Loading effects in deep silicon etching

Jani Karttunen[*a], Jyrki Kiihamäki[a], Sami Franssila[b]

[a]VTT Electronics, P.O.B 1101, FIN-02044 VTT, Finland

[b]Helsinki University of Technology, Microelectronics Centre, P.O.B. 3000, FIN-02015 TKK, Finland

ABSTRACT

Feature scale pattern dependencies (microloading and aspect ratio dependent etching) and chip and wafer level loading effects complicate the use of deep silicon etching in MEMS applications. They have major effect on uniformity and etch rate on the wafer scale and on a feature scale. The aim of this study was to find the limitations that these phenomenon set on deep silicon etching. Wafer scale (100 mm or 150 mm), chip scale (1-10 mm) and feature scale (1-1000 µm) structures were etched in pulsed ICP (Bosch process, SF_6/C_4F_8). Etched depths were 10-500 µm, and aspect ratios up to 20:1. Strong dependence of etch rate on loading was observed. On the wafer scale average etch rate was greatly reduced, from 5.4 µm/min (8% load, 2 mm feature size) to 1.7 µm/min (100% load). At same time uniformity deteriorated from excellent 2% to 35% which is too high value for practical applications. Chip pattern density did not affect etch rate on an isolated small chips (2.5 x 2.5 mm^2) but for 10 x 10 mm^2 chip 10% etch rate reduction was seen at high chip scale load. In this case wafer scale etchable area was 6%. We show that feature scale and wafer scale pattern dependies in ICP etching are strongly coupled.

Keywords: Plasma etching, pattern density, uniformity, MEMS, etch rate

1. INTRODUCTION

Single wafer etch processes must be characterized for various pattern effects: macroloading (etch rate depends on etchable area on the wafer scale), microloading (etch rate depends on etchable area on a chip or feature scale), and profile loading (sidewall angle depends on etchable area; typically array lines exhibit profiles different from isolated or edge-of-array lines). Basically they are caused by reactant depletion and can be reduced by reducing the etch rate, increasing the supply of reactants or by better control strategies [1, 2, 3]. In the ion activated, surface reaction limited, regime etching is less prone to loading effects than in mass transport limited regime[4].

RIE-lag (smaller etch rate in smaller features) and ARDE (aspect ratio dependent etch rate or etch rate decrease as a function of etch time for a given line width) are other important pattern effects in etching. In MEMS applications aspect ratios are similar to DRAM trench capacitors, 5:1 to 40:1, but feature sizes are one or two orders of magnitude larger, typically tens of microns. Much of the literature on RIE-lag and ARDE are about submicron structures [2, 5] and only recently has the importance of ARDE been recognized in MEMS-dimension structures [6].

The terms RIE-lag and microloading are often used interchangeably even though they are two are different phenomena. However, sometimes it is difficult to determine which effect is the dominant one: is the etch rate lower because of RIE-lag or because of loading effect? We have designed test structures and chip layouts in order to distinguish the two. We have characterized SF_6/C_4F_8 pulsed (time domain multiplexed) ICP etching (Bosch process) for loading effects at various scales: at feature scale (feature sizes range from tens of microns to hundreds of microns), at chip scale (1 to 10 millimeter) and at wafer scale (1-100% pattern density), and studied interactions between pattern effects.

2. EXPERIMENTAL

Silicon wafers of both 100 mm and 150 mm wafers were used. Both bulk and bonded SOI (Silicon On Insulator) wafers were used. Thermal, TEOS and PECVD oxides were used as masks. Mask oxide etching was done either in CHF3/CF4/He plasma or in BHF. Feature size enlargement due to undercutting in wet etching is usually not an issue for our feature sizes.

[*] Correspondence: Email: jani.karttunen@vtt.fi; Telephone: +358 9 456 6621; Fax: +358 9 456 7012

2.1. Pulsed ICP etching

Pulsed SF_6/C_4F_8 process from Bosch/STS was used[7]. In this process silicon is etched in a fast but isotropic fashion during SF_6 pulse, and C_4F_8 pulse is used to passivate the sidewalls of the etched features. This results in various problems not encountered in traditional etching: pressure is not a variable in the ordinary sense because of large overshoot when gases are switched, and the etch profile is undulating due to pulsing. The baseline process (DEEP) has 13/7 ratio of SF_6 and C_4F_8 pulses and 45 mtorr pressure. Alternative process ORTOB2 has 12/12-pulse ratio and similar pressure as the baseline process, whereas SLOWNEW has 15 mtorr pressure and 5/5-pulse ratio. Our baseline process is maximized with respect to etch rate (7 µm/min maximum etch rate; 4.8 µm/min average rate in etching through the 380 µm wafer), and some profile non-idealities (barreling, retrograde profile) are accepted in trade-off[8].

2.2. Test structures

Macroloading test structure consists of large squares of different sizes (2,4,6,8 millimeter edge length) surrounded by a 100 µm wide trench at 400 µm distance. Wafer stepper lithography was utilized to vary the percentage of exposed area on the wafer. Etched depths were measured from cross section samples by SEM. Sample weighing was also used for etch rate determination. Uniformity of etching for the 100% loading was measured using special wafers with ultralow TTV values of less than 1 µm, and etched thickness was calculated from nominal original thickness from SEM cross section samples.

For microloading, traditional arrays vs. isolated lines structures of various line widths and array sizes were employed. But a new type of structure with nested broken rings, with self-similar shapes but varying local pattern density was used in majority of test runs. The structure is shown on figure 1. Local loading varied from 0.6% to 40%. This test pattern was replicated at three different scales, with maximum edge lengths of 10 mm, 5 mm and 2.5 mm. For a given nested structure, the loading percentage was constant determined by both number and width of trench. The same structure is also used for determination of ARDE, because the shape of trench is unaltered and the line width of inner rings is decreasing.
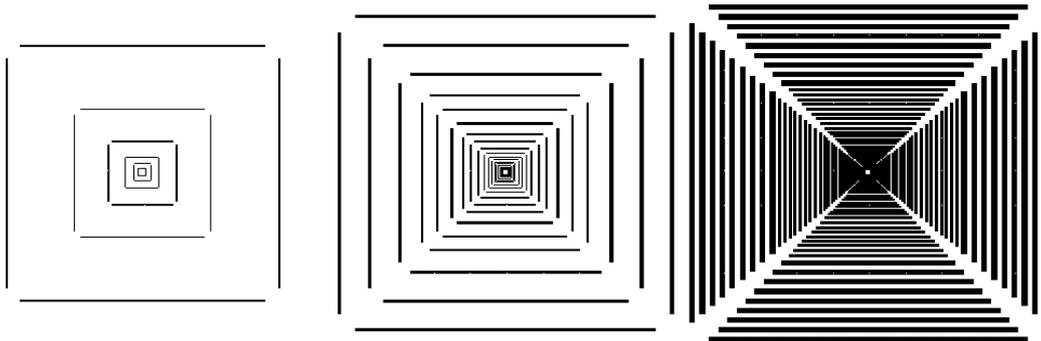
Figure 1 Nested ring test structures with 1.15 % (left), 6.91 % (middle) and 41.46 % (right) local loading.

The etch rate of mask oxide was measured from large area test sites on both macro- and microload layouts. Additionally, it was measured at the center of the nested ring microload structure, thus establishing oxide etch rate at varying local loads. Nanometrics AFT 4150 reflectometer was employed. Selectivity between silicon and oxide was calculated from silicon rates (for a fixed feature size) determined from SEM cross sections and oxide rates measured by reflectometry at adjacent sites.

3. RESULTS AND DISCUSSION

3.1. Macroloading

Wafers with very small etchable area (ca. 1%) etched for a short period approximated zero load. This gives the inherent maximum etch rate available under a chosen set of process parameters. Zero load etch rate was extrapolated to be 7.3 µm/min. Etch rate for 100% load was 1.66 µm/min from mass loss measurement, 1.5 µm/min from SEM cross sectioning.

Results are plotted in figure 2 against inverse of etch rate (ER). Large squares (2 mm) and lines (100 µm) have practically the same etch rate. The fitting line is according to a loading model by Mogab[1]. Even though the model was originally developed for batch reactors, the basic assumptions are valid for any radial reactors. G is the etchant generation rate, β is a proportionality factor to etching species, τ is the mean lifetime of active specie, V is reactor volume, and A_w is etchable area on the wafer.

$$ER = \frac{\beta\tau G}{1 + \beta\tau \dfrac{A_W d}{V}} \quad , \tag{1}$$

The model gives very good fitting. In large loads it has to be considered that etch rate non-uniformity across the wafer is significant (this will be discussed in more detail in following chapters) but in fig. 2 average etch rates across the wafer have been used. Another important consideration is that the maximum aspect ratio was only 3:1. In such low aspect ratio structures ARDE is insignificant and therefore etch rates of 2 mm squares and 100 µm lines are almost identical.
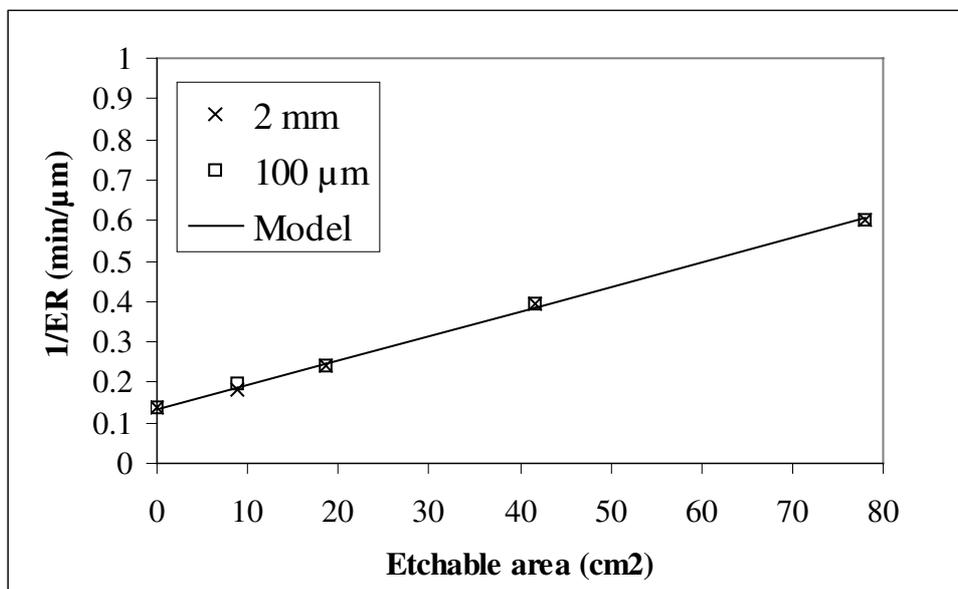
Figure 2 Inverse of etch rate in the function of loading for 2 mm square and 100 µm line and results fitted to model of Mogab.

3.2. Chip scale loading

In figure 3a we present the influence of nested ring size (2.5, 5, 10 mm edge length) on the etched depth of a 32 µm line. The etching time was 90 minutes. For the 2.5 mm rings etched depth is independent of local load percentage, but for 5 mm and 10 mm chips there is a local loading effect. The main reason for deviation of etched depths from trend line is etching uniformity across the wafer. In our etching system etch rate is typically higher at edge of wafer than at center[10].

In fig. 3b etched depths of 100 µm lines were measured from 100 mm^2 chips with a single square pattern in the middle (4, 16, 36, 64 mm^2) (the test structures were on different wafers in 3a and 3b. Etch time was 60 minutes.
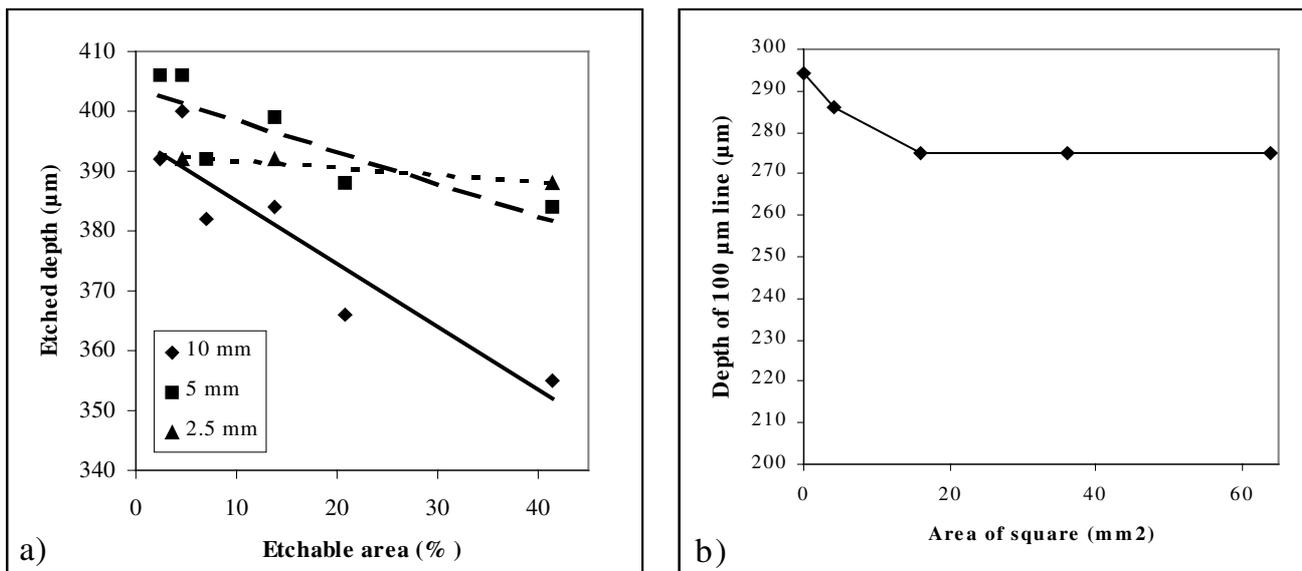
Figure 3 a) etched depth of 32 µm test line for different sized nested rings; b) etched depth of 100 µm test line adjacent to squares with different chip scale load (0 – 64 mm^2/100 mm^2 chip).

For a 2.5 mm chip size chip-scale etchable area does not have any effect on etch rate, for the 5 mm chip there is some reduction in the rate and for the 10 mm chip loading effect is large. From figure 3a we estimate that the reactant depletion distance for our baseline process is therefore of the order of 3-5 mm. Further evidence comes from fig. 3b: the effect is starting to appear when the square edge length is 4 mm.

3.3. Microloading

The traditional microload test structure of isolated and array lines gave a null result: no differences were found. The nested broken rings test structure has a much wider range of local loads, and as shown in figure 4, local load does indeed affect the etch rate. In figure 4a etched depths of different line widths with different local loading are plotted after 80 minutes of etching and in figure 4b identical structure is measured after 10 minutes of etching.

In the case of longer etching time, microloading has a minor influence to etched depth. Feature size effect is the dominant pattern effect. For 10 minutes etch time both microloading and ARDE can be seen and both have major effect on etched depth.

The traditional test structure of isolated and array lines fails to give information on microloading if it smaller than the depletion distance; and if its environment is not representative of real device design. Our nested ring structure involves local pattern density inherently, and it takes much more area but it is an essential structure for process characterization. Other investigators have used constant line patterns with varying spacings[9] but we feel that loading, which is an area effect, requires a truly 2-dimensional test structure. Additionally, the nested broken ring structure lends itself to easy SEM sample fabrication.
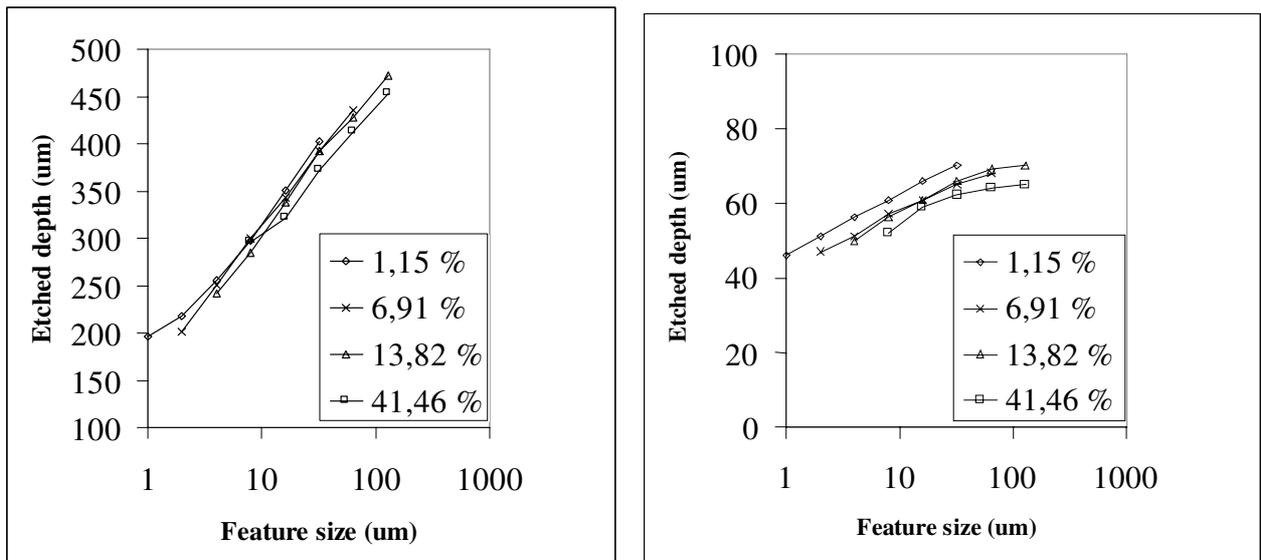
Figure 4 a) 80 min etch time: feature size effect dominates over microloading; b) 10 min etch time: microloading and feature size effect both have major effect on etched depth. Local load varies from 1.15% to 41.46%.

3.4. Uniformity

In figure 5 normalized etch rates of 2 mm squares are plotted as a function of radial distance. The etching time was 60 minutes.
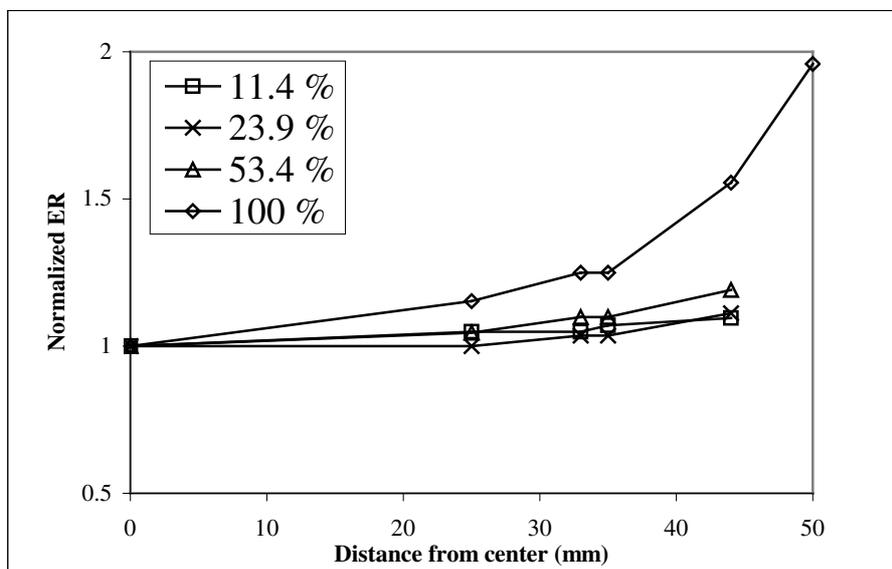
Figure 5 The effect of macroloading (11-100%) on radial uniformity. Normalized silicon etch rates.

From fig. 5 we can see that uniformity was ca. 5% for low load cases but increased to 35% for high loads. This sets serious limitations on device designs with large etchable areas.

3.5. Process variations

In order to improve uniformity, Bosch process modifications with variations in bias power, SF_6 flow, pressure, gas pulse ratios and durations were tried. Results for SF_6 flow and bias changes to our baseline process are shown in figure 6. Decreasing flow improves uniformity at the expense of etch rate, but at higher bias power improvement can be achieved at reasonably high etch rate.
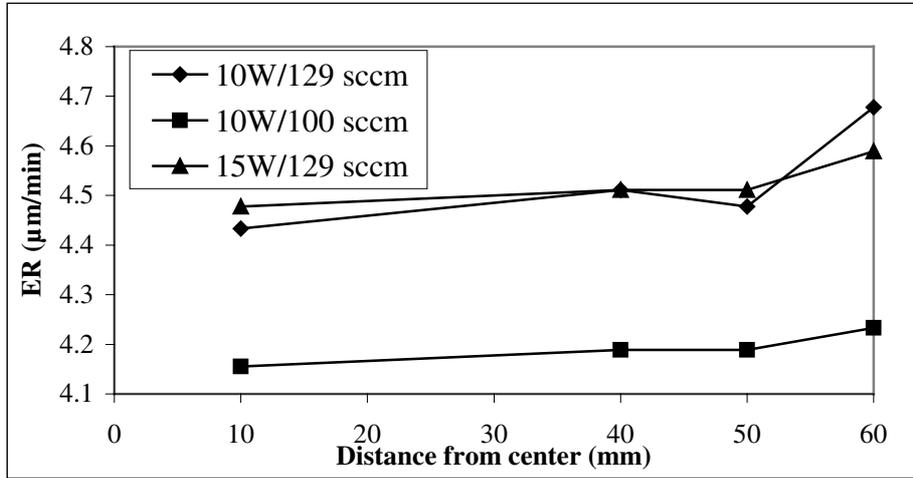
Figure 6 The influence of SF_6 flow and bias power on uniformity of 50 µm lines. Etchable area is 6 % on 150 mm wafer.

In fig. 7 alternative processes ORTOB2 and SLOWNEW[11] have been characterized with regards to uniformity. Improved uniformity has been achieved but at the expense of greatly reduced etch rate, as predicted by the model of ref. 1. In figure 9 DEEP indicates our baseline deep etching process, ORTOB2 is a process optimized for high selectivity and vertical sidewalls and SLOWNEW is a process optimized for small undercut and minimized ARDE.
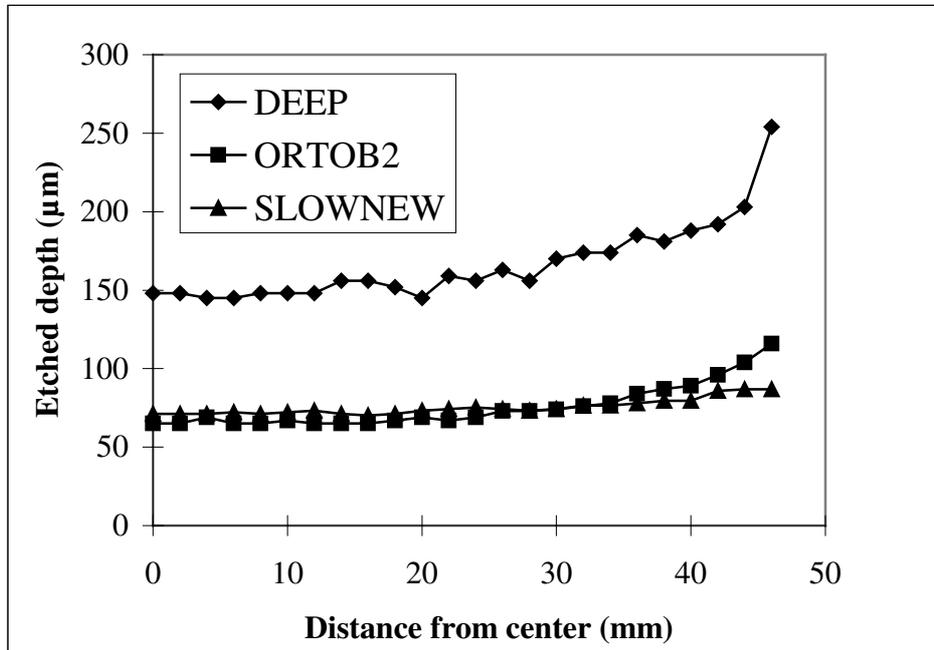
Figure 7 Etched depths and uniformity with three different processes: load 59%, time 60 min.

3.6. Oxide erosion and selectivity

Thermal, TEOS and PECVD oxides were used as masks and all were found to be practically identical as regards to oxide etch rate. Oxide etch rate in the function of silicon etchable area is shown in figure 8. Results are from same wafers than silicon etch rates showed in figure 5.
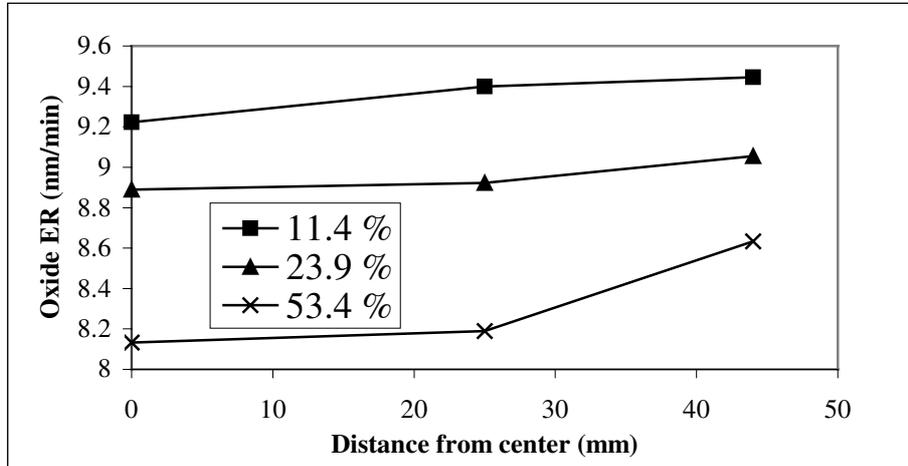
Figure 8 Oxide radial uniformity as a function of silicon macroloading.

It can be seen that oxide etch rate uniformity is much better than silicon etch rate uniformity. Oxide etching is not much affected by silicon load. Therefore oxide selectivity decreases mainly because of silicon ARDE. Because of this, selectivity needs to be defined for specific linewidths. This is plotted in figure 9.
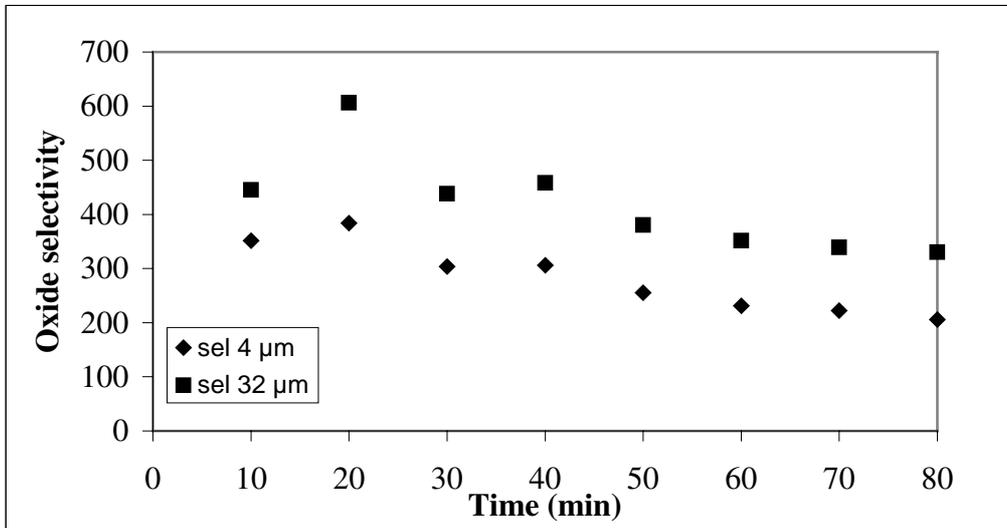
Figure 9 Oxide selectivity as a function of time for 4 µm and 32 µm wide silicon trenches.

Alternative process ORTOB2 with higher C_4F_8 ratio has been studied by us[11]. Oxide selectivities in excess of 1000:1 have been measured. However, the original oxide thickness and final oxide thickness are almost the same, which leads to large relative errors in oxide loss calculation and consequently for selectivity.

Buried oxide thickness needed to stop etching is much more complicated to determine: not only selectivity but silicon etch uniformity, RIE-lag (device design dependent) and oxide ARDE have to be taken into account. Overetch requirement is determined by silicon ARDE in low load case whereas across-the-wafer non-uniformity might be more important for higher loads.

We have observed a phenomenon that could be described as secondary microloading[5]: mask oxide etch rate is affected by local silicon load. This effect has no been seen in all experiments but only in a few series. Large area reference sites and local test sites with <15% local load had identical oxide losses, but at 20% local load there was a clear increase in mask oxide loss and at 40% local load the effect was significant, e.g. 45% increase after one hour etching. This was initially attributed to polymer confounding the oxide thickness measurement by the reflectometer, but remeasurement after oxygen plasma treatment gave unchanged results. We do not at the moment have an explanation to this secondary loading non-reproducibility but we speculate that polymer deposition in the reactor could cause it.

4. CONCLUSIONS AND SUMMARY

It is well known that pulsed (time domain multiplexed) deep silicon etching suffers from serious loading effects typical to reactant transport limited etching. We have investigated the magnitude of the loading effect with test structure design in micro and macroscale. In macroscopic scale the loading effect is unavoidable but in microscale with clever design the adverse effects of loading can be alleviated to some extent. The behavior of loading effect in macroscopic scale is well described by the model of Mogab[1] and the etch rate is dramatically decreased when the etchable silicon area exceeds 20%. Our measurement data on microloading shows that features larger than 2-3 mm or chips larger than ca. 10 mm^2 with high etchable area have an impact on etch rate of adjacent features or features within the same chip. The silicon etch uniformity is severely affected by loading. The radial etch non-uniformity causes extra requirements for etch selectivity against etch stop layer. To be able to correctly predict the etch depth of arbitrary features the models for pattern depended etching[8] and loading effect should be coupled in both micro and macro scales.

ACKNOWLEDGEMENTS

REFERENCES

1. C.J. Mogab, "The loading effect in plasma etching", J.Electrochem.Soc. **124**, pp. 1262-1268, 1977.
2. R.A. Gottscho, C.W. Jurgensen, and D.J. Vitkavage, "Microscopic uniformity in plasma etching", J.Vac.Sci.Technol. B **10**, pp. 2133-2147, 1992.
3. O.D. Patterson and P.P. Khargonekar: "Reduction of loading effect in reactive ion etching using real-time closed loop control", J.Electrochem.Soc. **144**, pp. 2865-2872, 1997.
4. K.P. Giapis, G.R. Scheller, R.A. Gottscho, W.S. Hobson, and Y.H. Lee, "Microscopic and macroscopic uniformity control in plasma etching", Appl.Phys.Lett. **57** (10) p. 983, 1990.
5. K.P. Muller, K. Roithner and H-J. Timme, "Selectivity and Si-load in deep trench etching", Microelectronic Eng. **27**, pp. 457-462, 1995.
6. H. Jansen, M. de Boer, J. Burger, R. Legtenberg and M Elwenspoek, "Black silicon method II: the effect of mask material and loading on the reactive ion etching of deep silicon trenches", Microelectronic Engineering **27** pp. 475-480, 1995.
7. A.M. Hynes, H. Ashraf, J.K. Bhardwaj, J. Hopkins, I Johnston, and J.N. Shepherd, "Recent advances in silicon etching for MEMS using the ASE(TM) process", Sensors and Actuators A, **74**, pp. 13-17, 1999.
8. J. Kiihamäki and S. Franssila, "Pattern shape effects and artefacts in deep silicon etching", J. Vac. Sci. Tech. A **17** pp.2280-2285, 1999.
9. C. Hedlund, H-O. Blom and S. Berg, "Microloading effect in reactive ion etching", J.Vac.Sci.Techol. A, **12**, p. 1962, 1994.
10. K. Richter, M. Orfert, S. Howitz, and S. Thierbach, "Deep plasma silicon etch for microfuidic applications", Surface and Coatings Technology **116-119** pp. 461-467, 1999.
11. J. Kiihamäki, S. Franssila "Deep silicon etching in inductively coupled plasma reactor for MEMS", Physica Scripta. **T79**, pp. 250-254, 1999.