be integrated, a step towards vertical or three-dimensional integration is seen as necessary.\textsuperscript{45-47} All these issues are addressed by silicon-on-insulator technologies.

In silicon-on-insulator (SOI), devices are built in a thin monocrystalline silicon film on top of an insulator. One of the few SOI processes employed for commercial purposes today is silicon-on-sapphire (SOS), where sapphire (Al$_2$O$_3$) serves as the insulating film as well as the substrate for CMOS circuits. The high costs and complexity associated with SOS technology has so far prevented it from being widely used. Short term objectives of current SOI development efforts are to provide thin films of dielectrically isolated silicon with good semiconductor properties at reasonable costs, particularly to prevent latch-up in advanced bipolar and CMOS circuits. Long term goal of these efforts is the production of vertically integrated circuits.\textsuperscript{48}

The methods by which SOI has been obtained are oxidation of porous silicon (FIPOS), seeded or unseeded recrystallization of molten polysilicon, separation by implanted oxygen (SIMOX), lateral solid-phase epitaxial growth, wafer bonding and wafer thinning, heteroepitaxial SOI, and epitaxial lateral overgrowth. These technologies will be briefly reviewed in the following subchapters, with a focus on silicon epitaxial lateral overgrowth as it was one of the main subjects of this research.

2.5.1 Full Insulation by Porous Oxidized Silicon (FIPOS)

High quality SOI layers can be fabricated by the method of oxidation of porous silicon. Here n-type islands are surrounded by p\textsuperscript{+} regions formed by conventional processing such as ion implantation, diffusion or epitaxy. Then the substrates are etched electrochemically such that porous silicon is formed in the p\textsuperscript{+} layer while the n-type regions are basically not altered. Since porous silicon oxidizes about one order of magnitude faster than monocrystalline
### Table 1

SOI technologies

<table>
<thead>
<tr>
<th>SOI Technology</th>
<th>Usable SOI area</th>
<th>3-D candidate</th>
<th>SOI perfection</th>
<th>Particularities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laser recrystallization</td>
<td>small</td>
<td>yes</td>
<td>entrainment defects</td>
<td>Technology to be applied during IC processing; careful engineering required to prevent device degrading; submicron MOS; low throughput.</td>
</tr>
<tr>
<td>Electron-beam recrystallization</td>
<td>small</td>
<td>yes</td>
<td>entrainment defects</td>
<td>Technology to be applied during IC processing; vacuum environment, careful engineering required to prevent device degrading; submicron MOS; low throughput; pseudo line scan.</td>
</tr>
<tr>
<td>Graphite-strip heater recrystallization</td>
<td>larger up to full scale</td>
<td>no</td>
<td>carbon contamination; can be made defect free</td>
<td>Technology to be applied in early stage of IC processing; higher throughput; lower defect density; submicron MOS; wafer warpage.</td>
</tr>
<tr>
<td>Tungsten-halogen lamps recrystallization</td>
<td>larger up to full scale</td>
<td>no</td>
<td>can be made defect free</td>
<td>Technology to be applied in early stage of IC processing; well suited for thicker SOI layers (&gt; 10 μm); higher throughput; wafer warpage.</td>
</tr>
<tr>
<td>SOI by O⁺ ion implantation and annealing; SIMOX</td>
<td>full wafer</td>
<td>no</td>
<td>dislocations to be prevented; oxide precipitates to be annealed</td>
<td>A technology for wafer suppliers before IC processing; most adapted to existing silicon technology; submicron MOS.</td>
</tr>
<tr>
<td>Full insulation of porous oxidized Si; FIPOS</td>
<td>small islands fully isolated</td>
<td>no</td>
<td>not dislocation free</td>
<td>A technology with a limited applicability due to the smallness of the islands; a technology before IC processing; submicron and thicker layers.</td>
</tr>
<tr>
<td>Lateral solid-phase epitaxial growth; L-SPEG</td>
<td>small stripes</td>
<td>yes</td>
<td>twins</td>
<td>A technology with a limited applicability due to the smallness of the stripes; processing IC compatible; submicron Si-layer.</td>
</tr>
<tr>
<td>SOI by wafer bonding and thinning</td>
<td>full wafer</td>
<td>yes</td>
<td>comparable to bulk silicon</td>
<td>Advanced technology of wafer production; thicker SOI layers (1 - 5 μm); technology for wafer suppliers.</td>
</tr>
<tr>
<td>Heteroepitaxy</td>
<td>substrate scale</td>
<td>probably yes</td>
<td>so far not defect free due to lattice mismatch and dilatation</td>
<td>Technology to be evaluated; IC compatibility dependent on substrate material.</td>
</tr>
</tbody>
</table>
silicon, oxide forms mainly underneath the n-type silicon islands during the subsequent oxidation. Material created by this method can have defect densities comparable to those of bulk-Si. It has successfully been used to build VLSI circuits. The major disadvantage of this method of creating SOI is its incompatibility with established IC fabrication steps. Also, the buried SiO$_2$ does not exhibit good insulating characteristics and the SOI islands are stressed by the oxide layer, inducing defects. The islands fabricated so far by this method were narrow stripes, of only a few hundred microns width. This method is not suited to obtain a continuous film of SOI over an entire wafer area.

2.5.2 Polysilicon Recrystallization

A method to create SOI which has received much attention is the recrystallization of polysilicon deposited over an amorphous insulating layer. The deposition of polysilicon by chemical vapor deposition is a well established procedure employed in almost all modern bipolar and MOS processes. Polysilicon is usually deposited at temperatures between 580 °C and 630 °C such that grain sizes between 10 and 100 nm result. Inside the grains, polysilicon has a regular crystal structure. By heating the polysilicon up to its melting point over an entire wafer area or locally restricted, and subsequently cooling it down, monocrystalline silicon is formed out of the melt over insulating material, which usually is SiO$_2$. The melting can be accomplished by different means.

As heat sources lasers, electron beams, high-intensity lamps, and graphite strip heaters are used. The required high temperatures for the melting of the silicon leads to large horizontal temperature gradients and can damage underlying layers, making vertical integration problematic. The unseeded melting of polycrystalline silicon leads to a random crystal orientation of the resulting SOI layer. Seeding of the polysilicon by contact with the silicon substrate is required to obtain a preferred crystal orientation. To reduce stress induced defects, pattern relief structures can be employed. Although considerable effort has been invested into the development of this technology, only functional models of integrated circuits making use of this process have been demonstrated.
2.5.3 Separation by Implanted Oxygen (SIMOX)

In the SIMOX technique, an oxide is formed underneath a silicon layer by high energy, high dose implantation of oxygen into a silicon wafer. Implantation energies are typically 150 keV to 200 keV and a dose in the order of $10^{18}$ atoms cm$^{-2}$ is required. During the implant wafers are held at temperatures between 400-600 °C to keep the silicon out of the amorphization range due to \textit{in situ} annealing. Following the implant, a high temperature (>1300 °C) anneal step is carried out to restore the implantation damaged silicon surface layer and to eliminate precipitates of SiO$_2$ by dissolving them and by diffusing oxygen atoms into the buried oxide layer. Even after the anneal, the SOI layer is saturated with oxygen atoms up to its solid solubility limit, and dislocations with densities of typically $10^9$ cm$^{-2}$ are present. While oxygen precipitates can be dissolved, the dislocations cannot be annealed out, even if anneal temperatures are close to the melting point of silicon. High temperatures are not only required to anneal implantation damage in the SOI layer but also to form a good quality, low leakage buried oxide. Anneal temperatures of about 1300 °C result in oxide leakage currents of nA/cm$^2$, while thermally grown oxide has leakage currents of typically pA/cm$^2$. To achieve lower defect densities in the SOI layer, different schemes have been devised. The channeling of implants leads to smaller implant damage of the silicon surface layer. Several subsequent low-dose implantations with successive anneals lead to a low-dislocation SOI layer. By these means defect densities could be reduced to $10^5$ cm$^{-2}$.

SIMOX has to be considered a 2-dimensional SOI technology. It is unlikely that more than one active device layer can be created by this method. SIMOX is most suited for very thin (100 nm) SOI layers, as they are required in SOI CMOS, where SOI layers thinner than the depletion layer are needed. In this case high defect densities are not as critical, since junction areas are very small. Hall mobilities in SIMOX SOI layers are comparable to those in bulk silicon, and acceptable low interface state densities and leakage currents have been demonstrated.

One of the most serious drawbacks of this technique is the high dose of oxygen implant required to form the buried oxide. Even with modern high current implanters the implantation time is in the order of hours. This leads to high wafer costs.
2.5.4 Lateral Solid-Phase Epitaxial Growth

In lateral solid-phase epitaxy (L-SPEG) polycrystalline or amorphous silicon is deposited on an insulating material. In the case of a polycrystalline layer, a channeled silicon ion beam is employed to amorphize most of the polysilicon and to leave grains with major crystallographic axes aligned along the channeling direction. The structures are then subjected to a heat treatment during which these grains act as a seed for the solid-phase epitaxial regrowth. Instead of amorphizing polysilicon by channeled silicon beams, openings in the insulating layer to the silicon substrate can act as seeds for regrowth of the amorphous silicon films.67-70 During the epitaxial regrowth the temperature has to stay below the temperature at which spontaneous recrystallization of amorphous silicon to polysilicon takes place. This leads to very slow epitaxial regrowth and limits the width of SOI films to a few microns. So far L-SPEG has not been used for integrated circuit or device fabrication.

2.5.5 Heteroepitaxial SOI

Since it is difficult to grow or deposit monocrystalline silicon on top of an amorphous insulating layer, SOI research in the past concentrated on finding an insulating crystalline material that would be compatible to silicon. In particular, lattice constants have to be close, the dilations as a function of temperature have to be matched, and the material surface has to be susceptible to atomic bonding with silicon. Sapphire has properties that come close to these requirements, and commercially available SOI wafers are mostly silicon-on-sapphire (SOS).71 Problems arise however from the imperfect match of the lattice constants, the brittleness of the sapphire, and the outdiffusion of aluminum into the SOI film. Due to lattice strain and stress, minority carrier lifetimes in SOS are small, baring the fabrication of bipolar transistors on these substrates. Other materials have been investigated, among them spinel, zirconia, and calcium fluoride.72,73 So far, none of these have been used in commercial applications.
2.5.8 Wafer Bonding and Wafer Thinning

When two clean and flat surfaces of high finish are brought close to each other dipole forces are induced into the two surfaces and Van der Waals bonding occurs. The bonding forces are strong under shear and pressure, but weak under pulling. Van der Waals bonding can be converted towards a stronger chemical bonding by annealing of the structure at elevated temperatures. One can bond wafers with an oxidized surface, thereby burying the oxide. The wafer chosen to later contain active devices can then be thinned to a specified thickness above the buried oxide, usually a few microns, resulting in the desired SOI structure. The key problem in this technology is the thinning of the "active" wafer with high geometrical precision and low damage to its surface. To obtain the required high uniformity one oftentimes introduces stopper areas on the wafer. Upon reaching these stoppers, which can be made of a different material or differently doped silicon, thinning effectively stops.

This technology requires facilities that are not usually found in IC fabrication laboratories. Problems also seem to persist with the reliability of the bonding.

2.6 Epitaxial Lateral Overgrowth

Epitaxy is a process where material is deposited onto a crystalline substrate or seed, and where the crystalline configuration of the structure is maintained. Epitaxy can be classified into four different kinds: Molecular Beam Epitaxy (MBE), Vapor Phase Epitaxy (VPE), Liquid Phase Epitaxy (LPE) and Solid Phase Epitaxy (SPE). Vapor phase epitaxy has found the widest acceptance because impurities can be well controlled and crystalline perfection of the grown layers is obtainable. MBE allows for an even better control of layer thicknesses, but suffers from low throughput and high system cost due to the required high vacuum levels. It has however drawn considerable attention lately, particular for the growth of III-V and II-VI structures. LPE is mostly employed for the growth of III-V compounds as GaAs and InP. SPE has been briefly reviewed in a previous subchapter.

VPE has been used since the early stages of device fabrication because it was the only way by which one could obtain a thin layer of lightly doped material on a heavily doped substrate with a sharp transition in the doping profile. Such a structure was desirable because it improved bipolar transistor and later bipolar IC performance.
Lately VPE has been successfully employed in CMOS technologies to reduce latch-up effects. A relatively new branch in VPE is selective epitaxy where the deposition of silicon is confined to certain areas on a wafer. Deposition proceeds from seed openings in a masking oxide (Fig. 2.16). When selective epitaxy is continued even after the growth surface has reached the wafer (oxide) surface, lateral growth occurs. This epitaxial lateral overgrowth forms silicon on the insulating masking oxide and can therefore be considered an SOI technique.

Selective epitaxy and epitaxial lateral overgrowth constitute the key processing steps in the technology developed in the course of this work. This subchapter therefore contains a somewhat detailed review of the principles of silicon VPE from silane and the chlorosilanes.

2.6.1 Basic Terminology

Chemical Vapor Deposition processes, to which VPE is counted, proceed in general according to the following scheme:

1) Reactants are transported to the surface
2) Reactants are adsorbed on the substrate surface
3) A thermally induced chemical reaction takes place at the surface
4) Reaction products are desorbed from the surface
5) Reaction products are transported away from the surface

In general, the various steps of CVD can be grouped into the kinetics of the chemical reactions taking place at a heated surface and the transport of a silicon containing gas compound (i.e. SiH₄, SiH₃Cl, SiH₂Cl₂, SiHCl₃, SiCl₄) towards and away from that surface, including the flow dynamics particular to a given reactor geometry. It is difficult to coherently model all important aspects of both transport phenomena and kinetics in silicon epitaxy. The kinetics and transport are closely linked and interact. Therein lies the difficulty in analyzing and modeling CVD reactor systems. Most researchers therefore have focused either on the kinetics of the process alone, or have simplified the system by limiting considerations to steps 1 and 3, namely gas transport towards the reaction surface and reaction rates there. Not even considered here are homogeneous (gas phase) reactions that can occur during or prior to reactant transport to the reaction surface.

A simple model was introduced by Grove that explains the basic dependencies of growth rates on operating parameters.
Figure 2.16: Selective epitaxy and epitaxial lateral overgrowth
Figure 2.17: Basic model of the epitaxial growth process\textsuperscript{72}

The following notation is used for the different parameters:

- $h_g$: gas-phase mass-transfer coefficient
- $C_s$: concentration of gas at the surface
- $C_g$: concentration of gas far away from the surface
- $F_1$: flux of reactants towards surface
- $F_2$: flux of reactants absorbed on the surface
- $k_s$: surface reaction rate constant

Then, under the assumption of a linear concentration gradient from the gas bulk to the surface one can write

\[ F_1 = h_g (C_g - C_s) \]  \hspace{1cm} (2.1)

and

\[ F_2 = k_s C_s \]  \hspace{1cm} (2.2)

At steady state conditions the two fluxes have to be equal and therefore

\[ C_s = \frac{C_g}{1 + \frac{k_s}{h_g}} \]  \hspace{1cm} (2.3)

The growth rate of the film is given by

\[ V = \frac{F}{N_1} \]  \hspace{1cm} (2.4)
with \( N_1 = 5 \times 10^{22} \text{ atoms/cm}^2 \) for Silicon.

One oftentimes likes to write the concentration of reactants in the gas bulk in terms of its mole fraction of the total gas concentration:

\[
C_g = C_T Y
\]

with \( Y \) as the mole fraction of reaction species and \( C_T \) as the total number of gas molecules/cm\(^3\) in the bulk gas. Then one can rewrite (2.4) as

\[
V = k_s \frac{h_g}{k_s + h_g} \frac{C_T}{N_1} Y
\]

This is the growth rate as predicted by the Grove model. Two limiting cases can be observed when either \( k_s \) becomes small compared to \( h_g \) or vice versa. The first case is called a surface reaction rate controlled deposition; in the latter case the deposition rate is determined by the rate with which the reactants arrive at the surface and is called the mass transfer controlled regime:

\[
V = C_T k_s \frac{Y}{N_1}, \quad h_g \gg k_s \quad \text{(surface reaction rate controlled)}
\]

\[
V = C_T h_g \frac{Y}{N_1}, \quad k_s \gg h_g \quad \text{(mass transfer controlled)}
\]

For thermally activated reactions the surface reaction constant can be written in the form

\[
k_s = k_0 e^{-\frac{E_a}{kT}}
\]

The dependence of the mass transfer coefficient \( h_g \) on temperature is usually much weaker than that of the surface reaction constant as shown in Fig. 2.18. Thus one can obtain an indication for determining whether operation proceeds in the mass flow controlled (region B in Fig. 2.18) or surface reaction rate controlled (region A) regime by performing depositions at various temperatures. Oftentimes the operating point lies in the transition region between the two regimes.

The Grove model contains two parameters that have direct influence on the growth rate, one of which is the mass transfer coefficient, the other being the surface reaction rate constant. In more sophisticated models, the latter can be a function of the gas concentrations at and away from the surface, while the former is closely linked with the flow dynamics of a given reactor system. For this reason the next chapter will concentrate on transport mechanism in
Figure 2.18: Growth rates of silicon films as a function of temperature for various silicon gas sources.

reactor systems, while the chapter after that will deal with the kinetics of various CVD epitaxy systems.

2.6.2 Transport and Fluid Dynamics

For the flux of reactants from the bulk gas towards the surface the Grove model gave

$$F_1 = h_g (C_g - C_s)$$

(2.1)

The gas phase mass transfer coefficient $h_g$ can be computed from experimentally more accessible parameters if fundamental laws of fluid dynamics are considered. When gas or fluid streams over a flat stationary plate a velocity profile perpendicular to the plates' surface will develop. The gas or fluid right at the surface will move very slowly while far away from the surface the gas or fluid velocity will be the unperturbed free field velocity.

Prandtl developed a model for the behaviour of fluids and gasses streaming along a stationary plate in his boundary layer theory. It describes a boundary layer of thickness $\delta(x)$ as a function of position, free stream velocity and
gas viscosity for a gas that flows along a plate. The thickness of the boundary layer is defined as the distance between the surface and the point where the gas velocity is 0.99 times the free field velocity $U$.

![Diagram of a boundary layer in gas flowing over a flat plate](image)

**Figure 2.19:** Development of a boundary layer in gas flowing over a flat plate and expanded view of boundary layer.

The boundary layer thickness can then be calculated from

$$\delta(x) = \sqrt{\frac{\mu}{dU}}$$

(2.8)

The average boundary layer thickness over the whole length of the plate is given by integrating (2.8) over $L$:

$$\delta = \frac{1}{L} \int_0^L \delta(x) dx$$

(2.9)

$$= \frac{2}{3} L \sqrt{\frac{\mu}{dUL}}$$

$$= \frac{2L}{3 \sqrt{Re_L}}$$
where

\[ \text{Re}_L = \frac{dU}{\mu} \]  

(2.10)

is called the Reynolds number for the gas. In general flow regimes with \( \text{Re}_L \) smaller than 2000 are called laminar while higher Reynolds numbers are considered to describe turbulent flow.

In modeling CVD systems the boundary layer is often assumed to be stagnant, i.e. velocities parallel to the surface are zero. The existence of such a stagnant layer close to the susceptor in a horizontal epitaxy reactor was demonstrated by Eversteijn \(^{102}\) et al. and later by Ban and Gilbert \(^{103}\) with TiO\(_2\) smoke experiments. Eversteijn developed a model in which the reactor is conceptually separated into two regions, one in which the bulk gas flows with a finite velocity \( U \), the other of which is stagnant. Reactant transport is thought to occur only by diffusion of the species through the stagnant layer. From Fick’s First Law one can deduce that

\[ F_1 = D_g \frac{(C_g - C_s)}{\delta_s} \]  

(2.11)

with

- \( D_g \) = diffusion coefficient of the active species
- \( \delta_s \) = thickness of stagnant layer.

A paper by Bloem\(^{104}\) presents a similar approach, but special attention is devoted to the temperature dependence of diffusion coefficients, which is claimed to be important because of the high temperature gradient above the surface.\(^{105}\) The average thickness of \( \delta \) can be substituted into (2.11) to yield an estimate for \( h_g \):

\[ h_g = \frac{D_g}{\delta_s} = \frac{3}{2L} \sqrt{\text{Re}_L} \]  

(2.12)

One observes a dependence of \( h_g \) on \( U^{1/2} \); \( U \) in turn is proportional to the gas flow rate. In some instances this has indeed been experimentally confirmed for epitaxial depositions at higher temperatures where \( k_s \gg h_g \).\(^{106}\)

It has been noted that there are high temperature gradients above the susceptor, leading to buoyancy forces that could be responsible for the onset of turbulent flow at much smaller Reynolds numbers.\(^{95,107}\) This has been confirmed by TiO\(_2\)-smoke experiments,\(^{103}\) and the use of the ratio of \( \frac{\text{Gr}}{\text{Re}^2} \) as a criterion for laminar flow has been devised, where \( \text{Gr} \) is called the Grashof number (\( \text{Gr} = \frac{\text{Inertia} \times \text{Buoyancy}}{\text{Viscous Force}} \)).
To accurately compute mass transport in real reactors one has to give up many of the simplifying assumptions made in early modeling attempts that tried to achieve a closed form analytical solution. For example, the stagnant boundary layer through which there is a linear concentration gradient of reactant species in many instances does not accurately describe the actual situation. Entrance effects are oftentimes ignored when a fully developed flow profile over the entire deposition area is assumed. Most models ignore gas phase reactions, and few consider gas phase diffusion and surface reactions to go on in series. Because of their simpler geometries, most researchers have studied epitaxial silicon growth in horizontal (RF and hotwall) reactors or vertical (radiant heated cylindrical or barrel) reactor configurations. For more complicated reactor forms like the pancake type reactor the use of numerical methods is imperative.

2.6.3 Chemical Reactions in Si-Epitaxy

Four sources are available for Si-Epitaxy that can easily be transformed into the gas phase, namely SiCl₄, SiHCl₃, SiH₂Cl₂ and SiH₄. The first three belong to the group of chlorosilanes and the reduction processes are quite similar for these gases. Silane has not been used widely for epitaxy due to its highly unstable behaviour and its tendency to reduce in the gas phase, leading to the formation of silica dust and successive contamination of the wafers. Epitaxy from silane can however proceed at much lower temperatures than from any of the chlorosilanes. A silane system is also easier to model since the decomposition of silane is essentially irreversible under normal operating conditions. No HCl is set free by the decomposition of silane so that one can operate at higher reactant concentrations than with the other gases without entering the etching regime.

The most widely used reactant in the past was SiCl₄ which has a rather low vapor pressure, is chemically stable and easy to handle. However, SiH₂Cl₂ and SiHCl₃ make it possible to operate at about 100 °C - 200 °C lower temperatures for comparable crystalline morphology and growth rate as compared to SiCl₄ systems. It has been shown that the efficiency of the reaction, defined as the ratio of deposited Si to the amount of reactant gas entering the reactor, is highest for SiH₂Cl₂ and lowest for SiCl₄. Nowadays, most reduced temperature epitaxy processes, including most SEG and ELO deposition systems, employ SiH₂Cl₂ as the silicon source gas. It has however recently been demonstrated that in the temperature range of 800 °C - 920 °C SiHCl₃ can be
used to obtain silicon SEG of high chemical purity.

The epitaxial deposition of silicon from silane and the chlorosilanes has been extensively investigated by many researchers.\textsuperscript{119,128-135} The overall reaction for the deposition of silicon from DCS can be written as

\[ \text{SiH}_2\text{Cl}_2 \rightleftharpoons \text{Si} + 2 \text{HCl} \] (2.13)

Thermodynamical calculations and mass spectrometric investigations suggest that \text{SiH}_2\text{Cl}_2 rapidly and nearly completely dissociates in the gas phase above 800 °C. The deposition of Si is believed to be driven by the free energy gain of the reaction

\[ \text{SiCl}_2 + \text{H}_2 \rightleftharpoons \text{Si} + 2 \text{HCl} \] (2.14)

at the hot wafer surface. Another possible mechanism is the reaction\textsuperscript{129,136}

\[ 2 \text{SiCl}_2 \rightleftharpoons \text{Si} + \text{SiCl}_4 \] (2.15)

However, the below equilibrium partial pressures of \text{SiCl}_4 observed during epitaxy\textsuperscript{128} and the abundance of adsorbed hydrogen at the growth surface, as calculated by Chernov\textsuperscript{137} may favor reaction (2.14).

A detailed investigation of silicon deposition from DCS has been conducted by Claassen and Bloem.\textsuperscript{138} They proposed a growth mechanism in which DCS rapidly dissociates into \text{H}_2 and \text{SiCl}_2 in the gas phase at temperatures above 800 °C. Rate determining surface reactions below 1000 °C were identified as being either the surface diffusion of adsorbed \text{SiCl}_2 molecules to atomic steps or kink sites or the chemical reaction at the step to remove the chlorine atoms from the attached \text{SiCl}_2 molecules. In the model development, the following steps were considered:

1. Gas phase diffusion of DCS
2. Gas phase reactions
3. Surface adsorption of different species
4. Surface reactions and silicon incorporation at steps

Claassen and Bloem then derived the growth rate expression

\[ R = \frac{k_4k_{15}p\text{H}_2p\text{SiCl}_2\theta}{(k_{-4} + k_{14})(k_{-14} + k_{15}p\text{H}_2)} - K_0\frac{p^2\text{HCl}}{p\text{H}_2} \] (2.16)

with the fraction of free surface sites \( \theta \) given by

\[ \theta = \frac{1}{1 + K_4p\text{SiCl}_2 + K_5p\text{HCl}/p^{1/2}\text{H}_2 + K_8p^{1/2}\text{H}_2} \] (2.17)
The first term of equation (2.16) describes the growth reaction of the reduction of SiCl$_2$ at the surface while the second term accounts for the etching reaction of silicon by HCl. The parameters in the above equations are the constants for the following chemical equations (* denotes adsorbed species or a free surface site, st denotes a step or kink site on the surface, and cr denotes species incorporated into a crystalline lattice):

$$\text{SiCl}_2(g) + * \leftrightarrow \text{SiCl}_2^*$$  \hspace{1cm} (2.18)

$$\text{SiCl}_2^* \leftrightarrow \text{SiCl}_2(\text{st})$$  \hspace{1cm} (2.19)

$$\text{SiCl}_2(\text{st}) + \text{H}_2(g) \leftrightarrow \text{Si(cr)} + 2 \text{HCl}(g)$$  \hspace{1cm} (2.20)

$$\text{HCl}(g) + * \leftrightarrow \text{Cl}^* + \frac{1}{2} \text{H}_2(g)$$  \hspace{1cm} (2.21)

$$\frac{1}{2} \text{H}_2(g) + * \leftrightarrow \text{H}^*$$  \hspace{1cm} (2.22)

$$K_e = \frac{R_{\text{etch}}}{P_{\text{H}_2}^2}$$  \hspace{1cm} (2.23)

with $K_4 = k_4/k_{-4}$, etc. It has to be noted that the etch expression was derived from experiments where no Si-containing species were introduced into the reaction ambient during etching. The reaction rate $R$ is governed by the adsorption rate of SiCl$_2$ at the surface ($k_4$), the surface diffusion of adsorbed SiCl$_2^*$ towards a step ($k_{14}$), and the chemical reaction there ($k_{15}$). The fraction of free surface sites for the adsorption of SiCl$_2$ is determined by the amount of already adsorbed SiCl$_2^*$ ($K_4$), adsorbed Cl$^*$ ($K_5$), and adsorbed H$^*$ ($K_6$).

With this model and appropriate parameter fitting, Claassen and Bloem were able to model the experiments they had conducted with good accuracy. It seems however that some effects observed in epitaxy such as the reduction of growth rate fall off with increasing HCl in the growth ambient once the etching regime is entered and the local loading effect in selective epitaxial growth are difficult to explain with this approach. In particular, the limitation of growth contributing ad-species to SiCl$_2$ seems somewhat arbitrary and restrictive. Stassinos et al. postulated the existence of SiH$_2$ as a second ad-species in addition to SiCl$_2$ by the reaction.
This reaction is suggested to govern the linear decrease in growth rate with HCl addition in the presence of hydrogen. In an inert atmosphere the gas phase reaction (2.20) is assumed to control the linear decrease in growth rate as was also suggested by Claassen and Bloem.

2.6.4 Applied SEG and ELO

It has been shown that there has to be a higher degree of supersaturation for the nucleation of silicon on SiO₂ and Si₃N₄ as compared to that for nucleation on silicon surfaces. Thus, by keeping the supersaturation below a critical value it is possible to selectively deposit Si on Si substrates masked by either silicon nitride or silicon oxide. Crystal growth theories, as discussed by Bennema and van Leeuwen, explain the initiation of growth by the adsorption of silicon at the growth surface. Adsorbed atoms form little clusters; these clusters are thermodynamically unstable until they reach a certain critical size. Thereafter it is energetically more favorable for them to remain in the solid phase than in the vapor phase.

The adsorption energy for Si on foreign substrates is generally higher than that for Si on Si; thus it becomes possible to operate at a point where the nucleus size on the foreign material is held below the critical value while over-critical size nuclei can form on the Si-growth surface. The process is a balance between reasonable growth rates and polysilicon nucleation on the masking material, which most often is SiO₂. The onset of nucleation on the mask is a function of temperature, pressure, mask material, mask material cleanliness, and the Cl/Si ratio in the reactor ambient. The ratio of masked to unmasked area on a wafer can have an indirect effect on the extent of nucleation.

The chemical system most often employed in SEG and ELO is SiH₂Cl₂-HCl-H₂ at reduced pressure and temperatures between 800 °C and 1000 °C. Other systems such as silicon-iodine and silicon-bromine have also been used. Most researchers introduce the etching and growth species into the reaction chamber simultaneously to maintain selectivity. Others have worked on schemes employing repeated growth and etch cycles of short duration, making use of the induction time of nucleation on SiO₂ which is in the order of 20 seconds.

The first problem encountered in developing a working selective epitaxial deposition process is that of the mask material degradation by the in situ pre-clean cycle typically employed in non-selective epitaxy. A pure silicon surface

\[ \text{SiCl}_2 + 2 \text{H}_2 \leftrightarrow \text{SiH}_2 + 2 \text{HCl} \]
is very reactive and almost instantly forms compounds when exposed to gases in the atmosphere even at room temperature. The occurring reactions typically result in a thin "native" oxide layer, or in the presence of reactive carbon containing species, in silicon carbon compounds. The crystalline perfection of an epitaxial layer strongly depends on that of the starting substrate surface. Residual native oxide or carbon on the growth surface would lead to high defect density epitaxial material. To remove the native oxide layer and possibly carbon an \textit{in situ} high temperature bake of a few minutes duration in hydrogen usually precedes the epitaxial deposition. The native oxide is removed by the reaction

\[
\text{SiO}_2 + \text{Si} \leftrightarrow 2\text{SiO(v)} \tag{2.25}
\]

in the presence of low partial pressures of oxygen and water vapor and high temperatures. The removal of carbon is more difficult, and temperatures in excess of 1200 °C are required.\textsuperscript{145} The same mechanism that helps in removing native oxide can also attack the masking oxide in selective epitaxy. This was indeed observed\textsuperscript{146,147} as an undercutting of the masking oxide close to the seed area as shown in Fig. 2.20. The amount of undercut is a function of time and temperature as depicted in Fig. 2.21. It is therefore mandatory to keep temperatures during the entire epitaxy cycle, including the preclean period, below about 1000 °C. This precludes the removal of carbide from the silicon surface. Special wet chemical cleaning procedures have therefore been developed to obtain a clean silicon surface prior to the deposition cycle by researchers working on silicon molecular beam epitaxy where problems are similar.\textsuperscript{148} Usually, carbon is removed by organic and inorganic solvents. Then a thin (1 nm) oxide is repeatedly chemically grown and removed, for instance in boiling HNO\textsubscript{3} solution, resulting in a very clean native oxide layer. This native oxide layer is decorated at its active sites with unstable adsorbants, such as Cl atoms, before carbon atoms are adsorbed. The Cl atoms desorb easily at temperatures below which the native oxide film is etched away, thereby effectively protecting the silicon surface from carbon contamination on its way to the reactor chamber. The oxide can be removed prior to the growth \textit{in situ} at temperatures down to 850 °C in a hydrogen ambient. At these temperatures the undercut process is so slow that it does not pose a problem for VLSI with minimum dimensions around 1 μm.

The addition of HCl to the growth ambient leads to a change in growth rate of the \{100\} planes relative to the \{110\} planes.\textsuperscript{149} The anisotropic growth rates cause a distinct facet formation which was investigated among
Figure 2.20: Undercutting of the SiO₂/Si interface at 1150 °C during a conventional in situ preclean in H₂ at 50 Torr \(^{226}\).

Figure 2.21: Arrhenius plot of undercutting length for preclean in H₂ at 50 Torr \(^{226}\).
others by Ishitani et al. and Drowley et al.\textsuperscript{150,151} To minimize facet formation, growth is devised to proceed from a (100) wafer surface where the masking oxide pattern should be aligned to the \textlangle100\rangle directions. As deposition is carried out at lower temperatures less HCl is needed to maintain selectivity. The same is true for wafers where masking oxide coverage is small. Then very little facet formation and good epi-Si surface planarity can be obtained without compromising selectivity and crystalline quality.

The quality of the epitaxial material is also related to the orientation of the seed sidewall oxides and the substrate surface orientation.\textsuperscript{43,85,152} In general, it was found that along \textlangle100\rangle oriented sidewalls the occurrence of stacking faults in the selectively grown epitaxial layer is minimized. Furthermore, growth from (100) substrates is of higher quality than that from (111) substrates. It is believed that stacking faults are generated as a strain relief mechanism. For seed window side walls along a \textlangle100\rangle direction on a (100) substrate, the directions along which dislocations can be easily nucleated do not coincide with the lines of high internal stress. Shear strain measurements on ELO material showed that samples with high stacking fault densities had almost undetectable strains while those that were free of stacking faults exhibited considerable shear strains. The strains were probably generated from thermal gradients during cooling or from differences in the thermal coefficients of expansion between silicon and the overgrown SiO\textsubscript{2}.

Initial attempts to employ SEG for the fabrication of CMOS circuits detected a problem where "edge" diodes, those that were at least in part defined by the SEG/masking oxide interface, exhibited increased leakage compared to "edgeless" diodes, those that did not touch the SEG/oxide interface.\textsuperscript{42} The high leakage currents were related to crystalline defects at the interface,\textsuperscript{153} but the exact nature of the problem was not understood. Prior to this study, there hadn't been any investigations regarding Si/SiO\textsubscript{2} interfaces that were not generated by the thermal oxidation of silicon but rather by the epitaxial overgrowth of silicon over the oxide. Some researchers observed a reduction in leakage currents when epitaxial depositions were carried out at lower temperatures.

If the film is grown longer than necessary to fill the void created by the etching of the SiO\textsubscript{2} mask it will not only grow vertically but horizontally as well, forming an SOI structure. It was initially believed that it would be possible to achieve higher horizontal growth rates than vertical ones, or in other words, to obtain aspect ratios of much greater than one. It was assumed that the enhanced diffusion of adatoms on the mask material close to the seed area
towards the laterally growing surface would increase the amount of Si available for growth there as compared to that available to the top surface. This expectation has not become a reality despite reports of aspect ratios between 6:1 to 1000:1.\textsuperscript{154} Results reporting high aspects ratios should be carefully re-examined since there is a possibility that the thin masking oxides that were necessary to achieve such ratios were degraded during the prebake preceding the epitaxial growth, thereby enlarging the seed area.

Jastrzebski \textit{et al.} investigated and compared devices build in ELO material, SEG, and bulk silicon.\textsuperscript{155} They determined that minority carrier lifetimes and mobilities were comparable between SEG and ELO material, i.e. that the quality of the silicon as it grows over the oxide does not deteriorate. Bulk control devices however exhibited characteristics that were significantly better than that of devices build in epitaxial silicon, as shown in Table 2.

\begin{table}[h]
\centering
\caption{Summary of device and material properties\textsuperscript{155}}
\begin{tabular}{llll}
\hline
 & Hole Mobility & N-Region Lifetime & P-Region Lifetime \\
\hline
Bulk Control & 200 cm\textsuperscript{2}/Vs & 315 $\mu$s & 50 $\mu$s \\
Homoepitaxy & 170 cm\textsuperscript{2}/Vs & 50-100 $\mu$s & 1 $\mu$s \\
20 $\mu$m ELO & 160 cm\textsuperscript{2}/Vs & 10-20 $\mu$s & 1 $\mu$s \\
6 $\mu$m ELO & 180 cm\textsuperscript{2}/Vs & 5-15 $\mu$s & \\
\hline
\end{tabular}
\end{table}

Recently Liu\textsuperscript{156} did a similar study, comparing the performance of MOS transistors build in ELO material and the SEG material directly above the seed region. He grew silicon selectively from seed windows, spaced 10 $\mu$m apart, such that the growth fronts merged, resulting in a planar overall surface everywhere. This was done in a radiant heated barrel reactor at 900 °C and 50 Torr and an HCl/DCS ratio of 2. The thickness of the buried oxide stripes was 400 nm. About 8 $\mu$m of vertical growth was required to obtain a planar surface, which was then etched back to about 1 $\mu$m by plasma etching. After the plasma etch the film was further thinned to about 0.6 $\mu$m by wet etching to remove plasma damage. A row of discrete NMOS transistors of various sizes was built on the SEG seed region, and comparable transistors were built on the ELO region above the oxide stripes. Both types of transistors exhibited a channel mobility of 410 cm\textsuperscript{2}/V-sec. The subthreshold slope
and breakdown voltages were 120 mV/dec and 8.1-8.3 Volts, respectively, for both types of transistors. This demonstrated that the electrical properties of SOI by ELO were about the same as the ones of the SEG bulk region.

2.7 Summary

In this chapter material pertinent to the thesis work has been reviewed. It was demonstrated that many basic CMOS circuits contain inverter-like structures as their basic building blocks. Furthermore, some examples were given where a good understanding between circuit demands and technological possibilities were imperative. The proposed new 3-D technology includes stacked transistor inverters as a basic structure and lends itself for applications such as SRAMs, CAMs and clocked CMOS.

Basic issues in CMOS technology were reviewed as they are important in the development of a new process. Emphasis was placed on device isolation concepts since this is a problem that becomes increasingly more important with shrinking device dimensions.

SOI technologies were shown to provide a remedy for most of the difficulties with which conventional planar processes are plagued today. So far however, most SOI technologies introduce new problems that are not easily overcome.

Selective epitaxial growth and epitaxial lateral overgrowth were introduced as new concepts to fabricate SOI structures and provide device isolation. Principles of silicon epitaxy were reviewed, as well as some more involved issues concerning modeling of epitaxial processes and practical considerations in the application of SEG and ELO.

Most of the material presented in this chapter served as background knowledge for the design of the new 3-D MOS process which will be described in the following sections.
CHAPTER 3

PROCESS DEVELOPMENT I: TEST STRUCTURES

3.1 Introduction

The ultimate objective of this project was to develop a three-dimensional CMOS technology in which PMOS and NMOS transistors were stacked above each other, sharing the same polysilicon gate. The required second silicon active device layer was to be grown on top of SiO₂ by ELO.

This technology contained many new and unproven steps compared to conventional VLSI silicon processing. For that reason, each of the new processing steps had to be developed and investigated separately, and then optimized to fit into the process as a whole. Once the single steps were developed, devices were fabricated and evaluated. Measurement results from these devices were employed to refine the fabrication process. When good device characteristics are obtained circuits can be fabricated to evaluate the potential of a new technology for VLSI applications. In our laboratory, the complexity of the circuits built was however limited by yield and environment induced defects during processing. For that reason the most complex structure attempted to be fabricated in this study was a stacked 3-D PMOS transistor.

The development of a new technology has to proceed from process design to device fabrication to circuit fabrication. In the following sections the test structures used for this work will be described and the mask layout for their fabrication will be discussed.

3.2 3-D CMOS Process Flow

The proposed three-dimensional shared gate stacked CMOS inverter is shown in Fig. 3.1. The fabrication sequence begins with the implementation of the standard bottom (NMOS) transistor. Since aspect ratios of much greater than one are difficult to obtain, the gate length of the NMOS
transistor has to be as small as possible to facilitate planarization of the ELO structure to be grown over the gate. The amount of vertical growth can be estimated from gate polysilicon thickness, minimum required spacing between seed hole and gate, and minimum required overgrowth over the gate on the side opposite to the seed hole.

In the Purdue solid state laboratory a misalignment of 1 μm between successive layers was considered the best that could be done on a routine basis. The limits of UV lithography in this laboratory were determined to be between 2 μm and 3 μm, mostly due to the lack of a highly stable, controlled environment. Thus a total vertical growth of (2(1 μm misalignment) + 0.1 μm gate oxide + 0.4 μm polysilicon + 3 μm gate length + 2 μm overgrowth beyond the polygate) 7.5 μm was needed to completely overgrow the oxidized gate and to leave some room on the side opposite of the seed hole, to allow for the implementation of the stacked PMOS transistor source region.

After the formation of the NMOS transistor, the polysilicon gate is thermally oxidized. The seed holes are opened in the drain region of the NMOS transistor and epitaxial growth is carried out. The resulting ELO island, which is about 7.5 μm high, has then to be planarized back to a height of about 0.5 μm above the poly gate top surface. Subsequently a threshold voltage
adjustment for the top device and the source and drain implant are performed. An insulating oxide layer is deposited and contacts are opened. Finally metal interconnects are formed. The process flow is as follows:

1. 0.8 μm field oxidation
2. Active area masking and etching
3. NMOS gate oxidation (100 nm)
4. 300 nm polysilicon deposition and then the phosphorus doping
5. Polysilicon masking and etching
6. NMOS arsenic source and drain implant
7. Polysilicon thin (100 nm) oxidation
8. Seed window opening and 7.5 μm selective epitaxial growth
9. Overgrowth planarization
10. PMOS threshold voltage adjust implant
11. PMOS boron source and drain implants
12. Plasma oxide coverage
13. Contact opening
14. Metallization

The major steps of the process are sketched in Fig. 3.2.

Since the first processing steps are for the fabrication of the conventional NMOS bottom device only, this part was omitted during the development work and a revised fabrication scheme was used to create a 3-D stacked capacitor:

1. 0.8 μm field oxidation
2. Heavy phosphorus implant to assure accumulation
3. 1.5 μm polysilicon deposition and phosphorus doping
4. Polysilicon masking and reactive ion etching
5. Polysilicon thin (100 nm) oxidation
6. Seed window opening and 7.5 μm selective epitaxial growth
7. Blanket epitaxial Si phosphorus implant
8. Contact implant anneal and oxidation
9. Contact opening
10. Metallization

This revised process resulted in a structure equivalent to the 3-D CMOS inverter except for the source and drain implants. It facilitated the investigation of the novel interface generated by the epitaxial lateral growth of silicon over oxidized polysilicon gates. The electrical characteristics of that interface were considered crucial with regard to the performance of the stacked (PMOS)
Figure 3.2: 3-D CMOS process flow
After Seed Window Opening

After Epitaxy

After Planarization

Figure 3.2 cont'd.
Figure 3.2 cont'd.
transistor. The PMOS transistor was later fabricated by extending the above process sequence by planarization and source and drain implants.

3.3 Test Device Structures

Examining the three-dimensional CMOS inverter (Fig. 3.1) and the fabrication process sequence it can be seen that many steps are required that are not part of conventional processing:

1. The fabrication of the top transistor gate oxide by oxidizing the polycrystalline silicon gate.
2. The ELO to create a second active device layer covering the oxidized polysilicon gate.
3. The fabrication of the ELO/polyoxide/polysilicon interface with good electrical characteristics by growing silicon on oxide, instead of oxidizing silicon.
4. The planarization of the ELO from a height of about 8 \( \mu \text{m} \) down to 1 \( \mu \text{m} \) with good uniformity over an entire wafer area.
5. The fabrication of the top PMOS transistor by implanting source and drain regions into the planarized ELO.
6. The integration of all of the above steps into a conventional NMOS process to create the stacked inverter without the introduction of unwanted interactions.

To develop each of the new processing steps two mask sets were designed: one dedicated to the investigation of ELO growth and ELO planarization, and the other for device development. The following section will describe the ELO and planarization mask set. Thereafter the structures designed for device development and the associated mask will be presented.

3.3.1 ELO Growth and Planarization

Selective epitaxial growth and epitaxial lateral overgrowth are relatively new processing methods that had just recently been established in the Purdue laboratory. The prime concern in epitaxial growth is good growth rate uniformity over a wafer and from wafer to wafer, good morphology of the material grown, and crystalline perfection of the deposited silicon. In SEG and ELO there is additional concern regarding the selectivity of growth and the aspect ratio, i.e. the ratio of horizontal to vertical growth rate. Although there have
Figure 3.3: Epitaxy and planarization test mask
been reports in the literature that aspect ratios can be as high as 1000:1,\textsuperscript{154} it is generally believed that ratios close to one are more realistic for the growth of device quality material. Equal growth rates in vertical and horizontal directions were found during the initial stage of this investigation, and no further effort was expended to achieve aspect ratios in excess of one.

It has been established that there exist local loading effects\textsuperscript{80} where growth rates on a wafer or a certain wafer area depend on the ratio of exposed to masked silicon. Furthermore, because of the low aspect ratios a planarization procedure had to be developed to obtain 7 $\mu$m lateral overgrowth of 1 $\mu$m thick silicon. An enlarged reproduction of the mask that was designed for the investigation of local loading during ELO, growth uniformity during ELO, and the planarization of structures of different width and spacing as well as the accurate profile measurement of the structures is shown in Fig. 3.3. This mask contains various arrays of lines with different width and spacing. In general the amount of planarization or step reduction is a function of these two parameters. Four different spacings were available, 20 $\mu$m, 30 $\mu$m, 40 $\mu$m, and 280 $\mu$m. For the first 3 spacings there were 4 lines each for 5 different line widths, namely 5 $\mu$m, 10 $\mu$m, 15 $\mu$m, 20 $\mu$m, and 30 $\mu$m. All lines were 20 $\mu$m long to facilitate profile measurements. The chosen values for spacing and widths resemble those that would be used in real devices. A grid structure of 3 $\mu$m by 10 $\mu$m dots was also included to simulate an array of actual devices.

The mask provided a second level for active area so that it was possible to check the dependence of planarization on recessed epitaxial growth. For that purpose there were 5 $\mu$m lines embedded in the middle of 10$\mu$m, 20 $\mu$m, 30 $\mu$m, 40 $\mu$m, 50 $\mu$m, 100 $\mu$m, 150 $\mu$m, and 200 $\mu$m active areas.

3.3.2 Electronic Device Structures

The fabrication of electronic devices involves many processing steps. A mask set suitable for the development of a process has to allow for structures by which each processing step can be evaluated and optimized independently. It must be possible to characterize each film of material with regard to the properties that are important for its function in device structures. In particular, the sheet resistance and thickness of doped regions, polysilicon films, and epitaxial films have to be measurable. Contact resistance as a function of contact-window size, metal step coverage, field oxide thickness and field oxide threshold voltage, gate oxide threshold voltage and thickness, and polysilicon oxide thickness and leakage have to be assessed. For many of these
Figure 3.4: 3-D capacitor cross section

Figure 3.5: 3-D capacitor electrical equivalent
measurements there exist simple standard test structures which will not be further discussed here.

Devices are created by forming junctions or interfaces between materials of different composition. The quality of these junctions and interfaces often times determines the performance of the electronic device they constitute. For MOS transistors, the interface between the insulator and the bulk silicon in which a modulated conducting channel is formed is of prime importance. Its properties determine gain, stability with regard to temperature and gate bias, lifetime, and noise sensitivity of the device.

Hence, the second stage in process design, after having developed a suitable procedure to selectively deposit epitaxial films, polysilicon films, and polysilicon oxide films, was the fabrication and evaluation of the ELO/polyoxide/polysilicon interface. To investigate these interfaces the MOS capacitor is one of the simplest and most powerful tools. MOS capacitors for material and interface evaluation purposes are usually planar structures formed by deposition of gate material on top of an insulator which in turn is on top of the silicon substrate. Since in this study an interface had to be characterized that was formed by growing a "substrate" laterally over an insulator, which in turn buried a gate electrode, different considerations applied in designing a suitable test structure.

![Figure 3.6: Threshold voltage as a function of doping density](image)

Figure 3.6: Threshold voltage as a function of doping density
First it has to be noted that it is very difficult to create a truly insulated single crystalline silicon layer on top of an amorphous insulator by epitaxial lateral overgrowth. Since the epitaxial growth proceeds from a seed, there is an electrical and physical connection between the wafer substrate and the epitaxial lateral overgrowth. The aspect ratio is generally about one, and thus it is not practical to create structures that grow more than about 10 \( \mu \text{m} \). Although in principle it would be possible to remove the epitaxial silicon in the seed area, thereby forming true silicon-on-insulator, it is very difficult to etch anisotropically through a 10 \( \mu \text{m} \) layer of silicon. If reactive ion etching (RIE) was to be employed, a suitable masking material would have to be found which could withstand the necessary extensive etching times. There is also reason to believe that extensive RIE would alter the electronic properties of the interface. Because of these difficulties, and since no suitable RIE system was available at the Purdue Laboratory, a different approach was taken where true insulation of the ELO was not necessary for the extraction of the ELO/polyoxide/poly interface characteristics.

The structure chosen as the test vehicle is shown in Fig. 3.4. Polysilicon is deposited over a thick field oxide, oxidized to form the top insulator, and then a seed window is cut into the field oxide close to the polysilicon gate. The subsequent epitaxial lateral overgrowth buries the oxidized polysilicon gate and forms an upside-down MOS capacitor. A simple electrical equivalent circuit of the entire structure is given in Fig. 3.5. Only the capacitor \( C_{\text{ELO}} \) formed by the ELO and the oxidized polysilicon gate is of interest. However, the capacitance that can be measured consists of the sum of \( C_{\text{ELO}} \), the capacitance between the polysilicon gate and the wafer substrate (\( C_{\text{Substrate}} \)), and the capacitance between the metal pad and the substrate (\( C_{\text{Pad}} \)).

To reliably extract \( C_{\text{ELO}} \) from the total capacitance, the bottom capacitance was held small by means of a thick field oxide, and the threshold voltages of the bottom devices were fixed such that these devices were always in accumulation for any bias voltage of interest (see Fig. 3.6).

The test mask also included stacked PMOS transistors structurally similar to the ELO capacitor, but with source and drain implants into the planarized ELO material. To evaluate the NMOS part of the process, NMOS transistors with varying width to length ratios were provided. Both PMOS and NMOS transistors were integrated to form the proposed 3-D CMOS inverter and a novel dynamic CAM cell.
Figure 3.7: 3-D capacitor mask layout
Figure 3.8: PMOS stacked transistor layout

Figure 3.9: 3-D CMOS inverter layout
Figure 3.10: CAM circuit diagram

Figure 3.11: Dynamic CAM mask layout
Figure 3.12: Electronic device characterization mask
3.4 Process Simulation

The fabrication of complex electronic devices is a long and expensive process. Process simulation programs can help in designing an initial set of implant doses, diffusion and oxidation times, and temperatures, with almost instant feedback to the design engineer. In later stages of the process development, simulations can be used to investigate the influence of parameter changes made in the fabrication sequence. Although the accuracy of the simulations is at most as accurate as the equipment dependent parameters that are put into the simulation input deck, they can nonetheless provide valuable information in the understanding of interactions between different processing steps. The process simulator SUPREM III\textsuperscript{158} was used in this work to first investigate the simple 3-D capacitor as shown in Fig. 3.4, and then to develop a set of parameters for the fabrication of the 3-D CMOS inverter as shown in Fig. 3.1.

3.4.1 3-D Capacitor Simulation

For the fabrication of the 3-D capacitor the only impurities intentionally introduced into the wafer were phosphorus and arsenic. It was mentioned that selective epitaxy grown on (100) substrates had superior qualities compared to growth from (111) wafers. Therefore the wafer starting material chosen was (100) oriented phosphorus doped silicon, with an approximate dopant concentration of \( N_D = 10^{14}\text{cm}^{-3} \). To obtain a large negative threshold voltage for the poly/field oxide/substrate capacitor a threshold adjust implant was performed prior to the field oxidation after which the dopant concentration at the field oxide/substrate interface was about \( N_D = 10^{19}\text{cm}^{-3} \). According to Fig. 3.6 this resulted in a threshold voltage of less than \(-30\text{V}\).

Following the polysilicon deposition, implantation, and thermal oxidation there were additional impurities pushed close to the field oxide/silicon bulk interface, since the segregation coefficients for arsenic and phosphorus are greater than one and diffusion coefficients for these dopants through oxide are small, as shown in Fig. 3.13 and Fig. 3.14. Three different temperatures were simulated for the selective epitaxial growth: \( 860^\circ\text{C} \), \( 900^\circ\text{C} \), and \( 950^\circ\text{C} \). It was taken into account that growth rates and hence deposition times were different for the different temperatures. For instance, at \( 860^\circ\text{C} \), \( 7\mu\text{m} \) of growth would require 70 minutes of deposition, while at \( 950^\circ\text{C} \) less than half that time was sufficient. The total concentration of active impurities after the contact implant and anneal at the polysilicon oxide/ELO interface is
Figure 3.13: Simulation of 3-D capacitor, polysilicon after P-implant

Figure 3.14: Simulation of 3-D capacitor, polysilicon after oxidation
Figure 3.15: Simulation of 3-D capacitor, 860°C ELO before contact anneal

Figure 3.16: Simulation of 3-D capacitor, 860°C ELO after 920°C contact anneal
Figure 3.17: Simulation of 3-D capacitor, 900 °C ELO before contact anneal

Figure 3.18: Simulation of 3-D capacitor, 950 °C ELO before contact anneal
Figure 3.19: Simulation of 3D capacitor, ELO after 1000 C contact anneal.
5.9\times10^{15}\text{cm}^{-3}, 6.5\times10^{15}\text{cm}^{-3}, \text{and } 8.8\times10^{15}\text{cm}^{-3} \text{ for epitaxial deposition at } 860^\circ\text{C}, 900^\circ\text{C}, \text{and } 950^\circ\text{C}, \text{respectively. It was therefore concluded that deposition temperature had no significant impact on dopant outdiffusion near the polyoxide/ELO interface. This is also evident from the following consideration: The diffusion coefficients for the impurities boron, arsenic, and phosphorus have activation energies of about 4 eV. This means that diffusion at 850^\circ\text{C} \text{ is about 30 times slower than at } 950^\circ\text{C}. However, in expressions for junction depths the square root of the diffusivity is usually the dominating term. This value of about 5 is close to the factor of 4 which is observed for the ratio of maximum selective growth rates at 950^\circ\text{C} \text{ and } 850^\circ\text{C}. The lower growth rates at smaller temperatures require longer deposition times to obtain the same amount of growth. Thus the smaller diffusivities are compensated for in part by the longer epitaxial growth times.}

The dopant distribution throughout the structure changes significantly following a high temperature anneal step as shown in Fig. 3.19. After 10 minutes at 1000^\circ\text{C} \text{ the interface dopant concentration has increased to } 1.3\times10^{16}\text{cm}^{-3}.

### 3.4.2 3-D CMOS Inverter Simulation

The three-dimensional shared gate CMOS inverter can be derived from the 3-D capacitor by changing the substrate to p-type, providing a thin gate oxide for the bottom device, and by implanting source and drain regions. The associated processing steps were simulated numerically. Figures 3.20 to 3.23 show the resulting dopant profiles for the four cross sections X1 to X4 as sketched in Fig. 3.1. The field oxide was grown to 0.8 \mu m thickness and the substrate doped to about 10^{18}\text{cm}^{-3} \text{ at the interface. The high dopant concentration at the field oxide surface that can be seen in Fig. 3.20 stems from the (boron) stacked transistor source/drain and contact implant.}

The arsenic implanted source/drain regions for the bottom (NMOS) device are about 0.7 \mu m deep (Fig. 3.21). This is more than would be desired for a 3 \mu m technology where typical values are in the range of 0.2 \mu m to 0.3 \mu m. The deep drive-in occurs mostly during the selective epitaxial growth and the stacked (PMOS) transistor source/drain drive-in. This dopant redistribution is an area of concern in designing sub-micron SEG processes.

The cross section of the shared drain region shows some outdiffusion of the arsenic drain into the SEG silicon (Fig. 3.22). The stacked (PMOS) transistor boron implanted drain reaches down to the arsenic implanted
bottom (NMOS) transistor drain, forming a p⁺/n⁺ diode about 0.7 \( \mu \text{m} \) from the top transistor surface. This diode can be useful in some circuits specifically designed around this process.

For the CMOS inverter, the gate oxide thickness was reduced to 50 nm, as is typical for a 3 \( \mu \text{m} \) process. This leads to some problems at the stacked transistor electronic interface due to outdiffusion of phosphorus out of the heavily doped polysilicon gate, as shown in Fig. 3.23. The peak dopant density directly at the interface was about \( 10^{18} \text{cm}^{-3} \) which would lead to a large threshold voltage. By counterdoping with boron it was possible to reduce the net active dopant concentration to the \( 10^{16} \text{cm}^{-3} \) range.
Figure 3.20: Simulation of 3-D CMOS inverter, field oxide region

Figure 3.21: Simulation of 3-D CMOS inverter, source/drain region
Figure 3.22: Simulation of 3-D CMOS inverter, common drain (SEG) region

Figure 3.23: Simulation of 3-D CMOS inverter, gate/ELO region
CHAPTER 4

PROCESS DEVELOPMENT II: EXPERIMENTAL

4.1 Introduction

Since device fabrication works with systems that are complex and cannot easily be described by mathematical means the introduction of new processing steps is by nature empirical. New process technologies have a higher probability of being successful if many of its steps have been proven elsewhere. In designing new processes, one has to concentrate on single steps and find criteria and methods by which the results of that step can be evaluated. Finally, all of the building blocks can be assembled to fabricate the intended new device structure.

The process technology developed in this work contained many procedures that are not part of established VLSI processing sequences. These steps have been identified in the previous chapter. This chapter deals with the experimental procedures employed to obtain a set of fabrication methods to build the 3-dimensional structures by ELO as discussed before. Four main areas of research were covered:

1. The fabrication of a reliable, high dielectric strength polysilicon oxide by thermal oxidation
2. The selective growth of epitaxial silicon films at low temperatures
3. The exploration of the parameter space for epitaxial depositions in the Purdue Solid State Laboratory epitaxy reactor
4. The investigation of the new interface created by the lateral epitaxial overgrowth of oxidized polysilicon.

Furthermore, to fabricate the 3-D CMOS inverter or 3-D PMOS stacked transistor a planarization scheme had to be developed.
4.2 Polysilicon

Polycrystalline silicon has been used in semiconductor manufacturing for over two decades. It's importance as gate and interconnect material for MOS processes has recently been complemented by applications in bipolar processes as well. Polycrystalline silicon played a key role in the development of high density, complex structure MOS devices. It allowed for the fabrication of self-aligned transistors, greatly reducing parasitics and increasing circuit density. More complex structures than were previously possible with metal gate electrodes were developed because polycrystalline silicon was compatible with common high-temperature silicon processing and could be conformally deposited over severe topography. In this section some of the more important aspects of polysilicon deposition, doping, and oxidation are reviewed, as far as it concerns the fabrication of state-of-the-art MOS devices.

Polycrystalline silicon is commonly deposited in a reactor operating at low pressures in the range of 100 mTorr. The deposition takes place by decomposition of a silicon-containing gas at the heated wafer surfaces. Silane is the most widely used source gas and pyrolyzes at high enough temperatures according to

\[
\text{SiH}_4(g) \rightarrow \text{Si}(s) + 2\text{H}_2(g) \quad (4.1)
\]

The depositions are usually carried out at temperatures between 580 °C and 650 °C. At lower temperatures deposition rates become unpractically small while at higher deposition rates and in regular commercial reactor systems homogeneous (gas phase) reactions occur, resulting in a loosely adhering film and bad deposition uniformity. Also, as temperature is increased above 650 °C film surfaces become very rough. Deposition rates depend on the temperature as shown in Fig. 4.1 and the silane partial pressure. Since depletion of the source gas occurs as the gas moves down the tube over the wafers, a temperature profile has to be set such that uniform deposition is achieved from the wafers closest to the gas inlet to the wafers closest to the outlet. The temperature profile may have to be ramped by as much as 30 °C, thereby unfortunately also changing some of the polysilicon characteristics which are sensitive to deposition temperature.

The \textit{in situ} doping of polysilicon is possible but can change the deposition rate. Boron doping increases the deposition rate while arsenic and phosphorus doping decreases it (see Fig. 4.2). The \textit{in situ} doping of polysilicon has not found widespread use because of difficulties in maintaining film uniformity.
Figure 4.1: Polysilicon growth rate as a function of temperature\textsuperscript{227}

Figure 4.2: Polysilicon growth rate as a function of doping\textsuperscript{74}
over a wafer and from wafer to wafer.

The fabrication of good quality SiO₂ layers on top of polysilicon is essential for the fabrication of stacked gate structures. Such configurations were first used for EPROMs and EEPROMs. Oxides grown thermally on polysilicon exhibit breakdown characteristics that are considerably worse than that of oxides grown on single crystalline substrates. While thermally grown oxides on substrates typically have breakdown field strengths of about 8 MV/cm, oxides on polysilicon may break down at 2 to 4 MV/cm. This causes problems in scaling processes that rely on polysilicon oxides as a gate dielectric.

One of the most extensive studies of the oxidation of polycrystalline silicon was done by E.A. Irene et al.. They found that the polyoxide was not uniformly thick, with thinner oxide near to the previous polysilicon grain boundaries. They determined that intergranulary oxidation occured and led to an increase in surface roughness. When the polysilicon was doped with phosphorus prior to oxidation the surface was smoother than that of oxide on undoped films, and stress in the polysilicon film was significantly reduced. This was attributed to the enhanced mass flow of silicon in the presence of phosphorus.

High conductivity in polyoxides was first reported by DiMaria and Kerr. The relationship between enhanced conductivity in polyoxides and surface asperities and oxide thickness undulations was then established by Anderson and Kerr. This understanding led to the development of methods that aimed at the fabrication of polyoxides with few and shallow asperities. It has been demonstrated that one can increase the quality of the oxide by oxidizing a layer of amorphous silicon instead of polycrystalline silicon. The small grain size prior to oxidation results in rapid recrystallization during the first few moments at the elevated temperature of the oxidation cycle. The surface however remains smooth, with roughness in the order of the small cluster size of the amorphous silicon. Shinada et al. studied the breakdown field dependence on the phosphorus concentration in the polycrystalline film and demonstrated an increase in the critical field with a pre-oxidation anneal. One of the highest strength polyoxides were reported by Alvi et al.. They employed rapid thermal processing for the growth of polyoxide films.

While there is general consent that the critical electric field depends on the doping concentration in the polysilicon prior to oxidation, the optimum values vary depending on the kind of processing used. Shinada et al.
found their highest strength oxide at a phosphorus concentration of $6 \times 10^{20} \text{cm}^{-3}$ while Alvi et al. found a peak at $1.4 \times 10^{20} \text{cm}^{-3}$, independent of dopant species, by employing rapid thermal processing. All authors found that oxidation at higher temperatures, above $1100 \degree \text{C}$, yields higher critical electric fields. They attributed this to a partial reflow of matter at the surface as well as operation under diffusion-limited conditions that reduce the influence of crystal grain orientation dependence on the oxidation rate.

In this study experiments were conducted to investigate the breakdown electrical field strength of oxidized polysilicon with regard to dopant type, doping density, and polysilicon deposition temperature in our laboratory. No differences were observed between arsenic and phosphorus doped poly. For a doping concentration of $7 \times 10^{20} \text{cm}^{-3}$, critical electric fields of about 4 MV/cm were measured, (from I-V characteristics of Fig. 5.7). This field was estimated from the voltage at which the current density reached $1 \text{nA/mm}^2$ divided by the measured thickness of the oxide. The resulting poly sheet resistance was about $50 \Omega/\square$ and thereby suited for circuit applications.

### 4.3 Oxide Degradation During SEG

In SOI by ELO, the polysilicon gate oxide of the top device functions at the same time as the masking oxide. The integrity of this thin oxide is essential for the fabrication of the stacked shared gate structures as illustrated in Fig. 3.1.

Investigations of the effect of post-oxidation anneals (POA) on the breakdown characteristics of thin, thermally grown oxides showed that a considerable number of oxide defects can be induced when the anneal is carried out in an oxygen free or oxygen deficient ambient. Such conditions can exist in epitaxial reactor systems, where the amount of residual water vapor and oxygen has to be minimized to obtain material of good crystalline quality. Initial experiments with polyoxides exposed to selective epitaxial growth ambients at $950 \degree \text{C}$ showed that oxides less than about 150 nm thick had extremely small breakdown fields, as shown in Table 3. In an additional experiment the same basic behaviour was observed for oxides on bulk substrates. This prompted a systematic study to understand the mechanism of this oxide deterioration and to find a solution so that work towards fabrication of a three-dimensional stacked device could continue.

In this study the effect of $\text{H}_2$, $\text{H}_2-\text{HCl}$, and $\text{SiCl}_2\text{H}_2-\text{HCl-}\text{H}_2$ ambients at $950 \degree \text{C}$ and 150 Torr, conditions typical for SEG, on the electrical breakdown
characteristics of thin, thermally grown oxides was investigated. Metal-oxide-semiconductor capacitors were built on 100 Ω-cm, n-type <100>-oriented silicon wafers. The wafers were cleaned in solutions of hot H₂SO₄ - H₂O₂ and buffered HF and then oxidized in dry O₂ at 1100 °C to form oxides from 60 to 120 nm thick. Oxide thicknesses were measured by ellipsometry. The wafers were divided into four groups, each group containing oxides of nominal 60 nm, 80 nm, 100 nm, and 120 nm thickness. To investigate the effect of the various gas components of the ambient used in SEG, these groups were subjected to different processing steps.

The first set of four (group I) was metallized with a layer of less than 300 nm of Al-Si in a sputtering system, patterned, and annealed at 450 °C in dry N₂ for 20 minutes. These wafers served as a standard against which the other wafers were compared.

The second set of wafers (group II) was placed into an inductively heated, pancake type epitaxy reactor. The temperature was ramped to 900 °C, with H₂ as the only gas introduced into the reaction chamber. After a 5 minute bake at 900 °C at atmospheric pressure, the temperature was raised to 950 °C and the chamber pressure lowered to 150 Torr. H₂ was still the only gas introduced into the chamber. The wafers were subjected to these conditions for 20 minutes, after which they were taken out of the reactor. The wafers were metallized, patterned, and annealed as described for the control set (group I).

<table>
<thead>
<tr>
<th>Oxide Thickness</th>
<th>Leakage</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 nm</td>
<td>2.6 mA/cm²</td>
</tr>
<tr>
<td>60 nm</td>
<td>1.8 mA/cm²</td>
</tr>
<tr>
<td>100 nm</td>
<td>2.3 mA/cm²</td>
</tr>
<tr>
<td>200 nm</td>
<td>49 pA/cm²</td>
</tr>
</tbody>
</table>

Table 3
Polyoxide leakage at 5V Bias after 40 min. exposure to SEG ambient at 950 °C and 150 Torr.
The third group of wafers (group III) was subjected to the same procedure as the second group, with the exception that HCl was introduced into the chamber in addition to the H₂ during the 20 minute bake at 950 °C at 150 Torr. The HCl partial pressure was about 3 Torr. These wafers were also metallized, patterned, and annealed.

The fourth set of wafers (group IV) was placed into the epitaxy reactor and a complete selective epitaxy run was carried out. The epitaxy cycle consisted of an in situ 5 minute H₂ bake at 900 °C at atmospheric pressure. Then the temperature was ramped to 950 °C and the pressure lowered to 150 Torr. HCl and SiCl₂H₂ (DCS) were added as reactant gases. The partial pressures of these gases were in the order of 1 Torr. All runs were carried out for 20 minutes such that no polysilicon nucleated on the oxide surface. After the epitaxy the wafers were metallized, patterned, and annealed as described above.

Some wafers were prepared at Harris Semiconductor, Applied Materials, and at the Institut für Mikroelektronik for control purposes. They were subjected to the same conditions as were the four groups already mentioned, with the exception that radiant heated barrel reactors were employed to perform the epitaxy.

For all our runs, the water vapor content in the H₂ carrier gas was below 1 ppm. Temperatures were measured with an optical pyrometer and corrected for both the emissivity of the substrates and the absorption in the bell jar glass windows.

The breakdown field distribution for each wafer was acquired by stressing the test capacitors with a stair case current ramp and detecting the voltage at which irreversible breakdown of the oxide took place.¹⁷¹,¹⁷² To compute the field, the voltage was divided by the respective oxide thickness. The test device was a rectangular capacitor with an area of 9x10⁻⁴ cm². Each wafer contained 150 of these capacitors. The oxide quality was expressed in terms of the two statistical parameters, \( \bar{E} \) the average breakdown field, and the defect density \( D \) given by the following relations ¹⁷³-¹⁷⁵

\[
\bar{E} = \frac{1}{N} \sum_{i=1}^{N} E_i
\]

\[
D = \frac{1}{A} \ln Y
\]

where \( A \) is the gate electrode area and \( Y \) is the fraction of devices with breakdown fields greater than 6 MV/cm. The defect densities calculated from the
Figure 4.3: Yield by processing group for oxides of different thickness
Poisson distribution (4.3) were within 5% of those calculated by the more exact binomial distribution formula, for the observed yields and sample size of this study.

**Table 4**

Average breakdown field $\bar{E}$, defect density $D$, and standard deviation of breakdown field $\sigma$ for the four different groups of 80 nm oxides.

<table>
<thead>
<tr>
<th>Group</th>
<th>$\bar{E}$ [MV/cm]</th>
<th>$D$ [cm$^{-2}$]</th>
<th>$\sigma$ [MV/cm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (I)</td>
<td>8.1</td>
<td>29</td>
<td>0.94</td>
</tr>
<tr>
<td>$H_2$ bake (II)</td>
<td>5.28</td>
<td>703</td>
<td>3.12</td>
</tr>
<tr>
<td>$H_2$ + HCl (III)</td>
<td>7.72</td>
<td>129</td>
<td>2.25</td>
</tr>
<tr>
<td>Selective Epitaxy (IV)</td>
<td>2.27</td>
<td>1433</td>
<td>3.36</td>
</tr>
</tbody>
</table>

Figure 4.3 plots the yield for the four different groups of wafers. As expected, the control wafers (group I) yield almost 100% good devices and an $\bar{E}$ of 8 MV/cm. The yield decreases noticeably for wafers baked in $H_2$ and HCl at 950 °C and 150 Torr for 20 minutes (group III), and even more when SiCl$_2$H$_2$ (DCS) was added to the ambient (group IV). Thinner oxides are more severely affected than thicker ones. The low yield observed for the wafers of group II (just $H_2$ bake) might have been caused by problems in the cleaning process before oxidation.

In Figs. 4.4 through 4.7 the breakdown field distributions for the four groups of 80 nm oxides are given. One can observe a slight increase in devices with low breakdown fields after the $H_2$-HCl bake. The only peak in the distribution originates from the good devices at fields of about 8 MV/cm. After 20 minutes of selective epitaxy the distribution has shifted. More than half of the devices are short circuited. There are a few capacitors with breakdown voltages somewhere between 0 and 8 MV/cm; about 30% of the devices breakdown at fields greater than 6 MV/cm. Qualitatively equivalent results were obtained from the wafers processed in the industrial facilities. However, there the yield was generally higher. In the following discussion these results will be correlated with those obtained by other researchers in related areas.
Figure 4.4: Breakdown field distribution for 80-nm control oxides (group I)

Figure 4.5: Breakdown field distribution for 80-nm H₂ bake oxides (group II)
Figure 4.6: Breakdown field distribution for 80-nm H₂ + HCl oxides (group III)

Figure 4.7: Breakdown field distribution for 80-nm H₂ + HCl + DCS oxides (group IV)
Figure 4.8: Regions of silicon dioxide growth as a function of oxygen partial pressure and temperature

Figure 4.9: Regions of silicon dioxide growth as a function of water vapor partial pressure and temperature
On a clean silicon surface, etching occurs in the presence of oxygen or water vapor when the O\(_2\) or H\(_2\)O partial pressures are smaller than the equilibrium partial pressure of SiO\(_2\) which is a strong function of temperature, as illustrated in Figs. 4.8 and 4.9. At 950 °C, partial pressures of O\(_2\) or H\(_2\)O of less than about 0.5 mtorr lead to etching of exposed silicon by formation of volatile SiO.

In the Si-SiO\(_2\) system, formation of SiO can occur via the disproportionation reaction

\[
\text{Si} + \text{SiO}_2 \rightarrow 2 \text{SiO(v)}
\]

for sufficiently low O\(_2\) and H\(_2\)O partial pressures at the reaction site. Such conditions have been achieved in the investigation of post-oxidation anneals of thermally grown SiO\(_2\) by Hofmann et al. They reported that between 750 °C and 1100 °C oxide decomposition was non-uniform, initiated at nucleation sites and resulting in voids in the oxide. Intentional damage to the silicon surface by Ar and Cu implantation before oxidation increased the number of voids in the oxide after annealing. A similar observation was also made by Honda et al. when investigating the effect of Fe impurities at SiO\(_2\)/Si interfaces. Some defects are not always electrically detectable. Lopez reported on defects that had electrical characteristics comparable to that of defect free oxide, but were chemically unstable.

In silicon epitaxy one likes to minimize the amount of H\(_2\)O and O\(_2\) in the growth ambient, to obtain good crystalline quality material. Here the reaction described by (4.4) is used to remove native oxide from the growth surface during the in situ precleaning step. In SEG, this can result in an undercut of the masking oxide layer at the Si-SiO\(_2\) interface.

At 950 °C, the following conditions must be met for reaction (4.4) to occur at measurable rates: the H\(_2\)O or O\(_2\) partial pressures have to be lower than 0.5 mtorr, silicon must be supplied to the reaction site, and the reaction product SiO must be able to leave the reaction site. For the POA investigations, Hofmann et al. suggested that in their case the last two conditions were met by enhanced diffusion of Si or SiO at defects in the oxide. Tabe demonstrated uniform etching of oxide films in ultra-high vacuum when he supplied silicon by sputtering from a polysilicon target. For the removal of native oxide and undercut of masking oxide, silicon is available from the substrate.
In the Purdue epitaxy system, all three conditions were met. Water content in the feed gas was always less than 1 ppm (typically 0.4 ppm). For a deposition pressure of 150 Torr, that would translate into a water vapor partial pressure of about 0.2 mtorr, which at 950°C is less than the equilibrium partial pressure of SiO.

The oxides that were treated in the epitaxy reactor with just H₂ or H₂ and HCl can be compared to those that went through the POA procedures as described by others. Here oxide decomposition would begin at weak points in the oxide, that may be caused by pre-oxidation substrate defects as well as post-oxidation contamination. It is not entirely clear whether the oxide decomposition starts at the Si-SiO₂ interface or at the surface. In either case, the proposed enhanced diffusion of either Si or SiO at electrically undetectable defect sites may be the reason for the observed void formation.

The oxides that went through a regular selective epitaxy, with DCS as the silicon source gas, exhibit a large number of defects. Here additional silicon is supplied to the oxide surface from the gas phase, by decomposition of the silicon source gas DCS at the hot surface. The abundance of Si can accelerate reaction (4.4), leading to higher defect densities. In addition, the reaction product SiO is easily transported away from the reaction site at the oxide surface.

Taking the equilibrium partial pressure of SiO as the dividing line between the region of oxide decomposition and oxidation, Fig. 4.10 shows the critical reactor pressure as a function of deposition temperature, with the carrier gas moisture content as a parameter. The regions right of the curves denote the conditions under which defects can be induced by oxide decomposition. One can see that the critical conditions fall right into a range where selective epitaxial depositions are commonly carried out. For depositions at 150 Torr and a moisture content of 1 ppm, the maximum allowable deposition temperature would be about 920°C.

This study gave evidence that thin, thermally grown silicon dioxide films decompose during selective epitaxy, similar to the defect formation seen after high-temperature annealing. The etching of oxide, which was non-uniform, was enhanced significantly by the presence of DCS as a silicon source gas. The water vapor and oxygen free environments necessary to obtain high crystal quality epitaxial silicon pose a problem for thin (<200 nm) oxide films during SEG and ELO. The critical partial pressures of water vapor or O₂ for oxide etching to occur depend exponentially on temperature. For smaller
Figure 4.10: Critical conditions for oxide decomposition by SiO formation to take place
temperatures smaller partial pressures are required for oxide decomposition to take place. Thus, the problem of oxide degradation could be solved by depositing the epitaxial silicon at lower temperatures or by chemically stabilizing the masking oxide layer by nitridation. This last method does however require equipment that is not available in most VLSI laboratories and was therefore not considered any further.

![Graph]

Figure 4.11: Yield for 3-D capacitors with epitaxy grown at 860°C, 900°C, and 950°C

To test the proposed explanation of the oxide degradation mechanism stacked capacitors as shown in Fig. 3.4 were built where the epitaxy was carried out at three different temperatures, namely 860°C, 900°C, and 950°C. The respective yields are plotted in Fig. 4.11. It can be seen that there are a large number of functional devices when the epitaxy was carried out at 860°C. At 900°C the yield dropped considerably, and at 950°C there were no functional devices. For the 860°C and 900°C runs operating conditions were such that they were just outside the shaded areas in Fig. 4.10. For the epitaxy run done at 950°C conditions were such that the operating point was well inside the respective shaded area.
4.4 Silicon Epitaxy

As had become evident from the study on gate oxide integrity, a good quality, intact gate oxide for the stacked device necessitated the development of an SEG procedure by which silicon of good crystalline quality could be grown at low temperatures. As epitaxy is carried out at lower temperatures, material quality usually deteriorates. To successfully develop the proposed three-dimensional technology a detailed understanding of the mechanisms that cause this deterioration as well as an understanding about the limitations in SEG with regard to low temperature were essential. For that reason the relationship between material quality, pre-clean procedures, deposition temperatures, deposition pressures, and epitaxial silicon quality was studied.

A second area of concern was the determination of deposition parameters for selective epitaxy as well as an understanding of the deposition process itself. Selectivity, growth rate uniformity, temperature dependence, and loading effects need to be considered when exploring the parameter space. This part of the study was conducted in collaboration with the School of Chemical Engineering at Purdue.

4.4.1 Crystalline Quality

Future high performance integrated circuits will require smaller geometries in lateral as well as vertical dimensions and abrupt transitions in dopant concentration between adjacent single crystalline layers. These requirements cannot be met by conventional high-temperature (>1000 °C) epitaxy. Goals in developing a reduced-temperature epitaxy are to minimize dopant outdiffusion and autodoping, to reduce dopant redistribution and temperature related stressing, minimize pattern distortion, achieve good uniformity, and to obtain defect free material. For the process developed in this work low-temperature epitaxy was essential to maintain the integrity of the stacked device gate oxides.

A considerable amount of research has been done on low-temperature epitaxy. Meyerson et al. 182,183 have demonstrated that device quality material can be obtained at temperatures as low as 750 °C in an ultra-high vacuum epitaxy system. Donahue and Reif 184 employed plasma in their studies of low-temperature epitaxy. L. Vescan et al. 185 achieved good quality material in a specially designed ultra-high vacuum reactor down to 760 °C. In commercially available reactor systems, Borland and Drowley34 and Drowley and
Turner\textsuperscript{186} were able to grow device quality epitaxial layers in the pressure range of 10 to 100 Torr and at temperatures of about 850 °C.

In this study experiments were conducted that elucidate the importance of the background water vapor and oxygen levels in the epitaxial growth environment, as previously demonstrated by Meyerson for an ultra-high vacuum reactor. The results from Ghidini and Smith\textsuperscript{176,177} on \(\text{Si}/\text{O}_2/\text{H}_2\text{O}/\text{SiO}_2\) equilibrium data were applied to define an operating range for commercially available epitaxy reactor systems where good crystalline quality material can be grown.

In their investigation of the interaction of \(\text{H}_2\text{O}\) and \(\text{O}_2\) with \(\text{Si}(111)\) and \(\{100\}\) surfaces, Ghidini and Smith showed that there exist critical pressures for water vapor and oxygen below which etching of silicon by the reactions

\[
\text{Si(s)} + \text{O}_2 \rightarrow 2\text{SiO(g)} \quad (4.5)
\]

and

\[
\text{Si} + \text{H}_2\text{O} \rightarrow \text{SiO(g)} + \text{H}_2(g) \quad (4.6)
\]

takes place. For partial pressures of water vapor or oxygen higher than these critical pressures, oxidation of silicon occurs. For temperatures greater than about 1000 °C, the critical partial pressures for water vapor, \(P_c(\text{H}_2\text{O},T_s)\), are one order of magnitude higher than those for oxygen, \(P_c(\text{O}_2,T_s)\). However, for temperatures less than 950 °C, \(P_c(\text{H}_2\text{O},T_s)\) and \(P_c(\text{O}_2,T_s)\) are about equal. \(P_c(\text{O}_2,T_s)\) can be expressed as

\[
P_c(\text{O}_2,T_s) = 4.4 \times 10^{12} \exp(-3.93\text{eV}/k_B T_s) \text{ torr} \quad (4.7)
\]

where \(k_B\) is Boltzmann's constant and \(T_s\) is the silicon surface temperature in degrees Kelvin.\textsuperscript{177}

For a given deposition pressure, one can thus compute a critical preclean and growth temperature as a function of moisture and oxygen content in the \(\text{H}_2\) carrier gas, provided it is the dominant source for these contaminants. This is illustrated in Fig. 4.12. For example, if one has a known moisture plus oxygen content of 1 ppm in the \(\text{H}_2\) carrier gas, and the reactor can be operated at a minimum pressure of 25 Torr, the minimum deposition temperature to obtain good crystal quality material would be 875 °C.

Employing the basic chemical equilibrium data for the \(\text{Si}/\text{O}_2/\text{H}_2\text{O}/\text{SiO}_2\) system, as interpreted in Fig. 4.12, experiments were conducted to demonstrate the importance of the parameters mentioned for epitaxial preclean and growth. The substrates used were 3" <100> crystal orientation n-type silicon
Figure 4.12: Critical temperature vs carrier gas moisture content, with reactor chamber pressure as a parameter
wafers, which were oxidized to 80 nm thickness, and patterned to open seed areas. After the wet etching in buffered HF solution, the substrates were cleaned in a hot hydrogen peroxide/sulfuric acid solution followed by a rinse in de-ionized water. They were then dipped in a dilute HF solution to remove native oxide, rinsed again in de-ionized water, dried, and transferred into the epitaxy reactor not later than one hour after the wet clean. The hydrogen carrier gas introduced into the reactor was cleaned by a catalytic converter and dryer positioned close to the reactor chamber. The dew point of the hydrogen before it went into the reactor was monitored for all runs. Temperatures inside the reactor chamber were measured by means of a pyrometer, corrected for the emissivity of silicon. Dichlorosilane was used as silicon source gas, and HCl was added to achieve selective growth. Material characterization was done by scanning electron microscopy (SEM) and Secco defect etching.

Two sets of experiments were carried out. The first set was to investigate the effect of water vapor and oxygen partial pressures on the preclean. The second set was to explore the limits of low defect density crystal growth for the reactor.

The preclean consisted of a 5 minute bake in hydrogen at the temperature and pressure indicated. The bake was followed by a 30 second in situ HCl etch to emphasize the effects of any residual native oxide. This led to uniform etching in areas where the native oxide was removed, and gave the appearance of a pitted surface in case of incomplete native oxide removal. Table 5 gives the results for the matrix of preclean experiments. Note that the last three entries are at temperatures above the minimum, as shown in Fig. 4.12, and result in excellent cleaning. Figure 4.13 and Fig. 4.14 show SEM pictures of the surface after a preclean above and below the critical temperature, respectively, for the given preclean pressure and moisture content. The surface in Fig. 4.14 was etched by the HCl only in some places where oxide existed and is therefore pitted.

For all growth experiments, the preclean conditions were the same and such that a clean un-pitted surface resulted prior to growth. Growth was carried out for 30 minutes. Table 6 presents the results for various growth runs. Note that the first entries are well above the minimum required temperature. It is observed that good crystal quality epitaxy (in this case selective) was also grown at 890 °C and 150 Torr. At 860 °C a high number of hillocks due to local oxidation during the growth was visible.
Figure 4.13: SEM of a seed precleaned at 150 Torr, above the required minimum temperature (point 2 in Fig. 4.12)

Figure 4.14: SEM of a seed precleaned at 760 Torr, just below the required minimum temperature (point 1 in Fig. 4.12)
Figure 4.15: SEM of growth at 150 Torr, above the required minimum temperature (point 4 in Fig. 4.12)

Figure 4.16: SEM of growth at 150 Torr, below the required minimum temperature (point 3 in Fig. 4.12)
Figures 4.15 and 4.16 show growth above and below the critical temperature, respectively, for a deposition pressure of 150 Torr and the moisture content in that particular run. Note that the numbers in Fig. 4.12 refer to Figures 4.13 to 4.16. Points 2, 3, and 4 have to be related to the curve for 150 Torr, while point 1 has to be related to the curve for 760 Torr.

The experimental results strongly support the assumption that the limitations of growth in commercial reactor systems at low temperature are indeed governed by the water vapor and oxygen background pressure during preclean and growth. In order to establish a reduced temperature epitaxy deposition procedure one has to first determine the background partial pressure of moisture and oxygen in the reactor chamber. In case of a leak-tight system, the moisture content in the chamber can be approximated by measuring the dew point of the carrier gas, which is often easier than the measurement of H₂O and O₂ partial pressures inside the reactor. Depending on the low pressure limit of the reactor, one can then pick the minimum deposition temperature from Fig. 4.12. Low partial pressures of oxygen and water vapor can be achieved by low pressure deposition systems, point-of-use purification of the gases, or both. For commercial reduced pressure CVD systems, the lower temperature limit at 0.1 ppm residual moisture and oxygen, and a minimum deposition pressure of 10 Torr, would be about 810 °C. This agrees with the deposition at the lowest temperature reported to obtain defect free silicon in a commercial epitaxy reactor system. The low temperature limit for the Purdue reactor was determined to be about 890 °C.

### Table 5

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Pressure</th>
<th>Moisture Cont.</th>
<th>Visual Appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td>950 °C</td>
<td>760 Torr</td>
<td>0.42 ppm</td>
<td>pitting</td>
</tr>
<tr>
<td>950 °C</td>
<td>150 Torr</td>
<td>0.21 ppm</td>
<td>good</td>
</tr>
<tr>
<td>920 °C</td>
<td>150 Torr</td>
<td>0.26 ppm</td>
<td>good</td>
</tr>
<tr>
<td>880 °C</td>
<td>150 Torr</td>
<td>0.14 ppm</td>
<td>good</td>
</tr>
</tbody>
</table>
Table 6

Growth experiment matrix.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Pressure</th>
<th>Moisture Cont.</th>
<th>Hillocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>950 °C</td>
<td>150 Torr</td>
<td>0.21 ppm</td>
<td>4 cm⁻²</td>
</tr>
<tr>
<td>920 °C</td>
<td>150 Torr</td>
<td>0.26 ppm</td>
<td>3 cm⁻²</td>
</tr>
<tr>
<td>890 °C</td>
<td>150 Torr</td>
<td>0.14 ppm</td>
<td>50 cm⁻²</td>
</tr>
<tr>
<td>860 °C</td>
<td>150 Torr</td>
<td>0.22 ppm</td>
<td>&gt;1000 cm⁻²</td>
</tr>
</tbody>
</table>

Table 7

Defects delineated by Secco etching

<table>
<thead>
<tr>
<th>Prebake Temperature</th>
<th>Prebake Pressure</th>
<th>Deposition Temperature</th>
<th>Deposition Pressure</th>
<th>Defects cm⁻²</th>
</tr>
</thead>
<tbody>
<tr>
<td>930 °C</td>
<td>150 Torr</td>
<td>860 °C</td>
<td>150 Torr</td>
<td>10⁵</td>
</tr>
<tr>
<td>930 °C</td>
<td>150 Torr</td>
<td>890 °C</td>
<td>150 Torr</td>
<td>10⁴</td>
</tr>
<tr>
<td>950 °C</td>
<td>150 Torr</td>
<td>900 °C</td>
<td>150 Torr</td>
<td>10⁶</td>
</tr>
<tr>
<td>950 °C</td>
<td>150 Torr</td>
<td>950 °C</td>
<td>150 Torr</td>
<td>10³</td>
</tr>
<tr>
<td>900 °C</td>
<td>760 Torr</td>
<td>950 °C</td>
<td>150 Torr</td>
<td>10⁴</td>
</tr>
<tr>
<td>900 °C</td>
<td>760 Torr</td>
<td>1000 °C</td>
<td>150 Torr</td>
<td>10⁵</td>
</tr>
</tbody>
</table>
4.4.2 Growth Rate Characteristics

In the last sections two key issues were investigated that were absolutely crucial for 3-D CMOS device fabrication by ELO: the fabrication of high dielectric strength polyoxides that hold up to the environment of SEG, and the epitaxial growth of monocrystalline silicon with few defects at low temperatures. The experiments conducted in this context helped to define regions of operation for the Purdue epitaxy reactor. In addition to masking oxide integrity, morphology, and silicon crystalline quality, the dependence of growth rates on reactor parameters such as feed gas composition, temperature, total gas flow etc. was viewed as important for the process design.

Initial experiments had indicated that there existed gross non-uniformities in growth rates across wafers. Such non-uniformities are difficult to tolerate in device and circuit fabrication. The experiments discussed in the following attempted to gain a better understanding for the cause of growth rate variations across a wafer. The goal of this research was to define a set of parameters where non-uniformities were minimized.

4.4.2.1 Reactor Description

The Gemini I epitaxy reactor in the Purdue Solid State Laboratory is an inductively heated pancake reactor with capability for reduced pressure operation, manufactured by Gemini Research Inc.. A cross section is shown in Fig. 4.17. The round graphite susceptor measures 46 cm in diameter and can accommodate 21 three inch wafers and 2 two inch wafers. The susceptor is factory coated with a silicon carbide film. A polysilicon film of at least 10 \( \mu \text{m} \) thickness is deposited across the susceptor before any epitaxy is carried out. This film is periodically removed and redeposited to remove contaminants and to inhibit the outdiffusion of contaminants from the graphite susceptor. Wafers are placed on the susceptor in recessed spherically grooved pockets. During operation the susceptor rotates counterclockwise at 8 rpm. The temperature of the susceptor is controlled by a closed loop system consisting of a pyrometer, focused on the susceptor, the temperature control unit where set points, ramp rates etc. can be entered, and the 500 kW high frequency generator that delivers its power over a water cooled transmission line to the induction coils embedded in quartz glass and placed underneath the rotating susceptor. The distance between the induction coils and the susceptor can be adjusted to optimize the temperature profile across the susceptor. The pyrometer is adjusted for an emissivity of 0.8, that of silicon. The minimum
Figure 4.17: Cross section of the Purdue Gemini I reactor
temperature that can be measured is about 650 °C.

The reaction chamber is composed of a stainless steel bottom plate and a mobile bell jar made out of quartz glass. For loading or unloading of wafers the bell jar is lifted up to give access to the susceptor. During operation the bell jar is air cooled while the bottom plate and walls of the reactor housing are water cooled.

Exhaust gases are pumped off through an outlet in the bottom plate underneath the susceptor. The vacuum pump has a throttle valve regulated by a pressure controller. During operation, the pressure inside the reaction chamber can be reduced down to 150 Torr. This lowest-pressure limit is given by the onset of arcing and plasma generation of the carrier gas by the induction coils at the high voltages and powers employed. Arcing can lead to a partial destruction of the quartzware and the rf generator.

Four gases are connected to the reactor: dry nitrogen, hydrogen, dichlorosilane, and hydrochloric acid. The hydrogen is purified by a catalytic converter and drier to remove oxygen and moisture before it enters the reaction chamber. The other gases are introduced as they come out of their storage tanks.

The growth experiments were carried out on (100) and (111) n-type silicon substrates. The wafers were initially cleaned in a hot H2SO4/H2O2 solution. After a thorough rinse in de-ionized water and a blow-dry with dry nitrogen they were transferred to an oxidation furnace. Oxidation temperatures were usually 1000 °C and oxidation took place in a steam atmosphere. Subsequently the substrates were patterned. For the (100) wafers, the edges of the rectangular pattern were oriented along the <100> directions and all wafers were placed in the reactor such that the [001] direction was aligned parallel to the radial of the round susceptor. The seed areas were opened by wet etching in buffered HF solution. Just prior to the epitaxy the wafers were cleaned again in organic solvents followed by a soak in a hot H2SO4/H2O2 solution. The wafers were rinsed off again in de-ionized water and dipped in a dilute HF solution to remove native oxide. After an additional rinse in de-ionized water and drying with dry nitrogen the wafers were transferred directly into the epitaxy reactor. In a hydrogen ambient at atmospheric pressure, the wafers were heated up to 920 °C or the following deposition temperature, whatever was higher, and subjected to a 5 minute bake to remove any native oxide from the wafer surface. Subsequently the susceptor temperature was brought to the desired deposition temperature, HCl was introduced, and
the substrates were etched for 30 seconds. Thereafter the chamber pressure was reduced to 150 Torr and DCS was added to commence with the deposition. The typical duration of growth was 20 minutes. In some instances this scheme was slightly modified; in particular the cleaning cycle was modified in later stages of the investigation and then always took place at the reduced pressure of 150 Torr.

For each of the wafers, SEG/ELO growth measurements were taken at 12 different points using a Tencor Alpha-Step profilometer and averaged. The absolute experimental error for growth measurements using this instrument was estimated to be 10 nm. The seed window geometries and locations within each die measured for intra-wafer or inter-wafer growth rate comparisons were identical.

4.4.2.2 Reactor Feed Composition Dependence

In the past various researchers have investigated the dependence of growth rates on the ratio of HCl/DCS partial pressures in the feed for different reactor types and geometries. For the pancake type reactor, we have studied this dependence for a total gas flow of 60 standard liters per minute (slm), a temperature of 950 °C, and a deposition pressure of 150 Torr.

Most investigations reported in the literature proceeded by fixing the DCS flow rate while varying the HCl flow. The resulting growth rates project a linear dependence between growth and the HCl/DCS flow rate ratio in the feed gas. Qualitatively, as the amount of HCl is increased growth rates decrease. The increase of DCS causes a corresponding increase in growth rates. The HCl/DCS ratio is often taken as the figure of merit; the higher this ratio the less growth is expected. However, we found that the HCl/DCS ratio cannot be used to compare results from experiments where the DCS flow rate was changed. In such cases even higher ratios of HCl/DCS can lead to higher growth rates instead to lower growth rates, as shown in Fig. 4.18. M. Kastelic140 has instead devised the measure as the quantity HCl²/DCS which is a more accurate figure to compare results from different experiments, as illustrated in Fig. 4.19.

For the pancake type reactor investigated we found a considerable change in growth rates along the radial direction across the susceptor (Fig. 4.20). Growth rates dropped basically independent of feed composition from the outside of the susceptor towards the inside. A maximum existed close to the perimeter of the susceptor, and a minimum close to the center. The absolute
Figure 4.18: Growth rate vs. HCl/DCS ratio of feed

Figure 4.19: Growth rate vs. HCl²/DCS ratio of feed
Figure 4.20: Growth rate profile for various feed gas compositions

Figure 4.21: Non-uniformity as a function of feed gas composition
difference between maxima and minima was only a weak function of the absolu­
tate growth rate. However, as soon as net zero growth or etching was
approached the slope in the curve went to zero and the profile became much
more uniform.

The amount of non-uniformity decreased as growth increased, as is evi­
dent from Fig. 4.20. If non-uniformity is defined as

\[
\text{%Non-Uniformity} = \frac{\text{GR}_{\text{max}} - \text{GR}_{\text{min}}}{\text{GR}_{\text{max}} + \text{GR}_{\text{min}}} \times 100
\]

where \( \text{GR}_{\text{min}} \) and \( \text{GR}_{\text{max}} \) represent the minimum and maximum growth rates
respectively, measured along the radial of the susceptor, and if this quantity is
plotted vs. the HCl\(^2\)/DCS ratio of the feed, as shown in Fig. 4.21, it turned
out that the most uniform films were obtained right in the transition region of
selective growth and polynucleation on the masking oxide. Uniformity was
worst close to average net zero growth.

Selective epitaxial growth can be viewed as the result of a deposition
reaction and a competing etching reaction by HCl. To investigate the contri­
bution of the etching reaction to growth rate non-uniformity independently an
experiment was conducted where no DCS and only \( \text{H}_2 \) and HCl were intro­
duced into the reaction chamber. The resulting etch rate profile is illustrated
in Fig. 4.22. Etch rate uniformity appears to be much better than growth rate
uniformity when growth is carried out on a comparable wafer under selective
conditions.

4.4.2.3 Masking Oxide Dependence

Preliminary experiments had indicated that there was a dependence of
growth rates on masking oxide thickness. Even though the technological impli­
cations of this effect weren’t viewed as too significant, some effort was
expended to understand the nature of this dependence and to investigate a
possible dependence of growth rate uniformity on oxide thickness.

In this series of experiments, each epitaxial run included three groups of
two wafers with oxides of 80 nm, 460 nm, and 870 nm thickness. One run
included wafers with oxides of 80 nm, 300 nm, 460 nm, 550 nm, 730 nm, and
870 nm thickness. Oxide thickness was determined by ellipsometry and
profilometer scans. In some of the runs (111) oriented wafers were also
included for comparative purposes.
To observe the behaviour of growth in an "active area" type region, with a recessed thin (80 nm) oxide enclosed by thick (900 nm) field oxide, and to investigate the possibility of uniform growth out of seeds defined by masking oxides of different thickness, the structures illustrated in Figs. 4.23 and 4.24 were included on the test mask. The width of the thin oxide region varied from 50 to 400 \( \mu \text{m} \), to investigate the effect of local oxide thickness on surface diffusion. The seed windows for these structures were kept at a fixed size. In an additional experiment, the thick and thin oxides in Fig. 4.23 were interchanged, as illustrated in Fig. 4.24.

To distinguish the effects of radiant heat transfer and heat transfer by conduction, some wafers had thick oxides (> 350 nm) on their backside, which increased the heat resistance from the susceptor to the wafer front surface. In addition, some wafers were included that were just 0.5 mm thick, while all others were 0.7 mm thick. In all cases the test pattern was identical.

The experimental error was dominated by the accuracy of the oxide thickness and epitaxial growth rate measurements. The total relative measurement error was less than ±2 percent.
Figure 4.23: Oxide thickness characterization test pattern cross section

Figure 4.24: Oxide thickness characterization inverted test pattern
Figure 4.25 shows the growth rate plotted as a function of the oxide thickness for several different reactant partial pressures. In these experiments the oxide was of uniform thickness across the measurement pattern. The growth rate dropped noticeably with increasing masking oxide thickness from 80 nm oxides to the next larger value. The differences in growth rates between all following data points were considerably smaller.

The total amount of growth rate reduction was a function of the input HCl and DCS gas partial pressures and thereby of the absolute growth rate, as shown in Fig. 4.25. The absolute change in growth rate with oxide thickness became larger with higher growth rates, i.e. smaller values of F, which was defined as the HCl mole fraction squared to the mole fraction of DCS. However, the relative (percentage) difference between the 80 nm growth rate and the 870 nm growth rate was larger for smaller growth rates, i.e. larger values of F. For the lowest value of F, the growth rate was 0.28 μm/min for 80 nm oxides and it dropped to 0.19 μm/min for the 870 nm oxide; a change of 30 %. The variation between the 80 nm oxide and the 300 nm oxide amounted to 25 %. For the highest HCl input partial pressure, (also highest value of F in Fig. 4.25) the variation between the thickest and thinnest oxides was 50%. A consistent dependence of growth rate uniformity on masking oxide thickness was not observed.

Oxides grown on the wafer backsides did not influence growth rates, even though oxide has a much smaller thermal conductivity than silicon. In addition, wafer thickness did not have any measurable effect on growth rates.

For the structure in Fig. 4.23, growth rates from seeds in the thin oxide windows were the same as those for seeds in the thick field oxide. Growth rates for the thick oxide seeds in the structure of Fig. 4.24 were the same as those for seeds in the thin field oxide. For either structure, the "active area" window size did not have any measurable effect on growth rates. Growth rates for the oxide structures in Fig. 4.23 and 4.24 were not affected by the different oxide thickness in the windows, but instead determined by the large area field oxide thickness.

A significant temperature difference can exist between the wafers and the susceptor, the susceptor being up to 40 °C hotter than the wafers. Also a steep temperature profile in the vertical direction can exist above the susceptor. Heat transfer to the wafer proceeds by radiation and conduction, while heat transfer from the wafer is by convection, conduction, and radiation.
Figure 4.25: Growth rate as a function of masking oxide thickness

Figure 4.26: Growth rate as a function of masking oxide thickness compared to a model where only thermal conduction is considered
The dependence of growth rates on oxide thickness, as plotted in Fig. 4.25, must be related to the wafer surface temperature, since all wafers of each curve were subjected to identical conditions. Also, the percentage change in growth rates for wafers with thin oxides to wafers with thick oxides was larger for smaller growth rates, being consistent with the observed temperature sensitivity of growth rates, which is largest for small growth rates.\textsuperscript{140}

Three mechanisms can be identified for heat transfer: convection, conduction, and radiation. We can assume that heat transfer to the wafer from the heated susceptor is independent of the front surface oxide thickness. Thus, any change in front surface temperature has to be due to a difference in heat transfer through and away from the wafer. In the following we will discuss each mechanism independently.

It is unlikely that convection plays a major role in the observed effect. Since the reactor is operated at reduced pressure and high gas flow rates, forced convection, which is equal for all wafers, dominates over free convection. Furthermore, the convective heat transfer rate from wafers with lower surface temperature would always be smaller than from those with higher surface temperature, thereby counteracting the observed effect. Differences in convection thus cannot account for the increased drop in temperature over wafers with thicker oxides. Lower surface temperatures lead to smaller convective heat transfer rates, which in turn reduces the temperature drop across the thermal resistance of the substrate. For simplicity and as a worst case assumption, we can therefore treat the convective heat transfer rate away from the wafer surface as independent of oxide thickness in the temperature range of interest.

Considering heat transfer by thermal conduction, one has to investigate the thermal conductivity of the wafers perpendicular to the wafer surface. Assuming a simple one-dimensional model based on Fourier's law, one can write for the thermal resistance

\[ R = K \left[ \frac{t_{\text{oxide}}}{k_{\text{SiO}_2}} + \frac{t_{\text{substrate}}}{k_{\text{Si}}} \right] \]  

(4.9)

Here $K$ is a constant, $k_{\text{SiO}_2}$ and $k_{\text{Si}}$ are the thermal conductivities for oxide and silicon, respectively, and $t_{\text{oxide}}$ and $t_{\text{substrate}}$ are the thickness of the oxide and the silicon substrate thickness, respectively. Approximating the temperature difference between the wafer front and the wafer backside as about 30 °K, one can obtain a heat transfer rate $q_k$ according to
\[ q_k = \frac{\Delta T}{R} \] (4.10)

where \( \Delta T \) is the temperature drop across the entire wafer perpendicular to its surface. This rate is the same as the rate of heat transfer by convection, which as a worst case we could treat as being constant. Thus, \( \Delta T \) and \( R \) are in a direct proportion to each other. More accurate of course, as \( R \) becomes greater, \( q_k \) becomes smaller. If the growth rate is expressed as an Arrhenius type equation, with an activation energy \( E_A \), one can write

\[
GR = C \exp \left[ \frac{-E_A}{k(T_{\text{susceptor}} - q_k R)} \right] \] (4.11)

where \( C \) is a constant, \( k \) is the Boltzmann constant, \( T_{\text{susceptor}} \) is the susceptor temperature, \( q_k \) is an approximated average heat transfer rate, and \( R \) is the thermal resistance as defined in equation (4.9). With thermal conductivity data from a reference,\(^{188}\) and an estimated worst case activation energy of 100 kcal/mol, normalized growth rates according to this simple model are compared with measured data in Fig. 4.26. It is apparent that differences in thermal conductivity cannot by themselves account for the drastic change in growth rates that is observed. Furthermore, if a temperature drop due to heat conduction through the wafer was the primary source for surface temperature differences, the wafers with backside oxides should have exhibited lower growth rates, and the thin wafers should have had higher growth rates; however, this was not the case.

Convective and conductive heat transfer cannot account for the drop in growth rates seen for wafers with masking oxides thicker than 300 nm. The only mechanism left is radiative heat transfer. However, since both materials, oxide and silicon, are translucent materials, and since the radiative properties of a surface depend very much on surface conditions which are difficult to evaluate, analytical expressions for radiative heat transfer from oxidized wafer surfaces are difficult to obtain. It is interesting to note that wave lengths in the infrared spectrum are of the same order as the oxide thickness.

From the experiments with locally varying masking oxide thickness (Figs. 4.23 and 4.24) it was concluded that the observed growth rate dependence is a bulk effect and not due to local variations in temperature and concentration of reactants. Surface diffusion lengths, at the employed pressures and temperatures, are in the order of 10 to 50 \( \mu m \). From ELO nucleation experiments, it was known that surface diffusion of adatoms towards the growth site contributes significantly to growth.\(^{189}\) Yet, even for 400 \( \mu m \) wide thin or thick oxide
regions we did not observe any change in growth rate. We believe that the thin or thick oxide window area is too small to cause local temperature differences that would lead to a measurable change in growth rate.

In conclusion, the wafer surface temperature is strongly influenced by the radiative heat transfer properties of the masking oxide layer. It is therefore important to consider wafer surface films, i.e. film material and thickness, when transferring growth parameters from one process to another or when comparing growth data between different systems. For technological applications it is safe to assume that it would be possible to grow epitaxial silicon from seeds masked with oxides of different thickness on the same wafer with good uniformity.

4.4.2.4 Temperature Dependence

The temperature dependence of non-selective growth from DCS has been studied by many authors. The deposition process is usually described as being of Arrhenius type with associated activation energies between 6 and 45 kcal/mole. From experiments conducted in our reactor an activation energy of about 20 kcal/mole can be calculated (Table 8). This implies that even above 950 °C depositions will be sensitive to temperature.

### Table 8

Temperature dependence of growth from 0.36 vol % DCS in H₂ at 150 Torr.

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Growth Rate (μm/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>900</td>
<td>0.134</td>
</tr>
<tr>
<td>950</td>
<td>0.189</td>
</tr>
<tr>
<td>1000</td>
<td>0.209</td>
</tr>
</tbody>
</table>

Figure 4.27 illustrates the growth profile obtained for depositions at various temperatures under non-selective conditions on bulk wafers. First, uniformity was much better than for depositions on masked wafers under selective conditions. Second, the observation made before that growth rates increased towards the perimeter of the susceptor was reversed for depositions carried out
Figure 4.27: Bulk (unmasked) silicon growth rate profiles at 920 °C, 950 °C, and 1000 °C

Figure 4.28: Selective epitaxy growth rate profile on unpatterned wafers at 950 °C and 1000 °C
at 950 °C and 1000 °C. When deposition was again carried out on bulk wafers, but under selective conditions, the growth profiles became somewhat less uniform, but still exhibited much less difference of growth between the inner and outer pocket positions than observed for selective depositions on masked wafers, as is illustrated in Fig. 4.28.

For masked wafers under selective conditions, growth tended to become more uniform as deposition temperature was increased (Fig. 4.29). In this graph, depositions conducted at 860 °C and 1000 °C required different feed gas compositions to maintain selectivity, and thus only the growth rate profile and not absolute growth rates can be compared directly.

To calibrate the temperature controller of the epitaxy reactor, an experiment was carried out where pieces of germanium were placed on bare silicon wafers that were then placed inside the reactor. One such wafer was put in an inner pocket while another was placed in an outer pocket position. The susceptor was slowly heated up and the temperature reading was recorded at which the germanium melted on the silicon wafers. The melting point of germanium was taken as 936 °C. From these experiments it was concluded that there existed a radial temperature profile of between 5 and 15 °C across the susceptor, and that the temperature reading through the pyrometer indicated temperatures that were about 20 °C lower than the actual temperature.

4.4.2.5 Total Flow Dependence

In non-selective deposition of silicon from DGS one can readily identify a temperature regime in which the growth process is controlled by reaction kinetics and one where growth rates are limited by mass transport to the reaction site. If growth rates are under kinetic control they are much less sensitive to mass transport phenomena such as total gas flow.

The effect of total flow on growth rates in selective epitaxial depositions was studied for three different flows, where the partial pressures of all gas species in the feed were held constant. The results, illustrated in Fig. 4.30, indicated that growth rates decreased with increasing total flow while growth uniformity did not change significantly. This was contrary to what would be expected from a mixed flow reactor where depositions are carried out under kinetic control. The reduction of gas residence time and conversion should lead to increased deposition rates with an increased total flow. Furthermore, if a boundary layer is present, its thickness would decrease with the reciprocal of the square root of free space gas velocity, enhancing mass transport to the
Figure 4.29: Selective epitaxy growth rate profile on patterned wafers at 860 °C, 920 °C, 950 °C, and 1000 °C

Figure 4.30: Effect of total gas flow on selective epitaxial growth rate profile
wafer surface and increasing reaction rates if reaction were mass transport controlled.

**Table 9**

<table>
<thead>
<tr>
<th>Flow Rate (slm)</th>
<th>$\frac{P_{HCl}}{P_{DCS}}$</th>
<th>$\frac{P_{HCl}^2}{P_{DCS}}$ (Torr)</th>
<th>Growth Rate (μm/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_2$</td>
<td>HCl</td>
<td>DCS</td>
<td>Inner</td>
</tr>
<tr>
<td>30</td>
<td>0.62</td>
<td>0.11</td>
<td>5.64</td>
</tr>
<tr>
<td>60</td>
<td>1.24</td>
<td>0.22</td>
<td>5.64</td>
</tr>
<tr>
<td>90</td>
<td>1.86</td>
<td>0.33</td>
<td>5.64</td>
</tr>
<tr>
<td>120</td>
<td>1.80</td>
<td>0.40</td>
<td>4.50</td>
</tr>
<tr>
<td>120</td>
<td>1.78</td>
<td>0.20</td>
<td>8.90</td>
</tr>
</tbody>
</table>

In the Purdue reactor this was not observed. It was concluded that the traditional concept of mass transfer limited and kinetically controlled growth regimes cannot be easily projected onto the system of highly reversible reactions taking place in SEG.

### 4.4.2.6 Area Dependence

Ishitani et al.⁸⁰ and Drowley ¹⁹⁰ had studied the influence of the ratio of exposed silicon to oxide covered area on a device scale and wafer scale. They discovered that growth rates were lower for areas where the fraction of exposed silicon was high, a term named "local loading". We have extended their studies to investigate loading effects also on a reactor scale, i.e. the effect of number of wafers on the susceptor on growth rates in a particular run.

The effect of oxide coverage over the entire susceptor area is shown in Fig. 4.31. It appears that growth rates increased as the exposed silicon and polysilicon area decreased. On a wafer scale level a similar effect was observed, as shown in Fig. 4.32. In addition, growth uniformity increased as oxide coverage on a wafer fell below about 40%. However, growth uniformity
Figure 4.31: Effect of masking oxide coverage on selective epitaxial growth rate profile

Figure 4.32: Effect of reactor loading on selective epitaxial growth rate profile
The existence of a temperature profile across the susceptor provided us with a unique opportunity to gain a better understanding of a possible mechanism for selective epitaxial growth of silicon from mixtures of DCS, HCl, and H₂ on wafers masked with SiO₂. Although there exist some plausible models for bulk non-selective epitaxial depositions from chlorosilane-hydrogen mixtures very little work has been done to extend or modify such models to account for the effects peculiar to selective growth. These effects are

1. The change in sensitivity towards changes in HCl concentration from the net growth regime to the net etching regime, as shown in Figs. 4.20, 4.30, and 4.33.

2. The difference in growth rate temperature sensitivity between bulk (unmasked) wafers and wafers largely covered with oxide, as illustrated in Fig. 4.31.

3. The effect of growth rate variations between small aperture seed windows and large aperture seed windows, as shown in Fig. 4.34.

4. The effect of enhanced or retarded edge growth, depending on feed gas composition.

Table 10

Effect of wafer oxide coverage on growth rate and uniformity

<table>
<thead>
<tr>
<th>SiO₂ Fraction</th>
<th>Growth Rate (μm/min)</th>
<th>% Non-Uniformity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>0.353</td>
<td>19.3</td>
</tr>
<tr>
<td>0.90</td>
<td>0.300</td>
<td>14.8</td>
</tr>
<tr>
<td>0.50</td>
<td>0.179</td>
<td>17.8</td>
</tr>
<tr>
<td>0.30</td>
<td>0.176</td>
<td>9.6</td>
</tr>
</tbody>
</table>

did not significantly increase with increasing amount of exposed susceptor silicon area.
Even the most elaborate growth models proposed for the SiH$_2$Cl$_2$-HCl-H$_2$ system for bulk depositions cannot account for all these effects. We conclude therefore that additional considerations have to be included in modeling selective epitaxial growth on masked wafers. For the temperature range of interest, i.e. below 1300 °K, it is assumed that reaction kinetics determine the rate limiting step of the growth process.

Bloem and Claassen in a series of experiments investigated bulk epitaxial depositions from silicon halides. $^{129,136,138,191}$ They concluded that in the absence of chlorine, growth proceeded via the formation of silicon nuclei from the adsorbed species Si$^*$ and SiH$_2^*$. Ad-species would be desorbed from the surface after a mean stay time $\tau$ which exponentially depended on the surface binding energy of the adatom or admolecule. During this time the ad-species would diffuse along the surface and randomly form nuclei with other admolecules or adatoms. These nuclei would decay unless they had reached a supercritical size after which their decay became energetically less favorable than their growth.

However, even small amounts of chlorine in the system would change this mechanism drastically. The prevailing ad-species would change to SiCl$_2$ and the amount of Si$^*$ was too small to nucleate to a supercritical size. Growth could now only commence at steps or kink sites on the silicon surface. At such sites the binding energy for Si-Si bonds was maximized and provided some of the energy to break up the Si-Cl bonds of the admolecule SiCl$_2$. Hence, the diffusion process of SiCl$_2^*$ towards kink sites was considered the rate limiting step. On foreign substrates such as amorphous SiO$_2$ or Si$_3$N$_4$, lattice steps are not available and it therefore is more difficult to pass the nucleation barrier. This is the main reason why selective epitaxial growth is possible at all.

The formation of SiCl$_2^*$ occurs rapidly with even small amounts of HCl in the growth ambient. However, there is still a sharp decrease in growth with increasing amounts of HCl beyond what would be expected by a simple superposition of etch rates and growth rates. This was explained by Claassen and Bloem through the desorption of SiCl$_2^*$ via reaction with HCl to SiHCl$_3$.

Bloems and Claassens model can explain most of the experimental evidence reported on the behaviour of bulk epitaxial depositions at reduced temperatures. To include selective epitaxial growth some extensions have to be made.

The crucial observation in SEG when compared to bulk epitaxy is that not any one point on the wafer has the exact same surrounding as another
Figure 4.33: Growth rate as a function of HCl/DCS ratio at 850°C and 30 Torr.

Figure 4.34: Effect of HCl flow on the thickness ratio between a 5μm and a 50 μm aperture on a patterned wafer.
one. There exist SiO$_2$ regions where essentially no growth takes place, where binding energies for adspecies are different, and where temperatures can be different. It is therefore hypothesized that SEG growth behaviour will approximate bulk growth rate behaviour only if oxide regions between adjacent seed openings are less than a few adspecies diffusion lengths apart. In all other cases, the difference in surface kinetics between oxide and silicon regions has to be considered to account for the experimentally observed growth rate behaviour.

We observed strong non-uniformities in growth rates for wafers mostly covered with oxide in the presence of a temperature gradient across the susceptor. When oxide coverage was less than about 40%, growth rates became more uniform. For bulk wafers, i.e. wafers with no oxide coverage, growth rates were quite uniform even in the presence of the same temperature gradient under identical deposition conditions. This suggests that it is not the surface reactions on silicon that are extremely sensitive to temperature variations but rather surface kinetics on the masking oxide. From our experimental results as well as that from other researchers the exact mechanism or step responsible for this sensitivity to temperature could not be identified. Possible candidates are the binding energy of SiCl$_2$ to SiO$_2$, the binding energy of H$_2$ to SiO$_2$, and the surface diffusion energy for SiCl$_2$ on SiO$_2$.

From depositions that were carried out under non-selective conditions on masked wafers we know of the existence of a nucleation-free zone around each seed window opening. This zone can be between 10$\mu$m and 90 $\mu$m wide. It can be conjected that ad-species in this zone diffuse towards the seed area while no ad-species diffuse from the seed area into the nucleation free zone, since nucleation energies of silicon on silicon are lower than that for silicon on silicon dioxide. Surface diffusion is thought of as a rather slow process, and while ad-species diffuse towards the seed window they can be desorbed by undergoing reactions with HCl. This is one of the most important steps in describing many of the effects that are not accounted for by traditional bulk epitaxy models.

Since surface diffusion of ad-species is assumed to contribute significantly to growth in the seed area, and while on statistical grounds it contributes to growth more at the edges of the seed area than towards its middle, the less desorption takes place the higher the growth rates would be at edges of structures or for structures with small apertures. When most of the ad-species on the masking oxide are desorbed, however, there will be relatively more ad-species in the middle of the seed area as compared to its edges, or more on
large area structures as compared to small area structures, since desorption from silicon is more difficult. As a result growth rates will be higher on large structures or in the middle of apertures in the masking oxide.

This conceptual model seems to be able to explain all of the effects in selective and non-selective silicon epitaxy reported in the literature and observed by us. While it does not go into the details of the chemical and physical processes present in growth ambients at silicon and silicon dioxide surfaces, it could serve as a starting point for a more formal treatment of the matter. As a model it can be succesful as long as it is not invalidated by experimental data. To be of more practical use to process engineers, a quantification of its different components would be desirable.

### 4.5 ELO Planarization

The planarization of the ELO silicon over the oxidized polysilicon gate for the fabrication of the stacked PMOS transistor was necessary because it was not possible to obtain horizontal to vertical growth rates much larger than one. This in turn leads to very high structures for the required widths. To facilitate the implantation of source and drain regions to form the stacked device the ELO must not be thicker than about 0.5 μm above the gate polysilicon oxide. It is therefore imperative to develop a procedure that can planarize the ELO material from a height of approximately 8 μm to 1 μm without significant loss in structure width.

The required amount of height reduction as well as the large size of the features to be planarized makes this processing step challenging. Initially, various isotropic and anisotropic wet etchants were tried but did not work because of poor selectivity between silicon and the masking material, mask material degradation, and horizontal etching, resulting in considerable width loss. Special planarizing materials such as spin-on glasses and polyimides that are routinely employed for smoothing severe topology failed to work here because of the desired tight control of the amount of planarization, which usually is a strong function of structure width. The objectives in the development work of a planarization procedure for ELO were to find a method by which the amount of planarization was not a function of structure size within the range of 3 to 20 μm, where standard processing techniques and equipment could be used, where good uniformity over a wafer and from wafer to wafer could be obtained, and were the non-uniformities stemming from the epitaxial deposition process would be eliminated or greatly reduced.
Figure 4.35: Planarization procedure by resist reflow and plasma etching
The first such procedure that was developed could account for all but the last requirement and is illustrated in Fig. 4.35. It will be briefly described here; for more extensive information the reader is referred to the work by Vedala. In this process, a thick coating of photoresist was applied to the wafer that had the ELO structures to be planarized on it. After a thermal reflow the resist was etched anisotropically in a plasma of oxygen, until all ELO silicon top surfaces were exposed while all other area was still covered with resist. Thereafter the wafers were subjected to a CF$_4$/O$_2$ plasma that etched silicon at a much faster rate than the resist. Once the desired ELO thickness was obtained, the remaining resist was removed and processing commenced.

Two main problems existed with this approach. The first problem came from the lack of an end point detection mechanism for the ELO plasma etch. It was thus imperative that very good repeatability from run to run could be obtained. This proved to be difficult. The second problem stemmed from the epitaxial non-uniformities and was principal in nature. Assuming a growth of 10 µm had to be planarized to 1 µm thickness, and growth varied by ±5 % after deposition, it would have been mapped to ±50 % after this planarization procedure. This mapping of growth rate non-uniformity was observed and resulted in a very low yield after planarization. In epitaxial depositions, growth rate non-uniformities of about ± 5 % are considered state of the art.

It became apparent that this scheme would not allow to continue work towards the stacked PMOS transistor. Therefore a different solution was sought after which is shown in Fig. 4.36. This approach followed closely a proposal from Kubota et al. at the IEDM 1987. Here a layer of nitride was deposited uniformly over the wafer to be planarized. Subsequently the wafer was polished chemo-mechanically, with a selectivity of polishing rates of silicon over nitride of about 60. Thus the nitride layer protected the ELO sidewalls and served effectively as a polishing stop. The uniformity of the ELO after polishing was theoretically determined by the uniformity of the CVD nitride layer which could be deposited with negligible thickness variations. However, even with this scheme there exist many practical problems. It required considerable expertise in polishing, polishing equipment, and a knowledge of the polishing agents to obtain the desired effect. For these reasons we did not try to develop this procedure in the Purdue Solid State laboratory but relied on the help of Harris Semiconductor Corporation.
After Epitaxy

After Plasma Nitride Deposition

After chemo-mechanical polishing

Figure 4.36: Planarization by preferential polishing
CHAPTER 5

ELECTRICAL EVALUATION

5.1 Introduction

In the last chapter the different fabrication steps necessary to build 3-D stacked devices were discussed. These methods included the formation of a high-dielectric strength polysilicon oxide, the definition of an operating range for the epitaxy reactor to maintain gate oxide integrity and grow low defect density material, and the exploration of the epitaxy reactor parameter space to achieve good uniformity growth across a wafer.

With the information available, three-dimensional stacked devices were fabricated. The main thrust of the research was directed towards the investigation of the thin interfacial layer between the ELO silicon and the amorphous polysilicon oxide film. These interfaces are of great importance to the properties of stacked MOS devices.

In the past SiO$_2$/Si interfaces have been extensively studied. However, they have been mostly formed by the growth or deposition of SiO$_2$ on Si, either by thermal oxidation of Si or chemical deposition methods rather than by deposition of crystalline Si over amorphous SiO$_2$, as done in this work. Only thermal oxidation of Si had given device quality insulating films, although recently good films by low temperature pyrolitic deposition were reported.$^{205}$

5.2 Interface States

Many of the major non-idealities in MOS transistors, such as 1/f noise, small gain, carrier trapping, and surface generated leakage currents are ascribed to surface or interface states.$^{206-208}$ These states are allowed energy levels in the band gap near the Si-SiO$_2$ interface.$^{209}$ The concentration of surface states is sensitive to details of the fabrication process; for conventional Si/SiO$_2$ interfaces by thermal oxidation it was found that one could minimize
the surface state density by a low temperature (<500 °C) anneal in H₂. One common model explains interface states as dangling bonds of incompletely oxidized Si at the growth interface. The hydrogen is believed to neutralize these bonds.

A good indication for high surface state densities is the distorted, spread out shape of the high frequency C-V curve of a MOS capacitor. However, to study the detailed distribution and to detect even low levels of surface states one has to employ methods that are sufficiently sensitive and accurate. Some of these methods are briefly reviewed in the following.

For extremely high interface trap densities the C-V plot will basically be flat since the high concentration of charge at the interface immobilizes ("pins") the Fermi level. The observed capacitance in such a case is often smaller than the accumulation capacitance $C_{ox}$. Fermi level pinning usually occurs for interface state densities close to and above $10^{13}$cm$^{-2}$eV$^{-1}$. At these levels, MOS transistors would exhibit gains close to zero.

At the other end of the spectrum, near ideal behaviour is observed for interface state densities of less than about $10^{10}$cm$^{-2}$eV$^{-1}$. Measurement techniques therefore usually try to cover this range.

Terman devised a method of computing $D_{it}$ from high frequency C-V data. For a given gate bias the difference in capacitance between an ideal, interface trap free capacitor, and the measured one is ascribed to surface states for states that can follow the ac signal:

$$C_{it} = qD_{it}(\psi_s)$$

(5.1)

The capacitance of the ideal structure can be theoretically calculated as a function of the surface potential $\psi_s$:

$$C_{HF} = C_s(\psi_s)\frac{C_{ox}}{C_s(\psi_s) + C_{ox}}$$

(5.2)

A given capacitance corresponds to the same amount of band bending. By setting equal the capacitance of the ideal and actual capacitors one can determine the surface potential for a measured capacitance for the real structure as a function of the applied gate bias. From basic principles the change in band bending caused by a small change in gate voltage can be written as

$$\frac{C_{ox}dV_G}{C_{ox} + C_{it}(\psi_s) + C_s(\psi_s)} = d\psi_s$$

(5.3)

From this equation and equation (5.1) $D_{it}$ is obtained as
This method involves numerical differentiation of measured data and is not very sensitive in detecting very low (<10^{10} \text{cm}^{-2} \text{eV}^{-1}) surface state concentrations. For the calculation of $C_s$ the actual doping profile of the measured capacitor needs to be known, making this evaluation very difficult in case of non-uniform dopant distribution throughout the wafer substrate.

Berglund\textsuperscript{212} devised a way of extracting surface state densities from low frequency C-V data. This method avoids the differentiation of data found necessary for the high frequency method but requires the computation of an ideal C-V low-frequency curve with an associated undetermined constant. The interface trap density is related to the measured capacitance by

$$qD_{it} = \left[ \frac{1}{C_{LF}} - \frac{1}{C_{ox}} \right]^{-1} - C_s$$

Equation (5.5)

$C_s(\psi_s)$ can again be calculated numerically, and can be related to $C_s(V_G)$ by

$$\psi_1 = \psi_{so} + \int_{V_{so}}^{V_G} dV_G \left[ 1 - \frac{C_{LF}(V_G)}{C_{ox}} \right]$$

Equation (5.6)

The choice of the integration constant $\psi_{so}$ directly influences $C_{it}$, making the accuracy of the procedure a function of the chosen value. This constant is related to details of the fabrication procedure resulting in fixed and mobile oxide charges and interface charges. This technique of determining interface state densities also requires knowledge of the dopant distribution throughout the space charge region, as was the case for the high-frequency method.

The MOS ac conductance technique first developed by Nicollian and Goetzberger\textsuperscript{213} is considered the most reliable and complete technique to measure surface state densities. We will not go into the details of the complex theory but rather sketch the procedures necessary for a characterization. The basic idea is that capacitor losses are directly related to interface trap phenomena. By measuring $\frac{G_P}{\omega}$ vs $\omega$ and $C(\omega)$ one can determine the essential parameters $D_{it}$, $\sigma_n$, and $\sigma_p$ by comparing measured data to the equation

$$\frac{G_P}{\omega} = \frac{1}{\sqrt{2\pi\sigma_s^2}} \int_{-\infty}^{\infty} \exp \left[ -\frac{(U_S-U_{so})^2}{2\sigma_s^2} \right] \frac{qD_{it}(U_S)}{2\omega \tau_m} \ln(1 + \omega^2 \tau_m^2) dU_S$$

Equation (5.7)
Computer programs have been written that automate the curve fitting and extraction of the parameters. This method involves a considerable amount of computation and is not suited for quick evaluations. For our purposes the conductance technique was not considered as suitable because of the difficulties in measuring conductance data with the required accuracy in the presence of high leakage oxides.

The capacitors for the interface evaluation of the novel polyoxide/ELO interface consisted of stripes, 450 μm long and 3 μm wide. This structure was necessary because of the limited amount of overgrowth that could be obtained. The design of a long narrow stripe effectively represents a distributed parameter transmission line and can only for small enough frequencies be approximated as a lumped circuit. Even then it has to be considered as an RC filter with a corner frequency above which considerable damping of the applied signal takes place. We observed experimentally that accumulation capacitance values became noticeably smaller once measurement frequencies exceeded about 1 MHz. The rather high leakage of the thin polysilicon gate oxide contributed a larger error to the conductance of the structure than to the susceptance. For these reasons the quasi-static method combined with high-frequency C-V was employed to obtain interface state densities for the 3-D stacked device.

5.3 Quasistatic \( D_{it} \) Measurement

The acquisition of true low-frequency C-V curves is difficult for semiconductors having long carrier lifetimes. For modern devices, measurement frequencies of less than 1 Hz would be required. Kuhn suggested instead to apply a slow, "quasi-static" voltage ramp to the gate of MOS devices and to measure the resulting displacement current to obtain the low frequency C-V characteristic. If the ramp rate is slow enough the device can maintain quasi-equilibrium.

To extract interface state densities versus bandgap energy by the low frequency method as developed by Berglund and the quasi-static technique as suggested by Kuhn, the measured data was to be compared with calculated data for a theoretical ideal capacitor. This leads to a problem with interface or oxide charges that lead to a shift of the C-V curve along the voltage axis. The shift depends on fabrication conditions and cannot be easily predicted. In
addition, a priori knowledge of the dopant profile throughout the space charge region is required.

Castagne and Vapaille\textsuperscript{215} proposed to compare the experimental low-frequency or quasi-static C-V curve with an experimental high-frequency curve of the same device to extract interface state data directly without the computation of an ideal C-V characteristic. The suggested method made use of the frequency response of interface states. At high enough frequencies interface trap levels cannot be emptied or occupied fast enough to follow the applied small signal and thus contribute to a change in capacitance.

5.3.1 Data Acquisition

An experimental set-up for interface state density extraction by the high-frequency/low-frequency or quasistatic method had been developed previously in our laboratory by J. Shields.\textsuperscript{216} In this set-up, the high-frequency measurements were made with an HP-4274A LCR Bridge at 100kHz, and the quasi-static data was acquired with an HP-4140B Pico-amperemeter. The measurement instruments were controlled by a desktop computer that also collected and plotted the data. The data could then be transferred to a miniframe computer for further analysis.

In the quasi-static method, the capacitance is evaluated from the displacement current caused by the change in voltage over the capacitor:

\[
C = \frac{I}{\frac{dV}{dt}} = \frac{I}{R}
\]  

(5.8)

where R is the gate voltage ramp rate. If there are leakage currents through the dielectric that are similar in magnitude or even larger than the displacement current, erroneous capacitance values are obtained by this technique. Since we needed to measure capacitors whose dielectrics were oxidized polysilicon, which exhibits leakage currents much larger than those of oxide thermally grown on silicon, this problem was of serious concern.

A modified quasi-static measurement technique that can obtain accurate capacitance measurements even in the presence of leakage currents was developed by Mego\textsuperscript{217} and subsequently implemented into the Keithley 595 Quasistatic C-V Meter, a commercially available instrument. In this instrument, the device under test is made part of the feedback loop of an integrator. A small voltage step is applied and the resulting integrator output voltage is measured three different times. The change in output voltage is related to the
Figure 5.1: Quasi-static C-V measurement instrumentation

\[ \Delta Q = C_X \Delta V; \Delta V_{OUT} = \frac{-\Delta Q}{C_F} \]

Figure 5.2: Quasi-static C-V measurement timing diagram

\[ C_X = \frac{Q_2 - Q_1}{t_{STEP}} \quad \frac{Q_2 - Q_2}{t_0} \]

\[ C_X^{(CORRECTED)} = C_X - \frac{(Q/t)(t_{DELAY \ TIME} + t_1)}{t_{STEP}} \]
Figure 5.3: HF-LF C-V measurement instrumentation set-up.
current through the device under test by

\[ I_x = \frac{C_l \Delta V_{out}}{\Delta t} = \frac{-\Delta Q}{\Delta t} \]  (5.9)

The current \( I_x \) may be composed of the displacement current through the capacitance to be evaluated plus some leakage current. To discern the two, three successive measurements of the charge or integrator output voltage are made as shown in Fig. 5.2. The slope in the curve is due to leakage current integration. If no leakage was present, the capacitance would be given by

\[ C_x = \frac{Q_1 - Q_2}{V_{step}} \]  (5.10)

The leakage current is determined by

\[ I_{\text{leak}} = \frac{Q}{t} = \frac{Q_2 - Q_3}{t_0} \]  (5.11)

Thus, the corrected capacitance can be computed by

\[ C_{x(\text{corrected})} = C_x - \left( \frac{Q}{t} \right) \left( \frac{t_{\text{delay}} + t_1}{V_{\text{step}}} \right) \]  (5.12)

The Keithley 595 was substituted for the HP 4110 Pico-amperemeter in the original experimental set-up and operated under control of the desktop computer. The software for data acquisition was modified to accommodate the quasi-static CV-meter. The C-V measurement instrumentation set-up is illustrated in Fig. 5.3.

### 5.3.2 Data Analysis

A simple electrical equivalent for an MOS structure is shown in Fig. 5.4. At high enough frequencies, only \( C_0 \) and \( C_s \) will contribute to the measured capacitance. \( C_0 \) is basically independent of bias, while \( C_s \) changes with the applied gate voltage. A perturbation of the ideal gate voltage, shown as \( \Delta V_{it} \), takes place due to interface traps for the high frequency limit as well as for the low frequency limit. At low frequencies, interface traps start to follow the applied signal and contribute a capacitance \( C_{it} \). To obtain an accurate picture of interface state densities, the high frequency measurement has to be done at frequencies where most of the surface states cannot follow the signal any more, yielding zero \( C_{it} \). On the other hand, the low-frequency or quasistatic measurement has to be slow enough that even the slowest interface traps can be
occupied fast enough to follow the applied gate bias. In practice, both limits cannot be reached, i.e. there will always be some very fast and very slow states that will not be detected by this method.

To simplify the analysis procedure the following assumptions were made:

1. A true high-frequency C-V characteristic is obtained at 100 kHz.
2. Oxide thickness and average dopant concentration in the semiconductor can be obtained experimentally with good accuracy.
3. Leakage currents are small enough to not significantly disturb the high-frequency C-V characteristic. For the quasi-static measurement, leakage is compensated for.
4. The voltage sources for the LCR bridge and the quasi-static C-V meter are synchronized.

Then, for the high-frequency limit one can write

$$C_{HF} = C_0 \frac{C_s}{C_0 + C_s}$$

and from this equation

$$C_s = C_0 \frac{C_{HF}}{C_0 - C_{HF}}$$

In the low-frequency limit, one can write
These equations can be solved for $C_{it}$, yielding

$$C_{it} = C_o \left[ \frac{C_{LF}}{C_o - C_{LF}} - \frac{C_{HF}}{C_o - C_{HF}} \right]$$

The interface state density per area and eV is then simply obtained by dividing $C_{it}$ by the capacitor area and unit charge:

$$D_{it} = \frac{C_o}{qA_G \cdot eV} \left[ \frac{C_{LF}}{C_o - C_{LF}} - \frac{C_{HF}}{C_o - C_{HF}} \right]$$

To correlate the interface state density with the band gap energy at which it occurs, a relationship between the high-frequency capacitance, the gate voltage, and the surface potential has to be established. For that purpose one can write for a depleted n-type semiconductor

$$C_s = \sqrt{\frac{q\varepsilon_s N_D}{2} \frac{q}{kT} \frac{e^{U_s - 1}}{\sqrt{e^{U_s - U_s} - 1}}}$$

A few $\frac{kT}{q}$ or more into depletion, with $U_s > 2U_F$ one can write (5.18) as

$$C_s \approx \sqrt{\frac{q\varepsilon_s N_D}{2} \frac{q}{kT} \frac{-1}{\sqrt{-U_s - 1}}}$$

Rearranging relationship (5.14) and squaring both sides one obtains

$$\left( \frac{C_o}{C_{HF} - 1} \right)^2 = \left( \frac{C_o}{C_s} \right)^2 = \frac{C_0^2 2kT}{q^2 \varepsilon_s N_D} \left( -U_s - 1 \right)$$

$$= \frac{-2C_0^2 kT}{q^2 \varepsilon_s N_D} \frac{1}{q} (1 + U_s)$$

The surface potential can be written as $\psi_s = \frac{kT}{q} U_s$, and equation 5.20 can be solved for $\psi_s$ as

$$\psi_s = \frac{-q\varepsilon_s N_D}{2} \left( \frac{C_o}{C_{HF} - 1} \right)^2 + \frac{kT}{q}$$

From this equation and the definition of $U_s$ the desired bandgap energy is
obtained as

\[ E_T = \psi_s - kTU_F + \frac{E_G}{2} \]  

Equation 5.18 is only valid in the depletion region and therefore the analysis procedure as described above can obtain useful results only over the respective gate bias range. As will be discussed further down, the highest accuracy using this method is obtained for states near midgap.

The electrical equivalent circuit for the 3-D stacked capacitor shown in Fig. 3.5 is composed of the capacitance under investigation, \( C_{ELO} \), and the parasitic capacitances \( C_{Pad} \) and \( C_{Substrate} \). Experimentally, only the total capacitance \( C_{total} \) could be measured. To obtain \( C_{ELO} \) the parasitic capacitances were subtracted from the measured data after the completion of the measurements. For that purpose, another similar capacitor structure but without any overgrowth was placed right next to the 3-D capacitor on the same die. The separation between these two structures was about 200 \( \mu \)m. The C-V for this non-ELO capacitor was taken first and recorded. Thereafter the C-V for the 3-D structure was taken. From geometrical dimensions the capacitance per unit area for the non-ELO device was computed and multiplied with the area of the ELO device facing the wafer substrate. The resulting "bottom" capacitance, i.e. \( C_{Pad} + C_{Substrate} \) was then subtracted from the total ELO capacitance to obtain \( C_{ELO} \). This procedure was repeated for each gate bias at which data was taken.

If the parasitic capacitances and \( C_{ELO} \) had similar values and threshold voltages, large errors could be introduced into the entire analysis. However, as was discussed in Chapter 3, the threshold voltages of the parasitic devices were fixed at large negative values such that these capacitances stayed in accumulation under all gate bias voltages of interest. Thus, a constant, basically frequency independent capacitance was measured for these structures. Also, the bottom field oxide was about four times as thick as the polyoxide. For similar areas, the accumulation capacitance of \( C_{ELO} \) was therefore about four times as large as the combined substrate capacitances.

The long narrow stripes had potential to introduce considerable error into the analysis procedure due to overetching, as shown in Fig. 5.5, thereby having the actual capacitor area deviate from its drawn dimensions. For that reason in many instances scanning electron micrographs were used to determine the actual width of the fingers that composed the 3-D capacitor. The thickness of the polyoxide could not be measured by ellipsometry but was acquired by
Figure 5.5: Overetched polysilicon finger structure

Figure 5.6: Total high-frequency capacitance exhibiting lateral effects
measuring a large area metal/polyoxide/polysilicon capacitor provided on each die for that purpose.

One of the major advantages of the high-low frequency method is that no a priori knowledge of the actual dopant profile throughout the semiconductor bulk has to exist. Furthermore, no theoretical calculations are necessary to extract interface state data. There are two major sources for error in this analysis. The first stems from the subtraction of two nearly equal quantities, and the other from the error in obtaining true quasi-static and high-frequency C-V characteristics.

The round-off errors in the interface state contribution to the capacitance, $C_{it}$, resulting from the subtraction of $C_{HF}$ and $C_{LF}$ can be derived as:

$$\frac{dC_{it}}{C_{it}} \approx \left( \frac{C_{HF}/C_{LF}}{1 - C_{HF}/C_{LF}} \right) \frac{dC_{HF}}{C_{HF}} + \left( \frac{1}{1 - C_{HF}/C_{LF}} \right) \frac{dC_{LF}}{C_{LF}}$$

(5.23)

As $C_{HF}$ and $C_{LF}$ approach each other, as is the case when the capacitor is biased close to or into the accumulation region, the relative error becomes very large. Also, for low interface state densities $C_{HF}$ and $C_{LF}$ will be very similar in the depletion region, thereby imposing a lower bound on the minimum detectable trap density.

Additional errors are introduced due to the use of a 100 kHz C-V curve for the high-frequency characteristic as some interface traps can still follow the signal at this frequency. For the quasi-static characteristics, errors are introduced due to the low current levels employed in the measurement. The actual capacitance measurement accuracy is reduced by the necessary subtraction of the parasitic bottom capacitances from the total capacitance. The Keithley capacitance meter has an accuracy of $\pm 1\%$ of the reading $+ 0.1\mu F$ in the 200 $\mu F$ range, excluding noise. For metallized wafers, the parasitic capacitances amounted to about 50% of the capacitance of interest, $C_{ELO}$. As a result, the total accuracy of measuring $C_{ELO}$ was less than $\pm 2.5\%$, noise related errors excluded. At the current levels employed, noise did however contribute noticeably to measurement fluctuations.

The relation between a certain interface state density and energy level in the band gap is obtained by comparison of the measured high-frequency capacitance with the theoretically computed one. Since for the theoretical computation idealized conditions such as no interface states and constant doping level throughout the semiconductor are assumed, attributing interface states to a certain energy level is dependent on how exact the actual structure
resembles the ideal one. In case of grossly varying dopant profiles this attribution will contain large errors.

In summary, $D_{it}$ can be measured most accurately for energies close to midgap, since in this region the difference between the high-frequency capacitance and the quasi-static capacitance are maximized. Furthermore, low doping in the semiconductor material minimizes errors due to the failure to obtain true low- and high-frequency characteristics. For a doping of about $10^{16}\text{cm}^{-3}$, the minimum detectable surface state density is in the range of $1-2\times10^{11}\text{cm}^{-2}\text{eV}^{-1}$.

5.4 SOI Interface Characteristics

To our knowledge, there has been no report in the standard literature on the electronic characteristics of a structure where the oxide is not thermally grown on silicon, but rather silicon is epitaxially grown over the oxide. Such interfaces also exist in advanced bipolar and proposed BiCMOS technologies that employ selective epitaxial growth for devices. Employing the measurement techniques described in the previous chapter, and drawing from the results of the basic fabrication process development work, the interface properties of the novel polysilicon/polyoxide/ELO structure were studied to extent the design work towards the fabrication of a 3-D stacked PMOS transistor as described before.

The substrates used for fabrication of the devices were (100) oriented 100-Ω-cm n-type 75 mm wafers. After initial cleaning, phosphorus was implanted to adjust the threshold voltage of the field oxide to about -40V in all but the first batch of wafers. The high-frequency C-V plot of $C_{\text{total}}$ and $C_{\text{ELO}}$ for devices from that first run are shown in Fig. 5.6. The threshold voltages for the top as well as the bottom capacitance were about -3 Volts. One can observe a dip in the high-frequency C-V curve around -5 Volts. The capacitances at gate biases smaller than -5 Volts were strong functions of the measurement frequency. It was concluded that the small width of the gate electrode fingers caused lateral ac current flow for the bottom capacitance $C_{\text{Substrate}}$ when it was biased into inversion. It is apparent that no useful interface state data can be extracted from such a structure. For all devices processed subsequently the large negative threshold voltage effectively prevented the occurrence of lateral effects.

Field oxide was grown to a thickness of 400 nm at 1100°C in steam in a conventional oxidation furnace. Polysilicon was deposited in an LPCVD
reactor at 580 °C to a thickness of about 300 nm. An implant dose of $10^{16}$ cm$^{-2}$ of phosphorus at an energy of 125 keV, in conjunction with the small grain size of the low temperature polysilicon, helped to maximize breakdown field strength of the thin (100 nm) polyoxide, grown subsequently at 1100 °C in oxygen. Between the implantation and the oxidation, the polysilicon was patterned by conventional lithography and dry etching such that gates were aligned along the <100> direction to minimize faceting in the subsequent epitaxial growth.

Seed windows were cut into the field oxide and the epitaxial lateral overgrowth was carried out.

![Figure 5.7: I-V curve for ELO capacitor with epitaxy grown at 900 °C](image)

The growth procedure began with a 5 min. prebake at a pressure of 150 Torr and a temperature of 930 °C, to remove native oxide from the seed area. Then the temperature was ramped to the growth temperature of 860 °C, 900 °C, or 950 °C, and the ELO was initiated. Dichlorosilane was employed as the silicon
source gas, HCl was introduced to prevent nucleation of polysilicon on the masking field and polysilicon gate oxide, and H₂ was used as the carrier gas. Growth was stopped when the polysilicon gates were completely overgrown with epitaxial silicon. Growth rates were typically in the order of 0.1 μm/min. The resulting ELO thickness was between 6 to 8 μm.

After the epitaxy the wafers received a blanket implant to form highly doped contact regions to the ELO, followed by a 10 minute wet oxidation at 920 °C to activate the implanted phosphorus. Contact window areas were patterned and wet etched. Metallization (Al-Si) was carried out on some wafers, while others remained without a metal layer for further processing. A detailed description of the processing sequence for the 3-D capacitor fabrication can be found in appendix A.

Before the time-consuming task of taking high-frequency and quasi-static C-V characteristics was begun, a yield map for each wafer, indicating all 3-D stacked capacitor devices that exhibited small leakage currents under 15 Volts bias stressing, was drawn. Wafers that were exposed to epitaxy runs at 950 °C showed no good devices (see also Fig. 4.11). Typical dielectric breakdown fields for good devices were 4 MV/cm, as shown in Fig. 5.7 for a device with 100 nm oxide thickness. Good thermal oxides grown on single crystalline substrates typically break down under fields in excess of 7 MV/cm.

High frequency C-V data was taken at 100 kHz and in some instances at 1 MHz between -8V and 8V for low-leakage devices, employing the measurement techniques described above. Low frequency data was acquired by the quasi-static method discussed previously. The total capacitance in accumulation was typically 18 pF.

A C-V characteristic for epitaxy deposited at 860 °C is shown in Fig. 5.8. The contribution of the fixed bottom polysilicon to substrate capacitance has already been subtracted. The equivalent curves for epitaxy done at 900 °C are illustrated in Fig. 5.9. For the wafers where epitaxy was deposited at 860 °C an average doping density at the polyoxide/ELO interface, and over the width of the depletion region, of \( N_D = 1.7 \times 10^{16} \text{cm}^{-3} \) resulted from C-V analysis. For the wafers where epitaxy was deposited at 900 °C a doping density at the polyoxide/ELO interface of \( N_D = 2.8 \times 10^{16} \text{cm}^{-3} \) was extracted. These impurities are mostly due to outdiffusion of phosphorus from the heavily doped polysilicon gate during and after the epitaxy, as was also indicated by the process simulations discussed in chapter 3. The threshold voltages were -2.1 Volt and -3.5 Volt for the 860 °C and 900 °C epitaxy wafers,
Figure 5.8: Capacitance $C_{ELO}$ for epitaxy grown at 860 °C

Figure 5.9: Capacitance $C_{ELO}$ for epitaxy grown at 900 °C
respectively. Apparent fixed charge densities \( (Q_f) \) were \(-1.7 \times 10^{11}\) cm\(^{-2}\) for the 860 °C wafers and \(4.0 \times 10^{10}\) cm\(^{-2}\) for the 900 °C wafers. Bulk control MOS (Al-Si) capacitors subjected to equivalent processing conditions exhibited apparent fixed charge densities of \(-5 \times 10^{11}\) cm\(^{-2}\). A summary of various parameters for the differently processed capacitors is shown in Table 11.

**Table 11**

Parameters of capacitors fabricated under different processing conditions.

<table>
<thead>
<tr>
<th>Param.</th>
<th>860 °C</th>
<th>900 °C</th>
<th>860 °C ann.</th>
<th>900 °C ann.</th>
<th>Al-Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Q_f) [cm(^{-2})]</td>
<td>(-1.7 \times 10^{11} )</td>
<td>(3.7 \times 10^{10})</td>
<td>(4.7 \times 10^{11})</td>
<td>(4.0 \times 10^{12})</td>
<td>(-5.0 \times 10^{12})</td>
</tr>
<tr>
<td>(V_{th}) [V]</td>
<td>(-2.09)</td>
<td>(-3.6)</td>
<td>(-5.86)</td>
<td>(-5.31)</td>
<td>(-0.718)</td>
</tr>
<tr>
<td>(V_{FB}) [V]</td>
<td>(0.63)</td>
<td>(-0.20)</td>
<td>(-2.55)</td>
<td>(-2.15)</td>
<td>(-0.15)</td>
</tr>
<tr>
<td>(C_{FB}) [pF]</td>
<td>(10.1)</td>
<td>(12.8)</td>
<td>(10.1)</td>
<td>(12.7)</td>
<td>(5.0)</td>
</tr>
<tr>
<td>(N_d) [cm(^{-3})]</td>
<td>(1.7 \times 10^{16})</td>
<td>(2.9 \times 10^{16})</td>
<td>(2.6 \times 10^{16})</td>
<td>(2.4 \times 10^{15})</td>
<td>(6.9 \times 10^{13})</td>
</tr>
<tr>
<td>(U_p)</td>
<td>(-14.49)</td>
<td>(-15.03)</td>
<td>(-14.94)</td>
<td>(-14.85)</td>
<td>(-9.0)</td>
</tr>
</tbody>
</table>

Interface state densities were extracted from the HF-LF C-V data, and are plotted in Fig. 5.10. The minimum interface state density of \(D_{it} = 2 \times 10^{11}\) cm\(^{-2}\) eV\(^{-1}\) is obtained near midgap, and is basically independent of the epitaxial deposition temperature. However, the epitaxy grown at 900 °C appears to lead to a flatter interface state density distribution over the band gap. It should be noted that data close to the band edges need to be treated with caution due to possible large errors in these regions. Furthermore, the correlation between interface states and energy at which they appear is most likely inaccurate, since for the positioning of the states in the energy gap a theoretical C-V curve with appropriate average and uniform doping was employed. As will be shown shortly, this assumption was not valid. Interface state densities below about \(10^{11}\) cm\(^{-2}\) eV\(^{-1}\) need to be treated with caution since they would be below the accuracy of the measurement procedure. For the interface trap density extraction it was assumed that the majority of the interface states had associated time constants much larger than those required to contribute to the capacitance at 100 kHz or 1 MHz.
Figure 5.10: Interface state density $D_{it}$ as a function of band gap energy for epitaxy done at 860 °C and 900 °C.

Figure 5.11: Capacitance $C_{ELO}$ for epitaxy grown at 860 °C, after 10 min. anneal in H$_2$ at 1000 °C.
Although the exhibited interface characteristics were much better than expected, and certainly sufficient to fabricate MOS transistors, we investigated the effect of a high-temperature post-epitaxy anneal on surface state densities. For that purpose, wafers were subjected to a 10 minute bake in pure H₂ at 1000 °C just before metallization. Respective C-V characteristics and Dᵣ data is plotted in Figs. 5.11, 5.12, and 5.13. A large negative shift in threshold voltage was observed. The negative shift may have been due to dopant outdiffusion out of the heavily doped polysilicon gate. Temperature bias stressing of the devices did not reveal any measurable shift of the C-V curve along the voltage axis. A remarkable decrease in the number of interface states is apparent for the weak depletion region. Here interface state densities are below the detection limit of about 10¹¹ cm⁻² eV⁻¹. The number of interface states in the weak inversion region however did not drop significantly. MOS transistors are usually operated under such biasing conditions.

The epitaxial overgrowth of oxidized polysilicon with monocrystalline silicon at temperatures around 900 °C leads to an interface that exhibits relatively good electronic characteristics. Normal Si/SiO₂ interfaces created by thermal oxidation of silicon, where interface state densities directly after oxidation increase with decreasing temperature typically have surface states in the low 10¹² cm⁻² eV⁻¹ range for dry oxidation around 900 °C. The as grown new ELO-device compares very favorably having 2x10¹¹ cm⁻² eV⁻¹. In the Purdue Laboratory, Al-Si MOS capacitors on (100) Si wafers exhibit about 10¹⁰ cm⁻² eV⁻¹ surface states after a postmetallization anneal.

For the 3-D structure it could be argued that trap densities should be quite high since the interface between the polyoxide and the ELO can consist of many crystallographically differently oriented planes due to asperities in the polyoxide. It has to be considered however that during the epitaxial growth, when bonding between the silicon and the oxide takes place, a vast supply of H₂ is available from the growth ambient. Chernov and Rusaién had analyzed adsorption layers in the Si-H-Cl system and found that under certain conditions equilibrium coverage of the silicon surface by H and Cl was nearly complete, leaving only few vacant adsorption sites. This could imply that sufficient H₂ is trapped during the overgrowth process at the ELO-Si/SiO₂ interface to neutralize interface traps even without subsequent annealing in appropriate ambients. Furthermore, the Cl could help in gettering impurities if it was present in a chemically active form at the low temperature employed during the epitaxy.
Figure 5.12: Capacitance $C_{ELO}$ for epitaxy grown at 900 °C after 10 min. anneal in H$_2$ at 1000 °C

Figure 5.13: Interface state density $D_{it}$ as a function of band gap energy for epitaxy done at 860 °C and 900 °C, after 10 min. anneal in H$_2$ at 1000 °C
The similar values of surface state densities at midgap for epitaxy runs done at 860 °C and 900 °C could have resulted from the equivalent postepitaxial processing, since during this fabrication stage wafers would have been annealed at equivalent temperatures. Surface state densities directly after the epitaxial deposition were not investigated due to the difficulties in making contact to the lightly doped (autodoped) ELO material.

Employing experimentally obtained average dopant levels and oxide thickness, high-frequency and low-frequency capacitances were computed theoretically and compared with measured data in Figs. 5.14 to 5.16. The theoretical computation was done using exact charge analysis and assumed a uniformly distributed dopant profile away from the interface as well as a uniform interface state distribution across the bandgap of $10^9$ cm$^{-2}$eV$^{-1}$. The underlying theory and a suitable program to carry out the analysis have been described elsewhere. The experimental C-V curves are quite different from the calculated theoretical form. It appears that at the onset of depletion, changes in capacitance with changing gate bias are much smaller for the experimental data than for the computed one. This can be caused by a highly non-uniform dopant profile close to the interface, as was also indicated by process simulations discussed in chapter 3, or by a high concentration of interface states, or both. At small gate potential changes from accumulation biasing towards depletion biasing, the large number of ionized donors near the interface requires only a small change in depletion layer width to balance the charge on the gate. As gate charge increases, the depletion layer extends into less and less doped material some distance away from the surface. At that point the change in depletion width becomes more sensitive to changes in gate charge and the C-V curve becomes steeper than the theoretical curve where an average dopant concentration was assumed. The low-frequency curve reveals that a significant number of interface states exist, since the minimum measured low-frequency capacitance is considerably higher than the minimum computed one.

5.5 3-D PMOS Transistors

The investigation on the polyoxide/ELO interface had revealed that reasonably good characteristics could be obtained and that modulation of the channel region between accumulation and inversion was possible. This was a prerequisite for building a conductance-modulated device.
Figure 5.14: Comparison of experimental and ideal LF C-V curve for epitaxy grown at 860 °C

Figure 5.15: Comparison of experimental and ideal HF C-V curve for epitaxy grown at 860 °C
Figure 5.16: Comparison of experimental and ideal LF and HF C-V curves for epitaxy grown at 900°C.

Figure 5.17: Scanning electron beam micrograph of a 3-D PMOS stacked transistor with 1.4 μm thick overgrowth.
It has been mentioned previously that SOI transistors can overcome some of the inherent limitations of bulk devices such as lateral isolation and latch up, short-channel effects, parasitic capacitances, and radiation tolerances. In the case of very thin SOI films, additional improvement is possible by volume inversion. In this case, the theoretical bulk depletion width is larger than the SOI Si thickness. Under appropriate biasing conditions, this leads to the inversion of the entire SOI layer. Thus, minority carriers, which are responsible for current transport in MOS transistors, are no longer confined to a narrow surface region, greatly enhancing device performance. However, for 3-D CMOS the SOI layer thickness would have to be less than about 0.3 μm which is difficult to achieve.

Another set of wafers was prepared, and processed equivalently to those employed in the fabrication of stacked capacitors up to the epitaxy. Thereafter a plasma nitride layer of 0.5 μm thickness was deposited and the wafers were send away for chemo-mechanical polishing planarization. This planarization resulted in ELO between 0 and 4 μm thick. The exact thickness above the polysilicon gate was difficult to evaluate since it was not known how much nitride was removed during the polishing. On two of the wafers, ELO thickness above nitride varied between 0.2 and 1.8 μm, and these wafers were used for further fabrication towards a stacked PMOS transistor.

A blanket phosphorus implant was carried out to set the ELO surface threshold voltage to a large negative value, thus preventing inversion and associated short circuits between the PMOS source and drain regions. This implant was not masked because the required negative resist did not allow for the resolution of the 3 μm wide features. Process simulation had however indicated that masking was not necessary and the n-doped layer could be counterdoped during the p+ source and drain implants. A plasma oxide of 0.25 μm thickness was deposited and a lithography with subsequent wet etching was carried out to define openings for the source and drain implanted regions. For the formation of the source and drain, a double implant was designed such that the predicted junction was 1.25 μm deep away from the ELO top surface and a doping concentration of about 10^20 cm^{-3} at the surface resulted after process completion to allow for low resistance contacts. The implants were annealed and driven in an oxidation furnace for about 20 min. at 920 °C. Thereafter lithography and wet etching was done for the contact regions. Since the polysilicon was still covered with some nitride from the planarization procedure, a nitride etch in hot boiling phosphoric acid was carried out, followed by another contact lithography and wet etch to remove the
remaining polyoxide in the polysilicon contact regions. To allow for eventual further high temperature processing, electrical testing commenced on a microprobe station without metallizing the wafers.

Fig. 5.17 shows an SEM microphotograph of a typical device. The thickness of the ELO above the polysilicon gate, which appears blurry in the foreground at the bottom of the picture is in excess of 1 \( \mu \text{m} \). The polysilicon is covered with about 0.3 \( \mu \text{m} \) of plasma oxide, and there is a 0.1 \( \mu \text{m} \) layer of plasma oxide on top of the ELO region that was not opened for contact and source/drain implant. The narrow line in the middle of the picture is the plasma oxide and nitride that masked the source and drain implants. This line was drawn as 3 \( \mu \text{m} \) and exhibits signs of severe overetching. The ELO surface appears rough in some areas from a "descum" plasma etch.

The typical electrical characteristics of this device are depicted in Fig. 5.18. This I-V plot indicates a short between the drain and source regions, most likely caused by lateral diffusion of the source and drain implants underneath the 0.8 \( \mu \text{m} \) masked channel region. For most such devices, no modulation of the source-drain current could be detected over a wide range of gate bias, i.e. from +15 V to -15 V. This doesn’t come as a complete surprise since the ELO thickness is in excess of the maximum allowable value of 1 \( \mu \text{m} \). A significant misalignment of more than about one \( \mu \text{m} \) between the source/drain regions and the polysilicon gate, even though not observed for this particular device, could have had the same effect. This misalignment was definitely observed on every wafer in some regions, due to uncontrollable thermal expansion of masks and wafers between successive masking steps. A second reason for the observed behaviour could have been that the low resistance short circuit masked out any high-resistance modulation of the conducting channel. The short circuit for the 20/5 transistors had values of a few hundred ohms while typical values for PMOS transistors are several thousand ohms at small drain voltages.

Since it appeared that source and drain were short-circuited at the ELO top surface, a silicon plasma etch was carried out to form a mesa above the polysilicon gate area. The ELO right above the channel region was thus about 0.3 \( \mu \text{m} \) higher than anywhere else. This procedure yielded some working devices, where no source-drain short circuits were observed and the channel current could be modulated by the buried polysilicon gate. An SEM micrograph of a working device is shown in Fig. 5.19. Here the seed hole was located at the right side of the polysilicon line which can be seen in the middle at the bottom of the picture. The ELO Si extended from the right side of the
Figure 5.18: I-V characteristics of bad 3-D stacked PMOS transistor with source-drain short circuit

Figure 5.19: Scanning electron beam micrograph of a 3-D PMOS stacked transistor with 0.8 μm thick overgrowth
polysilicon gate forming the drain, to the left side where the source region was located. The most important observation is that the source region was connected to the channel region only on small left-over strings of silicon. Furthermore, the thickness of the ELO above the polysilicon gate was less than 0.5 \( \mu m \). The channel masked region was misaligned to the right with respect to the gate by about 0.3 \( \mu m \).

I-V curves for this device are shown in Figs. 5.20 and 5.21. It is apparent that the current was foremost limited by the small effective channel width at the source. This prevented the transistor drain current from reaching saturation and also lead to very small currents for the transistor dimensions of \( 10 \mu m/3 \mu m \). The small effective channel length of less than 1 \( \mu m \) may have contributed to the observed I-V characteristics. Even though the I-V characteristics did not resemble that of a well behaved PMOS transistor, the drain current was modulated over 6 decades by the gate voltage.

Threshold voltages were a strong function of the applied drain voltage. Typical values were between 0 and 1.2 Volt. This could have been caused by the source and drain depletion regions which extended into a major portion of the ELO-SOI bulk material. The diode formed between the PMOS drain and the n-type substrate exhibited good characteristics but with a small reverse breakdown voltage, as shown in Fig. 5.22. Breakdown occurred at about 12 V, which indicated that the junction was in the ELO material which exhibited some defects.

An attempt was made to simulate the observed transistor behaviour with the SPICE circuit simulator. The experimental data could be approximated by a short channel transistor with large threshold voltage sensitivity on gate-bulk and gate-source bias, and large source and drain contact resistances as shown in Fig. 5.23. The simulation input file can be found in appendix D.
Figure 5.20: \( I_D \) vs \( V_{DS} \) characteristics of field effect modulated 3-D stacked PMOS transistor

Figure 5.21: \( I_D \) vs \( V_G \) characteristics of field effect modulated 3-D stacked PMOS transistor
Figure 5.22: I-V characteristics of PMOS-drain/substrate diode

Figure 5.23: SPICE simulation output of extreme short channel MOSFET
CHAPTER 6
SUMMARY AND PROPOSED FUTURE RESEARCH

The purpose of this research was to develop a suitable processing sequence for the fabrication of three-dimensionally integrated devices. The development of even a conventional 3-\mu m CMOS technology from scratch is a monumental undertaking where practical problems can be unsurmountable if no previous processing know-how is available. For that reason scaling of existing processes is preferred in industry over the design of a completely new process which is time consuming, risky, and extremely costly.

In this work, a completely novel approach to SOI and 3-D integration was taken. While the ultimate goal of this research project, namely the fabrication of well behaved 3-D stacked PMOS transistors was not reached, transistor action was demonstrated. It can be assumed that process optimization would lead to better device performance. Hence, 3-D integration by epitaxial lateral overgrowth of monocrystalline silicon may be a viable alternative to other SOI technologies.

The key processing step in this new technology was selective epitaxial overgrowth. It was discovered that typical epitaxial growth ambients degrade the thin insulating and masking oxides present on wafers subjected to them. This fundamental problem of SiO₂ disproportionation in the presence of low partial pressures of water vapor, oxygen, and silicon could be explained by the results of very recent related research on the effects of oxygen-deficient annealing ambients on thin gate oxides. Once understood, precautions were taken to prevent this degradation. As a result, lower epitaxial deposition temperatures were employed for selective silicon growth. This new understanding does not only apply to the case of 3-D ELO but to any Si/SiO₂ interface subjected to such conditions. It is particularly important for device isolation concepts making use of SEG.

Low temperature epitaxy was thus required to maintain the integrity of the masking and the insulating oxide. However, this raised the problem of crystalline quality of the epitaxially grown silicon. It was commonly known
that defect densities in epitaxial material increased rapidly as the epitaxial deposition temperature was lowered. Even though it had been demonstrated that residual oxidants in the growth ambient were partially responsible for this phenomenon, the exact relationship between the different components was not well understood by most researchers in the field. In this work a connection was established between the oxidant partial pressure in the growth ambient and deposition temperature. This relationship, derived from thermodynamical calculations for the Si-SiO$_2$-O$_2$-H$_2$O system, was confirmed by experimental evidence, both from experiments carried out as part of this work and from results reported in the literature on the oxidation behaviour of silicon under low oxidant partial pressures. For MOS transistors, crystalline quality is not as critical in terms of gain and speed as it is for bipolar devices; however, subthreshold leakage current is influenced by defects in the source/drain to substrate junctions and thereby can perturb the devices' switching behaviour significantly.

The results from the studies on oxide degradation and low temperature epitaxy indicated that suitable operating conditions for the fabrication of 3-D devices by ELO had to be a compromise between low epitaxial deposition temperature to maintain oxide integrity and a required minimum temperature to obtain defect free material. Since the relationship between the given and controllable system input parameters such as carrier gas moisture content, chamber pressure, and temperature was now understood, this compromise could be based on the data available. This subsequently led to typical operating temperatures for the Purdue epitaxy reactor system of about 900 °C, while it had been 950 °C before.

During the course of the investigation it became apparent that the ratio of horizontal to vertical growth rate or aspect ratio in selective epitaxial growth was essentially independent of operating conditions in the Purdue system and fixed at a value close to one. Thus it was necessary to planarize the overgrowth crystallites from heights of about 10 μm to about 0.5 μm above the polysilicon gate. It seemed helpful if the silicon crystallites would have been of same height prior to this planarization procedure. For that reason an investigation was carried out to better understand the growth rate behaviour in the SiH$_2$Cl$_2$ - HCl - H$_2$ system.

It was determined that selective epitaxial growth was extremely sensitive to temperature as long as wafers were mostly covered with masking oxide. When masking oxide coverage fell below about 40 % growth became much less sensitive to changes in temperature to the extend that non-uniformities
comparable to those obtained in bulk epitaxy, i.e. about ±3% were obtained. Surface reactions on the masking oxide with related high activation energies were considered to be mainly responsible for the extreme temperature dependence observed in selective epitaxy on wafers mostly covered with oxide.

For a uniform planarization of material 10 µm thick, to material 1 µm thick an initial non-uniformity of ±5% would map into a non-uniformity of ±50% for the final film. For the 3-D CMOS process this could not be tolerated and a planarization procedure had to be found that would tolerate initial large non-uniformities and yield final films of high uniformity.

The solution to this problem was found in chemo-mechanical polishing of the wafer with plasma deposited oxide or nitride as an etch stop. However, due to the lack of appropriate equipment and experience, the work was performed out of house. This step, first viewed as rather trivial, became one of the main obstacles in obtaining a working stacked PMOS device.

To evaluate the critical properties of the novel polyoxide/ELO interface, a new stacked capacitor structure was designed and built. For that purpose no planarization was necessary. The electrical test revealed reasonably good device behaviour, with surface state densities as low as $10^{11} \text{cm}^{-2}\text{eV}^{-1}$. It was thought that the abundance of hydrogen during interface formation helped in obtaining low surface state densities. Annealing in pure hydrogen at elevated temperatures did result in improvements of interfacial properties for the weak depletion region. It was determined that the interface created by the epitaxial lateral overgrowth of silicon over oxidized polysilicon did result in interfaces that were suited for the fabrication of MOS transistors, even though a further reduction in interface states was desirable. However, no further effort was expended to improve the electronic characteristics of that interface. However, an attempt was made to fabricate a stacked PMOS transistor. This attempt was partially successful as transistor action was demonstrated.

Three problems could be identified that were at least in part responsible for extreme difficulties in obtaining working stacked PMOS transistors. First, the planarization was not optimized. As a result, the thickness of the ELO material was in many areas of the wafer larger than about 1.5 µm, although this thickness was difficult to measure. The large thickness prevented the subsequent source and drain implant from reaching to the buried gate; thus, a conducting channel was not connected to the source and drain regions. Secondly, to form the top source and drain regions, an implantation mask had to be aligned to the polysilicon gate with high accuracy, i.e. better than 0.8 µm.
This was difficult to do technologically. The use of borosilicate glass masks with a thermal coefficient of expansion of 37 x 10^{-7} cm/cm °C and an estimated temperature variation of 5 °C between successive alignments resulted in changes of 1.5 μm for points initially 7 cm apart which clearly cannot be tolerated when alignment accuracies of better than 1 μm are required. Misalignment would result in gaps between the conducting channel and the source or drain implant. Thirdly, a high energy high dose implant was required to implant source and drain regions that would reach at least 1 μm down into the ELO substrate. The associated straggle and lateral diffusion in conjunction with the required slight overdeveloping and overetching of the masked backside channel region lead to the result of shortcircuited PMOS source and drain.

To obtain a higher yield and better performance of 3-D stacked PMOS transistors the mask set as well as the process have to be refined. Specifically, the mask used for the definition of the source and drain implanted regions for the top device need to be redesigned such that the necessary overetching and overdeveloping as well as lateral diffusion during annealing is compensated for. However, it would still be challenging to obtain good area definition in the 3 μm range by wet etching and in the presence of non-negligible environmental temperature and humidity changes. To that end, the use of better controlled processing areas and alignment equipment which are now available in the Purdue laboratory may be required. The second major problem was the ELO planarization. This step is not trivial and requires a lot of expertise in polishing techniques. The development of know-how in that area in house may become necessary to reduce turn-around times and to optimize the planarization step as circumstances require.

Practical aspects of the new technology have yet to be investigated more closely. This particularly concerns dopant redistribution during the epitaxy and during the top transistor source/drain drive in. The top transistor threshold voltage shift due to outdiffusion of the dopant used for the polysilicon implant and bottom transistor junction movement during the high temperature polysilicon oxidation are problematic. Here very recent research on plasma-deposited polysilicon oxides may be helpful. Additional research is needed to establish maximum selective epitaxial growth rates with no nucleation as a function of deposition pressure, deposition temperature, and wafer oxide coverage.
The redesign of the masks and the other process refinements had to be expected as there was basically no information whatsoever on 3 μm processing in the Purdue laboratory when work on this project began. However, it has been demonstrated that 3-D stacked active devices can be fabricated by ELO and it is hoped that as a result of this research 3-D CMOS SOI technologies will become a reality in the near future.
LIST OF REFERENCES


147. S. T. Liu, L. Chan, and J. O. Borland, "Reaction Kinetics of SiO$_2$-Si(100) Interface in H$_2$ Ambient in a Reduced Pressure Epitaxial Reactor," *10th Intl. Conf. on CVD*, October 1987, Honolulu.


176. G. Ghidini and F.W. Smith, "Interaction of $\text{H}_2\text{O}$ with Si(111) and (100)," *J. Electrochem. Soc.*, vol. 131, no. 12, pp. 2924-2928, December 1984.


APPENDIX A

3D CMOS Run Sheet for OG Cap. Evaluation.

1. Obtain an n-type substrate. 
   3" diameter  
   resistivity: 50-100 Ω-cm  
   orientation: (100)  

2. Mark wafer on backside  

3. Ultraclean wafer  
   5 min. TCA in US cleaner  
   5 min. acetone in US cleaner  
   2 min. methanol rinse  
   1 min. rinse under flowing DI  
   Blow-dry wafer with dry nitrogen  
   Mix 5 parts H₂SO₄ with 2 parts H₂O₂ and leave wafers in this solution  
   for at least 5 minutes  
   5 min. rinse under flowing DI. THIS IS VERY IMPORTANT. NO SHORTCUTS  
   HERE as it will have disastrous effects later on  

4. Implant threshold adjust  
   Species: Arsenic  
   Energy: 25 keV  
   Dose: 10⁻¹⁵ cm⁻²  

5. Ultraclean wafer  

6. Field Oxide (4000 - 4400 Å)  
   push: 3' N₂ @ 1100°C  
   purge: 5' N₂ @ 1100°C  
   drive-in: 35' H₂ burn @ 1100°C Tube #4  
   purge: 5' N₂ @ 1100°C  
   pull: 3' N₂ @ 1100°C  

7. Ultraclean or direct transfer to LPCVD  
   If wafer is directly transferred from oxidation furnace  
   to LPCVD tube without having it stored somewhere for more  
   than about one hour no ultraclean is necessary.  

8. LPCVD polysilicon (3000 Å)  
   Temperature: 580°C  
   SiH₄ flow rate: 50 sccm  
   Time: 60'  

9. Ultraclean wafer
10. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

11. Apply adhesion promoter HMDS
   Subject wafers to HMDS-vacuum for 5 min.

12. Apply AZ-4110 positive photoresist
   30" @ 5000 rpm
   Blow wafer off with dry nitrogen prior to applying resist.

13. Place wafer in prebake oven.
   Temperature: 90°C
   Time: 30'
   Use the big BLUE-M oven. DO NOT USE THE SMALL PREBAKE OVEN ON THE SPINNER BENCH.

14. Expose mask #2 (poly gate).
    setting on Cobilt: approx. 20
    This has to be adjusted such that development time is between 60 and 90".
    Clean mask before using it in ACE/METH/DI.
    Align wafer flat such that it is 45° off of the mask edge that faces you.

15. Develop AZ-4110 positive photoresist.
    Developer AZ400K: H₂O 1:4, 60-90"
    Try to obtain sharp corners and no ragged edges. This usually indicates optimized exposure and development times.
    Don't think it's o.k.! Overexposure and overdeveloping by just 0.5 μm leads to 50% linewidth loss for the most critical devices!

16. Inspect under microscope

17. Place wafer in hardbake oven.
    Temperature: 120°C
    Time: 20'

18. Plasma-etch polysilicon
    Use Technics system
    CF₄ mass flow: 30 on rotameter
    O₂ mass flow: 4 on rotameter
    Temperature: 40°C
    Pressure: 350 mTorr
    Power: 150 W
    Time: 1-4', depending on loading
    This step is no fun. It is very easy to overetch the 2 μm polygates.
    Stop the etch often and inspect ALL wafers. Change position of wafers before continuing. The Technics system is not very consistent. Don't assume because it took 3 min. last time it's going to take 3 min. this time! Eventually try wet etching (see thesis of John Denton) or a dedicated polyetcher which will hopefully become available soon.
19. Inspect.

20. Remove photoresist in solvent (ACE)

21. Ultraclean wafer.

22. Polysilicon implant
   Species: Phosphorus
   Energy: 125 keV
   Dose: $10^{16}$ cm$^{-2}$

23. Ultraclean wafer

24. Polysilicon oxidation (900 Å)
   push: 3' $N_2 @ 1100^\circ C$
   purge: 5' $N_2 @ 1100^\circ C$
   drive-in: 35' $O_2 @ 1100^\circ C$ Tube #5
   purge: 5' $N_2 @ 1100^\circ C$
   pull: 3' $N_2 @ 1100^\circ C$

25. Ultraclean wafer

26. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

27. Apply adhesion promoter HMDS

28. Apply AZ-4110 positive photoresist
   30'' @ 5000 rpm

29. Place wafer in prebake oven.
   Temperature: 90°C
   Time: 30'

30. Expose mask #3 (seed windows).
   setting on Cobilt: approx. 20
   This has to be adjusted such that development time is between 60 and 90''.
   This alignment is critical. Misalignment needs to be less than about 0.5μm. Good
   luck! The seeds are aligned to the polygates. Use the alignment markers for
   alignment. Set MASK SEPARATION to max. value. Align the center die visually
   first. Then align the die borders. Thereafter align the text CAPTEST SEED to the
   text CAPTEST POLY. Now the chances are minimized that you will align to the
   wrong alignment markers. Reduce MASK SEPARATION more and more as the
   alignment zeros in. If the THETA turn knob runs out of range, bring the wafer into
   contact with the mask and then turn the knob back a few turns. Thereafter separate
   mask and wafer again. It will be difficult at best to align both left and right field
   perfectly. Concentrate on just one field. Shut off vibration causing equipment such
   as the big airconditioners.
31. Develop AZ-4110 positive photoresist.  
   Developer AZ400K : H₂O 1:4, 60-90"  
   There is no margin for overdeveloping! Zero! If you do not see a little space between  
   the polysilicon and the seed hole edges, or if you see a lot of notches in the seed  
   hole lithography, DO IT OVER. Otherwise, the next week of work will be a waste  
   of time. Reflective notching sometimes occurs when either prebake temperature was  
   too low or exposure was too long. Try to reduce exposure time and lengthen  
   development time.  

32. Inspect under microscope  

33. Place wafer in hardbake oven.  
   Temperature: 120°C  
   Time: 20'  

34. Etch oxide in BHF  
   Estimated etch time: 4.5 min.  
   Record etch time:_______ min.  
   No margin for overetching!  

35. Inspect under microscope  

36. Remove resist in solvents  

37. Ultraclean wafer  

38. Do selective epitaxy  
   Temperature: 900°C  
   Time: 55'  
   HCl mass flow: 0.59 slm  
   DCS mass flow: 0.22 slm  
   H₂ mass flow: 60 slm  
   Wafers placed on susceptor outside.  
   There may be some nucleation. These values may require some  
   adjustment.  

39. Ultraclean wafer  

40. Blanket implant  
   Species: Phosphorus  
   Energy: 35 keV  
   Dose: 5*10^15 cm^-2  

41. Ultraclean wafer
42. Protective oxide (1000 Å)
   push: 3' N₂ @ 920°C
   purge: 5' N₂ @ 920°C
   anneal: 20' H₂ burn @ 920°C Tube #4
   purge: 5' N₂ @ 920°C
   pull: 3' N₂ @ 920°C
   Here the slow push and pull are critical to reduce stressing of the composite structure.

43. Ultraclean wafer

44. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

45. Apply adhesion promoter HMDS

46. Apply AZ-4110 positive photoresist
   30' @ 5000 rpm

47. Place wafer in prebake oven.
   Temperature: 90°C
   Time: 30'

48. Expose mask #5 (contacts).
   setting on Cobilt: approx. 20
   This has to be adjusted such that development time is between 60 and 90''.
   Alignment is to polysilicon. Can align to features.
   Misalignment allowed is 1.5 μm.

49. Develop AZ-4110 positive photoresist.
   Developer AZ400K : H₂O 1:4, 60-90''

50. Inspect under microscope

51. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 20'

52. Etch thermal oxide in BHF.
   Estimated etch time: 2 min.
   Recorded etch time: min.

53. Remove resist.
   The wafer could now be tested on the micro-manipulator probestation.

54. Ultraclean.
56. H$_2$ anneal (optional)
   Time: 10'
   Temperature: 1000°C
   Carried out in epitaxy reactor. This step didn't do too much in past experiments and may be omitted.

57. Ultraclean wafer

58. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

59. Apply adhesion promoter HMDS

60. Apply AZ-4110 positive photoresist
   30" @ 4000 rpm

61. Place wafer in prebake oven.
   Temperature: 80°C
   Time: 15'
   Notice different temperature and time for lift-off!

62. Expose mask #6 (metal definition).
   Setting on Cobilt: approx. 20
   This has to be adjusted such that development time is between 60 and 90''.

63. Develop AZ-4110 positive photoresist.
   Developer AZ400K : H$_2$O 1:4, 60-90''

64. Inspect under microscope

65. Aluminum deposition.
   Sputter to 5000Å thickness (20 min).

66. Inspect.

67. Metal lift off

68. Inspect.

69. Metal anneal
   Anneal: 15' N$_2$ @ 450°C Tube #8

70. Electrical testing.
APPENDIX B

3D CMOS Run Sheet for PMOS Transistor Fabrication

1. Obtain an n-type substrate.
   3" diameter
   resistivity: 50-100 Ω-cm
   orientation: (100)

2. Mark wafer on backside

3. Ultraclean wafer
   5 min. TCA in US cleaner
   5 min. acetone in US cleaner
   2 min. methanol rinse
   1 min. rinse under flowing DI
   Blow-dry wafer with dry nitrogen
   Mix 5 parts H$_2$SO$_4$ with 2 parts H$_2$O$_2$ and leave wafers in this solution
   for at least 5 minutes
   5 min. rinse under flowing DI. THIS IS VERY IMPORTANT. NO SHORTCUTS
   HERE as it will have disastrous effects later on

4. Implant threshold adjust
   Species: Arsenic
   Energy: 25 keV
   Dose: 10$^{15}$ cm$^{-2}$

5. Ultraclean wafer

6. Field Oxide (4000 - 4400 Å)
   push: 3' N$_2$ @ 1100°C
   purge: 5' N$_2$ @ 1100°C
   drive-in: 35' H$_2$ burn @ 1100°C Tube #4
   purge: 5' N$_2$ @ 1100°C
   pull: 3' N$_2$ @ 1100°C

7. Ultraclean or direct transfer to LPCVD

8. LPCVD polysilicon (3000 Å)
   Temperature: 580°C
   SiH$_4$ flow rate: 50 sccm
   Time: 60'

9. Ultraclean wafer

10. Place wafer in hardbake oven.
    Temperature: 120°C
    Time: 10'
11. Apply adhesion promoter HMDS

12. Apply AZ-4110 positive photoresist
   30" @ 5000 rpm

13. Place wafer in prebake oven.
   Temperature: 90°C
   Time: 30'
   Use the big BLUE-M oven. DO NOT USE THE SMALL PREBAKE OVEN ON THE SPINNER BENCH.

14. Expose mask #2 (poly gate).
   setting on Cobilt: approx. 20
   This has to be adjusted such that development time is between 60 and 90".
   Clean mask before using it in ACE/METH/DI.
   Align wafer flat such that it is 45° off of the mask edge that faces you.

15. Develop AZ-4110 positive photoresist.
    Developer AZ400K : H₂O 1:4, 60-90"

16. Inspect under microscope

17. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 20'

18. Plasma-etch polysilicon
    Use Technics system
    CF₄ mass flow: 30 on rotameter
    O₂ mass flow: 4 on rotameter
    Temperature: 40°C
    Pressure: 350 mTorr
    Power: 150 W
    Time: 1-4', depending on loading
    This step is no fun. It is very easy to overetch the 2 µm polygates.
    Stop the etch often and inspect ALL wafers. Change position of wafers before continuing. The Technics system is not very consistent. Don't assume because it took 3 min. last time it's going to take 3 min. this time! Eventually try wet etching (see thesis of John Denton) or a dedicated polyetcher which will hopefully become available soon.

19. Inspect.

20. Remove photoresist in solvent

21. Ultraclean wafer.
22. Polysilicon implant
   Species: Phosphorus
   Energy: 125 keV
   Dose: 10^{16} cm^{-2}

23. Ultraclean wafer

24. Polysilicon oxidation (900 Å)
   push: 3' N_2 @ 1100°C
   purge: 5' N_2 @ 1100°C
   drive-in: 35' O_2 @ 1100°C Tube #5
   purge: 5' N_2 @ 1100°C
   pull: 3' N_2 @ 1100°C

25. Ultraclean wafer

26. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

27. Apply adhesion promoter HMDS

28. Apply AZ-4110 positive photoresist
   30'' @ 5000 rpm

29. Place wafer in prebake oven.
   Temperature: 90°C
   Time: 30'

30. Expose mask #3 (seed windows).
   setting on Cobilt: approx. 20
   This has to be adjusted such that development time is between 60 and 90''.
   This alignment is critical. Misalignment needs to be less than about 0.5μm. Good luck! The seeds are aligned to the polygates. Use the alignment markers for alignment. Set MASK SEPARATION to max. value. Align the center die visually first. Then align the die borders. Thereafter align the text CAPTEST SEED to the text CAPTEST POLY. Now the chances are minimized that you will align to the wrong alignment markers. Reduce MASK SEPARATION more and more as the alignment zeros in. If the THETA turn knob runs out of range, bring the wafer into contact with the mask and then turn the knob back a few turns. Thereafter separate mask and wafer again. It will be difficult at best to align both left and right field perfectly. Concentrate on just one field. Shut off vibration causing equipment such as the big airconditioners.

31. Develop AZ-4110 positive photoresist.
   Developer AZ400K: H_2O 1:4, 60-90''
   There is no margin for overdeveloping!

32. Inspect under microscope
33. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 20'

34. Etch oxide in BHF
   Estimated etch time: 4.5 min.
   Record etch time: ______ min.
   No margin for overetching!

35. Inspect under microscope

36. Remove resist in solvents

37. Ultraclean wafer

38. Do selective epitaxy
   Temperature: 900°C
   Time: 70'
   HCl mass flow: 0.6 slm
   DCS mass flow: 0.22 slm
   H₂ mass flow: 60 slm
   Wafers placed on susceptor outside.
   The flow rates may have to be adjusted. Should get at least 9 μm of growth.

39. Ultraclean wafer

40. Plasma-nitride deposition as polishing stop
   Use Technics system
   SiH₄ mass flow: 19.5 sccm
   NH₃ mass flow: 43.6 sccm
   Temperature: 300°C
   Pressure: 400 mTorr
   Power: 50 W
   Time: 20' for 5000Å

41. Chemo-mechanical polishing (Harris Semiconductor)

42. Strip residual nitride in hot boiling phosphoric acid.
   Temperature: 190°C
   Time: 15 min.
   This is a dangerous procedure. Don't inhale the steam!

43. Rinse in DI for 5 minutes

44. Blow dry with dry nitrogen

45. Ultraclean
46. Backsurface threshold implant
   Species: Phosphorus
   Energy: 25 keV
   Dose: $5 \times 10^{13} \text{ cm}^{-2}$

47. Plasma-oxide deposition
   Use Technics system
   SiH$_4$ mass flow: 10 sccm
   NO$_2$ mass flow: 30 sccm
   Temperature: 300°C
   Pressure: 350 mTorr
   Power: 50 W
   Time: 10' for 2500Å

48. Ultraclean wafer

49. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

50. Apply adhesion promoter HMDS

51. Apply AZ-4110 positive photoresist
   30' @ 5000 rpm

52. Place wafer in prebake oven.
   Temperature: 90°C
   Time: 30'

   setting on Cobilt: approx. 20
   This has to be adjusted such that development
   time is between 60 and 90'.
   Alignment is to polygates. Misalignment 0.8 μm max. Good luck!

54. Develop AZ-4110 positive photoresist.
   Developer AZ400K : H$_2$O 1:4, 60-90'

55. Inspect under microscope

56. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 20'

57. Etch thermal oxide in BHF.
   Estimated etch time: 3.5 min.
   Recorded etch time: _______ min.

58. PMOS deep S/D implant
   Species: Boron
   Energy: 175 keV
   Dose: $5 \times 10^{15} \text{ cm}^{-2}$
59. PMOS shallow S/D contact implant
   Species: Boron
   Energy: 25 keV
   Dose: $5 \times 10^{15} \text{ cm}^{-2}$

60. Remove resist. This is not easy. Soak in hot Nophenol for half a day.

61. Rinse in DI.

62. Subject to $\text{H}_2\text{SO}_4$-$\text{H}_2\text{O}_2$ solution.

63. Rinse in DI.

64. Use LFE plasma stripper with oxygen at 400 Watts.
   Place wafer with front side towards end of tunnel all the way in.
   Etch until resist is all gone. This may take more than one hour.

65. Ultraclean wafer

66. Boron anneal and PMOS S/D drive in
   push: 3' $\text{N}_2$ @ 920°C
   purge: 5' $\text{N}_2$ @ 920°C
   drive-in: 20' $\text{H}_2$ burn @ 920°C Tube #4
   purge: 5' $\text{N}_2$ @ 920°C
   pull: 3' $\text{N}_2$ @ 920°C

67. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

68. Apply adhesion promoter HMDS

69. Apply AZ-4110 positive photoresist
   30" @ 5000 rpm

70. Place wafer in prebake oven.
   Temperature: 90°C
   Time: 30'

71. Expose mask #5 (contacts).
   setting on Cobilt: approx. 20
   This has to be adjusted such that development time is between 60 and 90". Alignment is to OG S/D

72. Develop AZ-4110 positive photoresist.
   *Developer AZ400K : $\text{H}_2\text{O}$ 1:4, 60-90"

73. Inspect under microscope
74. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 20'

75. Etch thermal oxide in BHF.
   Estimated etch time: 3 min.
   Recorded etch time:_____min.

76. Remove resist in solvents

   The wafer could now be tested on the micro-
   manipulator probestation.

77. Ultraclean.

78. H₂ anneal (optional)
   Time: 10'
   Temperature: 1000°C
   Carried out in epitaxy reactor

79. Ultraclean wafer

80. Place wafer in hardbake oven.
   Temperature: 120°C
   Time: 10'

81. Apply adhesion promoter HMDS

82. Apply AZ-4110 positive photoresist
   30'' @ 4000 rpm

83. Place wafer in prebake oven.
   Temperature: 80°C
   Time: 15'
   Note reduced time and temperature for lift-off

84. Expose mask #6 (metal definition).
   setting on Cobilt: approx. 20
   This has to be adjusted such that development
   time is between 60 and 90''.

85. Develop AZ-4110 positive photoresist.
   Developer AZ400K : H₂O 1:4, 60-90''

86. Inspect under microscope

87. Aluminum deposition.
   Sputter to 5000Å thickness (20 min).

88. Inspect.
89. Metal lift off
   20' in acetone, short ultrasonic pulses

90. Inspect.

91. Metal anneal
   anneal: 15' N₂ @ 450°C Tube #8

92. Electrical testing.
<table>
<thead>
<tr>
<th>Run</th>
<th>HCl</th>
<th>DCS</th>
<th>HC/DCS</th>
<th>Veh/25kg</th>
<th>Fold</th>
<th>Tax</th>
<th>Time</th>
<th>GR in</th>
<th>GR cut</th>
<th>Max</th>
<th>Min</th>
<th>GR int</th>
<th>Total</th>
<th>Area</th>
<th>Tarea</th>
<th>Remark</th>
<th>HOCa, Num. in, Num. out</th>
</tr>
</thead>
<tbody>
<tr>
<td>74</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.32</td>
<td>20</td>
<td>-0.308</td>
<td>-0.319</td>
<td>0.052</td>
<td>0.218</td>
<td>0.016</td>
<td>1.98</td>
<td>0.078</td>
<td>14.0</td>
<td>0.078</td>
<td>14.0</td>
<td>37.5</td>
<td>14.0</td>
</tr>
<tr>
<td>75</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.36</td>
<td>20</td>
<td>0.006</td>
<td>0.123</td>
<td>0.140</td>
<td>0.174</td>
<td>0.017</td>
<td>0.085</td>
<td>0.044</td>
<td>0.120</td>
<td>0.044</td>
<td>1.20</td>
<td>1.20</td>
<td>0.120</td>
</tr>
<tr>
<td>76</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.37</td>
<td>20</td>
<td>-0.301</td>
<td>-0.300</td>
<td>0.037</td>
<td>0.133</td>
<td>0.200</td>
<td>0.053</td>
<td>0.125</td>
<td>0.053</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>77</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.38</td>
<td>20</td>
<td>0.004</td>
<td>0.116</td>
<td>0.139</td>
<td>0.174</td>
<td>0.017</td>
<td>0.162</td>
<td>0.057</td>
<td>0.162</td>
<td>0.057</td>
<td>0.162</td>
<td>0.162</td>
<td>0.162</td>
</tr>
<tr>
<td>78</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.40</td>
<td>20</td>
<td>0.003</td>
<td>0.103</td>
<td>0.127</td>
<td>0.174</td>
<td>0.017</td>
<td>0.209</td>
<td>0.065</td>
<td>0.209</td>
<td>0.065</td>
<td>0.209</td>
<td>0.209</td>
<td>0.209</td>
</tr>
<tr>
<td>79</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.40</td>
<td>20</td>
<td>0.002</td>
<td>0.100</td>
<td>0.123</td>
<td>0.174</td>
<td>0.017</td>
<td>0.245</td>
<td>0.079</td>
<td>0.245</td>
<td>0.079</td>
<td>0.245</td>
<td>0.245</td>
<td>0.245</td>
</tr>
<tr>
<td>80</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.40</td>
<td>20</td>
<td>0.001</td>
<td>0.096</td>
<td>0.100</td>
<td>0.174</td>
<td>0.017</td>
<td>0.281</td>
<td>0.111</td>
<td>0.281</td>
<td>0.111</td>
<td>0.281</td>
<td>0.281</td>
<td>0.281</td>
</tr>
<tr>
<td>81</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.40</td>
<td>20</td>
<td>0.000</td>
<td>0.093</td>
<td>0.088</td>
<td>0.174</td>
<td>0.017</td>
<td>0.316</td>
<td>0.123</td>
<td>0.316</td>
<td>0.123</td>
<td>0.316</td>
<td>0.316</td>
<td>0.316</td>
</tr>
<tr>
<td>82</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.40</td>
<td>20</td>
<td>0.000</td>
<td>0.088</td>
<td>0.081</td>
<td>0.174</td>
<td>0.017</td>
<td>0.340</td>
<td>0.137</td>
<td>0.340</td>
<td>0.137</td>
<td>0.340</td>
<td>0.340</td>
<td>0.340</td>
</tr>
<tr>
<td>83</td>
<td>1.240</td>
<td>0.000</td>
<td>60</td>
<td>0.40</td>
<td>20</td>
<td>0.000</td>
<td>0.084</td>
<td>0.075</td>
<td>0.174</td>
<td>0.017</td>
<td>0.362</td>
<td>0.152</td>
<td>0.362</td>
<td>0.152</td>
<td>0.362</td>
<td>0.362</td>
<td>0.362</td>
</tr>
</tbody>
</table>

**APPENDIX C**
APPENDIX D

/* TSPICE input file for 3-D PMOS simulation

ugfrom=-3
ugto=1
ugstep=1
vdmax=-10

Title MOSFET ID vs VDS with gate voltage as parameter

Circuit vg
vds 4 0 v:dc=dcin
vg 1 0 v:dc=vg
ds 3 0 diode
dd 4 2 diode
rb 5 0 r: r=10e7
cg 1 0 c: c=10f
cd 2 0 c: c=10f
cs 3 0 c: c=10f
cb 5 0 c: c=10f
qpmos 3 1 2 5 pmos

Model pmos mfet2 : w=10u l=2.0u level=3 $
  ad=3e-11 as=3e-11 pd=20e-6 ps=20e-6 $
  pol=pmos vto=-2 kp=0.2u gamma=2 phi=0.6 rd=100 rs=100 $
  cbd=9e-10 cbs=9e-10 nsub=2e16 nss=2e11 tpg=-1 ld=0.2u uo=100 $
  tox=0.1e-6 cgso=4e-10 cgdo=4e-10 cj=220u mj=0.5 cjsw=500p $
  mjsw=0.33 pb=0.7 xj=0.5u delta=1 eta=1 theta=0.2

model diode d : is=1.0f rs=100 cjo=10p bv=0.1 ibv=15u pb=1
Endc

Var sweep vg:ugfrom ugto ugstep type=lin
DC analysis dcin:0 vdmax vdmax/100
keep i(vds,2)=id
Enddc
Endvar
probe id
plot id
Joerg Friedrich was born June 27, 1961 in Bochum, West Germany. He grew up in Ennepetal, West Germany, where he graduated from the local highschool in 1980. He then attended the Ruhr-University Bochum, West Germany, where among other things he worked on speech processing, layout verification systems, and high-speed bipolar circuit design.

In 1984 he received a scholarship from the German Academic Exchange Service to continue his studies in Electrical Engineering at Purdue University, West-Lafayette, IN. During his stay at Purdue he worked on the design of a microcomputer based HVDC simulator while also completing his diploma thesis work on CMOS adder circuit concepts. In August 1985 he graduated with the grade "very good" from the Ruhr-University Bochum with a diploma in Electrical Engineering.

During the fall of 1985 he attended the University of Florida in Gainesville, Florida, where he worked on the implementation of a power engineering circuit simulator. In January 1986 he returned to Purdue University to work towards a Ph.D. degree in Electrical Engineering. He was married to Baerbel Haberstroh in October 1988.

Mr. Friedrich has held several offices in various student groups. He is a member of the Electrochemical Society.