Finite element analysis for grinding of wire-sawn silicon wafers: a designed experiment

Z.J. Pei a,∗, X.J. Xin b, W. Liu a

a Department of Industrial and Manufacturing Systems Engineering, Kansas State University, Manhattan, KS 66506, USA
b Department of Mechanical and Nuclear Engineering, Kansas State University, Manhattan, KS 66506, USA

Received 20 November 2001; received in revised form 26 August 2002; accepted 2 September 2002

Abstract

Silicon is the primary semiconductor material used to fabricate microchips. The quality of microchips depends directly on the quality of starting silicon wafers. A series of processes are required to manufacture high quality silicon wafers. Surface grinding is one of the processes used to flatten the wire-sawn wafers. A major issue in grinding of wire-sawn wafers is the reduction and elimination of wire-sawing induced waviness. This paper presents the results of a finite element analysis for grinding of wire-sawn silicon wafers. In this investigation, a four-factor two-level full factorial design is employed to reveal the main effects as well as the interaction effects of four factors (wafer thickness, waviness wavelength, waviness height and grinding force) on effectiveness of waviness reduction. The implications of this study to manufacturing are also discussed.

Keywords: Design of experiment; Factorial design; Finite element analysis; Grinding; Lapping; Machining; Material removal; Semiconductor material; Silicon wafers; Slicing

1. Introduction

1.1. Silicon wafers and their manufacturing processes

Integrated circuits (ICs) are built on semiconductor wafers. Over 90% of semiconductor wafers are silicon [1]. About 150 million silicon wafers of different sizes are manufactured each year worldwide [2]. In 1999, the worldwide revenue generated by silicon wafers was $5.8 billion. Semiconductor devices built on these wafers generated $132 billion in revenues. Electronics systems using this circuitry enjoyed revenue of $988 billion [3].

To manufacture high-quality silicon wafers, a sequence of processes is required. Some typical processes are listed below [4–8].

1. Slicing, to slice single crystal silicon ingot into wafers of thin disk shape;

2. Flattening (lapping or grinding), to achieve a higher degree of parallelism and flatness of the wafer;

3. Etching, to chemically remove the damage induced by slicing and flattening without introducing further mechanical damage;

4. Polishing, to obtain a smooth wafer surface; and

5. Cleaning, to remove the polishing agent or dust particles from the wafer surface.

1.2. Slicing silicon ingots into wafers: ID sawing versus wire sawing

Until recently, internal-diameter (ID) sawing had been the dominant slicing method for the past three decades [9–10]. However, wire sawing is now fully established as the preferred method of slicing large diameter ingots. Due to its thinner kerf losses, wire sawing yields more wafers per unit length of crystal ingot than ID sawing. Although the time per cut in wire-sawing is much longer than in ID sawing, the overall throughput of wire-sawing is not inferior to ID sawing, because wire sawing can slice hundreds of wafers per cut while ID sawing cuts one wafer at a time.
A phenomenon associated with wire sawing is the waviness. The wire-sawing induced waviness is also called long cycle swelling or unevenness, or wavy stripes [11]. It has wavelength of 0.5–30 mm [12]. Fig. 1 shows a magic mirror picture [13–15] of a wafer exhibiting waviness.

The generating mechanism of this waviness is not yet fully understood. This has been the main reason that it is very difficult to eliminate waviness at wire-sawing process. If subsequent processes do not remove the waviness, it will adversely affect the flatness, especially the site flatness, of the wafers.

Waviness is not a problem for ID sawing. With modern ID sawing technology, the back of the wafer is ground before the wafer is sliced off from the ingot, providing a reference plane for subsequent processes.

1.3. Flattening wire-sawn wafers: lapping versus grinding

Conventionally, wafers after slicing will go through lapping operations for flattening. In a typical double side lapping operation, a batch of wafers (for example, 20 wafers) is manually loaded into a lapping machine. Aluminum-oxide slurry is injected between two metal plates that rotate in opposite directions [16].

Fig. 2 illustrates the wafer grinding process. Gridding wheels are diamond cup wheels. The workpiece (wafer) is held on a porous ceramic chuck by means of vacuum. The rotation axis for the grinding wheel is offset by a distance of the wheel radius relative to the rotation axis for the wafer. During grinding, the grinding wheel and the wafer rotate about their own rotation axes simultaneously, and the wheel is fed towards the wafer along its axis.

The advantages of grinding over lapping include: (1) fully automatic with cassette-to-cassette operation; (2) use of fixed-abrasive grinding wheel rather than loose abrasive slurry so the cost of consumables per wafer may be lower; and (3) higher throughput. For instance, it takes about 30–60 sec to reduce the wafer thickness (200 mm or 300 mm in diameter) by about 45–75 µm [8].

It has been reported that lapping can effectively remove waviness but grinding cannot [11–12]. The explanation for grinding’s ineffectiveness has been that, in grinding, the wafer is held onto a ceramic chuck by high-pressure vacuum and thus elastically deforms to conform to the chuck surface. The wafer will return to its original shape once the vacuum is released. A recent study [17] has shown that, under typical manufacturing conditions, the deformation of waviness peaks caused by the vacuum is less than one-tenth of that by grinding force. Therefore, it is the grinding force, not the vacuum, that causes the wafer to deform elastically and hence prevents eliminating waviness.

Fig. 3 shows magic mirror images of two silicon wafers processed by lapping and grinding respectively. Lapping cuts both sides of the wafer simultaneously while grinding cuts one side at a time. The total removal amount is the same for lapping and grinding. The polishing operation following lapping and grinding is to provide a shining surface required for magic mirror
inspection. The same amount of removal is polished off for both cases. It is obvious that lapping is very effective in eliminating the waviness while surface grinding is not. The difference between lapping and grinding in removing waviness has been made readily visible by the results of a finite element analysis (Fig. 4). It has been shown that, when the same total force is applied on a wafer, the elastic deformation of the wafer (measured by the displacement at the peaks of waviness) in lapping is much smaller than that in grinding. The amount of deformation in lapping is only 1/55 to 1/36 of that in grading [17].

1.4. Outline of this paper

Wafer grinding is potentially a more cost-effective process than lapping for flattening wire-sawn silicon wafers. However, its success in replacing lapping will depend on whether several technical barriers can be removed. One such barrier is the wire-sawing induced waviness. Several approaches to overcome this barrier have been proposed, including wafer grinding followed by a lapping [8], wax mounting [12], reduced vacuum [12] and use of “soft-pad” [18]. These approaches are summarized and discussed in [19].
This paper presents the results of a finite element analysis (FEA) into grinding of wire-sawn silicon wafers. In this investigation, a four-factor two-level full factorial design is employed to reveal the main effects as well as the interaction effects of four factors (wafer thickness, waviness wavelength, waviness height and grinding force). The output variable studied is the displacement of the waviness peaks. The results of this study have provided several implications to silicon wafer manufacturing from the viewpoint of reducing wire-sawing induced waviness.

There are four sections in this paper. Following this introduction section, section 2 describes the FEA model and the design of experiments. In section 3, the results of the finite element analysis will be presented and discussed. Finally, conclusions are drawn in section 4.

2. Finite element analysis model and design of experiments

2.1. Finite element analysis model

Commercial software, ANSYS, is used for this study. Critical issues encountered in the study include whether to treat the grinding process as a dynamic or static problem, what type of element to use, how to handle contact between the wafer and the supporting chuck, and how to simulate loading from the grinding wheel.

Grinding in actual production is a dynamic process. However, the focus of this study is the effectiveness of waviness removal, for which the wafer deformation under the impressing grinding wheel is presumably more important than the dynamic effect. Therefore, the grinding process is modeled as a static problem.

Since the thickness dimension of the wafer is much smaller than the diameter dimension, shell element, instead of 3D solid element, is used. This choice is capable of capturing the effect of in-plane loading and bending. It also avoids an excessively large number of degrees of freedom (DOF) and reduces computational cost.

Contact treatment in implicit FEA is usually rather time consuming, and may be difficult to converge if one or all contacting parts are not constrained properly. In this study, since the initial contact points between the wafer and the supporting chuck are known (the valleys of the waviness) and the range of deformation is small, DOF constraint is used to imitate the contact rather than using the contact algorithm in ANSYS. For the approach to be valid, all nodal displacements are monitored to ensure that none of the unconstrained nodes move below the chuck surface. This DOF constraint approach significantly reduces the computational cost.

Similarly, even though loading of the grinding wheel actually comes from contact, it is imitated by forces acting on nodes at the waviness peaks underneath the active grinding zone (the nominal contact area between grinding wheel and wafer, as shown in Fig. 2). The forces are distributed in such a manner that all peaks under the active zone are deformed to the same height.

From Figs. 1 and 3, it can be seen that the strips of wire-sawing induced waviness have approximately the same orientation but the wavelength and height change

Table 1

<table>
<thead>
<tr>
<th>Test</th>
<th>Wafer thickness</th>
<th>Waviness wavelength</th>
<th>Waviness height</th>
<th>Grinding force</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>7</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>9</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>13</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Fig. 5. One-fourth of FEA model for wafer grinding.

Table 2

<table>
<thead>
<tr>
<th>Factor</th>
<th>Unit</th>
<th>Low Level (–)</th>
<th>High Level (+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wafer thickness</td>
<td>μm</td>
<td>700</td>
<td>1000</td>
</tr>
<tr>
<td>Waviness wavelength</td>
<td>mm</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Waviness height</td>
<td>μm</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>Grinding force</td>
<td>N</td>
<td>80</td>
<td>200</td>
</tr>
</tbody>
</table>
irregularly. In this study, the FEA model is built with
the waviness profile simplified as sinusoids with uniform
wavelength and height, as shown in Fig. 5.

To generate the mesh, sinusoids with total length
longer than the diameter of wafer is extruded into a rec-
tangle wavy area along the depth direction. This wavy
area is then intersected with a cylinder with a diameter
of 200 mm to form a 3D wafer disk. The model is
meshed into Shell 63 triangular elements that have both
bending and membranes capabilities. For a typical mesh
of wafer there are 9330 elements and 4780 nodes. Elastic
modulus of silicon wafer is 135 GPa, and the Poisson's
ratio 0.3.

As shown in Fig. 5, the boundary conditions include:

1. DOF Constraints: All nodes at the valleys of waviness
are constrained from moving in \( Y \) direction (along
wafer thickness dimension) to simulate the support
from the ceramic chuck.
2. Forces: Grinding force is loaded at the peak nodes in
\( Y \) direction within the active grinding zone. Since the
grinding wheel is regarded as a rigid object, the
degree of freedom in \( Y \) direction of the loaded nodes
in the active zone has been constrained such that they
always move to the same height and the adjacent
nodes are always below the loaded nodes.
3. Pressure: The pressure is applied on the upper surface
of wafer to simulate the effect of vacuum.

2.2. Design of experiments

A \( 2^4 \) (four factors, two levels, 16 tests) full factorial
design is used for the experiments. Detailed description
of factorial design can be found in many textbooks such
as the one by DeVor et al. [20]. The experiment matrix
is shown in Table 1 and the factor levels are listed in
Table 2. The output variable is the displacement of the
peak nodes under active grinding zone.

3. Results and discussion

The test data are shown in Table 3. The software
called Design-Expert (Version 5, Stat-Ease Corporation,
Minneapolis, MN) is used to process the data. The sig-
nificant effects identified by the analysis are: wafer
thickness (A), waviness wavelength (B), grinding force
(D), two-factor interaction of wafer thickness and wav-
iness wavelength (AB), two-factor interaction of wafer
thickness and grinding force (AD), two-factor interaction
of waviness wavelength and grinding force (BD), and
the three-factor interaction of wafer thickness, waviness
wavelength and grinding force (ABD). Table 4 is the
ANOVA (analysis of variance) table [21] for the fac-
torial design.

3.1. Effects of wafer thickness

The final thickness is between 700–800 \( \mu m \) for silicon
wafers of 200 and 300 mm in diameter [4,22]. The wafer
thickness after the slicing operation is much thicker than the final thickness since allowance has to be made for subsequent processes.

The main effect of wafer thickness on the displacement of waviness peaks is shown in Fig. 6. It can be seen that when the wafer thickness increases, the displacement is reduced. In other words, it is easier to remove the wire-sawing induced waviness when the wafer is thicker. One practical implication of this result is discussed below.

In section 1, it was pointed out that surface grinding has many advantages over lapping but cannot effectively eliminate the waviness induced by wire-sawing operation. However, the lapping process can effectively remove the waviness. To utilize the benefits of both grinding and lapping, one approach (that has been proposed earlier) is to use surface grinding followed by lapping (as shown in Fig. 7) [8]. Grinding removes the majority of the removal amount and the task of eliminating waviness is left to the subsequent lapping operation. According to the result of this finite element analysis (the thicker the wafer, the easier to remove the waviness), putting the lapping operation prior to grinding would be more effective in removing the waviness.

The interactions of wafer thickness with other factors on the displacement of waviness peaks are shown in Fig. 8. The two-factor interactions of wafer thickness with waviness wavelength and grinding force are significant. At the high level of wavelength, the change in wafer thickness causes a larger change in the displacement of waviness peaks than at the low level of wavelength. Similarly, at the high level of grinding force, the change in wafer thickness causes a larger change in the displacement of waviness peaks than at the low level of grinding force.

Fig. 6. Main effect of wafer thickness.

Fig. 7. Lapping to remove waviness [after Vandamme, Xin and Pei, 2000].

Fig. 8. Interactions of wafer thickness and other factors.
in wafer thickness causes a larger change in the displacement of waviness peaks than at the low level of grinding force. There is no interaction effect between wafer thickness and waviness height.

3.2. Effects of waviness wavelength

Fig. 9 illustrates the variation of waviness peak displacement with waviness wavelength. With the augmentation of waviness wavelength, the peak displacement increases. Based on this result, waviness with a shorter wavelength is preferred to a longer wavelength. Therefore, it is recommended that the wire-sawing operation should produce wafers with a shorter waviness wavelength if possible.

The interactions of waviness wavelength with other factors on the displacement of waviness peaks are shown in Fig. 10. The two-factor interactions of waviness wavelength with wafer thickness and grinding force are significant. At the low level of wafer thickness, the change in waviness wavelength causes a larger change in the displacement of waviness peaks than at the high level of wafer thickness. For the interaction of waviness wavelength and grinding force, at the high level of grinding force, the change in waviness wavelength causes a larger change in the displacement of waviness peaks than at the low level of grinding force. There is no interaction effect between waviness wavelength and waviness height.

3.3. Effects of waviness height

The main effect and interaction effects of waviness height are shown in Figs. 11 and 12. It can be seen that none of these effects is significant. Within the range studied, the waviness height has little effect on the displacement of waviness peaks. Note that this does not suggest that wafers with high-amplitude waviness can be removed as effectively as wafers with low-amplitude waviness, because the output variable here is the displacement of waviness peaks.

3.4. Effects of grinding force

The main effect of grinding force on the displacement of waviness peaks is shown in Fig. 13. It can be seen that as the grinding force increases, the peak displacement is increased. In other words, it is easier to remove the wire-sawing induced waviness when the grinding force is lower. One practical implication of this result is illustrated below.
Typically, the grinding of silicon wafers is divided into two stages: coarse grinding and fine grinding. For coarse grinding, a grinding wheel with larger diamond grains is employed to remove the majority of the material at a higher feedrate. For fine grinding, a wheel with smaller diamond grains is used to obtain controlled surface roughness and subsurface damage. Normally, the grinding force in coarse grinding is much lower than that in fine grinding. This is consistent with grinding of other brittle materials [23–24]. Therefore, from the viewpoint of reducing waviness, only coarse grinding should be performed.

Consider the following two process arrangements for grinding wire-sawn silicon wafers. (1) Coarse grind and fine grind on one side, then flip the wafer to coarse grind and fine grind the other side. (2) Coarse grind the first side, then flip the wafer to coarse grind and fine grind the second side; and then flip the wafer again to fine grind the first side. From the preceding discussion, the second arrangement will be more effective in removing the wire-sawing induced waviness. This is illustrated in Fig. 14.

The interactions of grinding force with other factors on the displacement of waviness peaks are shown in Fig. 15. The two-factor interactions of grinding force with wafer thickness and waviness wavelength are significant. At the low level of wafer thickness, the change in grinding force causes a larger change in the displacement of waviness peaks than at the high level of wafer thickness. At the high level of wavelength, the change in grinding force causes a larger change in the displacement of waviness peaks than at the low level of wavelength. There is no interaction effect between grinding force and waviness height.

3.5. Three-factor interaction

The three-factor interaction of wafer thickness, waviness wavelength and grinding force is significant. As shown in Fig. 16, the best scenario for removing waviness is the combination of thicker wafer, shorter wavelength and lower grinding force. The worst case occurs with thinner wafer, longer wavelength and higher grinding force.

4. Conclusions

Surface grinding is potentially a more cost-effective process than lapping for flattening wire-sawn silicon wafers. However, its success in replacing lapping will depend on whether the wire-sawing induced waviness...
can be effectively removed. Four-factor two-level full factorial design is used to conduct a finite element analysis study into grinding of wire-sawn silicon wafers. The main effects and the two-factor interactions of these four factors (wheel speed, chuck speed and feedrate) on the effectiveness of removing the wire-sawing induced waviness are obtained.

The following conclusions can be drawn from this study:

1. The main effects of wafer thickness, waviness wavelength and grinding force are significant. It becomes more difficult to remove the waviness as the wafer thickness decreases, or as waviness wavelength increases, or as grinding force increases.
2. The two-factor interactions between wafer thickness, waviness wavelength and grinding force are also significant. The effect of wafer thickness is greatly enhanced at the high level of waviness wavelength and at the high level of grinding force. Also, the effect of waviness wavelength at the high level of grinding force is larger than that at the low level of grinding force.
3. Waviness height does not show any significant effects at either main effect or two-factor interactions.
4. The three-factor interaction of wafer thickness, waviness wavelength and grinding force is significant. The best scenario for removing waviness is the combination of thicker wafer, shorter wavelength and lower grinding force. The worst case occurs with thinner wafer, longer wavelength and higher grinding force.
Fig. 16. Three-factor interactions of wafer thickness, waviness wavelength and grinding force.

Acknowledgements

Financial support for this work was provided by the Advanced Manufacturing Institute at Kansas State University.

References