Plane Lapping Process of Silicon Wafers

DOBRESCU TIBERIU GABRIEL, PASCU NICOLETA ELISABETA,
GHINEA MIHAI, POPESCU ADRIAN
Machine and Manufacturing Systems Department
University Politehnica of Bucharest
313 Splaiul Independentei st. Sector 6 Bucharest
ROMANIA
tibidobrescu@yahoo.com

Abstract: - Lapping is a very complicated and random process resulting from the variation of abrasive grains by its size and shapes and from the numerous variables which have an effect on the process quality. Thus it needs to be analyzed by experiment rather than by theory to obtain the relative effects of variables quantitatively. In this study, the plane lapping was performed and analyzed by ANOVA table. As a result, effective variables and interaction effects were identified and disclosed. Also the optimal variable combination to obtain the largest percentage improvement of surface roughness was selected and confirmatory experiments were performed.

Key-Words: - Lapping, silicon wafers, ANOVA, orthogonal array, roughness, abrasive grain.

1 Introduction

Monocrystalline silicon is presently the most commonly used substrate material for the production of microelectronic components (microchips) worldwide.

Wafer diameter has increased steadily from less than 50 mm in the 1970s to 200 mm today and possibly to 300 mm in the near future.

Manufacturing of silicon wafers starts with growth of silicon ingots. A sequence of processes is needed to turn an ingot into wafers. This typically consists of following processes[1]:
- Slicing, silicon ingot into wafers of this disk shape;
- Edge profiling (chamfering), to chamfer the peripheral edge portion of the wafer;
- Flattening (lapping or grinding), to achieve a high degree of parallelism and flatness of the wafer;
- Etching, to chemically remove the damage induced by slicing and flattening without introducing further mechanical damage;
- Rough polishing, to obtain a mirror surface on the wafer;
- Fine polishing, to obtain final mirror surface;
- Cleaning, to remove the polishing agent or dust particles from the wafer surface.

The machining of this material is critical to high quality standards. With the development of new components, the electronic industry now has higher standards for the total thickness variation (TTV) and the wafer warp (less than 1 µm on a wafer with a diameter of 200 mm). In the field of surface finishing, the competitive technologies of double-sided lapping are applied.

Lapping is a finish method used to obtain good surface quality. Important variables lapping efficiency are abrasive grain size, lapping pressure, lapping speed, quantity of lapping compound supplied and viscosity of the compound etc. Comparison of the effects of variables on the overall process efficiency is not yet clear owing to the complexity and randomness of the process. A generalized and simplified model of the lapping process is shown in Fig. 1, in which two motion of abrasive grains in the lapping process [2] are represented.

Using the experimental design method is a typical approach to efficiently and logically identify characteristics of any complex process by experiment. In this method, experiments are generally designed by orthogonal array and the results are analyzed by ANOVA (ANalysis Of VAriance). A number of studies related to the experimental design method have been reported in the field of quality control and statistics. In this study, experimental designed by Taguchi's orthogonal array [4] were performed.

2 Double-side lapping

Double-side lapping is an established process in modern wafer production. Following the wafering operation by ID cut-off grinding, lapping is used for eliminating the sow marks, reducing the depth of damage and in particular for improving the form accuracy. Nowadays, for VLSI and ULSI application the following specifications are
demanded: thickness tolerance +/- 3 µm, parallelism 2 µm (in certain cases 1 µm) TTV, flatness < 1 µm and roughness R_z < 1 µm [2].

A common feature of all lapping processes is that the material removal takes place by the action of loose abrasive grains which are dispersed in a watery or oily lapping fluid.

The workpiece shape is produced by the relative motion between the lapping wheel pieces surfaces coupled with the action of the loose lapping compound located between these two active surfaces.

The engagement of the lapping grains takes place non-directionally and stochastically so that directional abrasion grooves are generally avoided. For large volume wafer production, mostly high powered special purpose double-sided lapping machines with planetary drive systems are used. The primary features of these modern lapping machines are [1]:
- Programmed and gauging controls which allow pressure-controlled operation;
- The infinitely variable speed adjustment of the drive motors (soft-start control);
- The application of grooved, temperature-controlled lapping wheels.

The necessity to gradually increase the lapping pressure in parallel with ensuring a soft-start control may be attributed to the extreme sensitivity of the wafers to brittle fracture.

3 Lapping abrasives
The lapping grains most commonly used for silicon wafers are: corundum (Al₂O₃), silicon carbide (SiC) or specially prepared mixtures such as Al₂O₃ and ZrO₂. Diamond and boron carbide abrasives play a secondary role due to their higher costs. The grain size range from 100 µm (rough lapping), to 5 µm (fine finishing) [3]. For single-step lapping operations in general grain size of 12 to 15 µm are generally applied [5].

The material removal during silicon lapping takes place predominantly through cracking, splitting and brittle fracture of the crystalline material [6, 7]. This is caused by abrasive grains which roll between the wafer surface and the lapping wheel this rolling action causes the sharp corners to press into the material (Fig.1). During lapping, the total number of coarse active lapping grains (gram 1 and 2 in Fig.1) dictates the material removal rate. An abrasion by scratching and scraper formation is undesirable in general because of the increased risk of wafer fracture and the greater depth of damage. This is partly caused by lapping grains which are anchored temporarily in the lapping wheel, (grain 1 in Fig.1).

Distinct scratches are produced occasionally by Si particles which break away from the wafer periphery. In order to avoid this edge-chipping, the wafer circumference is ground with a profiled diamond wheel, so called edge rounding. The temporary anchoring of abrasives in the lapping wheel can generally be prevented by the use of hard wheel materials (cast iron) [8, 9, 10].

4 Experiment
4.1 Experiment design
Two-level fractional factorial design, which is used in this study, is especially efficient in finding out important variables having an effect on the process performance.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Factor</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Grain size</td>
<td>#320, #600</td>
</tr>
<tr>
<td>B</td>
<td>Lapping pressure</td>
<td>4.144 [N/cm²], 0.828 [N/cm²]</td>
</tr>
<tr>
<td>C</td>
<td>Number of supply of lapping compound</td>
<td>41, 20</td>
</tr>
<tr>
<td>D</td>
<td>Lapping speed</td>
<td>30 rot/min, 10 rot/min</td>
</tr>
</tbody>
</table>

Table 1.

Experiments at two levels of each control variable were conducted: grain size, lapping pressure, number of lapping-compound-supply and
lapping speed. Control variables and their levels are shown in Table 1.

Percentage improvement of surface roughness before and after lapping was taken as an evaluation or response variable. Taguchi's fold-overtype orthogonal array is shown in Table 2.

In this study, only four main effects (i.e. A, B, C, D) and three two-factor interaction effects (i.e. AB, AC and BC) were considered in the experimental design. All the other effects were considered to be trivial.

The first column in Table 2 represents the standard order of experimental treatment combinations (tc). The number of combinations is eight, which is the smallest number required to satisfy resolution IV design with four variables. The other columns represent $2^3 - 1 (=7)$ contrasts: four main effects are assigned to the first, second, fourth and seventh columns and interaction effects to the third, fifth and sixth columns. In the Table 2, “+” represents high-level and “-” low-level. Though the levels in the experiment are represented as numerical quantitative, they can also be considered to have qualitative meaning, "high" or "low".

<table>
<thead>
<tr>
<th>Treatment combination (tc)</th>
<th>Column number and contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>A B AB AC BC ABC</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>-</td>
</tr>
<tr>
<td>a</td>
<td>+</td>
</tr>
<tr>
<td>b</td>
<td>-</td>
</tr>
<tr>
<td>ab</td>
<td>+</td>
</tr>
<tr>
<td>c</td>
<td>-</td>
</tr>
<tr>
<td>ac</td>
<td>+</td>
</tr>
<tr>
<td>bc</td>
<td>-</td>
</tr>
<tr>
<td>abc</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 2

4.2 Experimental results

Experiments were performed by a plane-parallel lapping machine MELCHIORRE SP3/600/2PR. All of the silicon wafers have an initial surface roughness of $R_a \ 0.63 \ \mu m - R_a \ 0.8 \ \mu m$.

<table>
<thead>
<tr>
<th>Column number</th>
<th>$R_a$ improvement [%]</th>
<th>$y$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$y_1$</td>
<td>$y_2$</td>
</tr>
<tr>
<td>(1)</td>
<td>47.64</td>
<td>48.43</td>
</tr>
<tr>
<td>a</td>
<td>37.41</td>
<td>37.46</td>
</tr>
<tr>
<td>b</td>
<td>55.54</td>
<td>54.77</td>
</tr>
<tr>
<td>ab</td>
<td>53.89</td>
<td>51.14</td>
</tr>
<tr>
<td>c</td>
<td>42.26</td>
<td>42.86</td>
</tr>
<tr>
<td>ac</td>
<td>39.04</td>
<td>36.27</td>
</tr>
<tr>
<td>bc</td>
<td>80.84</td>
<td>81.29</td>
</tr>
<tr>
<td>abc</td>
<td>69.32</td>
<td>69.84</td>
</tr>
</tbody>
</table>

Table 3

Four variables were considered as control variables: grain size (A); lapping pressure (B); number of lapping-compound- supply (C); and lapping speed (D). Experimental levels for each control variable are shown in Table 1. Surface roughness was measured before and after lapping using profilometer HOMMEL-TESTER Typ TR. Experimental results are shown in Table 3, in which $y_i$ (i=1, 2, 3, 4, 5, 6, 7 - number of silicon wafers) are percentage improvements of surface roughness and $y$ is the averaged value of the seven.

In this study, we are interested in the efficiency or percentage improvement of surface roughness, and equations used are as follows:

\[ y = \frac{\Delta K_a (\text{Improvement of surface roughness})}{K_a (\text{surface roughness before lapping})} \times 100 \]  
\[ \Delta K_a = K_a \text{of before lapping} - K_a \text{of after lapping} \]

5 Analysis and discussion

In lapping experiments with SiC abrasives, it was found that higher material removal rates can be achieved by increasing the average path velocity further by higher lapping pressures and by the use of grains of larger size.

The coarser F320 SiC grain (average grain size $d_k = 30 \ \mu m$) with $v_1 = 10 \ \text{m/min}$ approximately produces only a doubling of the removal rate as
compared with the fine F600 - grain (12 µm). On the other hand, with the same F600 abrasives, the removal rate can be increased by a factor of 2.5 to 3 if the path velocity is increased from 10 m/min to 30 m/min.

Lapping with F600 abrasives, a velocity of \( v_2 = 30 \) m/min and pressures of between 3 and 5 N/cm\(^2\) have resulted in an average depth of damage of \( H_D = 10 - 11.5 \) µm. When using the F320 lapping grain under otherwise similar conditions a doubling of the depth of damage (\( H_D = 20 \) µm) results. The results described above imply that with fine-grained SiC abrasives, both high material removal and a final smooth finish operation can be realized if the lapping process is controlled in an appropriate manner. Thus reloading operations from one lapping machine (pre-lapping with coarse grained abrasive) to another (finishing with fine abrasives) can be avoided in many cases.

The efficiency of plane lapping of silicon wafers, are the percentage improvement of surface roughness, was increased at coarse grain size. Lapping pressure shows the largest effect of all variables considered in this study and the efficiency of lapping increased dramatically at high-level compared with that at low level.

Number of lapping-compound-supply has little effect on the response. But it has significant interaction with lapping pressure and, so, should be treated as an important control variable.

That is, the effect of lapping pressure which is most significant can be affected by the level of number of lapping-compound-supply.

It can be predicted, from the experimental results, that if the lapping compound is supplied continuously the efficiency of lapping will become much higher.

Lapping speed has no effect and the efficiency at low speed is a little higher than that at high-level.

6 Conclusion
The plane lapping process of silicon wafers with plane parallel lapping machine MELCHIORRE SP/600/2PR have been characterized qualitatively by analyzing the effect of four control variables, namely: abrasive grain size, lapping pressure, number of lapping-compound-supply and lapping speed, on the percentage improvement of surface roughness as measure of efficiency using the experimental design method and the following results were obtained.

(1) Today, double-side lapping of the wafer is competitive technologies for improving the surface quality of a wafer after slicing.

(2) Lapping pressure has a significant effect on the efficiency of plane lapping of silicon wafers;

(3) Number of lapping-compound-supply should be treated as an important variable even though it has shown no effect on the efficiency because it interacts with lapping pressure.

(4) Optimal variable combination on lapping efficiency has proved to be low-levels for abrasive grain size and lapping speed, and high-levels for lapping pressure and number of lapping-compound-supply.

(5) The total thickness variation (TTV) of a 200 mm wafer can be reduced to 1 µm.

References: