Spring 2018

In-situ Measurement of Plasma Etch Rate Using EtchRate Wafer

Razieh Mahzoon
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Mahzoon, Razieh, "In-situ Measurement of Plasma Etch Rate Using EtchRate Wafer" (2018). *Master's Theses*. 4913.

DOI: <https://doi.org/10.31979/etd.7898-qekp>

https://scholarworks.sjsu.edu/etd_theses/4913

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

IN-SITU MEASUREMENT OF PLASMA ETCH RATE USING ETCHRATE WAFER

A Thesis

Presented to

The Faculty of the Department of Physics and Astronomy

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Razieh Mahzoon

May 2018

© 2018

Razieh Mahzoon

ALL RIGHTS RESERVED

The Designed Thesis Committee Approves the Thesis Titled

IN-SITU MEASUREMENT OF PLASMA ETCH RATE USING ETCHRATE WAFER

by

Razieh Mahzoon

APPROVED FOR THE DEPARTMENT OF PHYSICS AND ASTRONOMY

SAN JOSÉ STATE UNIVERSITY

May 2018

Peter Beyersdorf, Ph.D.

Department of Physics

Ramen Bahuguna, Ph.D.

Department of Physics

Kenneth Wharton, Ph.D.

Department of Physics

ABSTRACT

IN-SITU MEASUREMENT OF PLASMA ETCH RATE USING ETCHRATE WAFER

by Razieh Mahzoon

This paper presents the "EtchRate Wafer," a wireless monitoring silicon wafer that was invented by KLA-Tencor Corporation and patented in September 2014. This monitoring wafer measures the film thickness change in the etching plasma chambers while the etching process is running. The monitoring wafer has three sensors to measure the uniformity of thickness change across the wafer. The sensor is an optical stack that mainly consists of a collimator, an optical window, coated mirror, optical filter, and an optical detector. To measure the etch rate, the specific target material is coated on the optical window to be able to simulate the same test as in the plasma chamber. By placing the monitoring wafer instead of the regular target wafer in the chamber, the etch rate is measured by the use of the interference phenomenon. The transmitted light from the material layer produces an interference pattern on the optical detector, which can be analyzed to compute the layer thickness. Monitoring the changes in the interference pattern in the etching or deposition processes can help to measure the film layer thickness as a function of time. This thesis reviews the semiconductor fabrication processes include plasma etching processes and the interference phenomenon that is the basic concept of EtchRate sensor function. The EtchRate wafer and its sensors are described in detail and two methods for mathematical analysis are discussed. The Zemax modeling for the EtchRate sensor is also presented. In conclusion, it is possible to measure the plasma etch rate in situ by EtchRate wafer which is a convenient and time-saving method.

DEDICATION

This thesis is dedicated to my inspiring parents for their endless love, support, and encouragement and to my husband, Yusuf, for his kindness and devotion, and for his continuous support.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF FIGURES IN APPENDICES	xii
1. INTERFERENCE	1
1.1 Intensity Distribution of Interference	1
1.2 Interference Circumstances	4
1.2.1 Spatial coherence	6
1.2.2 Temporal coherence	9
1.2.3 Perfect incoherent/coherent light source	13
1.3 Different Kind of Interferometers	15
1.3.1 Off-axis interferometers	15
1.3.2 On-axis interferometers	15
2. SEMICONDUCTOR INDUSTRIES	22
2.1 Semiconductor Fabrication Processes	24
2.1.1 Electrical chip design and physical layout	25
2.1.2 Oxidation	25
2.1.3 Lithography	25
2.1.4 Etching	26
2.1.5 Film deposition	26
2.1.6 Chemical mechanical planarization	27
2.1.7 Doping of semiconductors	27
2.1.8 Ion implantation	28
2.1.9 Diffusion	28
3. ETCHING THEORY	30
3.1 The Terminologies in the Etching Processes	30
3.1.1 Etch rate	30
3.1.2 Uniformity	31
3.1.3 Selectivity	31
3.1.4 Etch profile	31
3.1.5 Damage	32
3.2 Etching Techniques	32
3.2.1 Wet etching method	33
3.2.1.1 Advantages of wet etching	34
3.2.1.2 Disadvantages of wet etching	34
3.2.2 Dry etching methods	34
3.2.2.1 Plasma etching	35
3.2.2.2 Reactive ion etching	35

3.2.2.3	Ion beam	35
3.2.2.4	Beam methods	35
3.3	Plasma Etching Principles	36
3.3.1	Plasma environment	36
3.3.1.1	DC Glow discharge regions	39
3.3.1.2	DC vs. RF plasmas	40
3.3.2	Plasma etching process	40
3.3.3	Chemical reactions in plasma etching	44
3.3.4	Plasma optical emission spectra	45
4.	ETCHRATE WAFER.....	48
4.1	Etch Monitoring Wafers	48
4.2	KLA-Tencor Wireless EtchRate Wafer	50
4.2.1	Etch rate sensor structure	51
4.2.1.1	Photodetector	53
4.2.1.2	Optical wedge or slab	54
4.2.1.3	Etch rate interferometer mirror	55
4.2.1.4	Band-pass filter	55
4.2.1.5	Collimator	55
4.2.1.6	Aperture	55
4.2.1.7	Thin film layers	56
4.2.2	EtchRate wafer working mechanism	56
4.2.2.1	Data collection	56
4.2.2.2	Contrast in EtchRate sensor	57
4.2.2.3	Data interpretation in the wedge design	58
4.2.2.3.1	Measuring cross-correlation	61
4.2.2.3.2	Measuring thickness by use of phase shift	63
4.2.2.4.	Wedge surface quality	64
4.2.3	Difference between wedge and slab	64
	SUMMARY	66
	REFERENCES	67
	APPENDICES	69
A.	Mathematical Analysis.....	69
A.1	Cross-correlation.....	70
A.2	Error Function	72
B.	Zemax Simulation	75

LIST OF TABLES

Table		Page
1.	Constructive and Destructive Interference Equations for Different Phase Shifts	17
2.	The Kind of Gas Plasma Required for Etching Each Material	45

LIST OF FIGURES

Figures	Page
1. Superposition of waves: a) Constructive interference, b) Destructive interference.	1
2. a) Interference of two optical waves. The resulting pattern on screen illustrates the black and white fringes. b) The fringes on the screen correspond to destructive (dark) and constructive (bright) interference.	4
3. Interference setup and fringes pattern in Young experiment.	6
4. Schematic Diagram of Young interference setup.	7
5. Michelson interferometer setup (“Michelson Interferometer”, n.d.).	10
6. Splitting the measurement arm and the reference arm by beamsplitter in Michelson interferometer.	12
7. Fringes pattern in Michelson interferometer.	13
8. The intensity of interference wave for perfectly coherent waves sinusoidally changes as a function of phase difference.	14
9. The intensity of interference wave for perfectly incoherent waves is constant as a function of phase difference.	15
10. Schematic of on-axis interferometer.	16
11. Reflection from a medium of higher index of refraction causes 180-degree phase change.	16
12. Direction of the electric field and wave vector for reflected and refracted beams in the interface.	18
13. Cross-section of the incident, reflected, and transmitted beams.	20
14. Czochralski silicon crystal growth.	23
15. Single crystal silicon ingot. b) Silicon ingot sliced to Si wafers.	23
16. Silicon wafer in different standard sizes.	24

17.	Copper interconnects in an IBM chip.	29
18.	Schematic of the etching process.	30
19.	A bad selective etch (a) vs. highly selective etch (b).	31
20.	Different etch profiles.	32
21.	A wet etching device consists of a bath container with a stirrer, controllable hot plate, and slits to hold the wafers.	33
22.	a) A cross-sectional view of a typical plasma chamber. b) Plasma chamber window.	36
23.	Schematics of a plasma chamber: A free electron is released from a neutral molecule by its interaction with a photon.	37
24.	Schematics of a plasma chamber: Acceleration of free electrons toward the anode.....	37
25.	Schematics of a plasma chamber: Electrons collide with the gas molecules in elastic and inelastic collisions.	38
26.	Schematics of a plasma chamber: All generated electrons are accelerated toward the anode, and making breakdown.	38
27.	Schematics of a plasma chamber: Emitting secondary electrons by the collision of positive charges with the cathode.	39
28.	Schematics of a plasma chamber: Current distribution in a DC glow discharge.	40
29.	Blocking mechanism of etching.	42
30.	A cross-section of 90 nm generation Cu interconnects. Light gray regions in all layers are Cu interconnects. Dark regions of MT 1-7 are low-k dielectric material; MT 8 is SiO ₂ , and MT 9 is polymer.	43
31.	Optical emission spectra lines for Ar plasma in the range 690–850 nm.	46
32.	An example of plasma emission lines.	47
33.	Relative irradiance in a plasma environment for two different recipes.	47

34.	Fabrication of memory chips on a silicon wafer.	48
35.	EtchRate sensor with slab window design.	52
36.	EtchRate sensor with wedge window design.	52
37.	A typical linear optical array.	53
38.	Optical response plot of a typical linear array detector.	53
39.	Optical transmission of BK7, UV fused silica and quartz (Optical materials, n.d.).	54
40.	Signal output. Line A represents a reference waveform (Bias); Lines B and C represent measured waveforms (O'Brien & Jensen, 2014).	57
41.	The incident beam and the reference beam paths passing through a wedge.	58
42.	The first fringe is destructive interference.	59
43.	Movement of interference wave by changing the thickness.	59
44.	The first maxima of reference wave (Black) and measurement wave (Brown) are chosen to measure the cross-correlation.	60
45.	Cross-correlation plot: F and G have phase shift of π	62
46.	Cross-correlation plot: F and G have arbitrary phase shift.	62
47.	Cross-correlation plot: F and G have phase shift of zero.	62

LIST OF FIGURES IN APPENDICES

Figures	Page
A1. Simulation of reference and measurement waves in 128-pixel linear array detector.	69
A2. Calculation of correlation function for two waves- Equations.	70
A3. Calculation of correlation function for two waves- Numerical values...	71
A4. Plotting correlation function, reference and measurement waves.	71
A5. Calculation of error function for two waves- Equations.	72
A6. Calculation of error function for two waves- Numerical values.	73
A7. The value of error goes to zero for offset value of -7.	73
A8. Solver parameters to find the offset value.	74
B1. Schematic for a 300nm coated wedge with the slope of 3 μm in 10mm. The scale is 10:1 and all units are in mm.	75
B2. Schematic for a non-coated wedge with the slope of 3 μm in 10mm. The scale is 10:1 and all units are in mm.	76
B3. Screenshot for the Interference Modeling in Zemax Software	77
B4. A simple modeling for EtchRate sensor in Zemax.	78
B5. The light paths are visible in the Zemax layout configuration.	78
B6. The zoom view of Zemax layout configuration clearly shows the split of reflected and transmitted rays for both surfaces.	79
B7. The interference pattern for the non-coated wedge (left) and 300nm coated wedge (right)	79
B8. The interference pattern across the y-axis for the non-coated wedge (The measurement beam-left) and 300nm coated wedge (The reference beam-right).	80

B9.	The interference patterns for simulated reference and measurement waves.	80
-----	---	----

CHAPTER 1

Interference

Interference is an optical effect, which occurs when two or more waves overlap. Interference is based on the superposition principle, which is adding wave functions mathematically. Figure 1 illustrates the superposition of two waves in both constructive and destructive interference.

Figure 1. Superposition of waves: a) Constructive interference, b) Destructive interference.

1.1 Intensity Distribution of Interference

For two monochromatic planar waves given as:

$$\vec{E}_1 = \vec{E}_{01} e^{i(\vec{k}_1 \cdot \vec{r} - \omega t + \theta_1)} \quad (1)$$

$$\vec{E}_2 = \vec{E}_{02} e^{i(\vec{k}_2 \cdot \vec{r} - \omega t + \theta_2)} \quad (2)$$

where E_{01} and E_{02} are the amplitude of two electric fields, ω is the optical frequency of waves which is equal for both waves, and k_1 and k_2 are wave numbers ($k=n\omega/c$).

$k \cdot r - \omega t + \theta$ is the term for the phase of the electric field. θ_1 and θ_2 are additional phases added to the electric field to keep the solution general. (Hecht, 2017, p. 402)

E total, the resulting sum of the two waves, is:

$$\vec{E}_{total} = \vec{E}_1 + \vec{E}_2 = [\vec{E}_{01}e^{i(\vec{k}_1 \cdot \vec{r} + \theta_1)} + \vec{E}_{02}e^{i(\vec{k}_2 \cdot \vec{r} + \theta_2)}]e^{-i\omega t} \quad (3)$$

Most light detecting instruments measure time averaged light intensity. The intensity for an electric field can be calculated by the ‘square-law’:

$$I_{total} = \frac{1}{2}\epsilon_0 c \langle |\vec{E}|^2 \rangle \quad (4)$$

$$I_{total} = \frac{1}{2}\epsilon_0 c \langle |\vec{E}|^2 \rangle = \frac{1}{2}\epsilon_0 c \langle \vec{E} \cdot \vec{E}^* \rangle \quad (5)$$

$$= \frac{1}{2}\epsilon_0 c \langle (\vec{E}_1 + \vec{E}_2) \cdot (\vec{E}_1^* + \vec{E}_2^*) \rangle \quad (6)$$

$$= \frac{1}{2}\epsilon_0 c (\langle |\vec{E}_1|^2 \rangle + \langle |\vec{E}_2|^2 \rangle + \langle \vec{E}_1 \cdot \vec{E}_2^* \rangle + \langle \vec{E}_2 \cdot \vec{E}_1^* \rangle) \quad (7)$$

$$\langle \vec{E}_1 \cdot \vec{E}_2^* \rangle = \langle \vec{E}_{01} \cdot \vec{E}_{02}^* \rangle e^{i(\vec{k}_1 \cdot \vec{r} - \omega t + \theta_1) - i(\vec{k}_2 \cdot \vec{r} - \omega t + \theta_2)} \quad (8)$$

$$= \langle \vec{E}_{01} \cdot \vec{E}_{02}^* \rangle e^{i((\vec{k}_1 - \vec{k}_2) \cdot \vec{r} + (\theta_1 - \theta_2))} \quad (9)$$

$$\langle \vec{E}_2 \cdot \vec{E}_1^* \rangle = \langle \vec{E}_{02} \cdot \vec{E}_{01}^* \rangle e^{i(\vec{k}_2 \cdot \vec{r} - \omega t + \theta_2) - i(\vec{k}_1 \cdot \vec{r} - \omega t + \theta_1)} \quad (10)$$

$$= \langle \vec{E}_{02} \cdot \vec{E}_{01}^* \rangle e^{-i((\vec{k}_1 - \vec{k}_2) \cdot \vec{r} + (\theta_1 - \theta_2))} \quad (11)$$

And,

$$\frac{1}{2}\epsilon_0 c \langle |\vec{E}_1|^2 \rangle = I_1 \quad \text{and} \quad \frac{1}{2}\epsilon_0 c \langle |\vec{E}_2|^2 \rangle = I_2 \quad (12)$$

So,

$$I_{total} = I_1 + I_2 + \sqrt{I_1 I_2} e^{i((\vec{k}_1 - \vec{k}_2) \cdot \vec{r} + (\theta_1 - \theta_2))} + \sqrt{I_2 I_1} e^{-i((\vec{k}_1 - \vec{k}_2) \cdot \vec{r} + (\theta_1 - \theta_2))} \quad (13)$$

$$I_{total} = I_1 + I_2 + \sqrt{I_1 I_2} (e^{i((\vec{k}_1 - \vec{k}_2) \cdot \vec{r} + (\theta_1 - \theta_2))} + e^{-i((\vec{k}_1 - \vec{k}_2) \cdot \vec{r} + (\theta_1 - \theta_2))}) \quad (14)$$

in which $\Delta\phi = (\vec{k}_1 - \vec{k}_2) \cdot \vec{r} + (\theta_1 - \theta_2)$ is a phase difference arising from a combined path length and initial phase-angle difference.

Since $\cos \Delta\phi = \frac{1}{2}(e^{i\Delta\phi} + e^{-i\Delta\phi})$, then:

$$I_{total} = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \Delta\phi \quad (15)$$

This equation is called the interference equation (Saleh & Teich, 2007).

Therefore, the intensity of the resulting interference wave in position $\vec{r}$ is the addition of the intensities of two waves plus an interference term, which oscillates with the phase difference between the two waves.

In Equation 15, the maximum intensity is when the cosine function is maximum (+1) which it shows constructive fringes.

$$\Delta\phi = 0, \pm 2\pi, \pm 4\pi, \dots \quad (16)$$

For destructive interference, the minimum intensity occurs when the cosine function is -1:

$$\Delta\phi = \pm\pi, \pm 3\pi, \pm 5\pi, \dots \quad (17)$$

In constructive interference, the amplitude of the resulting wave is greater than the individual ones, and reversely in destructive regions, the amplitude of interference wave is smaller than two waves. Figure 2 shows the interference pattern of two optical point source. Black and white fringes correspond to destructive and constructive interference respectively.

Figure 2. a) Interference of two optical waves. The resulting pattern on screen illustrates the black and white fringes. b) The fringes on the screen correspond to destructive (dark) and constructive (bright) interference.

The “optical path difference” of two waves, δ , is defined as the difference in distance that the waves have traveled. An optical path difference of λ corresponds to a phase difference of 2π , so the relation below does exist:

$$\frac{\delta}{\lambda} = \frac{\Delta\phi}{2\pi} \quad (18)$$

So,

$$\Delta\phi = \frac{2\pi}{\lambda} \delta \quad (19)$$

1.2 Interference Circumstances

The interference phenomenon can be observed when these conditions have been met:

- Two interfering waves have similar polarizations. Otherwise, the waves’ electric fields are in different planes, and don’t interfere.

- The waves are monochromatic, which means they are at a single wavelength.
- Two waves have some degree of coherence, which means that their phase difference has been constant. For example, if the phase difference between two waves is ‘ π ’, this phase difference should be constant over the time.

The degree of coherence can be measured by “spectral visibility of the fringes (contrast)” which is calculated by luminance difference divided by the average of luminance:

$$V = \frac{I_{Max} - I_{min}}{I_{Max} + I_{min}} \quad (20)$$

The contrast is maximum when two waves’ maxima interfere at the same position, and the fringe contrast will disappear when the maxima of one wave coincides with the minima of the second one. In other words, for a good contrast, the difference in intensity between a light and dark line should be large.

Coherence length is defined by the maximum path difference along two beams of light such that they generate an interference pattern when recombined. The coherence time is the elapsed time interval between the passage of these two waves that can interfere. By exceeding the coherent length of two sources, the contrast of their interference drops significantly. The coherence length (L) must be longer than the path length of the interferometer. Coherence length is:

$$L = \frac{\lambda^2}{\Delta\lambda} \frac{2 \ln 2}{\pi n} \quad (21)$$

where λ is the wavelength of beams, $\Delta\lambda$ is the bandwidth of the source, and n is the refractive index of the medium.

1.2.1 Spatial coherence. In Young interferometer, spatial coherence between two waves is the reason of interference. Since wavefronts are dividing, this setup is called wavefront-splitting interferometer (Hecht, 2017).

A monochromatic light source is emitting on a screen with double slits. As shown in Figure 3 two waves are passing from the identical and very closely spaced slits and then interfere at the right side of the screen. In all regions that two wavefronts have overlay, the interference pattern is produced. The dark fringes show the minimum intensity of the interference wave, which correspond to destructive interference; the bright fringes observed on the screen correspond to constructive interference, which has the maximum intensity on the screen.

Figure 3. Interference setup and fringes pattern in Young experiment.

Figure 4 shows the schematic diagram for Young experiment. The slits are named S_1 and S_2 , and the waves are combining on point P, which is in distance y from the perpendicular plane passing the middle of slits.

Figure 4. Schematic Diagram of Young interference setup.

r_1 and r_2 are the distance of point P from slits 1 and 2 respectively. As the path difference $r_2 - r_1$ increases, the degree of correlation between two waves decreases and therefore the contrast in the interference pattern drops.

In constructive interference, the path difference is zero or an integer, m , multiple of the wavelength. Therefore, for Young experiment the path difference can be written as:

$$r_2 - r_1 = m\lambda \quad (22)$$

$$r_1^2 = r^2 + (d/2)^2 - dr \cos((\pi/2) - \theta) = r^2 + (d/2)^2 - dr \sin \theta \quad (23)$$

$$r_2^2 = r^2 + (d/2)^2 - dr \cos((\pi/2) + \theta) = r^2 + (d/2)^2 + dr \sin \theta \quad (24)$$

$$r_2^2 - r_1^2 = (r_2 + r_1)(r_2 - r_1) = 2dr \sin \theta \quad (25)$$

Since $L \gg d$, therefore:

$$r_1 + r_2 \approx 2r \quad (26)$$

In addition:

$$r_2 - r_1 = 2dr \sin(\theta)/2r = d \sin \theta \quad (27)$$

So,

$$OPD = m\lambda \quad (28)$$

$$\delta = d \sin \theta = m\lambda \quad (29)$$

The first maximum of constructive interference fringe is when the point P is on the center of observation screen. So the angle between two waves would be zero and so does the path difference. Therefore, on the right side of equation “m” which is called ‘order number’ is zero (the zeroth-order) for the first central fringe. By moving farther from the center spot, each bright ring shows another order.

$$\delta = d \sin \theta = m\lambda \quad \text{for } m = 0, \pm 1, \pm 2, \pm 3, \dots \quad \text{for constructive interference} \quad (30)$$

For destructive interference, which produces dark fringes on the screen, the waves are 180 degrees (π) out of phase, which means half of wavelength in distance. Therefore, the interference equation, in this case, would be:

$$\delta = d \sin \theta = (m + 1/2)\lambda \quad \text{for } m = 0, \pm 1, \pm 2, \pm 3, \dots \quad \text{for destructive interference} \quad (31)$$

To find the location of each fringe, y on the screen, since the angle between two waves is so small $\sin \theta$ can be replaced by y/L in interference equations:

$$\sin \theta \cong \tan \theta = \frac{y}{L} \quad (32)$$

$$y_m = m \frac{\lambda L}{d}, \quad \text{for } m = 0, \pm 1, \pm 2, \pm 3, \dots \quad \text{for constructive interference} \quad (33)$$

$$y_m = (m + 1/2) \frac{\lambda L}{d}, \quad \text{for } m = 0, \pm 1, \pm 2, \pm 3, \dots \quad \text{for destructive interference} \quad (34)$$

Considering Equation 19 and 29, the optical phase different for Young experiment is:

$$\Delta\phi = \frac{2\pi}{\lambda} \delta = \frac{2\pi}{\lambda} d \sin \theta = \frac{2\pi}{\lambda} d \frac{y}{L} \quad (35)$$

The interference wave equation for Young interferometer would be:

$$I_Y = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos\left(\frac{2\pi}{\lambda} d \sin \theta\right) \quad (36)$$

Considering (32),

$$I_Y = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos\left(\frac{2\pi y}{\lambda L} d\right) \quad (37)$$

I_1 and I_2 are the Intensity of waves from two pinholes recombining. y is coordinate for the combining point in the screen and L is the distance between screen and pin holes.

By using this fact that in Young experiment two waves have come from the same source, so they have the same amplitude $E_0 = E_{01} = E_{02}$ and the same intensity I_0

Therefore, the Equation 37 can be written as:

$$I_Y = 2I_0 + 2I_0 \cos\left(\frac{2\pi y}{\lambda L} d\right) = 4I_0 \cos^2\left(\frac{\pi y d}{\lambda L}\right) \quad (38)$$

1.2.2 Temporal coherence. Michelson interferometer is an example of waves interference, in which the optical waves have temporal coherence (Figure 5). Temporal coherence describes the correlation between waves observed at different moments in time. Since in this setup the amplitude of the light is separated into two, it is called amplitude-splitting interferometer (Hecht, 2017)

Figure 5. Michelson interferometer setup (“Michelson Interferometer”, n.d.).

As the Figure 5 shows, the Michelson interferometer consists of a coherent monochromatic light source (laser), a 50/50 beamsplitter, two mirrors, and a screen. The beam emitted from the light source is hitting a 50/50 beamsplitter, which is oriented at the angle of 45° to the beam. It divides the incident beam into two equal-intensity beams at a 90° angle. Each identical transmitted wave travels toward a mirror; one mirror is fixed, and another is movable. The first transmitted beam travels into mirror M1, and it is reflected back to the beamsplitter. Again, by passing from beamsplitter the beam is split, and 50% of the returning beam is reflected toward the screen, and another half goes toward the light source. Similarly, the second beam reflects from mirror M2, passes beamsplitter for the second time, and divides into two perpendicular waves. One goes toward the screen, and another passes straight through the beamsplitter. Two beams on the screen interfere and make fringes. Whether they interfere constructively or

destructively, bright and dark fringes appear. The contrast is unity when the intensities of the interfering beams are equal.

In Michelson interferometer setup the beam which hits the fixed mirror is called ‘reference arm’, and the other beam which reflects from the movable mirror is named ‘measurement arm’. Since 50/50 beamsplitter is used the reference beam is at the same intensity as the measurement beam, and thus a maximum level of interference occurs. Since the two interfering beams of light were split from the same initial beam, they were initially in phase. If the two paths are equal, then the recombined waveforms at the screen are correlated, so they produce interference fringes. However, as the difference $L_2 - L_1$ increases, the degree of correlation decreases as well as the contrast in the interference pattern.

In Michelson experiment setup, unlike the Young experiment, the two beams are positioned to overlap completely at the same point on the screen. Since the optical path length is changed by twice the translation distance of the mirror M_2 , it can be written as:

$$OPD = 2L_1 - 2L_2 \quad (39)$$

In this setup in addition to the path difference, there is a time delay between two paths. An additional phase shift of π is introduced (more information in 1.3.2) because beam 1 is reflected off the inner side of the beamsplitter while beam 2 is reflected off the outer side (Hecht, 2017). Figure 6 illustrates the measurement arm and the reference arm splitting in Michelson interferometer.

Figure 6. Splitting the measurement arm and the reference arm by beamsplitter in Michelson interferometer.

So for Michelson interferometer, the Equation 19 can be written as,

$$\Delta\phi = \left(\frac{2\pi}{\lambda}\delta\right) - \pi = \left(\frac{2\pi}{\lambda}(2)(L_1 - L_2)\right) - \pi \quad (40)$$

If the optical phase difference ϕ of the two beams is an integer, multiple of 2π radians, then the light at all points will combine “in phase,” and the interference is constructive to make a “bright” fringe. If the optical phase difference is an odd-integer multiple π radians, then the light recombines “out of phase” to destructively interfere and make “dark” fringes.

Therefore, for constructive interference:

$$\Delta\phi = 2m\pi \quad \text{Where } m = 0, \pm 1, \pm 2, \pm 3, \dots \quad (41)$$

By substituting (40) in (41):

$$\left(\frac{2\pi}{\lambda}(2)(L_1 - L_2)\right) - \pi = 2m\pi \quad (42)$$

Therefore:

$$OPD = 2(L_1 - L_2) = \left(m + \frac{1}{2}\right)\lambda \quad (43)$$

In addition, for destructive interference:

$$\Delta\phi = \pm(2m + 1)\pi \quad \text{Where } m = 0, 1, 2, 3, \dots \quad (44)$$

By substituting (40) in (44):

$$\left(\frac{2\pi}{\lambda}(2)(L_1 - L_2)\right) - \pi = \pm(2m + 1)\pi \quad (45)$$

Therefore:

$$OPD = 2(L_1 - L_2) = m\lambda \quad (46)$$

In Michelson experiment, the maximum interference order is produced at the center of circular pattern (Figure 7).

Figure 7. Fringes pattern in Michelson interferometer.

If one of the beam paths is tilted compared to the other one, then the optical phase difference of the recombined beam linearly varies across the interference pattern. This fact converts the circular fringes to linear fringes look like the Young experiment.

Considering Equation 15, in Michelson interferometer the interference wave result is:

$$I_M = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos\left(\frac{2\pi}{\lambda}(2)(L_1 - L_2) - \pi\right) \quad (47)$$

$$I_M = 2I_0(1 + \cos\left(\frac{2\pi}{\lambda}(2)(L_1 - L_2) - \pi\right)) \quad (48)$$

Which I_1 and I_2 are equal I_0 are the intensity of two waves before combining.

1.2.3 Perfect incoherent/coherent light source. In a perfect coherence light source, which has the maximum contrast, two inference beams should have the same

intensity ($I_1 = I_2 = I_0$). Therefore, in Equation 38 the maximum intensity is $4I_0$, and the minimum intensity can be zero. Figure 8 demonstrates the interference wave intensity as a function of phase difference for two perfectly coherent waves.

Figure 8. The intensity of interference wave for perfectly coherent waves sinusoidally changes as a function of phase difference.

An incoherent light source emits the waves which do not maintain in constant phase relationship. In this kind of light source, the phase difference $\Delta\phi$ changes rapidly. Therefore, the interference fringes also move rapidly, and the interference term ($\cos \Delta\phi$) average out to zero. Therefore, in incoherent waves, no interference pattern is produced, and Equation 37 turns to:

$$I_{Total} = I_1 + I_2 \quad (49)$$

An incandescent lamp such as tungsten light bulb is a perfectly incoherent light source, which its emitted waves are out of phase and also in different wavelengths. Figure 9 shows the plot of intensity vs. phase difference for such an incandescent lamp.

Figure 9. The intensity of interference wave for perfectly incoherent waves is constant as a function of phase difference.

1.3 Different Kind of Interferometers

1.3.1 Off-axis interferometers. In most interference experiment setups, the two interfering beams are separated from each other and combine in the detector plane. In this type of interferometers, which is called “Off-axis” interferometers, the reference arm has a large angle with measurement arm (Ferraro, Wax, & Zalevsky, 2011, p.257). For example, in Michelson interferometer, two interference beams are perpendicular to each other, and in Young interferometer, the distance between the slits determines the angle between them. Therefore, the reference beam and the measurement beam are never on one axis.

1.3.2 On-axis interferometers. Another type of interferometer setup is the on-axis interferometer. On on-axis interferometers, reference arm and measurement arm are on one axis, or one is slightly (Fraction of a degree) beveled. In the on-axis interferometer, a slab of glass is used as a beamsplitter to make a reference arm and a measurement arm at the plane of detection (O’Brien & Jensen, 2014). In Figure 10, the interference beams are exaggeratedly shown. The incident beam hits to the slab and divides to the transmitted beam and the reflected beam on the inner surface of the slab.

The reflected beam reflects for the second time on the top surface of the slab. This reflected beam (Measurement beam-number 1) interferes with the first reference beam shown as number 2.

Figure 10. Schematic of on-axis interferometer.

The main issue with this kind of interferometer is that the reference and the measurement beams are not equal in magnitudes and hence the contrast of the interference is greatly reduced.

Summarizing the wave phase change at the medium interface (Figure 11), reflected light from a surface would experience a 180-degree phase change when it reflects from a medium of higher index of refraction and has no phase change when it reflects from a medium of the smaller index.

Figure 11. Reflection from a medium of higher index of refraction causes 180-degree phase change.

Therefore, as Table 1 states, the phase change in the mediums interface determines the equation for constructive and destructive interference.

Table 1

Constructive and Destructive Interference Equations for Different Phase Shifts

	Destructive	Constructive
Phase change= π	$\delta(OPD) = m\lambda$	$\delta(OPD) = (m + \frac{1}{2})\lambda$
Phase change=0	$\delta(OPD) = (m + \frac{1}{2})\lambda$	$\delta(OPD) = m\lambda$

Knowing that in on-axis interferometer (Figure 10) the measurement arm reflects twice in the wedge/slab, the phase change happens twice. Therefore, for constructive interference:

$$\delta = m\lambda \quad \text{for } m = 0, \pm 1, \pm 2, \pm 3, \dots \quad \text{for constructive interference} \quad (50)$$

Also because the measurement arm passes the wedge's/slab's thickness twice optical path difference (OPD) would be $2nt$; where n is the refractive index of air which is the environment both beams are met; and t is its thickness. So:

$$2nt = m\lambda \quad \text{for } m = 0, \pm 1, \pm 2, \pm 3, \dots \quad \text{for constructive interference} \quad (51)$$

$$2n\Delta t = \Delta m\lambda \quad \text{or} \quad \Delta t = \frac{\Delta m\lambda}{2n} \quad (52)$$

Therefore, in the on-axis interferometer, the distance between two peaks ($\Delta m=1$) is related to material thickness change which is $\frac{1}{2}$ wavelength.

As described in section 1.2.3, to have the maximum contrast the intensity of two interference beams should be equal to each other. In on-axis interference, the intensities of reference and measurement beam are determined by reflectance and transmittance coefficients calculated from Fresnel's equation. Fresnel equations describe the fraction of the reflected and refracted light from slab's interface. Figure 12 shows the schematic for reflected and refracted beams in the interface. The incident beam has angle θ_i with the normal line to the interface while θ_r is the angle of reflected beam. θ_t is the angle of refracted beam with the normal, which obeys the Snell's law.

$$n_i \sin \theta_i = n_t \sin \theta_t \quad (53)$$

Figure 12. Direction of the electric field and wave vector for reflected and refracted beams in the interface.

The Fresnel Equations for perpendicularly polarized (s-polarized) light are described as below (Hecht, 2017):

$$r_{\perp} = \left(\frac{E_{0r}}{E_{0i}} \right)_{\perp} = \frac{[n_i \cos \theta_i - n_t \cos \theta_t]}{[n_i \cos \theta_i + n_t \cos \theta_t]} \quad (54)$$

$$t_{\perp} = \left(\frac{E_{0t}}{E_{0i}} \right)_{\perp} = \frac{2n_i \cos \theta_i}{[n_i \cos \theta_i + n_t \cos \theta_t]} \quad (55)$$

For parallel-polarized (p-polarized) light the Fresnel Equations are (Hecht, 2017):

$$r_{\parallel} = \left(\frac{E_{0r}}{E_{0i}} \right)_{\parallel} = \frac{[n_t \cos \theta_i - n_i \cos \theta_t]}{[n_i \cos \theta_t + n_t \cos \theta_i]} \quad (56)$$

$$t_{\parallel} = \left(\frac{E_{0t}}{E_{0i}} \right)_{\parallel} = \frac{2n_i \cos \theta_i}{[n_i \cos \theta_t + n_t \cos \theta_i]} \quad (57)$$

The radiant flux density (W/m²) or irradiance is:

$$I = \frac{c\varepsilon_0}{2} E_0^2 \quad (58)$$

While ε_0 and c are permittivity (electric constant) and velocity in the vacuum.

For the boundary of two medium, “R,” reflectance factor, is relating to the ratio between reflected power and incident power:

$$R \equiv \frac{P_r}{P_i} \quad (59)$$

The incident power, P_i , (or radiant flux) is the energy per unit time flowing in the incident beam, and therefore it is the power arriving at the surface over A . In similar, P_r and P_t are the power in the reflected beam and the power being transmitted through area A , respectively. Based on Figure 13 if I_i , I_r , and I_t were the incident, reflected, and transmitted flux densities, and the cross-sectional area of the incident, reflected, and transmitted beams were, $A \cos \theta_i$, $A \cos \theta_r$, and $A \cos \theta_t$, the incident, reflected, and transmitted power are $I_i A \cos \theta_i$, $I_r A \cos \theta_r$, $I_t A \cos \theta_t$ respectively. Therefore, reflectance factor is written as:

$$R \equiv \frac{I_r A_r}{I_i A_i} = \frac{I_r A \cos \theta_r}{I_i A \cos \theta_i} \quad (60)$$

Figure 13. Cross-section of the incident, reflected, and transmitted beams.

By using Equation 5 and substantiation of $\varepsilon_0 = \frac{1}{\mu_0 c^2}$, flux density would be:

$$I = \frac{n}{2c\mu_0} E_0^2 \quad (61)$$

Which μ_0 is vacuum permeability (or magnetic constant). Converting the Equation 61 for medium n_i , I_r/I_i leads to:

$$\frac{I_r}{I_i} = \frac{n_r E_{0r}^2}{2c\mu_0} \cdot \frac{2c\mu_0}{n_i E_{0i}^2} = \frac{E_{0r}^2}{E_{0i}^2} \quad (62)$$

Because the angle of incidence on the interference is equal to the angle of reflection, the beam's area does not change on reflection. Therefore, "R," power reflectance factor becomes:

$$R = \left(\frac{E_{0r}}{E_{0i}} \right)^2 = r^2 \quad (63)$$

For obtaining the transmission factor, T, the ratio between transmitted power and incident power is calculated as:

$$T \equiv \frac{I_t A_t}{I_i A_i} \quad (64)$$

And,

$$\frac{I_t}{I_i} = \frac{n_t E_{0t}^2}{2c\mu_0} \cdot \frac{2c\mu_0}{n_i E_{0i}^2} = \frac{E_{0t}^2}{E_{0i}^2} \quad (65)$$

At the interface, the wavefronts of the incident and transmitted wave are tilted at different angles θ_i and θ_t , respectively, and the area of the incident and transmitted beams are not the same.

$$T = \frac{n_t \cos \theta_t |E_t|^2}{n_i \cos \theta_i |E_i|^2} = \frac{n_t \cos \theta_t}{n_i \cos \theta_i} |t|^2 \quad (66)$$

For reflection at normal incident angles $\theta_i = 0$, so the Fresnel equation would be:

$$R = \left(\frac{n_i - n_t}{n_i + n_t} \right)^2 \quad (67)$$

$$T = \frac{4n_i n_t}{(n_i + n_t)^2} \quad (68)$$

For the interface of air/glass which light enters the glass ($n=1.5$) from air $R=0.04$ and $T=0.96$.

In conclusion, for all different interference types have made due to either spatial coherency or temporal coherency the interference phenomena are very sensitive to optical path length. In addition, the number and contrast of the fringes can show the details about the interference setup and its source.

CHAPTER 2

Semiconductor Industries

Semiconductors are materials with conductivity between a conductor and an insulator. Their conductivity can be controlled by applying voltage and dopant concentration. The main usage of semiconductors is in bipolar transistors¹, MOS transistors², CMOS³, IC chips using in memories, microprocessors, and ASIC⁴s. Wiring places a crucial role in creating such devices. There is no way to connect these shrunk-sized devices other than microfabrication known as "semiconductor manufacturing."

The fabrication of integrated circuits (ICs) needs hundreds of sequential process steps, which start with using a uniformly doped silicon wafer. The processes to make wafer from silicon can be briefed as:

- Obtaining metallic grade silicon (MGS) from silicon dioxide (sand).
- Reacting MGS powder with HCl to form SiHCl₃ (TCS) in liquid.
- Purifying TCS by vaporization and condensation.
- Reacting TCS to H₂ to form high purity electronic grade silicon (EGS).
- Melting EGS and pulling the single crystal silicon ingot (Figure 14).
- Polishing and slicing the silicon ingot to make silicon wafers (Figures 15 and 16).

¹ Bipolar transistors can amplify electric signal, mainly used for analog systems.

² Also called MOSFET; electrically controlled switch, mainly used for digital systems.

³ Complementary Metal-Oxide-Semiconductor is the manufacturing process used to create processors.

⁴ An application-specific integrated circuit is an integrated circuit (IC) developed for a particular use.

Figure 14. Czochralski silicon crystal growth.

Figure 15. a) Single crystal silicon ingot. b) Silicon ingot sliced to Si wafers.

Figure 16. Silicon wafer in different standard sizes.

2.1 Semiconductor Fabrication Processes

The most important process steps used in the semiconductor fabrication are electrical chip design, physical layout, oxidation, film deposition, lithography (photoresist coating, exposure, development), etching, chemical mechanical planarization, doping of semiconductors, ion implantation, and diffusion.

Combining these fabrication steps can produce complex semiconductor devices and circuits, which is called planar technology. Since each fabrication step is applied to the entire silicon wafer, it is possible to make many devices with high precision at the same time and in the same environment. For example, in one 8-inch silicon wafer, 500 chips may be made. According to the Intel Microprocessor Quick Reference Guide (2017), in the Intel® Core™ i7-940 Processor, 731 million transistors have been used. Therefore, it is beneficial to reduce the area of each IC to reduce the size of devices and the number of

metal interconnects because the result is more chips per wafer and a lower cost per chip. Knowing that some conditions in the processing environment could have some impacts on existing wafer structure, studying those conditions is vital. For example, fabrication steps which have heating (or cooling) can cause the thin film delamination due to thermal expansion coefficient mismatch (Schmidt, 1998).

2.1.1 Electrical chip design and physical layout. After designing the electrical chip, the physical layout design (called blueprint) which is a negative photo of the pattern is provided. By printing the layout design to a glass plate, the lithography photomask is fabricated. Photomasks are typically made from transparent BK7 or fused silica with the chrome metal absorbing film pattern to block the light and blanks area for transmission of light (More details on 2.1.3).

2.1.2 Oxidation. Oxidation is the process in which the top thin layer of the silicon wafer is converted into silicon dioxide by exposing it to a mixture of high purity oxygen and hydrogen in a high temperature. Silicon dioxide is a high quality insulator, that is used for several purposes, such as serving as a mask or serving as the gate in the metal-oxide-semiconductor transistors. The ability of silicon to form high-quality silicon dioxide is an important reason, that silicon is still a dominant material in IC fabrication.

2.1.3 Lithography. Lithography is used to transfer the physical layout design to the surface of the wafer. The pattern information is recorded on a layer of photoresist on the top of the wafer, while the wafer substrate is rotated in a spin coater at high speed. By exposing the photoresist to the UV light, its material property changes and it would be soluble in an organic solvent (positive photoresist). The pattern defined by the mask is

either removed or remained after development, depending on the type of resist whether it is positive or negative. Therefore, a single mask can sometimes be used for two steps.

A glass mask containing the physical layout is brought near the wafer. The mask is transparent in the regions needs to process, and opaque in the others (for positive photoresist). The combination of mask and wafer is now exposed to ultra-violet light. Where the mask is transparent, the photoresist becomes soluble.

Developing the photoresist mask is either with an acid or base solution to remove the exposed areas of positive photoresist. After removing unwanted areas, they are soft-baked at a low temperature to harden the remaining photoresist.

A well-designed mask is the one that transfers the pattern into the wafer evenly over the entire area. The mask materials are so important in each recipe, because choosing a wrong mask can cause interaction between the mask and the etching process parameters. Some examples of silicon etch mask materials are metal, oxides, and resist.

2.1.4 Etching. Etching is used to remove material layers in order to create patterns. The pattern is defined by the etching mask in a way that the material, which should remain, is protected by the mask. The etching process steps and methods are covered in chapter 3.

2.1.5 Film deposition. The IC fabrication process requires the multitude deposition of layers of materials such as silicon nitride, silicon dioxide, Si, Al, TiN, etc. to either act as buffers for a processing step or as the insulating or conducting layers. The two most important deposition methods are the physical vapor deposition (PVD) and the chemical vapor deposition (CVD). PVD which is also called sputtering, evaporates the

particles from a sputter target in a vacuum chamber. The heat for evaporation is delivered by electron-beam or ion-beam bombarding. The Aluminum interconnect layers are typically deposited using sputtering. CVD uses a gas-phase reaction on the substrate surface at high temperatures. Silicon nitride is a sample material which is deposited using CVD. Since producing high temperatures is the most limitation factor for applying chemical vapor deposition, plasma is used to force the chemical reaction in the presence of radio frequencies instead of high temperatures. This deposition process is called plasma enhanced chemical vapor deposition (PECVD). The uniformity of deposited material in PECVD is better than PVD.

2.1.6 Chemical mechanical planarization. In IC fabrication, the flatness of the surface is crucial. Processes like etching, deposition, or oxidation lead to a non-planar surface. Chemical mechanical planarization (CMP) should be included before the deposition of any metal layer on top of the insulating SiO₂ layer. This process uses a slurry compound, which has an abrasive component such as aluminum oxide or silica to microscopically polish the surface layer and leave a very flat surface.

2.1.7 Doping of semiconductors. Semiconductors are materials which are neither conductors nor insulators. But with some engineering, they can behave like both conductors and insulators depending on their conditions. This is why they serve as the basis for switching and amplification. Doping refers to the addition of impurities into a semiconductor region (substituting Si by neighboring elements in the periodic table) in a controllable manner in order to define the electrical properties of this region. Therefore, in doped crystal region electrons can be moved by applying an external voltage. The two

classes of doping are p-type and n-type, which refer to the induction of positive and negative charge carriers. Ionized acceptors provide free holes in a semiconductor (p-type) while ionized donors provide free electrons (n-type). For example, doping P or As in silicon helps the material to perform as electrical conductor, and doping Si with the neighbors which have one less electron than Si, such as B or Al, creates ionized acceptors.

2.1.8 Ion implantation. In the semiconductor industry, the main technology for introducing doping atoms to form devices and integrated circuits uses the ion implantation method. The main advantage of using this technique is that the amount and position of the doping can be controlled. In this method, accelerated ionized atoms or molecules with the electric field are bombarding the silicon substrate. These particles penetrate into the target material until they come to rest beneath the surface because of interactions with the silicon atoms. The energy of ions controls the distribution, and the ion beam current and the exposure time controls the dose of the dopants in silicon. Also, the acceleration of the ions determines how deep they will penetrate into the material.

2.1.9 Diffusion. After doping, diffusion is applied to anneal the crystal defects after ion implantation. Since the dopant concentration may not be same within different regions, the concentration gradient is a driving force to spread the particles from regions of higher concentration to regions of lower concentration at high temperatures.

Figure 17 shows the copper wiring of a computer chip which passed all the IC fabrication processes.

Figure 17. Copper interconnects in an IBM chip.

In conclusion, among all of the semiconductor fabrication processes, etching and film deposition are two important ones, which their accurate studying in details helps to improve the process significantly.

CHAPTER 3

Etching Theory

Etching is the process of removing material by either dissolution or physical treatment, and is one of the crucial steps in micro-fabrication processes. Cleaning the surfaces and removing mechanical damages are two usages of the etching process. Mostly, the etching procedure is to create well-defined structures in integrated circuits, micromechanical, and optoelectronic devices. It also is being used to obtain special shapes as required for the proper functioning of devices and to reveal the defects on the surfaces (Nishi & Doering, 2000).

3.1 The Terminologies in the Etching Processes

These terms are often used in etching process:

3.1.1 Etch rate. Etch rate measures how fast the material is removed from the wafer surface. Higher etch rate increases the etch depth, while lower etch rate improves the precision. It is described as the ratio of thickness change after etching to the time of process (Figure 18).

Figure 18. Schematic of the etching process.

$$Etch\ rate = \frac{\Delta d}{t} \quad (A^\circ/min) \quad (69)$$

3.1.2 Uniformity. The etch uniformity measures the repeatability of etching process across the wafer, at each feature, and between wafer to wafer. For measuring uniformity, the etched surface thickness is measured in different locations, and the standard deviation for all data points is calculated to show uniformity across the points.

3.1.3 Selectivity. Selectivity is the ratio of etch rate for different layers of materials. Since the etching process must entirely remove the top layer of a multilayer structure without damaging the underlying or masking layers, measuring selectivity is so important. For example, selectivity is used to determine the ratio of etch rate of photoresist relative to the underlying material layers. Figure 19a, shows a bad selective etch, which not only removes the top layer but also removes a portion of the bottom layer. A highly selective etch is illustrated in Figure 19b which etching leaves the underlying material unharmed.

Figure 19. A bad selective etch (a) vs. highly selective etch (b).

3.1.4 Etch profiles. Etch profiles are generally divided into anisotropic etching and isotropic etching. Anisotropic profiles produce vertical, sharp, well-controlled features (Figure 20–a,c), but isotropic etches undercut the masking layer and form round cavities with sloping sidewalls (Figure 20–b,d). This profile is called isotropic because it removes the substrate equally in all directions. The nanostructure of this vertical surface

is important in many applications, particularly waveguides. (Shearn, Sun, Henry, Yariv, & Scherer 2010)

Figure 20. Different etch profiles.

3.1.5 Damage. Any surface or substrate destruction acquired from the processing technique is called damage.

3.2 Etching Techniques

Etching techniques have been widely used in the semiconductor industry since the invention of the integrated circuit (IC) in 1958. This technology is used to remove selected layers of material from the silicon substrate. There are two basic types of etching for bulk (subtractive) processing of silicon: wet etching which has liquid-gas base etchant and dry etching process which does not use chemicals.

The dry and wet etchant can be selected on the basis of different characteristics such as selectivity of etch masks, exposed metallization, availability of etch-stop methods

(doping or electrochemical), etch rate, degree of surface roughness created, safety, availability of facility, cost, etc. (Kocacs, 1998).

3.2.1 Wet etching method. Wet etching removes the coated material layer when the substrate is immersed in the etchant solution that reacts with the selected film. The dissolved layer is carried away by the etchant solution. Figure 21 shows the components of a simple wet etching device.

Figure 21. A wet etching device consists of a bath container with a stirrer, controllable hot plate, and slits to hold the wafers.

In the wet etching process, the selection of etchant is the key. Different etchants are effective for various semiconductor materials involved. In addition, in the presence of other materials such as metals, oxides, and photoresists, choosing the appropriate etchant could be important (Nishi & Doering, 2000). Another factor for choosing the effective etchant is consideration of their behaviors. The etchant behavior could be either isotropic or anisotropic (Kovacs, 1998). For example, silicon used in semiconductors has a single-

crystal structure and exhibits different etch rates for individual crystal orientation. In the 1960s, silicon etch was found to be orientation-dependent and concentration-dependent in some chemical solutions (Bean, 1978).

3.2.1.1 *Advantages of wet etching.*

- It is cheap, and due to purely chemical nature, it does not damage the substrate.
- The etching process is fast, and it is applicable to features bigger than 1-micron size.
- The process has simple equipment, and there is high selectivity of the material to work with.

3.2.1.2 *Disadvantages of wet etching.*

- Its process control is difficult due to the temperature sensitivity of materials. In addition, the processes are sensitive to time, and chemical concentrations as the final dimensions may change with etching time and chemical concentration.
- It is hard to use with features smaller than one micrometer. Etch block can happen due to surface tension.
- A large amount of chemical liquids is needed, and the cost of disposing of the chemicals which most of them are toxic is high.
- The wet etching process has multiple washing steps to remove each layer residue.

Because of disadvantages of wet etching process, wet etching is usually used for some non-critical processes. For more accurate processes dry etching methods are used.

3.2.2 Dry etching methods. There are different types of dry etching methods used in industry, which use gas reactant species to etch the film. The main four are

plasma etching, reactive ion etching, beam etching, and ion beam methods. Since all the four methods are using the plasma chambers, any plasma treatment performed without the use of chemicals is commonly referred to as dry etching.

3.2.2.1 Plasma etching. In this kind of etching, the substrate is placed on a grounded electrode, and plasma is generated at high pressure. The pressure in a typical experiment is about 0.2 to 2 Torr.

3.2.2.2 Reactive ion etching. In this kind of etching, the substrate is placed on a powered electrode for ion bombardment, and plasma is generated at low pressure. The pressure in this kind of etching is about 10^{-4} to 10^{-1} Torr.

3.2.2.3 Ion beam. In this method, plasma is generated in a separate chamber and ions are accelerated towards the substrate. In this way, ions and radical fluxes can be controlled independently. The pressure in a typical Ion beam etching experiment is about 10^{-4} to 10^{-1} Torr.

3.2.2.4 Beam methods. In this method, plasma is generated in a separate chamber, and mainly radicals are accelerated towards the substrate. Pressure is about 10^{-4} to 10^{-3} Torr.

Generally, substrates are placed in plasma as the reactive gas and the layer to be etched is removed by chemical reactions with the plasma or physically by ion bombardment. The volatile product is carried away with the chamber gas flow. All these three steps; plasma flow above, reaction of the layer, and movement of reactant products away from the surface can change the etch rate. So it can be one of the difficulties to

work with plasma etching rather than using wet etch since in wet etching it is easier to keep the etch rate low or close to constant.

3.3 Plasma Etching Principles

3.3.1 Plasma environment. A plasma is a partially ionized gas that is composed of equal numbers of positive and negative charges mixed with a different number of unionized neutral molecules. It is created in a vacuum chamber, which contains a constant flow of a gas at reduced pressure around 1mbar. To produce a glow discharge, a combination of low pressure and RF induced gaseous discharge are used in the plasma chambers (Figure 22).

Figure 22. a) A cross-sectional view of a typical plasma chamber. b) Plasma chamber window.

In the plasma environment occasionally a free electron can be released from a neutral molecule either by a random collision with other particles or by interaction with a photon

radiation (Figure 23). By applying a large voltage to the plasma, any free electrons, which may be present would be accelerated toward the anode (Figure 24).

Figure 23. Schematics of a plasma chamber: A free electron is released from a neutral molecule by its interaction with a photon.

Figure 24. Schematics of a plasma chamber: Acceleration of free electrons toward the anode.

As Figure 25 shows, since this voltage is so large, the electrons attain a high kinetic energy due to their low masses. Electrons begin to collide with gas molecules in two types of elastic and inelastic collisions. In the elastic collision kinetic energy is conserved, so all the kinetic energy of electron converted to kinetic energy of the gas molecule. However, it does not influence the velocity of molecules significantly because of the great mass difference between electron and molecule (Mass of electron = $9.11 \times 10^{-31}\text{ Kg}$, Mass of Argon = $6.64 \times 10^{-26}\text{ Kg}$).

In inelastic collision, the electron can excite the gas molecules to emit a photon or ionize them by removing an electron. Therefore, in this collision the electron kinetic energy converts due to the conservation of the total energy.

Figure 25. Schematics of a plasma chamber: Electrons collide with the gas molecules in elastic and inelastic collisions.

This emitted photon is the glow that commonly the plasma discharge is recognized with (Details on 3.3.4). In addition, the generated electrons are accelerated toward the anode, collide with new neutral molecules, and make the breakdown again. If the voltage were sufficient, the breakdown is continuing, and the plasma becomes filled with positive and negative charges (Figure 26).

Figure 26. Schematics of a plasma chamber: All generated electrons are accelerated toward the anode, and making breakdown.

On the other side, all positive charges are accelerating toward the cathode and with colliding to the negative electrode; they can emit the secondary electrons (Figure 27).

Figure 27. Schematics of a plasma chamber: Emitting secondary electrons by the collision of positive charges with the cathode.

The production of all the free electrons consisting of the electrons generated by the cathode and the electrons, which have made in the plasma environment makes a steady state self-sustaining discharge.

Some electron can go out of the system by colliding to chamber walls, recombining with positive ions and attaching to neutral molecules to make the negative ions.

3.3.1.1 DC glow discharge regions. The plasma discharge area between two nodes are globally neutral, but it contains regions of net positive and negative charges (Figure 28). Glow discharge regions are:

1. Cathode Dark Space (Crooke's Dark Space)
2. Negative Glow
3. Faraday Dark Space
4. Positive Column
5. Anode Dark Space

Figure 28. Schematics of a plasma chamber: Current distribution in a DC glow discharge.

3.3.1.2 DC vs. RF plasmas. Radio-frequency induction is used to start and maintain the stabilized plasma. A radio frequency AC is passed through a coil of wire, and by electromagnetic induction, the voltage induces across the ends of the other wire. The frequency of operation is often at 13.56 MHz, which is a band reserved for industrial use by the Federal Communications Commission in the United States.

3.3.2 Plasma etching process. In the plasma etching process, the wafer is loaded in a vacuum plate. The chamber is evacuated to a vacuum pressure of 0.1-0.2 Torr by a mechanical vacuum pump. The plasma process is begun and continues with RF-induced gaseous discharge while the carrier gas is presented into the chamber, raising the chamber pressure to about 0.3-1.2 Torr. The plasma excites the carrier gas atoms and molecules and ionizes them. In this time-varying field, electrons in the plasma oscillate between the anode and the cathode plates. Collisions of rapidly moving electrons with the slowly moving ions cause further ionizations. The potential between the plasma and the silicon substrate causes the electric field that drags the ions down to the substrate

(Forward power). The unmasked silicon bonds to the carrier gas (such as Fluorine) atoms to create the volatile etch product, which is carried away in the gas stream.

Generally, the plasma etching process etches the specimen in two ways; first by chemical bonding between specimen atoms and carrier gas ions and second by the mechanical milling of accelerated ions. The chemical etching makes an isotropic etching while the mechanical milling is anisotropic (Shearn et al., 2010).

There are different plasma etching mechanisms and recipes, one can use for different etching application. One of the main mechanisms is “Blocking Mechanism.” In this procedure which is shown in Figure 29, the main goal is to keep the etching anisotropic. SF_6 gas is injected into the chamber to start etching. After a few seconds, flow of SF_6 stops and then C_4F_8 gas is injected into the chamber as the passivation step. A thin polymeric layer of C_4F_8 forms on the substrate, on both the vertical and horizontal surfaces. The etching depth and thickness of passivation layer are determined by etching time and passivation step time, respectively. When the coating is complete, the C_4F_8 flow stops and again SF_6 is injected as the etching gas. Since ion bombardment is mainly in the vertical direction, the ions first mill away the horizontal passivation layers and the polymer deposition on sidewall blocks etch process on walls. Therefore, passivation layer protects the exposed vertical silicon while the unmasked horizontal silicon is etched away. The sequence of flowing these two gases continues until the etching process is terminated. (Takahashi, Itoh, Nakamura, & Tachibana, 2000).

Figure 29. Blocking mechanism of etching.

Cliteur reports that “When SF_6 (as the etching gas) is injected into the chamber while C_4F_8 acts as the passivation gas, the plasma ionizes and radicalizes the gas molecules to create a mixture of SF_x and F_y ions and neutrals, where x and y range from 0 to 6 and 1 to 2, respectively” (Cliteur et al., 1999).

By changing the etching process parameters, the etch rates, selectivity, etch roughness, and the etch profile can be controllable. These parameters are not limited to Inductively Coupled Plasma (ICP) power, forward power, temperature, chamber pressure, and gas flow rates.

- The ICP power controls the amount of ionization directly. Therefore, by increasing the ICP power, the chemical etch rate will increase both vertically and laterally.
- By increasing forward power which is the electric field between the plasma and the table electrode, the movements of ions increase, so the etching will increase due to increasing the silicon milling. This manner effects on increasing vertical etch rate while the lateral etching also slightly increases due to ions angular distributions.

- In some plasma etching methods changing the chamber temperature can affect the etch rate. If the polymer passivation layer is sensitive to temperature effects, even small temperature fluctuations change the etching profile.
- Since the chamber pressure controls the amount of gas in the chamber, it effects on the ionization and furthermore the etch rate.
- Changing gas flow rate changes the etch rate directly. While it has an optimum for increasing the gas flow, increasing the pressure of the chamber can be done by throttle valve controller placed between chamber and pump system.

The depth of the cavity can be controlled by etching time and known etch rate. Figure 30 shows a cross-section of Intel's 90 nm generation copper interconnects, which measuring depth of cavity plays an important role in it. "The combined MT1-MT9 interconnect stack provides high performance and high reliability and enables a completely lead-free product." (Moon et al., 2008)

Figure 30. A cross-section of 90 nm generation Cu interconnects. Light gray regions in all layers are Cu interconnects. Dark regions of MT 1-7 are low-k dielectric material; MT 8 is SiO₂, and MT 9 is polymer.

3.3.3 Chemical reactions in plasma etching. There are different kinds of recipes to use plasma chambers for dry etching purposes. The gases are selected based on the material, which is the target of etching and its compounds. These gasses are not limited to SF₆, CF₄, CHF₃, O₂, H₂, Ar, and He. A typical dry etching process is when Fluorine radical combines with silicon wafer to produce silicon tetrafluoride gas ($4F + Si \Rightarrow SiF_4$). To have Fluorine radical, CF₄ is used as the inert gas and with interaction with an electron Fluorine radical is produced ($CF_4 + e \Rightarrow CF_3 + F + e$). Si compounds are such as Si, SiO₂, Si₃N₄ and the byproduct Fluorine Chemistries are SiF₂ and SiF₄ which are volatile species.

Some gases are added in the plasma chamber to optimize the etching conditions:

- Hydrogen reduces fluorine concentration by combination to form HF.
- Adding Oxygen can increase fluorine concentration by combining to carbon (CO₂) to prevent its reaction by F and making CF₃.
- Argon increases plasma density which leads to increasing the Fluorine radicals' concentration.
- Adding He can cool the system and helps the photoresist to remain in place.

In Table 2 the kind of gas plasma required for etching different materials are listed.

Table 2

The Kind of Gas Plasma Required for Etching Each Material

Material	Kind of Gas Plasma
Si	CF ₄ , CF ₄ + O ₂ , CCl ₂ F ₂
poly-Si	CF ₄ , CF ₄ + O ₂ , SF ₆ + O ₂ , CF ₄ + N ₂
SiO ₂	CF ₄ , CF ₄ + H ₂ , HF*
	CCl ₂ F ₂ , C ₃ F ₈ **, C ₂ F ₆ + H ₂ **
Si ₃ N ₄	CF ₄ , CF ₄ + O ₂
Mo	CF ₄ , CF ₄ + O ₂
W	CF ₄ , CF ₄ + O ₂
Au	C ₂ Cl ₂ F ₄
Pt	CF ₄ + O ₂ , C ₂ Cl ₂ F ₄ + O ₂ , C ₂ Cl ₂ F ₃ + O ₂
Ti	CF ₄
Ta	CF ₄
Cr	Cl ₂ , CCl ₄ , CCl ₄ + Air
Cr ₂ O ₃	Cl ₂ + Ar, CCl ₄ + Ar
Al	CCl ₄ , CCl ₄ + Ar, BCl ₃
Al ₂ O ₃	CCl ₄ , CCl ₄ + Ar, BCl ₃
GaAs	CCl ₂ F ₂

There are other materials like aluminum and copper that cannot be etched by fluorine. Since AlF₃, the etch product, is not volatile, chlorine or bromine are used in aluminum etching. In this case, AlCl₃ is highly corrosive to the remaining aluminum film and must be removed quickly after plasma etching. Also, copper does not form any volatile compounds with known plasma etch gases, and therefore cannot be etched by reactive ion etching.

3.3.4 Plasma optical emission spectra. Since the light emitted by the plasma is due to the return of excited electrons to their ground state and the energies between the electron states are well defined for each element, each species will emit light at specific wavelengths, which will give the possibility to analyze the plasma. It means the excited photons produce the optical emission, which is characteristic of the composition of the

glow discharge gases. For example, argon emits different wavelengths depend on in which atomic layer the electron transition happens (Figure 31). Optical spectrometers can be used to monitor the different light wavelengths in plasma chambers for recognizing the type and amount of gases exist in the environment.

Figure 31. Optical emission spectra lines for Ar plasma in the range 690–850 nm.

Monitoring the optical emission of the plasma can be a good diagnostic for incorrect plasma striking, conditions or other changes in the plasma. If in a multiple gas recipe, the emission looks like only one of the gas species instead of the average of the colors, it can show the problem. “This happens when the other species are not being ionized, and thus will cause the process to take on a completely different character from a calibrated recipe.” (Shearn et al., 2010)

In Figure 32, an example of plasma emission lines is illustrated. These rich spectral lines are used in real etcher environments. Figure 33 shows the relative irradiance in a plasma environment when argon, oxygen, and SF_6 are present compared to the existence

of argon gas. In this picture, the plasma spectral lines are in such way that to resolve the required peaks, a filter with bandwidth less than 1 nm is needed. In addition, this narrow bandwidth filter should have good sideband rejection to discriminate against strong unwanted neighbor peaks.

Figure 32. An example of plasma emission lines.

Figure 33. Relative irradiance in a plasma environment for two different recipes.

CHAPTER 4

EtchRate Wafer

4.1 Etch Monitoring Wafers

Nowadays, many groups in manufacturing or micromachining research are using 300 mm (12 in) wafers. Other groups continue the older processes using 150 mm (6 in) silicon wafers. Typical 300 mm wafers are 750 μm thick and are used widely in the semiconductor industry for purposes such as producing memory chips (Figure 34).

Figure 34. Fabrication of memory chips on a silicon wafer.

A typical etch monitoring wafer in semiconductor fabrication plants (Fabs) is a sample wafer of the same size and type of the wafer which needs to be processed. This wafer is placed on the same plasma environment and the same plasma etch process conditions for the production wafers, but is used only for inspecting the quality of the process. The target layers are coated on the monitoring wafer. The time between the monitor test and the actual test is kept minimized to simulate the same conditions. The

thickness of the coated layer is measured before starting the etch process. Then, the monitoring wafer is sent through the etch tool and the thickness is measured after etching again. Therefore, the etch rate is calculated based on the thickness difference divided by the duration of the etch. The semiconductor fabs are set up to automatically use such an etch monitoring wafer easily with minimal man power. Multiple test wafers are often needed to adjust process conditions. Although these test wafers are cheap and easy to use, they measure the average etch rate and are not suitable to measure more complicated etch recipes. For such a recipe, the user needs to split the etching procedures into different steps.

A new method to measure the etch rate is using a temperature measuring system to monitor the silicon wafer surface temperature during plasma etching. This tool (EtchTemp wireless wafer), which is a current product of the KLA-Tencor company (KT), is the next-generation of in-situ⁵ plasma etch wafers (“EtchTemp Plasma, n.d.). During the etching process, all real-time temperature data are stored in memory, which is embedded inside the monitoring wafer. Therefore, by using this wafer, the user can save the data in the vacuum chamber wirelessly. For optically transparent films, the tools used are typically ellipsometers (e.g. Aleris or SpectraSight or WaferSight), and for metal films, the film step height measuring tools such as HRP-x50 are used (“Metrology Solutions”, n.d.).

⁵ On site, in the same place the phenomenon is occurring without isolating it from other systems.

4.2 KLA-Tencor Wireless EtchRate Wafer

The wireless EtchRate (ER) prototype wafer discussed here is an instrumented wireless silicon wafer which has the EtchRate sensors assembled on its surface. The EtchRate sensor measures the change in film thickness in situ in a plasma processing chamber and the wafer stores all process data in its memory. Although current etch monitor methods are familiar and inexpensive, they are not recording the thickness while the process is running. The EtchRate sensors are placed in three different positions on the wafer surface to measure the uniformity across the wafer. One sensor is located at the center of the substrate, one sensor is located at a mid-radius of the substrate and another one is at the edge. This EtchRate sensor measures silicon dioxide and silicon nitride removal rate as a function of time and location. The users will have the spatial and temporal etch rate data for the surface of the wafer. Since this wafer gathers the etch rate information in real time, it can be more valuable compared to today's fabs' best known method (BKM⁶) which is using etch monitoring wafers. Another benefit of using the wireless EtchRate wafer is that it reduces the time to test and analyze the data. Since EtchRate wafer presents the etch rate data as a function of time, mapping the etch sequences is possible.

The wireless ER wafer consists of a central processing unit (CPU), memory, battery, circuits which are configured to communicate wirelessly, and two semiconductor wafers which all the sensors and parts are sandwiched between them to provide noise isolation.

⁶ BKM abbreviation, which is widely used in semiconductor industry, is a process of making and checking designs using a proven workflow and checklist. ("Intel Quality System Handbook", 2014)

An off-line software is designed to process and analyze the stored data from the wafer and present spatial and time-resolved contour maps. An etch rate interferometer sensor, as described in Chapter 1, is an on-axis interferometer which the target layer is coated on the slab. The reference interference arm and measurement arm are collected with the sensor, and the cross-correlation is measured to determine a thickness change. Note that the measurement arm in this interferometer is being etched; therefore, the thickness change could be measured continuously. After the software starts data acquisition, the wireless ER wafer is inserted to the plasma chamber in the same conditions in which regular wafers are etched. During the etching process, the reference waveform and data waveform are stored in the ER wafer's memory and will be read after data collection is done.

4.2.1 EtchRate sensor structure. The etch rate interference sensor consists of a photodetector, a collimator, an aperture, a wedge or slab (based on needed resolution), a thin film coated as a 99% mirror, a thin film coated as a band-pass filter, and the etch target layer. Figures 35 and 36 show the two different designs that EtchRate sensor may have.

Figure 35. EtchRate sensor with slab window design.

Filtering the optical emission lines of plasma produces the monochromatic source for the sensor to be able to generate the interference pattern. In the first design (Figure 35), an optical slab is used to make interference fringes. The slab has a step at the middle to make two measurements in Part B and C. By choosing different optical filters in part B and C, measurements for two different lines of plasma can be done. In part A, light can pass through the slab to the detector without passing any optical filter.

Figure 36. EtchRate sensor with wedge window design.

In the second design, the optical slab is substituted with an optical wedge. The difference between these two designs will be discussed in 4-2-3.

4.2.1.1 Photodetector. The detector used in EtchRate sensor is a linear optical sensor array. This kind of detectors consists of numbers of photodiode detector, which line up in a row to form an array. Each photodiode makes a pixel, which is sensitive to the incident light intensity (Figure 37).

Figure 37. A typical linear optical array.

To select a right optical detector for the specific application, one should notice to the detector “optical response” (Figure 38).

Figure 38. Optical response plot of a typical linear array detector.

4.2.1.2 Optical wedge or slab. One of the main elements in EtchRate sensor is the optical wedge or slab (depending on the design) which produces the interference pattern. This optical element is a wedge or transparent window slab made of a material that allows transmission of light in the wavelength of interest. As shown in Figure 39, BK7, quartz or UV fused silica are materials that can be selected depending on the detecting wavelength. It is important that the thickness variation of this optical element be well characterized.

Figure 39. Optical transmission of BK7, UV fused silica and quartz (Optical materials, n.d.).

To choose the right material, its low thermal expansion and high resistivity to abrasion should be considered. For example, fused silica is a good choice for this application.

4.2.1.3 Etch rate interferometer mirror. As discussed in chapter 1, in an on-axis interferometer one of the mirrors is replaced by the internal reflection of the slab. Since the reference beam and the measurement beam are not of equal magnitude, the contrast of the interference is greatly reduced. To have the maximum contrast interference pattern, the bottom flat side of wedge/slab is coated to function as a 99% reflected mirror. The reflective side is the side towards the optical wedge.

4.2.1.4 Band-pass filter. EtchRate sensor includes a band-pass filter to select the desired optical emission line in the plasma. This filter, which is coated as a thin film on one side of the wedge/slab, limits the input light to the wavelengths in the range of interest to pass through.

4.2.1.5 Collimator. A collimator is designed to limit the angle of incident rays whether they come directly from the plasma above or passed through the coated filter. By limiting the incident angle, the Lambertian, incoherent illumination of plasma can interfere constructively and destructively after passing through the wedge.

4.2.1.6 Aperture. The aperture limits the opening through which the cone of light may pass. It determines the incident beams angle, size, and distribution to define the field of view.

4.2.1.7 Thin film layers. The thin film coated layers on the wedge or window surface are the same layers as the etching materials. It could be SiO_2 or Si_3N_4 which are mainly used in the etching processes. This thin film layer needs to be transparent to be able to produce an interference pattern. SiO_2 has excellent thermal stability, and since it is grown in-situ it is very adherent (Kovacs, 1998).

4.2.2 EtchRate wafer working mechanism.

4.2.2.1 Data collection. After setting the data collection frequency and integration time, the mission is started by use of ER wafer software. The ER wafer is inserted into plasma chamber in the same way that regular wafers insert. A vacuum chuck or holding stage is required to hold the ER device. After starting the mission, data is acquired at the data collection frequency. The data consists of a series of columns of values representing the intensity on each pixel as ADC unit (Analog to digital converter) during the time. For example, one row with 128 columns shows one second of 1Hz data collection for 128-pixel detector. Before any etching (or deposition) starts, two reference beams need to be collected. One is the intensity reference waveform measuring through an area not covered by the partial mirror to subtract a majority of the bias (Bias) and the second is the interference waveform through the wedge (REF). To reduce the noise in reference beam (REF), the average of multiple waveforms is measured to report as the reference signal.

When the etching procedure is started the coated layer on EtchRate sensor is being etched during the etching process. Figure 40 shows a set of data for etch rate wafer with the design illustrated in Figure 35. Each signal is an average of signals within integration

time (for example 250ms). Signal A is the reference signal (Bias), and signal B and C are the signals for two different thickness of the etched layer.

Figure 40. Signal output. Line A represents a reference waveform (Bias); Lines B and C represent measured waveforms (O'Brien & Jensen, 2014).

4.2.2.2 Contrast in EtchRate sensor. As described in chapter 1, the intensity of each pixel in wedge/slab interference is:

$$I = I_{Measured} + I_{Reference} + 2\sqrt{I_{Measured}I_{Reference}}\cos(\Delta\phi) \quad (70)$$

Based on Figure 41 and knowing that for glass-air interface $T=0.96$ and $R=0.04$ (chapter 1 page 25), and a 99% mirror is coated on bottom layer; $I_{Measured}$ and $I_{Reference}$ are:

$$I_{Measured} = I_{in} * 0.96 * 0.99 * 0.04 * 0.01 \quad (71)$$

And:

$$I_{Reference} = I_{in} * 0.96 * 0.01 \quad (72)$$

Therefore, the ratio between them is 4%.

$$\frac{I_{Measurment}}{I_{Reference}} = \frac{I_{in} * 0.96 * 0.99 * 0.04 * 0.01}{I_{in} * 0.96 * 0.01} \simeq 4\%. \quad (73)$$

Figure 41. The incident beam and the reference beam paths passing through a wedge.

Moreover, regarding to Equation 70 the maximum and minimum of interference pattern are when $\cos(\Delta\phi)$ is equal to one and -1 respectively.

$$I_{Max} = I_{Measured} + I_{Reference} + 2\sqrt{I_{Measured}I_{Reference}} \quad (74)$$

$$I_{min} = I_{Measured} + I_{Reference} - 2\sqrt{I_{Measured}I_{Reference}} \quad (75)$$

$$Contrast = \frac{I_{Max} - I_{min}}{I_{Max} + I_{min}} = \frac{2\sqrt{I_{Measured}I_{Reference}}}{I_{Measured} + I_{Reference}} \approx \frac{0.4I_{Reference}}{1.04I_{Reference}} = 0.38 \quad (76)$$

4.2.2.3 Data interpretation in the wedge design. Since we can measure the thin film thickness by using the number of fringes and their positions, studying the interference fringes helps to interpret the ER wafer data. According to O'Brien and Jensen (2014), by using the wedge instead of the slab in sensor design, thickness change causes ‘walking’ interference waves. The resulting interference fringes are sinusoidal waves across the linear detector which start with a dark fringe at the corner of the wedge (Figure 42).

Figure 42. The first fringe is destructive interference.

Before any thickness change, the signal from the reference wave has been stored. The first dark fringe is appearing at the beginning of wedge incline, so by changing wedge thickness the first node stays at the same position, which causes the waves to seem ‘walking’. It is assumed that etch rate is uniform across the detector length. By having uniform thickness change, the ‘walking’ fringes move across the array as a function of time.

Figure 43. Movement of interference wave by changing the thickness.

In Figure 43, the wedge is coated in 3 layers. The sinusoidal waves shown below the wedge are interference waves, which are started with dark fringe at the extension of layers' surfaces. The orange wave presents the reference wave for the un-etched layer. By etching the layer, the first dark in interference wave shifts toward the right; therefore, the interference pattern moves to the right. The red star shows the location of the first maxima of the reference wave (the orange wave). After the etching starts, the orange wave moves toward the right side; location 1 and later toward location 2. Two yellow stars show the movement of the red star after etching in two different thicknesses. In conclusion, by tracking the phase shift the material thickness change can be calculated. For this purpose, the 'Cross-Correlation' value for two waves would be computed. Cross-correlation is a standard method of measuring how two series are correlated. As shown in Figure 44, for calculating the cross-correlation of two waves, the shift of the first maxima in measured data from the first maxima in reference interference pattern can be measured. A fringe that has shifted an entire cycle would correspond to an optical path length change of one wavelength.

Figure 44. The first maxima of reference wave (Black) and measurement wave (Brown) are chosen to measure the cross-correlation.

4.2.2.3.1 *Measuring cross-correlation.* In signal processing, using cross-correlation function is the best way to find the similarity of two different signals. Considering two signals as $F(x)$ and $G(x)$, the cross-correlation of them is:

$$R(i) = F(x) \otimes G(x) = \int_{-\infty}^{+\infty} F(x)G(x-i)dx \quad (77)$$

$F \otimes G$ is function of the displacement of one relative to the other⁷. Using this function in ER wafer data analysis, it "scans" all the pixels of measurement waveform $G(x)$ until it finds a match for reference waveform $F(x)$. In the matching case, the integration of the product of two similar functions will be the maximum. The goal is to find the value of the displacement shift (i) that maximizes the correlation coefficient. The correlation coefficient of 1 means the beams are in perfect correlation; zero means there is no correlation between the waves; and -1 shows the perfect anti-correlation between the waves. Figures below show reference signal (F) and data signal (G) on the left plot as a function of time, their ratio, and their cross-correlation function in three situations of π phase shift (Figure 45), an arbitrary phase shift (Figure 46), and zero phase shift (Figure 47). The red dot in cross-correlation function plot shows the cross-correlation coefficient amount for that phase shift.

⁷ It also could be defined as function of time.

Figure 45. Cross-correlation plot: F and G have phase shift of π .

Figure 46. Cross-correlation plot: F and G have arbitrary phase shift.

Figure 47. Cross-correlation plot: F and G have phase shift of zero.

Another type of computing the displacement shift which is discussed in O'Brien and Jensen (2014) is calculating the error function (Equation 78). After storing reference beam and data beam in the ER wafer, the error function 'e' will be computed. Therefore, in real time this calculation can be done for each sensor location and at each sample interval for each data wave measurement.

$$e_M = \sum_{n=1}^{N-M} (R_n - P_{n-M})^2 \quad (78)$$

In this equation, R_n and P_n are the intensities of the reference beam and the measurement beam on pixel n of the linear array detector; N is the total number of pixels and M is the offset of the reference to the currently measured waveform. By minimizing the difference between the reference beam and measurement beam, the value of M , the offset, could be calculated.

4.2.2.3.2 Measuring thickness by use of phase shift. In data collection, for each column of measured waveform there is a column of error function coefficient (difference from reference wave). By integrating the error function and minimizing it with the solver, the offset (phase shift) between the reference wave and the measurement waves (M) could be computed. In appendix A, an example of computing error function and finding the phase shift of two sinusoidal waves are described.

As explained in 1.3.2 and Equation 52, by knowing the relative phase shift, the exact thickness of the layer will be calculated. For example, if there were 14 fringes over 128 pixels of the linear array detector, using on-axis interference Equation 52 the thickness change for each pixel phase shift in $\lambda=704$ would be:

$$\Delta t = \frac{\Delta m \lambda}{2n} = \frac{14}{128} \cdot \frac{704}{2 \cdot 1} = 38.5 \text{ nm/pixel} \quad (79)$$

Note that since the reference and the measurement waves are joining in the air, n , the refractive index would be 1.0. Now if the minimum of error function happens in the phase shift of 5, the thickness of the layer is:

$$t = 5 \times \Delta t = 5 \times 23.04 = 115.2 \text{ nm} \quad (80)$$

In this case, the sensitivity of the method for 14 fringes over 128 pixels with the wavelength of 704 nm is computed as 19.25 nm (half of pixel).

$$\Delta t_{min} = \frac{\Delta t}{2} = \frac{\Delta m \lambda}{4n} = \frac{14}{128} \cdot \frac{704}{4 \cdot 1} = 19.25 \text{ nm/pixel} \quad (81)$$

In the etch rate wafer data analysis, the real-time thickness calculation is plotted by time to show the thickness for removed or deposited material (or etch or deposition rate in nm/sec) in a contour map.

4.2.2.4 Wedge surface quality. Since the layer thickness calculation measures the overall thickness change of layer plus wedge, the surface uniformity of wedge should be precisely known. Any imperfection either scratches or surface flatness can cause an error in thickness measurement analysis. The imperfection of slab wedge surface quality and surface irregularity can also affect the interference pattern. Surface quality is the measurement of the surface imperfections caused in the manufacturing or handling process. The most common method of rating surface quality is U.S. Military Performance Specification (MIL-PRF-13830B), which uses “scratch and dig” numbers based on calibrated standards prescribed therein. The surface irregularity is measured in fractions of a reference wavelength, usually 632.8nm. A surface reported as $\lambda/20$ flat has a maximum peak to valley deviation of 632.8/20 or 30 nm (with one significant figure). The wedge or slab in EtchRate sensor has $\lambda/10$ irregularity, and it might change during the etching process.

4.2.3 Difference between wedge and slab. As it was shown in section 4.2.2.3, wedge interference pattern is moving during layer thickness change, but in the slab, each layer has its own interference pattern, which cannot be related to the other layer

thicknesses. Therefore, the way that the layer thickness is calculated is different. In the wedge design, the thickness is measured by cross-correlation method; but in a flat slab, since the sinusoidal wave in interference pattern shows the layer thickness, measuring any thickness better than a quarter wavelength is difficult. Resolution is reported as half of smallest measurement in which all 128 pixels show constructive fringes. (For example 176 nm for a 704 nm input light).

$$Resolution = \frac{\Delta t}{2} = \frac{\Delta m \lambda}{4n} = 1. \frac{704}{4*1} = 176 \text{ nm} \quad (82)$$

It is obvious that the resolution in flat slab design is so poor compared to 19.25nm (Equation 81) in wedge design.

SUMMARY

This paper is a review of designing EtchRate wireless silicon wafer at KLA-Tencor Corp. to temporally and spatially monitor the thickness change of wafer top layer during the plasma etch. EtchRate wafer can capture the thickness change process in situ while the regular etch rate uniformity measurements such as statistical process controls (SPCs), are time-consuming and also don't show the etch rate process during the time. Three EtchRate sensors are located on the wireless wafer platform to act like regular blank wafers which SPC machines work with. The EtchRate wafer is able to provide information about etch rate uniformity across the wafer during the plasma etch quickly and easily instead of sensing one single plasma parameter such as ions, electrons, photons, or RF-current. In addition, using the EtchRate wafer is very easy, and it will decrease the number of test cycles. It can also be used during deposition processes to monitor film thickness change. All of these benefits are highly valued by customers in this industry.

REFERENCES

- Bean, K. E. (1978). Anisotropic etching of silicon. *IEEE Transactions on Electron Devices*, 25, 1185-1193.
- Bellan, P. M. (2008). *Fundamentals of plasma physics*. Cambridge University Press.
- Bittencourt, J. A. (2013). *Fundamentals of plasma physics*. New York: Springer.
- Chapman, B. (1980). *Glow discharge processes: Sputtering and plasma etching*. Wiley.
- Cliteur, G. J., Suzuki, K., Tanaka, Y., Sakuta, T., Matsubara, T., Yokomizu, Y. & Matsumura, T. (1999). On the determination of the multi-temperature SF₆ plasma composition. *J. Phys. D*, 32(15), 1851-1856.
- Datta, M. (2003). Electrochemical processing technologies in chip fabrication: Challenges and opportunities. *Electrochimica Acta*, 48(20-22), 2975.
- Ferraro, P., Wax, A., & Zalevsky, Z. (2011). *Coherent light microscopy: Imaging and quantitative phase analysis*. Springer Berlin Heidelberg.
- Hecht, E. (2017). *Optics* (5th ed.). Pearson Education, Edinburgh Gate.
- Intel. (2017). *Intel® Core™ Processor Intel Microprocessor Quick Reference Guide*.
- Intel quality system handbook* (2014). Intel.
- Kovacs, G. T. A. (1998). *Micromachined transducers sourcebook*. WCB/McGraw-Hill.
- Lee, C. G. N., Kanarik, K. J., & Gottscho, R. A. (2014). The grand challenges of plasma etching: A manufacturing perspective. *Journal of Physics D: Applied Physics*, 47(27), 273001.
- Lu, X., Wu, S., Gou, J., & Pan, Y. (2014). An atmospheric-pressure, high-aspect-ratio, cold micro-plasma. *Scientific Reports*, 4, 7488.
- Moon, P., & et all. (2008). Process and electrical results for the on-die interconnect stack for intel's 45nm process generation. *Intel Technology Journal*, 21(2), 87-92.
- Nihtianov, S., & Luque, A. (2014). *Smart sensors and MEMS: Intelligent devices and microsystems for industrial applications*. Elsevier Science.
- Nishi, Yoshio, Doering, Robert, (2000). *Handbook of semiconductor manufacturing technology*. New York: Marcel Dekker.

- O'Brien, K., & Jensen, E. (2014). *Thickness change monitor wafer for in situ film thickness monitoring* US Patents.
- O'Brien, K., Jensen, E., Quli, F., & Sun, M. (2013). *Film thickness monitor* US Patents.
- Paul, K. C., Hatazawa, S., Takahashi, M., Cliteur, G. J., & Sakuta, T. (1999). Diagnoses of inductively coupled SF₆ and N₂ plasmas at atmospheric pressure. *Thin Solid Films*, 345(1), 134-139.
- Rubin, L., & Poate, J. (2003). *Ion implantation in silicon technology*. American Institute of Physics.
- Saleh, B. E. A., & Teich, M. C. (2007). *Fundamentals of photonics*. Wiley.
- Schmidt, M. A. (1998). Wafer-to-wafer bonding for microstructure formation. *Proc. IEEE*, 88(8), 1575-1585.
- Serway, R. A., & Jewett, J. W. (2013). *Physics for scientists and engineers*. Cengage Learning.
- Shearn, M., Sun, X., Henry, M. D., Yariv, A., & Scherer, A. (Eds.). (2010). *"Semiconductor technologies", advanced plasma processing: Etching, deposition, and wafer bonding techniques for semiconductor applications*. Laboratory of Applied Physics California Institute of Technology: Grym, Jan.10.5772/8564
- Takahashi, K., Itoh, A., Nakamura, T., & Tachibana, K. (2000). Radical kinetics for polymer film deposition in fluorocarbon (C₄F₈, C₃F₆ and C₅F₈) plasmas. thin solid films. *Thin Solid Films*, 374(2), 303-310.
- Takeuchi, H., Wung, A., Sun, X., Howe, R. T., & King, T. J. (2005). Thermal budget limits of quarter-micrometer foundry CMOS for post-processing MEMS devices. *IEEE Trans. Electron Devices*, 52(9), 2081-2086.
- Zernike, F. (1938). The concept of degree of coherence and its application to optical problems. *Physica*, 5(8), 785-795.

APPENDICES

A. Mathematical Analysis

The stored data in the etch rate wafer is the reference wave data and the measurement data for that specific thickness. To calculate the thickness, it is needed to compute the offset (in pixels) between two waves. Knowing the thickness change for each pixel phase shift (nm/pixel) (Equation 51 and 78) and offset, one can measure the thin film thickness. Here are two examples to compute the offset by calculating the cross-correlation and by using the error function for two sinusoidal waves.

An arbitrary reference wave is plotted in Figure A1 in 128 pixels. A measurement wave is also illustrated with the shift of 7 pixels to the left of reference wave (-7 in relative).

Figure A1. Simulation of reference and measurement waves in 128-pixel linear array detector.

A.1 Cross-correlation. To find the offset between two waves with cross-correlation function, CORR function in Excel is used. Figure A2 is the spreadsheet sample in which correlation function for two mentioned reference wave and measurement wave is calculated.

	B	C	D	F
2			=SIN(B9*3.14/2*180)	
3			=SIN(B10*3.14/2*180)	
4			=SIN(B11*3.14/2*180)	The Pixel # for first Maxima of correlation function
5			=SIN(B12*3.14/2*180)	=MATCH(MAX(E9:E29),E9:E29,0)-1
6			=SIN(B13*3.14/2*180)	
7			=SIN(B14*3.14/2*180)	
8	Pixel #		=SIN(B15*3.14/2*180)	
9	0	=SIN(B9*3.14/2*180)	=SIN(B16*3.14/2*180)	=CORREL(OFFSET(REFERENCE,\$B9,0,128-\$B9,1),OFFSET(meas,0,0,128-\$B9,1))
10	1	=SIN(B10*3.14/2*180)	=SIN(B17*3.14/2*180)	=CORREL(OFFSET(REFERENCE,\$B10,0,128-\$B10,1),OFFSET(meas,0,0,128-\$B10,1))
11	2	=SIN(B11*3.14/2*180)	=SIN(B18*3.14/2*180)	=CORREL(OFFSET(REFERENCE,\$B11,0,128-\$B11,1),OFFSET(meas,0,0,128-\$B11,1))
128	119	=SIN(B128*3.14/2*180)	=SIN(B135*3.14/2*180)	=CORREL(OFFSET(REFERENCE,\$B128,0,128-\$B128,1),OFFSET(meas,0,0,128-\$B128,1))
129	120	=SIN(B129*3.14/2*180)	=SIN(B136*3.14/2*180)	=CORREL(OFFSET(REFERENCE,\$B129,0,128-\$B129,1),OFFSET(meas,0,0,128-\$B129,1))
130	121	=SIN(B130*3.14/2*180)		=CORREL(OFFSET(REFERENCE,\$B130,0,128-\$B130,1),OFFSET(meas,0,0,128-\$B130,1))
131	122	=SIN(B131*3.14/2*180)		=CORREL(OFFSET(REFERENCE,\$B131,0,128-\$B131,1),OFFSET(meas,0,0,128-\$B131,1))
132	123	=SIN(B132*3.14/2*180)		=CORREL(OFFSET(REFERENCE,\$B132,0,128-\$B132,1),OFFSET(meas,0,0,128-\$B132,1))
133	124	=SIN(B133*3.14/2*180)		=CORREL(OFFSET(REFERENCE,\$B133,0,128-\$B133,1),OFFSET(meas,0,0,128-\$B133,1))
134	125	=SIN(B134*3.14/2*180)		=CORREL(OFFSET(REFERENCE,\$B134,0,128-\$B134,1),OFFSET(meas,0,0,128-\$B134,1))
135	126	=SIN(B135*3.14/2*180)		=CORREL(OFFSET(REFERENCE,\$B135,0,128-\$B135,1),OFFSET(meas,0,0,128-\$B135,1))
136	127	=SIN(B136*3.14/2*180)		
		REFERENCE WAVE	Measurement Wave	

Figure A2. Calculation of correlation function for two waves- Equations.

In Figure A2, cells B9 to B136 show the pixel numbers of the detector (0-127). Cells C9 to C136 are the sinusoidal wave indicating the reference beam. The reference beam is shifted -7 pixels in D2 to D129. The measurement wave is in D9 to D136 range. Correlation function between two waves is calculated in cells F9 to F136. The offset between these two waves can be calculated by finding the first maxima in correlation plot (Cell F5).

	B	C	D	F
2				0
3			-0.142848487	
4			-0.282767028	The Pixel # for first Maxima of correlation function
5			-0.416885772	7
6			-0.54245383	
7			-0.65689569	
8	Pixel #		-0.757864053	
9	0	0	-0.843287969	0.573326739
10	1	-0.142848487	-0.91141532	0.683838401
11	2	-0.282767028	-0.960848755	0.778932334
128	119	0.975582264	0.709553396	-0.441251965
129	120	0.996951681	0.601618195	0.115930595
130	121	0.997872756		0.571154939
131	122	0.9783266		0.811278458
132	123	0.938714119		0.920549159
133	124	0.879847801		0.969493793
134	125	0.802935045		0.991133471
135	126	0.709553396		1
136	127	0.601618195		
	REFERENCE WAVE		Measurement Wave	

Figure A3. Calculation of correlation function for two waves- Numerical values.

Figure A3 presents the numerical values for cross-correlation and its first maxima. Since the offset between these two waves is 7 pixels, the maxima for cross-correlation function is located at pixel 7 (Cell F5). Figure A4 plots correlation function, reference and measurement waves.

Figure A4. Plotting correlation function, reference and measurement waves.

A.2 Error Function. Using data Solver in Excel document, the user can find the number of pixel shift between two waves. Figure A5 is an Excel spreadsheet to calculate error function (Equation 77) and moreover finding the offset by the use of Data Solver. In Figures A5 and A6, cells B9 to B136 show the pixel numbers of the detector (0-127). Cells C9 to C136 are the sinusoidal wave indicating the reference beam. The reference beam is shifted -7 pixels in D2 to D129. The measurement wave is in D9 to D136 range. The shifted measurement wave (P_{n-M}) is generated in cells E9-E136, which reads the measurement wave by use of offset value (M) written in cell H5. Cells F9 to F136 are square of the difference between reference wave (R_n) and shifted measurement wave (P_{n-M}), considering that measurement wave is a portion of the shifted wave depend on the offset value. The error function, which is the summation of square of difference (Equation 77), is calculated in F5. While Figure A5 is showing all the described equations, Figure A6 shows the numerical value of the same sheet when the offset value set to -3.

	B	C	D	E	F	G	H
2		Measurment	=SIN(B9*3.14/2*180)				
3			=SIN(B10*3.14/2*180)				
4			=SIN(B11*3.14/2*180)				
5			=SIN(B12*3.14/2*180)	e(M)=	=SUM(F9:F136)	M=offset=	-3
6			=SIN(B13*3.14/2*180)				
7			=SIN(B14*3.14/2*180)				
8	Pixel #		=SIN(B15*3.14/2*180)	P(n-M)	$(R_n - P_{n-M})^2$		
9	0	=SIN(B9*3.14/2*180)	=SIN(B16*3.14/2*180)	=OFFSET(C9,\$H\$5,1,1,1)	=(C9-E9)^2		
10	1	=SIN(B10*3.14/2*180)	=SIN(B17*3.14/2*180)	=OFFSET(C10,\$H\$5,1,1,1)	=(C10-E10)^2		
11	2	=SIN(B11*3.14/2*180)	=SIN(B18*3.14/2*180)	=OFFSET(C11,\$H\$5,1,1,1)	=(C11-E11)^2		
128	119	=SIN(B128*3.14/2*180)	=SIN(B135*3.14/2*180)	=OFFSET(C128,\$H\$5,1,1,1)	=(C128-E128)^2		
129	120	=SIN(B129*3.14/2*180)	=SIN(B136*3.14/2*180)	=OFFSET(C129,\$H\$5,1,1,1)	=(C129-E129)^2		
130	121	=SIN(B130*3.14/2*180)		=OFFSET(C130,\$H\$5,1,1,1)	=(C130-E130)^2		
131	122	=SIN(B131*3.14/2*180)		=OFFSET(C131,\$H\$5,1,1,1)	=(C131-E131)^2		
132	123	=SIN(B132*3.14/2*180)		=OFFSET(C132,\$H\$5,1,1,1)	=(C132-E132)^2		
133	124	=SIN(B133*3.14/2*180)		=OFFSET(C133,\$H\$5,1,1,1)	=(C133-E133)^2		
134	125	=SIN(B134*3.14/2*180)		=OFFSET(C134,\$H\$5,1,1,1)	=(C134-E134)^2		
135	126	=SIN(B135*3.14/2*180)		=OFFSET(C135,\$H\$5,1,1,1)	=(C135-E135)^2		
136	127	=SIN(B136*3.14/2*180)		=OFFSET(C136,\$H\$5,1,1,1)	=(C136-E136)^2		
		REFERENCE WAVE, Rn	Measurement Wave				

Figure A5. Calculation of error function for two waves- Equations.

	B	C	D	E	F	G	H
2		Measurment	0				
3			-0.142848487				
4			-0.282767028				
5			-0.416885772	e(M)=	21.32685929	M=offset=	-3
6			-0.54245383				
7			-0.65689569				
8	Pixel #		-0.757864053	P(n-M)	$(R_n - P_{n-M})^2$		
9	0	0	-0.843287969	-0.54245383	0.294256157		
10	1	-0.142848487	-0.91141532	-0.65689569	0.264244528		
11	2	-0.282767028	-0.960848755	-0.757864053	0.225717183		
128	119	0.975582264	0.709553396	0.938714119	0.00135926		
129	120	0.996951681	0.601618195	0.879847801	0.013713319		
130	121	0.997872756		0.802935045	0.038000712		
131	122	0.9783266		0.709553396	0.072239035		
132	123	0.938714119		0.601618195	0.113633662		
133	124	0.879847801		0	0.774132153		
134	125	0.802935045		0	0.644704686		
135	126	0.709553396		0	0.503466022		
136	127	0.601618195		0	0.361944452		
		REFERENCE WAVE, Rn	Measurement Wave				

Figure A6. Calculation of error function for two waves- Numerical values.

To find the offset between two waves Data Solver is used. Obviously, changing the offset value changes the error value, and by setting the offset value to -7, which is the pixel offset between the reference wave and measurement wave, the error changes to zero. (Figure A7)

	B	C	D	E	F	G	H
2		Measurment	0				
3			-0.142848487				
4			-0.282767028				
5			-0.416885772	e(M)=	0	M=offset=	-7
6			-0.54245383				
7			-0.65689569				
8	Pixel #		-0.757864053	P(n-M)	$(R_n - P_{n-M})^2$		
9	0	0	-0.843287969	0	0		
10	1	-0.142848487	-0.91141532	-0.142848487	0		
11	2	-0.282767028	-0.960848755	-0.282767028	0		
128	119	0.975582264	0.709553396	0.975582264	0		
129	120	0.996951681	0.601618195	0.996951681	0		
130	121	0.997872756		0.997872756	0		
131	122	0.9783266		0.9783266	0		
132	123	0.938714119		0.938714119	0		
133	124	0.879847801		0.879847801	0		
134	125	0.802935045		0.802935045	0		
135	126	0.709553396		0.709553396	0		
136	127	0.601618195		0.601618195	0		
		REFERENCE WAVE, Rn	Measurement Wave				

Figure A7. The value of error goes to zero for offset value of -7.

The parameters of Solver are set in a way that it finds the offset value (Variable cell) by setting the value of error (Objective) to zero. For this optimization problem the constraint is the value of offset. This value should be an integer between -128 and zero. In addition, to solve the problem the offset value, cell H5, should have an initial value, such as -3 in Figure A5. Figure A8 shows the parameters of Excel Data Solver.

Figure A8. Solver parameters to find the offset value.

B. Zemax Simulation

To simulate the EtchRate sensor functioning, the interference of waves in a wedge is modeled with Zemax Optic Studio software. To begin with Zemax, the non-sequential mode was chosen. The model consists of a monochromatic collimated light, a quartz wedge, and a detector. The wavelength is set to 704nm, the slope of the wedge is $3\mu\text{m}$ in 10mm, and the number of pixels is set to 128 to simulate the parameters in main reference (O'Brien & Jensen, 2014). The wedge is designed by SolidWorks software. Since the SolidWorks modeling accuracy is not in nanometer, which is the resolution of the EtchRate sensor, a thick film layer is added to be able to solve with the preferred wavelength. Figure B1 and B2 show the dimensions of the designed wedge for 300nm coated wedge (Using to simulate the reference beam- before etching) and non-coated wedge (Using to simulate the measurement beam- after etching) respectively.

Figure B1. Schematic for a 300nm coated wedge with the slope of $3\mu\text{m}$ in 10mm. The scale is 10:1 and all units are in mm.

Figure B2. Schematic for a non-coated wedge with the slope of 3 μm in 10mm. The scale is 10:1 and all units are in mm.

In EtchRate sensor, the wedge bottom surfaces are coated with a 99% partial mirror facing toward the wedge. But since in the Zemax software the surface coatings are facing out of the object, the bottom surface of the modeled wedge is designed to reflect 1% and transmit 99%.

In Figure B3 a Zemax software screenshot is illustrated. At the first row in the non-sequential mode, the rectangular source is located with 2-degree tilt. The size of the light source is slightly bigger than the wedge. In the second row, the modeled wedge is imported as a STEP file, and its material is designed to be Fused silica with mirror coated bottom surface. The third row is the detector which is located below the wedge. Its size is the same as the wedge, and it has 128 pixels in the y-axis and 1 pixel in the x-axis. The wedge is placed in a way that its thicker edge is aligned with pixel 1 of the detector. Figures B4, B5, and B6 are different configuration windows in Zemax showing the EtchRate sensor modeling, 3D light paths layout, and the zoom view of 3D light paths layout.

Figure B3. Screenshot for the interference modeling in Zemax software

Figure B4. A simple modeling for EtchRate sensor in Zemax.

Figure B5. The light paths are visible in the Zemax layout configuration.

Figure B6. The zoom view of Zemax layout configuration clearly shows the split of reflected and transmitted rays for both surfaces.

After running the Zemax modeling for 300nm coated wedge (simulation of the reference beam- before etching) and non-coated wedge (simulation of the measurement beam- after etching) the interference patterns are plotted in Figure B7 and B8.

Figure B7. The interference pattern for the non-coated wedge (left) and 300nm coated wedge (right).

Figure B8. The interference pattern across the y-axis for the non-coated wedge (The measurement beam-left) and 300nm coated wedge (The reference beam-right).

By overlaying both measurement and reference beams in one chart (Figure B9), the phase shift between two interference waves is clearly observed.

Figure B9. The interference patterns for simulated reference and measurement waves.

Since in this simulation 13 fringes are observed over 128 pixels, rewriting Equation 79 shows that each pixel corresponds to 35.75 nm of the coated layer:

$$\Delta t = \frac{\Delta m \lambda}{2n} = \frac{13}{128} \cdot \frac{704}{2 \cdot 1} = 35.75 \text{ nm/pixel} \quad (83)$$

Therefore, a 300nm thin film shows about 8-pixel shift between the reference and the measurement waves ($300/35.75=8.39$). Studying Figure 43 shows that decreasing the thickness moves the interference fringes toward the thicker side of the wedge; therefore, it is expected that the reference beam in Figure B9 moves to the left side of the plot because the wedge thicker edge is aligned with pixel 1 of the detector. As it is illustrated in Figure B9 one of the maxima of the measurement beam happens at pixel 44, and there is a peak at pixel 52 for the reference beam. It shows that the reference beam has 8 pixels shift to the left after etching that matches with the calculation.