INVESTIGATION OF PROCESS TECHNOLOGY ELEMENTS

EECS 290N Integrated Circuit Technology Design
Fall 1990

by

Rodney Alley, Diane Hoffstetter, Heng-Yuan Hsu,
John Hutchinson, Chih-Chieh King, Angela Lee,
Kristen McNair, Rebecca Mih, Kaveh Niazi,
Stephen Parke, Perry Sheng, Zara Weng,
Kirt Williams, Clayton Yee, Hsi-Jen Yeh

Professors W.G. Oldham and A.R. Neureuther

Memorandum No. UCB/ERL M91/66

30 June 1991

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Silicon Epitaxial Lateral Overgrowth of Oxide

Rebecca Mih, Stephen Parke, Zara Weng

December 15, 1990

EE290n Project Report

Department of Electrical Engineering and Computer Sciences

University of California, Berkeley, CA 94720

ABSTRACT

This paper describes a systematic study of the growth behavior and electrical quality of epitaxial silicon films grown laterally over oxide in the UC Berkeley hot-wall reactor system. The growth behavior is compared over (100) and (110) oriented oxide shapes, consisting of lines, squares, and grids with varying pitches and oxide width-to-seed window ratios. The structural characteristics are investigated with optical, SEM, and TEM microscopy. PN diodes and MOS capacitors fabricated in the epi layer are used to analyze the electrically active defects and material lifetime, by using Electron Beam Induced Current (EBIC) contrasting and MOS deep-depletion transient analysis. The material quality is found to be strongly dependent on the orientation and shape of the buried oxide. Structures with (111) growth fronts exhibited macroscopic growth defects with a resulting high density of electrically active defects. However, structures with (110) growth fronts exhibited high quality overgrown epi with generation lifetimes as high as 31 microseconds. The intrinsic epi film quality was very high with a generation lifetime as high as 407 microseconds.

SECTION I: MOTIVATION

Epitaxial lateral overgrowth (ELO) of silicon over oxide was first demonstrated by Jastrzebski et al [1]. The basic scheme is shown in Fig.1. There are at least two classes of applications that require very high quality crystalline silicon over a buried oxide. The first is high density or high bandwidth structures. Lu et al [2] has described a 64Mb DRAM cell utilizing a trench storage capacitor which is overgrown by an epi layer (Fig.2). The cell's access MOSFET is then built in the ELO layer above the trench, thus reducing the cell surface area. Zingg et al [3] has described a stacked CMOS inverter which utilizes ELO to place the PMOSFET directly above the NMOSFET and thus share the polysilicon gate electrode (Fig.3). A multilayer image processing chip [4] which incorporates a photodiode array in the surface layer, an analog A/D converter array in the center layer, and a CMOS microprocessor on the bottom layer is shown in Fig.4. It demonstrates the density and bandwidth advantages available with 3-D structures. The second class of ELO applications is high performance silicon-on-isulator (SOI) devices. Oxide isolation of CMOS and BJT devices reduces parasitic capacitances, improves density, simplifies the process, and provides immunity to CMOS latchup and ionizing radiation upsets. In addition, it has been shown that extremely thin, fully-depleted SOI MOSFETs have higher transconductance and better short channel effects than their bulk counterparts. Shahidi et al [5] recently described a sub-0.5um SOI CMOS process fabricated by ELO and chemical-mechanical polish back (Fig.5). This concept can also be extended to high-performance lateral BJTs on SOI. The chem-mechanical polishing process is required to obtain a planar epi surface suitable for deep sub-um device fabrication. The SOI film thickness, important for fully-depleted MOSFETs, is controlled by the planarization stop thickness. A thickness of 100nm, with a tolerance of only 10nm, was achieved.

SECTION II: EPITAXIAL LATERAL OVERGROWTH ISSUES

The important issues involved in ELO are listed and discussed below:

1. Selectivity

2. Oxide edge orientation

3. Growth front shape

4. Overgrowth ratio

5. Intrinsic epi film quality

6. Reaction rate vs. transport limited growth

7. Buried oxide/epi interface state density

8. Buried oxide shape effects

1. Nucleation in the seed windows must be highly selective with respect to the exposed oxide surfaces. Polysilicon nucleation on the oxide is inhibited by adding HCl, and reducing growth pressure and temperature [6]. The HCl etches adsorbed nuclei from the oxide surfaces. These nuclei exist at small oxide defects created by non-ideal oxide processing. High temperature epi processes require repeated growth/etch cycles in order to maintain high selectivity. It is extremely important that the oxide surfaces be clean and free of particulates at the onset of epi growth.

2. Dislocations and stacking faults may occur along the edges of the buried oxide shapes. It has been shown that the edge defect density is greatly reduced by a reduction in growth temperature (Fig.6)[6]. The difference in thermal coefficients between oxide and silicon generates stress during the cooling period after growth producing edge defects. These defects are also reduced by increased HCl and lower pressure. For (110) oriented oxide edges, these edge defects can easily glide along the (111) growth front, since this is the silicon slip plane, and meet to form larger defects (Fig.7)[7]. However, defects formed along (100) oriented edges require a higher energy to move along the (110) growth front plane. Thus, the defects remain small and localized. ELO films grown over (100) oriented oxide edges have consistently exhibited much lower edge defect densities than (110) oriented oxide edges, as shown pictorially in Fig.8[8]. A 45 degree mask rotation or the use of wafers with a specially ground (100) flat is required to achieve this (100) orientation.

3. The shape of the advancing growth front must be controlled to prevent the formation of a void at the epi/oxide interface when two growth fronts coalesce (Fig.9). This void is the result of reduced lateral diffusion of the gas being transported to the growth front edge as the opening between the two fronts becomes very small (Fig.10). This problem can be greatly reduced by lowering the temperature to move toward surface reaction rate limited growth and away from mass transport limited growth. Also, for (100) oriented edges, the HCl concentration may be lowered in order to make the (101) growth rate slower than the (001) growth rate, thus keeping the growth front tapered at 45 degrees (Fig.11). This will allow more reactant gas to be transported to the

growth front edge. A steep growth front angle may also create seams through the epi layer where growth fronts meet. These seams should not contain electrically active defects or provide an enhanced dopant diffusion path. Shahidi et al [5] has shown these seams to be electrically inactive.

4. The overgrowth ratio is defined as the ratio of lateral-to-vertical epi growth dimensions. A high overgrowth ratio is desirable in order to fully overgrow device-sized oxide shapes without an excessive epi thickness. This ratio is a complicated function of the HCl chemistry and the oxide and seed window sizes, particularly if the growth is not reaction rate limited. Fig. 12 shows the geometrical interaction of the overgrowth ratio (R) and the growth front slope (F). R and F are extracted from SEM cross-sectional micrographs by measuring the epi thicknesses (T_{MAX} and T_{MIN}) and overlap distance (X_{OV}) for a given width oxide line.

5. The intrinsic epi film should have a minority carrier lifetime equal to or higher than the starting substrate. This is not a difficult issue in modern epi reactors.

6. It is desirable to maintain surface-reaction-rate-limited epi growth in order to minimize "loading" effects of both the reactor geometry and the local oxide pattern itself. Fig. 13 shows that a temperature of 850C is sufficient for a 100um oxide opening in a 30Torr system[9]. However, even lower temperature and/or pressure is required for sub-micron seed windows. At 850C and 1Torr, it is possible that small oxide shapes will exhibit a reduced overgrowth ratio and a sensitivity to pattern size.

7. The buried oxide/epi interfaces are of particular concern since they are not formed by conventional thermal oxidation and may have a higher interface state density. Different results have been reported to date. Junctions in contact with this interface had a leakage as high as 1E-9 A/um compared with 1E-13 A/um elsewhere (Fig.14)[10]. This leakage was reduced by (100) alignment, lower pressure, lower temperature, lower HCl, and gettering. On the other hand, Friedrich & Neudeck [11] have reported a surface state density of only 2E11/cm2/eV. They attributed this small number to the large H2 flow in the reactor during growth, which passivates the dangling interface bonds.

8. The interaction of different directional growth rates with the oxide shapes below can generate growth front collisions and stress leading to dislocations and stacking faults. In addition, the oxide corners will always be somewhat rounded due to photolithographic limitations, implying that the growth front passes through a continuum of planes at the corners. Some of these planes will have a faster growth rate than others, leading to a

complex situation. Fig. 15 shows the expected growth behavior for the oxide structures and orientations fabricated in this experiment. In particular, the (100) oriented oxide grids are expected to close slowly, with four (111) planes and associated defects coming together at each centerpoint. Meanwhile, the (110) oriented grids have (111) planes meeting along the entire length of each centerline in the x and y directions. The (110) oriented lines have a similar problem, except in one dimension only. These three structures are expected to have significant oxide edge generated defect densities.

SECTION III: EXPERIMENT DESCRIPTION

This experiment was designed in order to address each of the epi growth issues just described, although time did not permit all of these to be studied in sufficient detail. The experiment was based on the experience and results previously attained by C. Galewski, J. Lou, and W. Oldham with the UC Berkeley hot-wall reactor [12]. Since very high quality ELO films as thick as 3um had previously been attained, the epi growth recipe was not altered. Our strategy was to minimize process splits (the only split was epi thickness) while providing a wide variety of buried oxide structures on the test die.

Epi Processing

Two starting wafer types were used. E-series wafers were specially manufactured by Wacker Corp. These were (100) surface, CZ-grown, boron-doped wafers with 15-25 ohm-cm resistivity. The major flat was specially ground to the (100) direction. Reorder information is: SPEC# CZ04P, Product Line PB83, Product #s 3940/1&2, WA# 51903. D-series wafers were standard boron-doped prime wafers with identical specs, except that the major flat was along the usual (110) direction. Figs. 16 & 17 detail the experiment plan and process crossections. The starting wafers were cleaned and HF vapor treated as described by Galewski et al[12]. An initial 3.5um epi layer was grown so that the p substrate/n- epi junction would be well below the buried oxide. The epi process consisted of a 15 minute 1000C H2 prebake followed by an 8 hour growth at 850C. The growth flowrates were DCS=28sccm and H2=1000sccm at a pressure of 1 torr. Two wafers were set aside after epi#1 for use as intrinsic epi lifetime monitors. Next, a 5 hour dry oxidation at 1050C was performed, growing 196nm of buried oxide. I-line photolithography was then used to define the buried oxide. A 25% overexposure was required in order to compensate for problems with the emulsion mask. The overexposure eliminated most

0.5um squares from the wafers. At the same time, most square photoresist openings of 1.0um were still not exposed enough to be properly defined. Therefore, experiments with these structures were rendered non-functional. The oxide was then plasma etched using anisotropic conditions to achieve vertical sidewalls. The etch was done at 900W with CF4=40sccm and CHF3=45sccm for 30-35 seconds. Some wafers exhibited a "bullseye" pattern, with the center of the wafer underetched. In subsequent ELO growths of up to 12 hours, no epi was grown in the central oxide circles on these wafers, providing proof of the high selectivity of this process. Next, a sacrificial oxidation and etch was done to remove the plasma damaged silicon. The 70 minute, 900C oxidation yielded 19nm. After the HF wet etch, the BOX thickness was found by profilometry to be reduced to 170nm. The epitaxial lateral overgrowth step was performed with conditions identical to epi#1, except the H2 prebake temperature was lowered to 900C to avoid oxide undercutting. Growth times of 8, 10, and 12 hours were used, resulting in epi thicknesses of 3.5, 5.2, and 6.2 microns respectively. Step profiles measured over an isolated epi growth front are shown in Fig. 18. The growth rate initially increases with epi thickness and then saturates at 0.52um/hour for thicker epi. One or two wafers of each ELO thickness were set aside for microscopy.

Device Processing

In order to perform electrical film characterization, a simple, non-LOCOS process was used to build large area PN diodes, MOS capacitors, and PMOSFETs in the epi, over arrays of small buried oxide shapes. This process utilizes an n+ polysilicon ohmic contact to the as-grown n- epi surface and a low-energy BF2 implant to form p+ junctions. This implant depended on the gate oxide and polysilicon to serve as an implant mask. Double thickness (2um) KTI photoresist with 3.0 sec exposure time was necessary for all lithography steps following the epi process due to the 3.5-6.2um epi steps existing on the wafers. A 92nm gate oxide was grown at 950C for 6 hours. The gate oxide was patterned with the KTI process, followed by a two minute 5:1 BHF wet etch. The gate oxide mask was aligned to the existing BOX alignment mark. This alignment mark was somewhat difficult to use with accuracy due to the thick epi grown over it and associated facetting. A gross misregistration occured in regions of some wafers due to inadvertant interruption of the wafer stepper's chuck vacuum by the operator. A 20 second 10:1 HF dip immediately preceded the CVD deposition of 350nm of phosphorus-doped gate poly. The deposition conditions were: 650C, 407mtorr, 60 minutes. The n+ poly was patterned with the KTI process. As before, gross misregistration occured in some areas of some wafers. Wafer maps of

correctly aligned dies were created to aid in subsequent characterization. The gate poly was then plasma etched using a standard chlorine-based recipe. All wafers, except for one control, were implanted with 5E15/cm2 BF2 at 50keV without a prior sacrificial oxide. This implant was followed by a 20 minute N2 anneal at 900C. All wafers except one were electrically tested by directly probing the n+ poly and p+ diffusions. 500nm of aluminum was sputtered on the remaining wafer, and it was patterned and wet etched to create large metal pads for EBIC wirebonding.

Test Chip Design

A test chip was designed which consists of 11 main experiments, 12 small device array experiments, a large area diode and capacitor, SEM lines, and alignment marks, as shown in Fig. 19. Each main experiment contains four stripes, with a different buried oxide structure in each stripe: lines, grids, squares, and no oxide. A 200x200um PN diode, a 200x200um MOS capacitor, and a 200x200um PMOSFET, shown in Fig. 20, are built across each stripe, for a total of 12 devices in each main experiment. The experiments are designed with oxide pitches of 1.5, 2.0, 2.5, 3.0, 3.5, and 4.0um. With the exception of the 1.5um pitch, each pitch has two experiments associated with it: one with minimum oxide width and the other with minimum oxide space. The top row of five experiments has the oxide width fixed at 0.5um, with the seed window increasing from 1.5 to 3.5um moving from left to right. The middle row of five experiments has the seed window fixed at 1.0um with the oxide width increasing from 1.0 to 3.0um moving from left to right. The 3x4 set of small device array experiments contain large metal pads connecting an array of minimum size devices which are specifically oriented over seed areas, epi seams, or epi centerpoints depending on the experiment. These structures are meant to spatially locate any lifetime-limiting defects. They are only testable after metallization. A 400x400um MOS capacitor and 400x400um PN diode were included for DLTS and EBIC measurements. These two devices are built over 0.5um BOX lines with 3.5um seed windows. Also on the test chip is a set of 2mm long SEM lines with oxide linewidths continuously increasing from 0.5um to 100um. These are used to observe the growth front in various stages of coalescence. For each linewidth, a set of three identical lines with 1.0um seed windows were fabricated, plus an isolated line of the same width spaced a distance away.

SECTION IV: STRUCTURAL CHARACTERIZATION

Optical microscopy, scanning electron microscopy (SEM), and transmission electron microscopy (TEM) were used to characterize the structural perfection of the epi overgrowth. Optical microscopy revealed the macroscopic factors affecting epi overgrowth, such as the buried oxide orientation, shape, and width and the seed window size. Cross-sectional SEM was instrumental in understanding growth mechanisms and providing higher resolution detail of the topography of the overgrown layer, as well as the shape of the growth front and the structure of growth defects. Planar TEM was used to identify the type and crystallographic orientation of growth defects found along the buried oxide edges. Finally, SEM Electron Beam Induced Current (EBIC) was very useful in precisely locating electrically active structural defects and observing their density.

Optical and SEM Observations

There are several general comments to be made about the epitaxy. Most importantly, the (100) oriented wafers had far fewer buried oxide-generated defects than the (110) oriented wafers. Secondly, with either orientation, the epi grown over oxide grids had a jumbled, mostly random topography. The shape of the growth front was observed at various degrees of overgrowth using crossectional SEM through the designed SEM lines. Figs. P1-P3 show the 5um epi overgrowth over successively wider oxide lines. Fig. P4 shows the growth front all the way down to the buried oxide surface. From these photos, the growth front appears to be nearly a (110) plane with a slight amount of concavity. Also, there is a 300nm high vertical knee along the front's advancing edge. An overgrowth ratio R=10, and a growth front slope of F=1.2 were extracted from these photos. This means that an epi thickness of only 1.6um was required for growth fronts to meet over 3um wide oxide lines. However, 15um of epi would have been required to achieve total planarity over these lines.

Large, shallow crystallographic pits with varying densities were observed on most wafers. These pits, shown in Fig P5, are formed during epi growth and have dimensions on the same order as the epi thickness. A few also exhibit a very peculiar set of ridges or facets within the pit, as seen in the SEM micrograph. Surprisingly, none of these large pits was directly observed over a buried oxide pattern. The source of these pits may be SiO2 particulates on the original silicon surface which provide a local source of oxygen during growth. This oxygen combines with silicon atoms to form SiO gas which then evaporates, transporting silicon mass away from the surface.

Overall, for our reactor conditions, it is apparent that oxide width and shape are more important factors

in the ELO quality than seed window size. This indicates that a nearly reaction-rate-limited growth was achieved. However, closely spaced oxide lines did exhibit noticeably less lateral overgrowth than lines of the same width which were spaced further apart. This was observed in the SEM structure with 1um seed windows compared with 5 to 10um. This effect can also be seen in Fig. P6, where the epi thickness decreased slightly from the edge of a line structure (right) toward the center of the structure (left). The seed windows are 1.0um, with 2.5um wide oxide lines. This is a rather small effect, and in general it was observed that as long as the seed windows were wide enough to be uniformly etched through, good epitaxial overgrowth occurred.

The predicted effects of inside and outside oxide corners on ELO quality were observed in this experiment. Figs. P7-8 show photos of the full experiment set for a 5.2um epi E-series wafer (100 oriented). Buried oxide lines and squares consistently exhibited good overgrowth, while the oxide grids appeared somewhat jumbled, especially for the smallest seed windows (left side of P7). Photolithographic corner rounding served to complicate these effects. Figs. P9-10 show the patterned oxide shapes before epi overgrowth, revealing that both oxide squares and square openings became quite circular during processing. Although the epi over the oxide grids is obviously defective, it is difficult to determine the nature of these defects. SEM micrographs of these grids (Fig. P11) show that the overgrowth contains many truncated triangular growth defects (labelled by A's). These defects appear to start along a (110) line that intersects the center of a seed window below. The growth then propogates in the orthogonal (110) direction ending in another (110) line. Meanwhile, the sides of the defect are confined by (100) planes. Although this structure is unexpected and not yet understood, the direction is right for these defects to be associated with the (111) slip plane phenomena. In contrast to the grids, oxide squares of the same size were smoothly overgrown without defects. Fig. P12 shows that a small rounded dimple with (110) sidewalls appears above the center of each oxide square below.

The effect of wafer orientation on ELO quality was also evident in this experiment. Wafer E8 (100) and D5 (110) were part of the same 3.3um ELO epi run. Several major differences are evident in Figs. P13-P15. First, on D5 the minimum width oxide lines were well overgrown but appear to be segmented with ridges or facets orthogonal to the oxide lines. This behavior was not observed on wafer E8. Meanwhile, the oxide grids on wafer D5 look completely random, while those on E8 exhibit periodic grids for larger seed windows. In the oxide width experiment, there is a marked difference between the two wafer orientations. On wafer D5, the

wider oxide lines have large variations in overgrowth along the length of the line. The previously mentioned orthogonal ridges appear to be associated with the non-uniformity of the growth fronts. On wafer E8, however, the growth fronts were more uniform with none of the orthogonal ridges. The oxide squares exhibited good overgrowth on both E8 and D5. However, the squares on E8 exhibit two crossing major diagonals, while these are visibly absent on D5 oxide squares. This growth behavior is predicted by Fig. 15.

TEM Observations

The following planar view TEM micrographs of the epitaxial silicon were taken on a sample with 0.5um oxide lines oriented along the (100) direction. The ion milling (thinning) of the sample resulted in areas with no SiO2 remaining, since the milling rate of amorphous SiO2 is faster than that of crystalline silicon. Light areas are of overgrown single crystal, while the black areas are where the SiO2 used to be. The white areas are the silicon regions which were too thick to be transparent to the electron beam. The pictures were taken on a Phillips 301 microscope. All samples were tilted so that the diffraction patterns would exhibit the closest <100> orientation. Thus, any misorientation of the diffraction pattern is not caused by the varying sample thickness, but by stacking faults in the sample.

Figure TEM1

This picture shows the region surrounding a 0.5um oxide line. Along the oxide line edges, where the silicon dioxide used to be, there exists little triangular shaped microfacets. In the center of this photo is a large, somewhat triangular region which appears to be some type of line dislocation. From the Bragg diffraction pattern, the area appears to be oriented in the (100) direction with some slight misorientation indicated by the small secondary points close to the (100) points. This indicates that the defect may be a stacking fault.

Figure TEM2

Once again, there exists triangular regions in the bottom center and upper right of this picture. The silicon in this region generally exhibits (100) orientation.

Figure TEM3

Here, line dislocations are shown to exist in the epitaxial silicon growing over the oxide lines. The

diffraction pattern indicates a larger number of slight misorientations about the (100) direction, shown by the blurred diffraction points.

Figure TEM4

This TEM sample shows the region around a sharp oxide corner. The "cloudy" structures can be clearly identified as polysilicon by the diffraction pattern. Although this polysilicon could have been produced by the ion milling of the sample, these structures appeared several times throughout the sample, even along thick edges such as this one.

SEM-EBIC Observations

The scanning electron microscope electron beam induced current technique was used to correlate observed structural defects with the electrical behavior, in particular, with the carrier lifetime. It is well known that some structural defects, such as twin boundaries, are not necessarily electrically active. The goal was to characterize the epi quality, observe the defect state of the buried oxide/epi interface, and to determine if merging growth front "seams" displayed electrical activity.

Measurement Procedure

To perform the EBIC contrasting, the sample must either have a PN or Schottky junction. As the electron beam sweeps the sample surface, electron-hole pairs are generated. The beam-generated carriers give rise to a current in the junction. This current is fed into an amplifier, and the voltage output of the amplifier is fed into the CRT of the SEM display. Thus, the image in EBIC mode represents the electrical activity of the junction. Dark regions in the EBIC image are indicative of higher recombination current (less current is delivered to the load) due to defects in the material. The spatial resolution of EBIC is limited by the generation volume of the beam, which is approximately 1-3 microns wide. This technique is one of the few which can provide correlation between electrical data and structural characterization.

The SEM must have electrical feedthroughs, and the SEM CRT must be able to accept external inputs. Samples must have metal contacts. The easiest way to connect the diode to the current amplifier is to use a small sample stage which is attached to the SEM chuck. The sample is mounted on a standard 2-pin mount, and the pins are

wirebonded to the diode metal pads. It is important to test the diodes before loading them into the SEM. High leakage currents (100uA) are acceptable if the current amplifier has DC suppression capability to offset the leakage current. The accelerating voltage must be chosen with care, as this will determine the penetration depth of the beam, and therefore, the location of carrier generation. Carriers must be generated near or in the depletion region.

Results

Three diodes from wafer E8 were analyzed, each with a different buried oxide shape. The preliminary results are shown in Figs EBIC1-3, which contain both the standard SEI image and its corresponding EBIC image. The seed windows appear bright in the SEI image and dark in the EBIC image. Meanwhile, the buried oxide areas appear dark in the SEI image and white in the EBIC image. In this preliminary work, absolute current (contrast) levels from the diode were not measured. Thus, we are unable to make comparions of contrast levels between different photographs. In some samples the EBIC showed relatively light contrast over oxide areas, which indicates that growth over oxide is relatively free of "lifetime killers".

Fig. EBIC1 shows 2.5um oxide lines with 1.0um seed windows. Two types of defects are observed. First, large electrically active growth "segments" appear in several places. These appear light in the SEI image and dark black in the EBIC image with a one-to-one correspondence. They have very sharply defined sides, and the dark contrast in the EBIC image indicates enhanced recombination current at these defects. These defects fall predominantly over the seed windows, but some overlap onto the adjacent oxide line. The second type of defect seen is point defects, appearing as tiny black dots in the EBIC image. The dots have black contrast because the beam generated carriers are recombining through the point defects instead of being collected by the junction. These consistently appear only along the edges of the buried oxide lines. It is interesting to note that the generally uniform brightness of the buried oxide in the EBIC image is indicative of low surface recombination at the oxide/epi interface. Fig. EBIC2 shows 2.5um oxide squares with 1.0um seed window grids. The epi was thinner in the dimple at the center of each oxide square, and appears white in the EBIC image. No gross defects are observed in this structure and only a few point defects are observed. Finally, Fig. EBIC3 shows an oxide grid experiment. As seen before, the epi surface appears very rough, and not surprisingly, the EBIC image shows many large randomly-oriented electrically active regions (current sinks).

SECTION V: ELECTRICAL CHARACTERIZATION

General Measurements

The PN diodes exhibited a reverse bias breakdown voltage of only -2 Volts, as shown in Fig. 21. This was caused by p+ junction to p substrate punchthrough. This occured, even though the total epi thickness was around 8um, due to the very light background doping in the epi (3E13). Conventional MOS CV (Fig. 22) was used to determine the gate oxide thickness, epi doping, and fixed charge density. From Cox, a tox of 92nm was extracted, which agrees well with the value measured by ellipsometry. The extracted n-type epi doping concentration ranged from 3.3E13 to 1.6E14/cm3 across all the wafers. This value was lower than expected and indicates a very clean reactor system. For these low doping levels, a gate voltage of only -2.0V will create a steady-state depletion depth of 4.8-5.8um, and at Vg=-3.0V this will increase to 6.2-7.8um depth. It is clear that a small applied voltage can result in a depletion depth deep enough to enclose the buried oxide shapes. A problem with the MOS capacitors was found on all wafers which had received the boron junction implant. This problem, explained in Fig. 23, was due to implant-generated oxide traps. It did not prevent useful measurements from being taken, however. The flat band and threshold voltages were, -0.4V and -1.0V, respectively. These indicate very little fixed oxide charge in the gate oxide.

Lifetime Measurements

When a MOS capacitor is pulsed from accumulation into deep depletion, a large depletion region is instantaneously formed, which then recovers according to the bulk and surface generation rates of the minority carriers. This rate of recovery may be observed by measuring the capacitance of the MOS capacitor as a function of time. Fig. 24 shows an example of this recovery for both short and long lifetimes. The Zerbst analysis of this C-t response allows for the extraction of both the bulk generation lifetime and the surface generation velocity at the SiO2/epi interface. In our experiment, we tried to choose a gate voltage pulse such that the deep depletion region depth would not touch the buried oxide interfaces, since these would add additional surface generation and complicate the Zerbst analysis. The lifetimes measured over various buried oxide shapes and orientations provided a measure of the ELO film quality. The lifetime and surface generation velocity are determined from the slope and intercept of the Zerbst plot, as shown in Fig. 25. Most of our C-t data exhibited sur-

face limited-recovery as opposed to bulk-limited recovery. The intrinsic epi quality was measured on wafer E18, which contained no buried oxide (Fig.26). The surface generation velocity ranged from 0.11 to 0.18cm/sec. The bulk generation lifetime ranged from 264 to 407 microseconds, which indicates a very high quality epi film.

Wafer E13 has 5.75um of epi over the buried oxide, with an epi doping concentration of 2E14. For an applied gate pulse of -2V, the depletion depth should be less than 4.25um. Therefore, using this gate pulse and wafer, we believe the deep-depletion region did not reach the buried oxide patterns. Zerbst plots were constructed for each type of buried oxide shape: lines, grids, and boxes (example in Fig.27). In all cases, the lifetimes were lower than the intrinsic lifetime from wafer E18. For, overgrown lines, the lifetime increased linearly from 0.8usecs for 1.0um seed windows to 3.4usecs for 3.5um seed windows (Fig.28). This implies either: 1)Higher quality epi growth occurred with increasing seed window size, or 2)The depletion did reach the BOX interface and the lifetime exhibits a linear improvement as the amount of BOX area is reduced. For a fixed seed window and oxide width decreasing from 3.0um to 1.5um, the lifetime increased from 2.7usecs to 5.2usecs. This indicates that the ELO film quality decreases as wider oxide lines are overgrown. For oxide squares, the lifetime ranged from 2.6usecs, for 3.0um squares, to 31.0usecs for smaller 1.0um squares. This is the same dependence as that of the lines, but the lifetimes over oxide squares were generally several times higher than that over lines. In almost all cases, the lifetime over oxide grid patterns was too short to be measured, indicating very poor quality film over these patterns.

These lifetime measurements agree qualitativly with our structural characterization and provide useful quantitative information as well.

SECTION VI: CONCLUSIONS

Device quality epitaxial films with 31 microsecond lifetimes were successfully grown over buried oxide patterns. Oxide lines and squares, which contain only inside corners, exhibited high quality overgrowth when these shapes were aligned with a (100) direction. (110) aligned lines exhibited a high electrically active defect density in the overgrown film. Meanwhile, all oxide grids, which contain outside corners, exhibited very poor epi overgrowth and correspondingly low minority carrier lifetime. This was attributed to the coalesce of (110)

growth edges containing defects. Although, the quality of the buried oxide-to-epi interface was not directly characterized, both EBIC and deep-depletion C-V results indicate that this interface has a relatively low interface state density.

SECTION VII: ACKNOWLEDGEMENTS

We would like to thank Prof. William Oldham and Carl Galweski for helpful discussions and guidance. We especially thank J. Lou for his many hours of help in growing the epitaxial films. Tai Nyugen was instrumental in providing TEM support, as was Jeff Bailey for providing the EBIC equipment.

References

[1] L. Jastrzebski, et al, "Device characterization on monocrystalline silicon grown over SiO2 by the ELO (epitaxial lateral overgrowth) process," *IEEE Electron Device Letters, EDL-4, No 2, Feb. 1983.*

[2] N.C.C. Lu, et al, "A Buried-Trench DRAM Cell using a Self-aligned Epitaxy Over Trench Technology," *IEDM Digest, p. 588, Dec. 1988.*

[3] R.P. Zingg et al, "Novel fabrication technique for dual-gate MOS transistors," *Microelectronic Engineering, Vol 10, No 2, p. 115, Jan. 1990.*

[4] W. Krull, *IEDM SOI Short Course, Dec. 1990.*

[5] G. Shahidi et al, "Fabrication of CMOS on ultrathin SOI obtained by epitaxial lateral overgrowth and chemical-mechanical polishing," *IEDM Digest, p. 587, Dec. 1990.*

[6] L. Jastrzebski et al, "Issues and problems involved in selective epitaxial growth of silicon for SOI fabrication," *J. Electrochemical Society, Vol 136, No 11, p. 3506, Nov.* 1989.

[7] H. Yamamoto et al, "Lateral solid phase epitaxy of evaporated amorphous Si films onto SiO2 patterns," *Silicon-on-Insulator: Its Technology and Applications, KTK* Scientific Publishers, p. 187, 1985.

[8] T. Kamins "Silicon-on-insulator (SOI) technologies for integrated-circuit applications," *UC Berkeley, Oct. 1983.*

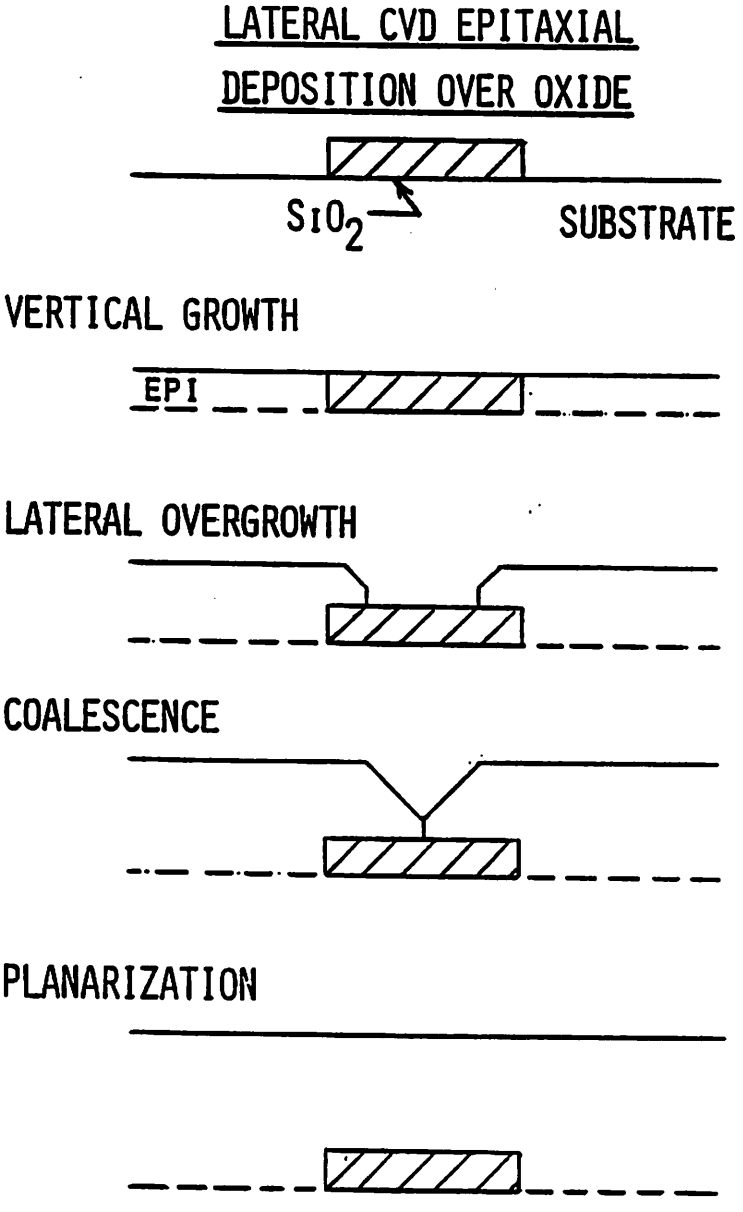
[9] R. Pagliaro, et al, *J. Electrochemical Society, May 1987.*

[10] L. Jastrzebski, et al, "Chemical Vapor Deposition 1987," *Proc. of the Electrochemical Society, PV87-8, p. 334, 1987.*

[11] J. A. Friedrich and G. Neudeck, "Interface characterization of Si epitaxial lateral growth over existing SiO2 for 3-D CMOS structures," *IEEE Electron Device Letters, Vol 10, No 4, p. 144, Apr. 1989.*

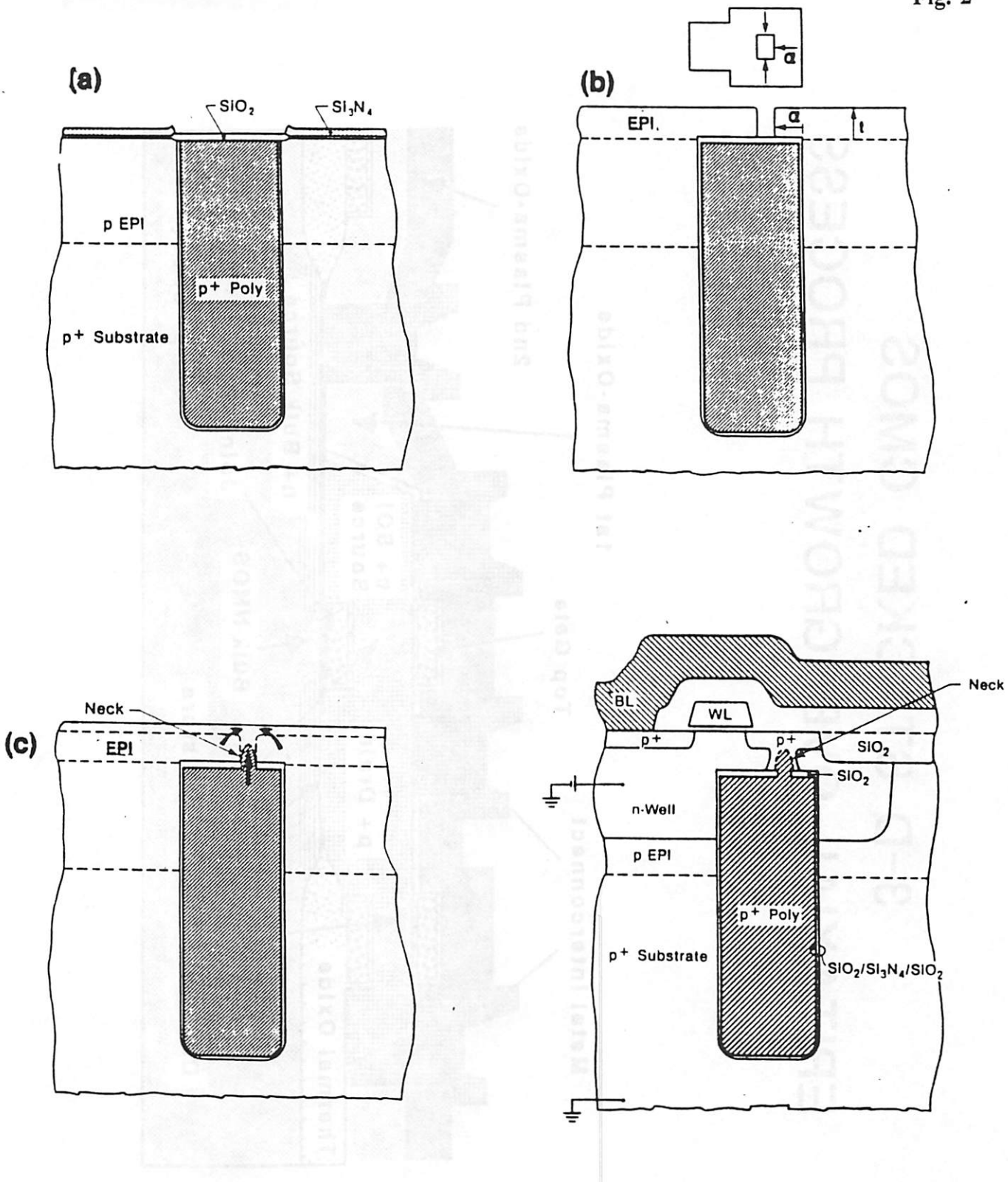
[12] C. Galewski et al, "Silicon Wafer Preparation for Low-Temperature Selective Epitaxial Growth," *IEEE Transactions on Semiconductor Mfg., Vol 3, No 3, Aug 1990.*

Fig. 1



REF: D D RATHMAN, ET AL, J. ELECTROCHEM. SOC. 129, 2303 (OCTOBER 1982)
L JASTRZEBSKI, ET AL, J. ELECTROCHEM. SOC. 130, 1571 (JULY 1983)

Fig. 2


(a)
SiO_2
Si_3N_4
p EPI
p+ Poly
p+ Substrate
(b)
a
EPI
a
t
(c)
Neck
EPI
BL
WL
Neck
p+
p+
SiO_2
SiO_2
n-Well
p EPI
p+ Poly
p+ Substrate
$SiO_2/Si_3N_4/SiO_2$

3-D STACKED CMOS EPITAXIAL OVERGROWTH PROCESS

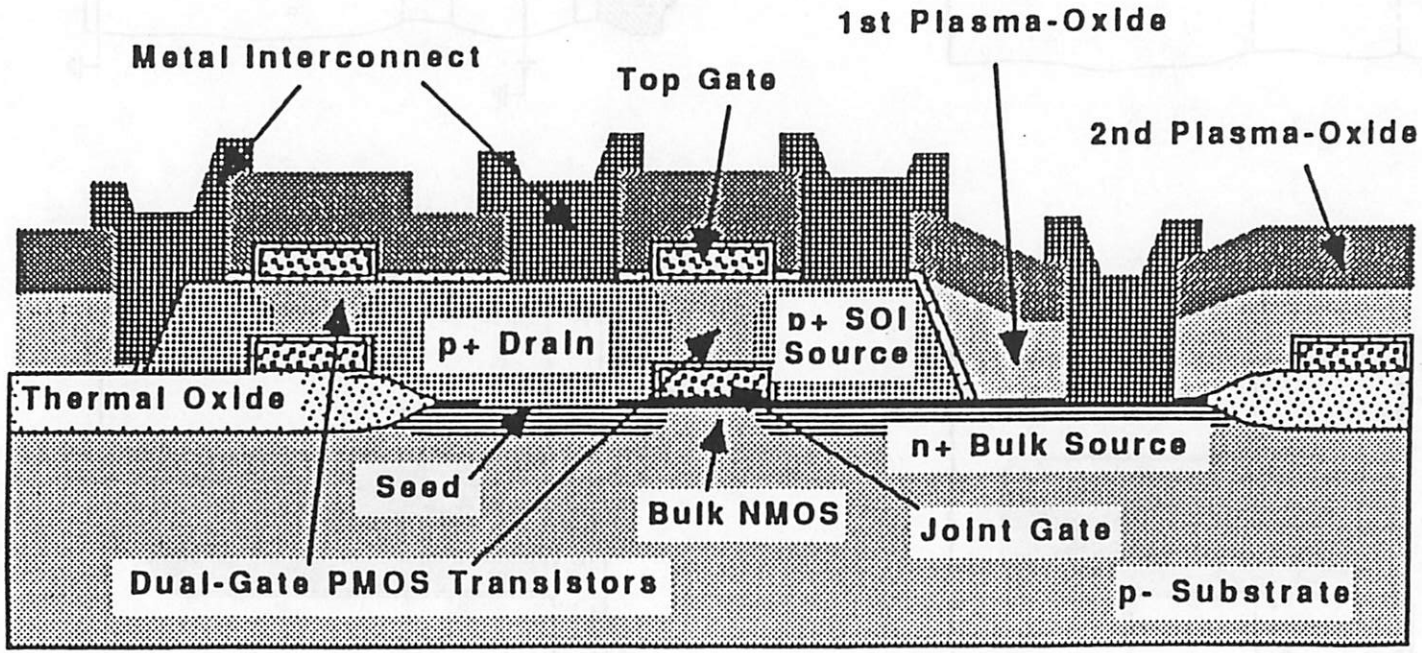
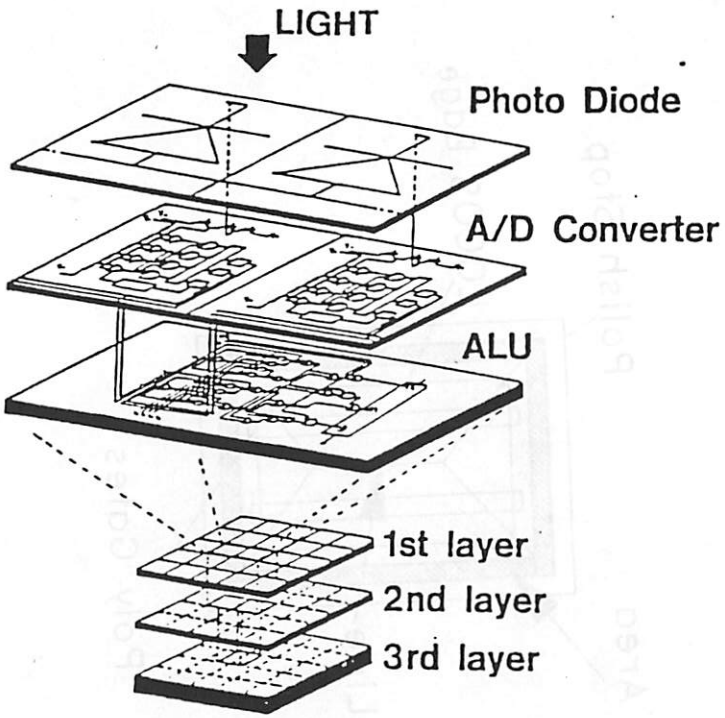


Fig. 3

IMAGE PROCESSOR IC IN 3-D



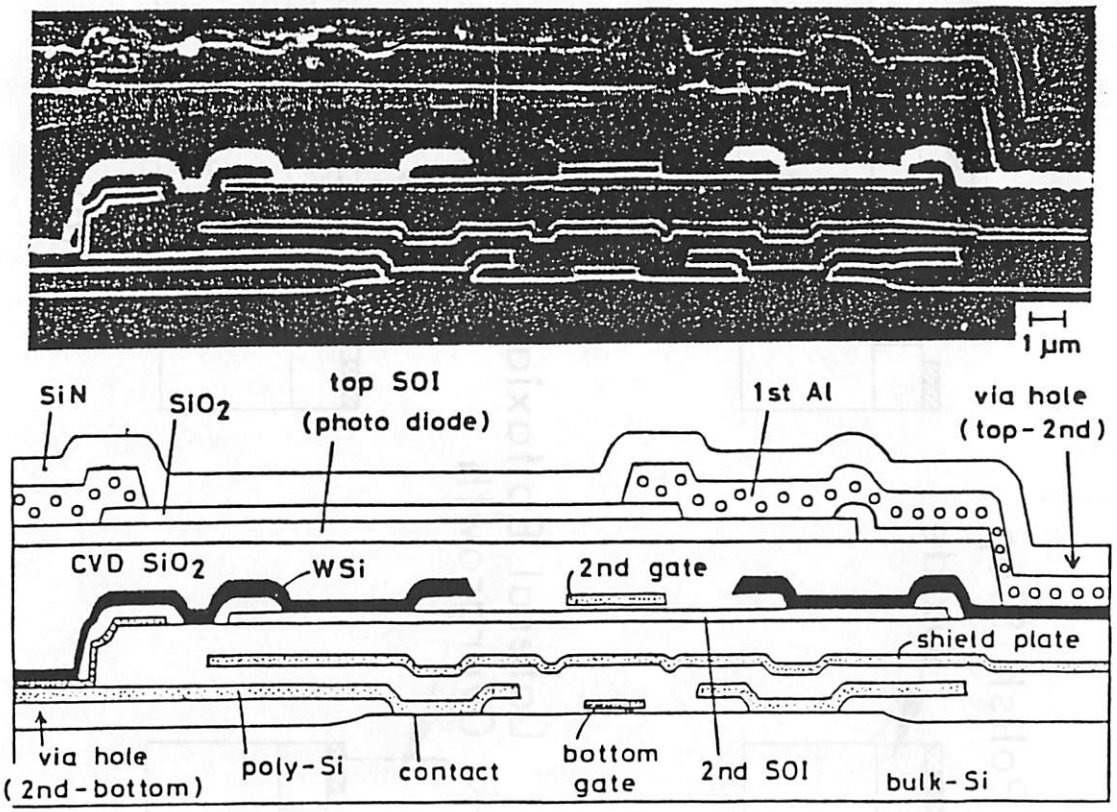
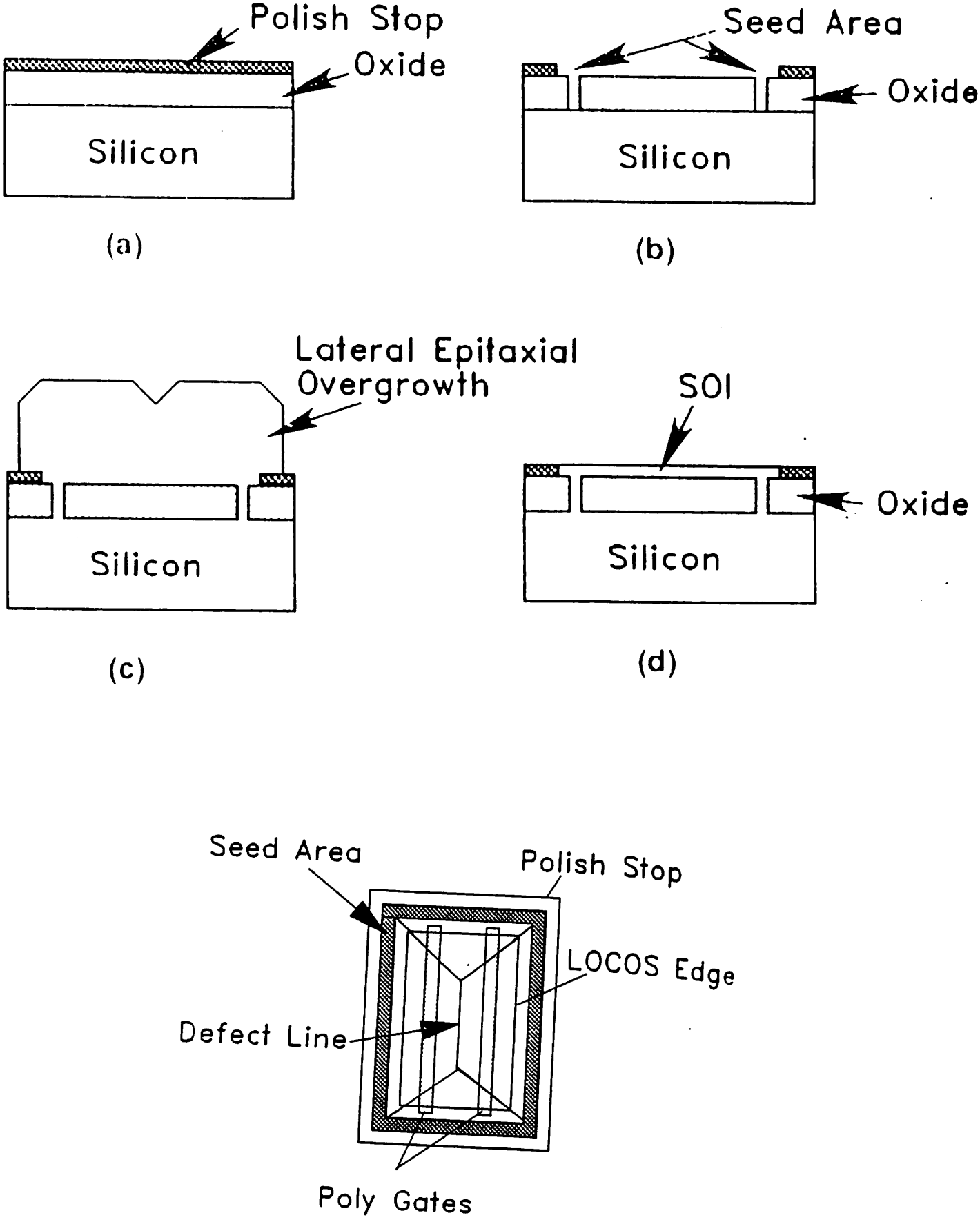
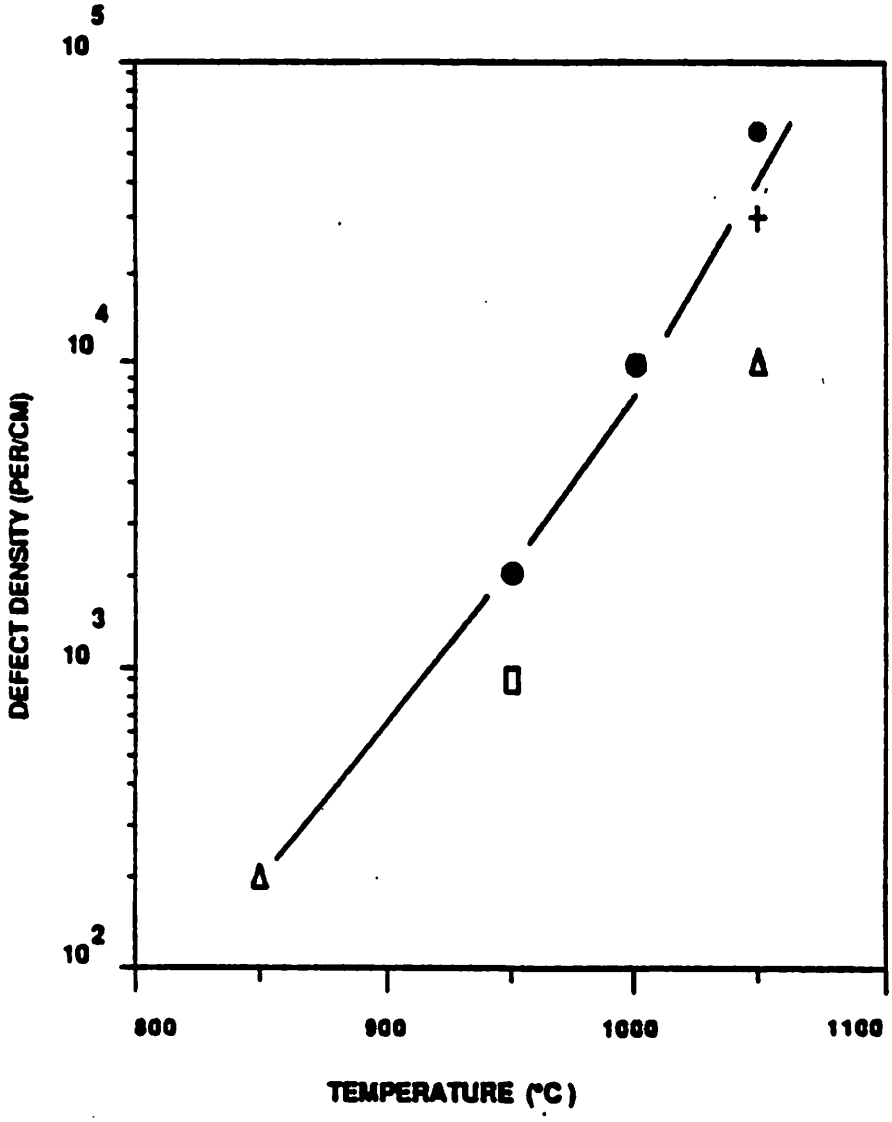


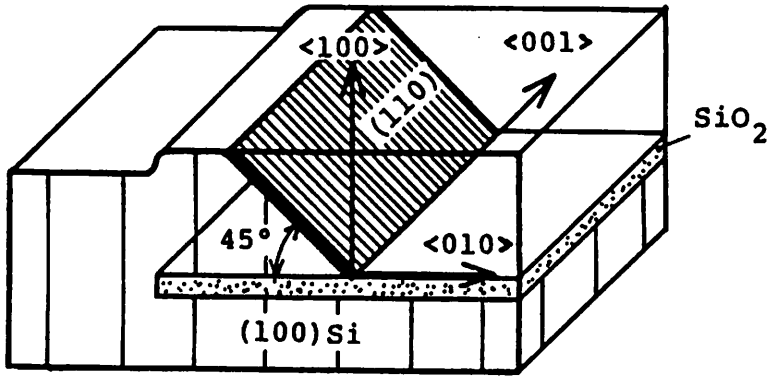
Fig. 4


Polish Stop
Oxide
Silicon
(a)
Seed Area
Oxide
Silicon
(b)
Lateral Epitaxial
Overgrowth
Silicon
(c)
SOI
Oxide
Silicon
(d)
Seed Area
Polish Stop
LOCOS Edge
Defect Line
Poly Gates



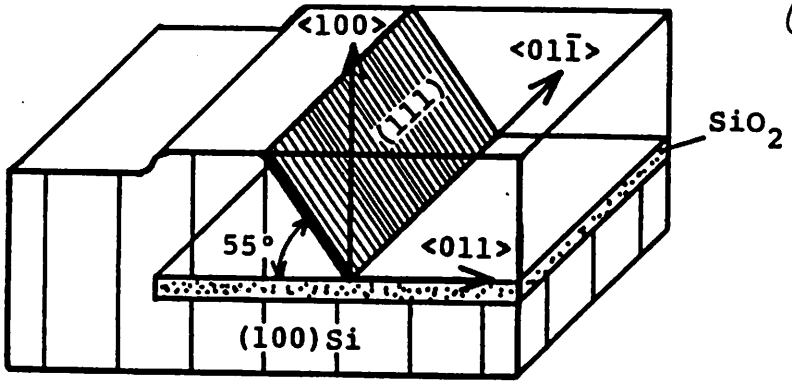
Defect density in selective epi as a function of growth temperature, +, Δ our results (TEM, etching, respectively, ● after [20], □ result on sample grown by J. Borland [43].

(a)



(100) ORIENTED

(b)



(110) ORIENTED

Fig. 8

MASK ALIGNMENT

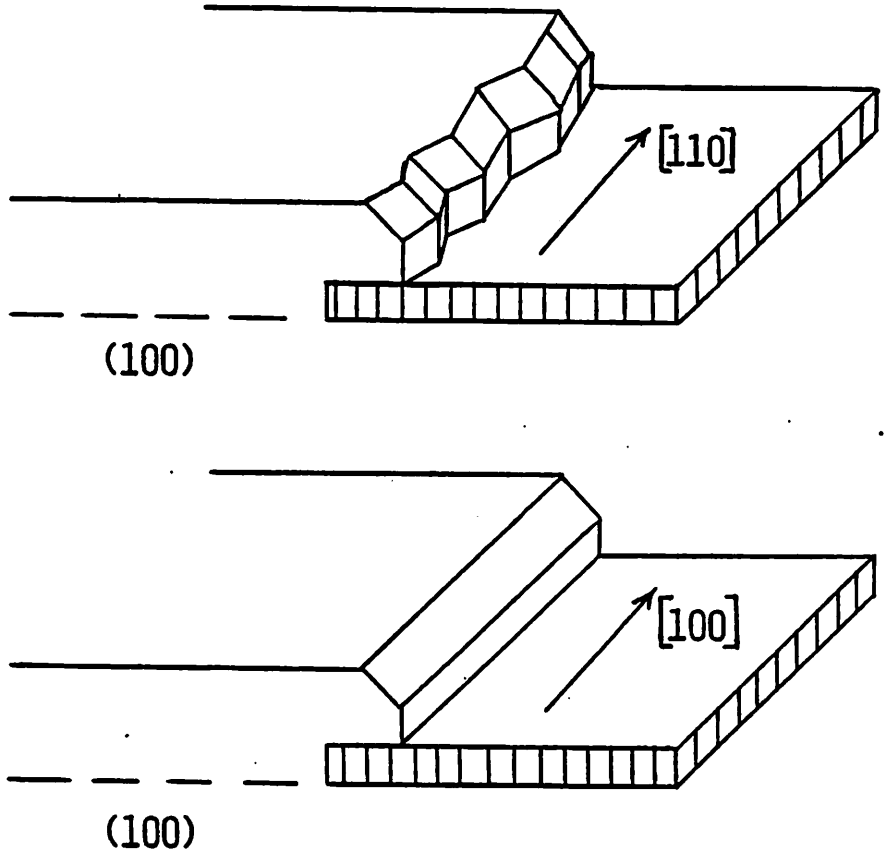


Fig. 9

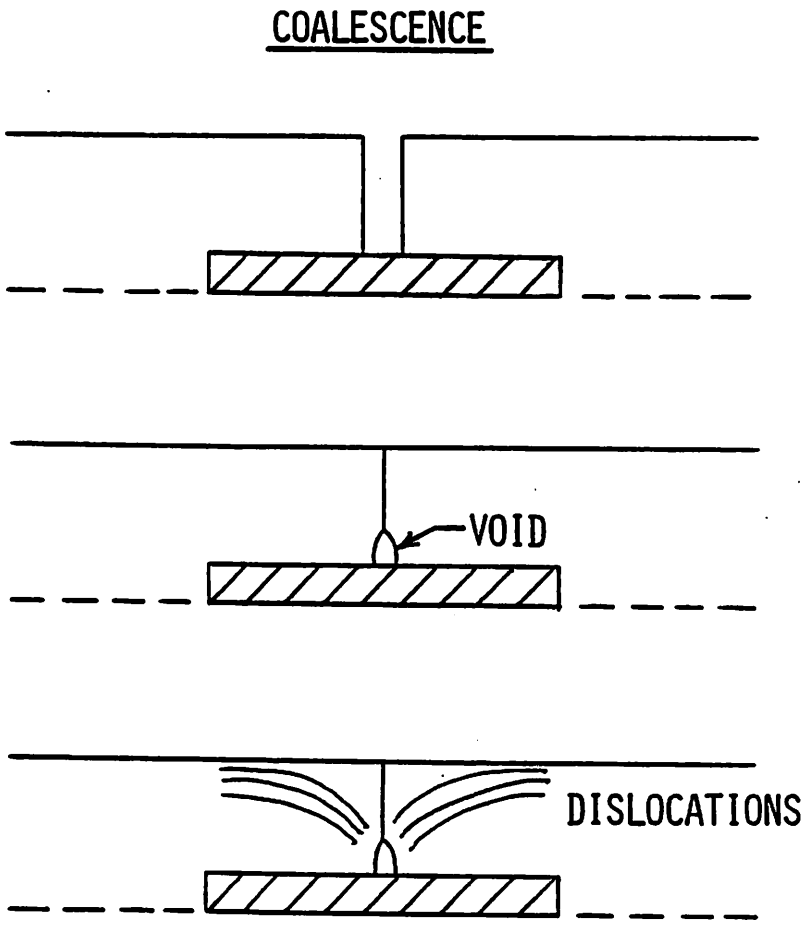


Fig. 10

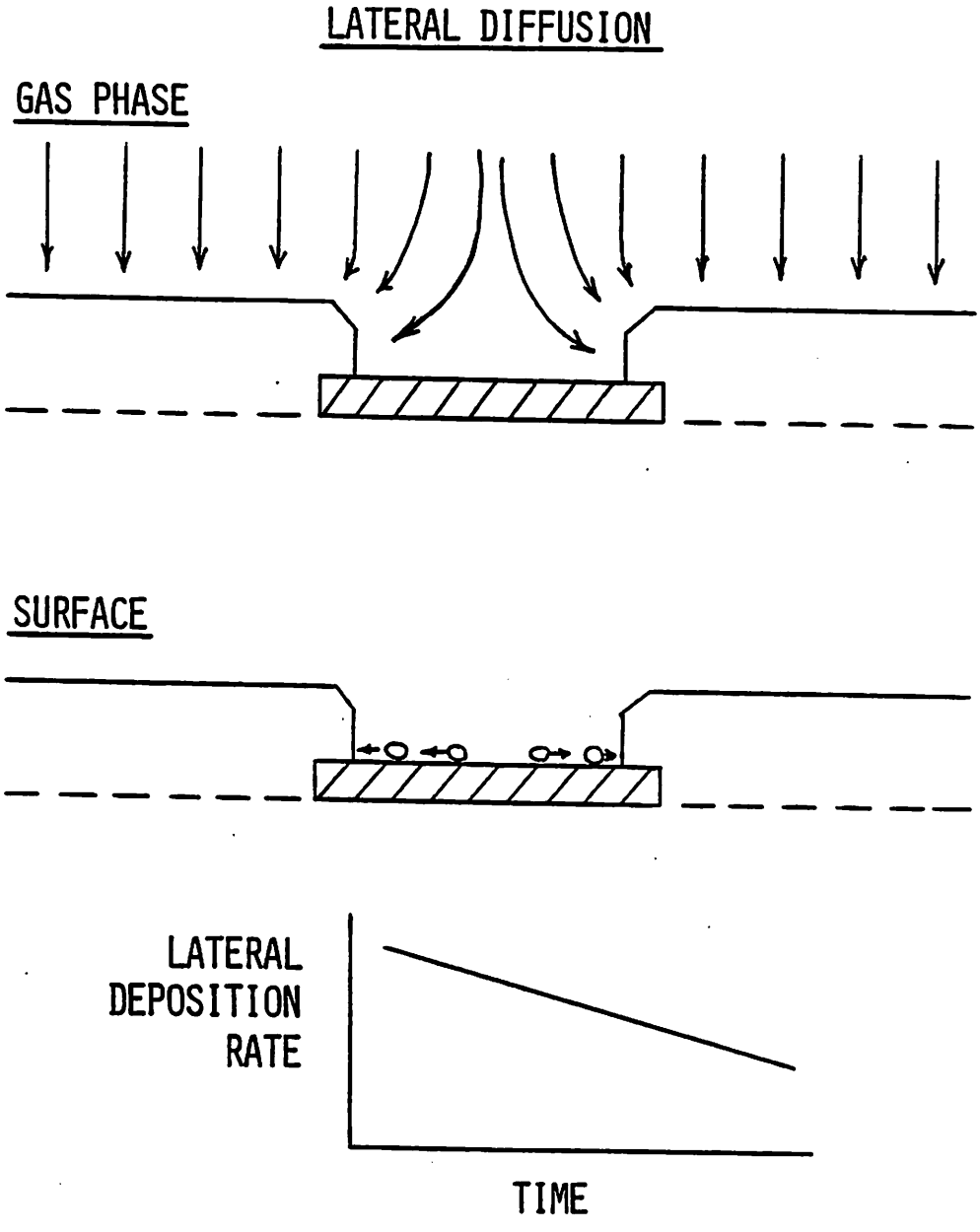
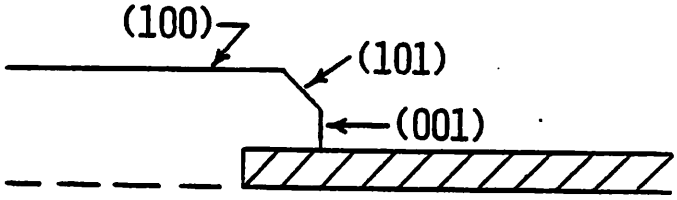

LATERAL DIFFUSION
GAS PHASE
SURFACE
LATERAL
DEPOSITION
RATE
TIME

Fig. 11

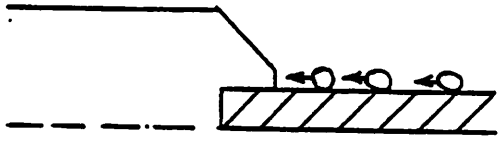
SHAPE OF LATERALLY ADVANCING GROWTH FRONT



BOUNDED BY SLOWEST GROWING PLANE

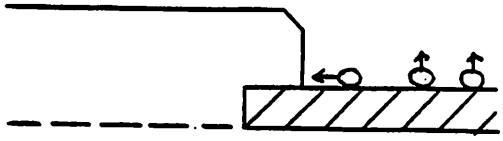
NO HCL

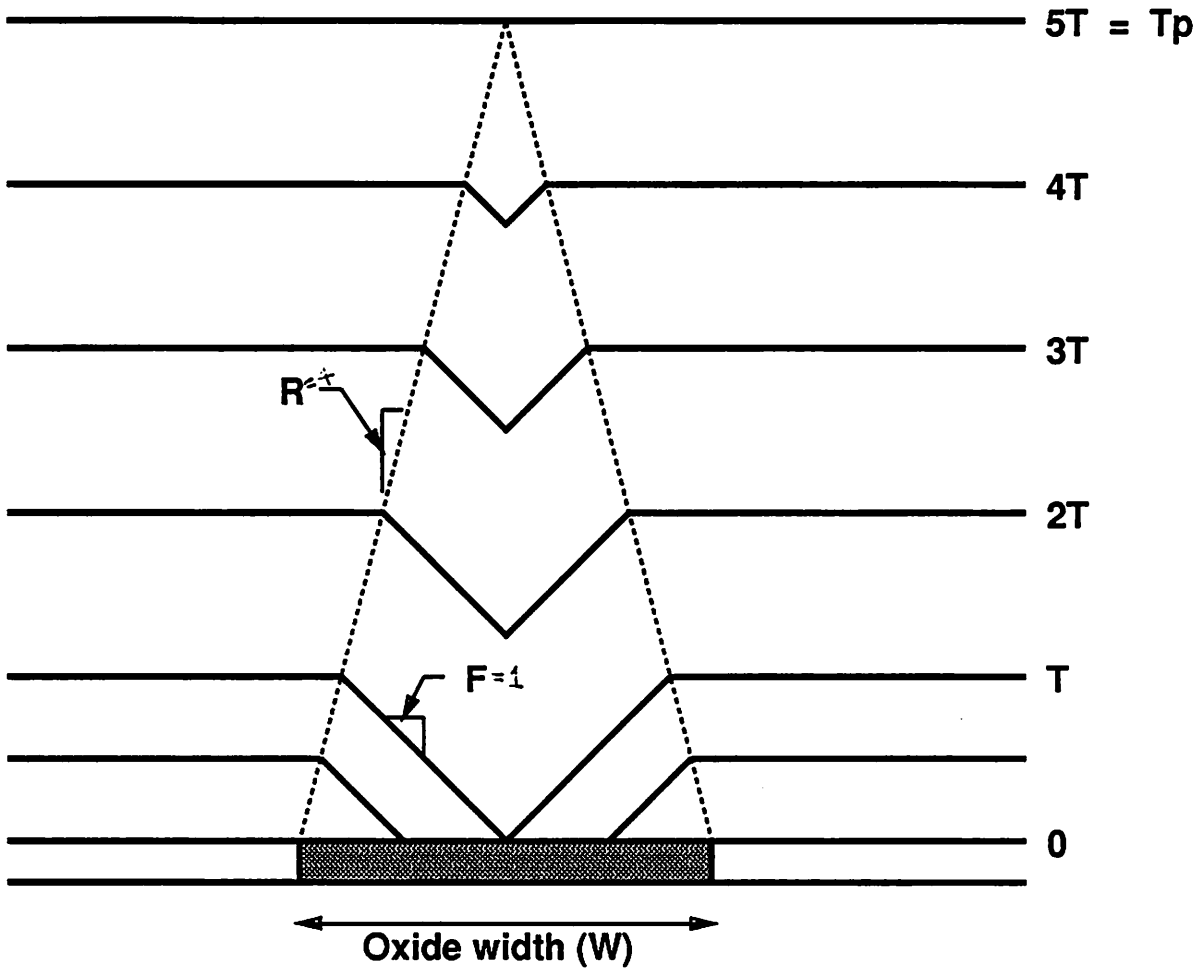
$R_{001} > R_{101}$



WITH HCL

$R_{101} > R_{001}$





$$T = \frac{W}{2} \times \frac{R \times F}{R + F}$$ Epi thickness when growth fronts meet

$$Tp = T \times \frac{R + F}{F}$$ Epi thickness when planarity complete

$$= R \times \frac{W}{2}$$

Example: UCB Process

R = 10

F = 1.2

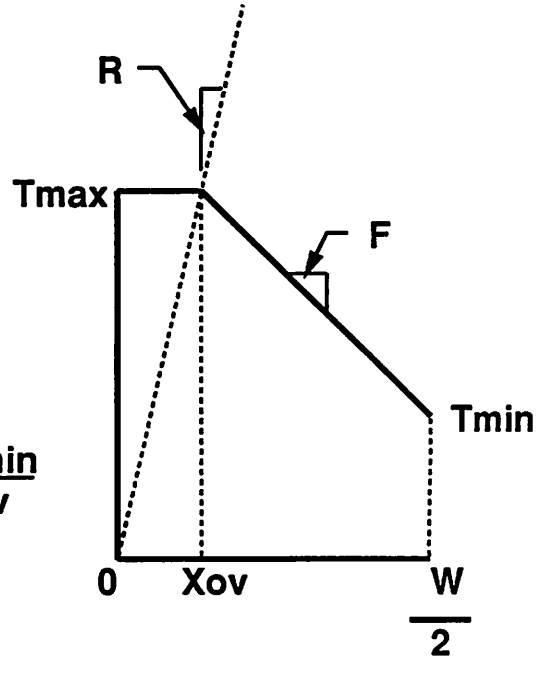
For W = 3.0um .....

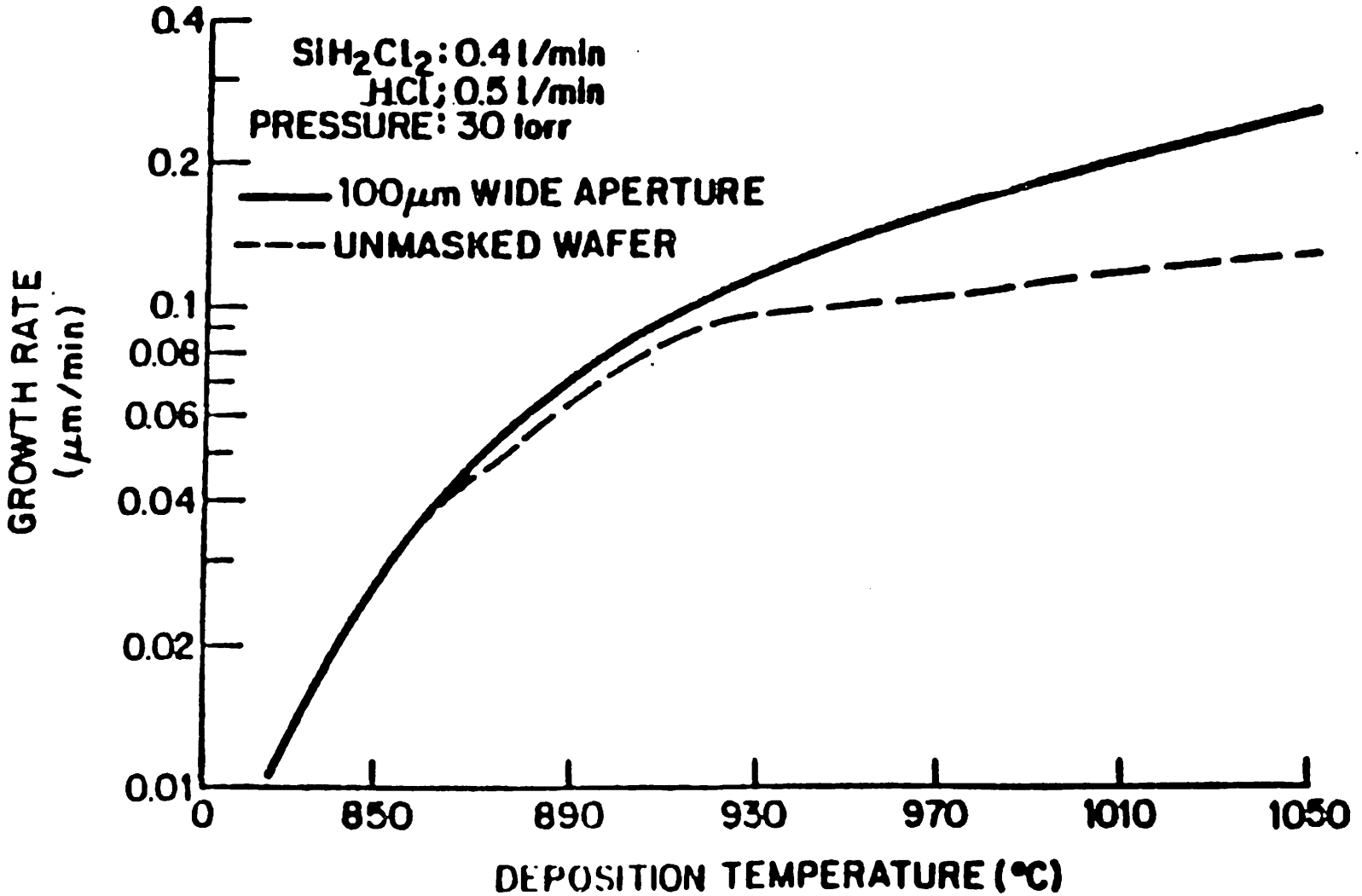
T = 1.6um

Tp = 15um !!

$$R = \frac{Tmax}{Xov}$$

$$F = \frac{Tmax - Tmin}{W/2 - Xov}$$





Growth rate in 100 μm wide aperature (solid line) and unmask (no oxide) silicon wafer (dashed line) as a function of growth temperature [40].

Fig. 13

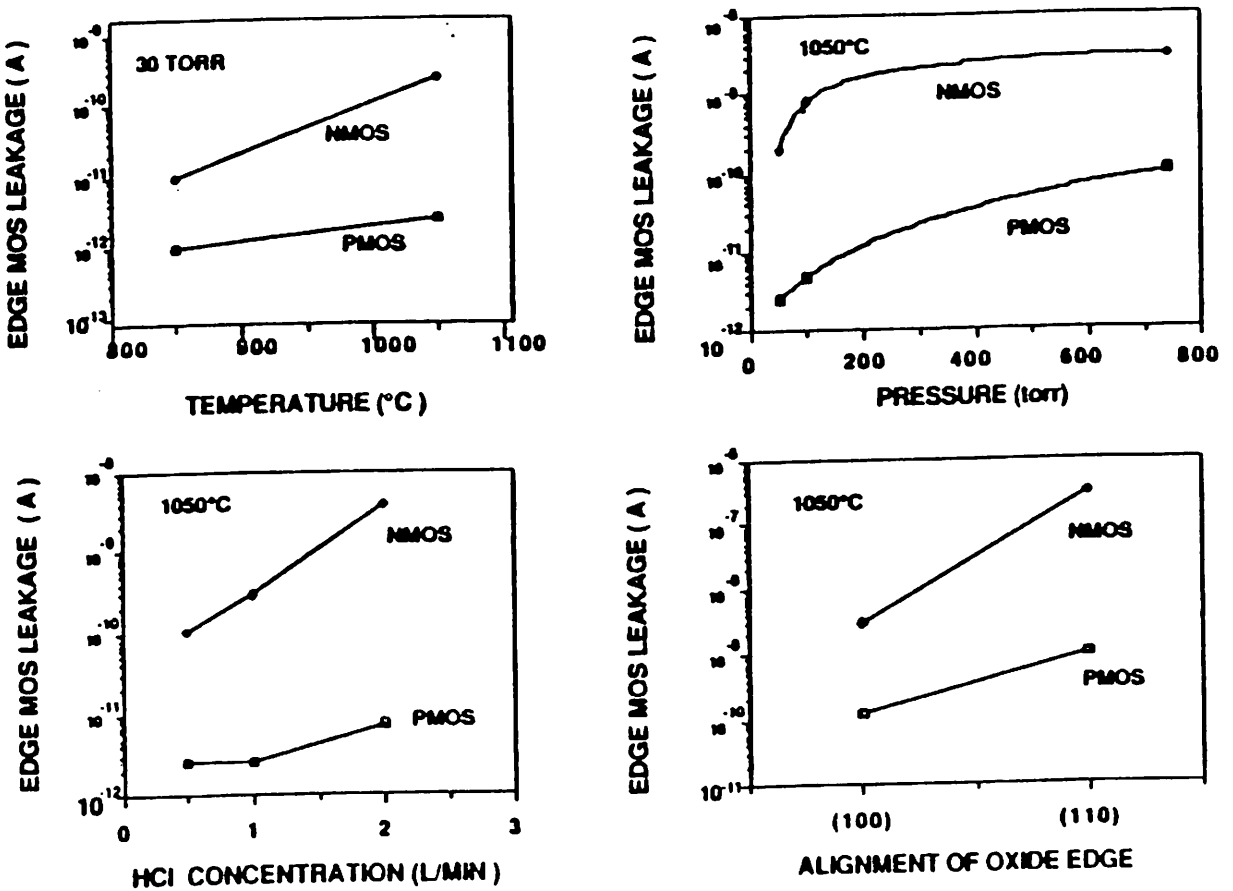


Fig. 9. Typical leakage current of edge MOS transistors made in selective epi as a function of a) deposition temperature, b) pressure, c) HCl concentration, d) orientation of oxide edge.

Fig. 15

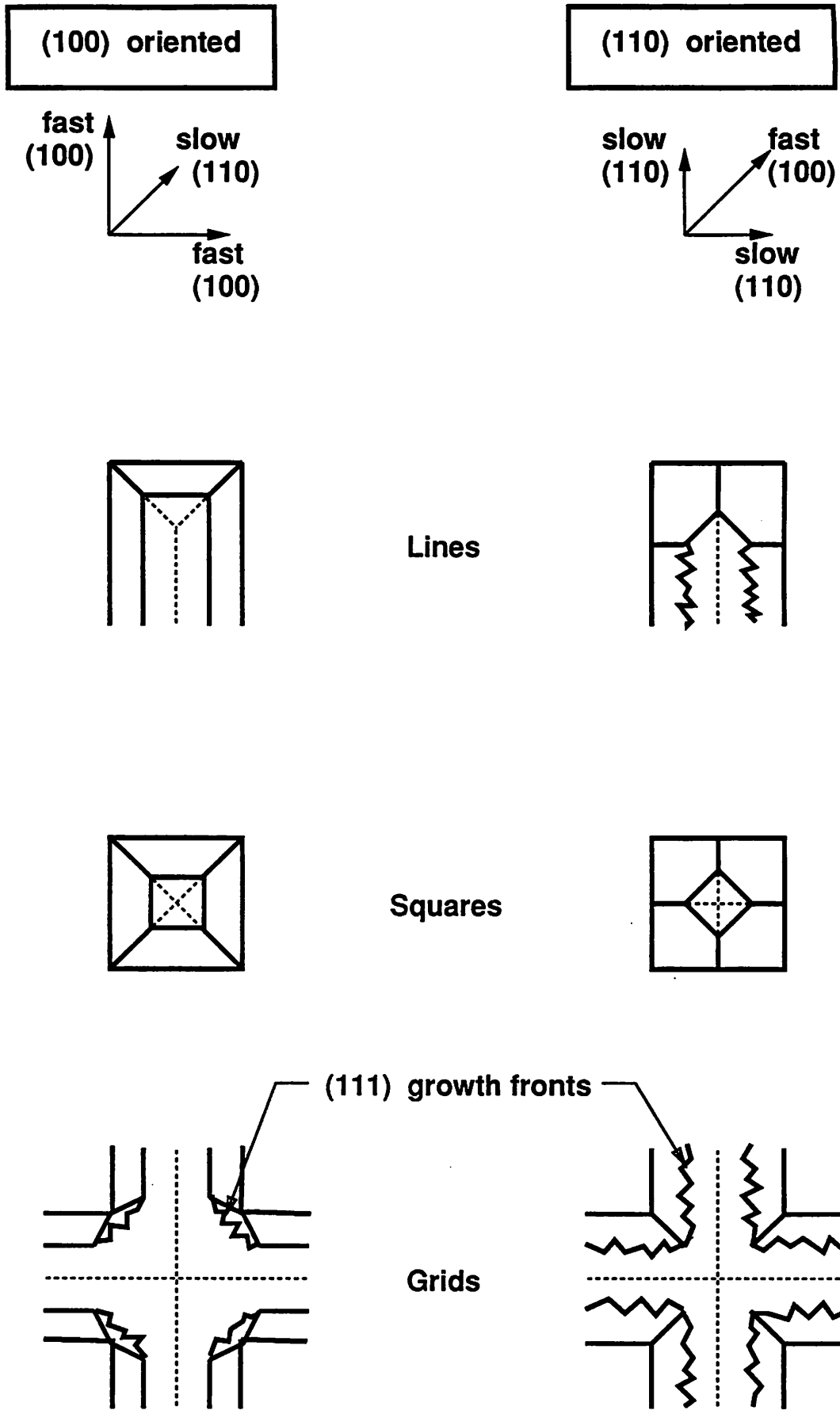
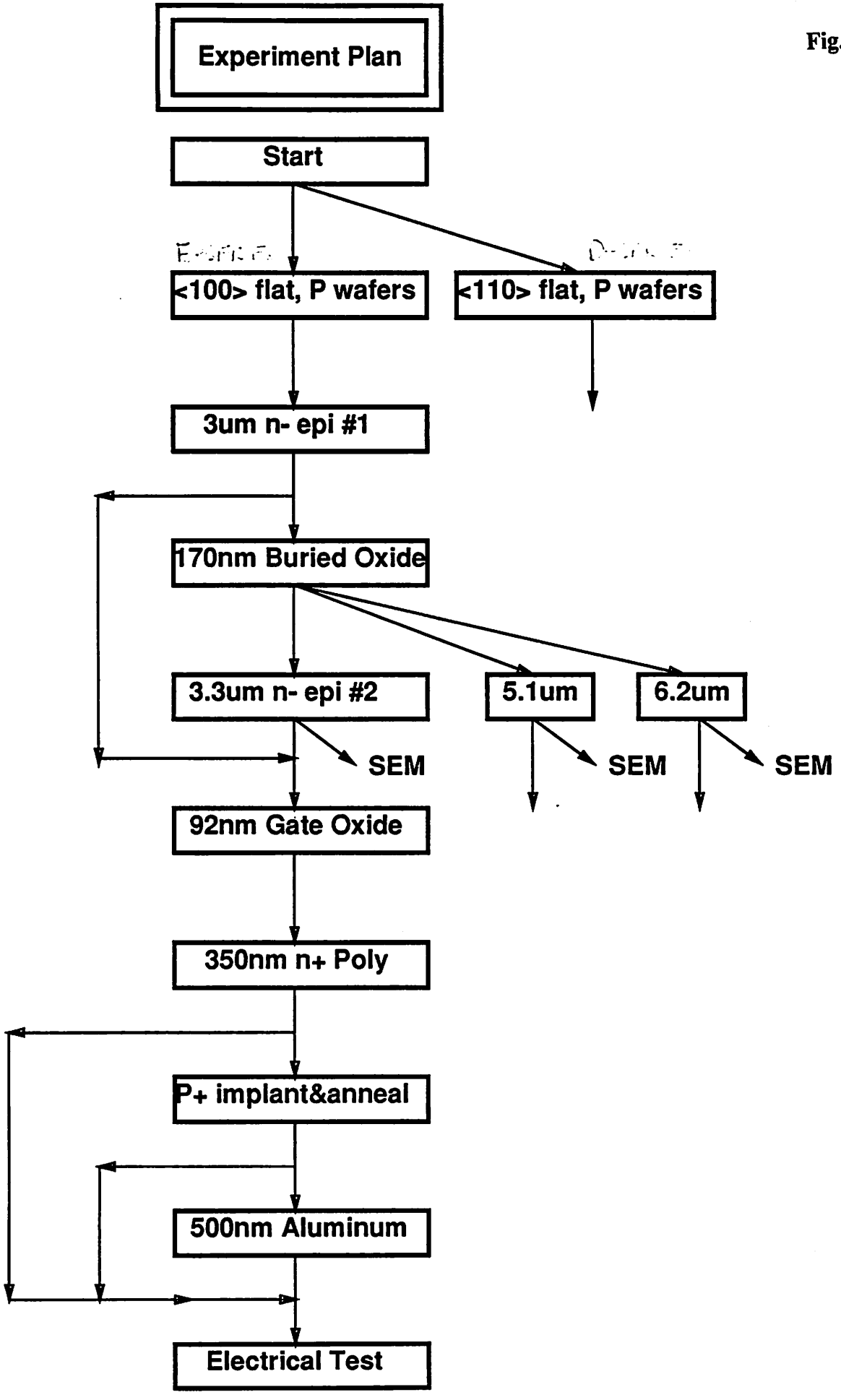
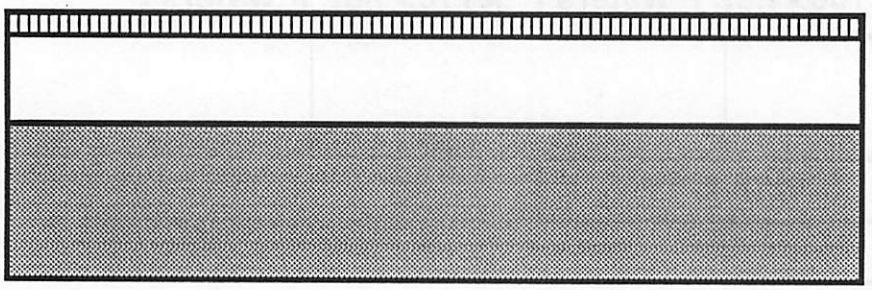

(100) oriented
fast (100)
slow (110)
fast (100)
(110) oriented
slow (110)
fast (100)
slow (110)
Lines
Squares
(111) growth fronts
Grids

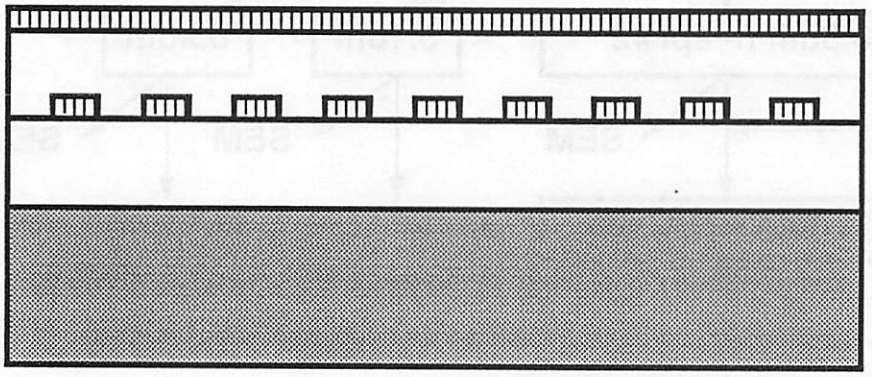
Fig. 16


Experiment Plan
Start
<100> flat, P wafers
<110> flat, P wafers
3um n- epi #1
170nm Buried Oxide
3.3um n- epi #2
5.1um
6.2um
SEM
SEM
SEM
92nm Gate Oxide
350nm n+ Poly
P+ implant&anneal
500nm Aluminum
Electrical Test

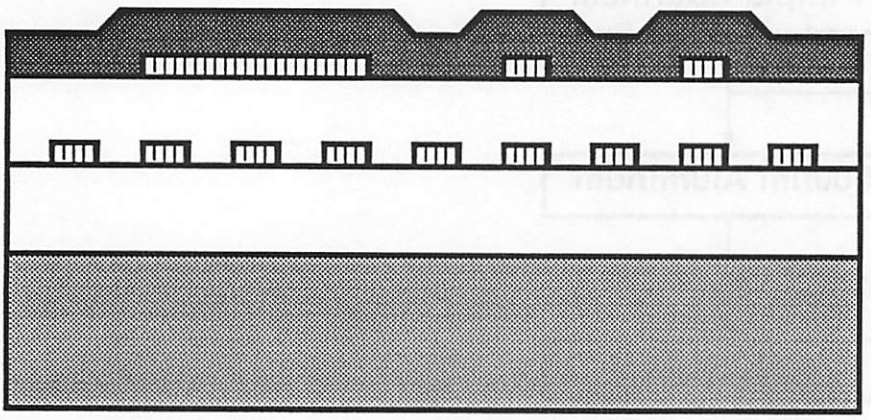
Selective Epi Lateral Overgrowth (SELO) Process



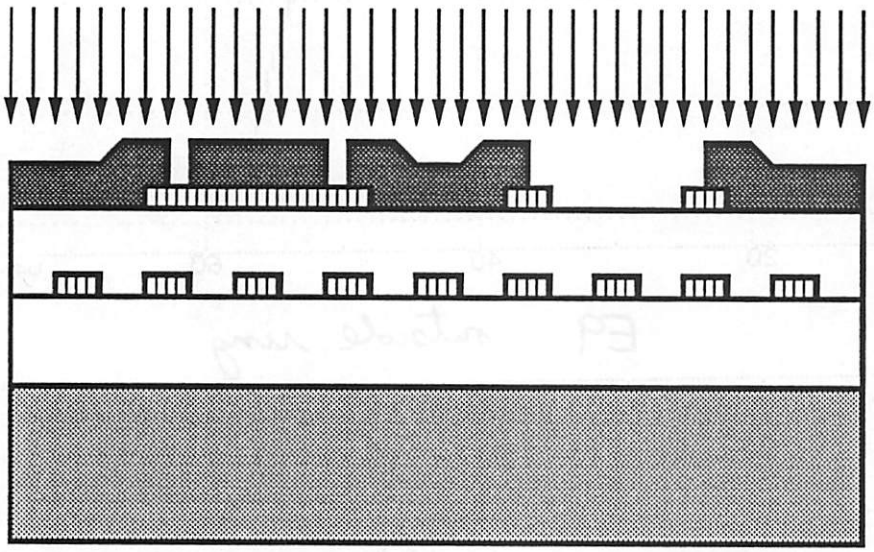
1. Grow 3.0um n-epi on p-sub
2. Grow 200nm buried oxide



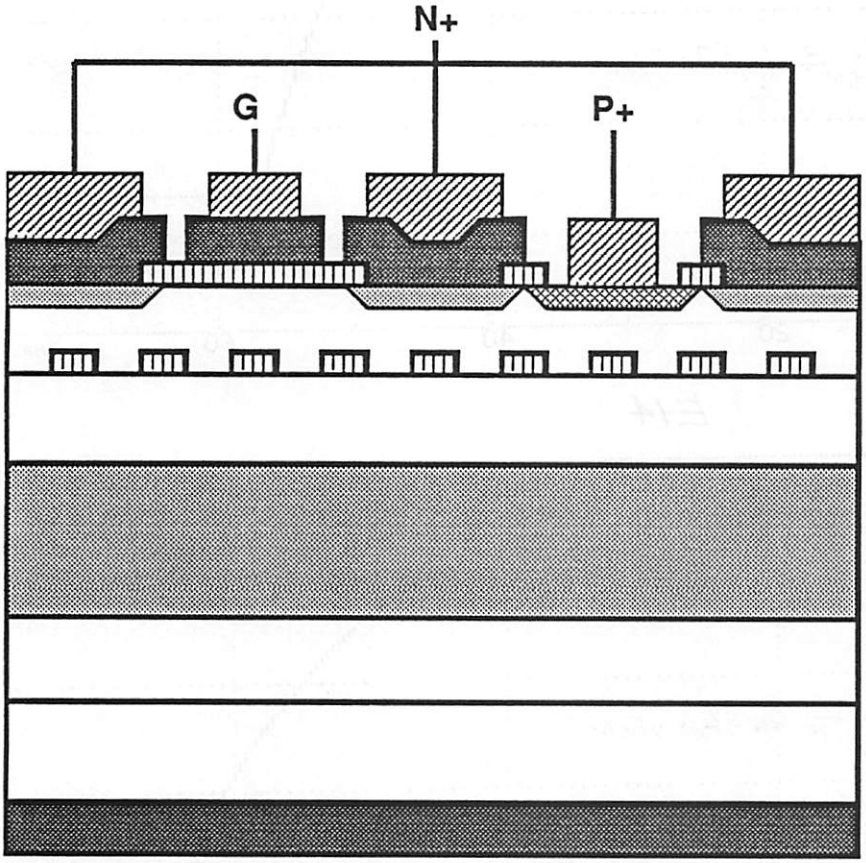
3. Mask & plasma etch oxide
4. Grow 20nm sacrificial oxide
5. Wet etch sac oxide
6. Pre-epi surface preparation
7. Grow 3-6um n-epi over oxide
8. Grow 100nm gate oxide



9. Mask & wet etch gate oxide
10. Deposit 350nm n+ poly



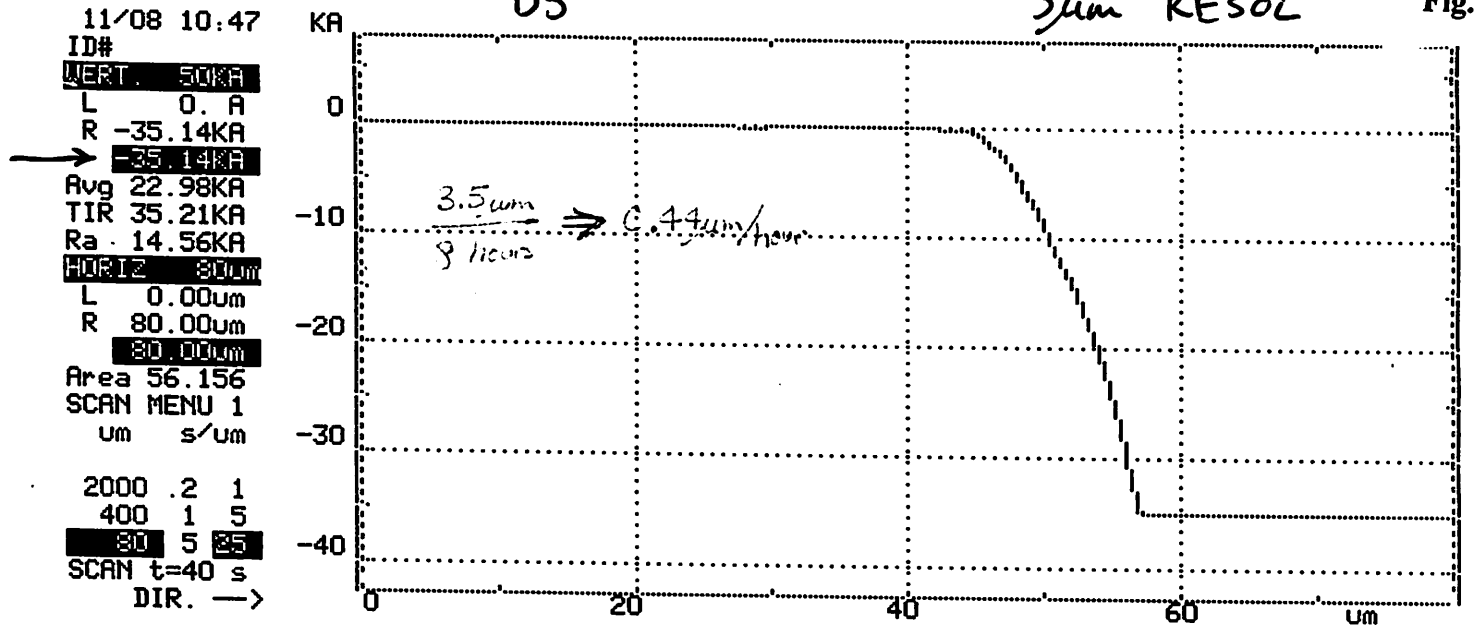
11. Mask & wet etch poly
12. Un-masked 50keV BF2 implant

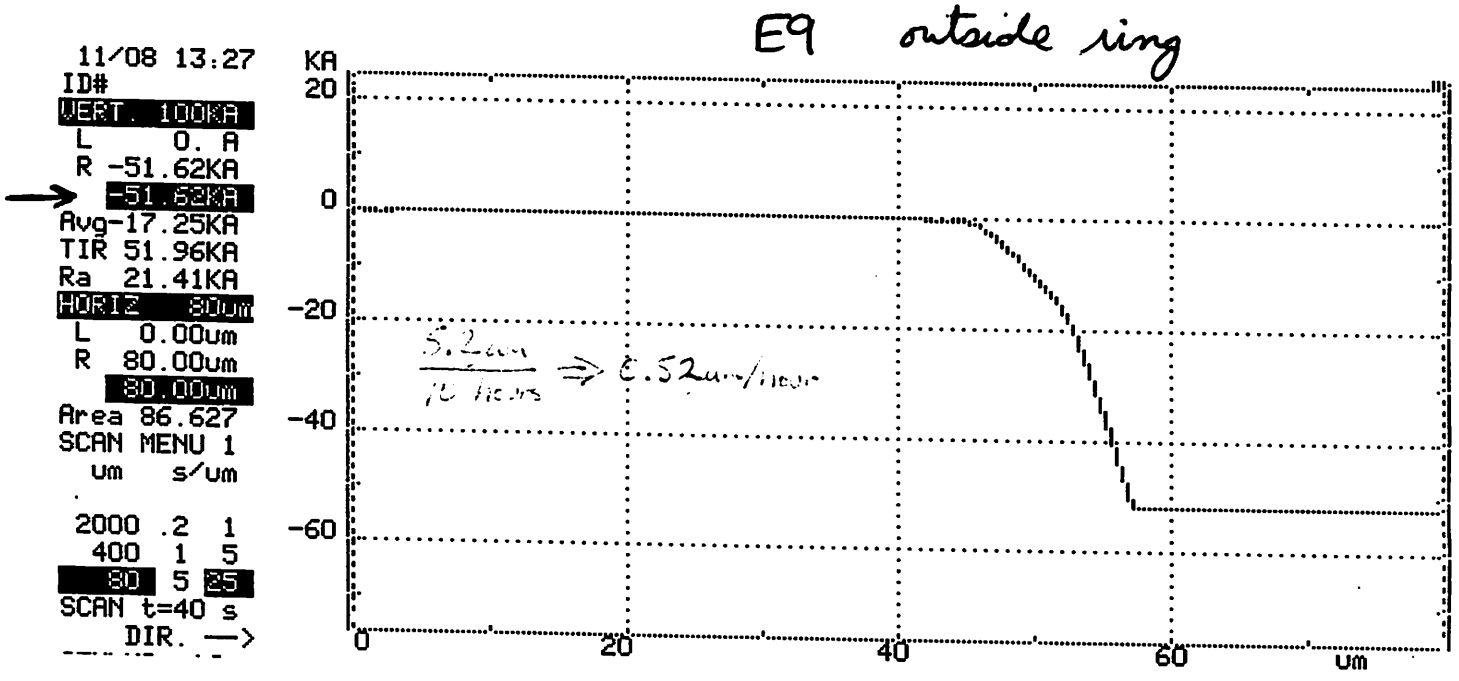


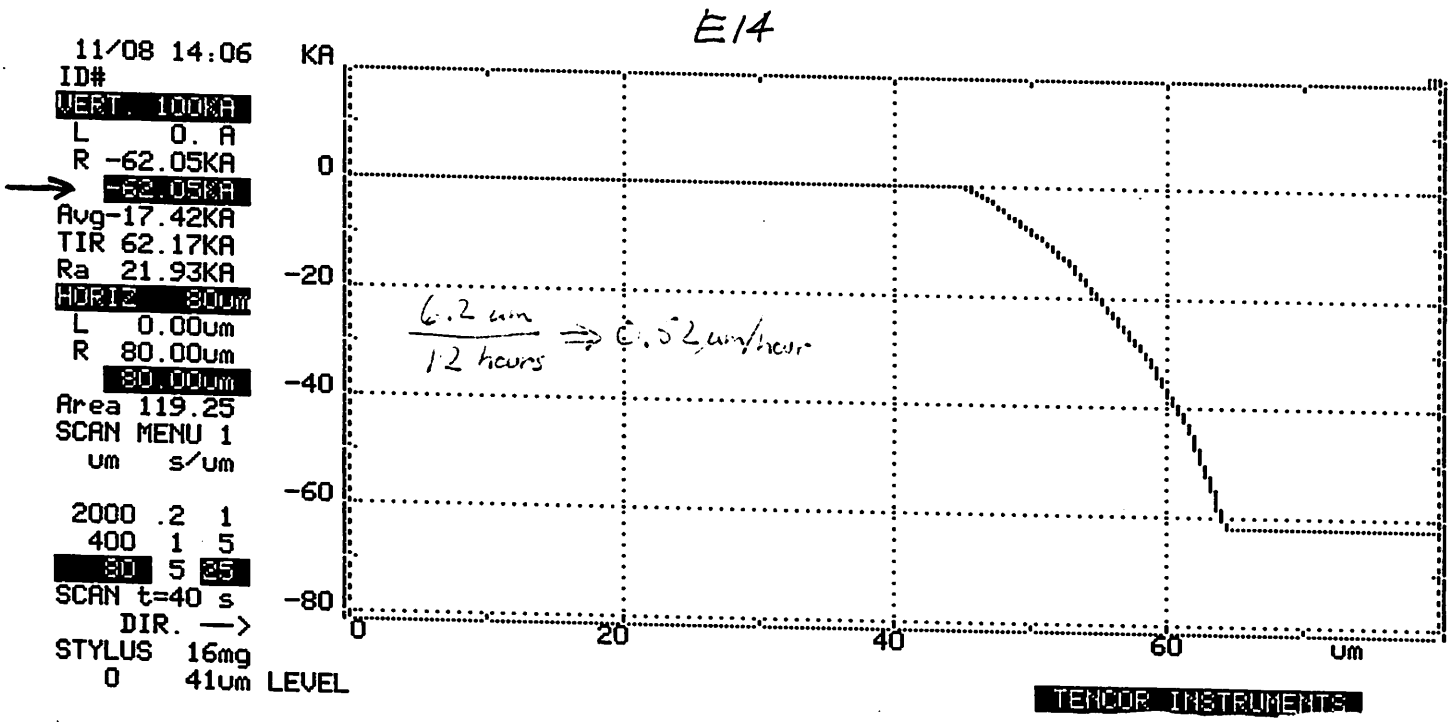
13. 20min, 900C junction anneal
14. HF dip
15. Sputter Aluminum
16. Mask & wet etch Aluminum

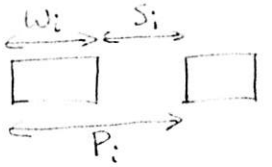
Backside n+ poly getter

Fig. 18


11/08 10:47
ID#
VERT. 50KA
L 0. A
R -35.14KA
-35.14KA
Avg 22.98KA
TIR 35.21KA
Ra 14.56KA
HORIZ 80um
L 0.00um
R 80.00um
80.000um
Area 56.156
SCAN MENU 1
um s/um
2000 .2 1
400 1 5
80 5 25
SCAN t=40 s
DIR. —>
KA
0
-10
-20
-30
-40
0
20
40
60
um
D5
5um RESOL
3.5um / 8 hours ⇒ 0.44um/hour


11/08 13:27
ID#
VERT. 100KA
L 0. A
R -51.62KA
-51.62KA
Avg-17.25KA
TIR 51.96KA
Ra 21.41KA
HORIZ 80um
L 0.00um
R 80.00um
80.00um
Area 86.627
SCAN MENU 1
um s/um
2000 .2 1
400 1 5
80 5 25
SCAN t=40 s
DIR. —>
KA
20
0
-20
-40
-60
0
20
40
60
um
E9 outside ring
5.2um / 10 hours ⇒ 0.52um/hour


11/08 14:06
ID#
VERT. 100KA
L 0. A
R -62.05KA
-62.05KA
Avg-17.42KA
TIR 62.17KA
Ra 21.93KA
HORIZ 80um
L 0.00um
R 80.00um
80.00um
Area 119.25
SCAN MENU 1
um s/um
2000 .2 1
400 1 5
80 5 25
SCAN t=40 s
DIR. —>
STYLUS 16mg
0 41um LEVEL
KA
0
-20
-40
-60
-80
0
20
40
60
um
E14
6.2 um / 12 hours ⇒ 0.52 um/hour
TENCOR INSTRUMENTS



Selective Epi Testchip Floorplan

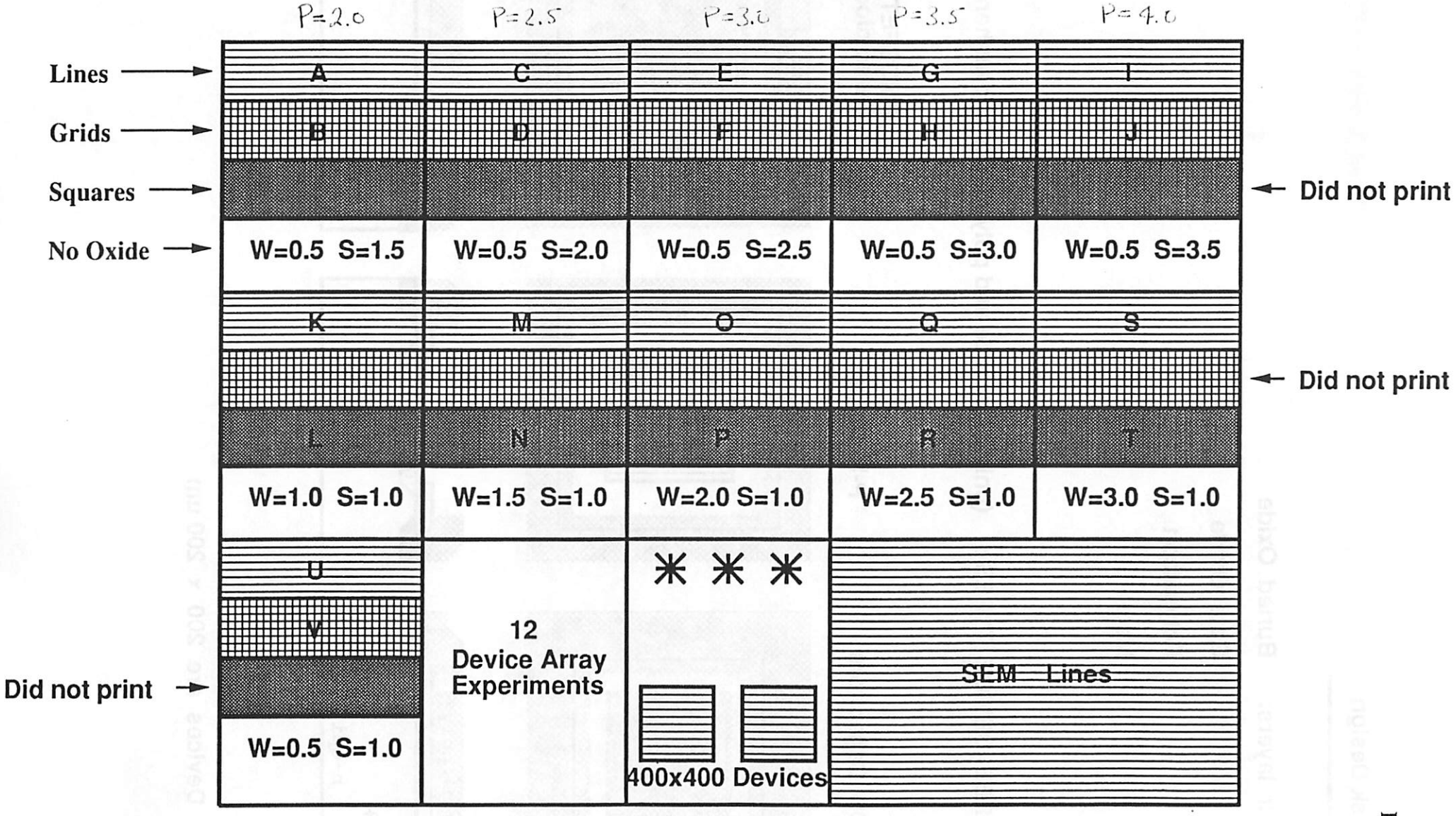


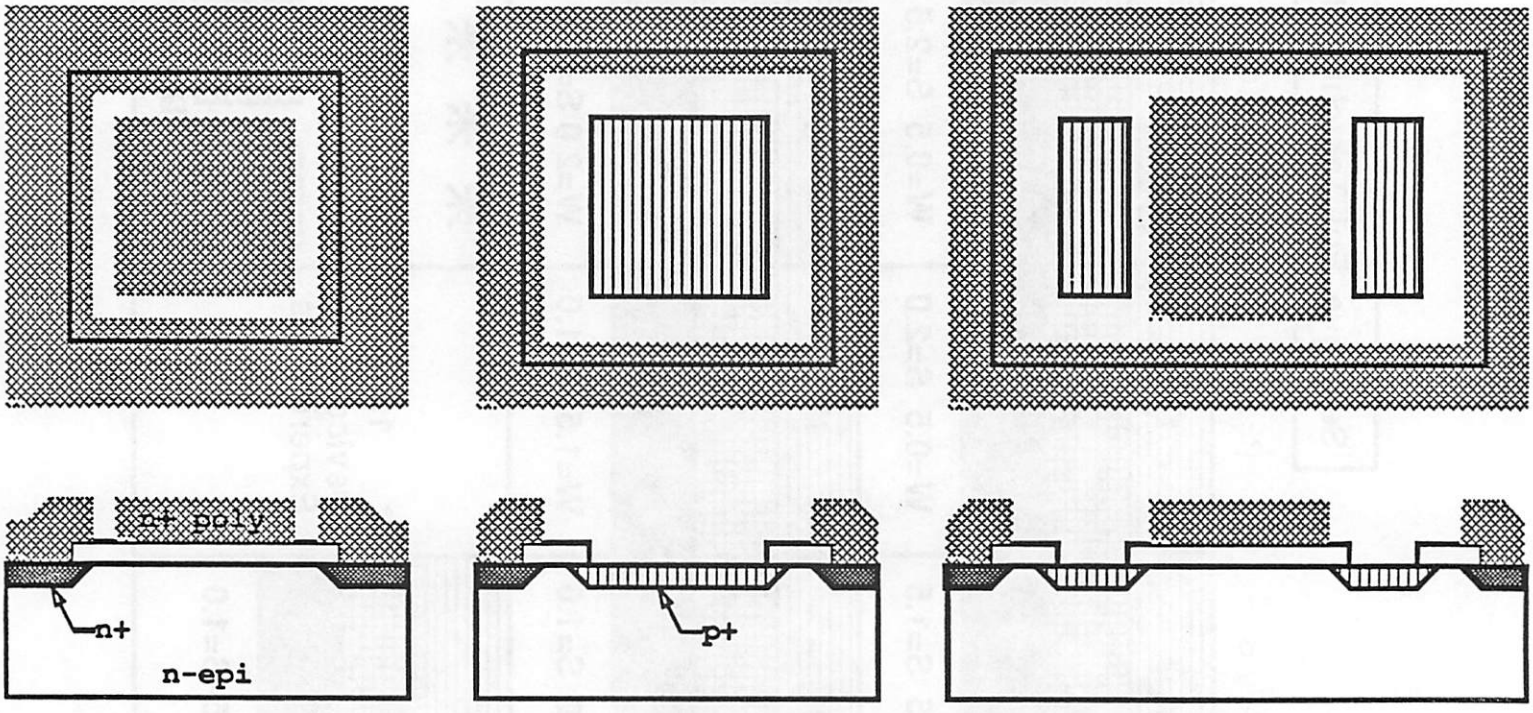
Fig. 19

Mask Design

Four layers: Buried Oxide
Gate Oxide
Polysilicon
Metal

THREE DEVICE TYPES : (only gate oxide and poly masks shown here)

MOS Capacitors | P-N Diodes | PMOSFETs / Gated diodes



All Devices are 200 x 200 um

Fig. P1

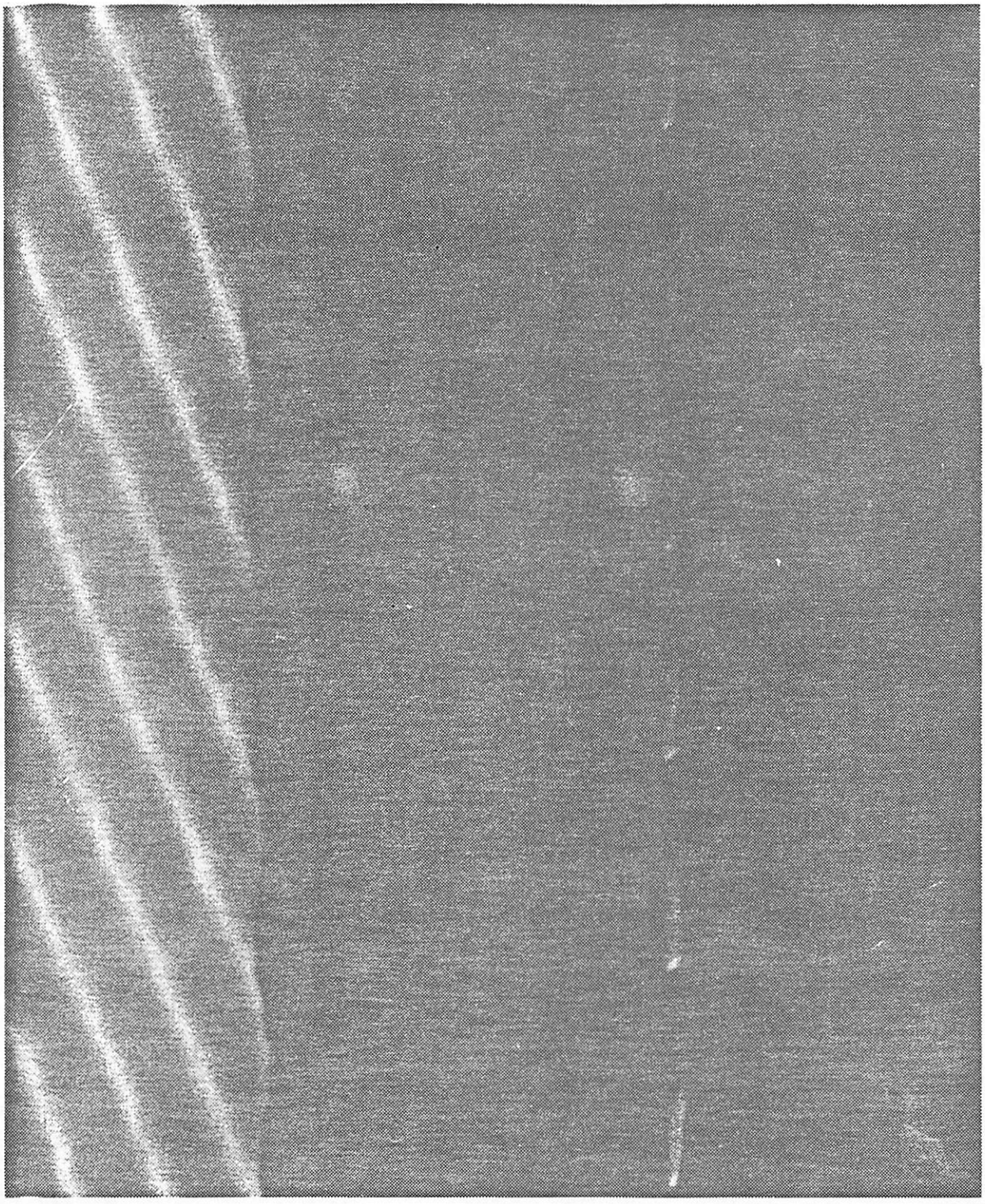


Fig. P2

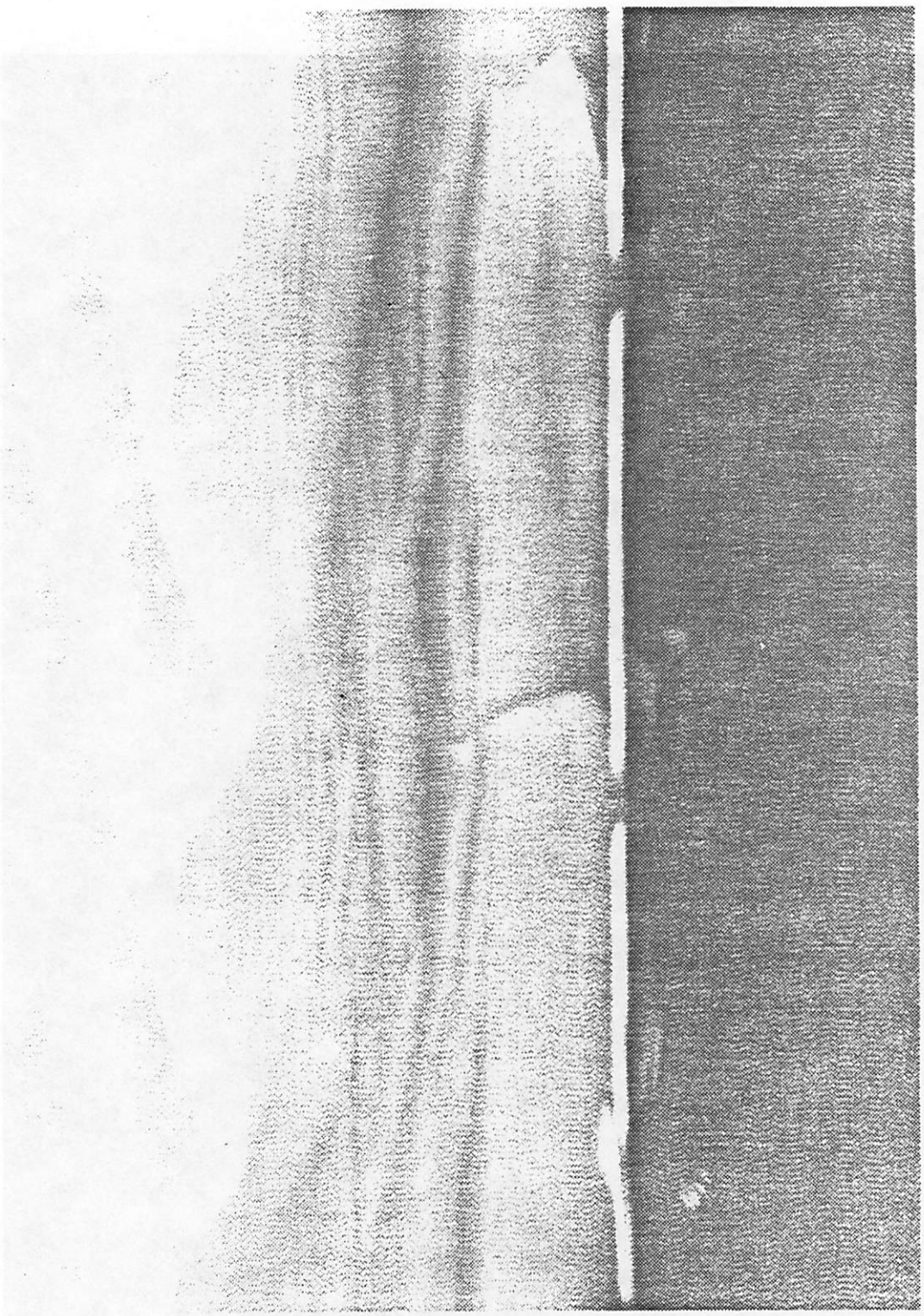
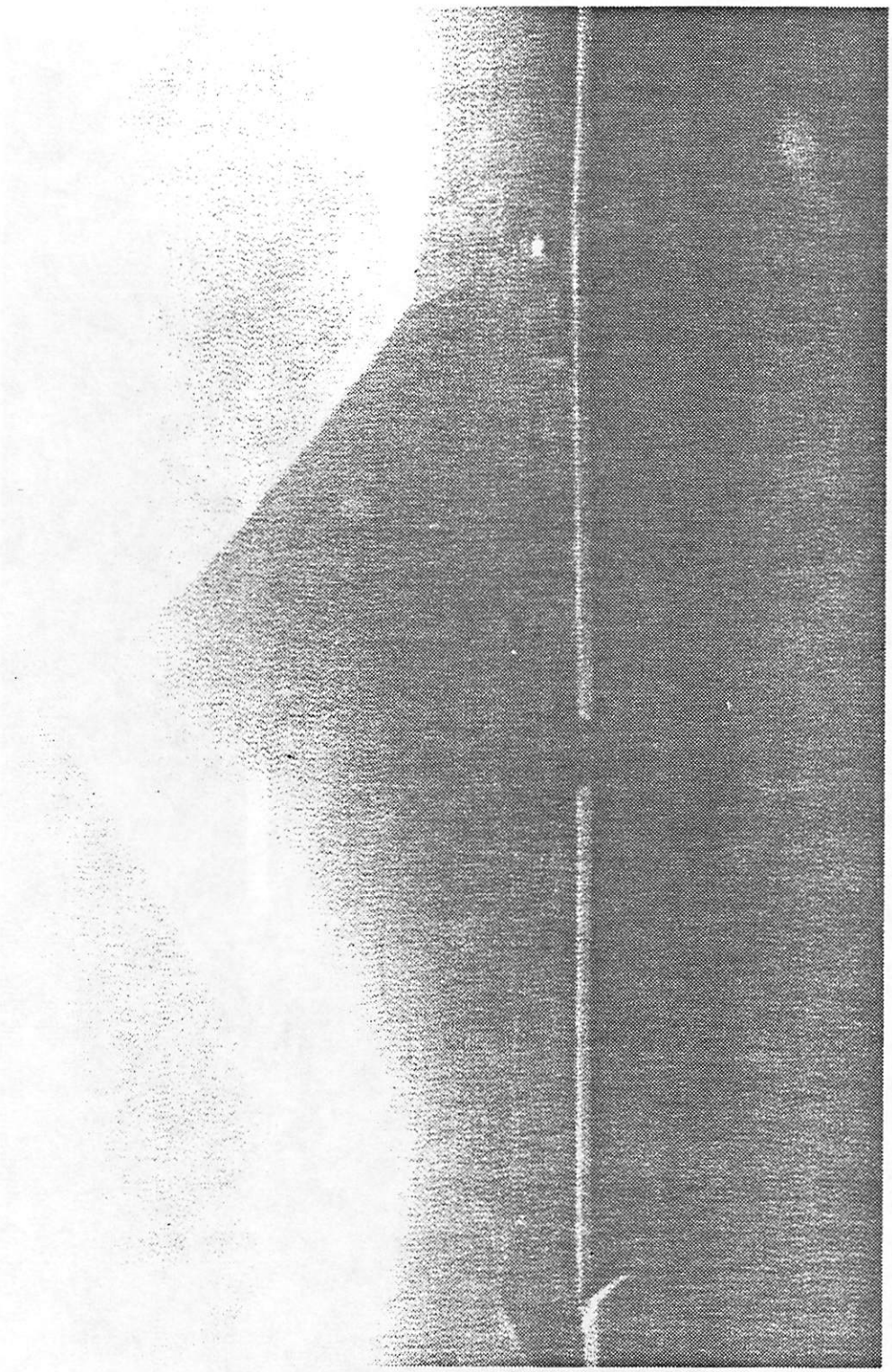
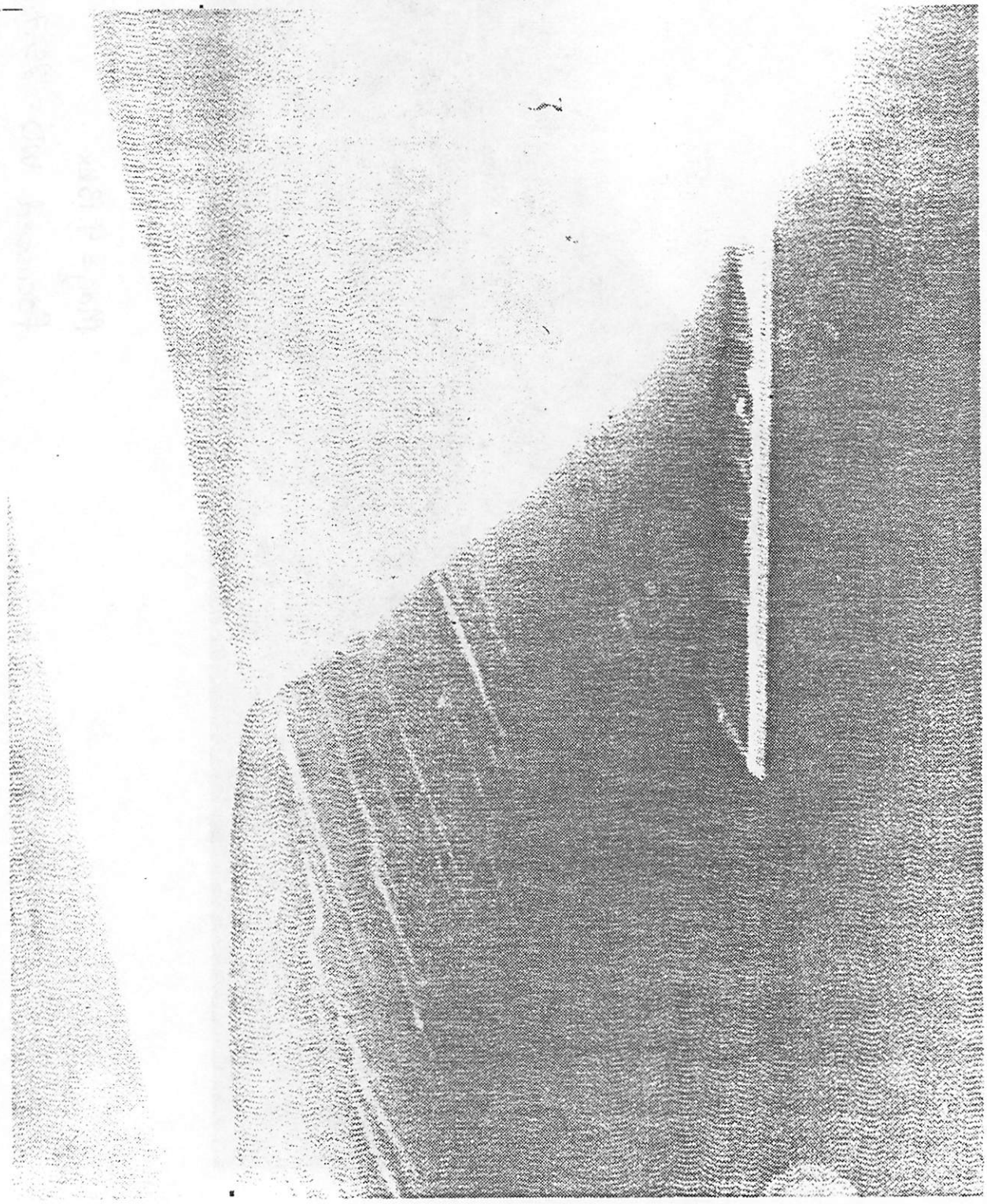

Accel voltage = 19.35 KeV
Focussed WD ≈ 25.6 mm
oxide ≈ 5.6 μm
Mag = 9.11 Kx
total length = 0.26μm

Fig. P3



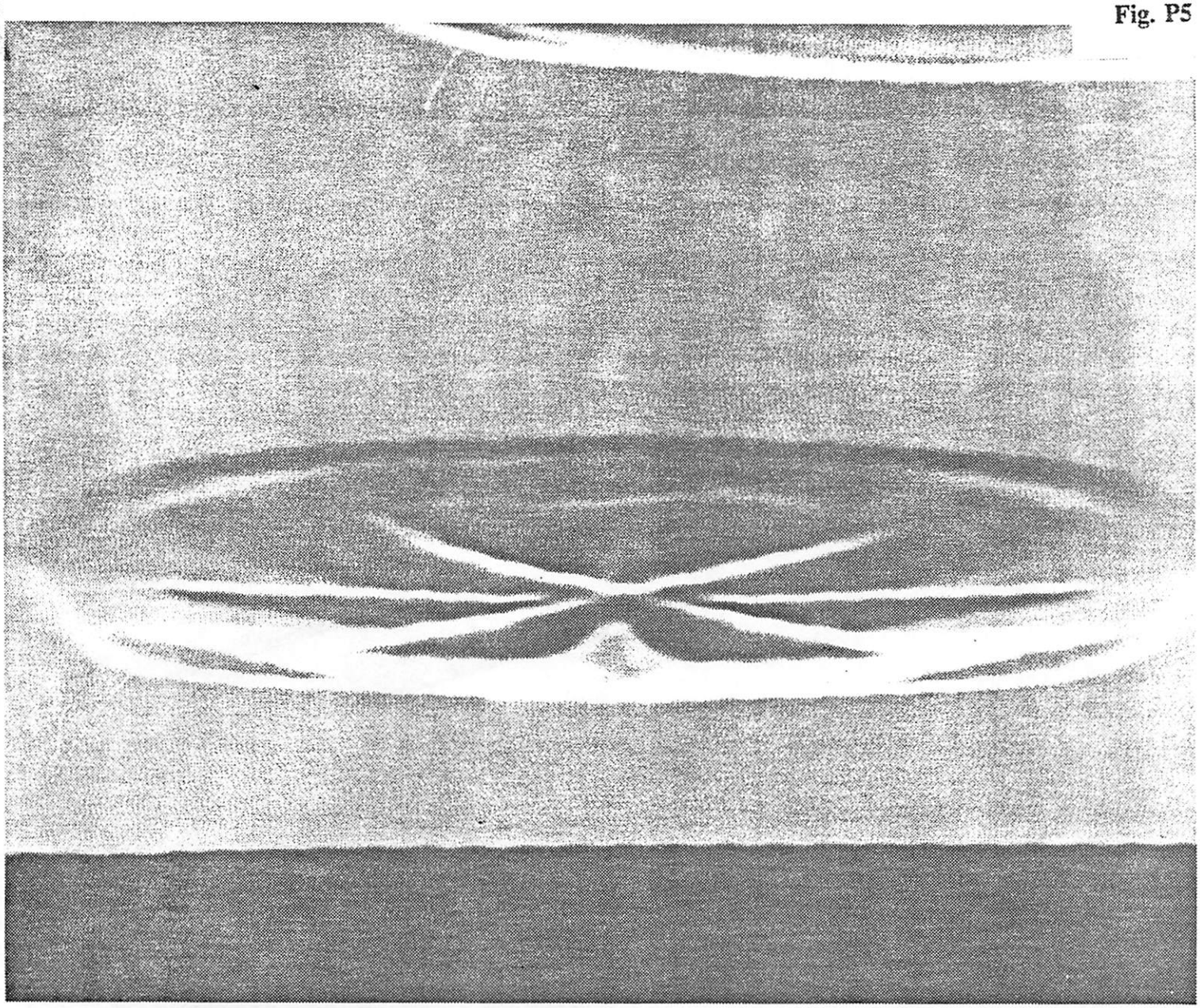
Mag = 9.13kx
focussed WD = 25.7 mm

Fig. P4

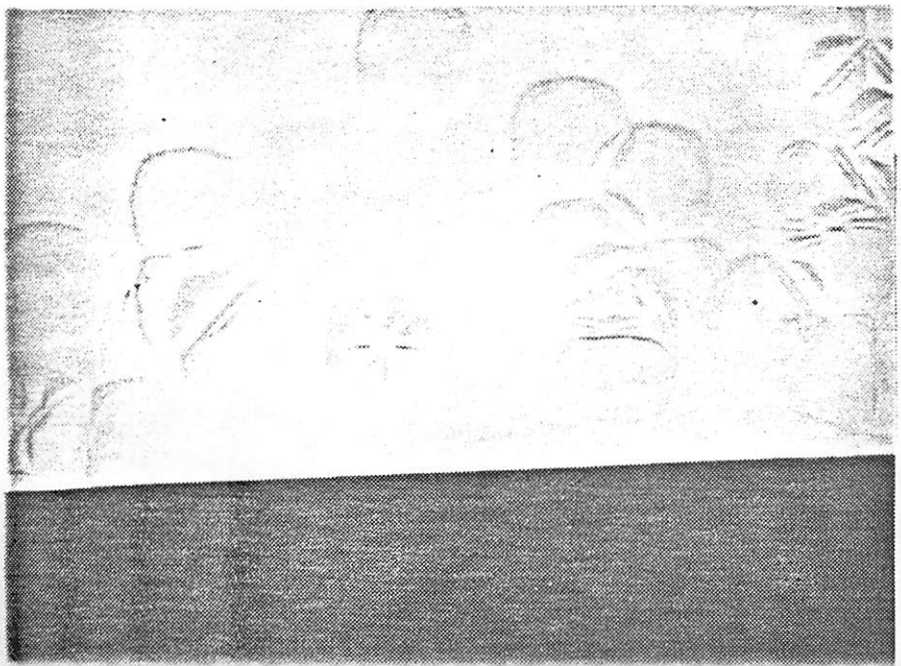


Mag = 9.13 KX
E11
Epi thickness = 5.1 μm

Fig. P5



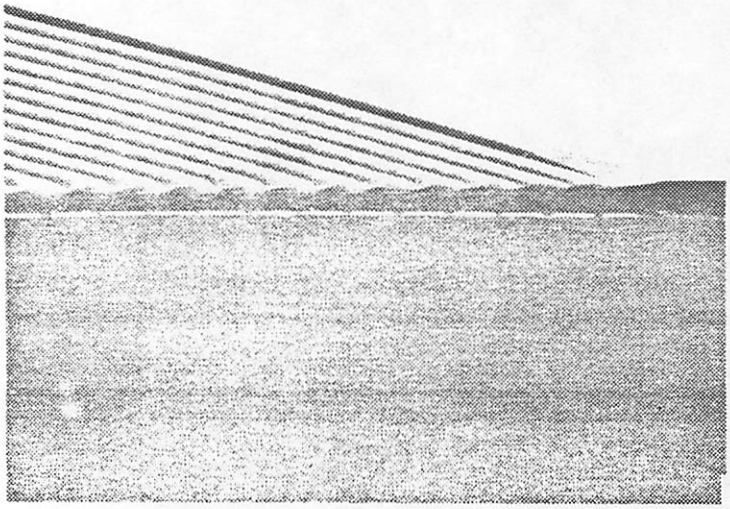
Mag = 8.38 kx



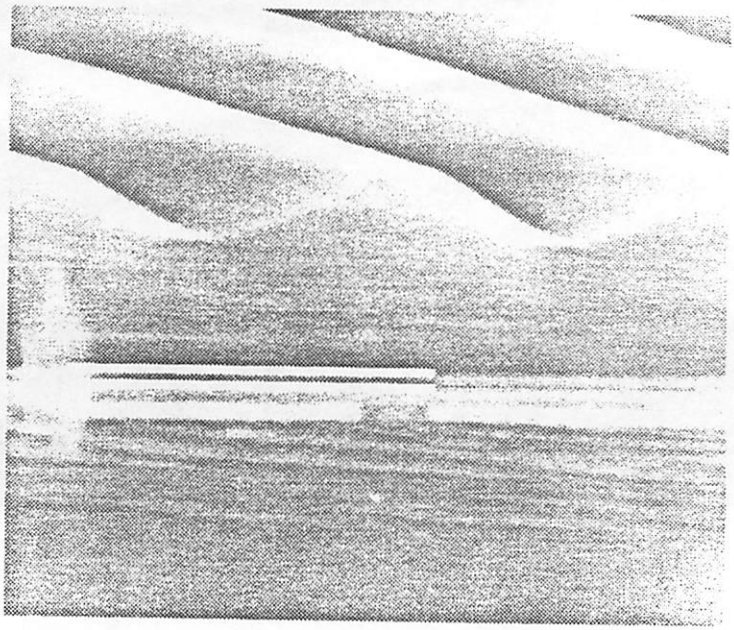
Mag = 100x

Fig. P6

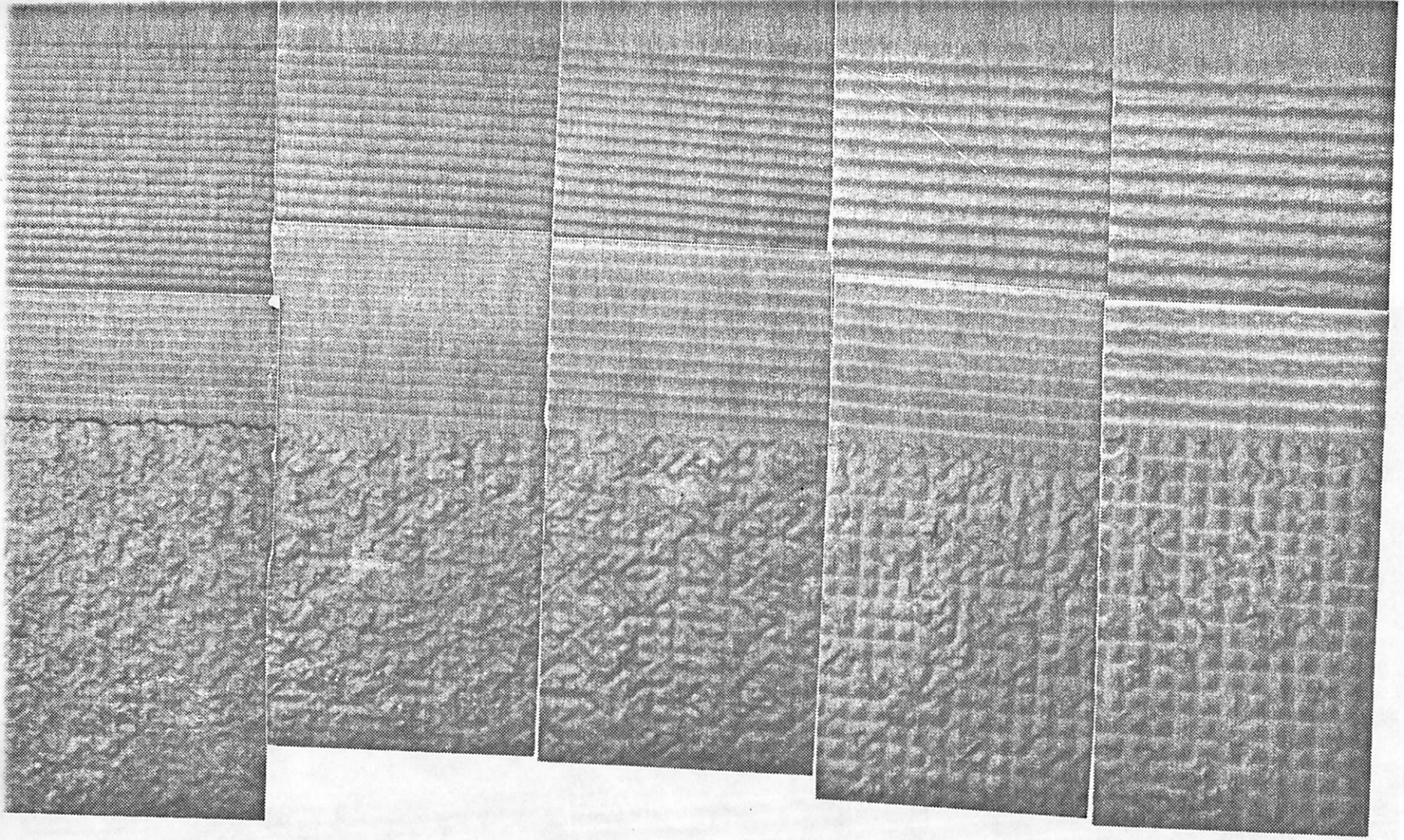
M= 1.78Kx



M= 13.16Kx



M=9.8Kx



E12
Epi thickness =
5.2 μm

THICK EPI
(100) FLAT

P=2.5
W=0.5

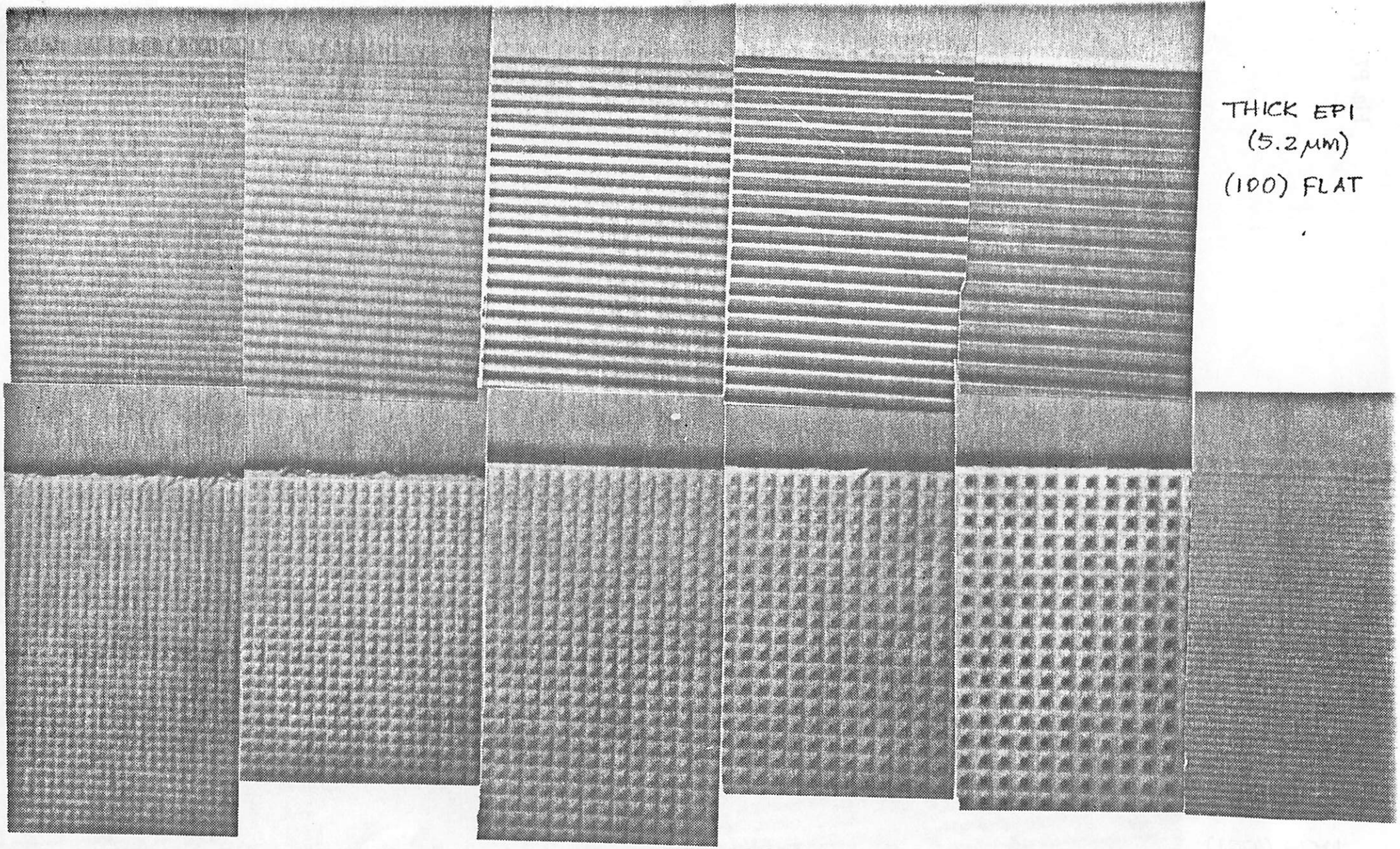
P=3.0
W=0.5

P=3.5
W=0.5
S=3.0

P=4.0
W=0.5
S=3.5

P=2.0 W=0.5
S=1.5

Fig. P7



P=2.0
S=1.0
W=1.0

P=2.5
S=1.0
W=1.5

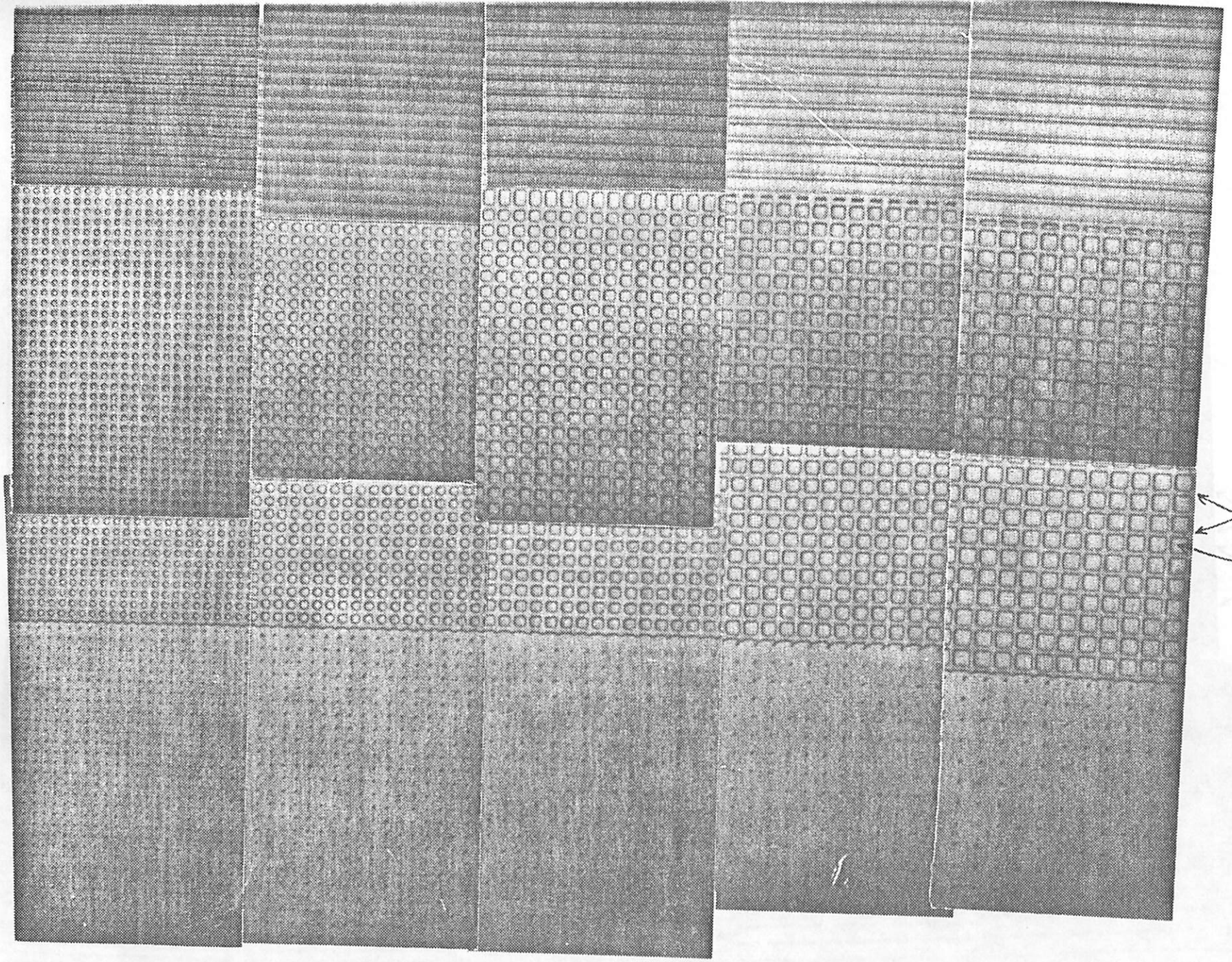
P=3.0
S=1.0
W=2.0

P=3.5
S=1.0
W=2.5

P=4.0
S=1.0
W=3.0

P=1.5
S=1.0
W=0.5

Fig. P1



P=2.0 ω=0.5 s=1.5

P=2.5 ω=0.5 s=2.0

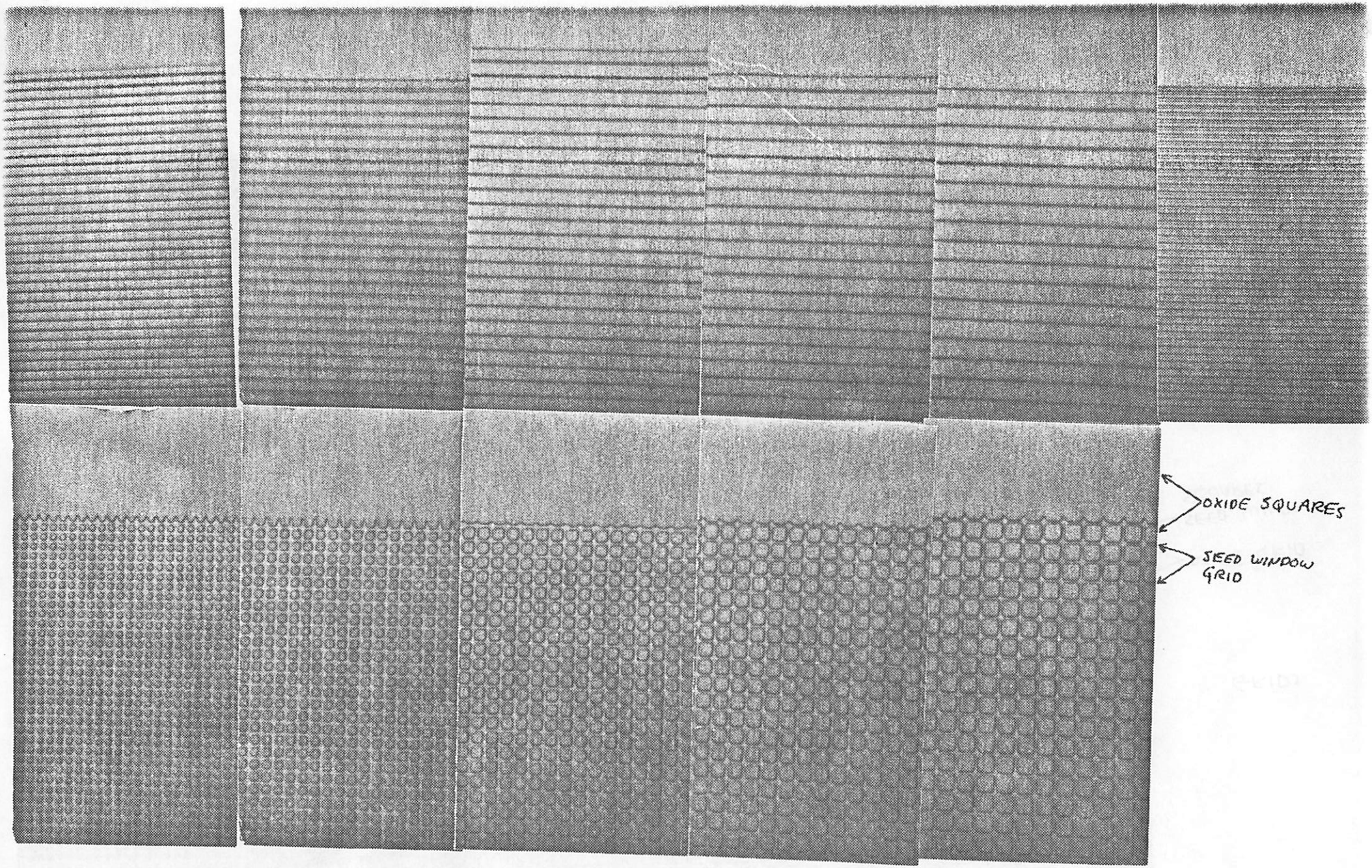
P=3.0 ω=0.5 s=2.5

P=3.5 ω=0.5 s=3.0

P=4.0 ω=0.5 s=3.5

Mag=100X

Fig. P9



P=2.0
S=1.0

P=2.5
S=1.0

P=3.0
S=1.0

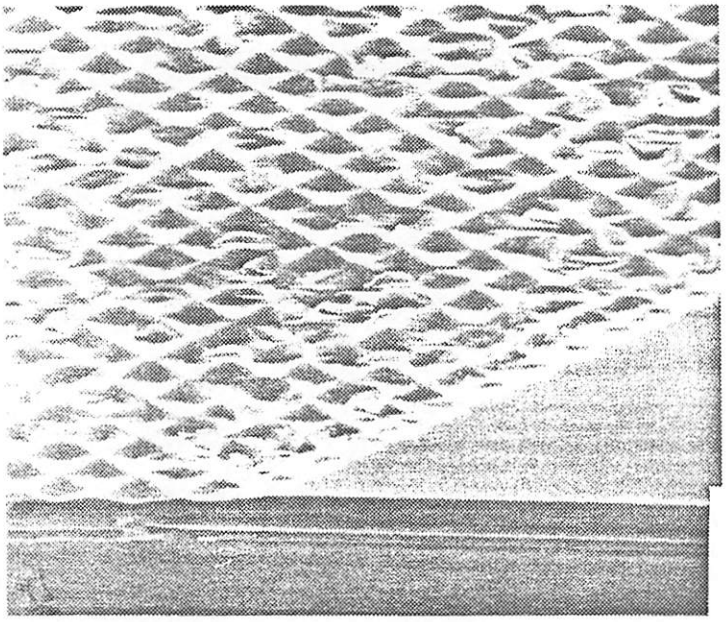
P=3.5
S=1.0

P=4.0
S=1.0

P=1.5
S=1.0

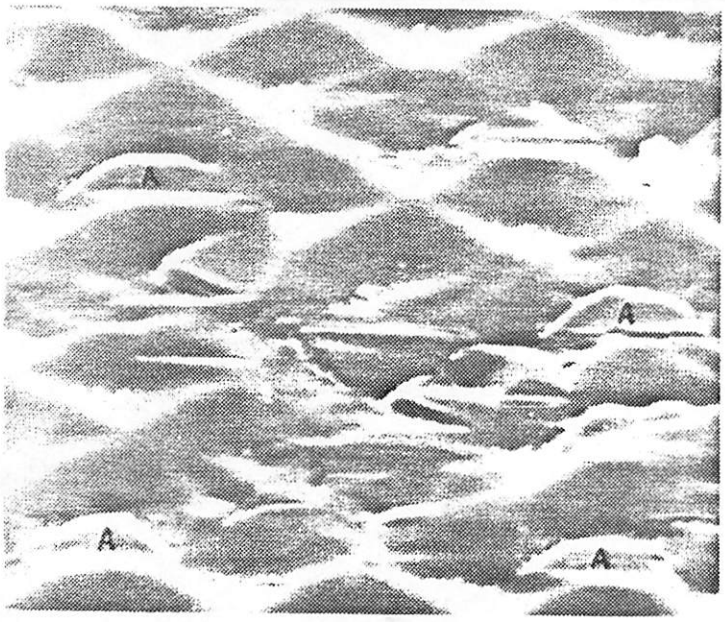
Fig. P10

M = 2.48 Kx



M = 7.76 Kx

M = 7.83 Kx



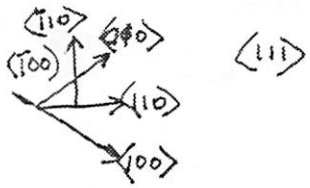
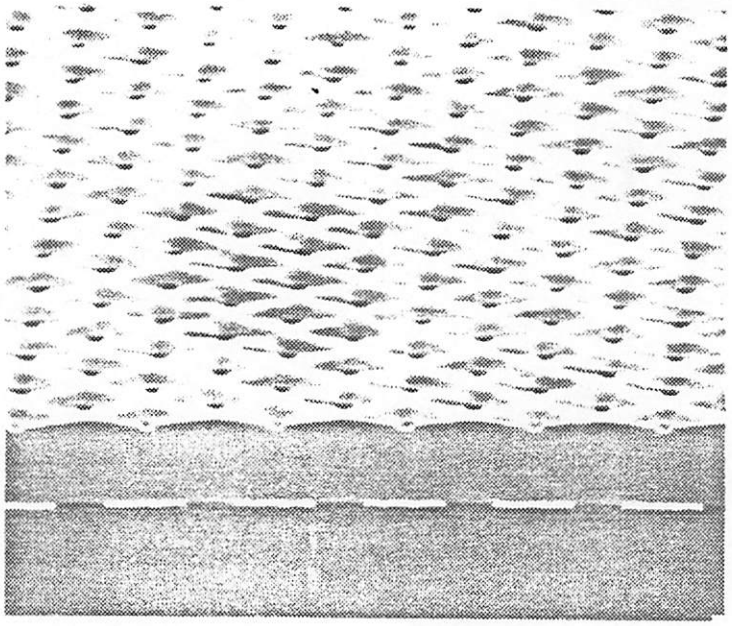
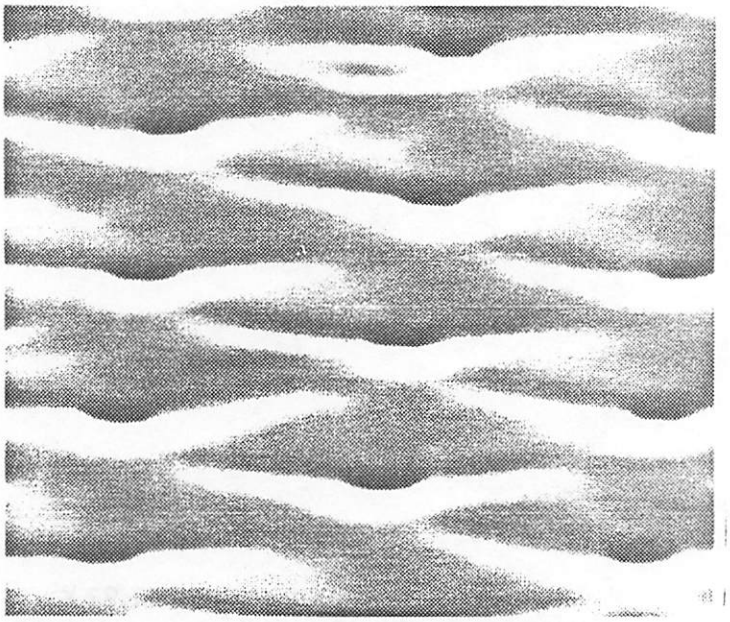


Fig. P1.

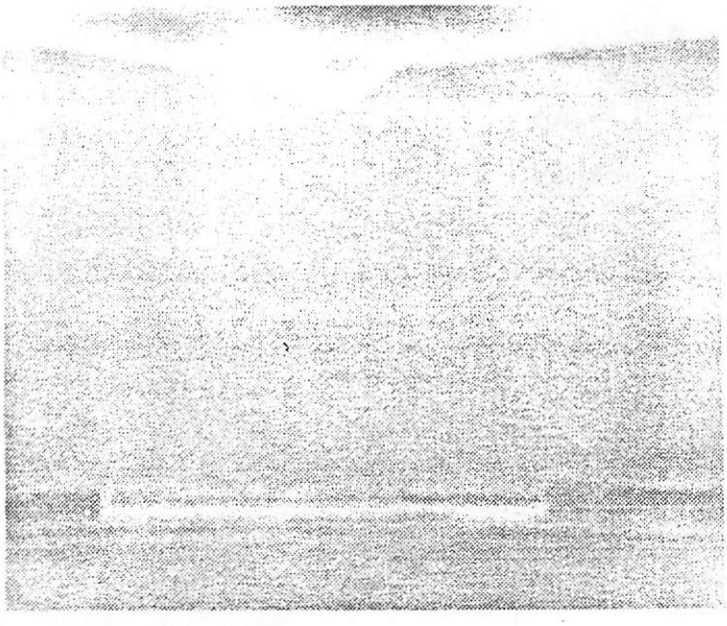
M = 4.63 Kx

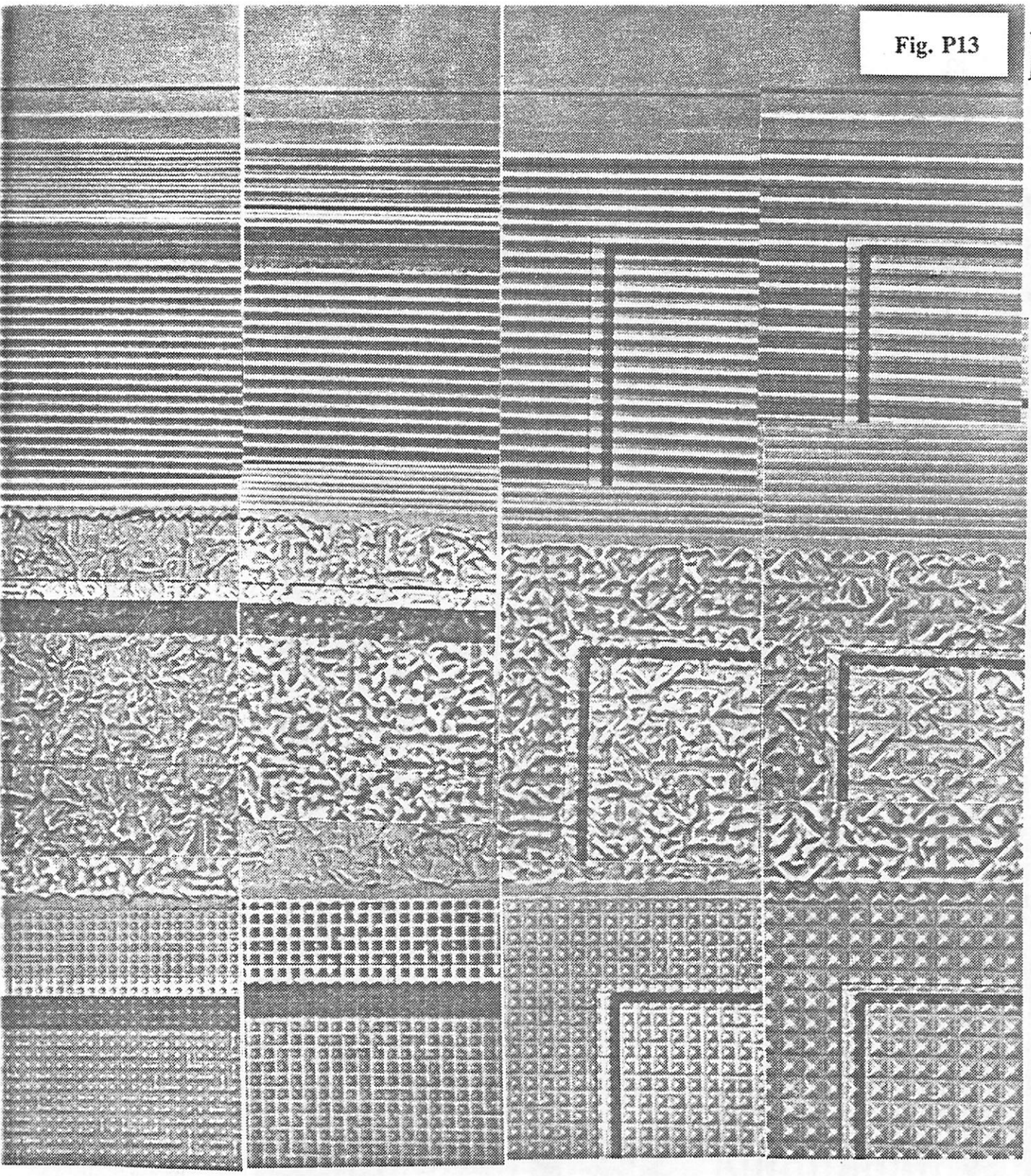


M=19.37 Kx



M= 25.9Kx





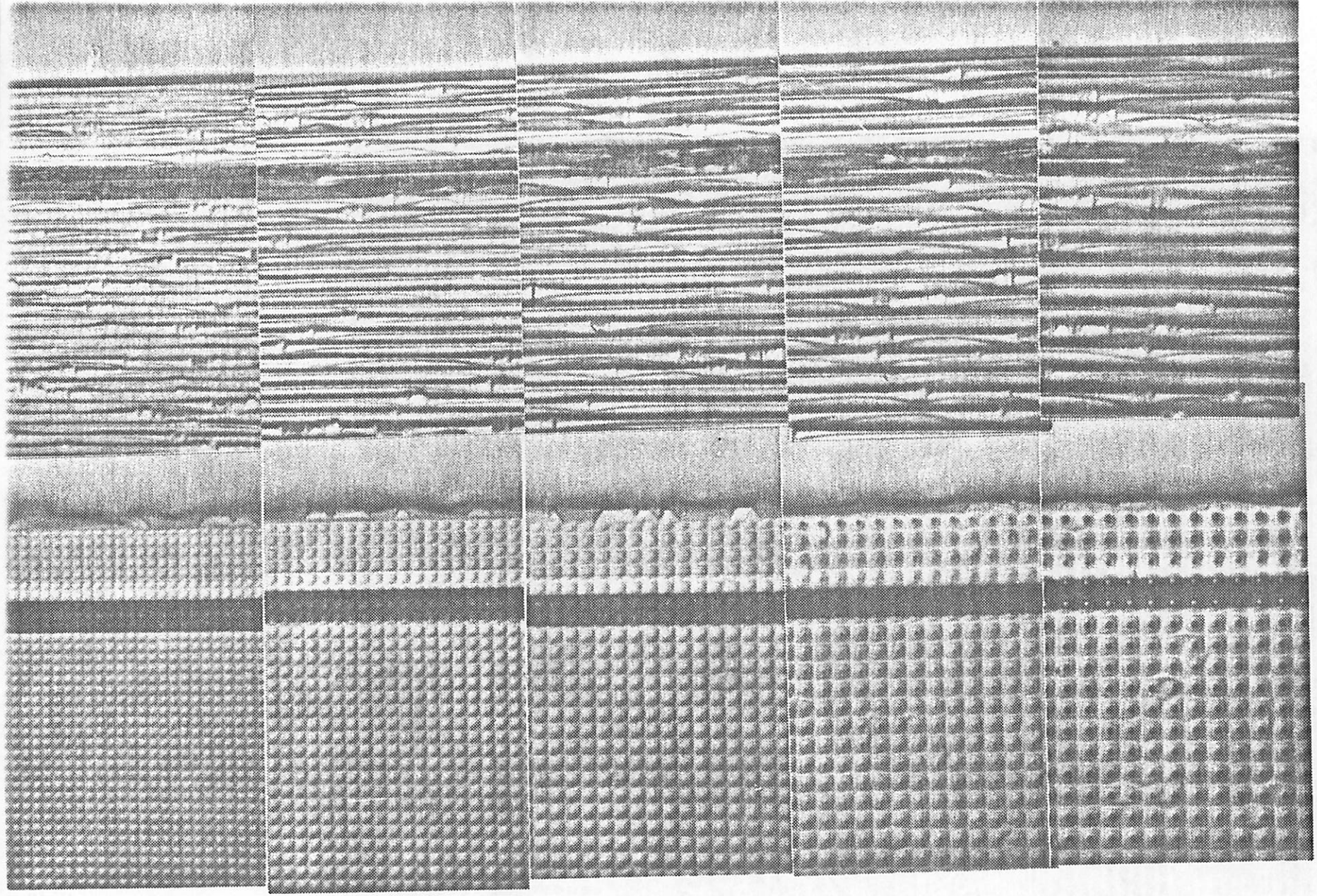
P=2.0
W=0.5
S=1.5

P=2.5
W=0.5
S=2.0

P=3.0
W=0.5
S=2.5

P=4.0
W=0.5
S=3.5

Mag=100X



THIN D5
EPI (3.3μm)
(110) FLAT

LINES

SQUARES

P=1.6
S=1.0
W=1.0

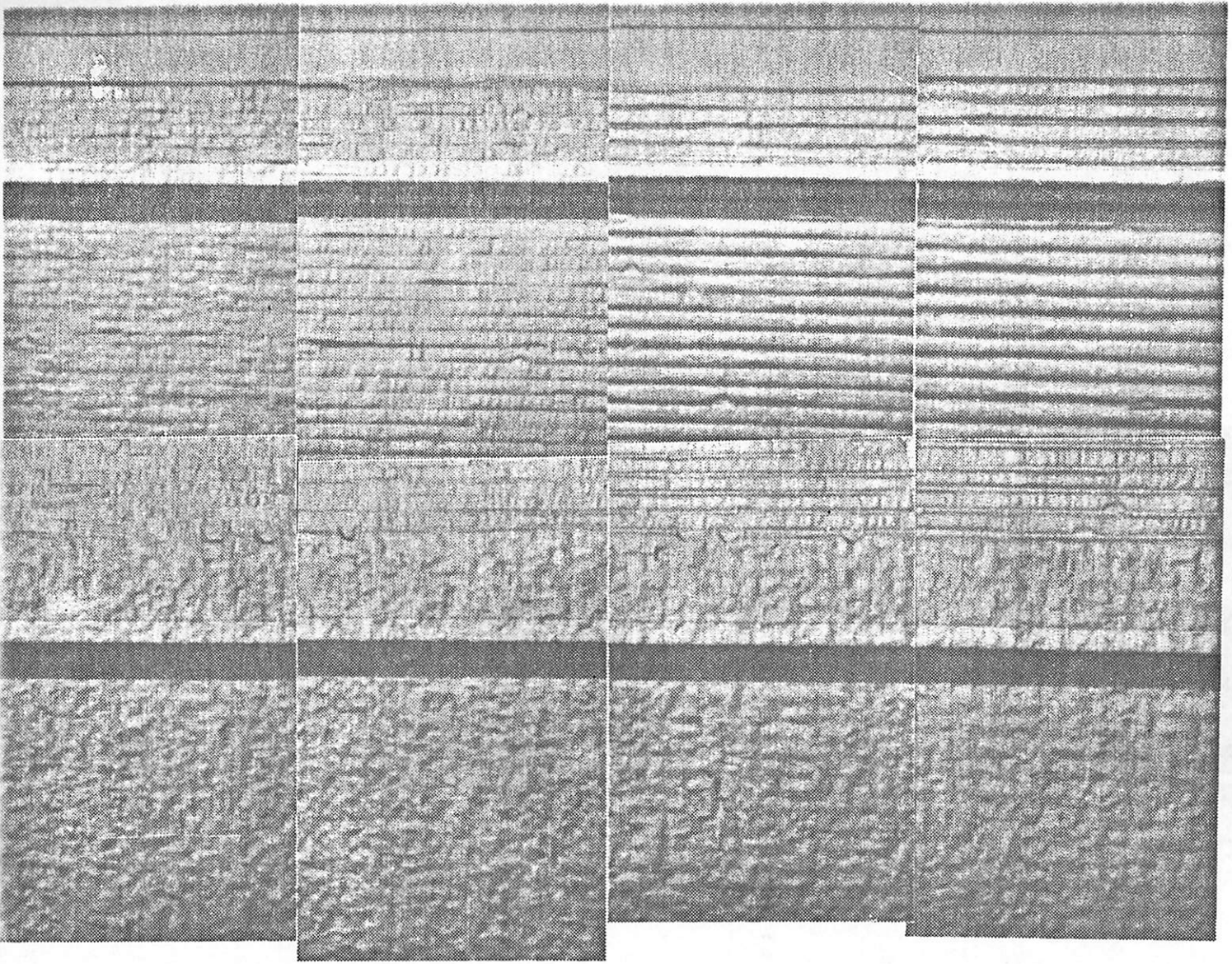
P=2.5
S=1.0
ω=1.5

P=3.0
S=1.0
ω=2.0

P=3.5
S=1.0
ω=2.5

P=4.0
S=1.0
ω=3.0

Fig. P14



DS

THIN EPI (33)

(110) FLAT

LINES

worse in smaller seeds

GRIDS

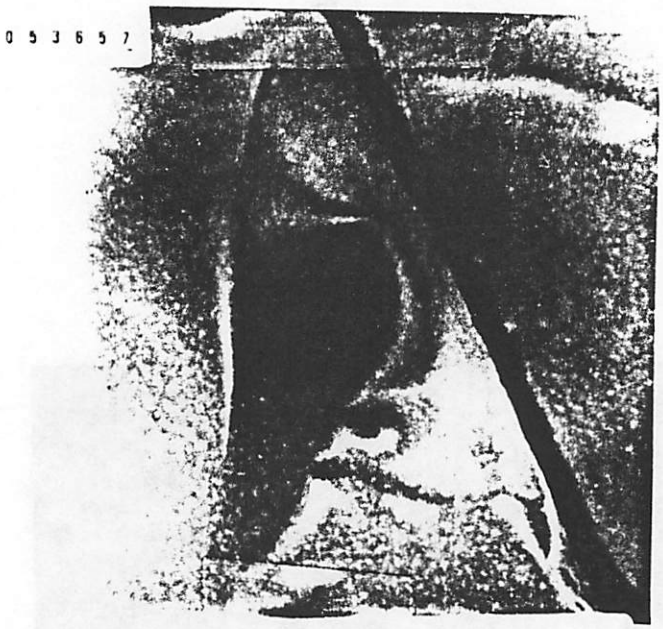
no visible pattern

P=2.0	P=2.5	P=3.0	P=4.0
w=0.5	w=0.5	w=0.5	w=0.5
s=1.5	s=2.0	s=2.5	s=3.5

Fig. P1.

MAG = 100λ

Fig. TEM1



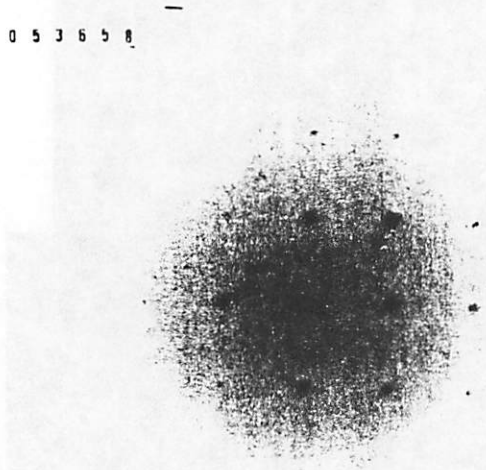
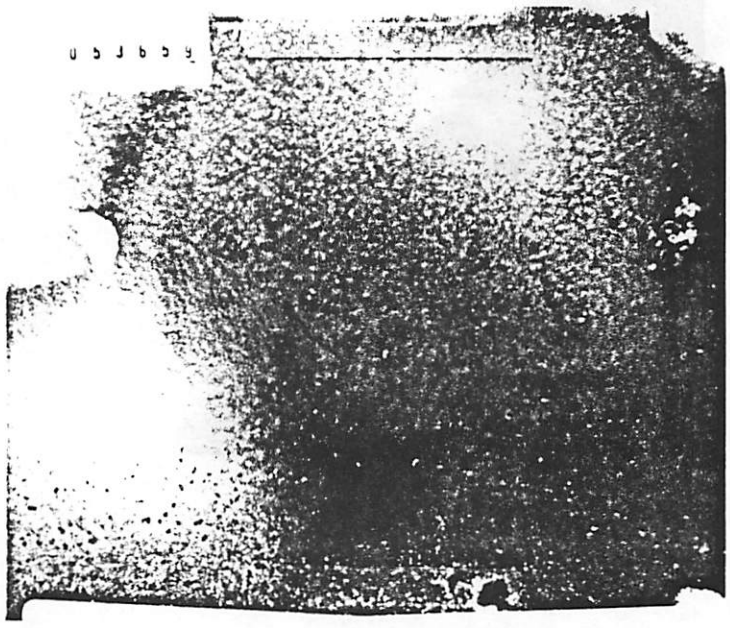


Fig. TEM2



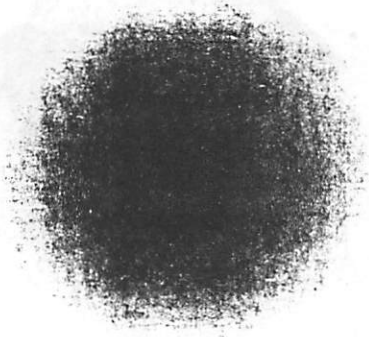
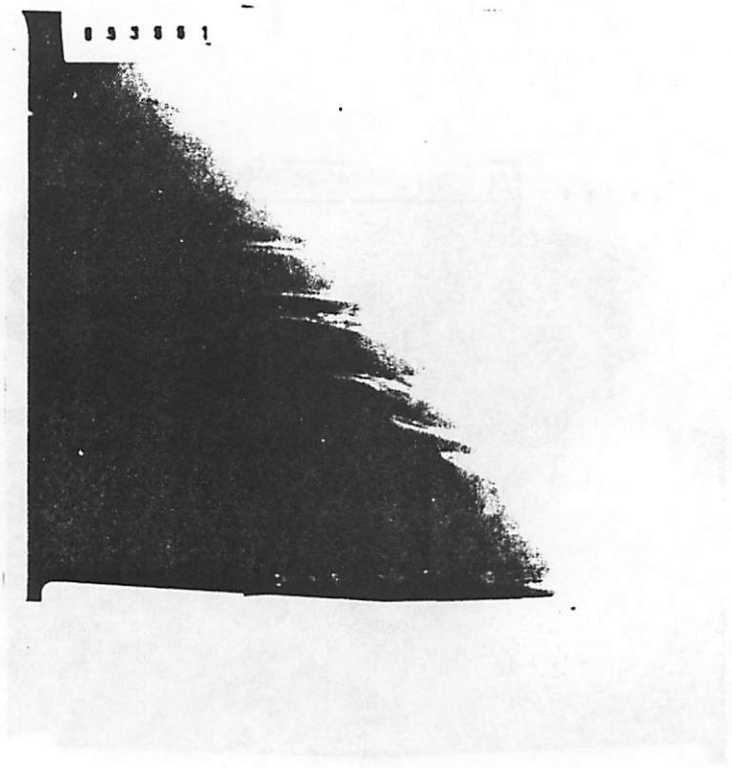
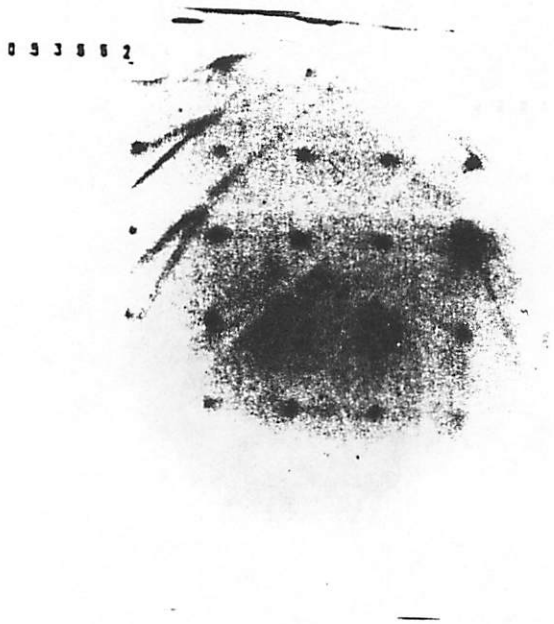
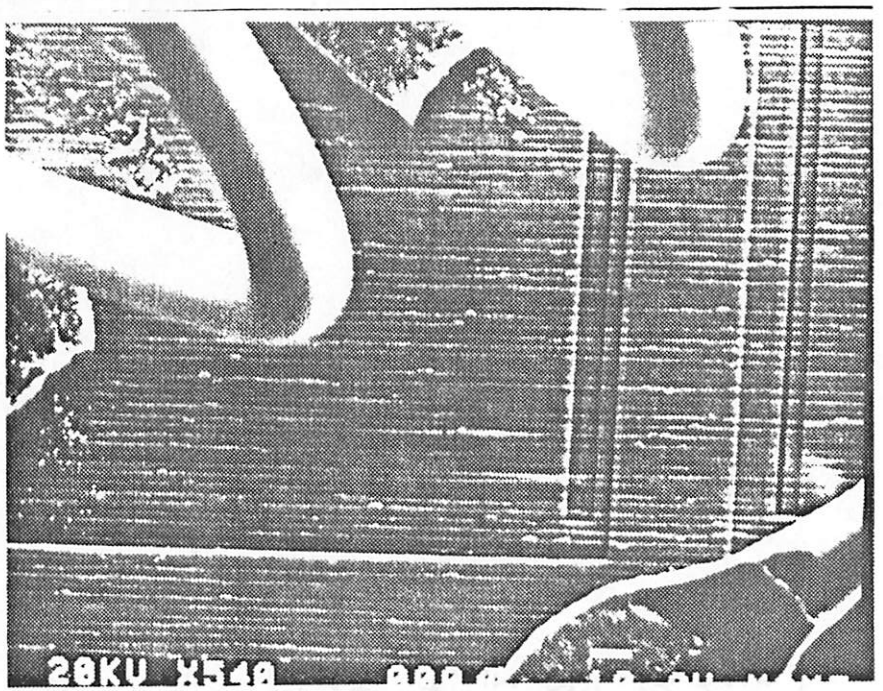
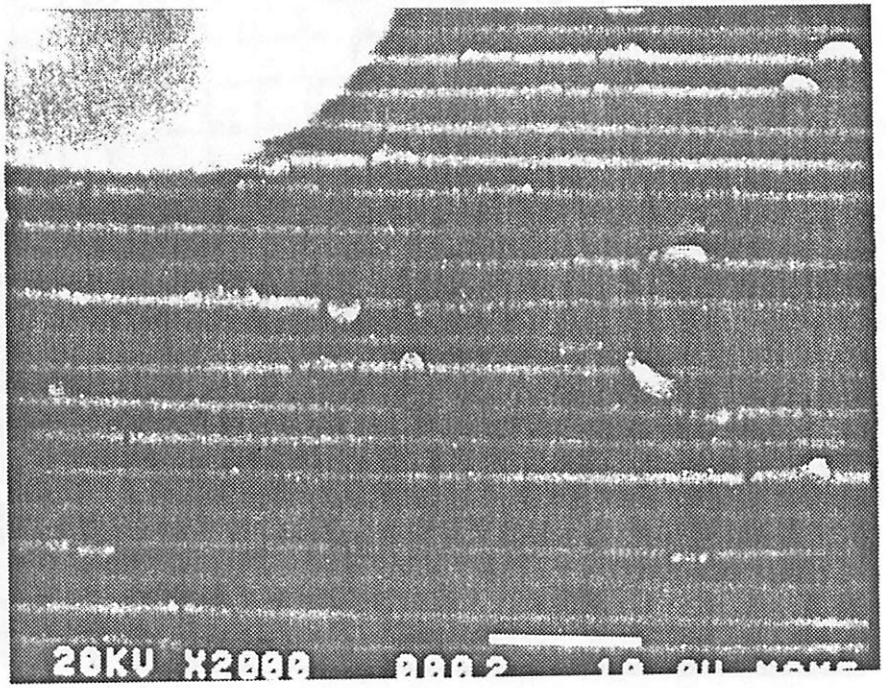


Fig. TEM3









E8 (100) FLAT THIN EPI (3.3μm)

P=3.5 (OXIDE VARIATION)
W=2.5
S=1.0

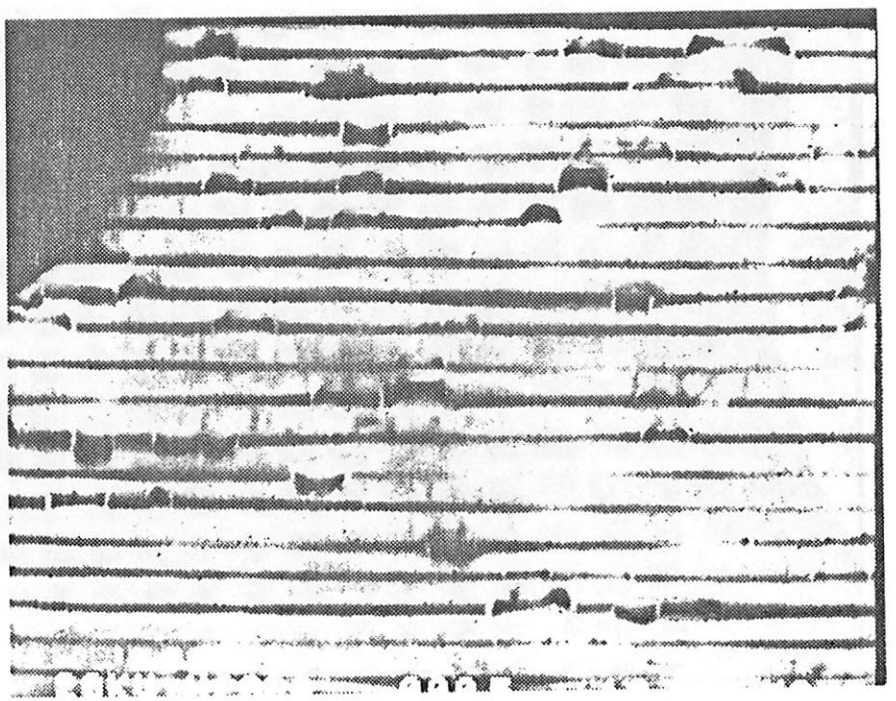
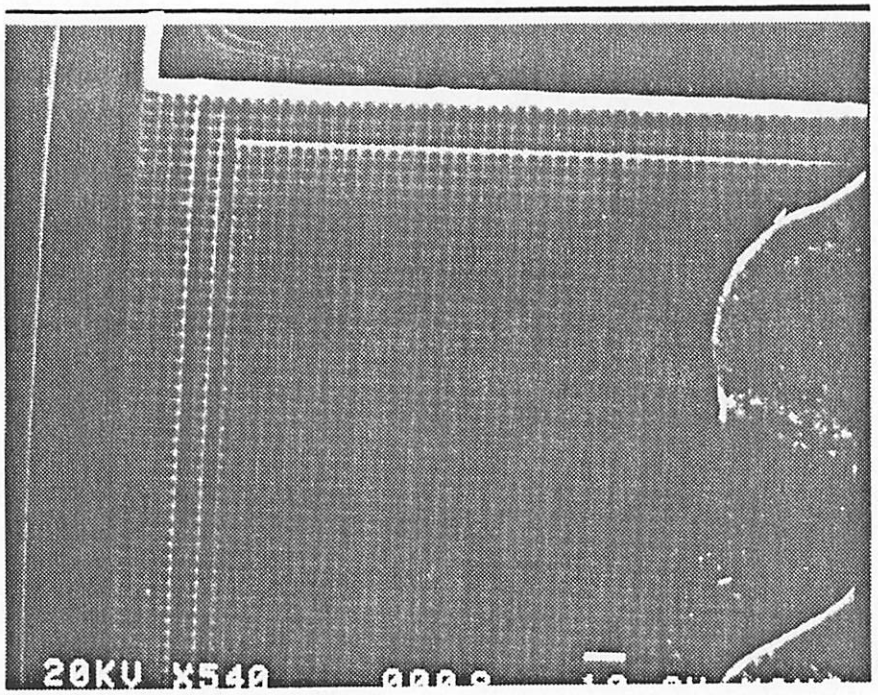
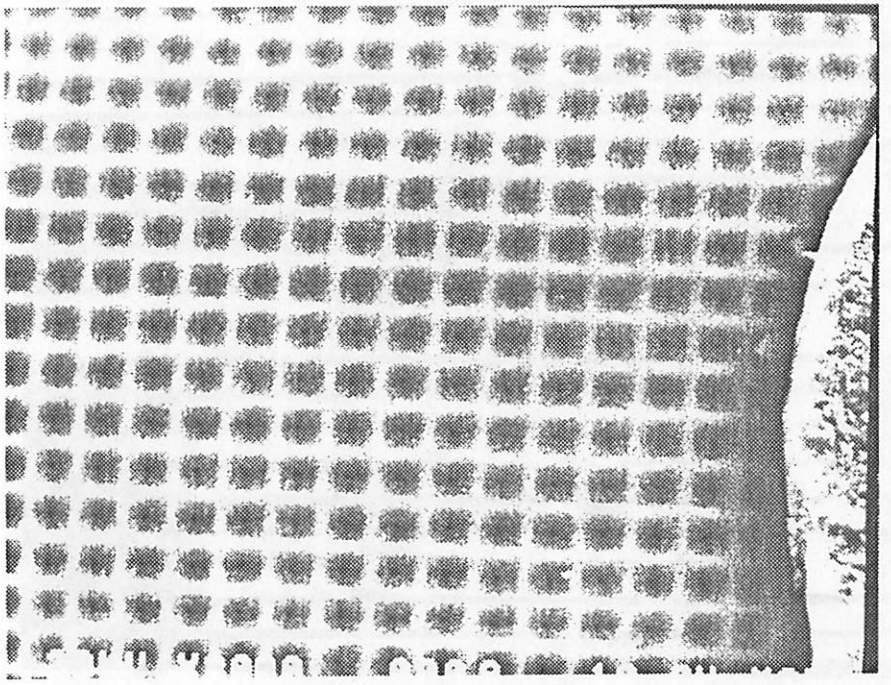


Fig. EBIC1





E8 (100) FLAT THIN EPI (3.3μm)

P=2.5
W=0.5
S=2.0
} SEED WINDOW EXPT.

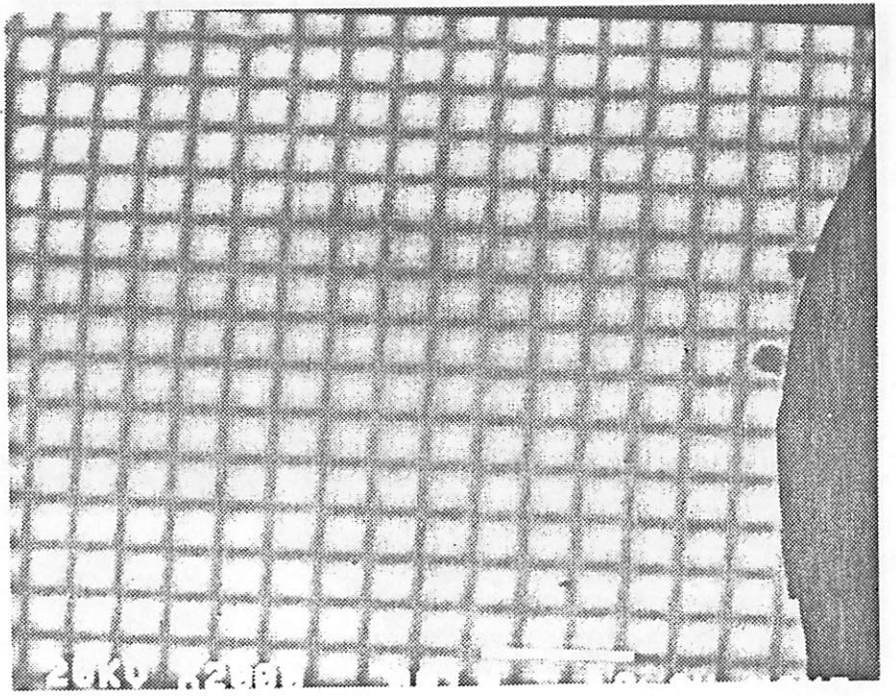
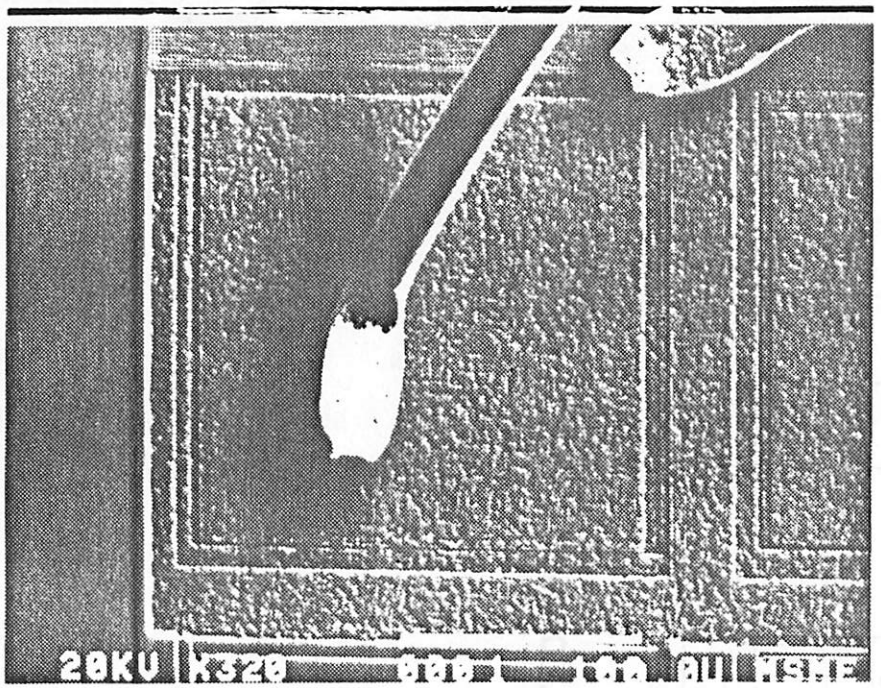
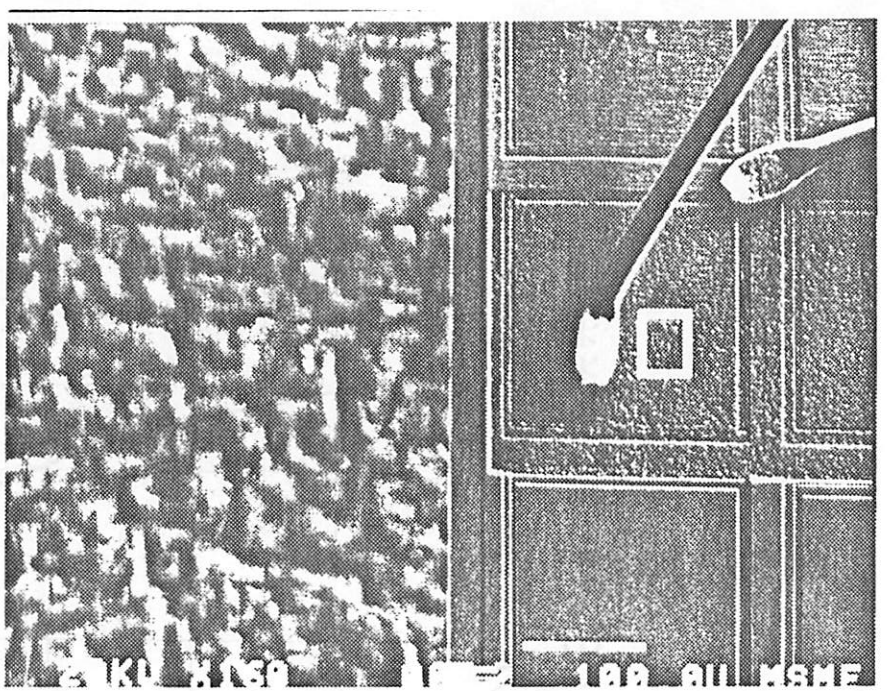


Fig. EBIC2



SEI



SEM EBIC
GRIDS
P= 2.0
W= 0.5
S= 1.5

E8 (100) FLAT
THIN EPI (3.3 μm)

SEED WINDOW EXPT.

EBIC

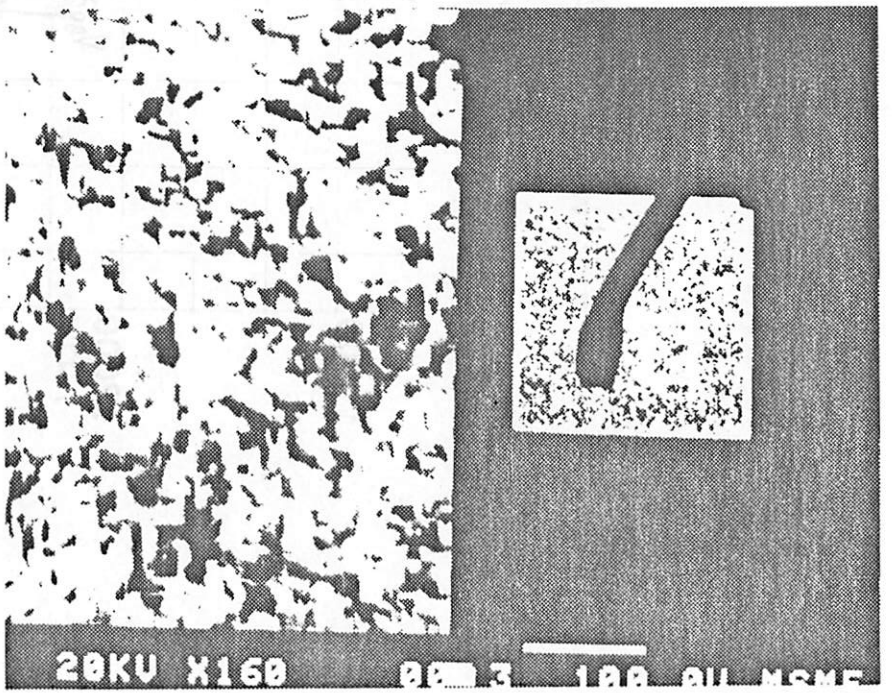


Fig. EBIC3

Fig. 21

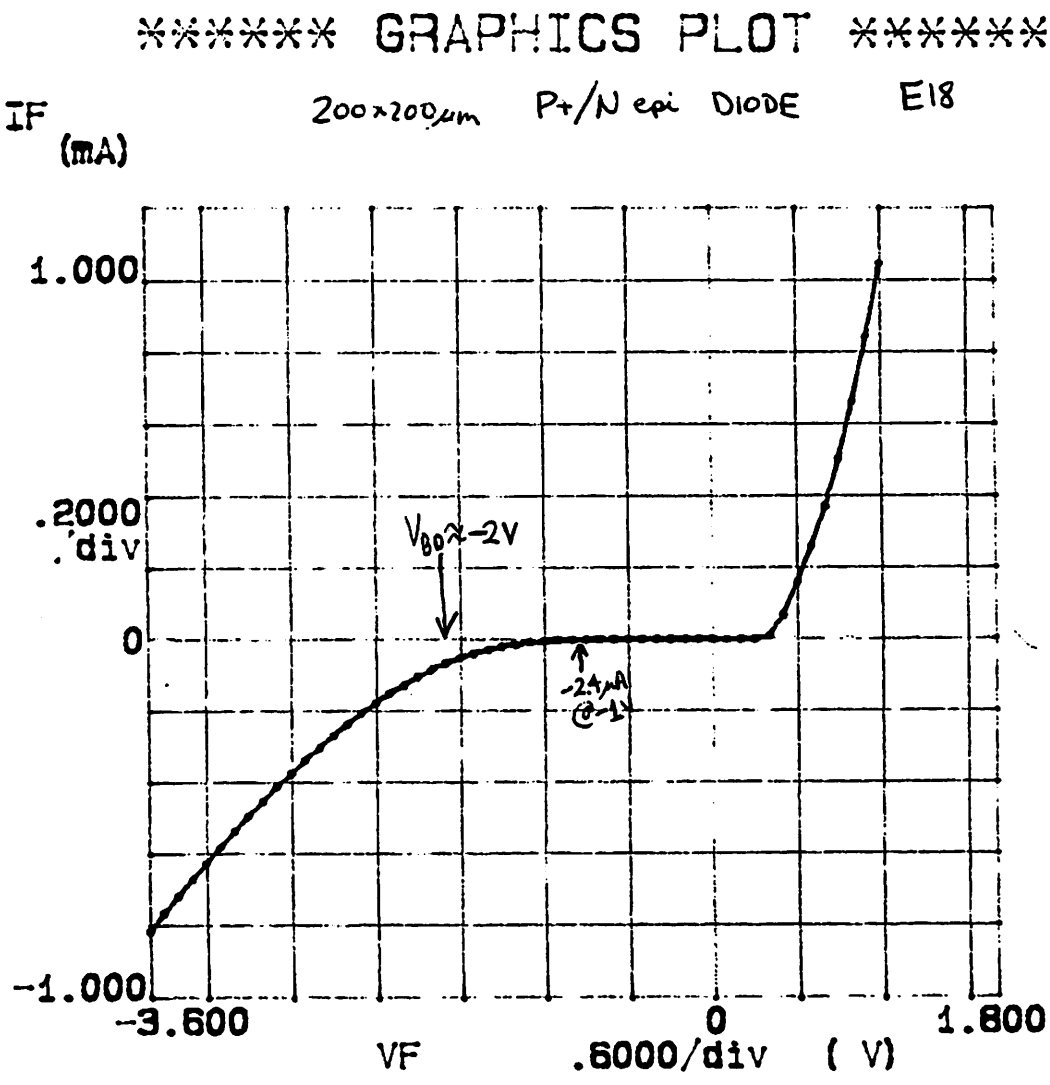
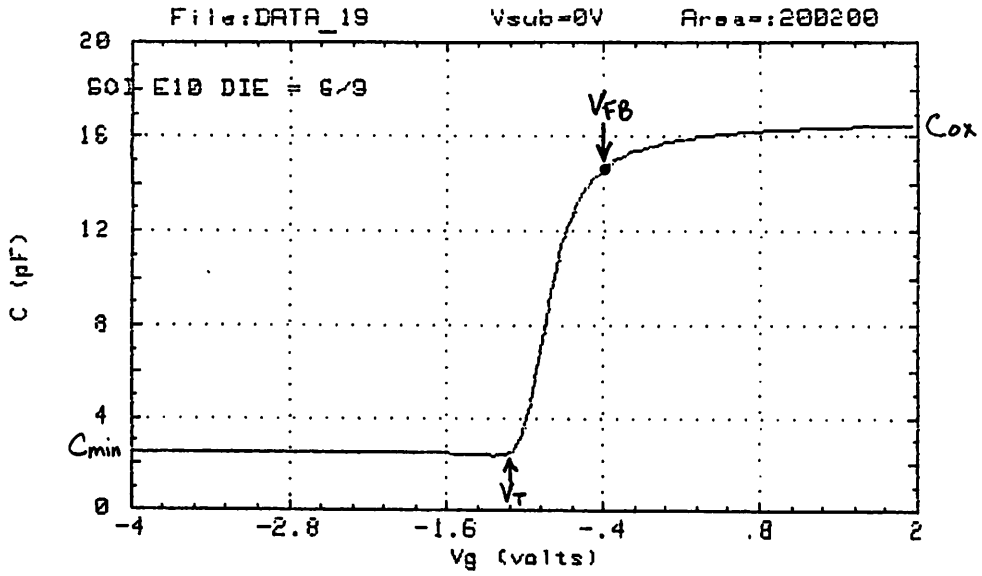
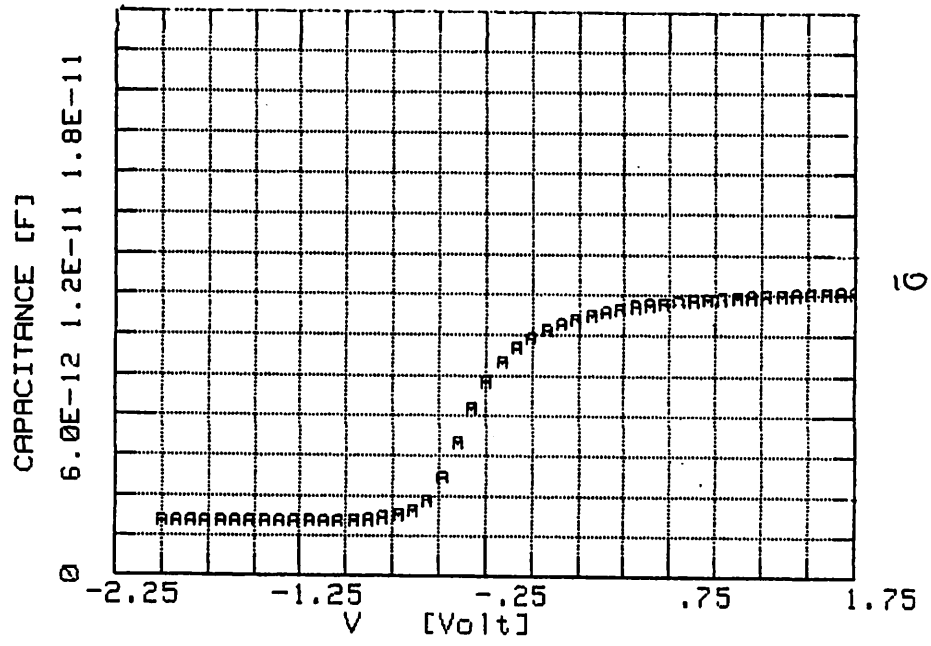

****** GRAPHICS PLOT ******
IF
(mA)
200x200μm P+/N epi DIODE E18
1.000
.2000
/div
0
-1.000
$V_{BD} \approx -2V$
-2.4μA
@-1V
-3.600
0
1.800
VF
.6000/div
(V)

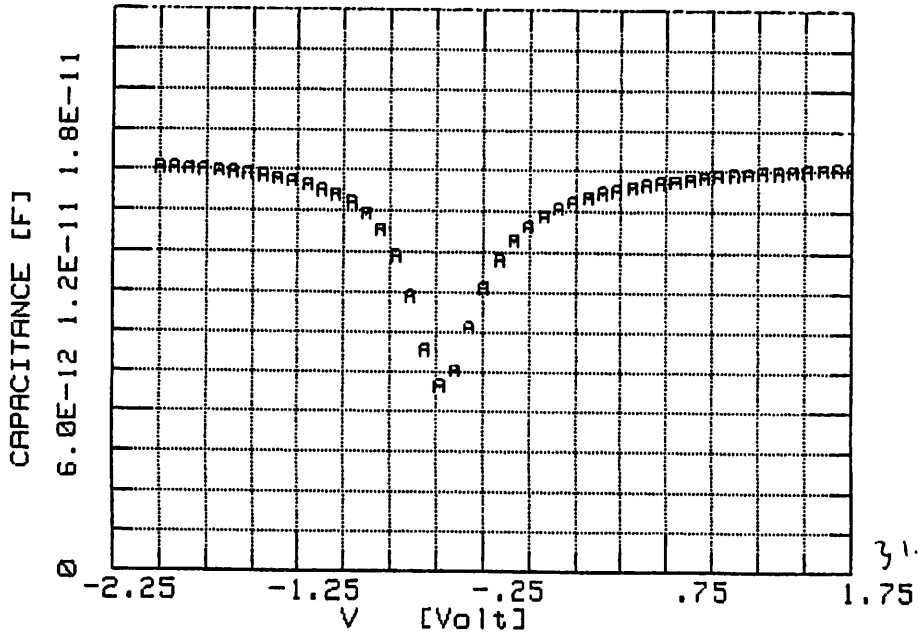
Fig. 22





10.59 +2V

light off

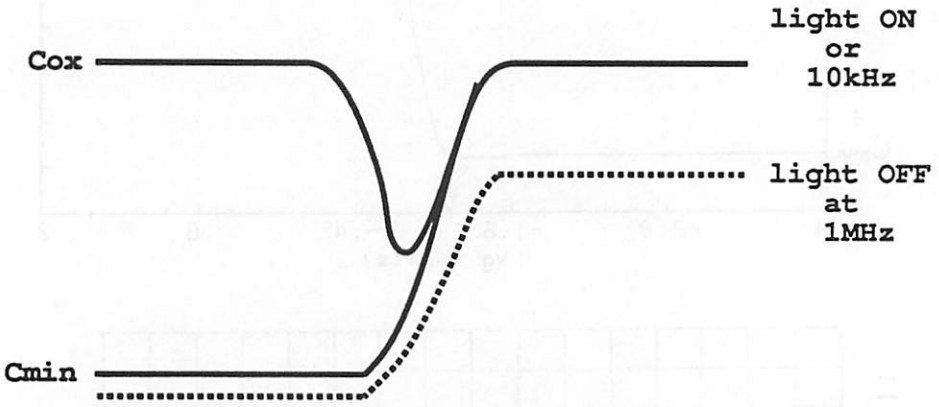


15,10 +2V

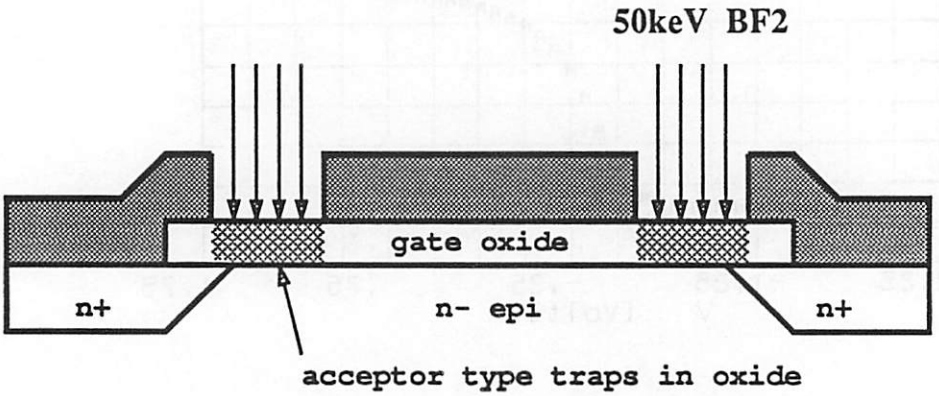
light on

31.5

MOS Capacitor Problem



All wafers receiving the boron implant exhibited light and frequency sensitive behavior in accumulation.



We believe the boron implant generated acceptor type traps in the gate oxide

These traps are neutral when filled (light on / low freq) and therefore do not affect the C-V curve.

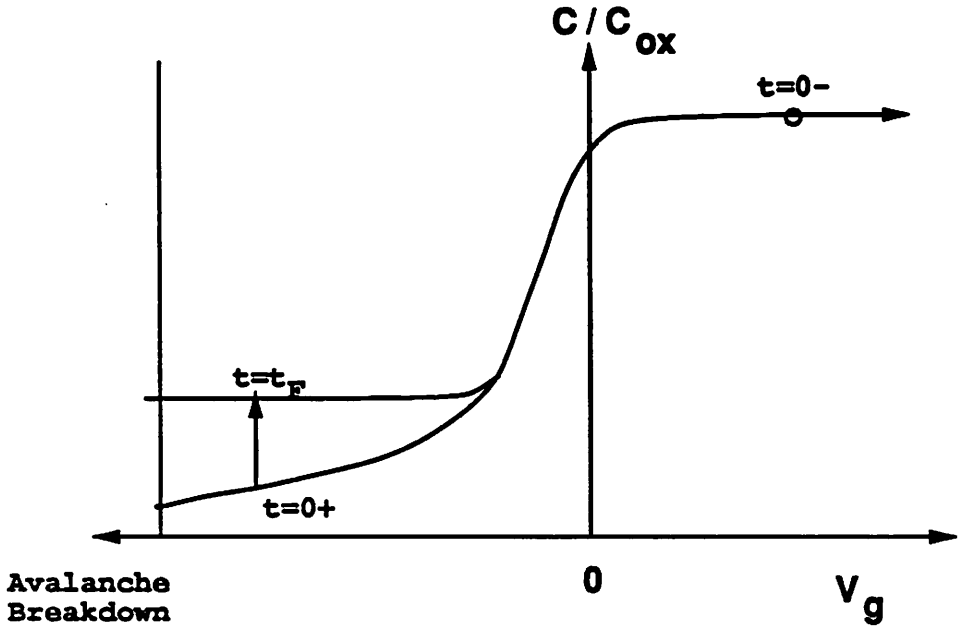
However, these traps cannot respond to 1Mhz with the light off, and remain negatively charged creating a depletion region in the silicon.

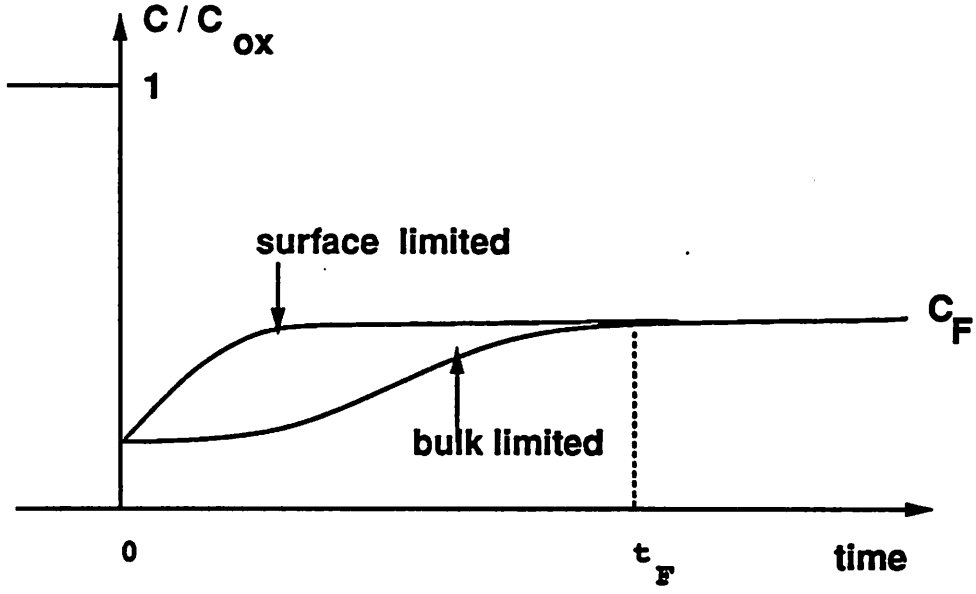
This depletion region affects the accumulation portion of the C-V

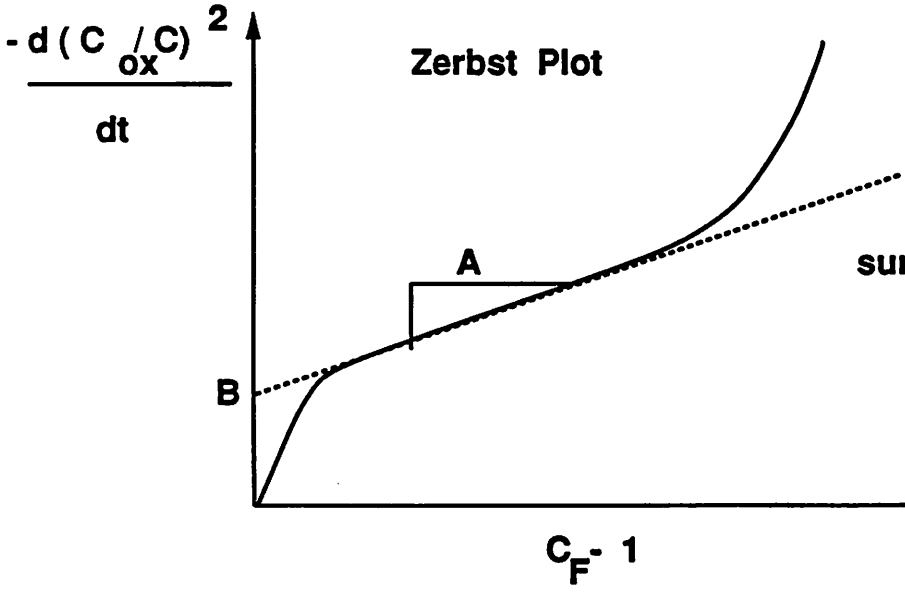
We were forced to take C-t data on the HP4280a at 1MHz with the light off.

Fig. 25

DEEP-DEPLETION RECOVERY ANALYSIS

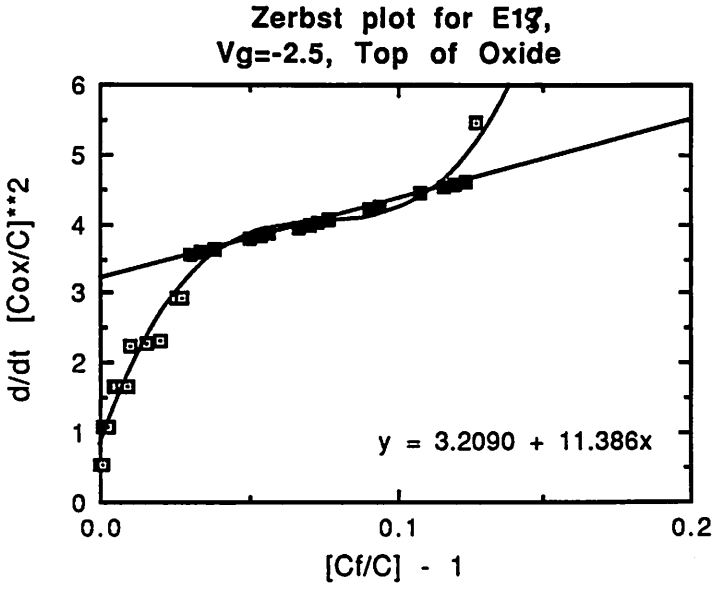


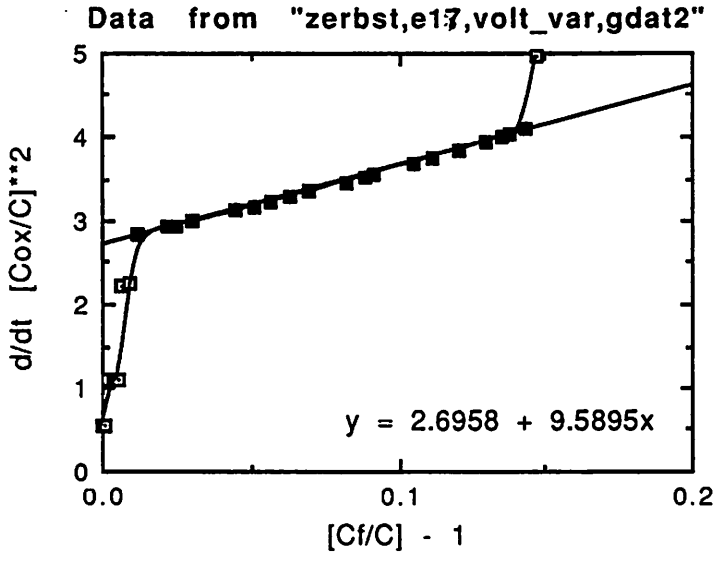


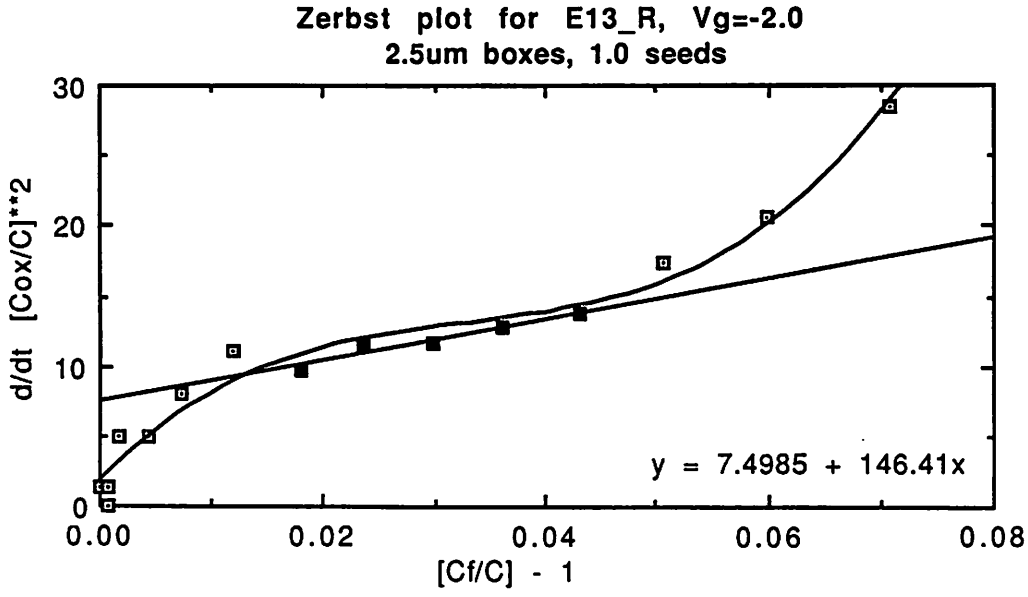


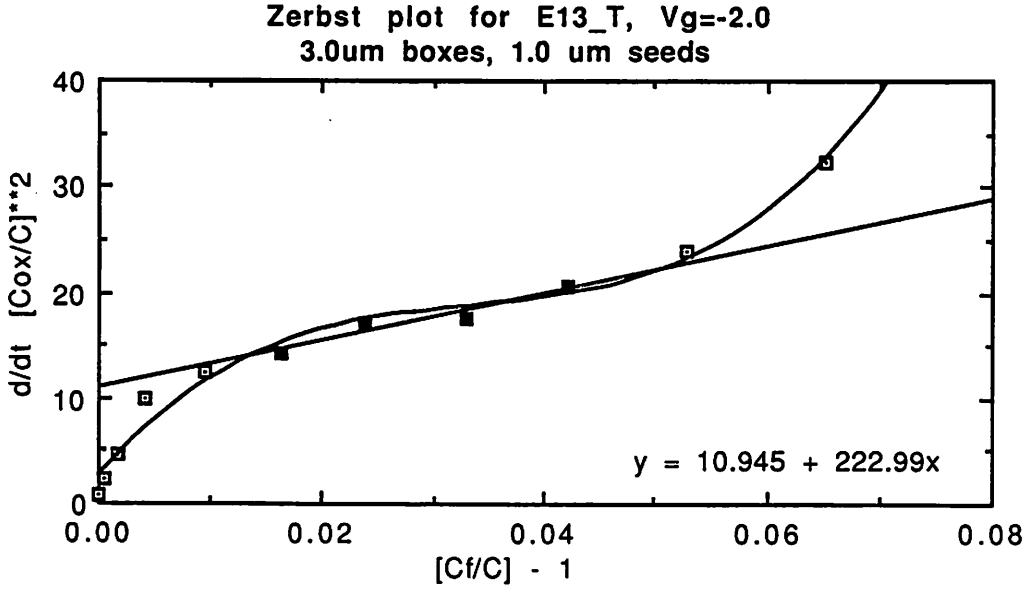
$$\text{bulk lifetime} = \frac{2 n_i C_{ox}}{A N_{epi} C_F}$$

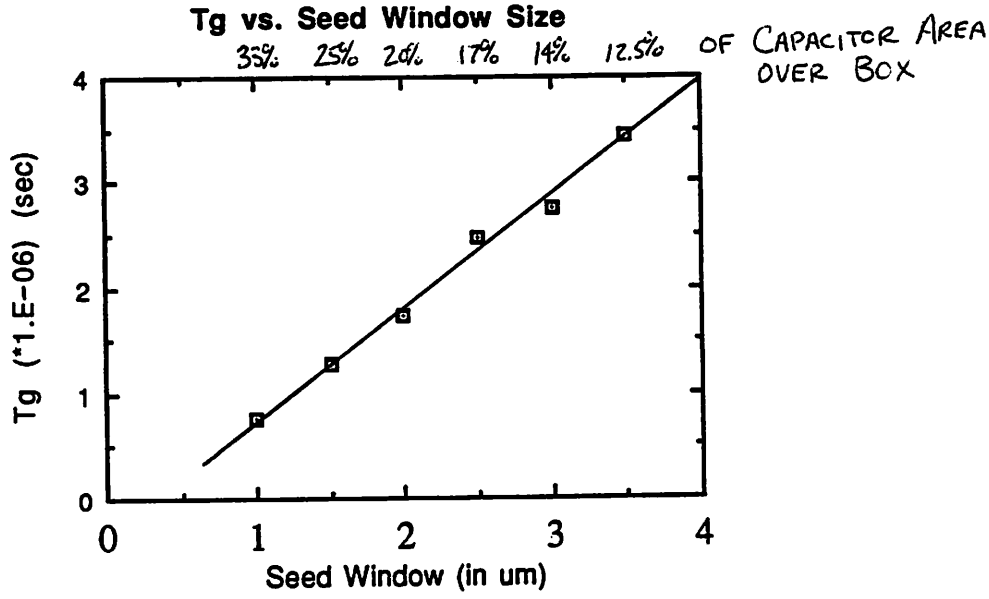
$$\text{surface recomb velocity} = B t_{ox} \frac{N_{epi} e_s}{2 n_i e_{ox}}$$

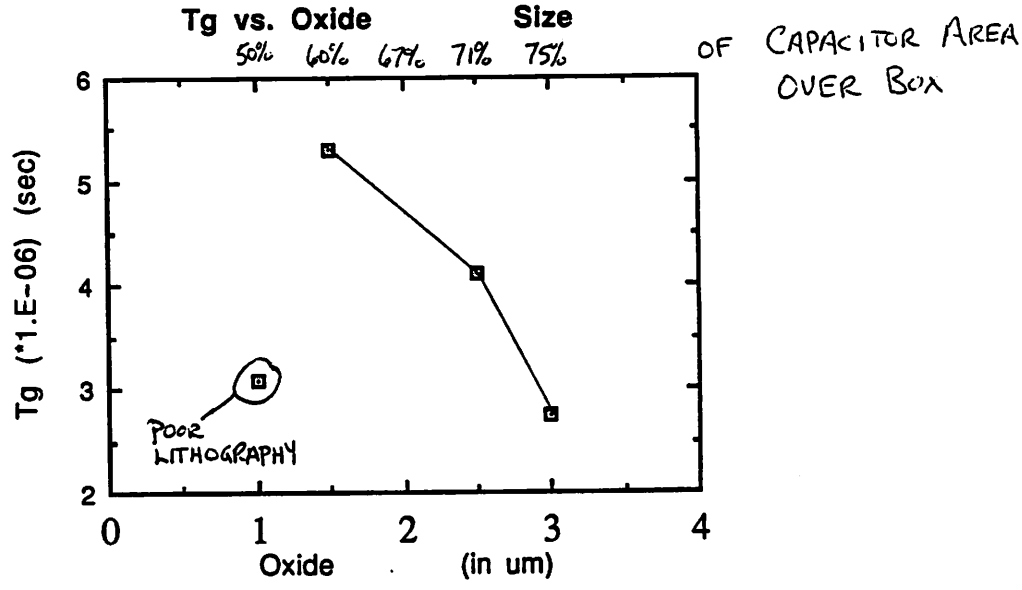

Zerbst plot for E17,
Vg=-2.5, Top of Oxide
d/dt [Cox/C]**2
[Cf/C] - 1
y = 3.2090 + 11.386x
0
1
2
3
4
5
6
0.0
0.1
0.2


Data from "zerbst,e17,volt_var,gdat2"
d/dt [Cox/C]**2
[Cf/C] - 1
y = 2.6958 + 9.5895x
0
1
2
3
4
5
0.0
0.1
0.2


Zerbst plot for E13_R, Vg=-2.0
2.5um boxes, 1.0 seeds
d/dt [Cox/C]**2
[Cf/C] - 1
y = 7.4985 + 146.41x
0
10
20
30
0.00
0.02
0.04
0.06
0.08


Zerbst plot for E13_T, Vg=-2.0
3.0um boxes, 1.0 um seeds
d/dt [Cox/C]**2
[Cf/C] - 1
y = 10.945 + 222.99x
0
10
20
30
40
0.00
0.02
0.04
0.06
0.08


Tg vs. Seed Window Size
35% 25% 20% 17% 14% 12.5% OF CAPACITOR AREA OVER BOX
Tg (*1.E-06) (sec)
Seed Window (in um)
0
1
2
3
4


Tg vs. Oxide Size
50% 60% 67% 71% 75% OF CAPACITOR AREA OVER BOX
Tg (*1.E-06) (sec)
Oxide (in um)
POOR LITHOGRAPHY
2
3
4
5
6

Liquid-Phase Deposition (LPD) of Glass

Rod Alley
Joe King
Clayton Yee

Professors William G. Oldham & Andrew Neureuther
Fall 1990 - EECS 290N

Abstract

The goal of this project was to implement an LPD oxide process in the Berkeley Microlab similar to that reported by Nippon Sheet Glass researchers. Deposition of thin oxides and the testing of their electrical characteristics was emphasized. Results indicate the desired material was not obtained. The material deposited was hydrophobic, exhibited self-limiting growth, and was a very leaky insulator. Part of the difficulty appears to have been formation of silicon hydride stain films in place of or in addition to the LPD oxide.

Background

The Background Publications List that immediately follows is referred to in this section of the paper only, and not in later sections. This reference list includes *all* LPD-related references that were discovered in English and other languages, even though the Japanese-language papers couldn't be obtained (the comments appended to these were obtained from DIALOG database abstracts). This list is intended to provide an indication of the present state-of-the-art and its antecedents, which we considered in defining this project. In brief, the LPD process has been used almost exclusively by one company in Japan in a variety of applications under development. This builds on earlier work at RCA and Honeywell in which the properties of saturated HF-silica-water solutions were investigated, and such solutions were used to selectively etch glass surfaces. Please note that citations in the later sections of this report refer to a different reference list found at the end of the report.

Liquid-Phase Deposition of SiO_2 is the deposition of flouride glass by precipitation onto wetted solid surfaces from a supersaturated silica *solution*; i.e., the silica is dissolved in the solvent, as opposed to being present as a *colloidal sol* (particles of silica suspended in a liquid medium), or a *gel* (a polymer network of silica particles in a liquid medium) [Iler, Chemistry of Silica, 1979]. This distinquishes LPD from other liquid-based methods of forming a glass which are commonly used in industry.

For our purposes, the first investigation of saturated HF-silica-water solutions to consider were those by Thomsen [1-5]. Two applications were found:

(1) selective etching of soda-glass surfaces to leave behind a surface depleted of alkali impurities. This modified the transmission properties of the glass and provided a more uniform substrate for subsequent optical coatings. Very tight control of the dissolved silica content was required to maintain selectivity [1-2,5] (we still don't have all these references, but have inferred their content from how they're cited elsewhere). One of the questions that may be answered by the references that weren't obtained is whether some silica *deposition* occurred during the so-called *etching* process.

(2) precipitation of the silica from the solution when KOH or NaCl where added. This wasn't so much an application as an observed effect, which was useful in determining solution make-up [3-5].

Two points were noted about the precipitation:

(1) when a rather large quantity of NaCl (much larger than would be necessary for our purposes) is added, the precipitation of silica was initiated which continued for hours.

(2) sodium fluosilicate was also precipitated. This is a systemic problem: metals added to the saturated solution will be precipitated as silicates or fluosilicates

[Iler,ib. id.] and incorporated in the deposited silica layer. Contamination of the solution is thus a problem. On the other hand, if the solution is continuously filtered, this precipitation can provide a very effective gettering mechanism for the solution.

The next work used the selective etching technique for antireflective coatings [6-8]. It's of interest because these researchers were trying to use the saturated chemistry as part of a manufacturable process, where monitoring and control were critical. There are three noteworthy points here:

(1) the saturated solution begins as a concentrated HF or H_2SiF_6 solution to which silica (SiO_2) is added to obtain saturation. Flourosilicic acid, H_2SiF_6, is preferred, no doubt because it is closer to saturation in its initial composition than is HF, and less time is thus required to obtain saturation.

(2) rather than a metal salt or hydroxide, small quantities of HF or $B(OH)_3$ are used to adjust the solution condition. The dissolved silica, in the form of fluosilicic acid (H_2SiF_6) is initially in the equilibrium

$$H_2SiF_6 + 2H_2O \rightarrow 6HF + SiO_2$$

Addition of more HF will reduce the degree of saturation. Addition of boric acid will consume some of the HF already in solution through the reaction

$$B(OH)_3 + 4HF \rightarrow \text{Products}$$

which will cause a temporary depletion of the HF and precipitation of SiO_2 via the previous reaction. Reference [5] describes the equilibrium and precipitation process in a greater detail which differs slightly from the explanation provided here, but the difference is not important for the purposes of this report.

(3) the condition of the solution can be determined by monitoring it's refractive index. Such in-process monitoring is necessary because the solution is being used for selective etching, which is an extremely sensitive process.

Most recently, researchers from Nippon Sheet Glass (NSG) have concentrated on the deposition that's possible with this chemistry [9-15] (note though, that [10] was done at Toshiba). Most of the chemistry details that are provided are the same as in [6-8]. Here a flouride glass is selectively deposited for microelectronics applications. Deposition occurs on a variety of glass surfaces as well as polysilicon, but not on photoresist. A deposition rate of ~200 Angstroms/hour is achieved by *continuously* adding boric acid to the solution. The film that results has a refractive index and permittivity somewhat lower than stoicheometric thermal silicon dioxide and other properties that approach those of thermal oxide as the deposited glass is annealed for longer times and at higher temperatures. Subsequent sections of this report will describe in greater detail the process and the properties of the LPD glass. Readers of this report who

have begun from the three English-language papers may be unaware of two points:

(1) an American patent was granted NSG for this process in 1984. The patent application is very detailed and should be very useful to anyone wanting to evaluate the LPD process in some application. On the other hand, the existence of a patent necessarily limits such application.

(2) a much greater number of publications exist in Japanese than in English. These publications apparently cover a variety of applications and include a review article. The abstracts of these publications are available in the DIALOG/INSPEC database.

Background Publications List

[1] S. M. Thomsen, U. S. Patent 2,490,662 (1949). *Unreviewed. RCA researcher.*

[2] S. M. Thomsen, *Title not provided*, in RCA Review, v. 12, (1951), pp. 143-?. *Unreviewed.*

[3] S. M. Thomsen, "High-Silica Fluosilicic Acids", in J. Amer. Chemical Society, v. 72, no. 6, (June 14, 1951), p. 2798.

[4] S. M. Thomsen, "Acidimetric Titrations in the Fluosilicic Acid System", Analytical Chemistry, v. 23, no. 7, (July 1951), pp. 973-975.

[5] S. M. Thomsen, "High-silica Fluosilicic Acids: Specific Reactions and the Equilibrium with Silica", J. Amer. Chemical Society, v. 74, (April-June 1952), pp. 1690-1693.

[6] H. Y. B. Mar, R. J. H. Lin, P. B. Zimmer R. E. Peterson, and J. S. Gross, "Optical Coatings for Flat Plate Solar Collectors" Final Report for ERDA under Contract No. NSF-C-957 (AER-74-09104), September 1975. *Unreviewed, Honeywell Systems and Research Center researchers, under government contract.*

[7] R. J. H. Lin and P. B. Zimmer, "Optimization of Coatings for Flat Plate Solar Collectors", Final Report for ERDA under Contract No. EY-76-C-02-2930.000,July 1977.

[8] R. J. H. Lin, J. C. Lee, and P. B. Zimmer, "Low-Cost Solar Antireflection Coatings", Annual Report for ERDA under Contract No. EM-78-C-04-5300, October 1979.

[9] Nippon Sheet Glass, "Surface treatment of glass containing alkali metal", in Japan Kokai Tokkyo Koho (April 1983), 3 pages. *Apparently, the first mention of the deposition technique. Japan Kokai Tokkyo Koho is frequently cited, and will be abbreviated to as JKTK below.*

[10] Nippon Sheet Glass, "Surface treatment of alkali metal-containing glasses", in Japan Kokai Tokkyo Koho (Sept. 1983), 3 pages.

[11] H. Kawahara, H. Nagayama, and H. Honda. "Silicon dioxide coating". U. S. Patent 4,468,420 (August 28, 1984), 12 pages. *This is a surprising patent document, with much more useful detail than is usually the case.*

[12] H. Nagayama, M. Hyodo, M. Misonoo, H. Honda, and H. Kawahara. "Silicon oxide films", in Japan Kokai Tokkyo Koho (Sept. 1985), 6 pages. *Japanese language, unreviewed, Nippon Sheet Glass researchers. The English-language abstract on DIALOG has a great deal of useful detail.*

[13] T. Aida, H. Nagayama, and H. Kawahara. "Formation of silicon dioxide coatings", in JKTK (Dec. 1986), 3 pages.

[14] T. Watanabe, K. Okumura, "Method for forming micropatterns", in JKTK, (June 26, 1987 (I can't follow the reference description; may have been published in 1989)), 4 pages. *Japanese language, unreviewed, Toshiba.*

[15] H. Kawahara, "Coating with silicon dioxide by LPD (liquid-phase deposition) method" in Kogyo Zairyo, v. 35, no. 9 (1987), pp. 54-57. *Japanese language, unreviewed, Nippon Sheet Glass researcher.*

[16] T. Goda, H. Nagayama, A. Hishinuma, and H. Kawahara, "Physical and chemical properties of silicon dioxide film deposited by a new process", in MRS Sym. Proc., v. 105 (1988), pp. 283-8. *Nippon Sheet Glass. Note that the listing for this paper in the '90 VLSI Sym. publication appears to be wrong.*

[17] H. Nagayama, H. Honda, and H. Kawahara, "A New Process for Silica Coating", in J. Electrochem. Soc., v. 135, no. 8, (August 1988), pp. 2013-2016. *Nippon Sheet Glass.*

[18] H. Nagao, Y. Ino, M. Misonoo, and H. Kawahara. "Surface roughening of glass plates" in JKTK (Dec. 1988), 5 pages.

[19] T. Aida, H. Nagayama, A. Hishinuma, and H. Kawahara. "Manufacture of silica films", in JKTK (Jan. 1989), 3 pages.

[20] A. Hishinuma, S. Hayashi, T. Aida, M. Kitaoka, and H. Kawahara. "Manufacturing of silicon dioxide film by anodic precipitation", in JKTK (Jan. 1989), 4 pages. *Japanese language,*

unreviewed, Nippon Sheet Glass.

[21] T. Aida, A. Hishinuma, H. Nagayama, and H. Kawahara. "Manufacture of silica films with phosphorus", in JKTK (Feb. 1989), 4 pages. *The phosphorus content is only 0.03 weight %.*

[22] S. Hayashi, A. Hishinuma, and H. Kawahara. "Manufacture of silicon oxide particles", in JKTK (June 1989), 3 pages.

[23] H. Tada, Y. Saitoh, K. Miyata, and H. Kawahara, "Effect of silica-coating over rutile titania particles on the suppression of their photocatalytic activity", in Shikizai Kyokaishi, v. 62, no. 7, (1989). *Japanese language, unreviewed, Nippon Sheet Glass researchers.*

[24] T. Aida, A. Hishinuma, and H. Kawahara. "Manufacture of grooved glass substrates for magnetic and optical disks and lenses", in JKTK (Oct. 1989), 4 pages.

[25] H. Kawahara, "Formation of SiO2 film by chemical reaction in aqueous solution", in Yoyuen oyobi Koon Kagaku, v. 33, no. 1, (1990), pp. 7-23. *Japanese language, unreviewed. It's particularly regrettable that this is in Japanese, because it's a review article with 23 references.*

[26] T. Homma, T. Katoh, Y. Yamada, J. Shimizu, and Y. Murao, "A New Interlayer Formation Technology for Completely Planarized Multilevel Interconnection by Using LPD", in 1990 Sym. on VLSI Tech. Digest, (1990), pp. 3-4. *Nippon Sheet Glass researchers.*

In addition to these, there were three Czech publications using a fluosilicic acid solution to precipitate sodium fluosilicates, and there's a good literature in oil industry publications, where HF is injected into the ground to make geological formations more porous, so that oil is more accessible from a particular well. These end up as saturated solutions (silica and alumina from the rock) and considerable work has been done on the chemistry of such systems.

Goals

The objectives were:

(1) to reproduce as closely as we could the LPD process described by Nippon Sheet Glass researchers in their 1983 patent description, modified as described in their 1988 JECS paper. We were to use this to deposit and characterize thin (less than 1000 Angstroms) oxide films.

(2) to characterize the effects of saturation temperature and of the boric acid concentration on the deposition rate, and the deposited film's electrical and material properties. This characterization was to be done around a central operating point identical to that used by NSG researchers in their 1988 paper.

(3) to characterize the compatibility of various materials with the LPD process; i.e., whether deposition can be done on a material, and whether that material will survive the deposition process.

(4) to attempt to obtain supersaturation and deposition by raising the solution temperature and without resorting to boric acid.

By accomplishing these objectives and documenting our methodology and results, we hoped to make the LPD oxide process accessible to other UCB microlab users.

Experimental Approach

Initially a simple apparatus for depositing thin films on wafers was constructed to verify the effectiveness of the chemistry and our experimental setup.

We then investigated basic parameters of the chemical process including: deposition rate, temperature dependence, and concentration dependence. We also characterized the basic parameters of the material including etch rate, breakdown voltage, dielectric constant.

The process's compatibility with other materials was also important. It's nucleation and adhesion on nitride, poly, photoresist, tungsten, and other materials were also investigated.

For application in micromechanics, the growth of thick films (greater than one micron) is very desirable. We planned to investigate the properties of thick films.

Experimental Plan

Based on the NSG work, we studied boric acid concentrations between 2.0×10^{-2} Molar and 3.5×10^{-2} M, deposition temperatures between 25°C and 40°C, and film thicknesses up to 1000 Angstroms. We decided to concentrate on thin films because of the slow deposition rate and our interest in the electrical performance of the films which had not previously been characterized.

Apparatus

Our initial setup, to prove the feasibility of the process, consisted of a large water bath tank containing a small 500 mL beaker of deposition solution. The large tank sat atop a combination stirrer/hot plate with its internal open loop temperature controller. This setup provided very poor temperature control for the deposition solution because there was no feedback of the solution temperature to the hot plate's temperature controller.

Our final setup (Figure 1), consisted of a large outer tank containing a temperature controlled water bath. Temperature is controlled by coils containing recirculating fluid circulating through a temperature controller. A one liter beaker containing the wafers, stir bar and acid solution is surrounded by the water bath. Both the beaker and outer tank had loose fitting covers to retard evaporation and to prevent significant illumination of the solution.

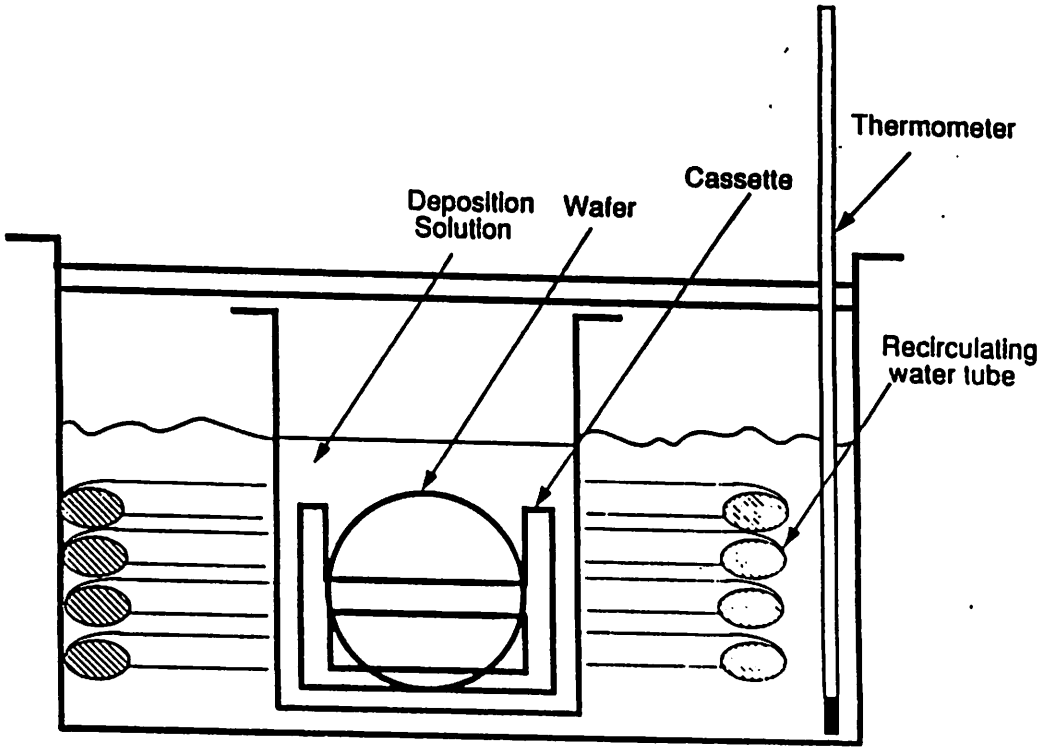
Figure 2, show a diagram of NSG's setup. This setup contains two features not found in our setup: 1) a continuously filtered and recirculated deposition solution in laminar flow across the wafer, and 2) a continuous drip source of boric acid. We add our boric acid all at once and the wafers either sit in a still solution or a swirling solution with a stirbar.

Process Flow

Our basic process steps were 1) mixing a saturated solution of H_2SiF_6, 2) adding boric acid to supersaturate the solution and 3) placing wafers in the solution for deposition. Our starting solution was 2.5 M H_2SiF_6. This differs from NSG's reported initial 4 M H_2SiF_6 solution. Our 2.5 M concentration is the same used by S.M Thomsen and our calculations show that commercially available 30 wt % H_2SiF_6 is approximately 2.6 M. Figure 3 shows a more detailed diagram of our process. Later in this paper, we provide a detailed recipe for mixing a supersaturated solution. It contains necessary volumes and weights and would be useful for someone who would like to run the LPD process.

After mixing the solution, we followed a standard sequence of tests to investigate as many deposition parameters as possible. Figure 4 illustrates a typical sequence for our deposition runs. The first wafers into the solution were either bare wafers or thermal oxide wafers. This first test checked the solution for saturation, i.e. oxide should neither etch nor deposit though in practice it usually deposited at a reduced rate. Next, we added boric acid to a low concentration, about 2×10^{-2} M, and ran another deposition test. We then added still more boric acid and ran deposition tests at higher boric acid concentrations. At this point we ran wafers for varying amounts of time. Finally, we tested the deposition selectivity by using wafers of different materials.

Experiment Setup



- A stir bar takes the place of the cassette during the mixing.

Figure 1. Our experimental setup.

1. Outer tank
2. Inner tank
3. Water
4. Heater
5. Stirrer
6. Inner tank - front compartment
7. Inner tank - middle compartment
8. Inner tank - rear compartment
9. Deposition sample
10. Circulation pump
11. Filter
12. Boric Acid

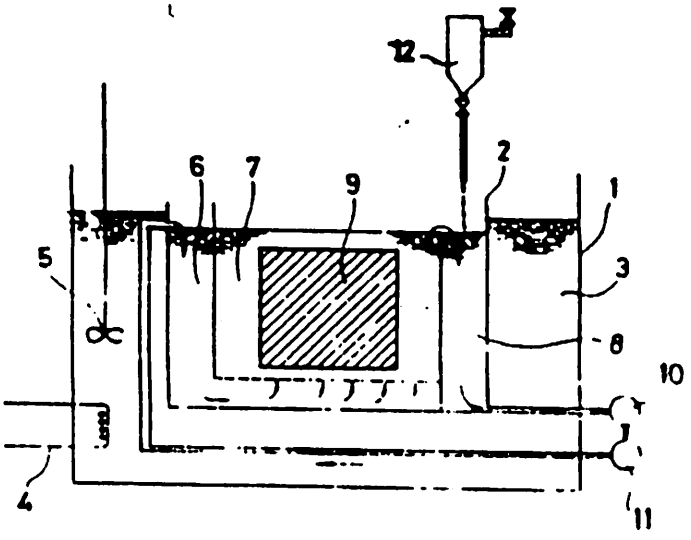


Figure 2. NSG's experimental setup. [3]

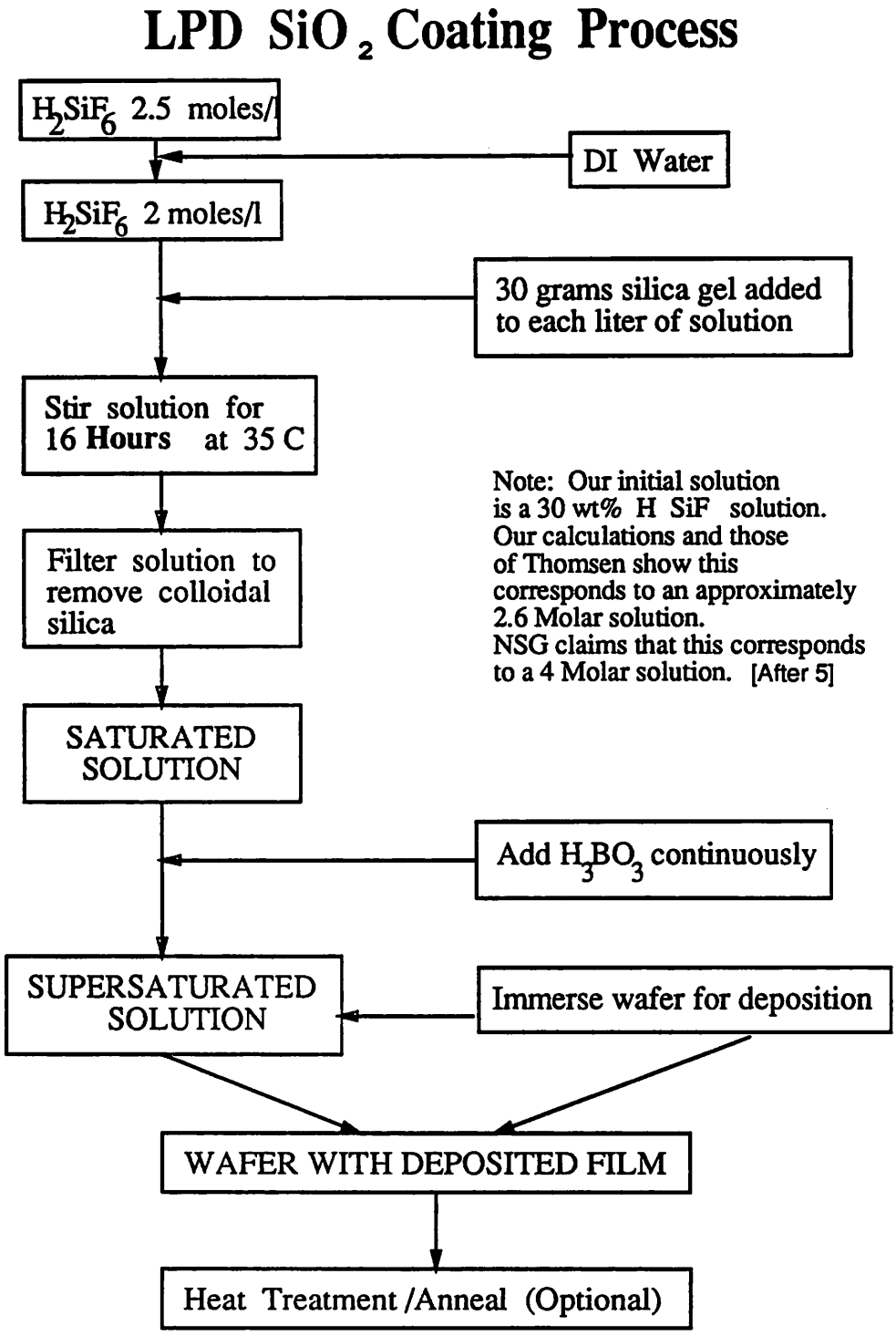


Figure 3. LPD process flow.

Test Sequence

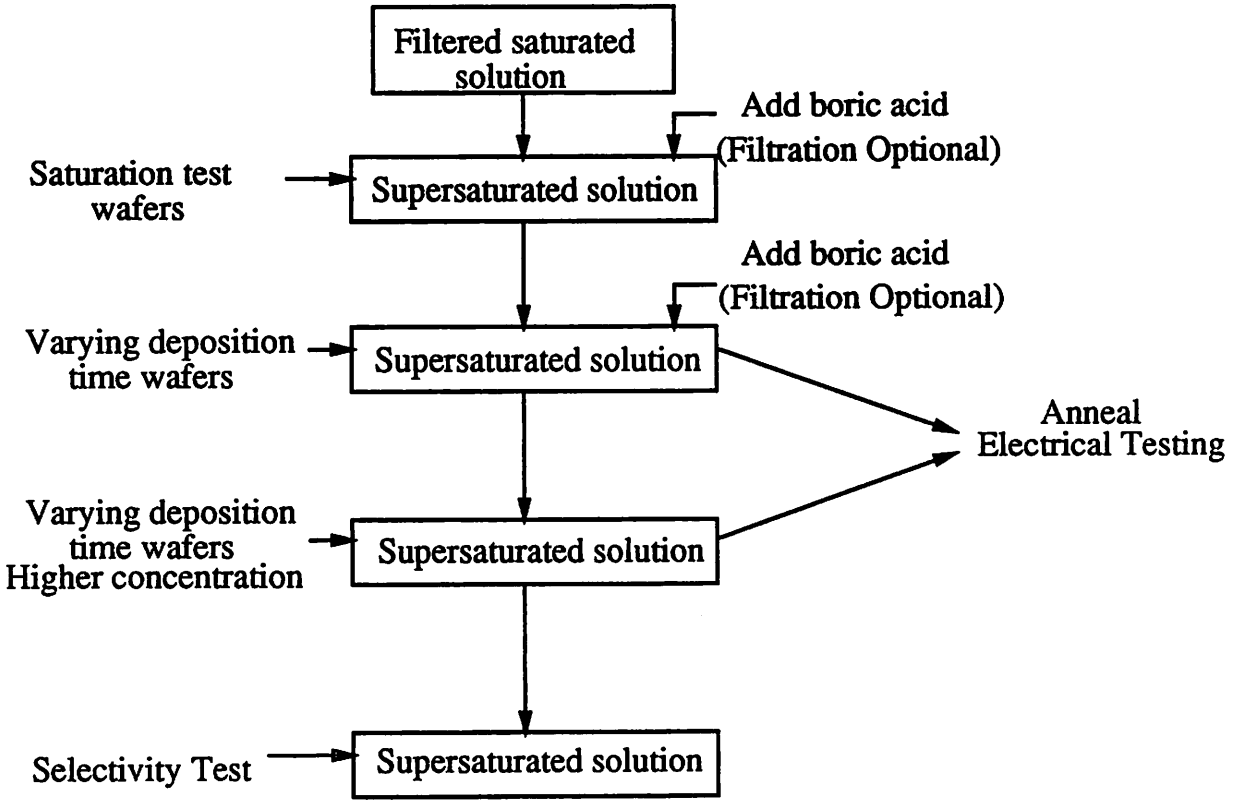


Figure 4

Liquid-phase Deposition Silicon Dioxide Recipe

- Measure volume V1 of 30% by weight H_2SiF_6 solution as delivered by vendor. V1 used should be in the range 750-800 ml, if the combination of stirrer, cassette table and 2" cassette are to be used. The volume can be smaller if the stirrer will not be used at the same time wafers are placed in solution.
- Add V2 = (0.3×V1 + 5 ml) of DI water to the solution. This equation assumes that that the original solution was 2.6 M (H_2SiF_6). An additional 5 ml of DI water is added to allow for the final volume being less than the sum of V1 and V2. The molarity of the final volume should be very close to 2.0 M (H_2SiF_6).

  Comment:

  The assumption of original molarity eliminates weighing the original solution, which is probably more trouble than it's worth, and introduces more variability than it avoids. 2.6 M is the average of the range (2.5-2.7 M) given by Thomsen, and approximately the same as the average found in our second deposition (2.66 M).
- Allow the 2.0 M solution to reach thermal equilibrium by leaving it in the temperature controlled bath for four hours prior to dissolving silica in it. The temperature controlled bath must be at the temperature that will be used for saturation. The stirrer should be used.

  Comment:

  The reason for this is concern that if the silica is added to the solution before it reaches the appropriate temperature, the solution may become saturated at a lower temperature than it will ultimately achieve, and the relaxation from thermally-initiated supersaturation may take many hours thereafter. There is some reason to think that this happened with the second saturated solution we made up (some deposition occurred prior to boric acid being added to the 35°C saturated solution).
- Add 35 grams of silica for every liter of (V1 + V2). 30 grams should be sufficient: the excess is to allow for a 5% error in H_2SiF_6 content, some hydration of the silica (which will exaggerate its weight), and to provide some undissolved silica residue during filtration as a visual verification that saturation was achieved.
- Allow this mixture to stay at the saturation temperature for 16 hours with the stirrer on. This time may be excessive, but we've no reason to take the trouble of finding the minimum time required.
- Filter this solution. Check for undissolved silica filtrate to verify saturation.

  Comment:

  At this moment, we use a 5 μm-pore size paper filters. This is much larger than the 1.2-1.5 μm-pore size used by NSG. It is a good

compromise between cost, filtration rate, and strength. Use of vacuum-assisted filtration reduced the filtration time to less than five minutes. This time reduction is important: the solution temperature undoubtedly drops during this step.

- Place one native oxide covered or recently HF-cleaned wafer and one wafer with a measured thickness of thermal oxide in the saturated solution to check for etching/deposition. Remove these after two hours, and before adding boric acid to the solution.
- H_2SiF_6 To obtain X molarity of $B(OH)_3$, weigh out $X \times V3 \times 61.84$ grams/mole of solid $B(OH)_3$. Usually X is about 2.8×10^{-2} M. The suggested procedure is:
  * use the same white teflon beak for weighing the $B(OH)_3$ as will be used for mixing it with water before adding this mixture to the saturated solution.
  * add 10 mL of DI water for every gram of $B(OH)_3$.
  * heat the beaker's contents on a hot plate at a moderate temperature for 10-15 minutes. Occasionally tap a corner of the beaker against the hotplate surface to aid mixing. DO NOT try to stir the mixture: the boric acid tends to stick to all available surfaces.
  * when the $B(OH)_3$ has completely dissolved take the hot mixture straight from the hotplate and add directly to the saturated solution.

  Comment:

  This is the only way we've found so far to get $B(OH)_3$ into solution and keep it there. Allowing the heated $H_2O/B(OH)_3$ mixture to cool somewhat lets the boric acid plate out on the beaker surfaces and at the liquid/air interface. The difficulty is that we take the saturated solution away from thermal equilibrium for a matter of perhaps an hour.
- Re-filter as necessary.
- Immediately after any filtration, check the deposition rate using a native-oxide-covered or recently HF-cleaned wafer.

Deposition Results

Introduction

After experimenting with mixing the solution and refining our apparatus in the first two runs, we used the recipe on on the previous pages to prepare our solution. We prepared and ran samples through a total of seven batches of solution. In each solution, we performed a series of experiments including a pre-boric acid saturation test, varying the boric acid concentration and temperature, and testing the selectivity on various materials.

Saturated Silica Solution

Immediately after preparing the saturated silica solution, but before adding boric acid, we placed several bare and thermal oxide wafers in the solution to check for deposition and etching of oxide. The results are summarized below:

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	A1	23	691	668	19.83	33.69	36	0.00E+00
4	A2	23	654	631	19.83	31.82	36	0.00E+00
5	A3	24	74	50	1.83	27.32	36	0.00E+00
6	B1	29	208	179	2.37	75.52	45	0.00E+00
7	B2	25	418	393	1.67	235.33	45	0.00E+00
8	C1	24	130	106	2.25	47.11	36	0.00E+00
9	ES1	11	56	45	2.00	22.50	27	0.00E+00
10	FS1	24	93	69	2.42	28.51	28	0.00E+00

Clearly, we had some film deposition. The rate, however is very erratic with no clear correlation with temperature. Visually, these films had color ranging from brown-black to shiny metallic. Ellipsometric measurements of the thicker deposited films indicate a refractive index of 1.45 to 1.46, identical to that of SiO_2. We assumed a refractive index of 1.46 in our measurements of thinner films. Brown-black stains are evident in many of the samples left in solution for many hours.

Boric acid concentration

After the addition of boric acid to the saturated silica solution, we successfully deposited films of various thicknesses. We used the ellipsometer to characterize the thickness and refractive index of our deposited film. The refractive index readings we consistently got were 1.45 to 1.46. This similarity to SiO_2's refractive index indicates the primary component of the thicker deposited films is probably SiO_2.

In one run, we added boric acid to a concentration of 2.0×10^{-2} M, ran some wafers and then added more boric acid to a final concentration of 2.8×10^{-2} M. The deposition results were as follows:

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	B3	22	188	166	2.47	67.21	45	2.00E-02
4	B4	22	205	183	2.47	74.09	45	2.00E-02
5	B5	21	195	174	2.47	70.45	45	2.00E-02
6	B10	23	250	227	2.47	81.95	45	3.00E-02
7	B11	23	49	26	0.92	28.26	45	3.00E-02

Compare with the results from another set (A) of wafers in 2.8×10^{-2} M solution and a third set (C) in 2.86×10^{-2} M solution:

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	A4	24	79	55	3.35	16.42	36	2.80E-02
4	A6	13	294	281	3.32	84.64	36	2.80E-02
5	A10	27	131	104	3.32	31.33	36	2.80E-02
6	A11	24	291	267	3.32	80.42	36	2.80E-02

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	C2	24	472	448	6.58	68.09	36	2.86E-02
4	C6	24	281	257	3.60	71.39	36	2.86E-02
5	C7	24	269	245	3.60	68.06	36	2.86E-02
6	C8	24	292	268	3.60	74.44	36	2.86E-02

In our last two runs, we varied the boric acid concentration between 2.3×10^{-2} M and 3.5×10^{-2} M.

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	E-S4	11	131	120	2.05	58.53	39	2.27E-02
4	E-S5	11	121	110	2.17	71.39	39	3.50E-02
5	F-S2	25	48	23	2.08	11.05	28	2.36E-02
6	F-S3	25	29	4	1.73	2.31	28	3.20E-02

Our data is inconclusive about the relationship between deposition rate and boric acid concentration. NSG found the deposition rate to be linearly proportional to boric acid concentration for concentrations between 2.0×10^{-2} M and 3.0×10^{-2} M.

Temperature

Our deposition runs were done at several different temperatures.

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	B10	23	250	227	2.47	81.95	45	3.00E-02
4	E-S4	11	131	120	2.05	58.53	39	2.27E-02
5	E-S5	11	121	110	2.17	71.39	39	3.50E-02
6	F-S2	25	48	23	2.08	11.05	28	2.36E-02
7	F-S3	25	29	4	1.73	2.31	28	3.20E-02

Here, there is some positive correlation between temperature and deposition rate. NSG found that deposition rate is linearly proportional to the deposition temperature for temperatures between 30°C and 40°C.

Thick Films

The thickest films we were able to deposit at the consistent 60-70 Angstrom/hour rate were about 600 Angstroms.

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	O3	13	419	406	8.38	48.45	35	2.80E-02
4	O4	12	421	409	13.20	30.98	35	2.80E-02
5	C3	24	606	582	22.27	26.13	36	2.86E-02
6	C4	24	620	596	22.27	26.76	36	2.86E-02
7	C5	24	605	581	22.27	26.09	36	2.86E-02
8	A13	14	483	469	11.97	39.18	36	2.80E-02
9	A14	14	450	436	11.97	36.42	36	2.80E-02
10	A15	14	469	455	11.97	38.01	36	2.80E-02
11	A16	14	468	454	11.97	37.93	36	2.80E-02
12	A17	14	478	464	11.97	38.76	36	2.80E-02

Our process self-limited at this thickness. This is in sharp disagreement with the NSG researchers who achieved film thicknesses up to 2 microns.

Selectivity

We attempted deposition on different material substrates.

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	O-Nitride1	729	708	-21	1.07	-19.63	36	2.80E-02
4	A-Nitride1	744	703	-41	2.90	-14.14	35	2.80E-02
5	E-N1	743	758	-15	2.00	-7.50	27	0.00E+00
6	A-PSG1	5296	5337	41	2.90	14.14	35	2.80E-02
7	A-Upoly3	-	-	321	2.90	110.69	35	2.80E-02
8	A-Dpoly2	-	-	400	2.90	137.93	35	2.80E-02
9	O-Wetox1	1046	1049	3	2.13	1.41	35	2.80E-02
10	A-Wetox12	1044	1044	0	0.00	0.00	36	2.80E-02
11	C-Wetox1	1049	1089	40	6.58	6.08	35	2.80E-02
12	E-X1	919	914	5	2.00	2.50	27	0.00E+00

From the above table, the material deposits well only on poly. It etches nitride and has little effect on thermal oxide. In contrast, NSG was able to deposit the LPD oxide on top of a thermally grown oxide. The material seems to deposit slowly on PSG.

A four day deposition test on aluminum and photoresist covered wafers revealed that neither material survived the solution.

The predeposition cleaning treatment also has a significant impact on the deposition rate.

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3	A6		Standard sink6 clean, 30 second 10:1 HF dip, no drying					
4	A6	13	294	281	3.32	84.64	36	2.80E-02
5								
6	A8		Standard sink6 clean, 30 second 5:1 BHF dip, no drying					
7	A8	27	132	106	3.15	33.65	36	2.80E-02
8								
9	A10		Standard sink6 clean, No HF dip, blown dry					
10	A10	24	131	104	3.32	31.33	36	2.86E-02

A8 and A10 show a markedly lower deposition rate. Perhaps the piranha and HF dips left the surface with hydrides or hydroxyls which affected the deposition rate. However, A6, which also went through the HF dip does not have a lower deposition rate.

Temperature induced deposition

The supersaturation of the silica solution increases with temperature. Consequently, the deposition of oxide should be enhanced. We attempted to deposit oxide by raising the temperature of the solution instead of using boric acid.

	A	B	C	D	E	F	G	H
1	Wafer	Thick-Bef	Thick-Aft	Thick-Net	Time (Hr)	Rate/Hr	Temp C	Boric (M)
2								
3				Oxide				
4	E-X1	919	919	0	2.00	0.00	27	0.00E+00
5	E-X1	919	919	0	1.95	0.00	38 -> 42.5	0.00E+00
6								
7				Nitride				
8	E-N1	758	743	-15	2.00	-7.50	27	0.00E+00
9	E-N1	743	688	-55	1.95	-28.20	38 -> 42.5	0.00E+00
10								
11				Silicon				
12	E-S1	11	56	45	2.00	22.50	27	0.00E+00
13	A2	23	654	631	19.83	31.82	36	0.00E+00
14	C1	24	130	106	2.25	47.11	36	0.00E+00
15	E-S2	11	327	316	1.95	162.05	38 -> 42.5	0.00E+00

It is interesting to note that the bare silicon wafer deposited at a very rapid 162 angstroms/hour during the ramped temperature test.

Electrical Characterization

The electrical characterization was done on LPD oxide wafers to examine the electrical properities and the possibility of using it in VLSI fabrication. MOS capacitor samples were fabricated and standard I-V and C-V testing were then performed on the samples.

1. Test Device Structure and Fabrication

The test structures used are primarily MOS capacitors with aluminum gates, as shown in Figure 5. Because 2" wafers are used in the deposition process, the gate mask for EE143 laboratory is used for gate definition mask. The available size of aluminum pads ranges from 2200 to 3000μm^2.

The process flow is listed below.

1. Standard clean of P-type <100> wafers (sink6).
2. Deposit 400-800Å, LPD oxide.
3. Piranha clean wafers with LPD oxide without HF dip (sink6).
4. Standard nitrogen annealing, 900°C, 20 minutes (tylan14).
5. Deposit 5000Å aluminum (CPA).
6. Gate-defining lithography (2" wafers, G-line, using EE143 Al mask).
7. Wet etch aluminum and photoresist stripping.
8. Deposit photoresist on front side and hard bake.
9. Back side etch in BHF (sink8) then strip photoresist.
10. Standard forming gas annealing, 400°C, 20 minutes (tylan13).

The process has been run twice with different oxide thickness (400-800Å), with/without nitrogen or forming gas annealing. Control wafers with thermal oxide of the same thickness were made to monitor the process. Electrical testing was done on both LPD and thermal oxide wafers to see if there was a difference.

2. Annealing

Annealing was used to improve the quality of oxide after deposition. 900°C, 20 minutes nitrogen annealing was used. Before the annealing step, the wafers were cleaned in piranha without HF dip (process step 3,4).

Decrease of the oxide film thickness was observed after piranha clean and annealing. By using optical measurement (ellipsometer), the decrease in film thickness after piranha clean is less than 10% and about 40% after annealing. This observation is consistent with the fact that a hydride film existed.

3. I-V Testing

Ramped voltage testing is used to check the dielectric strength and leakage of the LPD oxide. Figure 6 shows a representative I-V characteristic of the same as-deposited/grown thickness while the thermal oxide exhibits very low leakage current at low field and breakdown at about 50 volt, the LPD oxide is very leaky. At a 1 volt biasing, the leakage could be as large as 100μA. Wafers with annealing and without annealing don't show much difference in the I-V characteristics and wafers from different run show the same result.

The conduction mechanism of the leakage is further investigated. From Figure 7, we can find the polarity dependence of the current is very large. For negative gate voltage, which bias the substrate into the deep depletion region, the leakage is smaller than the other polarity (accumulation region) by at least 5 orders of magnitude. If the light is on, the leakage is about 10 times larger. From this observation, we conclude the leakage current is found to be limited by the generation of carriers.

Figure 7 shows the leakage current for capacitors with different aluminum sizes in linear scale with positive gate voltage voltage. From this figure, we see that when the bias voltage is larger than 5 volt, the leakage current is proportional to the square of the gate voltage. The slope of current to voltage square in Figure 7 is independent of the capacitor area. From this observation, we can attribute the leakage to the space-charge-limited conduction mechanism.

The magnitude of current for different capacitors in thermal oxide shown in Figure 5 is proportional to the area of aluminum pad. But in Figure 7, we don't find this for LPD oxide samples. Therefore, there exist some weak spots on the oxide and the leakage current flows mainly through these weak spots. This explains why the leakage current is not proportional to the capacitor area.

4. C-V Testing

C-V measurement is widely used in MOS systems to investigate the properities of interface states, mobile ion densities, and trapped carrier concentrations. But in the LPD oxide, because of the large leakage current, the carrier trapping is rather serious and flat band voltage shifting is easy to obtain after a short time of dc current flowing. Therefore, most C-V techniques can not be applied to this LPD oxide.

Figure 8 shows the high-frequency C-V curves of the test capacitors with different test signal frequencies. The capacitance in accumulation region is very close to the theoretical calculation using the thickness by optical measurement.

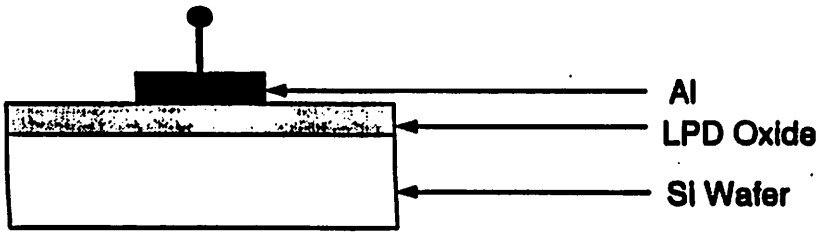


Figure 5. The MOS test structure.

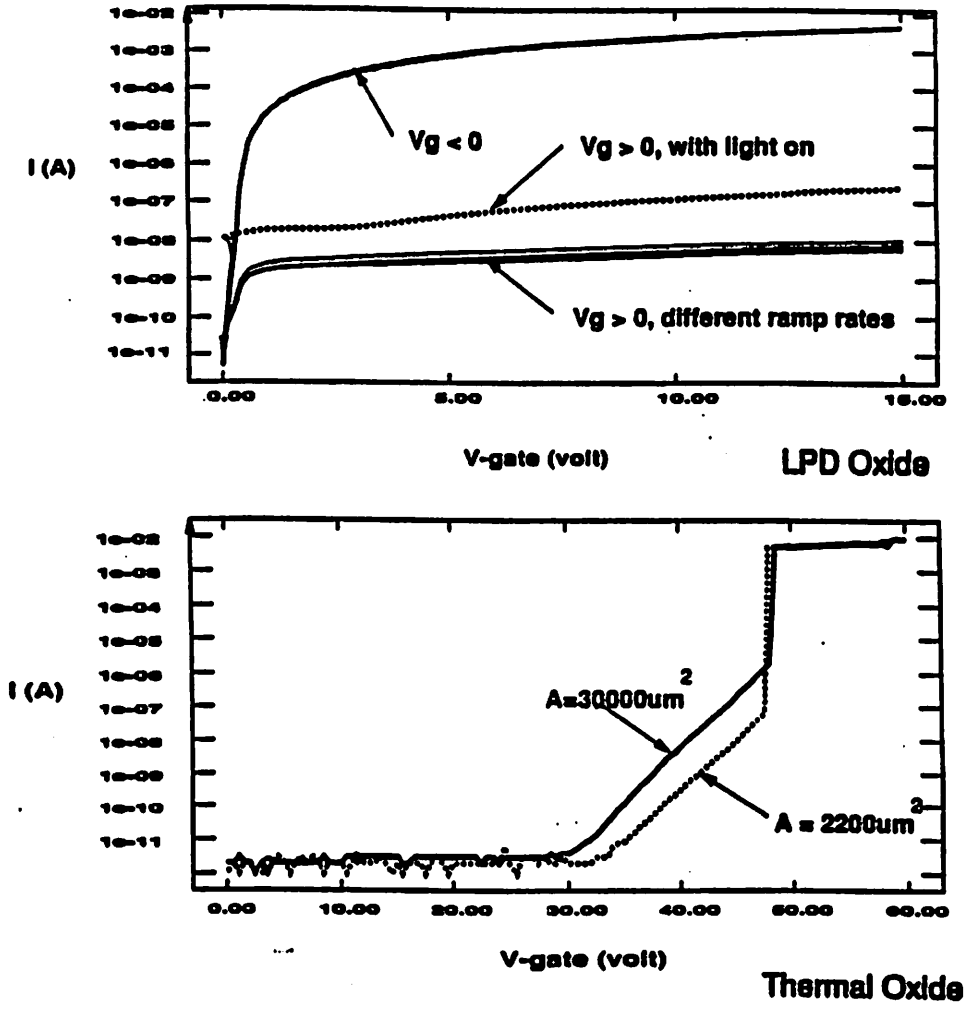


Figure 6. I-V curves of test structure with LPD and thermal oxide.

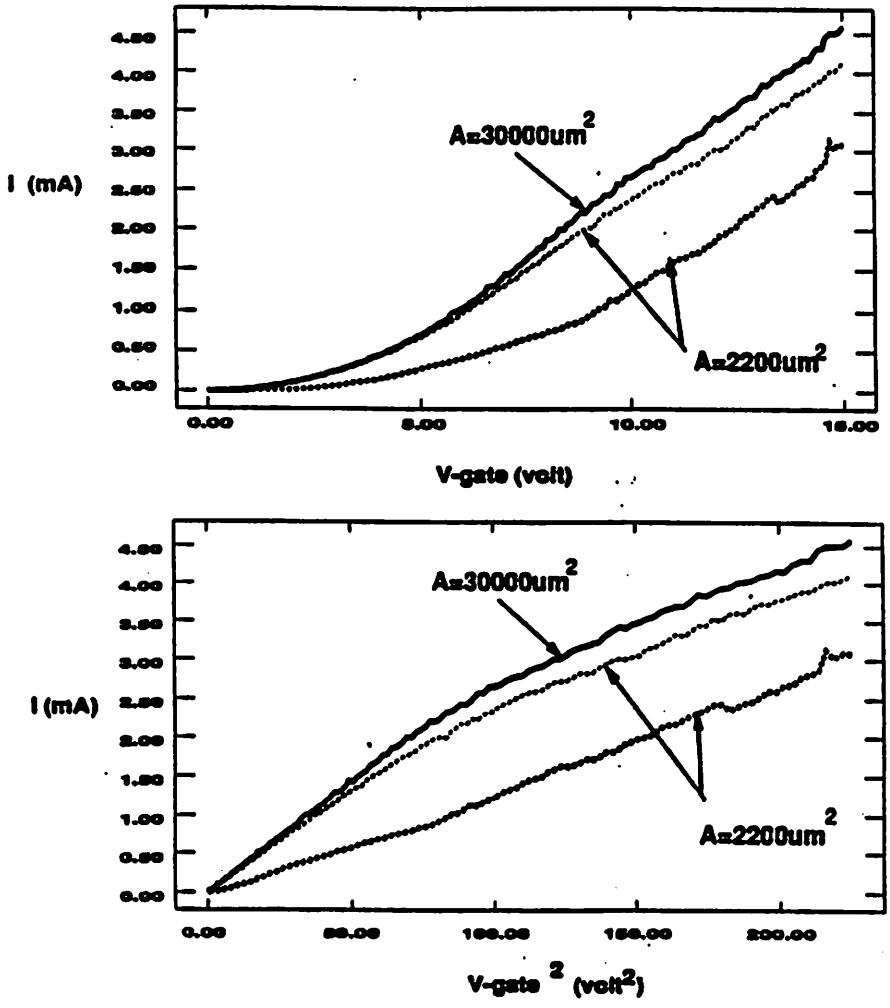


Figure 7. I-V curves of test structure with LPD oxide.

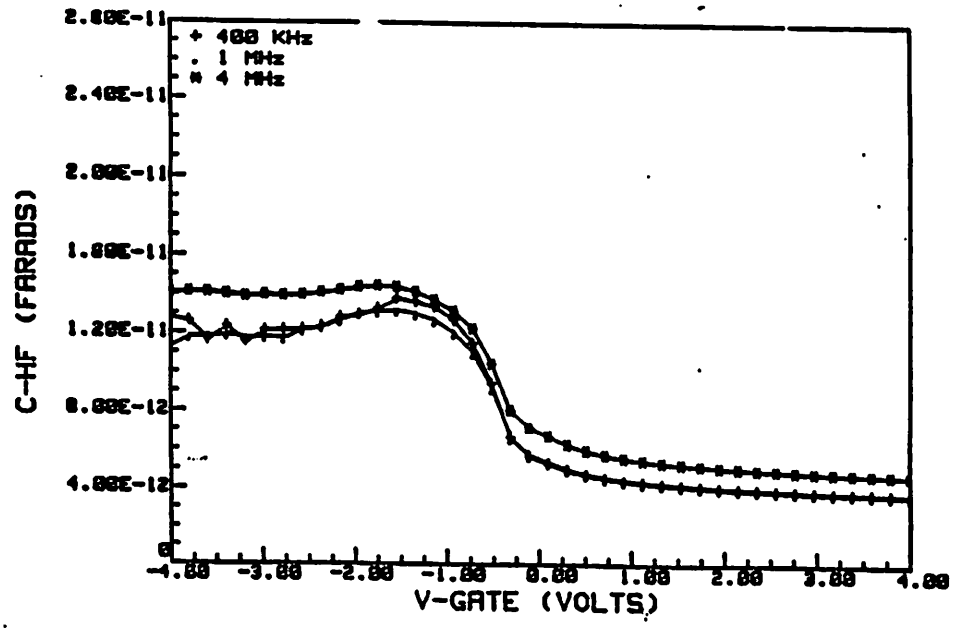


Figure 8. C-V curves of test structure with LPD oxide.

Liquid-Phase Deposited Oxide Characteristics

Our initial objective was to implement the LPD process reported by Nippon Sheet Glass researchers. The characteristics of this process and the deposited films were [3-6]:

- a rather concentrated (2 M H_2SiF_6 has the same HF content as 12 M HF, which corresponds to a roughly 3:1 HF solution) flourosilicic acid solution is saturated with silica and then perturbed by raising the temperature or adding boric acid.
- the saturated solution undergoes recirculating filtration and is kept in laminar flow as it goes past the silicon wafer surface.
- the deposition rate under typical operating conditions (2 M flourosilicic acid saturated at 35°C, with about 2×10^{-2} M boric acid) was 100-200 Angstroms/*hour*: this is a very slow process. Sustained deposition occurred as long as the boric acid used to initiate the deposition was replenished as necessary. Films of up to 0.8 μm thickness were reported.
- without resort to an anodic process, deposition would occur on doped polysilicon, bare silicon surfaces, and particularly silicon oxide or glass surfaces. Deposition would not occur on resist surfaces.
- the as-deposited film contains an undetermined amount of flourine as indicated by an FT-IR absorption peak at about 930 cm^{-1}. There is no very noticable absorption around 2100 cm^{-1}, where it should occur if there are Si-H groups in the film. With the exception of the flourine peak, the spectrum of an unannealed film looks very much like that of thermally grown silicon dioxide, and when annealed, is virtually identical to that of thermal oxide.
- the film's refractive index is 1.43-1.44 as deposited, approaching equivalence with that of thermal oxide after moderate annealing.
- as deposited, the film etches more rapidly than thermal oxide in p-etch and at about the same rate in BHF. P-etch rates approach those of thermal behavior of thermal oxide with annealing.
- it wasn't a result reported in the LPD-related literature, but one of our group anticipated that the LPD oxide might be hydrophobic, because of the incorporated flourine. Intentional flourine incorporation has this effect in commercial glass-making [7]. When this was apparently confirmed, it was taken as additional proof that a successful LPD oxide process had been obtained.

Stain Films on Silicon

"Stain" films can form on silicon surfaces placed in concentrated HF if these surfaces are the biased cathode of an electrochemical cell, or if some chemical species is present in the HF which can undergo reduction at the silicon surfaces. The surface

silicon changes its valence state and reacts with whatever is available and energetically favorable. In the case where no bias is applied to the silicon surface, a small background concentration of nitric acid added to the HF provides the species that is reduced. The film that is grown has the following properties [8,9]:

- the composition and formation rate are dependent on the surface doping. On all n-type and lightly doped p-type ($< 8\times10^{16}$ cm^{-3}) silicon, the film composition is between $HSiO_{1.5}$ and H_2SiO.
- the hydrogen exists as a hydride in the film, attached directly to the silicon atoms. This configuration is well known to be hydrophobic; it is responsible for the hydrophobic behavior of silicon surfaces etched in HF [10].
- the film growth is rapid but self-limiting. The self-limiting thickness depends on the nitric acid concentration and the silicon doping, varying from 1000 to 3000 Angstroms after a matter of minutes. These relatively thick films near the growth limit can have a rather rough texture.
- the stain film's refractive index is about 2.7.
- the FT-IR spectra is characterized by a relatively prominent Si-H absorption peak around 2100 cm^{-1} .
- depending on the specie which is reduced at the silicon surface, the formation of the stain film may be stopped with vigorous agitation or purifying the HF prior to adding nitric acid. These particular comments come from reference [8], and any such behavior is reported along with exceptions: stain film formation appears to be a touchy, not very repeatable process.
- since the staining is a growth or conversion rather than a deposition and requires the presence of elemental silicon, obviously no deposition occurs on non-silicon surfaces.

Materials Analysis

Materials results are limited to four areas, none of the four as thoroughly as anyone would like. These are qualitative evaluation of wetting behavior, refractive index, BHF etch rate behavior, and FT-IR spectra.

Until the seventh and last distinct saturated solution, we had "deposited" (if that's what happened) films on about 50 wafers which initially were either bare silicon or covered by a native oxide as delivered by the vendor. The table on the next page shows the qualitative wetting behavior of the wafers when divided into two thickness ranges, as observed during a DI water rinse right after the LPD deposition. A description of the grading criteria used is given. The apparent dependence on thickness may be due to a change in morphology (rougher surfaces dewet more slowly) or because the film composition is changing as the thickness increases. The wetting behavior of unannealed wafers also changes as a function of time exposed to room air after deposition. After 2-4 weeks in room air, no dewetting will take place: the entire wafer

Wetting Behavior of LPD Films on Silicon Substrates

	A	B	C	D	E
1	**Number of Wafers With**				
2	**Listed Behavior**				
3	*Thickness Range (Angstroms)*	*Dewets*	*Slowly Dewets*	*Doesn't Dewet*	*No Data*
4	0 to 400	20	0	1	10
5	400 to 800	5	8	2	1

Dewets - wafers belonging to this group dewet with little or no entrainment of DI water during rinse. Water beads up at a large contact angle on these wafers. Their behavior is nearly indistinquishable from that of freshly-etched silicon surfaces.

Slowly Dewets - water gradually spills off these wafers over a matter of seconds. Surfaces do not stay wetted. Behavior is similar to a silicon surface that has had a thin native oxide form on it.

Doesn't Dewet - water/solid contact angle is nearly zero. Most or all of the surface remains wetted. Hydroxylated, or silicon dioxide surfaces behave in this way.

No Data - data wasn't taken, for whatever reason.

surface will remain wetted. Annealing wafers for 30 minutes at 900°C will also completely eliminate the dewetting behavior.

The refractive index data of the unannealed films has been relatively consistent as measured on several wafers taken from different deposition solutions over the course of the semester. A summary is given in the table below. The data are obtained using an ellipsometer, and are necessarily limited to the thicker films so that convergence can be obtained. These data indicate very clearly that the thicker films on these wafers are very similar to silicon dioxide, and don't particularly resemble the hydride stain films. The only significant discrepancy occurs with the C-run; the lower refractive index may be due to porosity or a less dense oxide structure [11]. The second group of electrical test wafers came from this run.

	A	B	C	D	E	F
1						
2						
3			**Index of Refraction Statistics**			
4	*Wafer*	*Psi Range*	*Mean*	*Std. Dev.*	*Maximum*	*Minimum*
5	O14	23.64 to 24.04	1.454	0.006	1.46	1.446
6	O12	26.38 to 29.81	1.45	0.008	1.461	1.44
7	A1	26.30 to 31.43	1446	0.007	1.455	1.436
8	A2	25.10 to 29.37	1.457	0.014	1.477	1.444
9	B9	27.64 to 30.79	1.469	0.003	1.473	1.464
10	C3	23.96 to 25.58	1.39	0.003	1.394	1.385
11	C4	24,78 to 25.90	1.394	0.006	1.403	1.385
12	C5	24,62 to 25.06	1.392	0.007	1.401	1.384
13						
14		5 measurements for each wafer.				

The etch rate data is particularly scanty. We have data for two unannealed LPD films that differ in their qualitative wetting behavior, and a thermal oxide reference wafer, obtained with fresh 10:1 BHF shown in the etch rate table below. The wafers were etched within minutes of each other. The two LPD films had been deposited two days previously, in a solution that was giving inconsistently dewetting films. These are rather surprising data: it appears that these "LPD" films differ significantly from pure silicon dioxide. It's not clear whether the difference between the two LPD wafers is due to a true interwafer variability, or thickness-dependent composition or density variation. Owing to the variablity of the the deposition rates and dewetting behavior with this solution, no further etch rate tests were run.

	A	B	C	D
1	**Wet Etch Data, 10:1 BHF**			
2			***Rate Statistics***	(A/min)
3	*Comments*	*Nominal Initial Thickness*	*Mean*	*Std. Dev.*
4	Thermal Oxide	1039	482	20
5	Hydrophobic LPD Oxide, Unannealed	209	90	13
6	Marginally Hydrophobic LPD Oxide, Unannealed	413	330	73

The FT-IR spectra of Figures 9 and 10 were taken in transmission mode (through the film and its substrate) with no effort being made to delete the effect of the substrate by stripping the film and obtaining the substrate spectra by itself. This means that some of the effects noted could be due to the substrate. In particular, the Si-O absorption peaks could be due to interstitial oxygen dissolved in the substrate. Three spectra are shown with peaks identified for comparing annealed and unannealed LPD films, and a recent deposition to one that has been in room air for a month. The substrate shouldn't prevent the Si-H peaks around 2100 cm^{-1} being seen if they're there. Figure 10 shows representative wide-band and 1500-3000 cm^{-1} spectra from the unannealed, recent LPD sample. One would expect that even with the thin film, if this were a stain film the Si-H peak should be visible.

We did attempt during finals week to follow up on the suggestion that we use the in-room FT-IR system, with it's subtractive mode and convenience (we wouldn't have had to impose on someone else to take spectra for us, and could have had immediate access at all hours) to obtain more satisfactory results. Unfortunately, this in-room system was much less tolerant of the single-sided polished wafers we were using than is the Chemistry Department Nicolet 20X (we had been warned that we would be unable to use any of the single-sided polished wafers with films already on them, and found this to be quite true). We couldn't find a source of the expensive double-sided polished wafers to use in the in-room system on short notice and made no further efforts.

Discussion

This discussion must cover the technical questions answered or raised, and because this is a project report and we failed to obtain the results we had in mind, some analysis of what went wrong within our group.

Several strong indicators point to the presence of silicon hydride stain film: the self-limiting growth, the significant shrinkage when the film undergoes a 900°C anneal, and the reduced etch rate in 10:1 BHF of the unannealed film. The silicon hydride will decompose when annealed, leaving behind the silicon [8]. This would explain the high degree of leakage we found in electrical testing. The failure to detect any Si-H IR absorption may be due to the thin film thickness used. That the refractive index of the thicker films is so similar to SiO_2 argues that SiO_2 must be the primary component in the thicker films. A reasonable conjecture is that growth of the hydride and deposition of a silica film proceed simultaneously. Initially, while the silicon surface is exposed the hydride growth dominates, and the thinner "deposited" films are primarily hydride, which explains why these films are strongly hydrophobic. The hydrophobic nature of the surfaces may initially retard silica deposition, but eventually the deposition does occur, explaining the index data for those films thick enough to measure, and the declining hydrophobicity of the thicker films. This conjecture could be tested by etchrate and FTIR measurements on films of different thicknesses. We were unable to do complete this test.

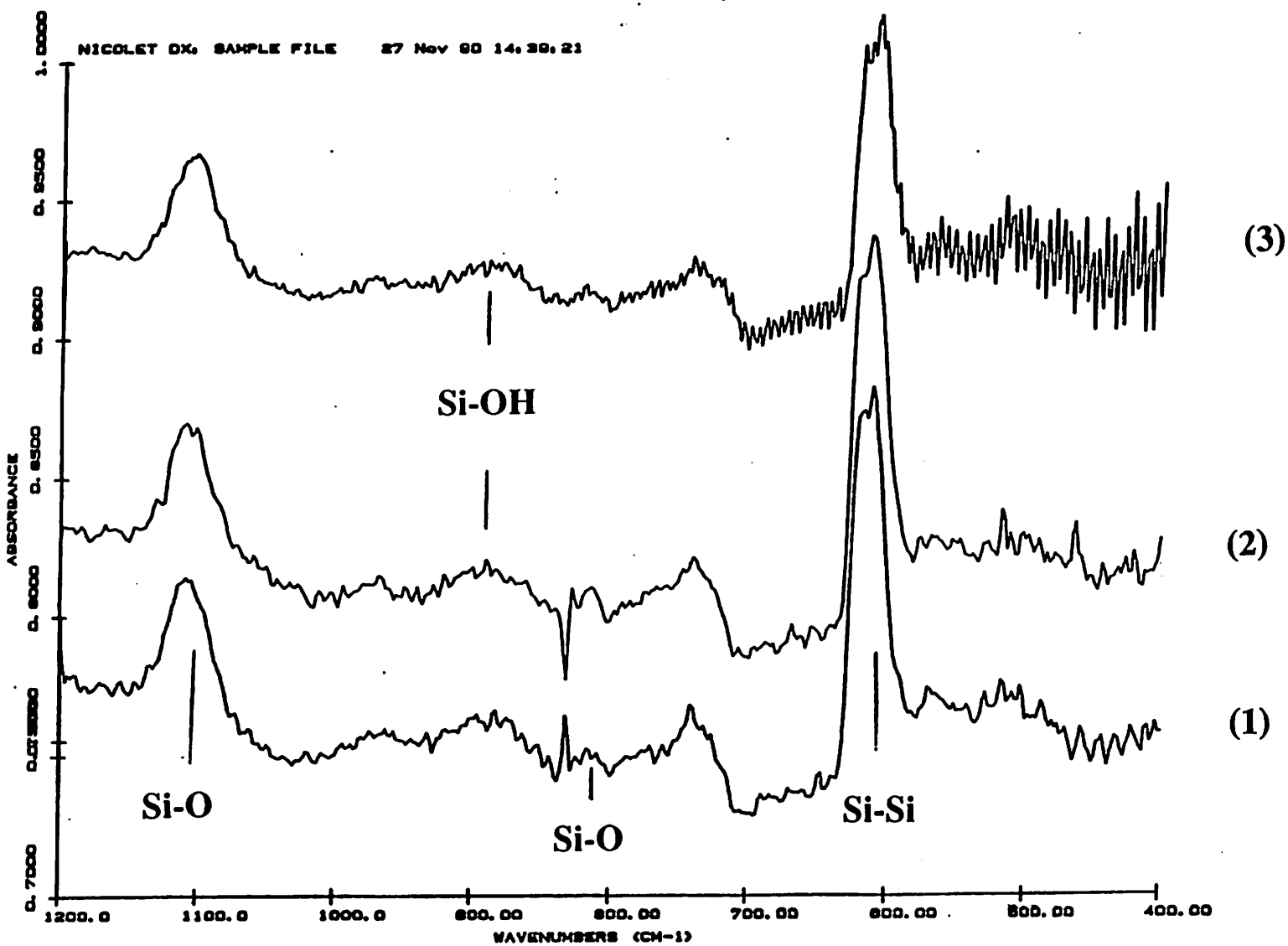
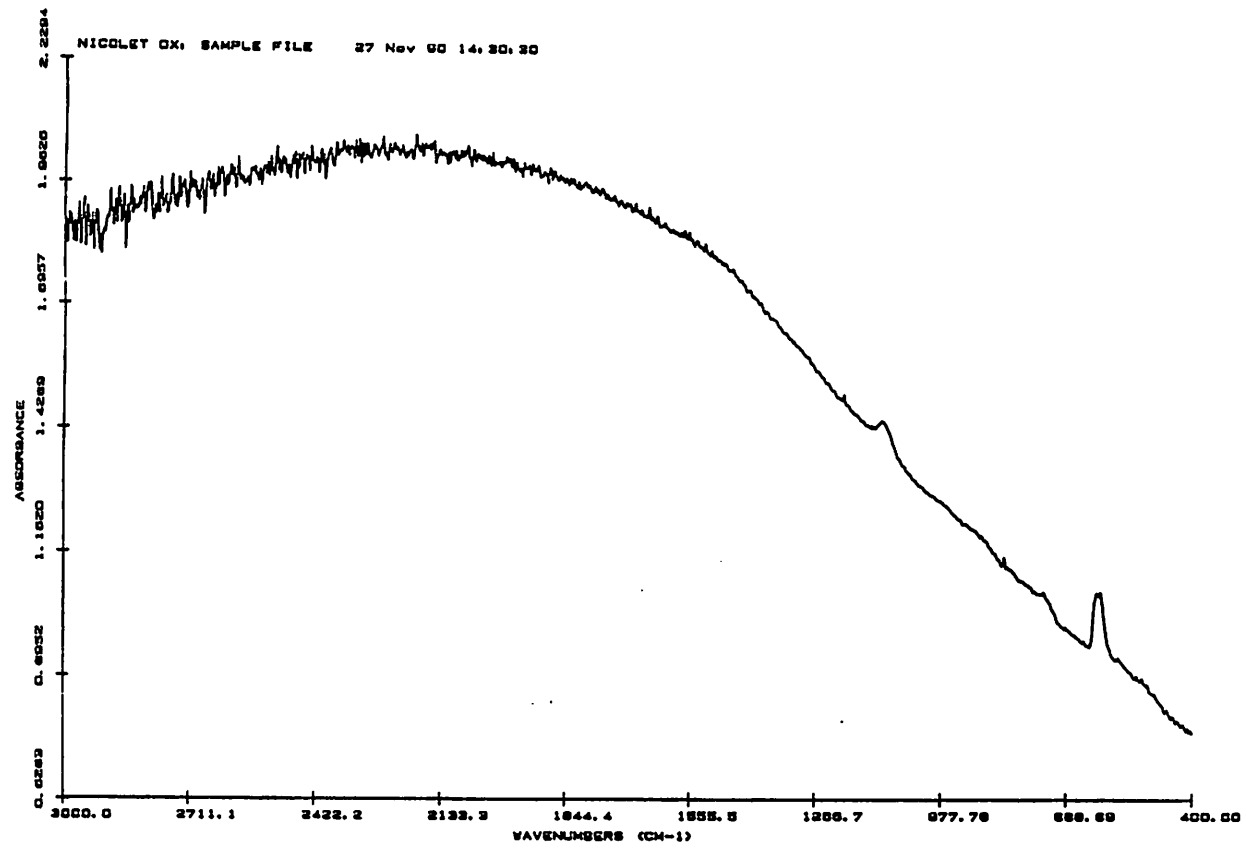


Figure 9: FT-IR Spectra of three LPD glass samples. (1) Sample LU, 135 Angstrom unannealed LPD film exhibiting marked hydrophobic behavior. Spectrum taken two hours after deposition. (2) Sample LA, 92 Angstrom (after anneal) LPD film that underwent a 30 minute 900 degree C N_2 anneal. Hydrophilic surface. Spectrum taken 30 minutes after anneal, which in turn occurred right after deposition. (3) Sample A11 with 291 Angstrom unannealed LPD film, 32 days after deposition (all in room air). Hydrophilic surface. Arbitrary absorption units and leveling.



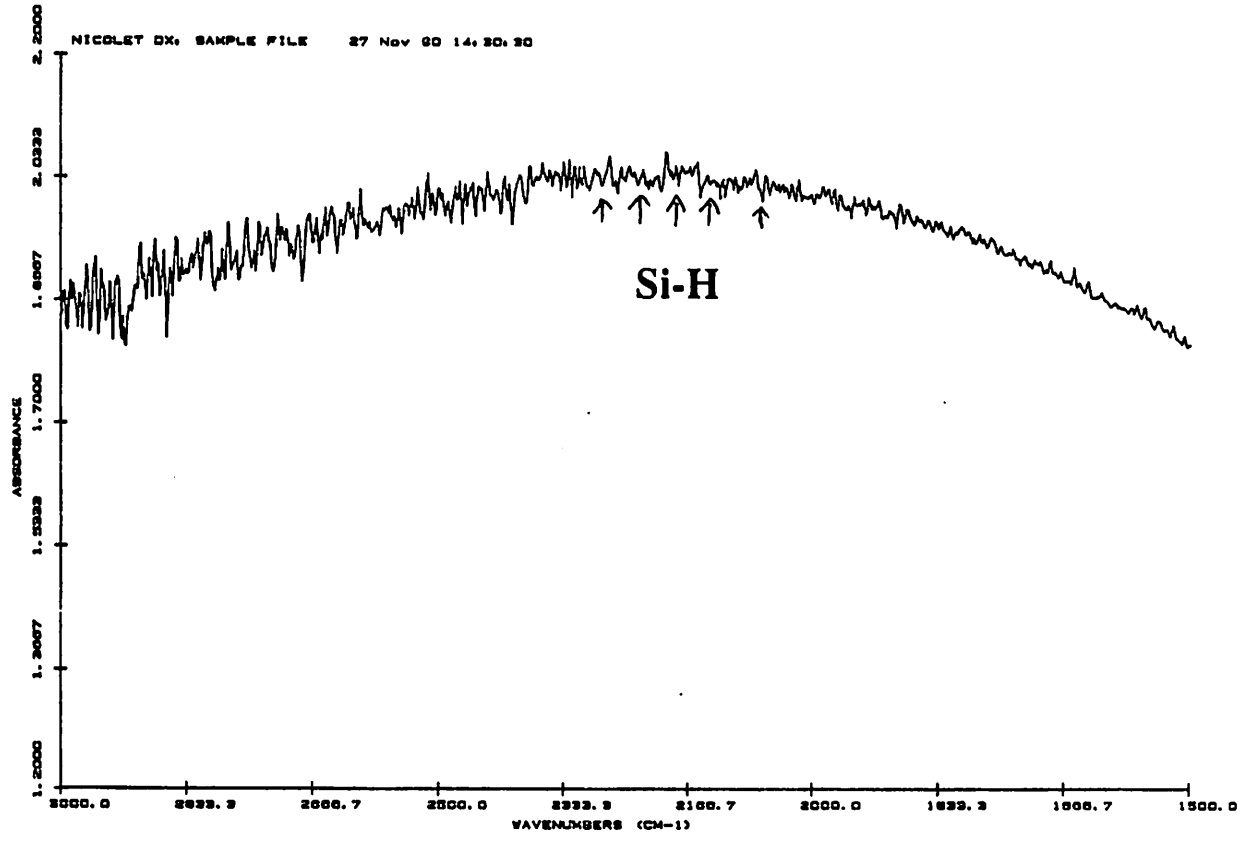


Figure 10: Broadband and detail of Si-H peaks (there are none). These spectra belong to wafer LU described in the previous slide, but are representative of the other two samples: hydride peaks weren't evident on any one of them. Arbitrary absorption units and leveling.

The previous conjecture doesn't explain why growth on silicon is self-limiting, or why little or no deposition occurs on oxide surfaces, though the two things are undoubtedly related; i.e., deposition on a moderately thick, continuous deposited silica layer stops for the same reason as deposition doesn't occur on oxides. This is the sharpest disagreement between our results and those obtained by NSG.

If the hydride is forming on the silicon surfaces placed in the solution, how may this be avoided? As noted earlier, the explanation of how stain films are formed in concentrated HF with small amounts of HNO_3 is a bit vague, with poorly explained exceptions to most of the rules given. We do note that the apparatus described in NSG publications continuously circulates and filters the saturated solution while the stain film descriptions mention that filtering the HF with large surface area silicon prevents stain film formation when the nitric acid is added later. The NSG patent disclosure describes the pains taken to maintain laminar flow of the solution over the surfaces where deposition is to occur. Stain film formation is very sensitive to agitation [8]. It may be that flow and filtering effects play an important part in assuring that the stain film is not formed, in addition to their role in the control of gross particulate and haze problems.

As described in an earlier section, our initial observations were that the material being formed on silicon and native-oxide covered wafers was in- deed the LPD oxide we were trying to obtain. The deposition rate was in the right range, "deposition" was initiated or at least accelerated by adding boric acid, refractive indices of films that were thick enough to measure were quite consistent with LPD oxide, and the deposited film was hydrophobic, which was taken as demonstrating flourine incorporation. We were hampered by being limited to a deposition on one wafer at a time, which prevented us from doing the sort of simultaneous deposition that would have revealed the self-limited film formation early in the semester. It wasn't until depositions were done with the third saturated solution that we understood this was happening. The self-imposed delay in obtaining any materials data probably wasn't significant: with the exception of the etchrate data, none of the other materials data would have had sufficient weight with us when compared to the refractive index data. The latter was pretty persuasive that we had obtained the material we desired at the very beginning. It took some time before we could accept this was the wrong view. There were three systemic reasons for this:

(1) an apparent early success, when any deviations from the NSG published results could still be blamed on a kludgy apparatus (we had deviated considerably from a detailed published specification of the NSG apparatus). Prior to the first deposition, the group felt that the process would require significant effort be made simply getting the deposition to work (when the first wafer that was removed from a saturated solution dewetted, the initial belief was that no deposition had occurred. When the process appeared to yield the desired result at the beginning, caution went down the aspirator and we lost that healthy sense of doubt that all is not right with the world (and LPD oxide).

(2) a lack of communication. We all worked on this, but we worked apart: when one or two of us had something else that demanded our attention, the remaining part of the team would attempt to do everything for a time. That tended to prevent the sharing of doubts and questions. In the same vein, we should certainly have communicated more with Professor Oldham.

(3) no awareness that anything like stain film formation could occur in the LPD solution: for that matter, no awareness that stain films existed.

Conclusions

- we have described our efforts to set up a LPD oxide process and the results of characterization work on the film that was deposited.
- in most respects, we did not obtain the material described by NSG researchers as LPD silicon dioxide. The material we obtained exhibited self-limiting growth/deposition and had poor electrical characteristics.
- we believe our problems were at least in part due to stain film formation. We were unable to identify with confidence the means to avoid this in the future, but have suggested points in which our deposition system differs from that reported by NSG, and which may be critical for the stain films.

Acknowledgments

We owe several people thanks for their help in this project. Rolfe Anderson was very helpful in several ways, with the initial rounding up of equipment, the FT-IR analysis later on, and during occasional conversations through the semester. Phil Guillory saved us much time while putting together our third and final version of the the temperature-controlled bath: he knew where all the surplus but still useful parts were stored, what fittings to use, and put the recirculating bath lines together. Ben Costello lent us the stir plate and allowed us to use a recirculating fluid temperature controller for most of the semester. We have Rosemary Spivey and Susan Kellogg to thank that we didn't have to spend half the class budget on flourosilicic acid alone. Most importantly, we thank Professor Oldham for his guidance, and in particular for suggestng that we could be dealing with a stain film.

Report References

[1] S. M. Thomsen. "High-silica Fluosilicic Acids: Specific Reactions and the Equilibrium with Silica", in *J. Amer. Chem. Soc.*, vol. 74, pp. 1690-1693, (1952)

[2] R. J. H. Lin, J. C. Lee, and P. B. Zimmer. "Low-Cost Solar Antireflection Coatings", Annual Report for ERDA under Contract No. EM-78-C-04-5300, October 1979.

[3] H. Kawahara, H. Nagayama, and H. Honda. "Silicon dioxide coating". U. S. Patent 4,468,420 (August 28, 1984), 12 pages.

[4] T. Goda, H. Nagayama, A. Hishinuma, and H. Kawahara. "Physical and chemical properties of silicon dioxide film deposited by a new process", in MRS Sym. Proc., vol. 105, pp. 283-8 (1988).

[5] H. Nagayama, H. Honda, and H. Kawahara. "A New Process for Silica Coating, in *J. Electrochem. Soc.*, vol. 135, pp. 2013-2016, (1988).

[6] T. Homma, T. Katoh, Y. Yamada, J. Shimizu, and Y. Murao. "A New Interlayer Formation Technology for Completely Planarized Multilevel Interconnection by Using LPD", in 1990 Sym. on VLSI Tech. Digest, pp. 3-4.

[7] D. M. Zirl and S. H. Garofalini. "Reactions on Modified Silica Surfaces", in *Journal of Non-Crystalline Solids*, vol. 122, pp. 111-120, 1990.

[8] R. J. Archer. "Stain Films on Silicon", in *J. Phys. Chem. Solids*, vol. 14, pp. 104-110, (1960).

[9] K. H. Beckmann. "Investigation of the Chemical Properties of Stain Films on Silicon by Means of Infrared Spectroscopy", in *Surface Science*, vol. 3, pp. 314-332, (1965).

[10] M. Grundner and R. Schulz, "The Surface State of Si (100) and (111) Wafers After Treatment with Hydroflouric Acid", in Proceedings, *Deposition and Growth: Limits for Microelectronics*, pp. 329-337.

[11] W. A. Pliskin. "Comparison of properties of dielectric films deposited by various methods", in *J. Vac. Sci. Technol.*, vol. 14, pp. 1064-1081, (1977).

Photodiodes for Stepper Characterization

Diane Hoffstetter
John Hutchinson

EE290N Project
December 17, 1990

Abstract

Techniques for characterizing wafer steppers through wafer integrated photodiodes were developed. A Berkeley Microlab compatible process was developed to produce photodiodes on silicon wafers and to pattern these photodiodes with narrow linewidth grating patterns, suitable for sampling an aerial image. Techniques for measuring photodiode signal while the wafer is on the stepper chuck were developed. The integrated photodiodes were used to determine best focus and simple vibration experiments were performed.

Photodiode Image Monitor

Key Issues

- Producing 0.2um spaces in aluminum films
- Measuring Photocurrent while on stepper
- Manual Alignment of Wafer to PC Board

Contributions

- Microlab compatible process for photodiode fabrication
- Process for patterning narrow spaces in aluminum films
- Techniques for measuring photodiode currents while on the stepper stage

III. Statement of Problem and Key Issues

As feature sizes shrink and approach the optical limits of wafer steppers, the need for accurate and repeatable focus and alignment becomes critical. Normally, best focus is subjectively determined by observing a test wafer with a matrix of focus positions exposed on it. Best focus can be more objectively determined by direct measurement of aerial images and maximizing contrast in a line/space grating.

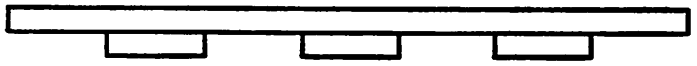
The basic problem to be explored in the project is to build a system which allows direct measurement of aerial images on a wafer stepper. The idea is to build photodiodes integrated on a wafer which have a sampling grating which scans an aerial image as the wafer is stepped across a mask with line/space patterns (Figure 3.1). A line and space mask is imaged onto the wafer plane and produces an aerial image that becomes more sinusoidial as the resolution limit of the stepper is approached. The photodiodes consist of a large area which is patterned with long and thin gratings. These narrow "slits" allow the aerial image to be sampled.

Key Issues:

- Patterning 0.2μm wide resist lines 500μm long and being able to lift off 1000A of aluminum sputtered over the lines.
- Exposing continuous lines in photoresist 0.5mm wide from the photodiodes to the flat of the wafer.
- Manually aligning wafer in the wafer stepper and being able to make contact to the photodiodes while the wafer is on the chuck.
- Getting enough current from the photodiodes. The dark current must be low. Low series resistance of the diodes is need by sintering the aluminum on the wafer but the sinter must be done at a low temperature so that junction spiking does not occur.

Figure 3.1

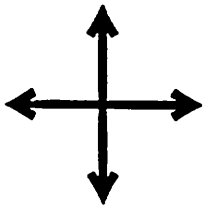
Direct Image Characterization

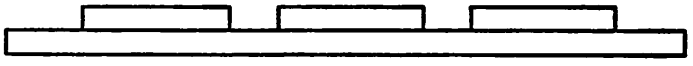


Mask

Imaging Lens

Aerial Image





Photodiodes
on Wafer

IV. State Of The Art: Literature Summary

There have been several articles on stepper characterization published in the literature. In particular, Brunner has produced several systems which sample an aerial image with narrow 'active areas' of fluorescent material or photodiodes in order to reconstruct the entire image. In the early work by Brunner [Brunner85], a test wafer was produced with fluorescent detector structures 0.8 microns wide. The green fluorescent signal was measured as a line/space grating illuminated the wafer. He attained good agreement with isofocal bias point, but measured contrast was much lower than predicted on SAMPLE. The monitor was used to determine best focus, overlay error and vibrational effects.

Later work by Brunner [Brunner86, Brunner88] continued to document use of the fluorescent stepper image monitor. In Brunner86, the same technique of fluorescent lines was used, but with a 0.4um wide structure on a 4um pitch. He measured overlay vector of 4 dies on the wafer to determine alignment vectors for translation, magnification and rotation. Best focus was found by minimizing the wdith of the aerial image, but again SAMPLE simulations showed a narrower image than measured. He documented the effects of barometric pressure on best focus and magnification and found the wafer cassette elevators to be a significant source of vibration. In Brunner88, Brunner states that the fluorescence signal decays with time due to photo-chemical degradation, and utilizes a new technique employing an ultra-flat wafer and poly on oxide lines of 0.4 micron width. The 80nm oxide is low reflectance at 436nm (g-line), so blue scattered light signal from poly is measured. Best focus is determined by maximizing the correlation of the measured image with SAMPLE calculated image. He documents the change in best focus versus lens heating and suggests a simple RC model, and also models bluring effects due to finite width of detector structures. The error in the detector is proportional to the ratio of detector width to the minimum feature size squared.

The most recent work by Brunner [Brunner90] documents a photodiode measurement system for the stepper monitor. This is the system used in GCA steppers called INSITU. Unfortunately, he does not provide any details on the photodiode structures, nor signal levels. He uses 1 micron width slits over a photodetector. He documents similar best focus effects to earlier work, but suggests a two body lens heating model.

The work by Brink [Brink90] uses a method similar to the scattered light setup in Brunner88, where diffracted light is measured by a detector near the image plane. Many detectors are used, but many measurements are possible.

References

[Brunner85], T. A. Brunner and R. R. Allen, "In Situ Measurement of an Image During Lithographic Exposure" IEEE EDL-6, July 1985, p329-331.

[Brunner86], T. A. Brunner and S. M. Stuber, "Characterization and Setup Techniques for a 5X Stepper," SPIE 633-106.

[Brunner88] T. A. Brunner, et. al., "A Stepper Image Monitor for Precise Setup and Characterization," SPIE 922-44.

[Brunner90] T. A. Brunner, et. al., "Stepper Self-Metrology Using Automated Techniques," SPIE 1261-286.

[Brink90], M. v.d. Brink, et. al., "Automatic On-line Wafer Stepper Calibration System," SPIE 1261-298.

V. Approach

Goal :

Develop a Berkeley Microlab compatible process to produce photodiodes on silicon wafers and to pattern these photodiodes with narrow linewidth grating patterns, suitable for sampling an aerial image.

Approach :

The processing of the photodiodes mainly follows standard lab processes. The trickiest part of the photodiode fabricaton was to make 0.2μm spaces in aluminum at a 2μm pitch. Since it is easier to make thin resist lines than it is to make thin spaces in resist, a lift-off method for producing the thin spaces was chosen over a subtractive process.

Methods for creating the 0.2μm width lines were explored. Overexposure and overdevelopment of the resist were considered but were found to be unrepeatable. A second method consists of ashing the resist in an oxygen plasma in technics-c at a power of 50watts. This was the method that was futher explored and later used in the final process.

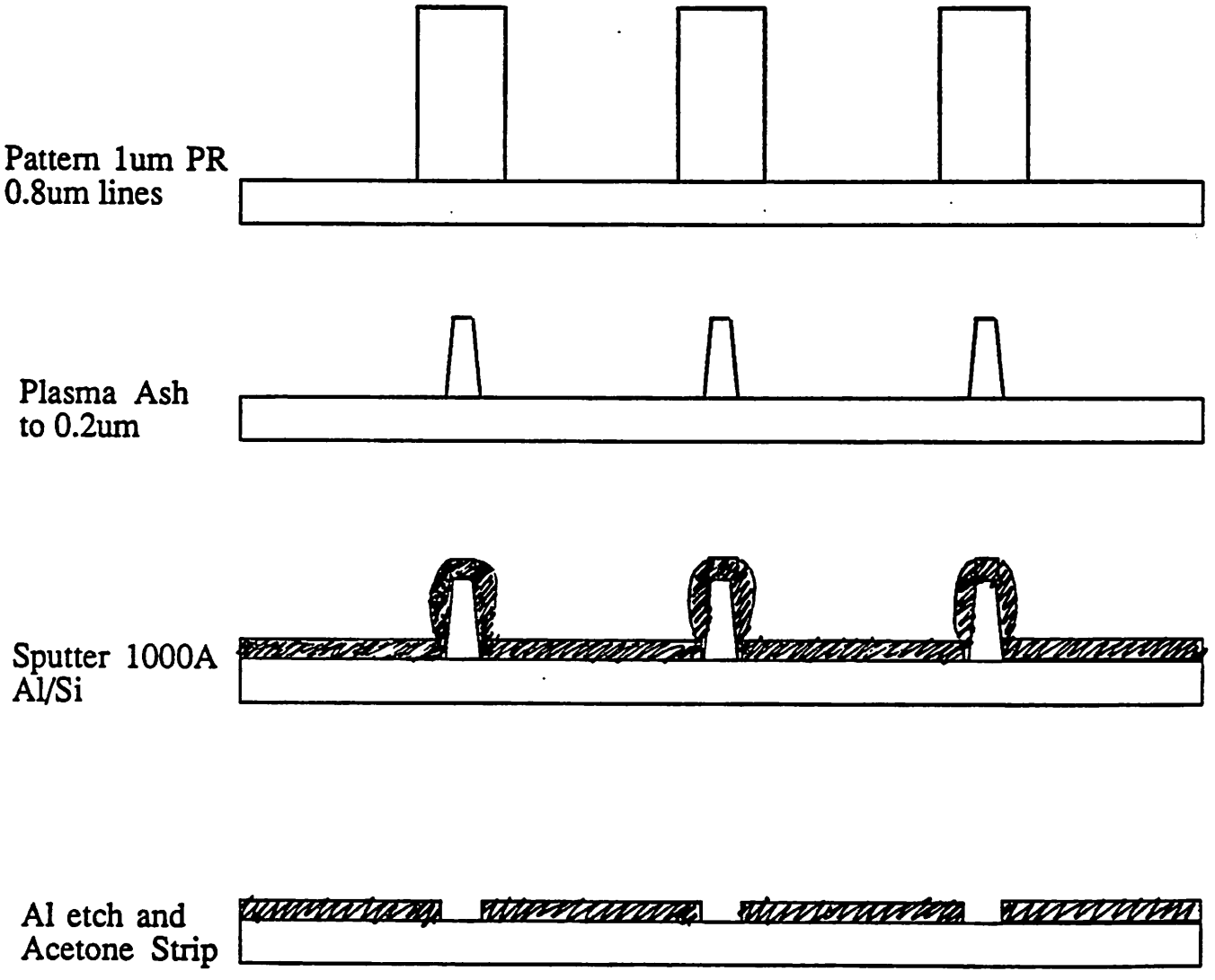
A test mask consisting of lines greater that 0.5mm long and at varying widths (0.1-1.5μm wide) and at various pitches (1-3μm pitch) was used. The 0.8μm lines printed repeatably. The 0.6μm lines also printed fairly well but resist thickness loss was noticed.

The wafer containing the printed resist lines was placed in technics-c for times of 1-10 minutes. SEM pictures showed that the resist lines were shrinking but the 0.8μm lines were still rather wide. Another wafer was put into the technics-c for three minutes at a time, rotating the wafer 90° afterwards. This was repeated until the 0.6μm lines started to disappear. The thickness of the resist left after the ashing was about 0.5μm.

Aluminum needed to be deposited over the resist lines. Evaporation has poor resist sidewall coverage and would have been useful for the lift-off, however, x-rays caused by the e-beam would have damage the photodiodes. Annealing the damage might cause the aluminum to spike in the silicon and short out the photodiodes. The CPA sputtering system would not cause as much damage as the e-beam and since the aluminum in the CPA contains some silicon, spiking would less likely occur if the wafers were annealed. The CPA unfortunately gives good step coverage over the resist and the slanting of the resist sidewalls due to the plasma ashing didn't help the situation. Therefore, only a very thin layer of aluminum was deposited (1000A). The aluminum underwent an aluminum etch at room temperature to remove some of the metal from the resist sidewall. Without this etch, the aluminum wouldn't lift off. Also, it was observed that the aluminum on the resist was cloudy and that the quality of the aluminum on the resist was poor and would etch faster. After removing most of the metal in acetone in an ultrasonic bath (~3 hours) the wafer was left in clean acetone overnight to remove residual photoresist. Finally, to really remove the residue in the spaces, a 1 minute descum was done. The liftoff process is schematically summarized in Figure 5.1.

Figure 5.1

Photodiode Aluminum Liftoff Process



VI. Description of Experiment

Experiment Plan

The major part of this project consisted of developing a Microlab compatible process to build photodiodes used for characterizing a wafer stepper. As stated above, the basic idea is to produce large area photodiodes which are then patterned with a line/space grating that is the same pitch as an aerial image to be sampled.

A cross section of the photodiodes produced is shown in Figure 6.1, and a top view is shown in Figure 6.2. The process used to produce the diodes is outlined in the process description section below, and the system for measuring the photodiode signal is outlined in the apparatus section.

The layout of the mask is as follows. Each die consists of nine photodiodes in a 3 by 3 grid (see figure 6.3). Each diode size is defined by the square oxide window opened above the diode. The size of the three diodes in the first row is 500um x 500um. In the second row, there is a 200um x 200um diode, a 100um x 100um diode and a 50um x 50um diode. The third row consists of three 20um x 20um diodes. The nine diodes are spaced such that the contact lines from all nine diodes run vertically and evenly spaced to the bottom of the die. The contact lines are 0.5mm wide and have a 0.8mm pitch.

A row of eight die are processed across the middle of the wafer. The contact lines are brought out from the photodiodes, across the wafer, to the flat of the wafer. From the contact lines, the signal can be taken off of the wafer using a PC board and an alignment bar.

Apparatus

Once the photodiodes are produced on the wafer, a system for measuring the photodiode signal is used while stepping the wafer under the line/space mask. This section describes the electrical connection to the wafer, the amplifiers and the stepper jobs used to step the wafer.

To measure the photodiode currents while the wafer was on the stepper chuck, a new alignment bar and a custom built PC board were used (Figure 6.4). The new alignment bar has a slot in which the PC board lays. The bar is attached to the chuck with two pressure fit pins extended from the bottom of the bar which match two holes in the chuck. The PC board consists of 40 0.5 millimeter lines on a 0.8 millimeter pitch, which exactly matches the size and pitch of the lines of aluminum extension lines on the wafer. The PC board is placed in the alignment bar slot, and affixed with rubber cement. Since the PC board is narrower than the slot, horizontal alignment is possible. At the end of the PC board which contacts the wafer, the space between each metal line has been cut out, producing "fingers" and the board has been thinned in that region giving the fingers flexibility. In addition, solder bumps were placed on each line. Wires were soldered into vias in the PC board, and these wires go to the transimpedance amplifiers which measure the photodiode current. The diode current to be measured is determined by which wire is selected. A ground connection is provided by another wire which is attached to one of the screws used to make the alignment bar. The alignment bar is in electrical contact with the chuck which is in intimate contact with the backside of the wafer when the vacuum is on. This provides the second connection to the photodiodes.

The photodiode current is amplified into a voltage by a transimpedance amplifier. The transimpedance amplifier consists of an LF356 JFET input operational amplifier with a 5 megaohm resistor in the negative feedback path. JFET input amplifier

provide low bias currents which provides more dynamic range for the amplifier. The diode current is fed into the inverting input of the op-amp and the non-inverting input is grounded. Hence, the photodiode is always at zero-bias and should be provide linear response and no dark current. The voltage output of the transimpedance amplifier can be measured with a voltmeter, oscilloscope or a chart recorder. The chart recorder was used the most frequently since it provided a permanent record of the intensity as the wafer was stepped underneath the line space mask.

The wafer was stepped underneath the lines space mask by using a standard stepper job. The die to be used was selected, and the wafer and alignment bar/PC board combination were placed on the stage. The wafer is first aligned using the manual alignment procedures outlined in the GCA stepper manual. The job consists of several passes all of which expose the same die multiple times which an increasing offset in the X or Y direction in a plug pass. Each exposure was set to four seconds, and the photodiode signal was recorded on the chart recorder during this entire four second interval.

Process Description

The following is a brief description of the process flow for fabrication photodiodes in the Berkeley Microfabrication lab. It is intended to provide some insight on why the process was developed the way it was. A complete process flow is given in appendix A.

The starting wafers are P+ prime <100> Si wafers with 1e19 cm-3 concentartion. One of the contacts to the photodiodes is via the substrate. Therefore, a low resistivity P+ wafer is desirable over a high resistivity P wafer.

One micron of N- epitaxial silicon was grown on the wafer. To compensate the epi to make it P-type, two boron ion implantations were done, one with a 5e11 cm -2 dose at 50keV and the second with a 5e11 cm-2 dose at 150keV. The wafer was then annealed at 1000C for 60 minutes. To determine the energy, dose, and anneal conditions for the ion implantation, SUPREM simulations were used. The final result produced a uniform concentration of ~1E16 cm-3 through the epi (Figures 6.5 and 6.6). The epi thickness and concentration needed to large enough so that the depletion region in the epi (caused by the N+ implant done later in the process) would not deplete the entire epi region.

The first photostep consists of putting the alignment mark on the two die that were used for alignment. This needed to be done separately from the N+ photostep. If the alignment marks would not have been put on the wafer before the implantation, they would not have been visible after the implantation since implantation does not leave a visible pattern. Therefore, subsequent photostep layers would not be able to be aligned to the N+ layer. To expose only the alignment marks on the wafer, the N+ mask was covered with electrical tape over the areas where there was no chrome, except for the alignment marks, and then the exposure was done. The wafers were put in silicon etchant for two minutes and after the photoresist was removed, the alignment marks were visible.

The next step was to implant shallow N+ junctions in the photodiode regions. An implant screen oxide 500A thick was placed over the wafer so that the peak concentration from the implant would fall at the silicon/oxide interface. The oxide used is a sputtered oxide that was deposited in the TOPGUN. The advantages of this sputtered oxide (as compared to CVD oxide in a tylan furnace or technics-b) is that the uniformity is excellent (<1% variation over the wafer) and the oxide is deposited at a low

temperature. The N+ activation anneal was done at a low temperature of 750C for 30 minutes to prevent diffusion of the dopants.

The insulation oxide used was 1000A of sputtered oxide. The oxide windows were etched using a wet etch The wet etch was used rather than a plasma etch because a plasma etch could damage the silicon surface. Also, a wet etch would provide sloped sidewalls allowing better step coverage for the metal contacts. The etch rate of the sputtered oxide in 10:1 BHF is 1350A/minute, faster than the etch rate of thermal oxide, ~500A/minute. The area of the oxide window overlaps the N+ region by 2.5um on a side. This was done to make sure that the metal that will be deposited over the photodiodes would not contact the P+ region and short N+ region to the diode P region.

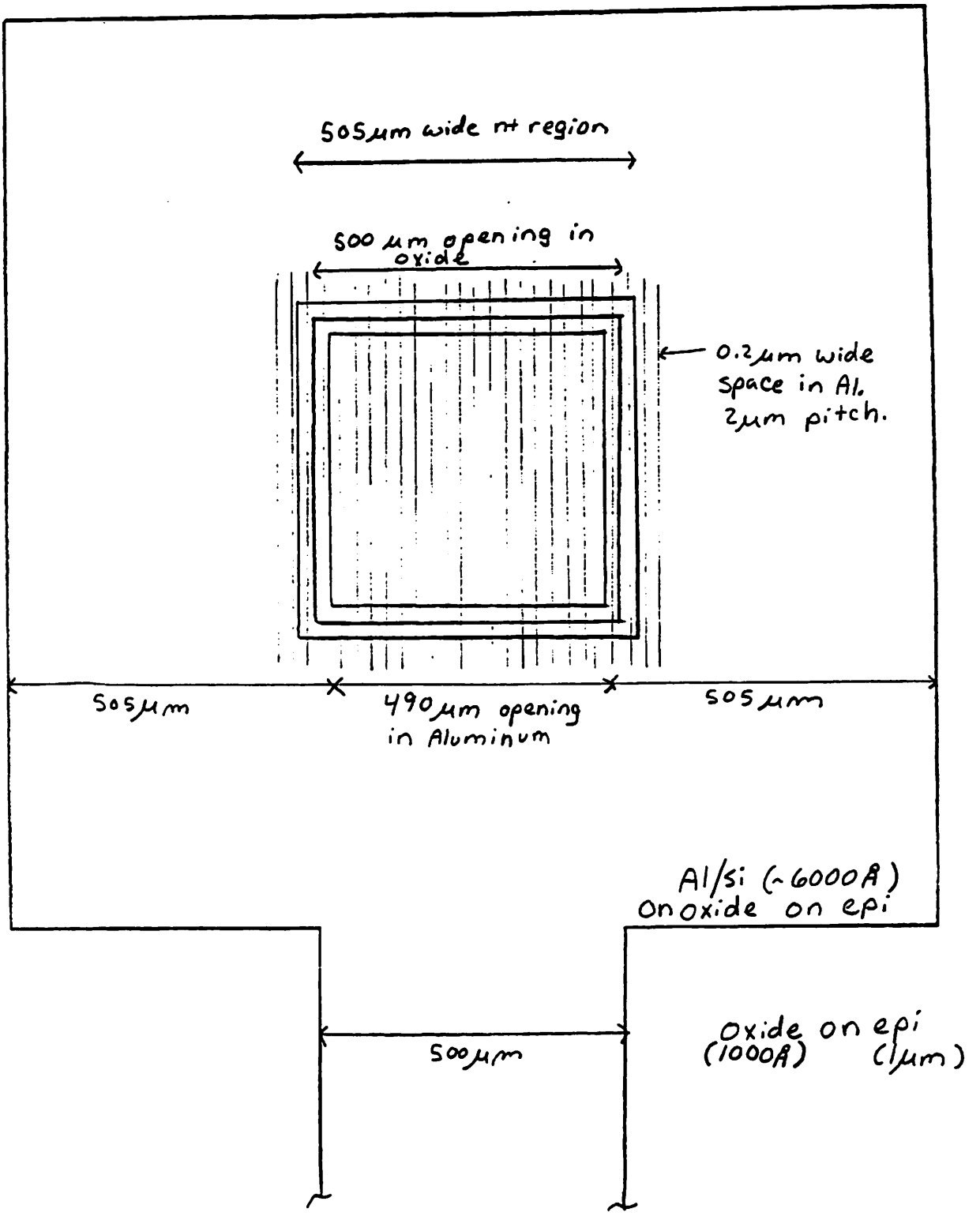
The original photodiode process was designed such that the metal contacts to the diodes would not actually contact the N+ region, but would have an opening identical to the oxide window opening over the diodes. The metal grating was then going to be used as contact between the N+ region and the contact metal. This original design posed some problems. The grating metal needed to be thin (as will be explained later) and thus step coverage of the aluminum over oxide might have been a problem. Also, the diodes should be tested to see if they work before the effort of putting the grating over them is done. Therefore, the contact layer was exposed twice, once with an x and y offset of +5um and once with an x and y offset of -5um. This caused a 5um wide rim along the edge of the N+ region to be exposed. Aluminum was deposited over the wafer and lifted off and thus aluminum was left along the 5um rim, contacting the N+ region. This aluminum could be thicker than the grating aluminum and therefore, step coverage over the oxide was not a problem. Also, the diodes could be tested before the grating pattern was applied. The contact metal lines, which extend to the edge of the wafer to provide electrical contact to the amplifiers were exposed on the wafer using the same resist as was used with the contact mask (before the post-exposure bake and development). Since the spacing of the die in the y direction is smaller that the die size, the die overlap in the y direction. Therefore, a continuous line was exposed from the photodiode to the flat of the wafer. After the lift-off, the aluminum would remain in the lines.

The grating pattern that is put over the diodes has a pitch of 2um and a line width of 0.8um. A linewidth of 0.8um was chosen since it is the smallest line width that can be patterned reproducibly on the GCAWS. It is important to make the line width as small as possible. The lines are plasma ashed until the width of the lines is ~0.2um. The smaller the lines are, the less time they will need to be ashed, and therefore, the less the resist will be ashed away vertically. It is important to have as much resist as possible left so that the lift-off can be done. The method for determining when the resist lines are ashed down to ~0.2um was to ash the 0.8 micron lines until 0.6 micron width "monitor" lines near the 0.8 micron lines almost disappeared. The lines that were initially 0.8um wide are then assumed to be about 0.2um wide. Since the resist has been ashed down to a height of about 0.5um (from ~1.1um), the aluminum that is to be deposited over it and lifted off needs to be kept thin however, it must be thick enough to block out the uv light from the wafer stepper. The thickness was chosen to be 1000A.

After the aluminum was lifted off, the wafer was descumed to remove any residual resist from the wafer. The aluminum was sintered using the SINT400 tylan recipe. This recipe contains two parts to the sintering process, both at 400C: a 5 minute fast sinter and a 15 minute slow sinter. The difference between the fast and

slow sinter is that the fast sinter has a higher flow rate of forming gas than the slow sinter. Since the photodiodes have shallow pn junctions, to prevent aluminum spiking into the silicon, the sintering recipe was cut short by aborting the tylan run when the fast sinter was complete. SEM photographs of the 0.2 micron lines are shown in Figure 6.7.

Figure 6.1



Top view of a photodiode

Figure 6.2

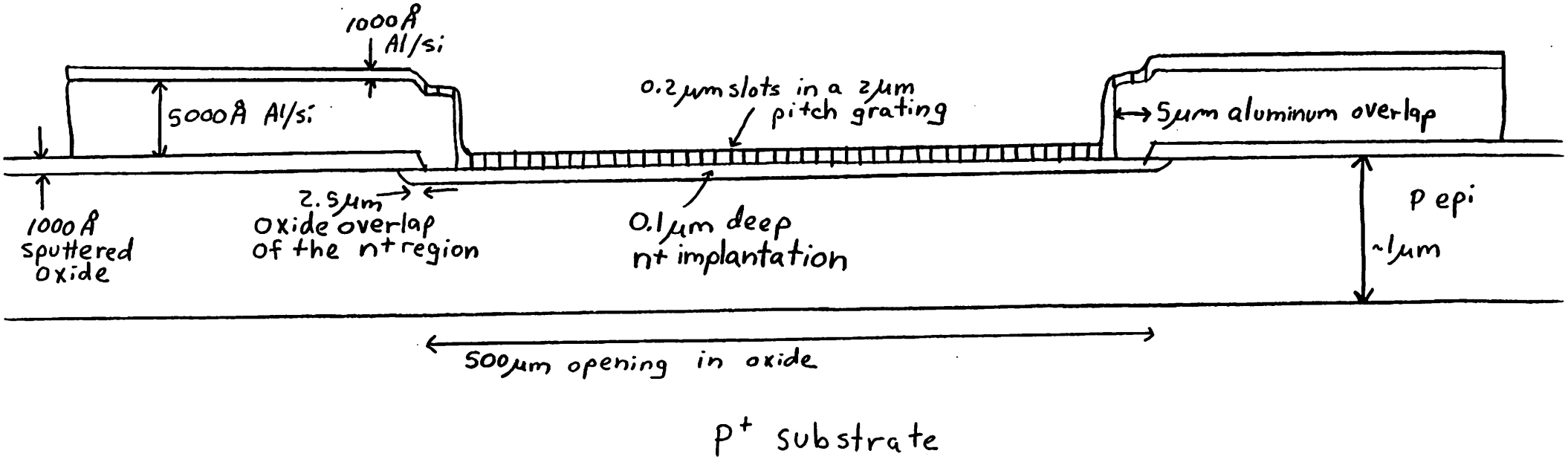

Side view of a photodiode
1000Å Al/Si
5000Å Al/Si
0.2μm slots in a 2μm pitch grating
5μm aluminum overlap
1000Å sputtered oxide
2.5μm oxide overlap of the n+ region
0.1μm deep n+ implantation
P epi
~1μm
500μm opening in oxide
P+ substrate

Figure 6.3

Photodiode Layout

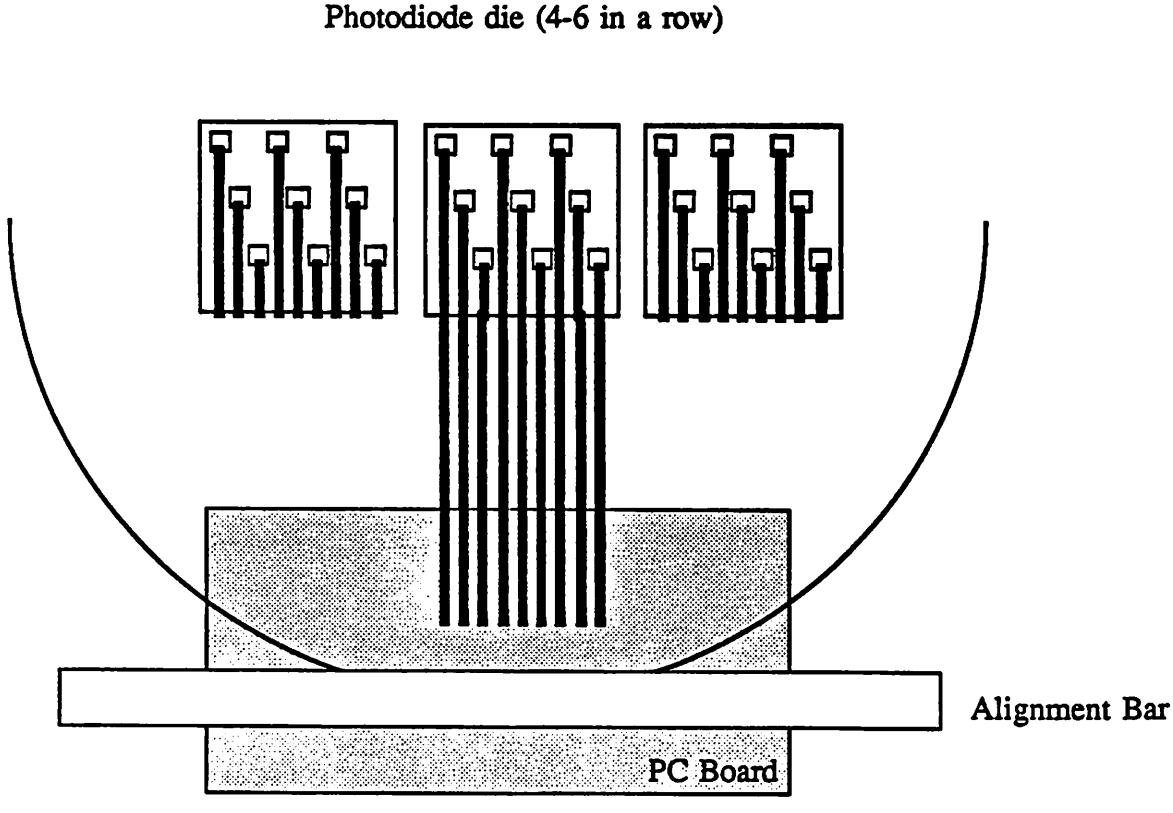
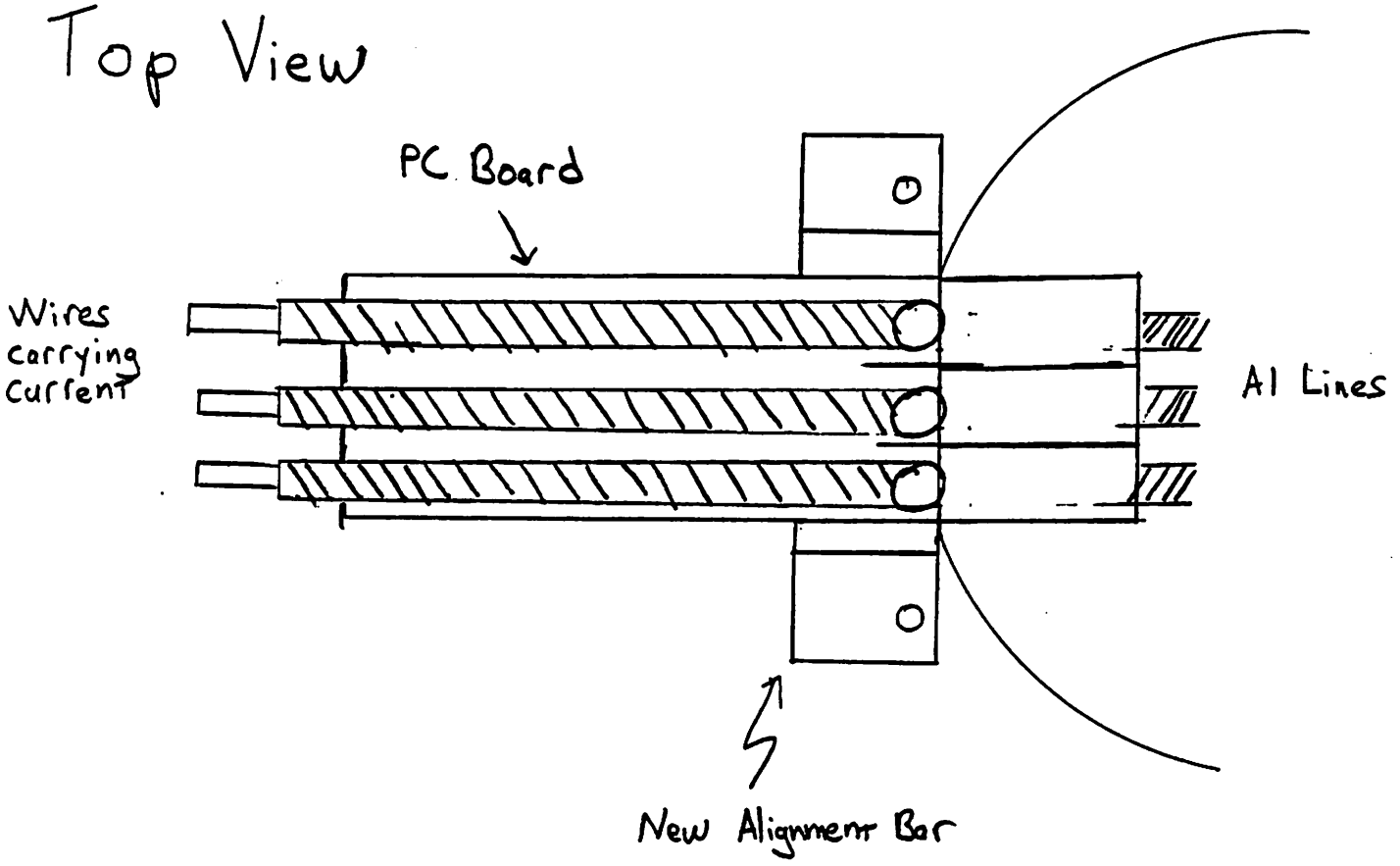


Figure 6.4

Electrical Connection

Top View



Side View

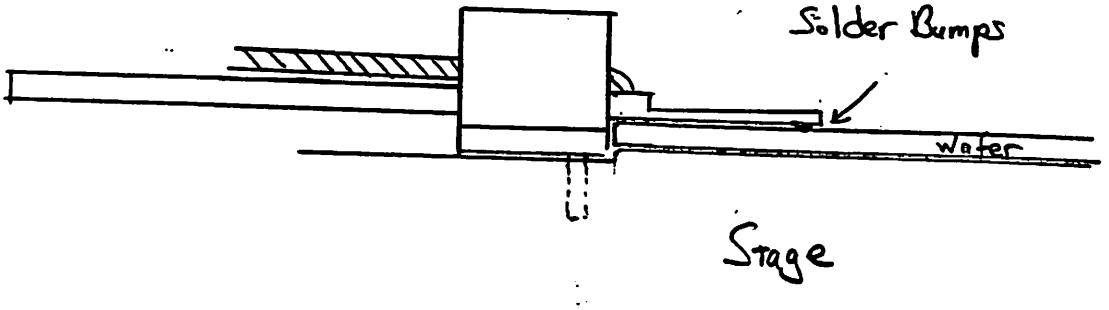


Figure 6.5

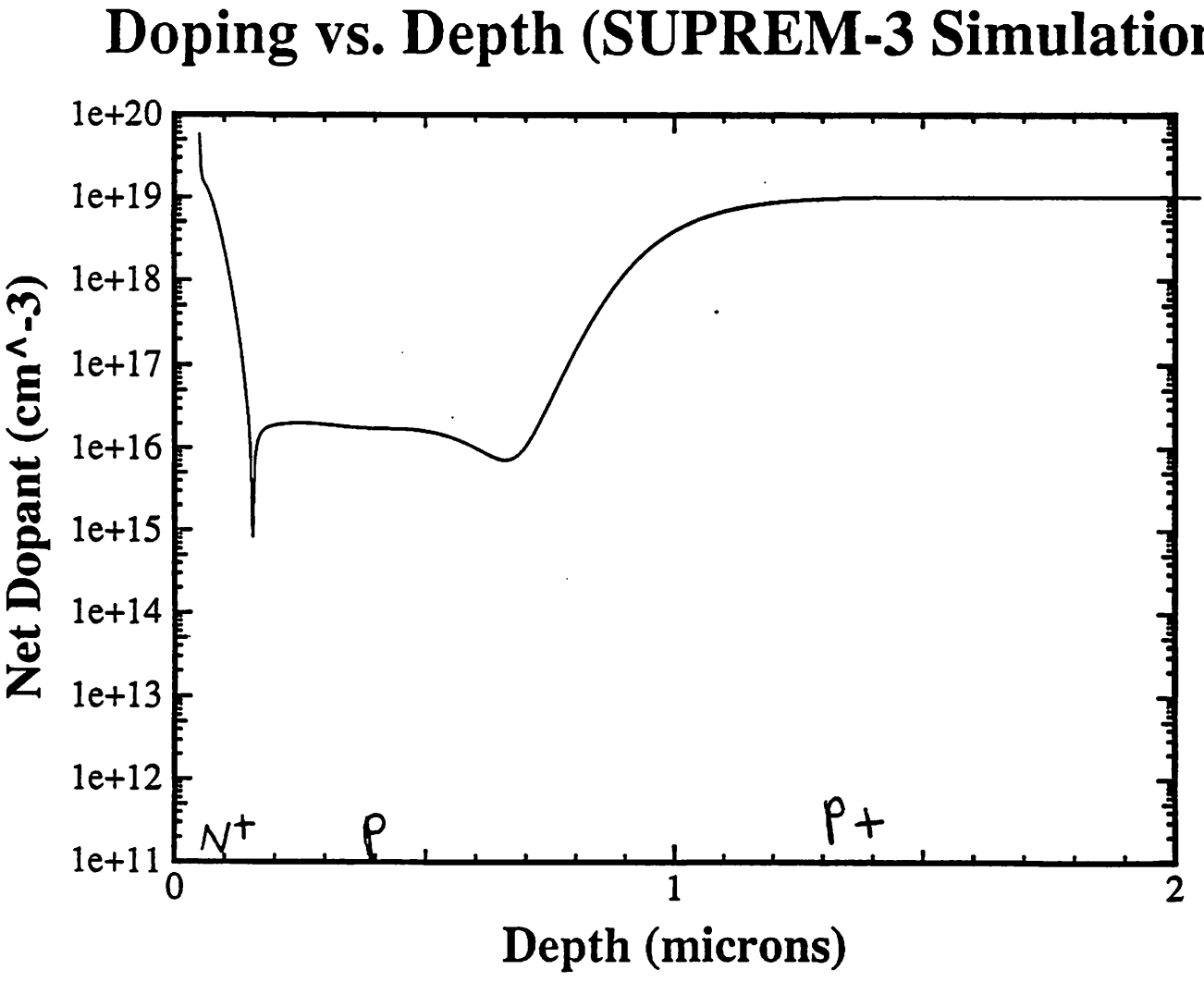
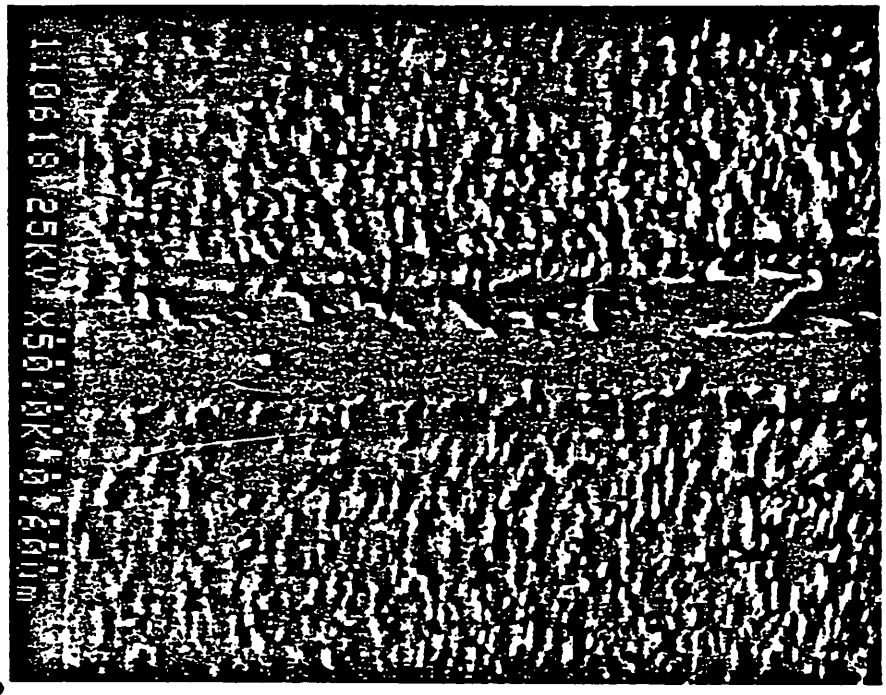

Doping vs. Depth (SUPREM-3 Simulation)
Net Dopant (cm^-3)
1e+20
1e+19
1e+18
1e+17
1e+16
1e+15
1e+14
1e+13
1e+12
1e+11
0
1
2
Depth (microns)
N+
P
P+

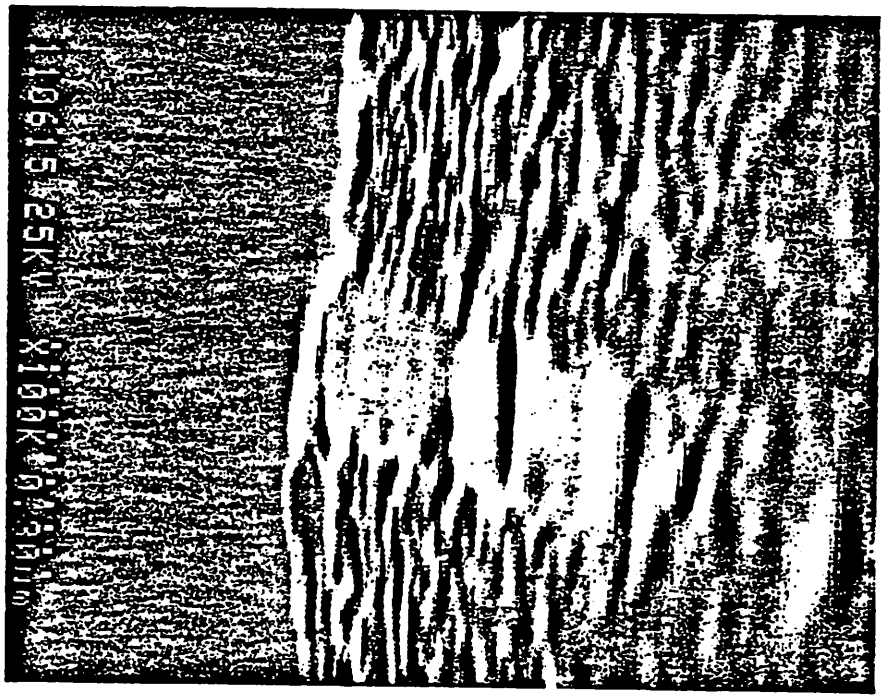
Fig 6.6
SUPREM input file

```
Title UCB Solar Cell Image Characterization Diodes
Comment John Hutchinson 12/16/90


Comment  Start with <100> Silicon, p doped to 1e19
Initialize  <100> Silicon Boron Concentration=1e19
  +                 Thickness=1  Spaces=300


comment Grow the n- epi on the p+ substrate
deposit silicon <100> phosphorous conc=3e13 thick=1
comment print concentration net chemical columns=1

comment print  layers

comment Implant with boron twice to compensate the n- epi to p type.
implant boron energy=50  dose=5e11
implant boron energy=150 dose=5e11

comment print concentration net chemical columns=1

comment Thermal cycle for the boron implant, 1000C, 60 min
diffusion nitrogen temperature=1000 time=60
comment print concentration net chemical  columns=1

comment deposit 500Ascreen oxide (TOPGUN) for implant
deposit oxide temperature=600 thickness=0.05

comment implant n+ contacts
implant phosphorous energy=40 dose=1e14

comment anneal n+ contacts
diffusion nitrogen temperature=800 time=30

print concentration net chemical columns=1
plot lpplot net active
comment print layers
comment end of the process

stop
```

Figure 6.7 - 0.2 um aluminum lines/spaces



0.8um x 2um pitch 60sec etch



0.8um x 2um pitch 30sec Aetch.

VII. Results and Interpretation

Electrical Testing Results

Once the first level of aluminum had been placed on the wafer, the DC characteristics of the diodes were measured with a HP4145 parameter analyzer in the Berkeley Device Characterization Lab. DC results showed that all of the devices performed as diodes. However, the turn on voltage was around 2 volts, which is rather high, and may have been due to a large series resistance. The metal had not been sintered at this point. Figure 7.1 shows a typical I-V curve for one of the 500 micron by 500 micron diodes. It should be noted that this figure plots the reverse voltage and current, not the forward voltage and current, due to the orientation of the probes in the probe station. The effect of illumination on the diodes is to produce a vertical offset of the I-V curve (Figure 7.2). This figure was produced by plotting I-V curves for differing levels of illumination caused by setting the probe station microscope lamp at different settings. Unfortunately, it was not possible to determine a responsivity (Amperes per watt) for the photodiodes because the illumination was white multiple wavelength light. The diodes also exhibited reverse breakdown starting around -5 volts, also an indicator of a large series resistance.

Once the grating level of aluminum had been patterned over the diodes, their DC characteristics were measured in the same apparatus. The unilluminated characteristics were virtually identical, but the illuminated current (Figure 7.3) measured was much less than the unpatterned diodes, possibly due to the decreased "active area" in the photodiodes, but this may also be due to the non-uniformity of the illumination in the probe station.

To reduce the large series resistance and high turn on voltage of the photodiodes, the wafers were sintered at 400C for 5 minutes in forming gas. The I-V characteristics measured with the 4145 (Figure 7.4) do show a decreased turn on voltage, and an increased breakdown voltage. None of the diodes measured exhibited "contact" or ohmic behavior. Ohmic behavior can be caused by junction spiking which was a major concern, since our N+ junction depth is only 0.1 microns.

GCA Wafer Stepper Test Results

Once the photodiodes had been electrically characterized on the bench, they were placed on the GCA wafer stepper stage for further test and measurement. The responsivity of a photodiode at a specific wavelength is computed by dividing the induced photocurrent in a device by the power of the incident radiation.

The first test involved an unpatterned photodiode, i.e. it had no grating patterns, illuminated by a clear mask with no line/space patterns. The photocurrent was measured with a Keithley ammeter with diodes at zero bias. The photocurrent was 2.5 uA, with an estimated incident power of 67 mW/cm2 over the photodiode area of 0.25 square millimeters, resulting in a 0.015 Amperes per Watt sensitivity at 365 nanometers. The second test involved a photodiode which did have a grating pattern, illuminated again by a clear mask, resulting in 1.6 uA of photocurrent, and since the estimated active area of the photodiode is reduced by one-eighth by the grating, the responsivity increases to 0.081 Amperes per Watt. The increase in responsivity may be attributed to less contact resistance since there is more metal in contact with the diode area, and hence more current carrying capability.

Next, the photocurrent was measured using a transimpedance amplifier and strip chart recorder as described above, and the wafer was stepped underneath the line/space

mask that was used to pattern the resist lines used in the ashing procedure. Using the same masks assures no magnification error of the mask relative to the wafer. First, the wafer was stepped in 0.1 micron increments in the X direction, or perpendicular to the direction of the line/space grating, and the signal was measured (Figure 7.5). This was at a focus setting of 317, where a GCA focus "unit" is 0.37microns of defocus per "unit", which was the best focus determined by a focus exposure matrix and wet development. The contrast of the image was 0.27, and decreased for both positive and negative amounts of defocus (Figure 7.6). However, the measured contrast was much lower than the contrast estimated by SAMPLE of 0.96. While stepping the wafer in the Y direction, or parallel to the line/space gratings, the measured signal was relatively constant, but there was some variation, possibly due to vibration of the wafer relative to the mask (Figure 7.7).

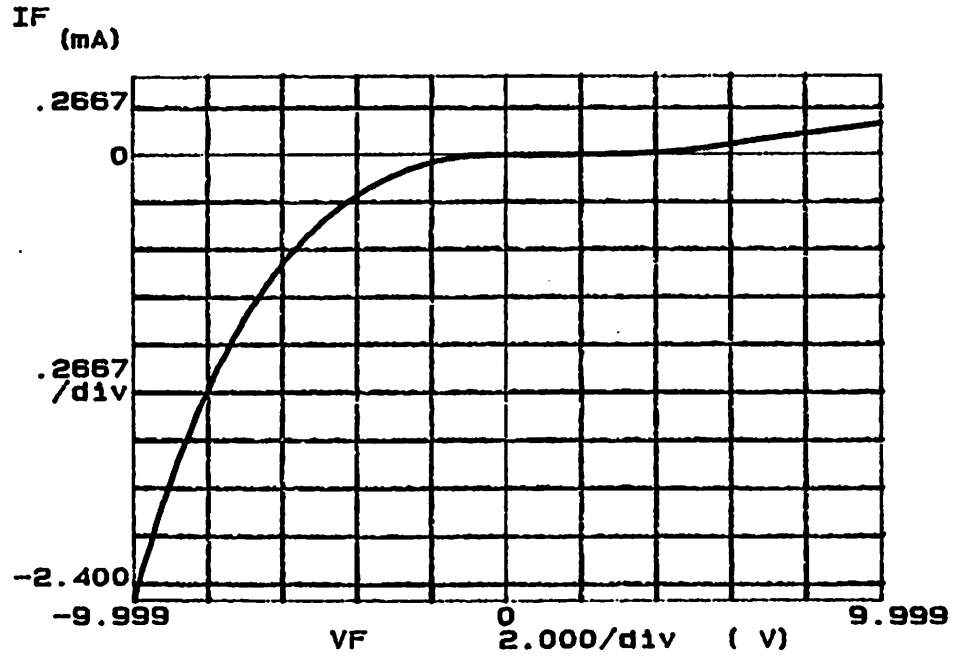
Vibration effects were qualitatively observed (Figure 7.8). Causing vibration in the lamp, the lens and the stage by tapping them resulted in large variations in the intensity during a single exposure. Vibration caused by tapping the stage resulted in the highest variation in measured signal.

Site3

Figure 7.1

500μm device
DC Characteristics

****** GRAPHICS PLOT ******



Variable1:
VF -Ch1
Linear sweep
Start -10.000V
Stop 10.000V
Step .1000V

Constants:
V -Ch2 .0000V

Site 3

Figure 7.2

1 2 (3) 4 5 6 7 8

200 μm
$I_{dark} \approx 5$ nA

200 x 200 μm Device DC characteristics for increasing illumination

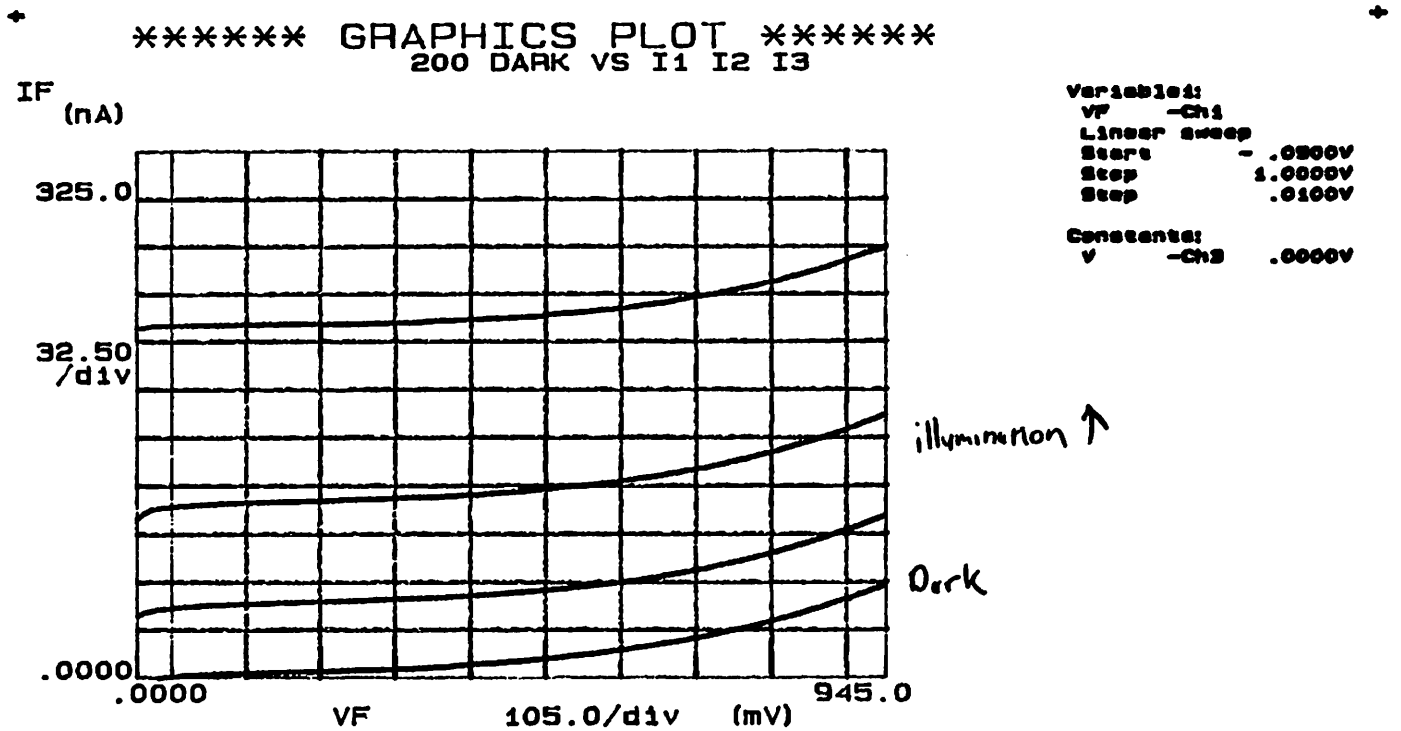
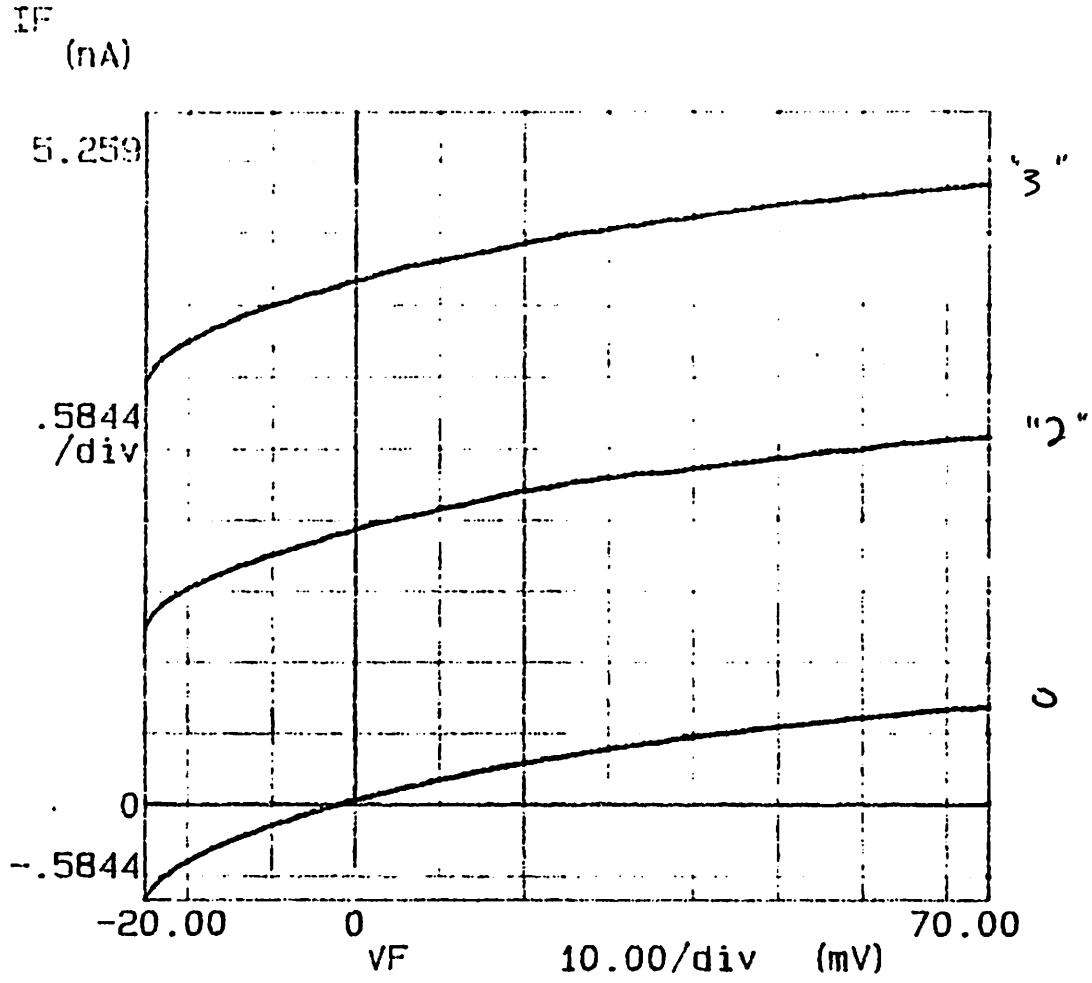


Figure 7.3 - DC characteristics of patterned photodiode versus increasing illumination

11/28/90

****** GRAPHICS PLOT ******



Variable1:
VF -Ch1
Linear sweep
Start - .0250V
Stop .0750V
Step .0010V

Constants:
V -Ch3 .0000V

"P1"

↑ illumination Die 5.A

No illumination

Figure 7.4- DC characteristics for photo diode after sintering

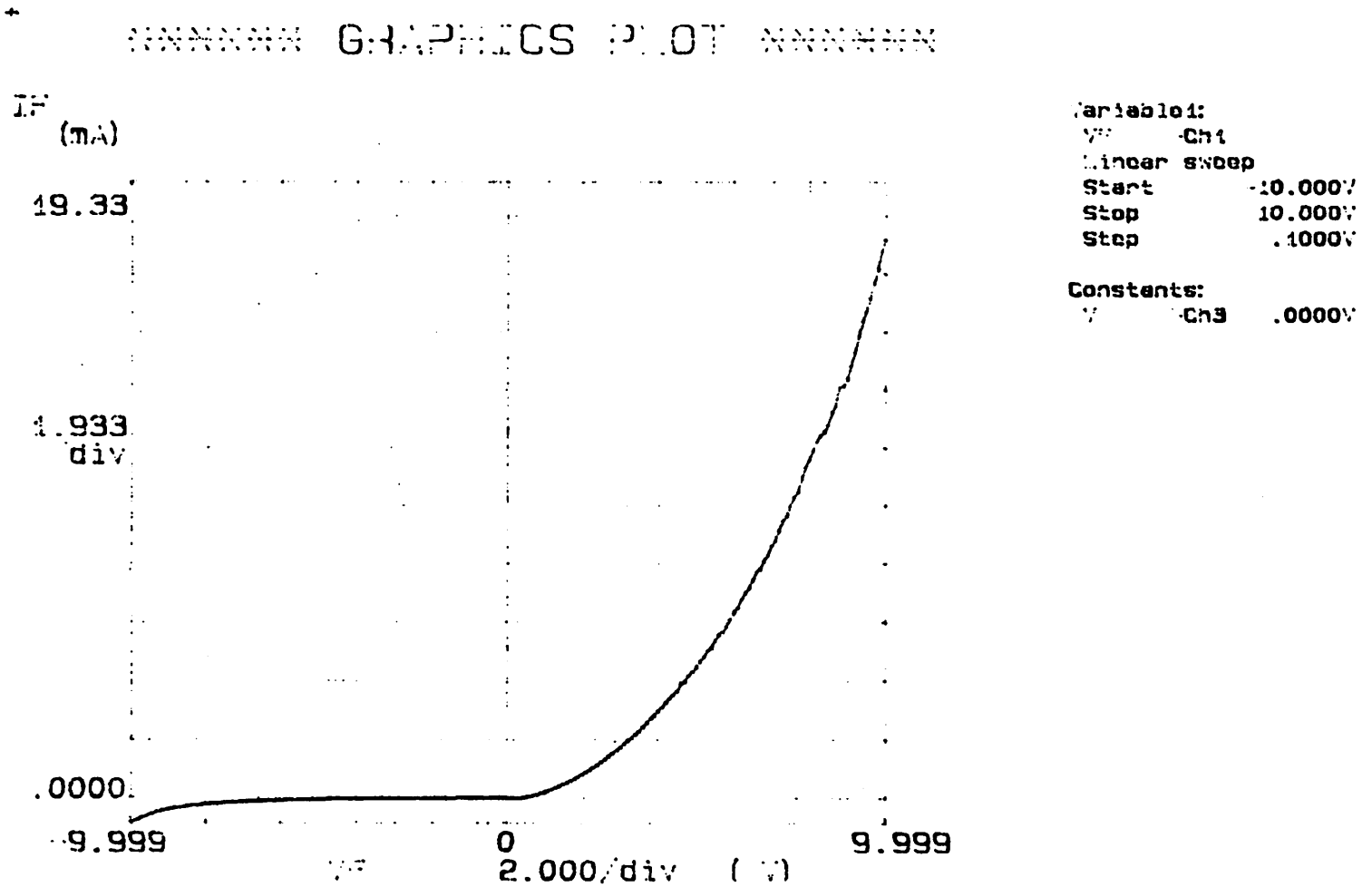


Figure 7.5- Photodiode Signal vs. Position

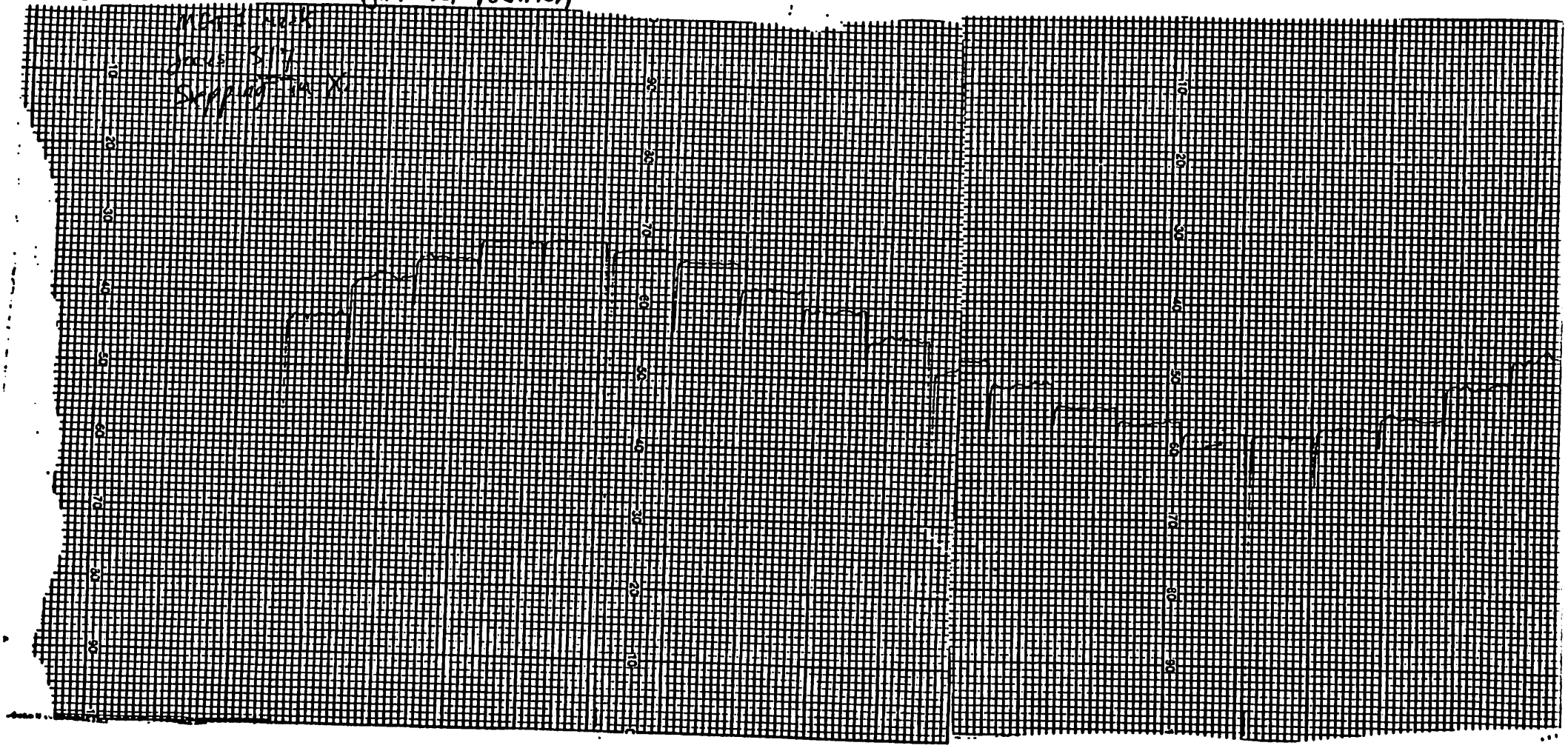


Figure 7.6- Current vs. Position

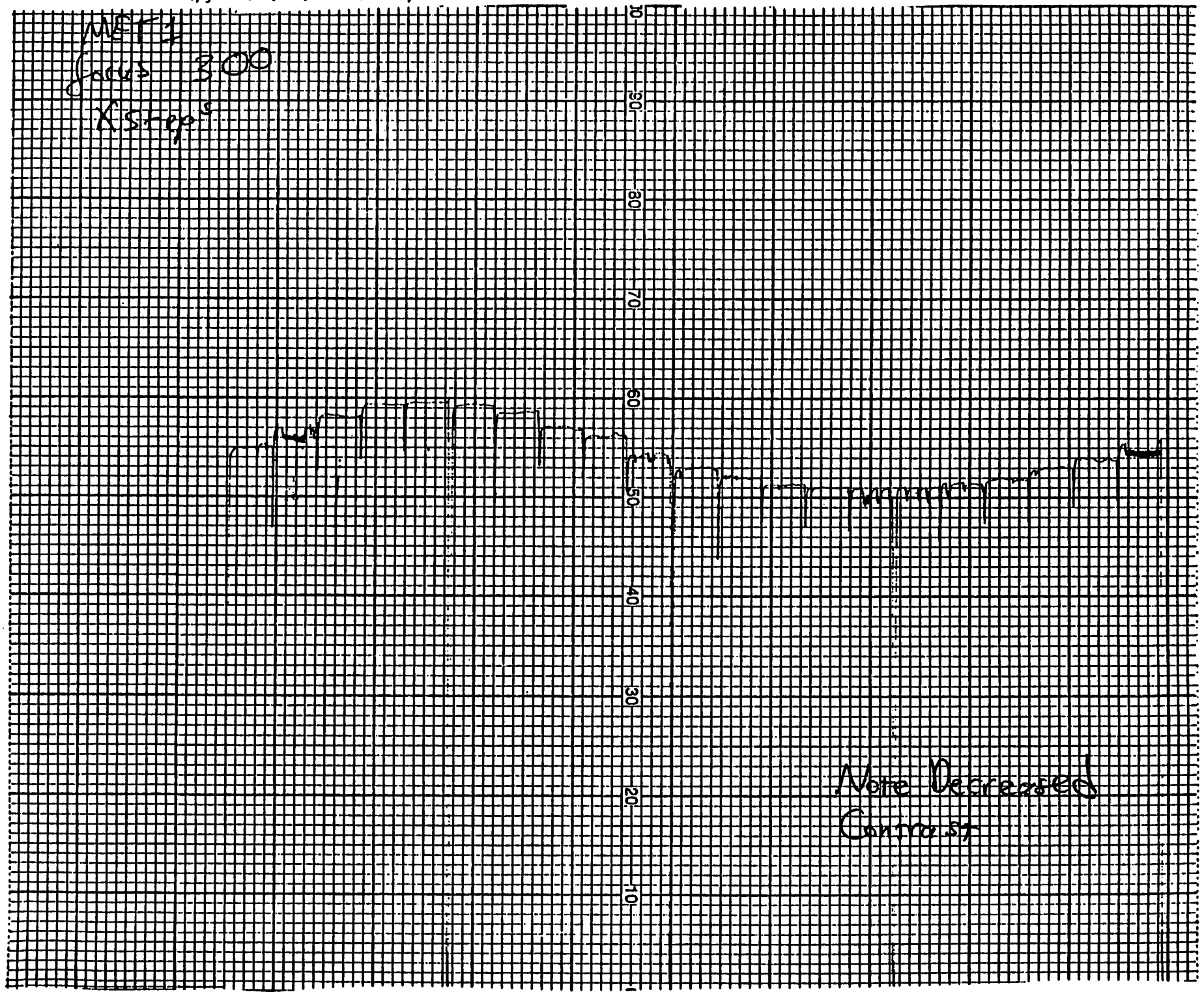


Figure 7.7 - Current vs. Position - Y stepping

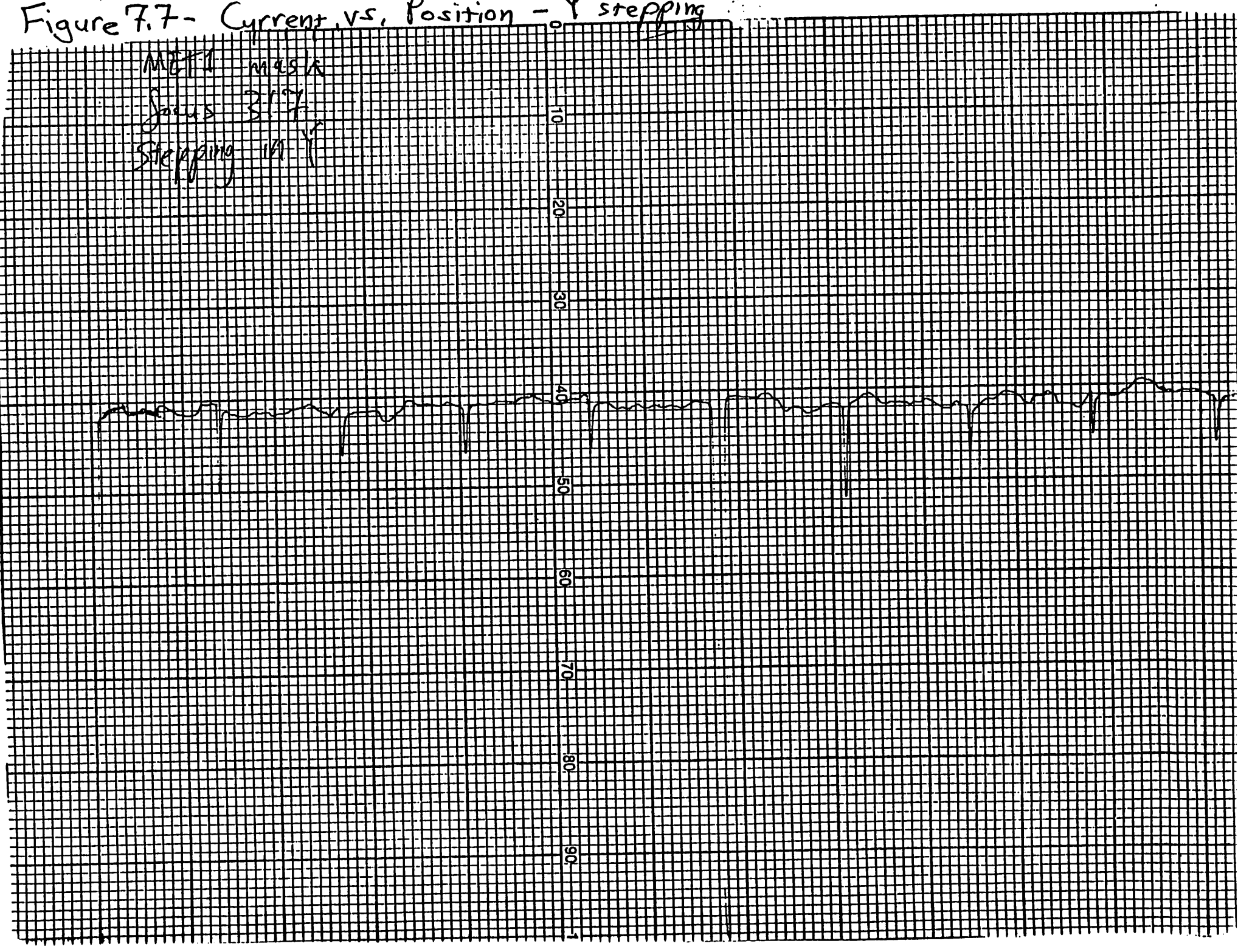
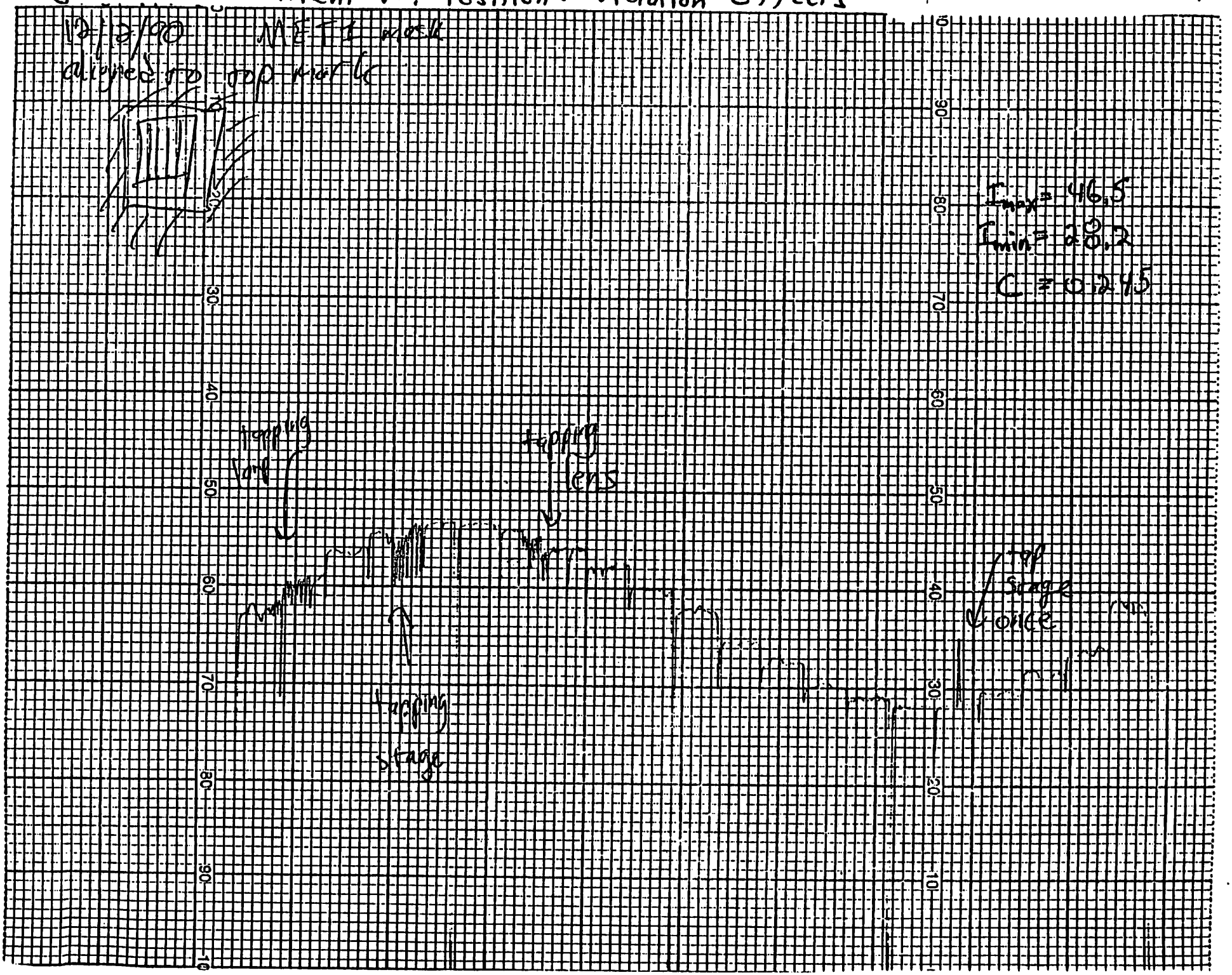


Figure 7.8 - Current vs. Position- Vibration Effects



VIII. Conclusions and Recommendations

Apparatus

In general, the apparatus for measuring the photodiode signal worked well, allowing easy and repeatable electrical connection. However, the spacing between the dies on the photodiodes is different than the 0.8 millimeter pitch of the contact lines, so only one die can be contacted at a time. In addition, the PC board to wafer contact, although usually less than 1 ohm, could be intermittent. Hence, several recommendations are in order. A vacuum system to connect the PC board to the wafer could result in intimate contact, and would allow theta alignment to proceed without concern for the wafer rotating relative to the PC board and becoming misaligned. Also, the PC board should be remade which accounts for the pitch of the lines and the pitch of the die, so that multiple die can be measured simultaneously. The PC Board should also have a "fan-out" pattern, so that soldering wires into the vias can take place in larger holes, instead of the tiny wire-wrap size vias in the current PC board.

Photodiodes

The responsivity of the photodiodes was excellent, and the signal measured was much higher than anticipated. It may be possible to use smaller diodes, as long as the liftoff problems of small areas is solved.

Stepper Characterization

The contrast measured with the stepper image monitor is much lower than contrast estimated from SAMPLE. This is a grave problem and needs to be studied more closely. The thickness of the aluminum grating, greater than 700 angstroms, is sufficient to stop 99.9% of the incident radiation at 365 nm, as long as the metal is free of pinholes and flaking. Despite the problems with contrast, the stepper image monitor can be determine the focus of maximal contrast and can be used to study the effects of different types of vibration on the image quality. In particular, the effects of moving the cassette elevator during exposure, turning the room fans on and off and other vibration sources can be studied.

In addition to focus and vibration studies, aerial images of features less than 0.8 microns can be measured and resolution limits of the GCA stepper explored. Misalignment vectors for translation, rotation and magnification can be measured and studied over time. Lens heating effects [Brunner88] can also be studied.

Appendix A

Photodiode process Steps

Diane Hoffstetter, John Hutchinson, Anton Pfau

Starting Material:

P+ Prime <100> wafers, concentration 1e19 cm-3.

1. Epi growth:

   1um of 3e13 cm-3 n-type epi, process per J.C. Lou.

2. Implant for p- epi:

   Boron, Energy=50keV, Dose=5e11 cm-2.
   Boron, Energy=150keV, Dose=5e11 cm-2.

3. P-epi diffusion step:

   Piranha clean in sink8, then sink6.
   Tylan2, NANNEAL recipe, 1000C, 60 minutes.

4. Level 0 Photostep - alignment marks:

   Mask: "PFAU 1"; Use electrical tape to mask out windows so that only the alignment mark is visible.
   Standard Olin Hunt Process, dehydrate, HMDS, EATON program#15.
   Orient the mask so the label is on the right side.
   Set step size so there are 8 X-steps (10.88), and 11 Y-steps (7.6).
   Expose only the alignment die (7,1 and 7,8) using a plug pass.
   Expose using GCAWS, current best focus and exposure.
   Post exposure bake, EATON program #4.
   Develop using MTI program #70.
   Descum in technics-c at 50 watts for 1 minute.
   Hard bake in VWR, 30 minutes minimum at 120C.

5. Etch Alignment marks:

   Sink 8 wet silicon etch, 2 minutes.

6. Implant Screen Oxide:

   Strip resist using MTI program 10.
   Ash using technics-c at 300 watts for 5 minutes.
   Piranha clean in sink8, then sink6.
   Deposit 500A screen oxide in TOPGUN per Dave Hebert's processes.

7. Level 1 Photostep - n+ implant region:

   Mask: "PFAU 1"; Remove the electrical tape from mask and use methanol to clean if there is residue. Orient mask as in step #4.
   Standard Olin Hunt process: dehydrate, HMDS, EATON program #15.
   Expose row 5 using GCAWS, current best focus and exposure.
   Post exposure bake, EATON program #4.
   Develop, MTI program #70.
   Descum in technics-c at 50 watts for 1 minute.
   Hard bake in VWR, 30 minutes minimum at 120C.

8. N+ Implant:

   Phosphorous, Energy=40keV, Dose=1e14 cm-2.

9. Resist strip and clean:

   Strip resist using MTI program #10.
   Ash using technics-c at 300 watts for 5 minutes.
   Piranha clean in sink8, then sink6.

10. N+ activation anneal:

    Tylan2, NANNEAL recipe, 750C, 30 minutes.

11. Insulation oxide:

    Strip Oxide in BHF, sink8, until dewets (~30 seconds).
    Deposit 1000A sputtered oxide in TOPGUN per Dave Hebert's process.

12. Level 2 Photostep - oxide windows:

    Mask: "PFAU 2"; rotated as per step#4.
    Standard Olin Hunt process: dehydrate, HMDS, EATON program #15.
    Expose row 5 using GCAWS, current best focus and exposure.
    Post exposure bake, EATON program #4.
    Develop using MTI program #70.
    Descum in technics-c at 50 watts for 1 minute.
    Hardbake in VWR 30 minutes minimum at 120C.

13. Etch oxide windows:

    1000A oxide etch, sink7, 10:1 BHF, 1 minute.

14. Strip resist and clean:

    Strip resist using MTI program #10.
    Ash using technics-c at 300 watts for 5 minutes.
    Pirahna clean in sink8.

15. Level 3 Photostep - contact metal:

    Mask: "PFAU 3" and "PFAU 4"; Orient as in step#4.
    Standard Olin Hunt process: dehydrate, HMDS, EATON program #15.
    For mask "PFAU 3" :
    Expose row 5 using GCAWS, current best focus and exposure, x-offset=5um and y-offset=5um.
    Expose row 5 using GCAWS, current best focus and exposure, x-offset=-5um and y-offset=-5um.
    For mask "PFAU 4" :
    Expose rows 1-4 using GCAWS, current best focus and exposure.
    Post exposure bake, EATON program #4.
    Develop using MTI program #70.
    Descum in technics-c at 50 watts for 1 minute.
    Hardbake in VWR 30 minutes minimum at 120C.

16. Deposit Aluminum:

    Deposit 5000A of Al/Si in the CPA at 1kWatt (low power).

17. Aluminum etch:

    Etch aluminum in aluminum etchant at room temperature for 90 seconds.

18. Contact metal lift-off:

    Place wafer in a beaker full of acetone and place beaker in ultrasonic agitator until most of the metal is lifted off (~1 hour). Remove actone in the beaker to get rid of metal flakes but do not let the wafer become dry. Rinse the wafer in acetone until most of the metal flakes are gone. Leave the wafer sit in fresh acetone for about 24 hours (covered) then put in ultrasonic agitator until all the metal is removed. Rinse the wafer well in DI water.

19. Level 4 Photostep - grating pattern:
Mask "MET 1"; Orient mask as in step#4.
Standard Olin Hunt process: dehydrate, HMDS, EATON program #15.
Expose row 5 using GCAWS, current best focus and exposure, x-offset=-120um and y-offset=-750um.
Post exposure bake, EATON program #4.
Develop using MTI program #70.

Descum in technics-c at 50W for 1 minute.
Hardbake in VWR 15 minutes at 120C.

20. Resist ashing:

    Plasma ash resist in technics-c at 50 watts for 3 minutes, rotate wafer 90deg. Repeat 3 times.
    Plasma ash resist for one minute, rotate wafer 90deg. Repeat until 0.6um resist lines start to disappear.
    (Total time ~13 - 16 minutes)

21. Deposit Aluminum:

    Deposit 1000A of Al/Si in the CPA at 1kWatt (low power).

22. Aluminum etch:

    Etch aluminum in aluminum etchant at room temperature for 90 seconds.

23. Aluminum grating lift-off:

    Follow procedure in step #18.

24. Descum:

    Descum in technics-c at 50 watts for 1 minute.

25. Sinter:

    Sinter in Tylan13, program SINT400, only use the fast sinter part of the program (5 minutes).

Process complete.

Silylation of Photoresist Using HMDS

EECS 290N Project, Fall 1990

A. Lee, K. Niazi, P. Sheng

ABSTRACT

A reasonable silylation process using a standard resist, not specifically tailored to silylation, and a parallel plate etcher has been established. By variation of the etch power and pressure a selectivity of unexposed:exposed etch rate of 19:1 was achieved. The achievable selectivity was found to be very strongly dependent on the amount of silicon uptake in the exposed and unexposed regions. The effects of changing the exposure dosage on the resist profiles were also found. A short dip in dilute hydrofluoric acid was shown to be very effective in removing grass. We also found that the degradation of the silylating agent (HMDS) is detrimental to the silylation process.

I. Statement of Problem

There is a continuous drive today to develop lithography processes that can meet the challenges such as ever-reducing line width and severe topography. However, some fundamental limitations have been realized. One of the major problems is the trade-off between reducing bulk effects and reducing reflection. In order to reduce bulk effects and to achieve vertical sidewalls, a very transparent resist is desirable. On the other hand, in order to minimize reflections, resists should be made as absorbent as possible. These two contradictory requirements are extremely difficult to satisfy simultaneously in standard, wet-developed resists. Also, in order to smooth severe device topography, thicker resist layers must be used. This means that the focus range must be comparably deep. However looking at the two limiting equations that Lin[2] defined

$$W \sim k_1 \frac{\lambda}{(NA)}$$

and

$$\Delta Z \sim k_2 \frac{\lambda}{(NA)^2}$$

one can see that for λ = 193nm, NA=0.42, k_1=0.5, k_2=0.35, the depth of focus parameters are:

$$W \simeq 0.29\text{um} \qquad \Delta Z \simeq 0.39\text{um}$$

versus

$$W \simeq 0.6\text{um} \qquad \Delta Z \simeq 1.23\text{um}.$$

Thus to obtain high resolution, there is a very narrow latitude for depth of focus.

II. Review of Previous Work

One of the techniques developed to solve these problems is the trilayer resist scheme first introduced by Havas, et al.[3] In this technique a thick (planarizing), hard baked (high thermal stability) resist is used. An intermediate layer, such as SiO_2, is deposited. A thin resist layer is put on top and patterned. The intermediate layer acts as an etch mask when the pattern of the top resist is transferred to the planarizing layer by dry etch. Most of the reflection problems can be resolved by using dyes in the planarizing layer or in the intermediate layer. However, an inherent problem with this is its complexity. Formation of interfacial layers, stress and cracks are sometimes encountered. Furthermore, the top thin layer is prone to pin-hole problems.

To resolve some of these problems, simplified single layer techniques have been proposed. In such a process the diffusion of the silylating (silicon incorporating) agent is enhanced in the exposed areas of resist. Consequently, when the resist is put in contact with the silylating agent, the bonding of the silicon to the resin of the resist (as shown in figure 1) occurs preferentially in the exposed regions. The wafer is then developed in an anisotropic oxygen plasma. The processing steps are summarized in figure 2.

The major advantages of such a scheme is that the silicon incorporation only takes place in the top 2000-3000 A of the resist. Thus the depth of focus requirements can be greatly relaxed and dyes may be incorporated to minimize reflection problems. Since the silylation step is typically done at elevated temperatures, the resist will also be thermally stable. Also, by using a properly optimized anisotropic plasma etch process, it is possible to vertically transfer the latent image to the bulk of the resist. In this way, well-controlled, vertical sidewalls are obtained.

To date, the best results of a single layer silylation has been achieved by the DESIRE process. DESIRE is an acronym for Diffusion Enhanced Silylating Resist and is commercially available through UCB Electronics. With this process, using an i-line stepper of NA=0.42, 0.4 micron lines have been printed which indicates a k-value as low as 0.44. The SEM photographs presented show these lines to have fairly vertical sidewalls.[4] A selectivity of 40:1 (unexposed:exposed etch rate) has been achieved.[5]

The DESIRE process uses PLASMASK resist, which has been specifically tailored to silylation.[6] The resin plays a very important role since it determines how much silicon can actually bond and become incorporated. It was shown that the silicon uptake depends directly upon the number of OH groups that are present. Usually 8 to 10% by weight of silicon is considered to be the minimum uptake for sufficient selectivity. Standard PLASMASK resist contains enough OH-groups to obtain an uptake of 11% by weight of silicon. Another important parameter is the sensitizer concentration which ensures that proper crosslinking occurs in the unexposed region to inhibit the diffusion of silicon. It is claimed that of all novolak-diazoquinone resists that were tested, only the PLASMASK resist had the chemical composition to preform well in the DESIRE process.[7]

III. Objective

The basic objective of this research was to develop a standard silylation process for resist which is not specifically tailored to silylation and achieve a 20:1 selectivity of unexposed:exposed etch rate. It was also our goal to achieve this down to a line width of 0.8 micron.

IV. Basic Process Flow and Apparatus

The basic process flow is schematically represented in figure 2. Bare silicon wafers are first cleaned using the standard cleaning procedure. MacDermid 1024, the photoresist used in this process, is spin coated and softbaked using the Eaton Wafer Track. The resist was spun on at 5000 rpm and soft baked at 100° C for 1 minute. Since the development takes place in a dry environment it is not necessary to prime the wafer before coating. The resist is then exposed using the GCA Stepper. The wafer is then exposed to HMDS (hexamethyldisilazane) vapors in the silylation apparatus which is shown in figure 3. The silylation apparatus consists of the vacuum chamber, hot plate for heating the wafer, baratron, inlet lines for HMDS and nitrogen, and pumping line. Since it is important that no condensation of the HMDS vapor occurs, the HMDS flask, HMDS inlet line, hot plate, and the top cover of the chamber was all heated to 90° C, 125° C, 100° C, 125° C respectively. The silylation process consisted of pumping down the chamber to a base pressure of 1 torr, vacuum/nitrogen purging of the chamber, backfilling the chamber with HMDS to a pressure of 110 torr, silylating for a specific length of time, followed by two purge cycles. The final process step is developing the resist in a reactive ion etcher. A quartz cover was placed on the powered electrode in order to prevent the sputtering of the electrode material onto the samples. The samples in turn were placed on the quartz cover. The powered electrode was driven at 13.56 MHz. The other electrode was grounded.

V. Approach

From the initial survey of the problem, we noticed that one of the challenges of this project was that the two key pieces of equipment to be used,

the silylation apparatus and the reactive ion etcher, was not fully characterized. Thus we spent some time in understanding and characterization of the equipment.

It was also realized that there are basically two parts to this project that are strongly coupled. The first part is the silylation. This phase includes the optimization of the exposure dose as well as various parameters of the silylation step. The latter step can be affected by large number of variables such as post-exposure bake temperature, post-exposure bake time, silylation time, silylation temperature, silylation pressure, HMDS temperature. It was found from a previous statistical experiment that the silylation temperature, time, and pressure have significant effect on the selectivity and critical dimensions.[8] The other part of the experiment is the dry developing by plasma etching. The variables affecting this phase are RIE power, pressure, and oxygen flow.

The basic approach was to start first by optimizing the plasma etching phase. Thus we began the investigation by first comparing the etch rates obtainable for fully silylated areas as opposed to those obtained for a completely unsilylated wafer. In this way, we were able to concentrate only in optimizing the plasma etching parameters to achieve the best selectivity without concern for effects of the silylating parameter. We chose to look at the effects of varying pressure and power of the etch.

Our next planned step was to see if we could reproduce these effects for large patterned areas on the same wafer.

Our next step after achieving the best selectivity for large areas was to look at small features (down to 0.8 micron lines). We expected that at this stage it would be necessary for us to look at the silylation parameters to achieve good critical dimension control for small features. Considering the

previously done statistical screening experiments, we chose to vary the exposure energy and silylation time. We then planned to study the achievable anisotropy and linewidth of the patterns by viewing the samples with a scanning electron microscope.

Finally we planned to look at the residue problem, quite often reported in the literature, and possible ways of removing and preventing the residue.

VI. Results and Discussion

A. Equipment Characterization

The parallel plate that we used was assembled by John Alonzo, a graduate student in the Chemical Engineering Department. He instructed us in the use of the etching equipment and shared with us the results from his own silylated resist experiments which gave us an indication of the location in etch-parameter space that we should be operating in. The etching setup was left as was, except for the addition of an automatic matching network obtained by Perry Sheng from his employers at TEGAL. This helped speed things up a bit, especially on those runs when many different etch parameters were used.

The Silylation apparatus, located in room GL4 in the microlab was assembled by Chris Spence. Although this equipment had been used for silylation in the past, and was in fact to be used for a study of the effect of resin chemistry on the silylation process, concurrently with our experiment, it had not been fully characterized, which is to say that the results of the silylation were not completely reproducible. Our trials and tribulations with the silylation, numerous as they were, were useful in helping us understand the pitfalls associated with this process.

Upon silylating our first samples we noticed a non-uniformity in the amount of silicon uptake across the wafer. This was worrisome because the amount of

silicon uptake is one of the parameters that determine selectivity, and we were hoping to find that this was a determined by the "silylation parameters." In an effort to understand this non-uniformity we looked at the uniformity of the hot plate temperature. The temperatures we measured were uniform or near-uniform across the hot plate. Based on our observation that the regions of largest silicon uptake were adjacent to the gas lines, we further explored this little understood correlation by rotating our samples and noting the orientations with respect to the inlets. This study was inconclusive partly because at about the same time we were alerted to a seeming degradation of HMDS with time, and we started changing our HMDS flask more regularly upon which both the non-uniformity and the failure to uptake silicon both disappeared.

Another phenomenon that we observed in our first weeks of silylation was the completely unexpected observation that the exposed silylated portions of the photoresist were at times thinner than the unexposed unsilylated portions of the photoresist. This was contrary to our picture of the silylation process in which the silylated areas swell as a result of the uptake of silicon. This thickness difference was, in fact, what we were using as an indicator for the amount of silicon uptake. A comparison of exposed and unexposed photoresist thicknesses for a wafer that had been exposed and baked but not silylated clarified the situation. The results showed that, prior to silylation, the exposed photoresist was approximately 100 A thinner than the unexposed photoresist, so, upon silylation, any silicon uptake failing to make up for this deficit will result in a silylated region which is thinner than the unsilylated region.

We feel that some of the problems we encountered: the variability of the silicon uptake as a function of silylation date, and the failure, at times, to silylate at all, are due to the degradation of the HMDS with time, perhaps due

to its hydrolization in the presence of water. Persons in the semiconductor industry who use HMDS, and whom were contacted stressed the importance of not exposing HMDS to moisture. Our HMDS samples on the other hand were routinely exposed to the atmosphere when we, or fellow users of the silylation equipment had to change the silylating agents for our own respective silylation experiments. To look for the suspected chemical changes that the HMDS had undergone we evaporated a drop of HMDS from a year-old sample, a 3 weeks old sample and a fresh bottle on a clean wafer. Not surprisingly the results looked different in each case: the old HMDS leaving the most residue, the fresh HMDS leaving hardly any at all. This may be a reasonable way to test for the possible degradation of HMDS.

B. Achieving Best Selectivity for Fully Silylated vs. Completely Unsilylated Photoresist

We started our experiment by etching photoresist which, through the use of a large exposure energy (exp. time 1.8 s.) and a long silylation time, was over-silylated. Our aim was to find the optimal region in the etch-parameter space. By looking at the ratio of the etch rates for unsilylated wafers (which had never been exposed to silicon) to the etch rates for exposed areas on the over-silylated wafers we were also measuring an upper limit on the etch-selectivity of unsilylated to silylated resist on the same wafer. We observed etch rates of 50 to 200 A/min for silylated resist (fig. 4), 600 to 1800 A/min for unsilylated (fig. 5). Samples etched at 150 Watts were burnt, and so we limited the power to the etcher to a maximum of 100 Watts. Our best selectivity of ~18 occurred at 20 mT, 100 Watts (fig. 6). At this point we were not interested in looking at small features, for which the etch selectivity was expected to depend on the

silylation conditions. That a smaller amount of silylation was required for any process involving small features was evident in the fact that etch rates for unexposed, and hence supposedly unsilylated, regions on our samples were too slow, due to the considerable silicon uptake for these regions (fig. 7).

C. Achieving Best Selectivity for Reduced Silylation

We decreased the amount of silylation by decreasing the exposure energy of the photoresist (exp. time 0.9 s.). Attempting to optimize the etch conditions for our new silylation conditions we arrived at 20 mT/50 Watts. The optimal etch conditions for the previous step, 20 mT/100 Watts resulted in poor selectivity, because the etch rate for the now less-silylated exposed photoresist was too large (fig. 8). A selectivity of 15 for this process showed promise and prompted us to proceed to the next phase of our experiment in which we attempted to optimize the selectivity further by looking at wafers patterned with small features.

Our results from this initial phase of the experiment indicated that the etch rate for our silylated photoresist is reasonably constant, and that it is increased more or less abruptly once the silylated layer has been removed. These results were consistent with our rough model of the exposed photoresist consisting of a uniform silylated later on top of a uniform un-silylated layer. So our rough model of the exposed photoresist consisting of a uniform silylated layer on top of a uniform un-silylated layer was consistent with the results.

D. Best Selectivity for Fine Features

The data obtained for the wafers patterned with fine features is shown in figure 9. The exposure time is 0.9 seconds and the wafer was silylated for 20 minutes. The etch rate for the exposed region was slow, an average of 70A/min.;

whereas the etch rate for the unexposed region was on average 1050A/min. This resulted in a selectivity of 15:1. It is also good that the etch rate of the exposed areas does not increase which indicates that even at the longest etch time of 9 minutes the SiO^X was not completely etched away. Since the increase in the etched thickness stops at about 8.5 minutes, we concluded that the development is complete.

From inspection with the optical microscope, we observed that the linewidths for the 0.9 second exposure, 20 minute silylation wafer looked broader than the spacing. The mask used has equal linewidth and spacing. We initially thought that the cause of this reduction in linewidth was a lack of silicon uptake, caused by either insufficient silylation time or lack of exposure energy. However, T. Rucker et al. of Intel Corp.[9] have reported that the linewidth of the resulting pattern is broadened for both too low energy as well as too high energy exposures. Furthermore, they have also reported that the linewidth is very sensitive to exposure energy for silylated samples, whereas for the wet developed sample the linewidth is relatively invariant with exposure energy changes. Thus we proceeded to try to achieve better linewidth control by varying the silylation time and the exposure energy.

Table 1 summarizes the results of the effect of varying silylation time on the etch rates and selectivity. It can be seen that for short silylation time the delta is smaller and the etch rate of the exposed region increases. From this we can conclude that when the silylation time is increased there is more silicon uptake in the exposed region (increased delta) which results in a slower etch rate of the exposed resist. Consequently with increased silylation time, the selectivity improves, provided that we do not oversilylate.

Exp. Dose 1.35 sec.		Etch 50W, 20mT		
		Etch Rate (A/min)		
Silylation Time	Delta	Unexp	Exposed	Selectivity
10 min	80	940	60	16
20 min	280	950	50	19

Exp. Dose 0.9 Sec.		Etch 50W, 20 mT		
		Etch Rate (A/min)		
Silylation Time	Delta	Unexp	Exposed	Selectivity
10 min	-40	970	110	9
20 min	80	1120	75	15

*Delta: (silylated thickness) - (unsilylated thickness) before etch

Table 1. Effect of Varying Silylation Time

E. Study of Achieved Linewidth, Anisotropy & Residue Removal using SEM

Observations were made using the scanning electron microscope for the wafers with the process conditions summarized in Table 2.

Power	Pressure	O_2 Flow
50W	20 mT	10 sccm

Silylation Time	Exp. Dose	Etch Time
20 min.	0.9 sec.	12 min.
10 min.	1.35 sec.	12 min.
20 min.	1.35 sec.	13 min.

Table 2. Process Conditions for Wafers used in SEM Study.

After the plasma developing, the wafers were then dipped in 10:1 HF solution for approximately 30 seconds.

Figure 10 shows a 0.8 micron width line. The striking feature of this image is the very dense grass that can be seen fairly uniformly across the field. This residue is often referred to as 'grass' because of its appearance. The height of the grass is approximately half that of the resist. This residue problem is apparently an inherent problem of the single layer silylation process and has been reported for such commercial processes as DESIRE.

There are presently two widely held hypotheses regarding the cause of the formation of grass. The first proposed mechanism is the limited silylation of the unexposed area of the photoresist. In this model the limited silylating agent's diffusion and reaction with the unexposed area may cause large enough silicon uptake to result in micromasking. The second proposed mechanism is the sputtering off and re-deposition of the SiO_x from the surface of the exposed area to the unexposed area and the formation of micromasks.

That grass can be greatly reduced by a dip in hydrofluoric acid has been reported by UCB Electronics, suppliers of DESIRE process[5] and was observed by us(fig. 11). HF does not attack photoresist or bare silicon. A possible explanation for the grass removal ability of HF is that it attacks the native oxide layer immediately under the grass. The thicker lines are protected from any similar lift-off effect by their large base dimensions. Overetching might be useful in the further reduction of residue as can be seen by comparing samples etched at 12 and 13 minutes (figures 11 & 12).

The SEM pictures allow one to compare the integrity of the resist as a function of the silylation parameters. With an exposure of 1.35 s, and a silylation time of 10 min. we see a %16.5 reduction of line width, and jagged line surfaces (fig. 13). An increase in the silylation time to 20 min. results in better critical dimension control and a smoother, flatter line surface. This

is entirely consistent with the larger uptake of silicon expected for the longer silylation time; the resulting mask being more etch-resistant.

The trapezoidal profile of the 2.0 micron line shown in figure 14 strongly indicates the presence of mask erosion. In figure 15 profiles of 3.0 micron lines for 0.9 sec. exposure, 20 min. silylation time is shown. The preserved linewidth in this sample suggests that the longer silylation times has increased the mask's resistance to etching (despite a reduction in exposure energy). In addition the straight and virtually horizontal upper half of the profile suggest that mask erosion has been reduced. The reasons for the curving of the profiles near the wafer surface are not clear, though it should be noted that similar profiles have been reported for oversilylated resist by Rucker et. al.[9]

VII. Conclusion

In this work, we have demonstrated that a reasonable silylation process can be established using a standard resist, not specifically tailored to silylation, and a parallel plate etcher. Table 3 summarizes the best process conditions determined in these experiments.

Silylation Parameters

Resist	Exposure	Silylation Time	Pressure	Temperature
MacDermid 1024	1.35 sec	20 min	110 T	110°C

Plasma Etch Parameters

	O_2 gas	Pressure	Power (13.5 MHz)	Time
Bulk	10 sccm	20 mT	50 Watt	9 min
Overetch	10 sccm	20 mT	50 Watt	4 min

Etch Rate Unexp Resist: 950 Å/min
Exp Resist: 50 Å/min

Selectivity: 19:1
Line width/Pitch: 1:1 (for 1 micron line)

Table 3.

By variation of the etch poser and pressure, we were able to achieve a selectivity of 18:1 for the etch condition of 50 W, 20 mTorr. These conditions allowed us to achieve a reasonably fast etch rate for the unexposed region and also to complete the etch without breaking through the SiO_x mask of the exposed region. The achievable selectivity was found to be also dependent on the amount of silicon uptake in the unexposed and exposed regions. The uptake of silicon varied predictably with the exposure dosage and silylation time. However, the effects of changing the exposure dosage on the resist profile — an increase of linewidth observed at both extremes of the exposure dosage — which was also reported in literature is not well understood currently.

Through this study, we have found conditions for the process on which a more detailed analysis of the effects of varying parameters can be based.

The degradation of HMDS is a problem that requires serious "looking into". That the HMDS undergoes degradation with time, was confirmed by our "evaporation" experiment, where the amount of residue left by HMDS samples was correlated to the age of the samples. The inadequate silylation of some of our samples was clearly due to this aging effect of the HMDS. Inexplicably, a gradual non-uniformity of silylation across the wafer seems to be a result of aged HMDS, also. For well-controlled, reproducible silylation runs, we can not over-emphasize the importance of the quality of the HMDS.

The residue problem is one for which there seems to be a very good solution: a quick post-etch dip in hydrofluoric acid. Although due to time constraints we were unable to study the effect of the concentration of the acid or length of immersion times, the results, as indicated by the SEM pictures are quite remarkable.

VIII. Recommendations for Further Research

A. Presilylation Bake

It has been previously reported that a prolonged bake before silylation considerably improved the achieved selectivity and profiles. As explained earlier, the selective silylation is a result of the difference in diffusion rate of the silylating agent into the exposed and unexposed areas. With a presilylation bake the diffusion rate in the unexposed region is decreased. This is currently thought to be a result of thermal crosslinking of unexposed naphthoquinone diazide in the resist during the bake which inhibits the silylating agent's diffusion. It has also been reported that the presilylation bake resulted in greatly reduced grass formation. From this it seems more likely that the grass formation is due to incorporation of silicon in the surface of the unexposed areas.

B. Fluorine Chemistry Pre-etch for Residue Reduction

It has been found that a short pre-etch consisting of flouring containing gas such as CF_4 was effective in reducing the residue. The fluorinated chemistry pre-etch helps to eliminate the silicon incorporated in the surface of the unexposed region. However the selectivity achievable in the oxygen etch has been found to be extremely sensitive to the presence of fluorine. One study has shown that a selectivity of 25:1 would reduce to 1:1 immediately after an etch containing fluorinated gas. The selectivity improved back to 25:1 only after a thorough isopropanol wipe-down of the system and 8 hours of oxygen plasma cleaning. From this it seems that a fluorine chemistry pre-etch probably must be done in a separate system in order to prevent the loss of selectivity.

C. Optimization of HF Dip

It has been shown that a post-develop HF dip was successful in removing the grass. However it is necessary to determine what the shortest duration of this dip is for the effective removal of grass. It is important to keep this step as short as possible to minimize the possible damage to underlying layers.

D. Understanding HMDS degradation

One important process implication we have encountered is the possible HMDS degradation and its detrimental effect to the silylation process. It is important to understand better the rate at which this degradation occurs. We have also suggested that a valve with a rotating stop cock be built in the flask containing the silylating agents so that the chemicals would not have to be exposed to air when the silylating agents are changed.

Acknowledgements

We would like to express our gratitude to Chris Spence who generously allowed us the use of his silylation apparatus. This work would not have been possible without his thoughtful suggestions and patient guidance. We would also like to express our appreciation to John Alonzo for allowing us the use of his etcher as well as for his technical assistance. We would also like to thank Kim Chan for her help in sample preparation and Richard Hsu for his help in SEM photography. Finally, we would like to express our gratitude to Professors. Neureuther, Oldham, and Lieberman for their suggestions regarding this research and their excellent lectures in class. We have certainly gained a lot from this experience.

References

1. G.N. Taylor, O. Nalamasu, L.E. Stillwagon, "Materials & Processes for Imaging at the Resolution Limits of Optical Lithography," Microelectronic Engineering 9 (1989) 513

2. B.J. Lin, Microelectronic Engineering, 6 (1987) 31

3. J.R. Havas et. al., U.S. Patent 3,873,361 (1973)

4. R. Roland et. al., "Thermal Crosslinking Dry Unexposed Naphthoquinone Diazides as Diffusion Inhibition Mechanism in the DESIRE Process" Proceedings of SPIE 920 (1988) 120

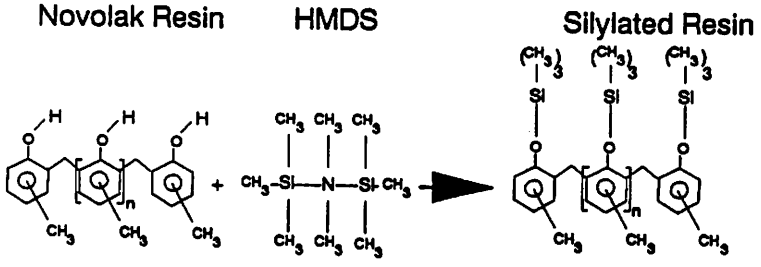
5. F. Coopmans et. al., "Dry Development with Reactive Ion Etching (RIE) and Magnetron Enhanced Etching (MIE) Technology," Proceedings of SPIE 772 (1987) 159

6. R. Roland, et. al., "Optimisation of Resist Composition for the DESIRE Process," Proceedings of SPIE 1262 (1990)

7. M. Templeton, K. Phan, E. Sum, Proc. of SPIE, Vol. 1086 (1989)

8. C.M. Garza et. al., "Manufacturability Issues of the DESIRE Process", Proceedings of SPIE 1086(1989)

9. Thomas Rucker, Chin Ting (Intel Corp., Santa Clara) et. al., "Processing Considerations for Dry Etching Silylated Photoresist"



Silylation Reaction of the novolak resin and HMDS

Fig1.

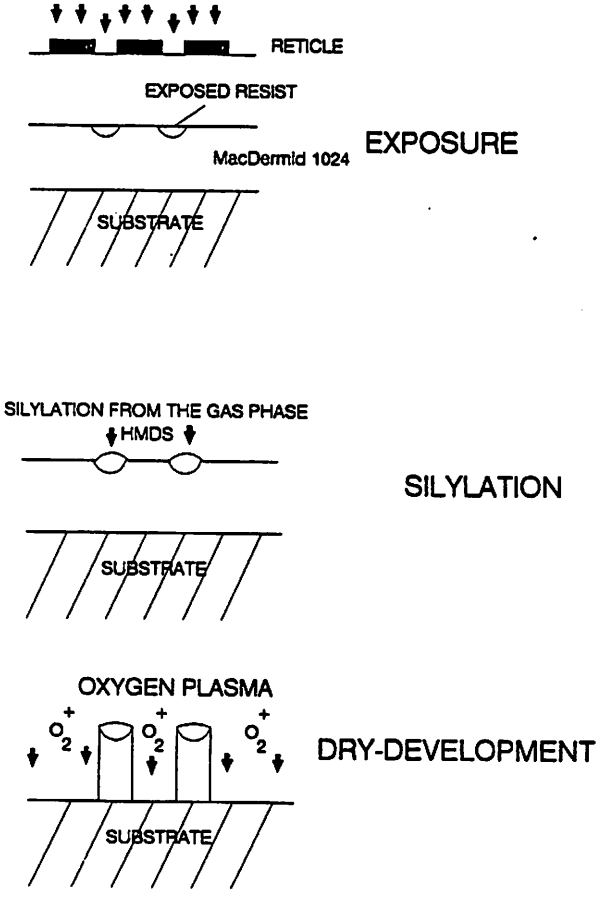


figure 2

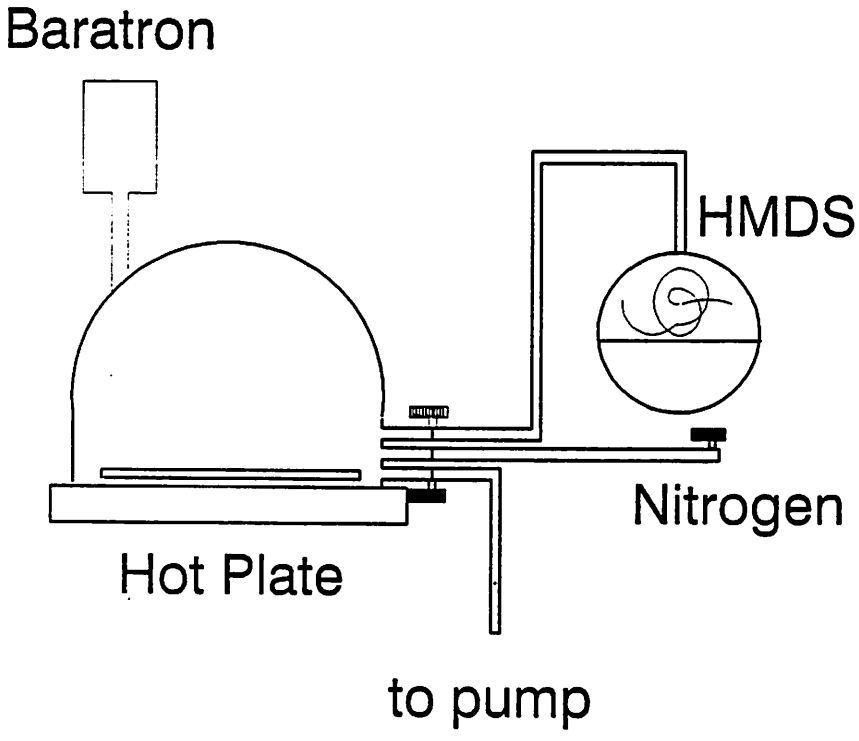


figure 3

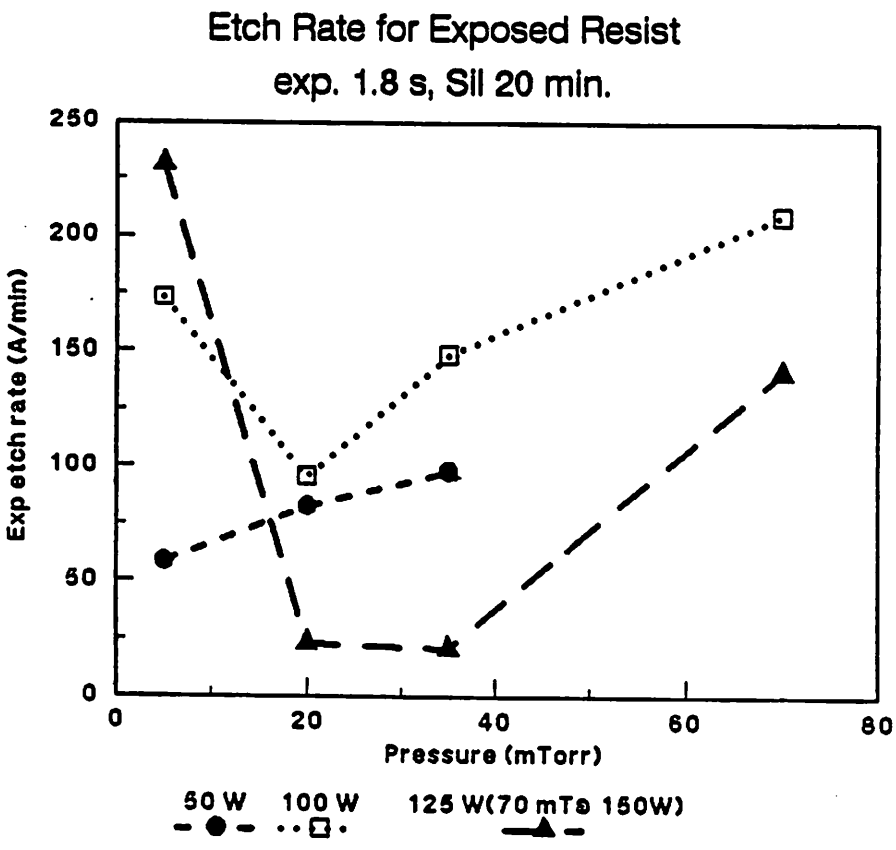


figure 4.

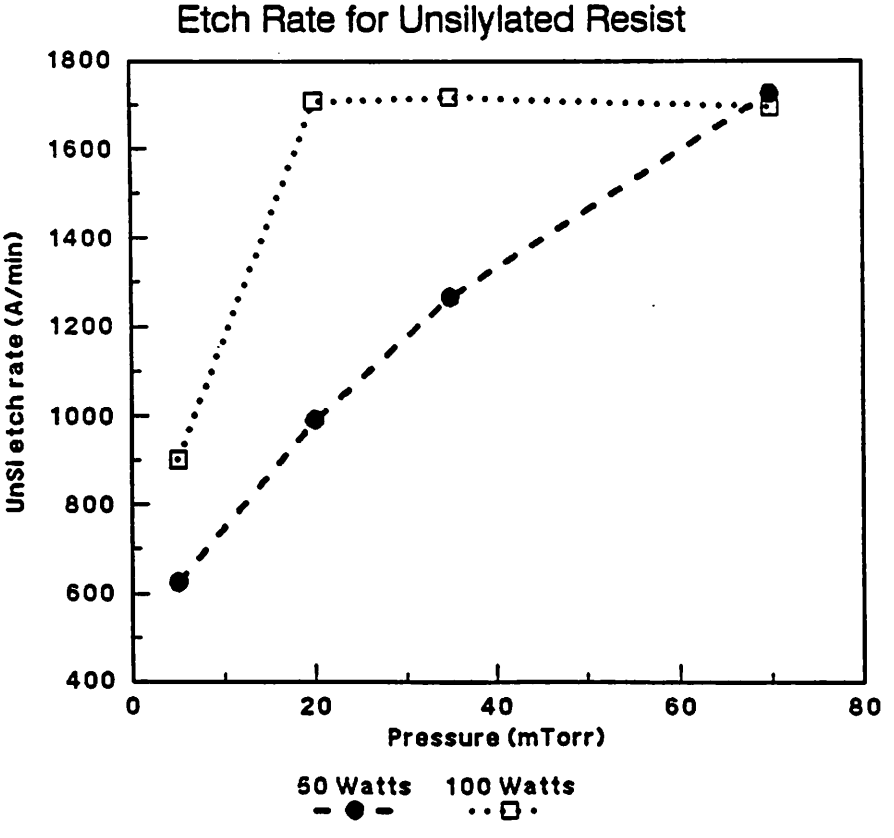


figure 5

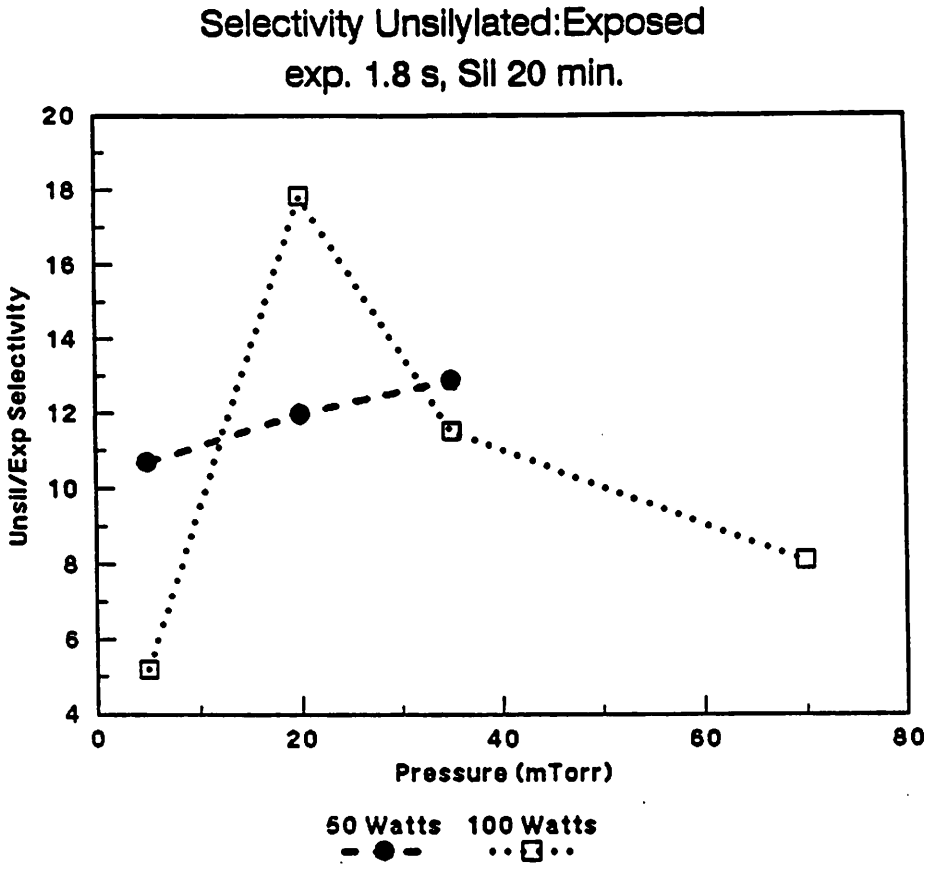


figure 6

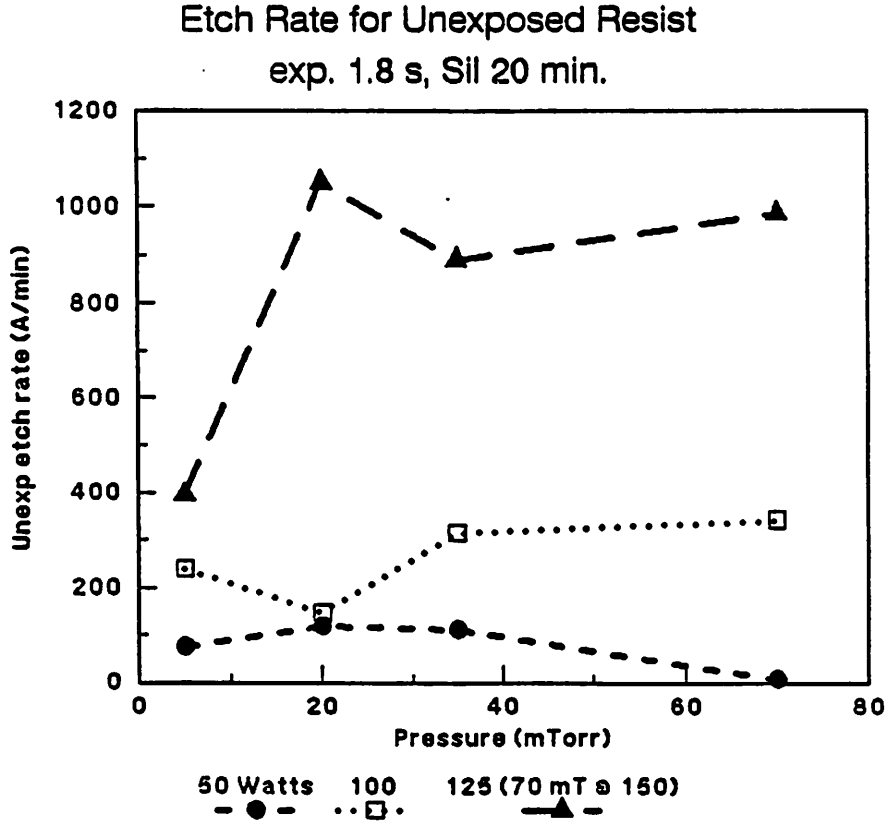


figure 7

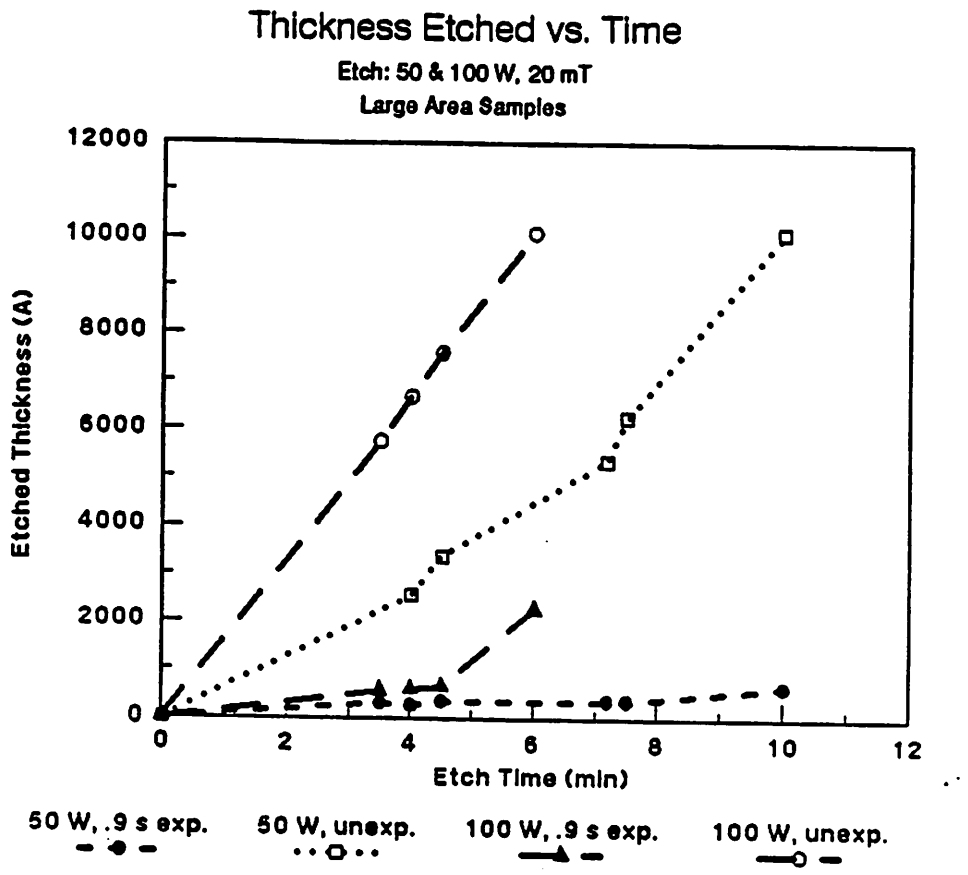


figure 8

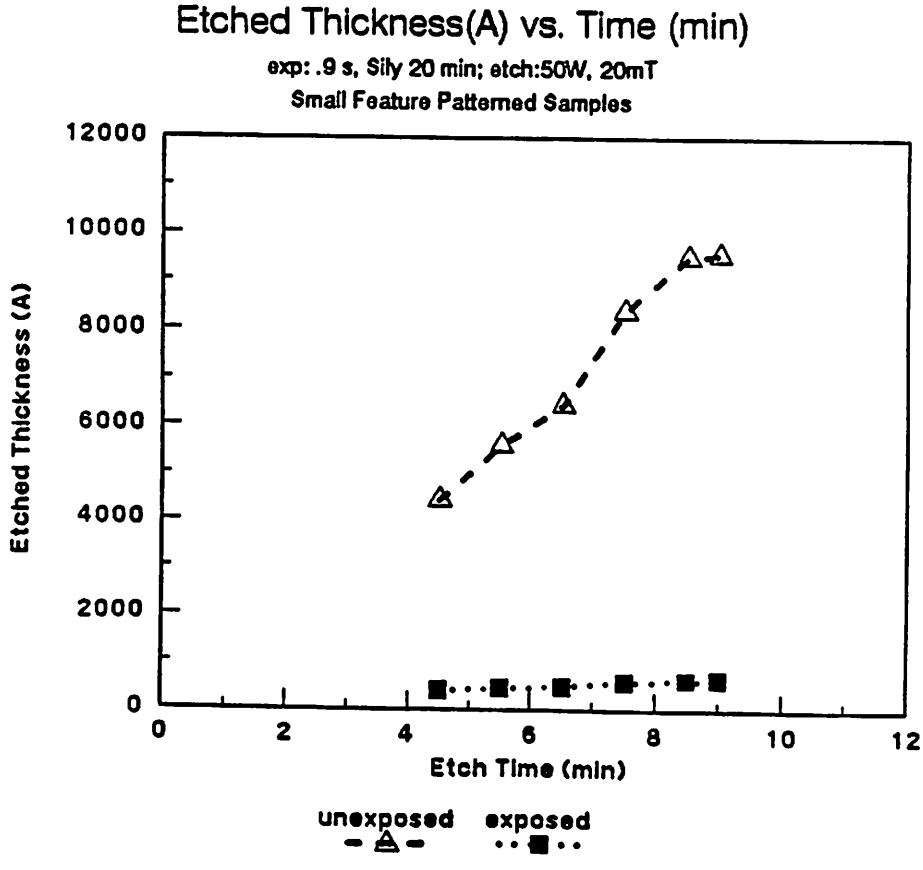


figure 9

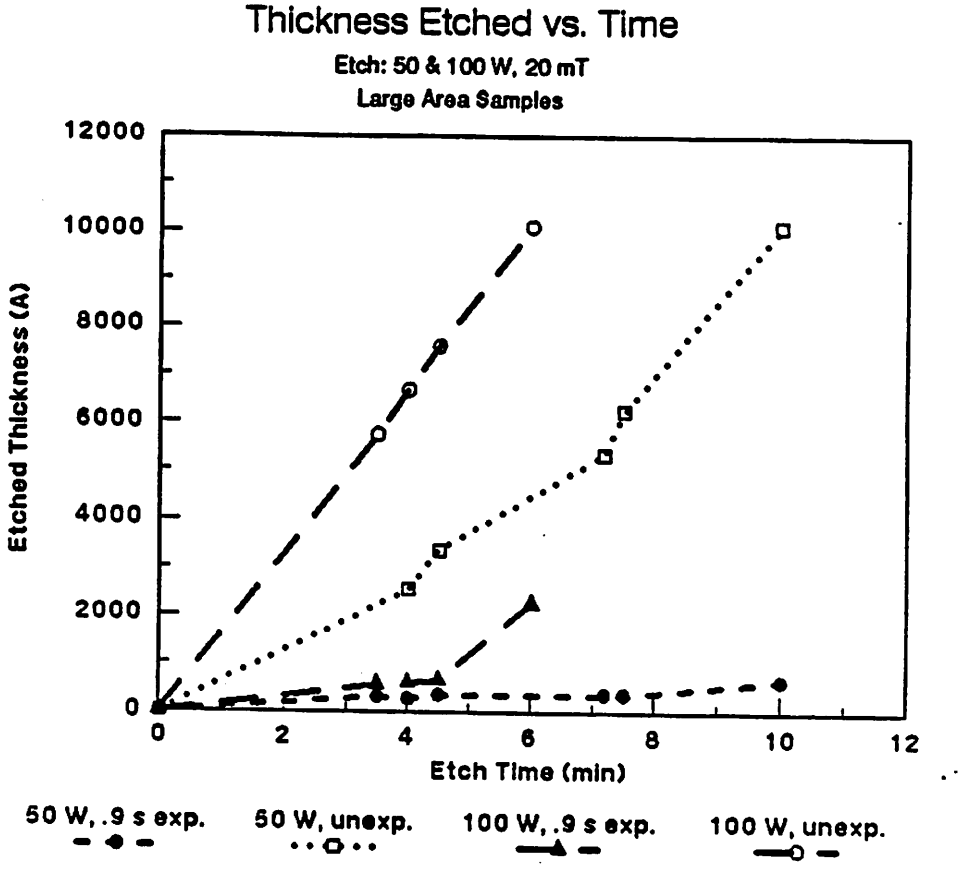


figure 8

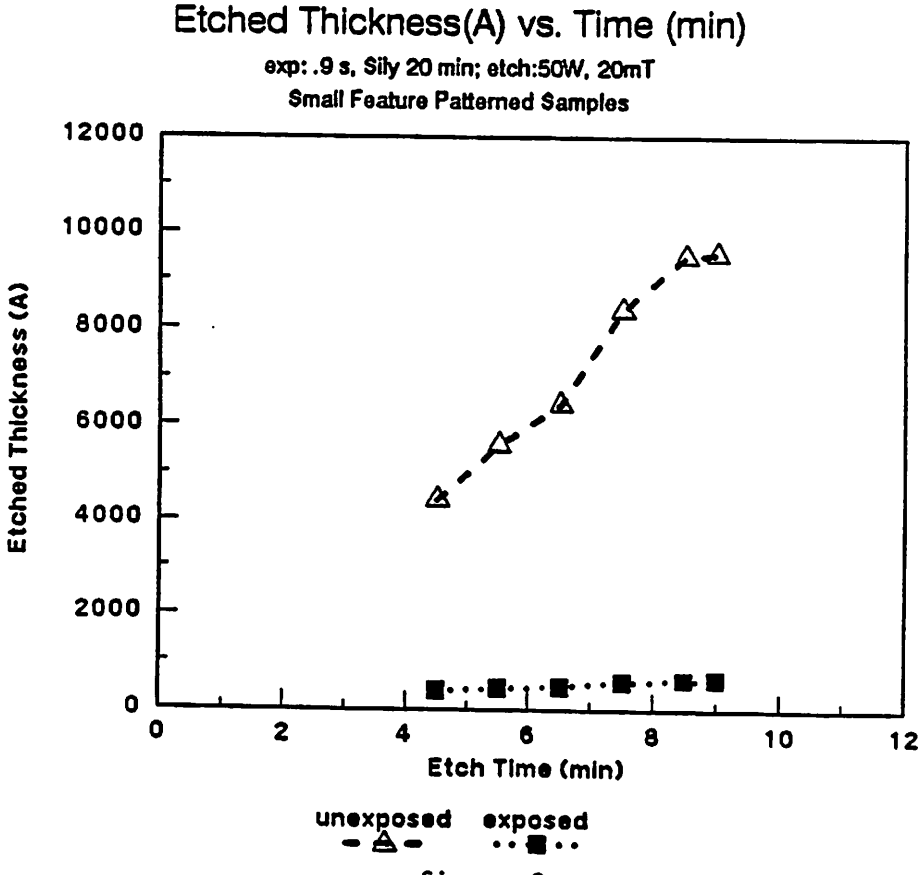
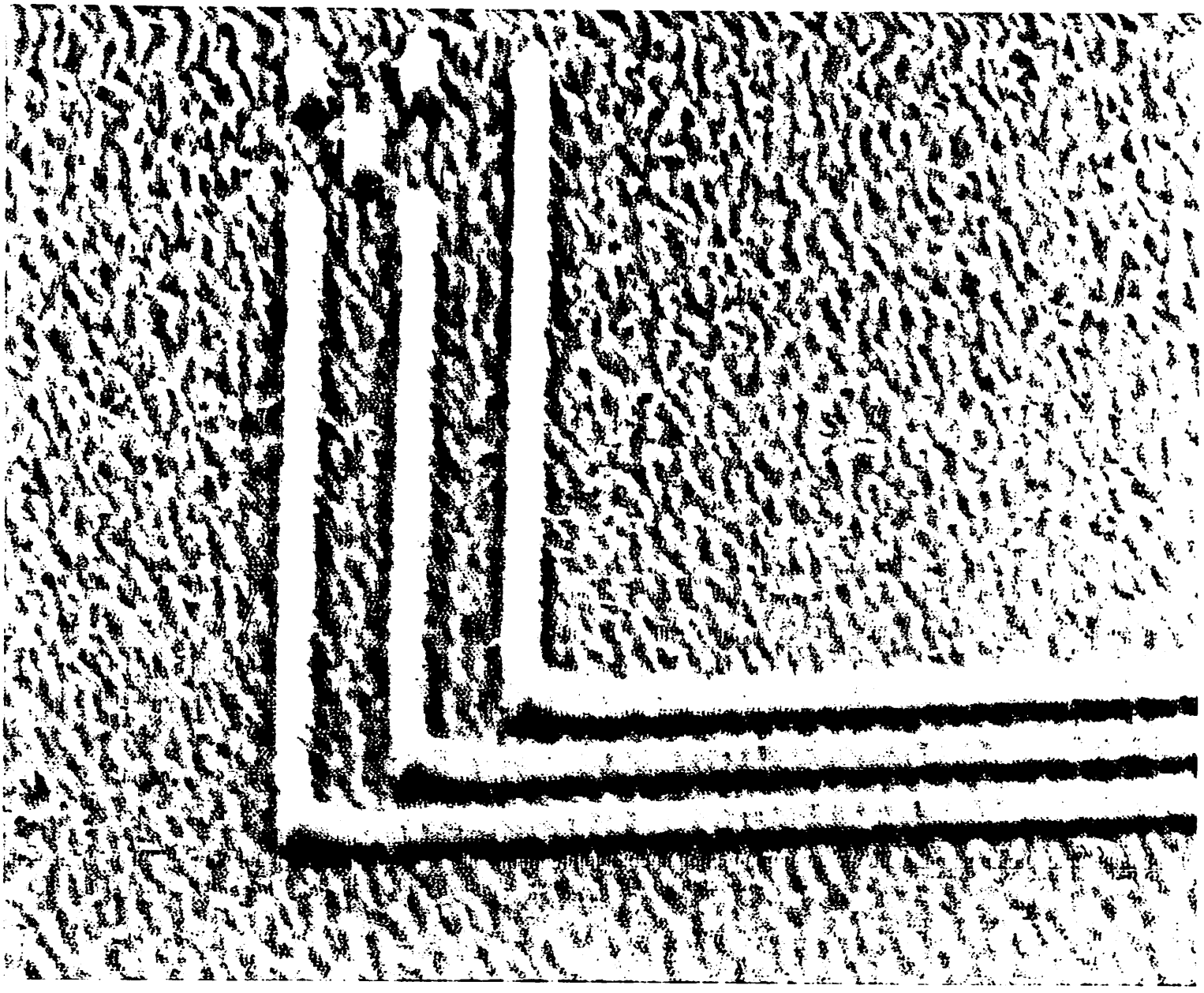


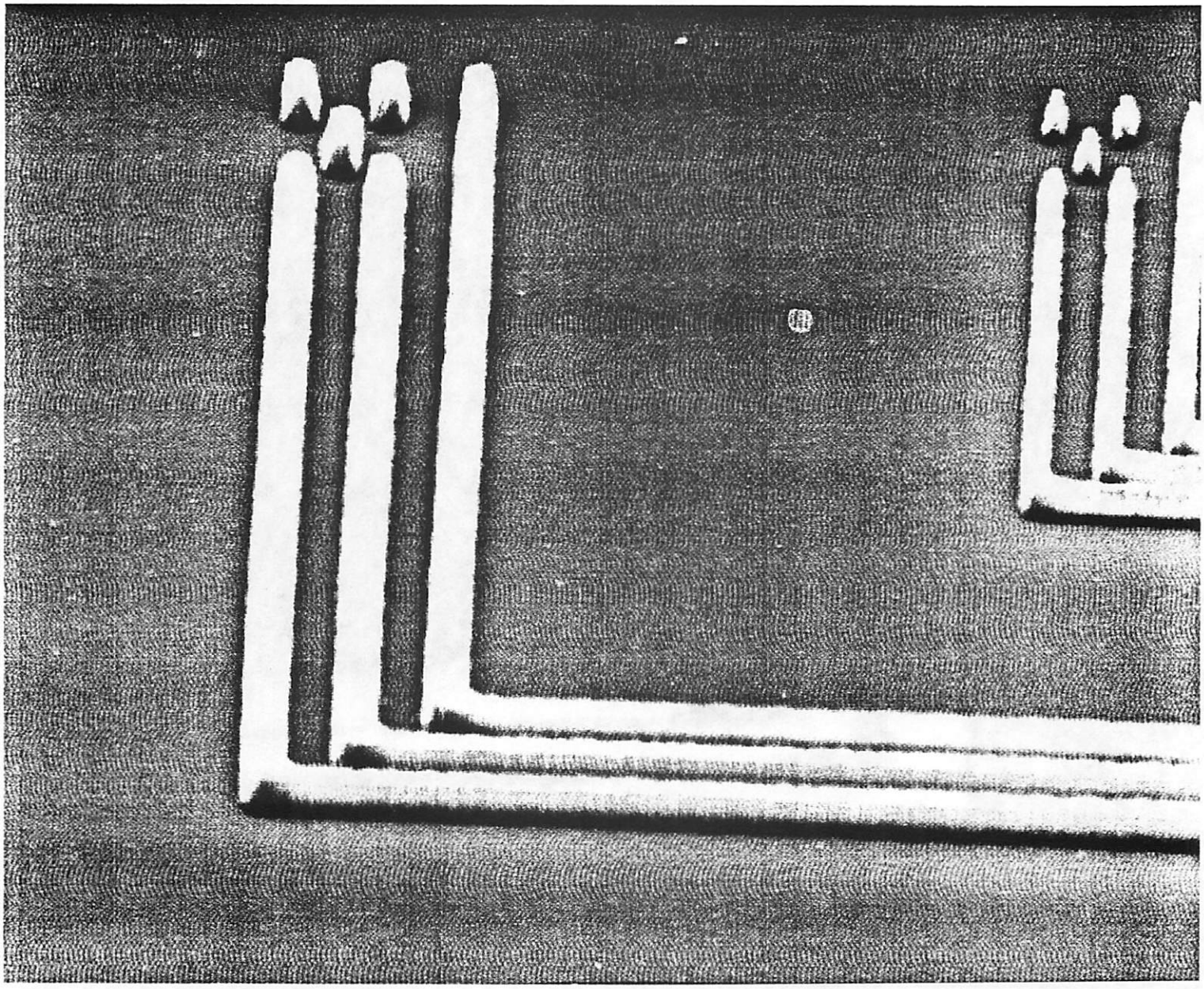
figure 9



Exposure	silylation time	etch time	line width
1.35 sec	10 min	12 min	0.8 μ

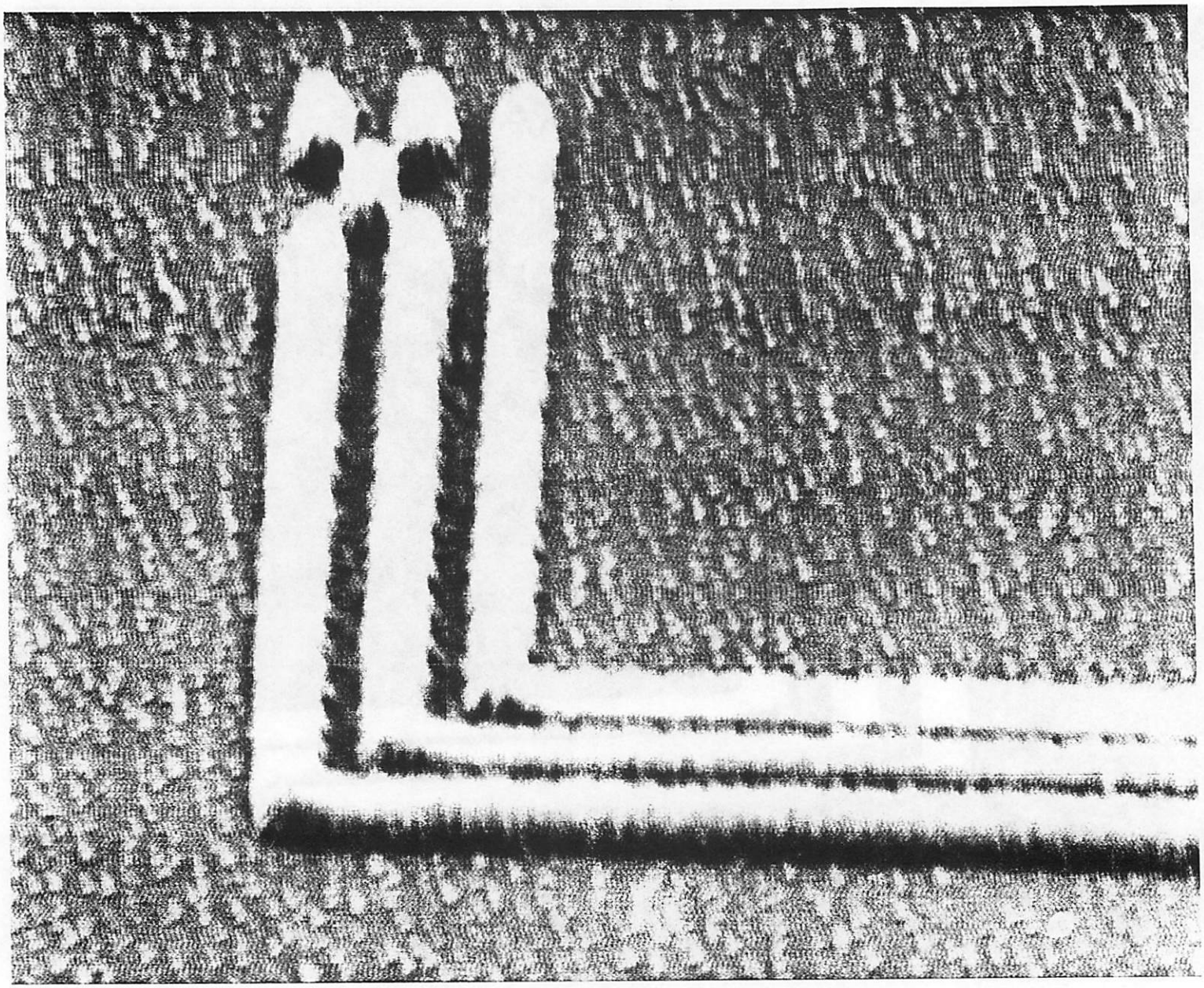
A-2

figure 10



Exposure	silylation time	etch time	HF 10:1	line width
1.35 sec	20 min	13 min	30 sec	1.0 μ

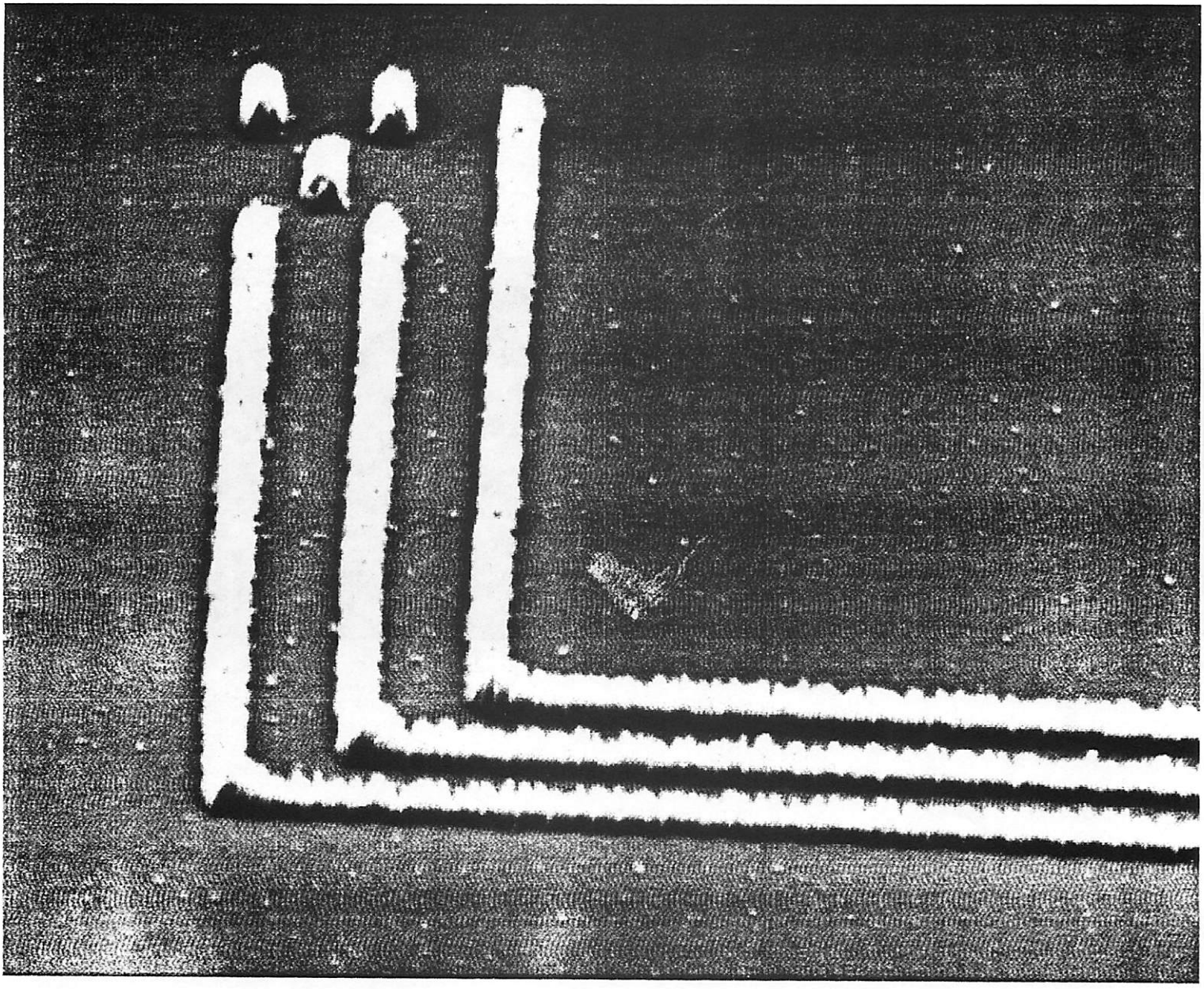
figure 11



Exposure	silylation time	etch time	HF 10:1	line width
0.9 sec	20 min	12 min	30 sec	0.8 μ

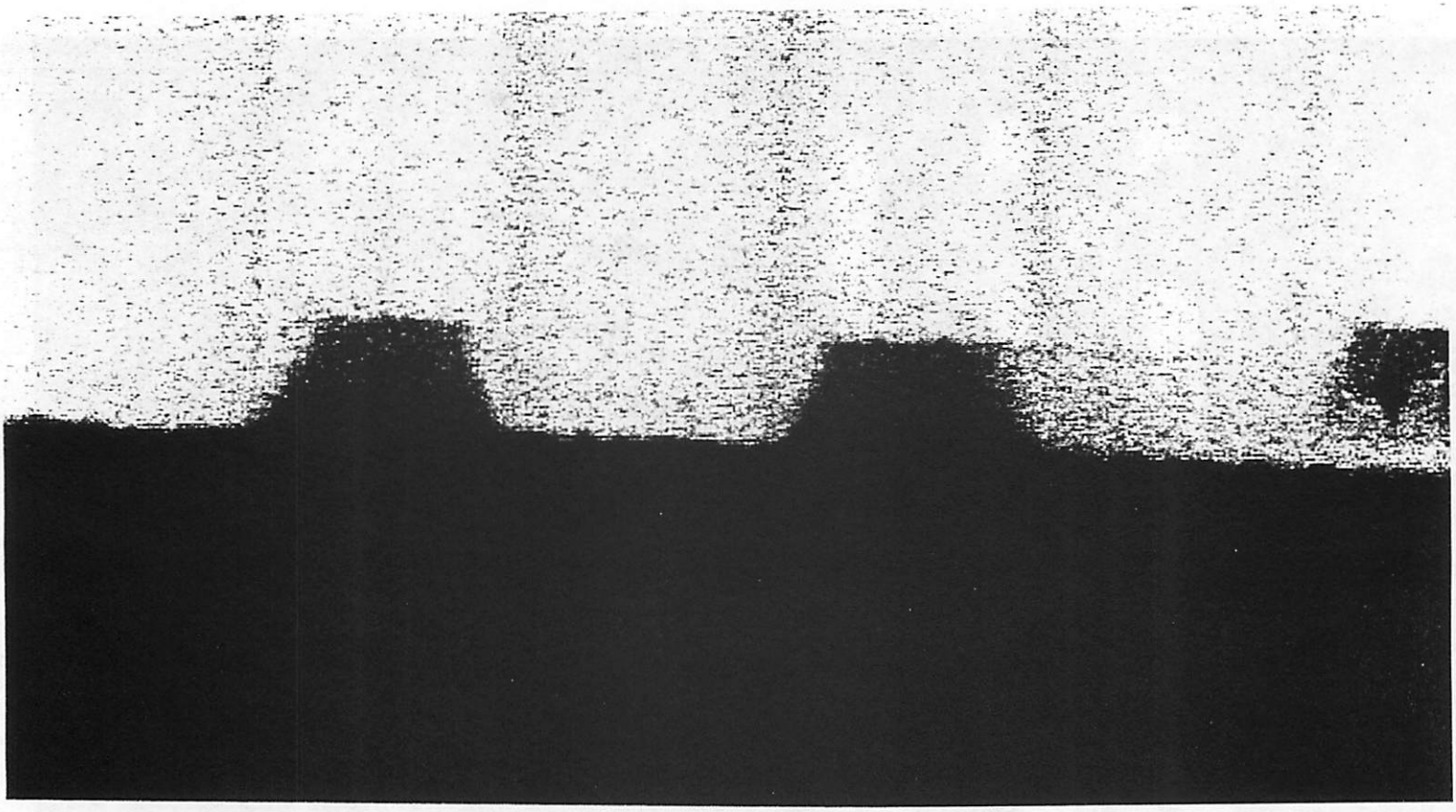
C-1

figure 12



Exposure	silylation time	etch time	HF 10:1	line width
1.35 sec	10 min	12 min	30 sec	0.8 μ

figure 13



Exposure	silylation time	etch time	HF 10:1	line width
1.35 sec	10 min	12 min	30 sec	2.0 μ

A-5

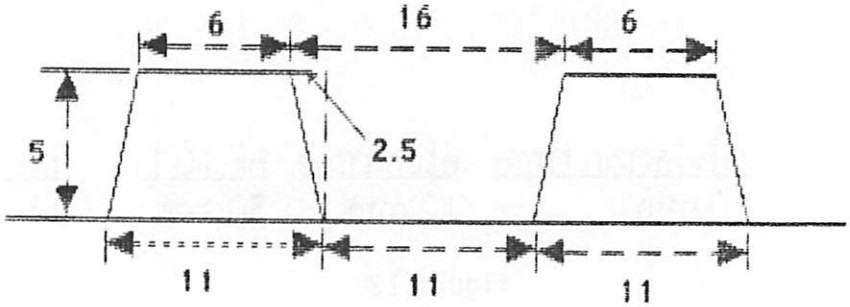
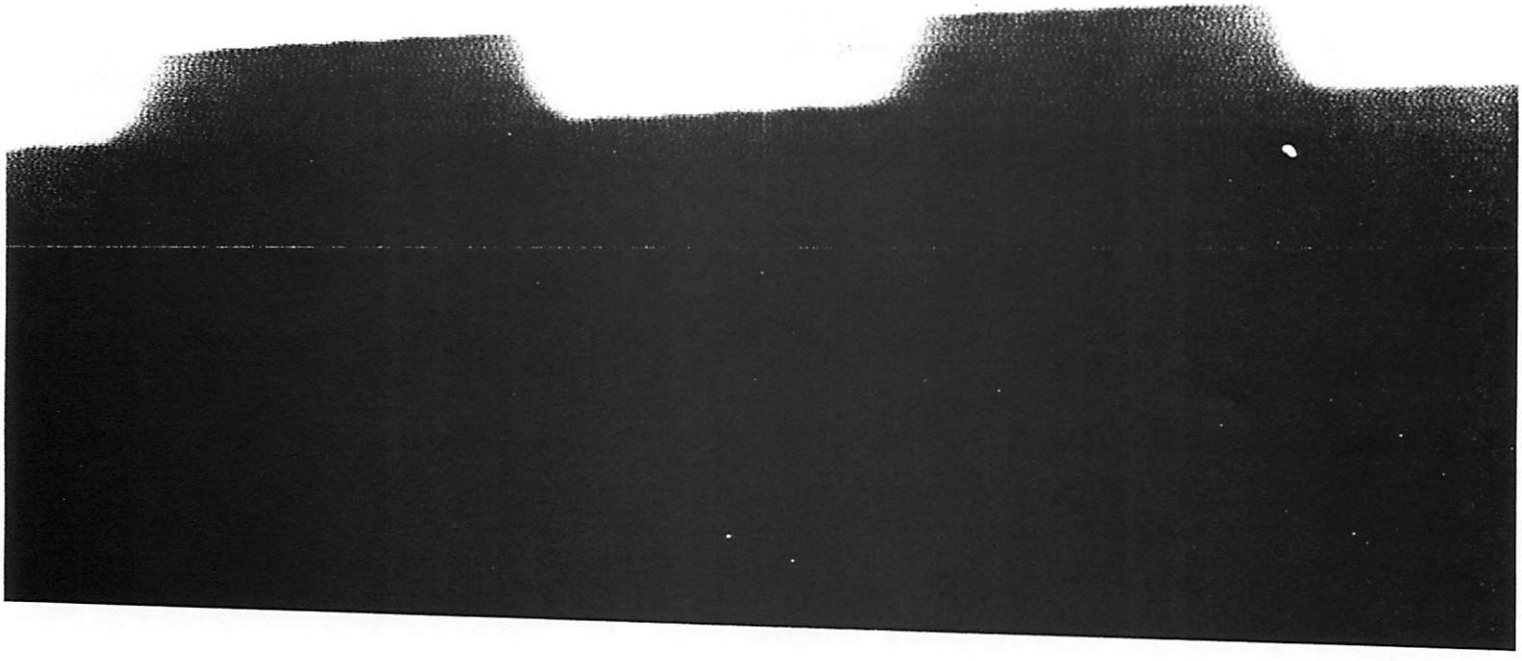


figure 14



Exposure	Silylation Time	Etch Time	Line Width
0.9 sec	20 min.	12 min.	3.0 u

figure 15

Selective Tungsten Etching

Heng-Yuan (Charles) Hsu
Hsi-Jen (James) Yeh

EECS-290N PROJECT REPORT

Prof. A. R. Neureuther
Prof. M. A. Lieberman
Prof. W. G. Oldham

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, CA 94720

Selective Tungsten Etching

by

Heng-Yuan (Charles) Hsu
Hsi-Jen (James) Yeh

ABSTRACT

Recently, chemical-vapor deposited (CVD) tungsten has attracted considerable attention as a suitable material for building microstructures by surface micromachining technology. CVD tungsten has also been considered as the interconnect material for developing a level-independent-interconnect, high density, and high speed Ultra-Large-Scale-Integration (ULSI) process and structures [1, 2]. In addition, as microsensor technology advances, CVD tungsten has increasingly been used as the interconnect material for developing an integrated CMOS/microstructure process [3]. However, in all these applications the ability to anisotropically and selectively etch tungsten is the key to process acceptance and adaptability to existing process flows. The ability to etch tungsten anisotropically is also crucial in achieving high density (in submicron range) tungsten interconnect structures.

An anisotropic reactive ion etch (RIE) process using SF_6 and existing UC Berkeley Microfabrication Laboratory (Microlab) equipment has been developed to etch tungsten. The process is applicable for resist mask and possibly for inorganic-metal mask as well. The best selectivity over photoresist obtained is approximately 5:1, and anisotropic etch (vertical sidewalls) has also been acquired. The effect of applied R.F. power and gas pressure on etch anisotropy, selectivity over resist, and etch rate was investigated. Etch characteristics of other etching gasses (CF_4, Cl_2, and $CHCl_3$) were also studied.

MAJOR ACCOMPLISHMENTS

- A UC Berkeley Microlab compatible tungsten-etching process has been developed in LAM 3 Aluminum etcher.
- Did not observe cross contamination or polymerization problems from mixing Cl and F chemistry.
- Achieved outstanding selectivity over photoresist with SF_6 - 5:1.
- Achieved near-perfect vertical sidewalls (anisotropic etch) with SF_6.
- Performed experimental tungsten etching using CF_4, Cl_2, and $CHCl_3$.

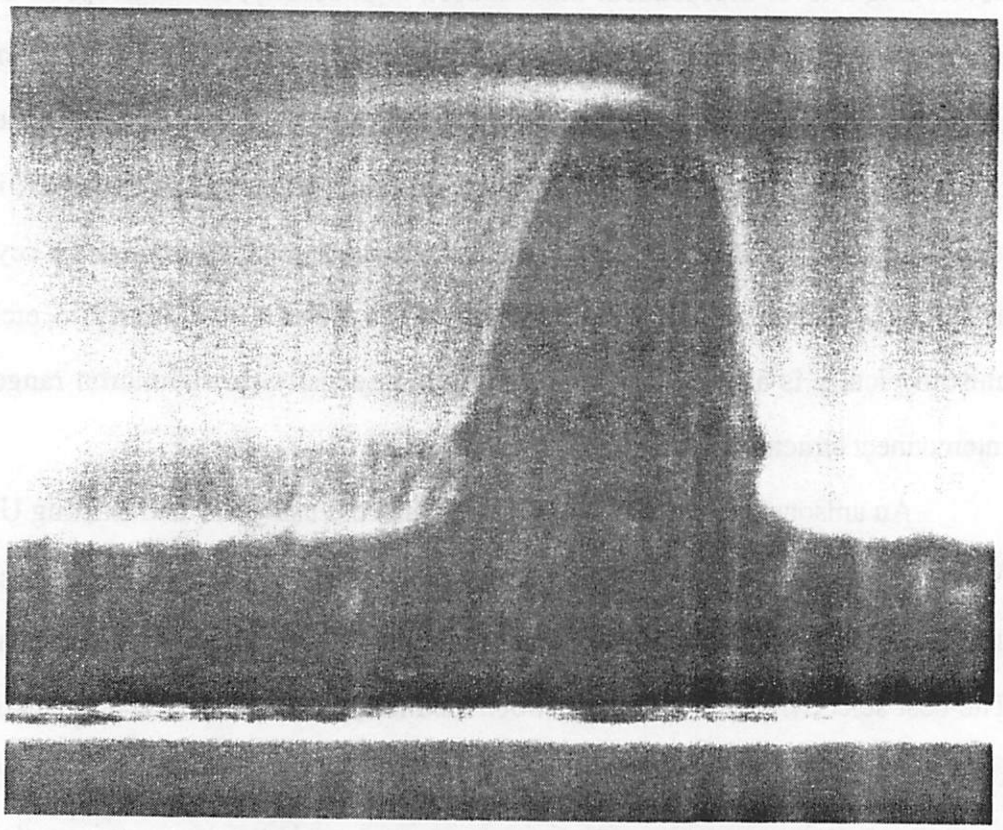


Fig. 1 SEM of vertical W sidewalls (75 W, 50 mT, SF_6 20 sccm)

STATEMENT OF PROBLEM AND KEY ISSUES

For the purposed tungsten-micromachining process, it is necessary to etch approximately 2 μm-thick tungsten films. The tungsten film is used to substitute for the structural polycrystalline silicon (polysilicon) to build tungsten microstructures such as micromotors and interdigitated-finger (comb) drives [4, 5]. Anisotropic etch (to form vertical sidewalls) is essential to the successful fabrication of the comb drives because the drive efficiency of the electrostatic comb drives will be degraded dramatically if the sidewalls of the comb fingers that form capacitive couplings are not parallel (Fig. 2) [6].

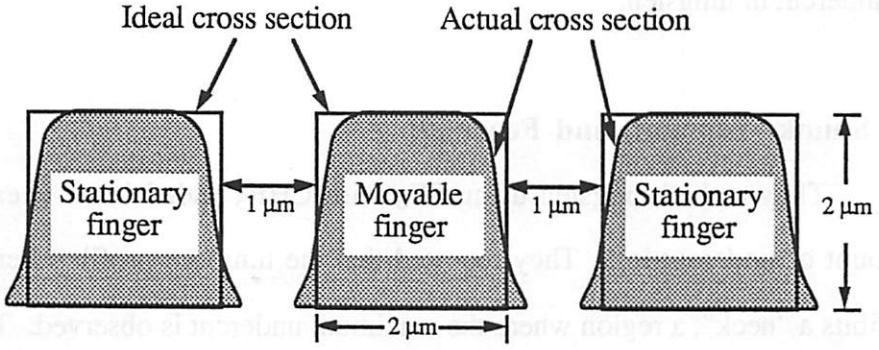


Fig. 2 Cross section of the comb fingers as a result of nonideal plasma etching.

There is an "almost-standard" 2 μm-thick photoresist process in the Microlab; that is twice the thickness of the "standard" 1 μm-thick photoresist process. The photoresist is spun on the wafer with thickness ranging form 1.9 to 2.0 μm. The resist thickness decreases to about 1.6-1.8 μm after the 120°C hardbake process. A selectivity of at least 1.25:1 over the photoresist mask layer is then required for etching a 2 μm-thick tungsten film. Using thicker photoresist poses other problems, and the most significant one is that the focus of the GCA wafer stepper degrades rapidly as the photoresist thickness increases. Developing this thick resist has problems with areas that may not be completely cleared due to the degraded focus of the stepper. Given the linewidth requirement (1 μm resolution), photoresist films thicker than 2 μm should be avoided.

SUMMARY OF PREVIOUS WORK IN LITERATURE

1. Tandon and Jones [1]

They etched tungsten using SF_6 and CF_4 etch gasses and varying the applied R.F. power, pressure, and gas flow rate. They have obtained data for tungsten and resist etch rates and the degree of anisotropy for their tungsten etching process. They concluded that by using SF_6, they could get the best selectivity (2:1) over photoresist with some undercut unless they operated at a very low pressure (5 mTorr). With CF_4, they were able to obtain an anisotropic etch at 5-10 mTorr pressure with very poor selectivity (1:2) over resist. Their study also suggested that the presence of carbon in the RIE system delayed the onset of undercut in tungsten.

2. Shunk, Tennant, and Feuer [9]

They etched tungsten using SF_6 with CHF_3 added in an attempt to reduce the amount of undercutting. They reported that the tungsten profile after etched with SF_6 exhibits a "neck", a region where the maximum undercut is observed. They also reported that different concentration of CHF_3 in SF_6 changes the amount and position of the neck.

3. Fischl and Hess [10]

They etched tungsten and tungsten silicide in Cl_2 and Cl_2/BCl_3 plasmas and obtained etch rate data from the experiments. They were able to obtain a tungsten etch rate of approximately 500 Å/min..

4. Beatty and Kessler [11]

They developed a process using a chromium mask for tungsten etch. Chromium (600 Å) was first patterned with a PMMA photoresist mask and an O_2/Cl_2 plasma; it was then used subsequently as a mask layer for a 2 μm-thick tungsten film. Using a CF_4/O_2 plasma, they were able to etch 0.9 μm wide lines.

APPROACH

It has been shown that CF_4 plasma etches tungsten anisotropically with very poor selectivity over photoresist, while SF_6 etches tungsten selectively but with a "neck" [7]. By mixing CF_4 and SF_6 we may be able to optimize both anisotropy and selectivity over photoresist in tungsten etching. This would involve flowing both gasses (CF_4 and SF_6) into the etching chamber in different gas compositions. Shunk, et al. have investigated a similar process with a gas mixture of SF_6 and CHF_3 [7].

A matrix of characterization experiments is then designed with variations in gas composition, substrate temperature, applied R.F. power, and chamber pressure. However, the plasma etcher to be characterized LAM 3 has very strict constraints on varying substrate temperature and obtaining low (5-10 mTorr) chamber pressure. In addition, the initial screening experiment shows that CF_4 plasma is not acceptable for tungsten etching under a high substrate temperature (60°C) and chamber pressure (50 mTorr) process condition in LAM 3. Thus we would only use SF_6 plasma for tungsten etching and reduce the etching characterization matrix with variations only in applied R.F. power and chamber pressure. A full-factorial experiment based on the characterization matrix is conducted.

Other gasses can also be tried in LAM 3 to etch tungsten. An expriment is designed to etch tungsten in a N_2+Cl_2+$CHCl_3$ plasma [2] based on the existing aluminum etch. The experimental result allows us to evaluate the effect of chlorine chemistry in tungsten etching. Another experiment has also been designed to examine the use of an inorganic-metal mask layer, where a 1000 Å-thick aluminum film is used as a mask layer over a 1 μm-thick tungsten film. Aluminum is not etched by fluorine because aluminum fluoride is not volatile; aluminum is only eroded slowly by physical etching in SF_6 plasma. However, an aluminum mask may not provide the carbon atoms necessary to form sidewall polymers for an anisotropic etch.

DESCRIPTION OF EXPERIMENT

1. Plasma etchers

The Berkeley Microlab has two plasma etchers that may be used to etch tungsten, namely, Plasma Therm PK-12 PE/RIE system and Lam Research Corp. AutoEtch 690 plasma etching system. A detailed description for each plasma etcher is presented in the following sections.

1.1. Plasma Therm PK-12 PE/RIE system

This plasma etcher, also called p-therm, has two etching modes: plasma etching (PE) and reactive ion etching (RIE). When the power is applied to the upper electrode, the system operates in plasma-assisted etching mode, in which good selectivity and a satisfactory degree of anisotropy can be obtained. When power is applied to the lower electrode, the system operates in reactive ion etching mode, which is characterized by vertical sidewalls, but less selectivity [3].

This system is a general-purpose plasma etcher in which a variety of materials are etched because of its control flexibility. There are five control parameters in the system: gas composition, gas flow rates, power, chamber pressure, and substrate temperature. Tungsten has been etched in p-therm by Weijie Yun [3] for his research project. Our major concern in using p-therm is that a host of very different materials are being etched in it constantly. Thus it is considered a "dirty" system. Different materials and gasses that were etched in p-therm prior to tungsten etching may have unpredictable effect on the outcome of our experiment. In addition, p-therm is not a user-friendly system to use. The setup time prior to actual etching can be as long as 30 minutes per wafer. In conclusion, if a process were developed in p-therm, it would not be considered as a "standard" Microlab process; tungsten would be just another "dirty" material that we etch in this general-purpose etcher.

1.2. Lam Research Corp. AutoEtch 690 plasma etching system

The AutoEtch 690, also called LAM 3, is a fully automated, single wafer, cassette-to-cassette plasma etching system. LAM 3 is configured in RIE mode; the lower electrode assembly (Fig. 3) contains a cathode which is electrically connected to the R.F. generator by the R.F. match assembly. The upper electrode assembly consists of the anode. The process gasses flow through a gas manifold, through the shower head hole pattern of the anode, and into the process chamber directly above the wafer. LAM 3 is equipped with a solid state, water cooled, 650 W ENI R.F. generator operating at 13.56 MHz [7]. The system's lower electrode is maintained at 60°C by a heated closed-loop deionized water recirculating system. The chamber vacuum manifold and chamber assemblies are heated and maintained at approximately 60-65°C to prevent condensation of etch by-products.

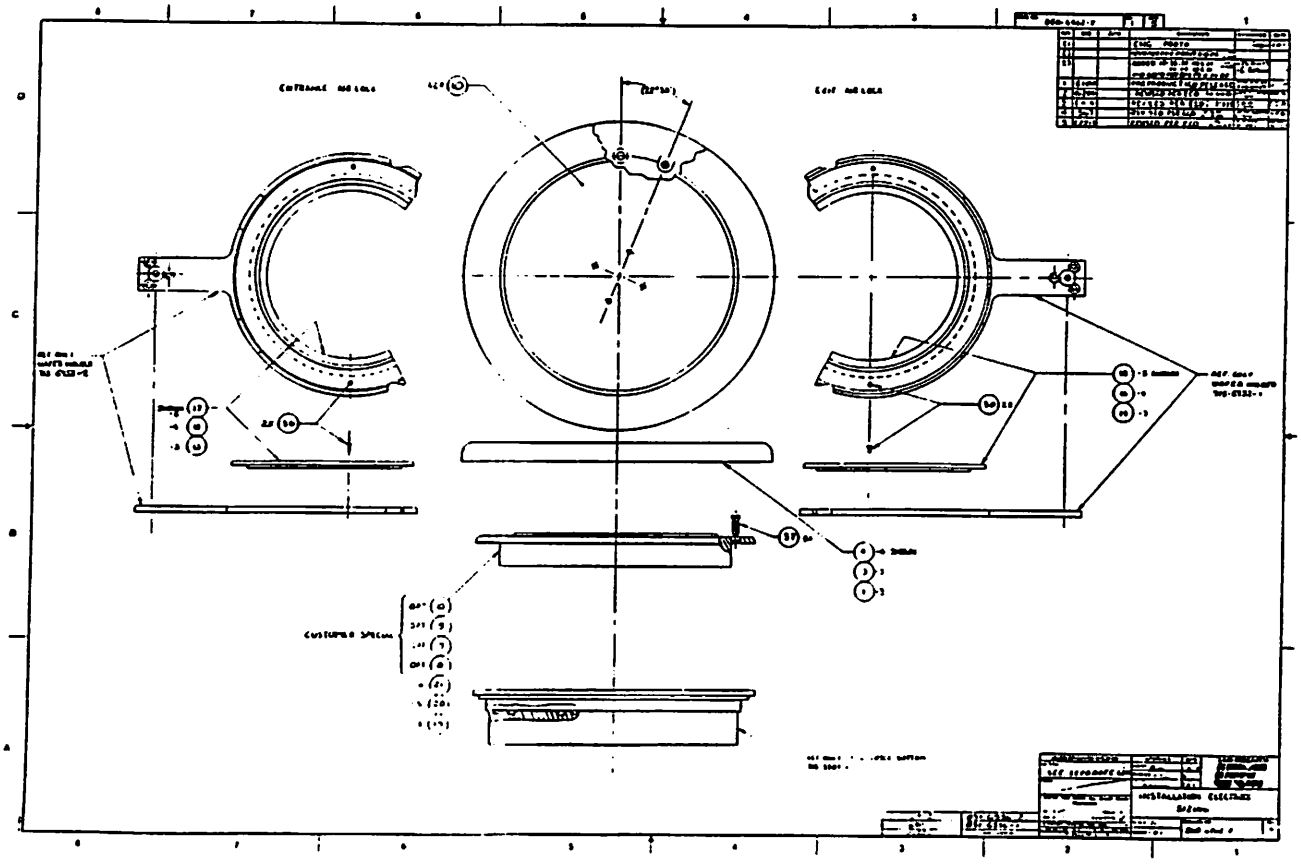


Fig. 3 Schematic of LAM 3 lower electrode assembly (source: AutoEtch 690 manual). Electrode dimension: 8" in diameter.

This etcher is originally configured to etch aluminum using chlorine chemistry. Aluminum is etched first by BCl_3 to remove the native Al_2O_3 layer, then it is etched in an N_2+Cl_2+$CHCl_3$ plasma. Of these three gasses, Cl_2 is the main etch gas for aluminum, N_2 is used as a diluent, and $CHCl_3$ is used to form the sidewall polymer which protects the sidewalls during etch leading to a vertical profile. These etching processes take place in the REACTOR chamber, and the etched wafer is transferred to the AIRLOCK chamber afterwards. Once the wafer is in the AIRLOCK chamber, a CF_4+O_2 plasma is used to passivate the exposed aluminum surface by exchanging chlorine for fluorine in the residual $AlCl_3$ which exists after etching. This is a very important process because $AlCl_3$ reacts with moisture in the air to form hydrochloric acid instantaneously which will then attack and corrode the aluminum film [8].

One of the major obstacles in characterizing tungsten etching in LAM 3 is to convince the Microlab staff that the purposed experiment will not destroy the dedicated aluminum etching. After much negotiation, the characterization experiment is allowed in LAM 3 on a trial basis provided that no cross contamination can be observed in the existing aluminum etch.

For tungsten etching, a mass flow controller (MFC) is installed for the SF_6 gas line. There already exists a SF_6 gas line in LAM 3 but is never used since SF_6 is not required for aluminum etching. A CF_4 gas line is spliced into the existing $CHCl_3$ gas line via a three-way valve. This new gas line shares the same MFC with the $CHCl_3$ gas line, and a correction factor of 1.2 must be applied for CF_4 flow since the MFC is calibrated for $CHCl_3$.

From the summary of references, we notice that tungsten etching has always been done at relatively low pressures, such as 5, 10, or 50 mTorr. We are unable to achieve such low pressure in LAM 3 due to a low-efficiency vacuum system. The base pressure for LAM 3 is about 4 mTorr, and we observe a 20 mTorr pressure reading in the chamber

even with a gas flow rate of only 1 sccm. After several trials, we have concluded that the lowest practical operating pressure is approximately 50 mTorr with a gas flow rate of 20 sccm. According to the references, this high-pressure operating condition should result in completely isotropic etch. Also, the 60°C substrate temperature can be detrimental to the selectivity over photoresist, as the photoresist might erode faster with increasing temperature by the Arrhenius dependence: $R_0 \exp[-E_A/kT]$, where E_A is the activation energy.

Since we are mixing fluorine and chlorine chemistries, contamination and polymerization in the chamber can pose an enormous problem. Fluorine mixed with chlorine can form polymers that would drastically change the etching characteristic of the plasma etcher. Since AlF_3 is not volatile, introducing fluorine into the REACTOR chamber may result in a destructive decrease in aluminum etch rate. On the other hand, mixing tungsten and aluminum in the same etcher presents little danger in contaminating the chamber since both metals are considered "dirty" and sputtered in the same Circuit Processing Apparatus (CPA) sputtering system.

In order to avoid any possible cross contamination, a pre-tungsten-etch and a post-tungsten-etch cleaning procedures are conducted each time the tungsten etch is performed. The pre-tungsten-etch cleaning step is necessary to fill the chamber with fluorine and remove any residual chlorine. The post-tungsten-etch cleaning step is essential to burn away the fluorine polymers and restore chlorine chemistry in the chamber. Furthermore, we etch an aluminum wafer before and after etching tungsten in LAM 3 to ascertain that the standard aluminum etching is not ruined.

2. Experimental plan

After depositing tungsten films on 4" silicon wafers, a series of screening experiments will be carried out to test the suitability of various gas mixture (SF_6 and CF_4)/applied R.F. power/chamber pressure combinations and to find an operating point

around which we can perform a full-factorial experiment. Moreover, we would like to make sure that tungsten etching in LAM 3 is feasible from these screening experiments. Since both the chamber assemblies and the lower electrode (substrate) are fixed at 60°C, the temperature variable is not included in the full-factorial matrix.

The intrinsic high operating pressure of LAM 3 will not allow us to work with chamber pressures below 50 mTorr. Thus the chamber pressure will vary from a low value at 50 mTorr to a high value at 150 mTorr with the center point at 100 mTorr. The low R.F. power setting is chosen at 75 W, giving a power density of 0.23 W/cm^2 when divided by the electrode area, and the high power setting is chosen at 175 W (0.54 W/cm^2) with the center point at 125 W (0.39 W/cm^2). This gives four experimental operating points with the combinations of the high and low settings in the applied R.F. power and the chamber pressure. Two more etches will be done at the center operating point to complete the full-factorial experiment. The complete set of full-factorial experiment will be performed twice, with different tungsten film thickness, to see if there is any change in tungsten etching characteristics due to different film thickness.

The tungsten etch rate and the photoresist etch rate will be monitored both at the center die and the outermost die on the wafer to obtain data in characterizing the etching uniformity of LAM 3. In addition, the etching selectivity over photoresist can also be derived from the etch rate data. Scanning electron microscope (SEM) will be employed to examine the degree of etching anisotropy.

Several experiments, which involve chlorine plasma and aluminum mask, will also be executed and characterized after the completion of the full-factorial experiment.

3. Process flow

The actual process flow to carry out the experimental plan will be described in detail in the following sections.

3.1. Wafer preparation

We start initially with 4" n-type, <100> orientation, silicon prime wafers. The process begins with depositing a layer of 2 μm-thick low-pressure chemical-vapor deposited (LPCVD) phosphosilicate glass (PSG) on the wafer. This is done in Tylan 12 with the standard recipe SDOLTOE. A layer of 1000 Å-thick LPCVD polysilicon is then deposited on top of the PSG layer in Tylan 11 with the standard recipe SDOPOLYG. This polysilicon layer is used as an adhesion layer, or seed layer, for tungsten since tungsten has problems sticking to the oxide layer.

Tungsten is then sputter deposited on the wafer using CPA 9900 sputtering system. The system consists of a main sputtering chamber and two load locks at each end, and has three stationary targets in the process chamber, titanium (Ti), aluminum (Al), and tungsten (W). Wafers are mounted on the stainless steel pallets in the right load lock, and the pallets are passed through the main sputtering chamber under the tungsten target on a constant-speed (controlled by users) conveyer. A planar magnetron tungsten target with purity of 99.95% is used for all the experiments. Once wafers are loaded into the system, the chamber and the load lock are pumped for about 30 minutes until the chamber pressure is about $5x10^{-7}$ Torr. To clean the chamber, a 15-minute pre-deposition chamber sputtering is completed in an argon plasma at 20 mTorr argon pressure and 2 kW of sputtering power. Wafers are then passed through the main sputtering chamber under the tungsten target for tungsten deposition. Since the track speed is linearly related to film thickness, a track speed of 10 cm/min. results in a 3000 Å-thick sputtered-on tungsten film, and a track speed of 3 cm/min. gives approximately 1 μm-thick of sputtered-on tungsten.

After tungsten has been deposited on the wafers, the Microlab standard photolithography using Olin-Hunt I-line photoresist will be used to form the masking layer for tungsten patterning. A "standard" 1 μm-thick photoresist is used for the 3000 Å-thick tungsten films, and a "double-thick" 2 μm-thick photoresist is used for the 1 μm-thick ones.

3.2. LAM 3 cleaning and tungsten etching recipes

As described in Section 1.2, a pre-tungsten-etch and post-tungsten-etch cleaning processes must be performed to avoid any possible chlorine-fluorine contamination during tungsten etching. All cleaning and etching processes, in LAM 3 recipe form, will be discussed in the following subsections.

A. Standard Al-etching recipe (monitoring Al etch)

A monitor aluminum etch is performed before and after the actual tungsten etch, and the standard Al-etching recipe is used for this process. The etching steps are described in section 1.2 of this report, and this standard recipe can be found in the Microlab Equipment Manual or on the WAND.

B. Pre-W-etch cleaning recipe

The pre-W-etch cleaning consists of generating an SF_6 plasma at a chamber pressure of 50 mTorr, an applied R.F. power of 250 W, and an SF_6 gas flow rate of 20 sccm for 5 minutes.

Recipe 1. Pre-W-etch cleaning recipe.

[REACTOR] RECIPE			[REACTOR] RECIPE			[REACTOR] RECIPE		
STEP # 01			STEP # 02			STEP # 03		
PRESS	50.	[MTORR]	PRESS	50.	[MTORR]	PRESS	0.	[MTORR]
RF LWR	0.	WATTS	RF LWR	250.	WATTS	RF LWR	0.	WATTS
BCL3	0.	SCCM	BCL3	0.	SCCM	BCL3	0.	SCCM
N2	0.	SCCM	N2	0.	SCCM	N2	0.	SCCM
CL2	0.	SCCM	CL2	0.	SCCM	CL2	0.	SCCM
CHCL3	0.	SCCM	CHCL3	0.	SCCM	CHCL3	0.	SCCM
SF6	20.	SCCM	SF6	20.	SCCM	SF6	0.	SCCM
COMPL [STABILITY OR TIME]			COMPL [TIME]			COMPL [RECIPE]		
WAIT	00:20	MIN:SEC	WAIT	05:00	MIN:SEC	WAIT	00:00	MIN:SEC

C. Tungsten etching recipe

After the pre-W-etch cleaning, the actual tungsten etching is performed in either SF_6 plasma or CF_4 plasma. An SF_6 plasma is used in the following sample recipe.

Recipe 2. Tungsten etching recipe.

[REACTOR] RECIPE			[REACTOR] RECIPE			[REACTOR] RECIPE		
STEP # 01			STEP # 02			STEP # 03		
PRESS	50.	[MTORR]	PRESS	50.	[MTORR]	PRESS	0.	[MTORR]
RF LWR	0.	WATTS	RF LWR	75.	WATTS	RF LWR	0.	WATTS
BCL3	0.	SCCM	BCL3	0.	SCCM	BCL3	0.	SCCM
N2	0.	SCCM	N2	0.	SCCM	N2	0.	SCCM
CL2	0.	SCCM	CL2	0.	SCCM	CL2	0.	SCCM
CHCL3	0.	SCCM	CHCL3	0.	SCCM	CHCL3	0.	SCCM
SF6	20.	SCCM	SF6	20.	SCCM	SF6	0.	SCCM
COMPL	[STABILITY OR TIME]		COMPL	[TIME]		COMPL	[RECIPE]	
WAIT	00:20	MIN:SEC	WAIT	03:00	MIN:SEC	WAIT	00:00	MIN:SEC

D. Post-W-etch cleaning recipe

After all tungsten etches have been completed, a post-W-etch cleaning process must be performed to restore chlorine chemistry in the etcher. The post-W-etch cleaning consists of generating a chlorine plasma in the chamber to burn away any fluorine polymers.

Recipe 3. Post-W-etch cleaning recipe.

[REACTOR] RECIPE			[REACTOR] RECIPE			[REACTOR] RECIPE		
STEP # 01			STEP # 02			STEP # 03		
PRESS	250.	[MTORR]	PRESS	250.	[MTORR]	PRESS	250.	[MTORR]
RF LWR	0.	WATTS	RF LWR	250.	WATTS	RF LWR	250.	WATTS
BCL3	0.	SCCM	BCL3	0.	SCCM	BCL3	0.	SCCM
N2	50.	SCCM	N2	50.	SCCM	N2	50.	SCCM
CL2	0.	SCCM	CL2	0.	SCCM	CL2	20.	SCCM
CHCL3	0.	SCCM	CHCL3	0.	SCCM	CHCL3	0.	SCCM
SF6	0.	SCCM	SF6	0.	SCCM	SF6	0.	SCCM
COMPL	[STABILITY OR TIME]		COMPL	[TIME]		COMPL	[TIME]	
WAIT	00:20	MIN:SEC	WAIT	02:00	MIN:SEC	WAIT	05:00	MIN:SEC

RESULTS AND INTERPRETATION

1. Screening experiment

The initial screening experiment was carried out on 3000 Å-thick tungsten films using pure CF_4 and SF_6 gasses. Etches done using CF_4 gas (at 75 W, 50 mTorr, 60°C, and 20 sccm) gave very low tungsten etch rates, averaging about 193 Å/min., and very poor selectivity over photoresist (~ 1:1.9). The SEM photo (Fig. 4) shows tapered tungsten sidewalls at approximately 45° angle, indicating severe resist erosion during the tungsten etch [9]. Initial SF_6 experiment, under identical etching conditions, resulted in a factor of five increase in etch rate (~ 1113 Å/min.) and a selectivity of approximately 1.5:1 over photoresist. The SEM photo shows a promising anisotropic etch. We noticed a radially nonuniform etch on the wafer during both etches; the dies near the edge of the wafer cleared sooner than the dies in the center of the wafer.

The primary equipment for measuring tungsten etch rates are Alphastep 200 (as200) and Nanospec. The optical thickness measurement (program #10: positive resist on silicon) from Nanospec gives us a value for the remaining resist thickness. However, the photoresist lies on tungsten, not silicon, and the resultant measurement is only accurate within 20%. Thus, Nanospec is only used for quick resist-thickness measurements to assure that there is still photoresist left on the wafer before it is sent back to LAM 3.

More accurate tungsten- and resist-thickness measurements are done with Alphastep. The initial resist thickness is measured using Alphastep before the tungsten etch. Then the film-thickness is measured on the exact same feature after the tungsten etch, which gives a value for the sum of the final resist and tungsten thickness. The photoresist is then stripped in a 300 W, O_2 plasma for 10 minutes, and the final tungsten thickness can be measured. This is the amount of tungsten etched in LAM 3. The amount of photoresist etched is obtained by subtracting the final resist thickness from the initial resist thickness. Dividing by the total etching time, we obtain both the tungsten and photoresist etch rates.

After the 3000 Å-thick tungsten film was cleared via anisotropic etch, the thin polysilicon layer (~ 1000 Å) and the underlying PSG layer were also attacked. The polysilicon film was etched away very quickly and isotropically by SF_6 plasma, thus we observed severely undercut polysilicon underneath the tungsten layer (Fig. 5). In fact, narrow tungsten lines were completely freed by this undercutting mechanism. We were able to free tungsten lines as wide as 4.5 μm in a 75 W, 50 mTorr, 60°C, and 20 sccm SF_6 plasma after a 4-minute overetch; this is equivalent to a polysilicon-undercut rate of approximately 1.1 μm/min.. We could clearly observe the freed tungsten lines because these freed tungsten lines curled out of the wafer plane due to the stress gradient in the tungsten films.

2. Full-factorial experiment using SF_6 plasma

2.1. Experimental results

The results for the full-factorial experiment are summarized in Tables 1 and 2, and the SEM photos are shown in Figs. 6-10. The best selectivity over photoresist is 5.1:1, as shown in Table 2, with the lowest power and the highest pressure settings, 75 W and 150 mTorr respectively. However, these etching conditions also give the poorest anisotropy, where the degree of anisotropy is calculated by dividing the vertical distance etched by the maximum horizontal distance etched. The resultant SEM photo shows a characteristic "neck" profile (Fig. 7) [9] when tungsten is etched with a high degree of isotropy. Vertical sidewalls are obtained with the highest power and the lowest pressure settings, 175 W and 50 mTorr respectively.

There is a clear trade-off between selectivity and anisotropy. High R.F. power increases the degree of anisotropy because it enhances the ion bombardment that creates the anisotropic etch, yet the same ion bombardment causes enhanced physical etching of the photoresist and decreases the selectivity. An increase in the chamber pressure increases the concentration of fluorine radicals, which attack tungsten isotropically but selectively over

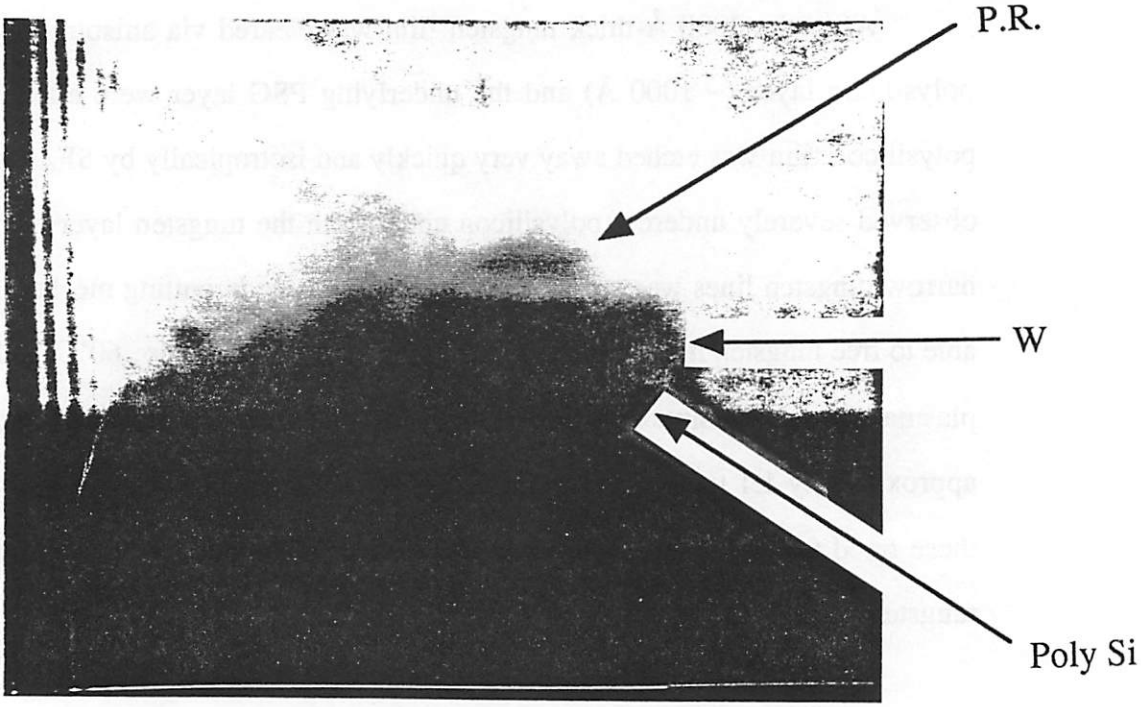


Fig. 4 SEM of CF_4 etch (75 W, 50 mT, CF_4 20 sccm).

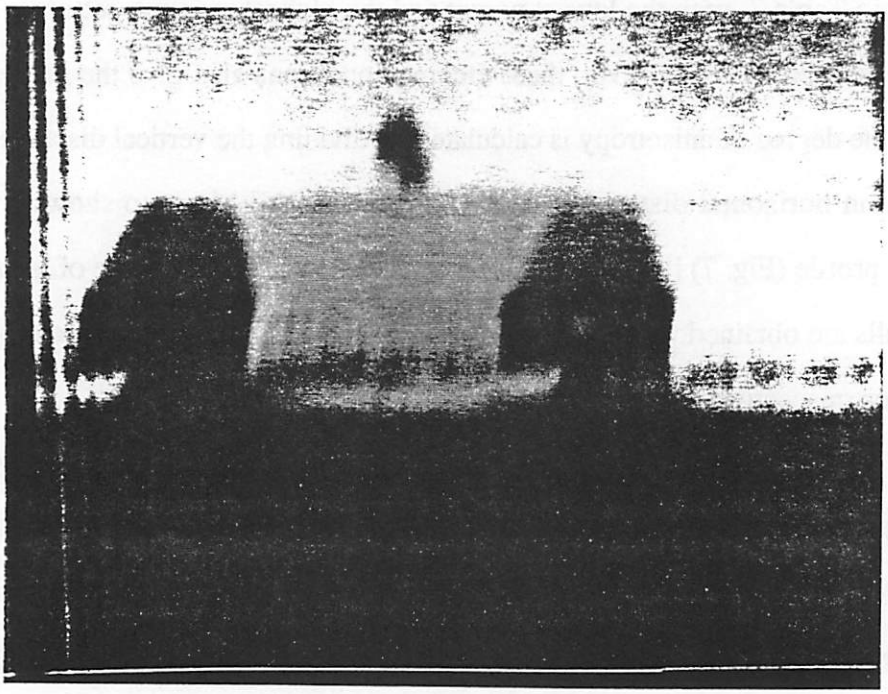


Fig. 5 SEM shows severe polysilicon-undercut underneath the tungsten layer.

photoresist. An increase in either power or pressure enhances the tungsten etch rate - enhanced ion bombardment or increased fluorine radical concentration will result in a higher tungsten etch rate. All these experimental results correspond to the prediction given in Professor Lieberman's lecture. In addition, we have noticed that the radial etching uniformity across the wafer is independent of the applied R.F. power, but improves with increasing chamber pressure. This may be due to the fact that the shower head hole pattern, through which the gasses flow into the chamber, only covers a circular area 4" in diameter; this arrangement may make the gas distribution more uniform at a higher chamber pressure.

We have also observed that thinner tungsten films (~ 3000 Å) are etched more slowly than thicker ones (~ 1 μm). This may be due to the fact that tungsten forms a thin oxide layer in air, which is slowly etched away by SF_6 plasma; thus, a thinner tungsten film may appear to be etched more slowly because it has more oxide per unit thickness of tungsten. Moreover, the stress in tungsten may also affect the etch rate. Thin tungsten films are compressive, and the tungsten atoms are being pushed together by the underlying layers; this may result in a slower etch rate. Thick tungsten films, on the other hand, are tensile, and the tungsten atoms are being pulled apart by the underlying layers; this may result in a faster etch rate. However, the strain is such a small fraction of unity, and the tungsten density cannot vary too much. Thus the stress in tungsten may not pose a significant effect in the tungsten etch rate.

2.2. Theoretical modeling and curve fitting

In theory, the ion flux in the etching chamber is given by

$$\Gamma_i = \frac{P_{rf}}{2\,\varepsilon_T} \propto P_{rf} \tag{1}$$

where P_{rf} is the applied R.F. power, and ε_T is the energy loss per electron-ion pair created. The ion bombardment energy has the following relationship with P_{rf}:

$$\varepsilon_i \propto V_{rf} \propto \sqrt{P_{rf}} \tag{2}$$

where V_{rf} is the R.F. voltage. And the fluorine radical flux is given by

$$\Gamma_F = \frac{A}{2\,\varepsilon_T}\left(\frac{n_0\,l}{2\,u_B}\right)^{1\,-\,\varepsilon_F/\varepsilon_\alpha} P_{rf} \propto p^{\alpha} P_{rf} \tag{3}$$

where n_0 is the neutral gas density, u_B is the Bohm velocity, p is the chamber pressure, and α is a constant between 0 and 1. The anisotropic etch rate of tungsten is strongly dependent on Γ_i, Γ_F, and ε_i while the isotropic etch rate of tungsten depends only on Γ_F. The photoresist etch rate depends only on Γ_i and ε_i. Figures 11-15 display the response surfaces fitted for the experimentally determined values of the tungsten and photoresist etch rates, the selectivity over photoresist, tungsten etching uniformity, and the anisotropy respectively.

Since there are only five experimental points available for each of the above measurements, nearly any function of P_{rf} and p can fit the data. There are total of six parameters in a full-quadratic expression involving P_{rf} and p, so one of the parameters must be dropped. The p^2 term is dropped from all but the etching uniformity 3-D plots because, as seen from Eq. (3), pressure only comes into the fluorine radical flux with an exponent less than 1. As seen from Figs. 11-12, the response surfaces are roughly linear for the tungsten and photoresist etch rates but with a significant $P_{rf} \cdot p$ term. The selectivity, in contrast, shows a visible curvature in its response-surface plots, indicating a significant P_{rf}^2 term. This result is expected since the selectivity is defined as the ratio of the tungsten etch rate and the resist etch rate, and the tungsten etch rate is strongly dependent on the fluorine radical flux, which in turn depends on P_{rf}, while the resist etch rate is not dependent on Γ_F at all. A response surface for the inverse of the anisotropy (Fig. 15) has also been obtained. The inverse of the anisotropy is graphed because one of the experimental points gives a nearly infinite degree of anisotropy. Since the degree of anisotropy is roughly equal to the

ratio of the tungsten anisotropic etch rate and the isotropic etch rate, its inverse is fitted with terms of negative powers in P_{rf} and p. Physical interpretation of this surface plot is obscure - it only serves as a visual aid for the reader. Finally, the response surface for the tungsten etching uniformity is shown in Fig. 14, and it shows a non-linear dependence on the chamber pressure which is a characteristic of the plasma etcher.

2.3. Analysis of variance

An analysis of variance was performed for the power and pressure dependence of the tungsten etch rate and the selectivity over photoresist. Unfortunately, we did not have a variance for every experimental point because the data are not comparable for different tungsten thickness and die positions - there is a systematic offset involved. Therefore, only the variance at the center point (125 W and 100 mTorr) is taken into account.

For the tungsten etch rate, a 99% confidence level is obtained for its dependence on P_{rf}, p, and the product $P_{rf} \cdot p$. However, a confidence level of less than 90% is found for the significance of curvature, which agrees with the observation made in Section 2.2. For the selectivity over photoresist, again a 99% confidence level is obtained for the significance of P_{rf}, $P_{rf} \cdot p$, and curvature terms, while a 95% confidence level is found for its dependence on p.

3. Other experiments

After the SF_6 etching process has been developed and characterized, other experiments involving chlorine etching and aluminum masks were conducted. Extremely discouraging results were obtained when trying to etch tungsten with chlorine chemistry. The following recipe was used:

Recipe 4. Chlorine etch recipe.

[REACTOR] RECIPE			[REACTOR] RECIPE			[REACTOR] RECIPE		
STEP # 01			STEP # 02			STEP # 03		
PRESS	250.	[MTORR]	PRESS	250.	[MTORR]	PRESS	0.	[MTORR]
RF LWR	0.	WATTS	RF LWR	250.	WATTS	RF LWR	0.	WATTS
BCL3	0.	SCCM	BCL3	0.	SCCM	BCL3	0.	SCCM
N2	50.	SCCM	N2	50.	SCCM	N2	0.	SCCM
CL2	20.	SCCM	CL2	20.	SCCM	CL2	0.	SCCM
CHCL3	20.	SCCM	CHCL3	20.	SCCM	CHCL3	0.	SCCM
SF6	0.	SCCM	SF6	0.	SCCM	SF6	0.	SCCM
COMPL	[STABILITY OR TIME]		COMPL	[TIME]		COMPL	[RECIPE]	
WAIT	00:20	MIN:SEC	WAIT	02:00	MIN:SEC	WAIT	00:00	MIN:SEC

The tungsten etch rate was extremely low (~ 100 Å/min.), which more or less agrees with the results obtained by Fischl and Hess [10]. The etched tungsten was not clearly visible form the SEM photo (Fig. 16), even when we applied 250 W R.F. power (0.77 W/cm^2) during the etch.

We began the aluminum-mask experiment but were unable to complete it by press time; a processing error cost us our first batch of wafers, and we did not have time to prepare another batch at the time of printing.

Selective Tungsten Etching Data Set #1

Power [W]	Pressure [mTorr]	W etch rate [Å/min.]	P.R. etch rate [Å/min.]	P.R. ething uniformity	Selectivity [W : P.R.]
75	50	in: 1113 out: ——	in: 831 out: 1212	0.81	in: 1.34 out: ——
75	50	in: 1029 out: ——	in: 599 out: 815	0.85	in: 1.72 out: ——
75	150	in: <2672 out: ——	in: 1460 out: 1616	0.95	in: <1.83 out: ——
125	100	in: 1689 out: ——	in: 1874 out: 2348	0.89	in: 0.90 out: ——
175	50	in: 1007 out: ——	in: 2091 out: 2960	0.83	in: 0.48 out: ——
175	150	in: 2391 out: ——	in: 3121 out: 3508	0.94	in: 0.77 out: ——

$$\text{Etching uniformity} = 1 - \frac{|\text{etch rate (out)} - \text{etch rate (in)}|}{[\text{etch rate (out)} + \text{etch rate (in)}]}$$

$$\text{Selectivity} = \frac{\text{W etch rate}}{\text{P.R. etch rate}}$$

$$\text{Anisotropy} = \frac{\text{vertical distance etched}}{\text{max. horizontal distance etched}}$$

Table 1 Full-factorial experiment - set #1.

EXPERIMENTAL DATA

Selective Tungsten Etching Data Set #2

Power [W]	Pressure [mTorr]	W etch rate [Å/min.]	P.R. etch rate [Å/min.]	W ething uniformity	P.R. ething uniformity	Selectivity [W : P.R.]	Anisotropy (SEM)
75	50	in: 1191 out: 1891	in: 299 out: 539	0.77	0.71	in: 3.98 out: 3.51	5:1
75	150	in: 2293 out: 2807	in: 553 out: 550	0.90	0.99	in: 4.15 out: 5.10	4:1
125	100	in: 2140 out: 2903	in: 1136 out: 1343	0.85	0.92	in: 1.88 out: 2.16	10:1
125	100	in: 1895 out: 2545	in: 1168 out: 1231	0.85	0.97	in: 1.62 out: 2.07	8:1
175	50	in: 1604 out: 2496	in: 1020 out: 1480	0.78	0.82	in: 1.57 out: 1.69	NO undercut
175	150	in: 3740 out: 4800	in: 3560 out: 2652	0.88	0.85	in: 1.05 out: 1.81	8:1

$$\text{Etching uniformity} = 1 - \frac{|\text{etch rate (out)} - \text{etch rate (in)}|}{[\text{etch rate (out)} + \text{etch rate (in)}]}$$

$$\text{Selectivity} = \frac{\text{W etch rate}}{\text{P.R. etch rate}}$$

$$\text{Anisotropy} = \frac{\text{vertical distance etched}}{\text{max. horizontal distance etched}}$$

Table 2 Full-factorial experiment - set #2.

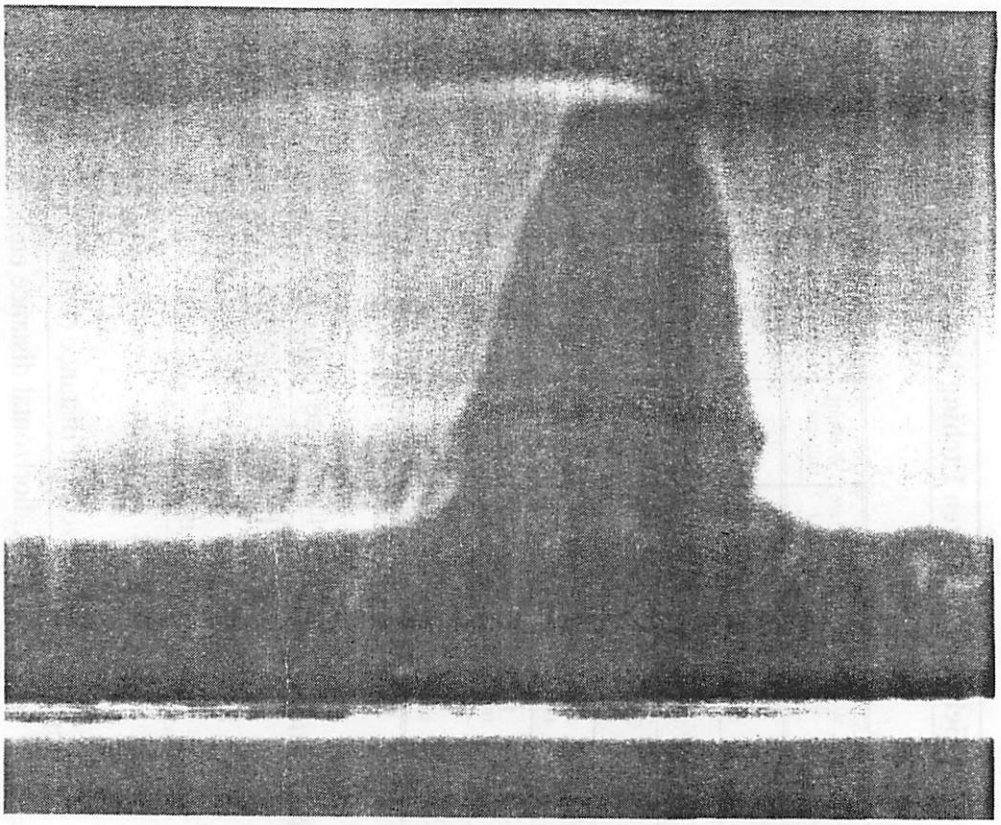


Fig. 6 SEM of SF_6 etch (75 W, 50 mT, SF_6 20 sccm).

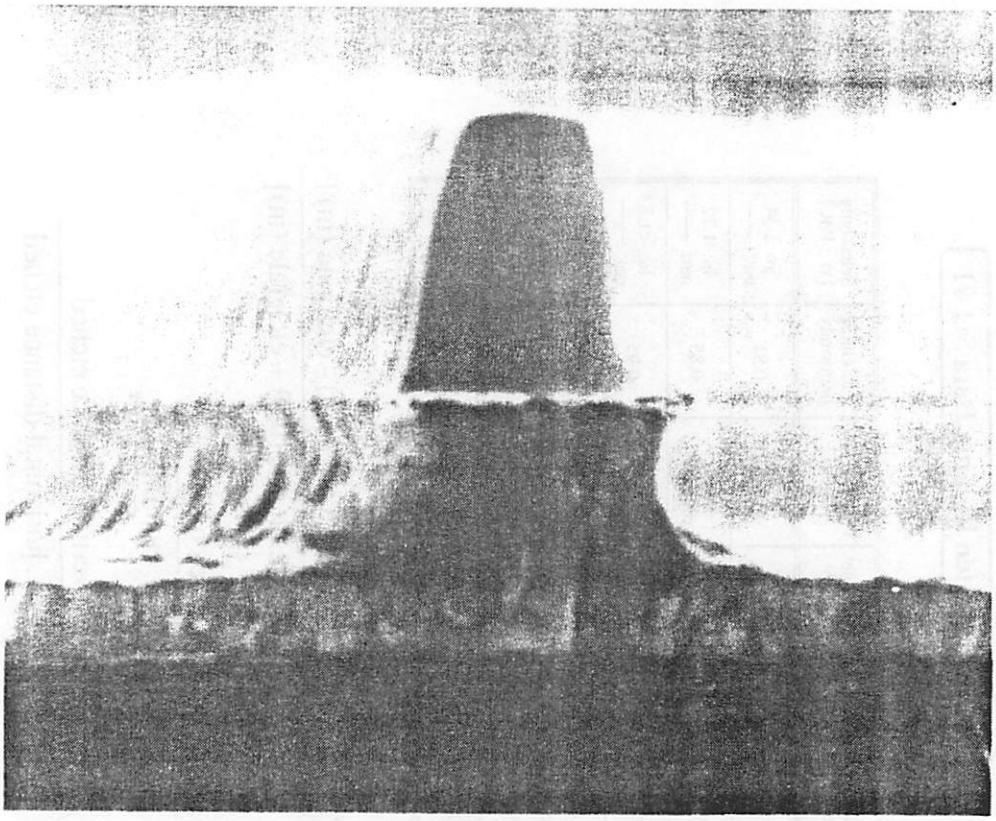


Fig. 7 SEM of SF_6 etch (75 W, 150 mT, SF_6 20 sccm).

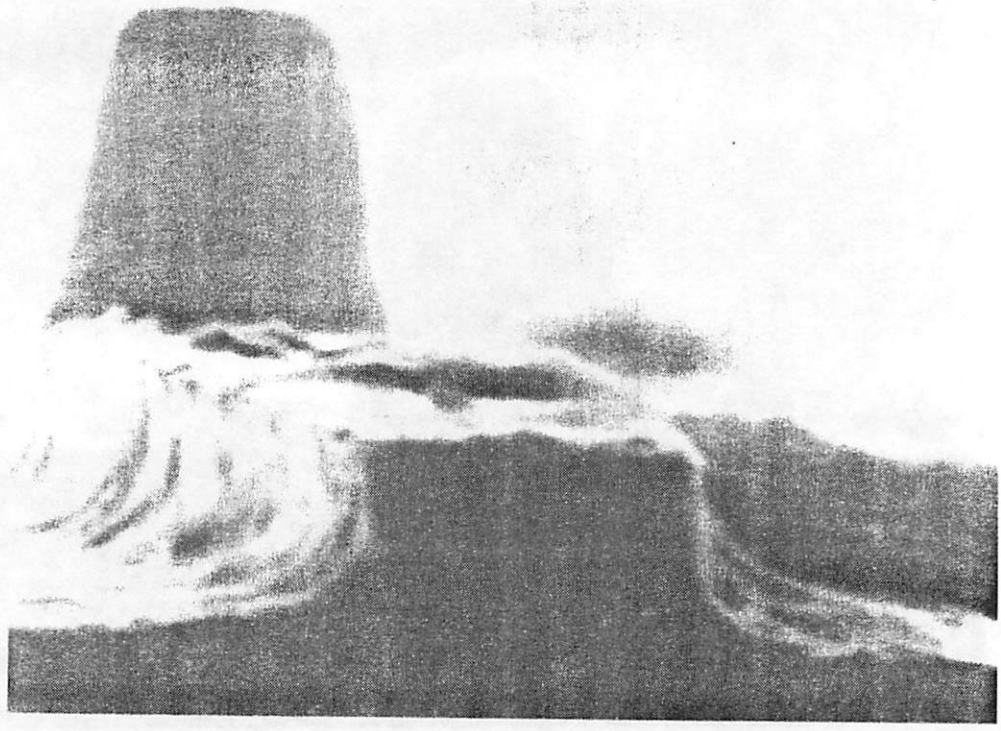


Fig. 8 SEM of SF_6 etch (125 W, 100 mT, SF_6 20 sccm).

Fig. 16 SEM of chlorine etch (250 W, 250 mT, Cl_2 20 sccm, N_2 50 sccm).

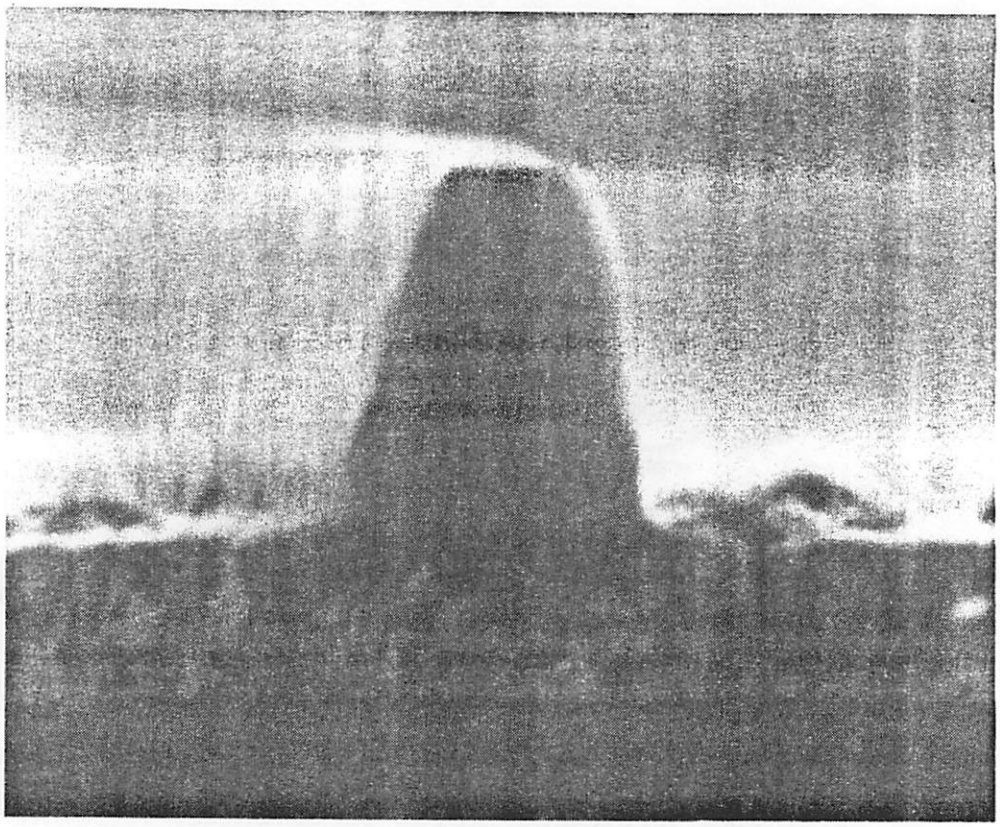


Fig. 9 SEM of SF_6 etch (175 W, 50 mT, SF_6 20 sccm).

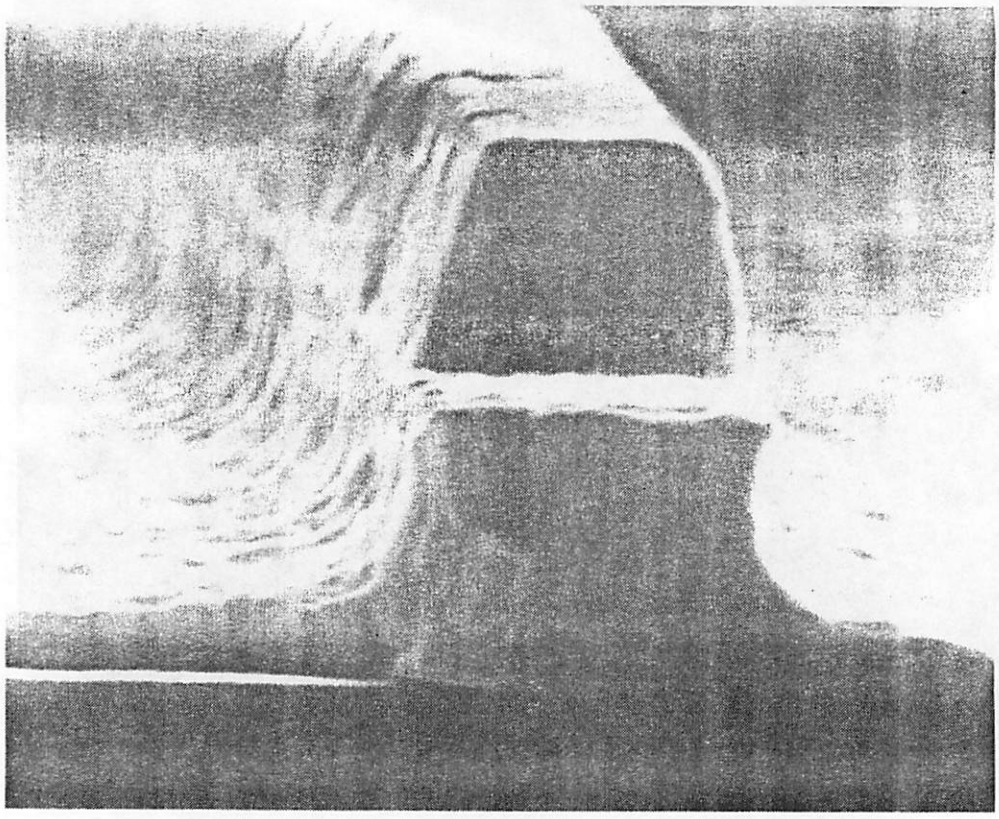
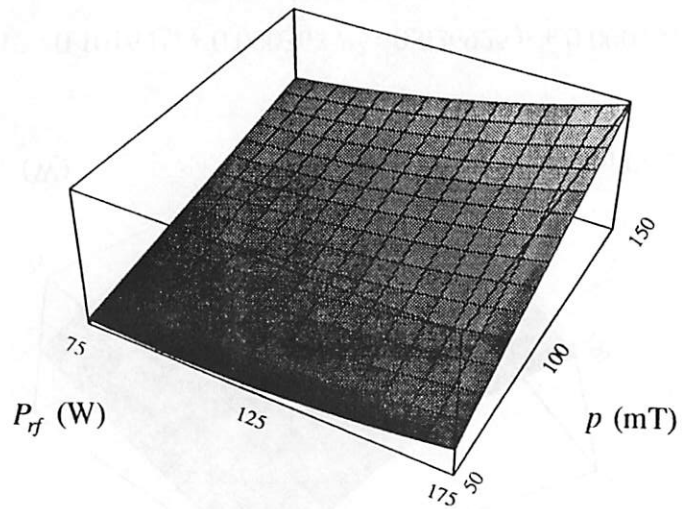


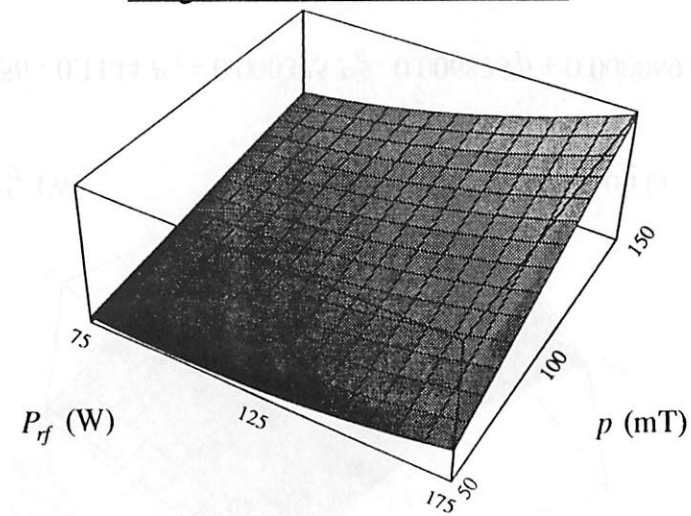
Fig. 10 SEM of SF_6 etch (175 W, 150 mT, SF_6 20 sccm).

Tungsten Etch Rate: center die



$$1710.25 - 19.94\,P_{rf} + 0.0756\,P_{rf}^2 - 3.265\,p + 0.1034\,P_{rf} \cdot p$$

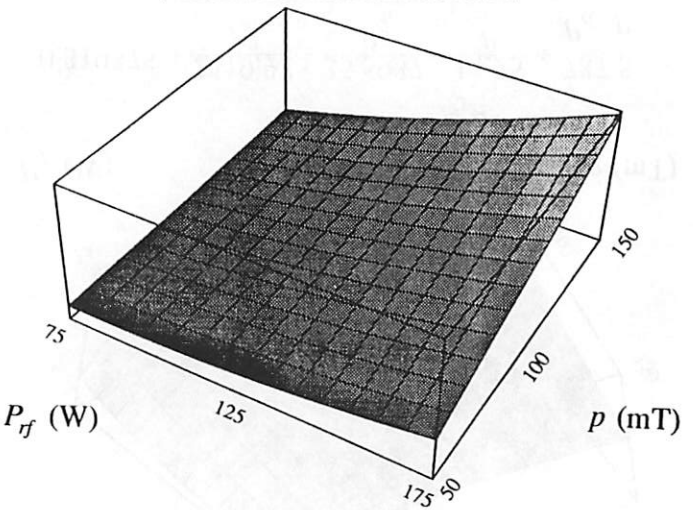
Tungsten Etch Rate: outermost die



$$2940.88 - 28.34\,P_{rf} + 0.1098\,P_{rf}^2 - 1.25\,p + 0.1388\,P_{rf} \cdot p$$

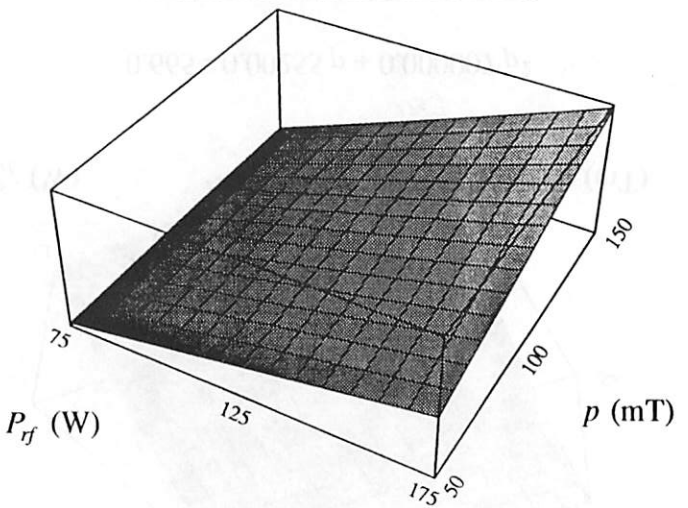
Fig. 11 Response surface of the tungsten etch rate.

Resist Etch Rate: center die



$$1570 - 24.82\,P_{rf} + 0.0824\,P_{rf}^2 - 14.605\,p + 0.2286\,P_{rf} \cdot p$$

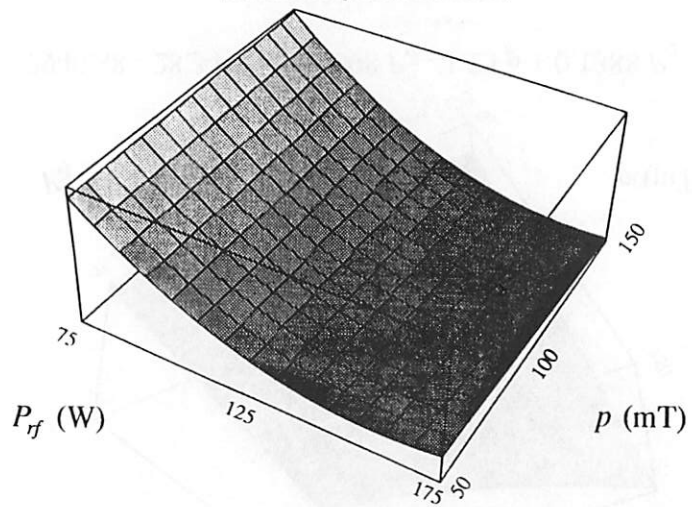
Resist Etch Rate: outermost die



$$358.938 - 1.78\,P_{rf} + 0.0073\,P_{rf}^2 - 8.5975\,p + 0.1161\,P_{rf} \cdot p$$

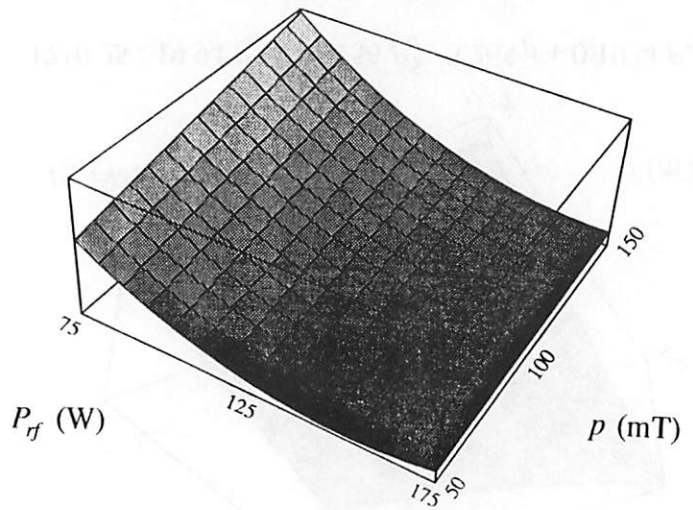
Fig. 12 Response surface of the resist etch rate.

Selectivity: center die



$$10.3656 - 0.1144\,P_{rf} + 0.000375\,P_{rf}^2 - 0.006875\,p + 0.000069\,P_{rf} \cdot p$$

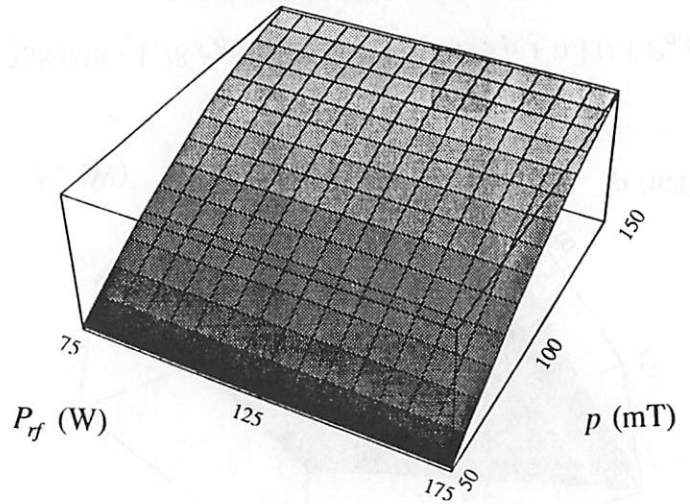
Selectivity: outermost die



$$8.29312 - 0.1016\,P_{rf} + 0.000363\,P_{rf}^2 - 0.026925\,p + 0.000147\,P_{rf} \cdot p$$

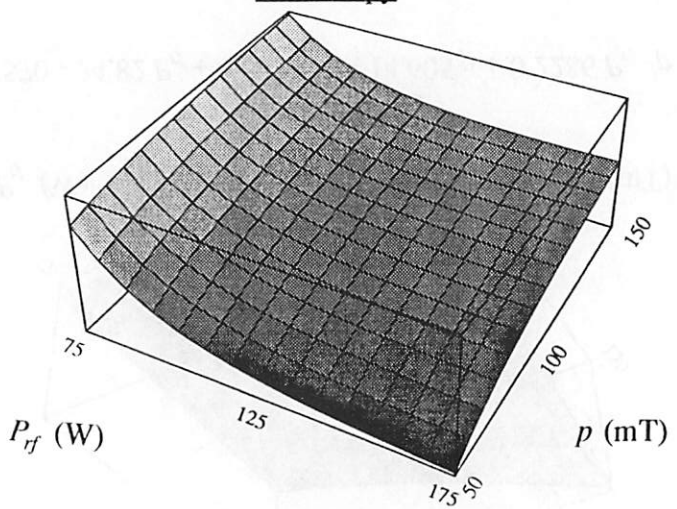
Fig. 13 Response surface of the selectivity over resist.

Tungsten Etching Uniformity



$$0.665 - 0.00255\,p + 0.000007\,p^2$$

Anisotropy



$$0.319375 + \frac{2419.92}{P_{rf}^2} - \frac{35.5937}{P_{rf}} - \frac{14.25}{p} + \frac{787.5}{P_{rf} \cdot p}$$

Fig. 14 Response surface of the tungsten etching uniformity. (top)
Fig. 15 Response surface of the inverse of the anisotropy. (bottom)

CONCLUSION AND RECOMMENDATIONS FOR FUTURE WORK

A completely Microlab-compatible tungsten-etching process has been successfully developed as the result of this class project. This newly developed tungsten-etching process utilizes the dedicated aluminum etcher LAM 3 for selective tungsten etching. Etch processes using SF_6 single source etch gas have been demonstrated to anisotropically and selectively etch 1 μm tungsten patterns with no trace of cross contamination in the etching chamber due to the mixing of chlorine and fluorine chemistries. The process is applicable for Olin-Hunt photoresist mask and possibly for inorganic-metal mask [11] as well.

Operating at a high temperature (60°C) and relatively high pressure (≥ 50 mTorr) region, we were able to obtain near-perfect vertical profiles or an approximately 5:1 selectivity over photoresist depending on the etching conditions. There is a trade-off, however, between anisotropy and selectivity.

The etching conditions which gave maximum degree of anisotropy and used the standard 1-μm photoresist technology are as follows: 175 W, 20 sccm SF_6, and 50 mTorr. These etching conditions are extremely useful for patterning thin (~5000 Å) tungsten films in CMOS applications. On the other hand, for very thick (≥ 2.5 μm) tungsten films, it may be prudent to use the 2-μm photoresist technology and the following etching conditions: 75 W, 20 sccm SF_6, and 150 mTorr since these etching conditions maximize the selectivity over photoresist (~5:1) at the expense of vertical sidewalls. For tungsten micromachining applications, when 2 μm-thick tungsten films need to be patterned, the following etching conditions will give the best compromising result: 125 W, 20 sccm SF_6, and 100 mTorr.

A possible future research direction is to use aluminum as the masking material and characterize the effect of inorganic-metal mask in tungsten etching. Another possible direction involves etching tungsten with a mixture of both chlorinated and fluorinated gasses. However, this experiment is not recommended since the possible polymer formation may destroy the aluminum etching performance in LAM 3.

ACKNOWLEDGEMENT

We would like to express our sincere appreciation to Professor A. R. Neureuther, Professor M. A. Lieberman and Professor W. G. Oldham for their guidance and encouragement in the course of this work. We are also grateful to Professor D. Flamm for his interest in this work and for his assistance in tungsten-etching literature research.

We wish especially to thank Dr. Simon Wong at Stanford University for his assistance in the CVD tungsten process. Thanks also to Valerie Wong and Jennifer Dockter for help with the SEM observations, and Tariq Haniff for preparing tungsten samples.

Also, special thanks go to Professor P. Ko and staff members in the Microfabrication Laboratory, especially Katalin Voros, Robert M. Hamilton, Debra Hebert, and Robert Norman, for their assistance throughout the tungsten-etching characterization process.

Finally and most importantly, we would like to express our gratitude to all other graduate students in the EECS-290N class for their encouragement and support.

REFERENCES

1. S. Tandon and G. Jones, "Reactive ion etching of tungsten in SF_6 and CF_4," *Tungsten and Other Refractory Metals for VLSI Applications*, **IV**, 165-174, (1989).

2. M. E. Burba, E. Degenkolb, S. Henck, M. Tabasky, E. D. Jungbluth, and R. Wilson, "Selective dry etching of tungsten for VLSI metallization," *Journal of the Electrochemical Society*, **133**, 2113-2118, (1986).

3. W. Yun, "CMOS metallization for integration with micromachining process," M.S. Report, Dept. of Electrical Engineering and Computer Sciences, University of California at Berkeley, May 1989.

4. W. C. Tang, T.-C. H. Nguyen, and R. T. Howe, "Laterally driven polysilicon resonant microstructures," *IEEE Micro Electromechanical Systems Workshop, Salt Lake City, Utah, February 20-22, 1989* , pp. 53-59.

5. W. C. Tang, T.-C. H. Nguyen, and R. T. Howe, "Laterally driven polysilicon resonant microstructures," *Sensors and Actuators, 20* (1989), pp. 25-32.

6. W. C. Tang, "Electrostatic comb drive for resonant sensor and actuator applications," Ph.D. Thesis, Dept. of Electrical Engineering and Computer Sciences, University of California at Berkeley, December 1990.

7. *AutoEtch 690 Plasma Etching System Operational Manual*, Lam Research Corporation.

8. *Berkeley Microfabrication Laboratory Equipment Manual*, section 7.8, Dept. of Electrical Engineering and Computer Sciences, University of California at Berkeley.

9. S. C. Shunk, D. M. Tennant, and M. D. Feuer, "Resolution enhancement by anisotropic lateral etching to form sub-0.10 μm refractory metal HIGFET gates," AT&T Bell Laboratories internal report, May 30, 1989.

10. D. S. Fischl and D. W. Hess, "Plasma-enhanced etching of tungsten and tungsten silicide in chlorine-containing discharges," *Journal of the Electrochemical Society*, **134**, 2265-2268, (1987).

11. C. C. Beatty and D. D. Kessler, "Multilevel tungsten integrated circuit metallization with chromium non-erodible mask and etch stop layers," *Fourth International VLSI Multilevel Interconnect Conference Proceedings, June 15-16, 1987*, pp. 163-166.

Measurement of Residual Stress and Strain in IC-Processed Thin Films

Kirt Williams and Kristen McNair

EECS 290N Class Project
U. C. Berkeley
Fall 1990

ABSTRACT

This project was designed to measure residual stresses and strains in thin films using two different types of test structures. The films studied were polycrystalline silicon, "low-stress" (silicon-rich) silicon nitride, selectively deposited LPCVD tungsten, sputtered tungsten, and sputtered aluminum.

The first type of test structure was patterned from free-standing doubly clamped beams, which are sensitive to compressive strain, and rings with beams across their diameters, which are sensitive to tensile strain. For a minimum magnitude of residual strain that depends on geometry, all beams in these structures longer than a critical length will buckle. An array of these structures were made, varying the beam lengths and film thicknesses to achieve sensitivity to several orders of magnitude of residual strain. Even with such a wide dynamic range, each film that was looked at either buckled all or none of the test structures, which only told us that the films had more or less strain than a certain value, or was near zero. It did show that the low-stress nitride indeed had a low stress, and that the normally low-stress poly could have a large compressive stress.

The second test structure was a two-layer cantilever beam, with one of the materials of a low or known residual strain and the second's to be determined. Differences in residual strain caused the beam to deflect into a circular arc. Measuring the radius of curvature of the beam using one of three methods, and taking into account bending of the beams on their own, residual stresses were calculated. This technique was not limited in range like the single-layer one was, allowing us to calculate a tensile residual stress for the tungsten of about 3GPa.

The simplest and most accurate of the three radius measurement methods is measuring long beams that have curled into circles directly. For larger radii of curvature, deflection measurements by focusing a microscope on different parts of a beam is appropriate. Together, these two methods cover a range of ρ = 10μm to 80cm. A third method using a laser beam to try to measure larger ρ's was unsuccessful.

Measurement of Residual Stress and Strain in IC-Processed Thin Films

ACCOMPLISHMENTS

- Designed a four mask set
- Designed a process flow that works with many different film combinations
- Fabricated single-film test structures sensitive to several orders of magnitude of compressive stress and an order of magnitude of tensile stress
- Fabricated two-film cantilevers to determine residual stress from radius of curvature
- Used three different methods to determine radius of curvature
- Designed and built a laser apparatus to measure radius of curvature
- Studied poly, low-stress nitride, and selective CVD tungsten
- Obtained residual stress results that were self-consistent

ISSUES

- The residual strains were never in a range that resulted in some test structures buckling and some not, so we don't know how much resolution can be attained
- The laser apparatus did not work

III. PROBLEM AND KEY ISSUES

Residual stresses, the unrelieved stresses remaining in thin films after processing, can have detrimental effects on the operation of integrated circuits and micromechanical devices. For example, if tungsten is excessively tensile and is used for long interconnect lines, these lines may peel off the die. In various sensor and actuator devices, silicon nitride is used to form a mechanical membrane. If the silicon nitride is too compressive, the membrane will buckle; if it is too tensile, the membrane will rupture. Before fabricating a device, one would like to characterize the stress in a thin-film to ensure that it meets the device requirements.

In the Berkeley Microfabrication Facility, thin film depositions vary from run to run despite use of the same recipe. Thus, the stress characteristics of a thin film cannot be guaranteed after determining the stress of the film resulting from one or two deposition runs. To be certain of the stress behavior of a thin film, one should characterize the stress immediately before or even during the deposition for the device fabrication. Of course, such frequent characterization could require more time than student processors are willing to spend. A simple method for determining stress behavior before doing the actual process deposition run might encourage students to use this method. Consequently, the number of runs wasted as a result of excessive thin-film stresses could be reduced.

Presently, a simple method of stress characterization, the wafer curvature method, exists in the Microlab. Its accuracy, however, is questionable, and it is not appropriate for all situations where residual stress needs to be determined. Using this method, one deposits ~0.5μm-3.0μm of the thin-film of interest on a test wafer either before or during the actual deposition. Using the Flatgage, which ultrasonically measures distances from a reference plane to a wafer, the processor generates a set of wafer deflection data. Next, he removes the film completely from one side of the wafer and repeats the Flatgage measurement for a second set of deflection data. He then takes the difference between the two sets of data at four points on the wafer rim and at the center. From the average change in deflection from the rim to the center of the wafer, the processor calculates the residual stress of the thin film. One problem with this method is that our test wafers have (100) orientation and thus have a Young's modulus that varies in different directions along the wafer surface. (111) wafers have a constant Young's modulus, but are not stocked in the Microlab. Consequently, a film with uniform stress behavior could cause non-uniform bowing of the wafer, since the wafer strain would vary with direction. A second problem is that this method assumes that the wafer bows spherically; however, the difference data often show that the height difference between the wafer center and rim is not constant. A third problem is that even if this method is accurate, film characteristics can vary from wafer to

wafer in the same deposition run. Finally, this method is inappropriate for selectively deposited films, whose characteristics can be different from the blanket-deposited films required for this technique.

As a result, a simple, more accurate method for stress characterization is desired. This method would require existing pieces of Microlab equipment or the addition of a small, inexpensive piece of equipment. In our project, we explore a few such methods.

IV. STATE-OF-THE-ART

[1] Borden, Peter G., "A simple technique for determining the stress at the Si/SiO2 interface," Appl. Phys. Lett. Vol. 36, No. 10, May 15, 1980, pp.829-830.

This technique involves etching a groove in silicon using the oxide layer as a mask. The oxide overhang, resulting from the silicon undercutting etch, is not constrained by the Si/SiO_2 bond and assumes a periodic shape. This periodicity results from the fact that once the underlying silicon is removed, the oxide is unstressed. By measuring the ratio of the lengths of the stressed and unstressed oxide, the strain required to conform the oxide to the silicon surface is found. The oxide/silicon interface stress and the strain in the silicon can then be calculated. The measured values are in good agreement with those obtained by the bowing and x-ray diffraction techniques. The technique is useable only when the oxide is under compressive stress.

[2] Howe, R.T., and Muller, R.S., "Stress in polycrystalline and amorphous silicon thin films," J. Appl. Phys., Vol. 54, No.8, August 1983, pp.4674-4675.

Stress in polycrystalline and amorphous silicon thin films deposited on oxidized silicon wafers is determined from the lengthening of the undercut edge of a silicon beam. The silicon film must be undercut sufficiently to allow the compressive strain to relax completely. The stress in the silicon film is given by

$$\sigma \approx -E\frac{(\pi A_0/\lambda)^2}{(1-\nu)}$$

The amplitude A_o and wavelength λ are measured using the known oxide thickness to calibrate the SEM magnification. The technique measures only compressive stresses, has a stress resolution of 1MPa, and a spatial resolution on the wafer of 250 μm. It is found that the large compressive stress in silicon films is greatly reduced by annealing at 1100C for 20min. in N2.

[3] Guckel, H., Randazzo, T., and Burns, D.W., "A simple technique for the determination of mechanical strain in thin films with applications to polysilicon," J. Appl. Phys. Vol. 57, No. 5, March 1, 1985, pp.1671-1675.

IC-fabricated doubley beams have resonant frequencies. The axial load, P, on these beams is nonzero because it is due to the built-in strain. Increases in P are observed to cause decreases in the self-resonance. Sufficiently large axial strains will produce a situation for which the smallest resonant frequency becomes zero. This condition causes buckling and is associated with a specific value of P which must be determined. Buckling for rectangular beams which are simply supported or clamped is only a function of beam geometry and not material parameters. Beams for strain field diagnostics are composed of a free standing portion and two symmetrical supports. The supports are fabricated from the same material as the beams, resulting in a flat, dumbbell-like structure for which clamped boundary conditions can be obtained and can be verified optically. Buckling is established by fabricating beams of similar topology but changing lengths and by observing beam deformation with an interference contrast microscope. Again, this technique measures only compressive stresses.

[4] Allen, M.G., Mehregany, M., Howe, R.T., and Senturia, Stephen D., "Microfabricated structures for the in situ measurement of residual stress, Young's modulus, and ultimate strain of thin films," Appl. Phys. Lett., Vol. 51, No. 4, July 27, 1987, pp.241-243.

This letter reports two types of microfabricated test structures and illustrates their use for the measurement of residual tensile stress, Young's modulus, and ultimate strain of a polymeric thin film. The first structure is a suspended thin-film membrane. The second is a released structure, an asymmetric structure patterned into the thin-film membrane prior to removal of the support. The released structure concept relies on the presence of residual tensile stress in the film. When released, the structure deforms due to the residual stress. Measurement of this deformation in conjunction with

appropriate mechanical models yields the ratio of residual stress to Young's modulus. Released structures having both T and H shapes are investigated. The ultimate strain can also be determined from released structures which fracture due to the residual stress.

[5] Stewart, R.A., and Kim, J., "Young's Modulus and Residual Stress of LPCVD Silicon-rich Silicon Nitride Determined from Membrane Deflection"

This paper describes a simple experimental setup for determining residual stress and Young's modulus by a membrane deflection technique. The pressure-deflection measurement setup was designed to provide quick and simple measurements without sacrificing too much precision. Vertical membrane deflection is measure by focusing on the center of the membrane with an optical microscope. Deflection vs. differential pressure plots were generated, and values for stress and Young's modulus were obtained from a least-squares fit to the measure data. Comparison of membrane deflection measurements with wafer curvature measurements shows that for LPDVD silicon-rich silicon nitride films, stress obtained by the latter method is consistently 50-60% higher than by the former method. Further studies need to be carried out to obtain a more complete characterization of mechanical properties of silicon-rich silicon nitride and other micromechanical materials.

[6] Burns, D.W., Ph.D. Thesis, Department of Materials Science, University of Wisconsin - Madison, May 1988.

Describes doubly clamped beams and ring-and-beam residual strain test structures.

[7] Sekimoto, M., Yoshihara, H., and Ohkubo, T., "Silicon nitride single- layer x-ray mask," J. Vac. Sci. Technol., 21(4), Nov./Dec. 1982. Contains the equation for wafer bowing used in conjunction with the Flatgage measurements:

$$\sigma_r = \frac{E_s}{3(1-\nu_s)}\left[\frac{H_s}{r}\right]^2\frac{Z}{H}$$

where E_s, ν_s, andH_s are Young's modulus, Poisson's ratio, and the thickness of the substrate, respectively, r is the radius on the wafer to the points on the rim where Z is measured, H is the film thickness, and Z is the change in the average deflection between the center and the rim.

[8] Roark, Raymond J., and Young, Warren C., "Formulas for Stress and Strain,", 5th edition, McGraw-Hill, 1975. Contains formulas relating stress, strain, deflection, and loading.

[9] Judy, Michael W., Cho, Young-Ho, Howe, Roger T., and Pisano, A. P., "Self-Adjusting Microstructures (SAMS)," to be published (authors with the U. C. Berkeley Berkeley Sensor and Actuator Center). Contains formulas relating residual stresses in two-layer cantilevers to residual stress.

[10] Hoffman, R. W., "Physics of nonmetallic thin films, "Mechanical Properties of Non-Metallic Thin Films," Case-Western Reserve University, Plenum Press, New York, 1976. Analysis of stresses in beams. Measurement techniques.

V. APPROACH

One important goal of our project was to design stress diagnostic structures and an apparatus for measuring the stress in the thin films forming these structures. Another goal was to demonstrate the feasibility of some diagnostic structures previously implemented as described in the literature. We wanted these structures to be simple enough for processors to use them either as a preliminary to, or in tandem with, their actual runs.

For the first stress-diagnostic structures, we designed simple, two-layer cantilever beams, as illustrated in Fig. 4. By having a two-layer beam with a compressive film as the first layer and a tensile film as the second layer, we hoped to observe a positive, upward curvature in the beam that we could then measure with our specially-designed apparatus. We wanted the measurement to be simple so that a processor could make a measurement on his/her diagnostic structures, then from a single calculation, or even a prepared table, determine the stress value. We decided that the two-layer cantilever beams and the laser curvature measurement apparatus met these requirements. One advantage of the laser beam apparatus is that it can measure a larger radius of curvature. The laser apparatus also met the cost requirement for our project. We also had two other optical methods for measuring smaller radii of curvature: measuring radius of curvature by inspection for cantilever beams that curled into circles, and measuring the upward end deflection of cantilever beams from the difference in focus under a microscope. These two methods cover a different curvature range. Any magnitude of stress (or strain) can be calculated from these structures if the radius of curvature can be found and the residual strain in one film is known or negligible.

From our literature search, we chose the doubly clamped beams and ring and beam diagnostic structures implemented by researchers at the University of Wisconsin-Madison, as illustrated in Figs. 1 and 3. These diagnostic structures are attractive because their required processing is simple, as is the measurement technique necessary for determining thin-film strain. The doubly clamped beams are designed to determine the compressive strain of a single-layer film. Since clamped beams longer than a critical length will buckle, we can determine the critical buckling length by observing, through a microscope, the shortest beam in a range of beam lengths. Knowing the critical buckling length, we can then calculate the compressive strain of the thin film. The main advantage of this technique is that the single measurement needed is made visually with a microscope. The disadvantages of this technique are that the critical buckling length is accurate only within the increment between beam lengths, and that the method determines only compressive, not both tensile and compressive, strains.

The ring-and-beam diagnostic structures are designed such that tensile stresses in the rings

create compressive stresses in the central beam. Rings with a radius larger than the critical radius will have buckled beam. We can determine this critical radius by observing the smallest ring with a buckled central beam within a range of ring sizes. From the critical radius, we can calculate the tensile strain. Again, the advantage of this technique is that it requires only a visual measurement using an optical microscope. One disadvantage is that the strain calculation gives a factor of 4 uncertainty and assumes a value for Poisson's ratio of the material. Another disadvantage is that this method determines only tensile, not both compressive and tensile, strains.

VI. DESCRIPTION OF EXPERIMENT

1.0 Design of Test Structures

1.1 Single-Layer Structures: Doubly Clamped Beams and Ring-and-Beam Structures

The doubly clamped beams (Fig. 1) and rings with beams across their diameters (Fig. 2), taken from Burns' work [3], were designed to look at residual strain using single layers of thin films. (Residual stresses are simply found for all the structures in this project by multiplying by the effective Young's modulus, $E' = E/(1-\nu)$ [3]). They function in a similar manner: structures are patterned in a thin film, residual stress cause beams to buckle, and the buckled beams are observed under a microscope (the microscope's numerical aperture must be small enough that the buckled regions appear dark; most Microlab microscopes will work for this). In order to buckle, these structures must be firmly attached to the substrate, but raised from it. This is accomplished by terminating the structures at their endpoints by rectangular pads larger than four times the widest beam, made from the same layer of film. These pads are connected to the substrate by thick rectangular anchors of phospho-silicate glass (PSG).

1.1.1 Doubly Clamped Beams

The doubly clamped beams (Fig. 1) detect compressive residual stresses. (By convention, tensile stresses are positive and compressive ones negative. Thus, a compressive stress with a larger magnitude is more negative.) For a given residual stress ε_r, beams longer than a critical length L_{cr} will buckle according to the Euler buckling criterion $\varepsilon_r < -C\pi^2(h/L_{cr})^2$, where h is the film thickness and C is the coefficient of restraint, a dimensionless parameter describing how much the endpoints of the beam are constrained from moving or tilting out of plane. According to Roark and Young [8], C is 1 for ends that are free to tilt and 4 for firmly clamped ends. Burns [3] describes using wide pads to obtain clamped boundary conditions, but does not mention C and instead uses

$$\varepsilon_r < -\pi^2\left[\frac{h}{L_{cr}}\right]^2, \tag{1}$$

which corresponds to C=1. We used the above equation, although we recognize that this may be an error. Note that ε_r is independent of material parameters, including, according to Burns, variation in strain through the thickness of the film.

Our beams were laid out with widths of both 10 and 20μm for each length to test for unexpected width effects. Lengths were varied from 10 to 60μm in increments of 5μm. This 60x range of

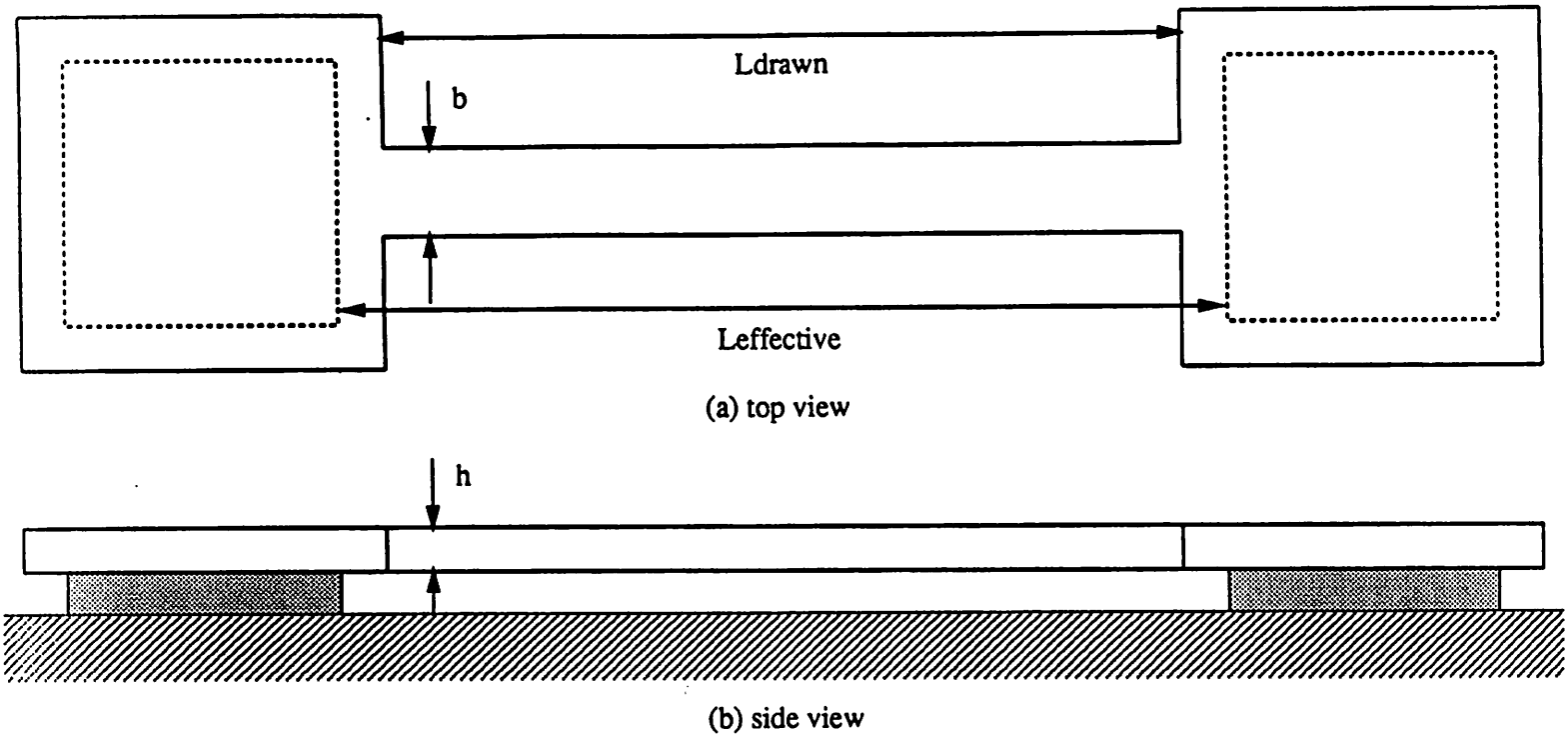


Figure 1: Doubly Clamped Beam

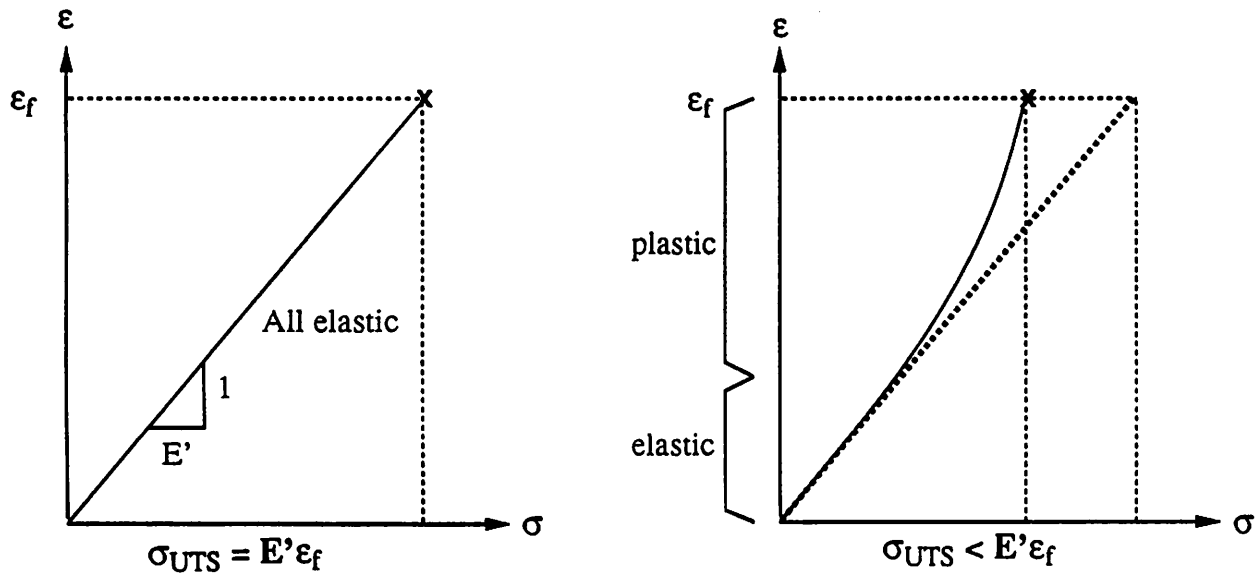


Figure 2: Elastic vs. Plastic Deformation and Fracture

possible lengths for L_{cr} corresponds to a range of 3600x (=60^2) in detectable residual strain. Varying the film thickness from 0.5 to 1μm extends this range by a factor of 4, for a total detectable compressive strain of 14,400x.

In addition to compressive strain, the doubly clamped beams can be used to look at *tensile* strain at fracture ε_f. If the residual strain ε_r is too great, all beams, regardless of length, will break and can easily be detected under a microscope. ε_f is then less than ε_r, but by an unknown amount. An additional uncertainty is due to the effect of stress concentration. At points like the corners where beams are attached to pads, the effective stress is several times greater (by the stress concentration factor k_f) than elsewhere, so fracture will occur at these points. Thus $\varepsilon_f < \varepsilon_r/k_f$.

The ultimate tensile strength σ_{UTS} for materials fracturing in the linear elastic region (Fig. 2a) is simply $\sigma_{UTS} = E'\varepsilon_f$, so an upper limit for σ_{UTS} can be found. Brittle films such as polysilicon and nitride fall into this category. Materials that plasticly deform before breaking have a σ_{UTS} that is less than $E'\varepsilon_f$ (Fig. 2b). Bulk metals such as aluminum exhibit this behavior, although thin films may not due to microcracks present on their surfaces. Thus an upper limit for σ_{UTS} for all thin films can be determined, but it is unknown whether the true σ_{UTS} is near this value or orders of magnitude less.

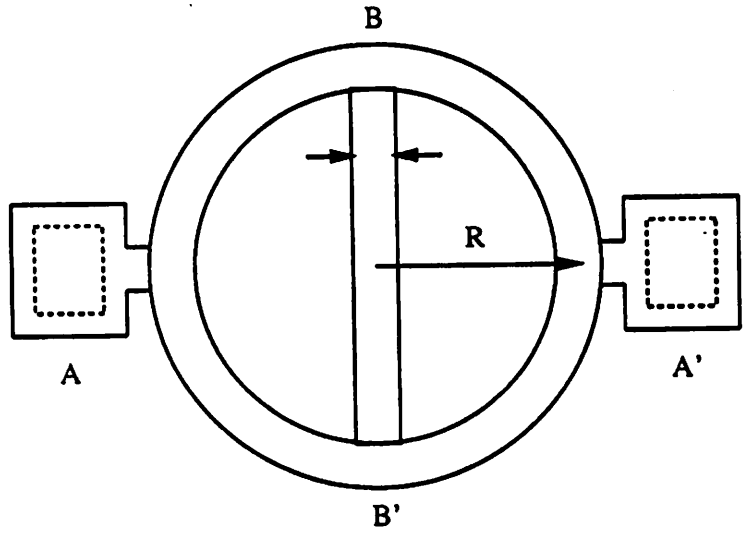
1.1.2 Ring-and-Beam Structures

The ring-and-beam structures (Fig. 2) respond to tensile residual stresses. Because the ring is attached to the substrate at its ends A and A', tensile stresses cause the points B and B' to be pulled together. This, in turn, buckles the beam (of length $\approx 2R$) in the middle if it is long enough. The residual strain falls in a range [3]

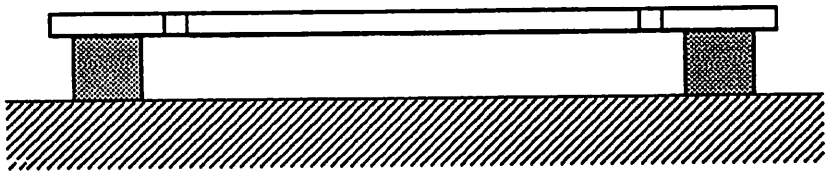
$$\frac{\pi^2 h^2}{12 g(R)(2R_{cr})^2} < \varepsilon_r < \frac{\pi^2 h^2}{3 g(R)(2R_{cr})^2}, \tag{2}$$

where R_{cr} is the smallest ring radius for which bucking will occur and $g(R)$ is the ratio of contraction of points B and B' to extension of points A and A'. $g(R)$ values as a function of radius with the other dimensions as listed below have been tabulated and appear in Fig. 3. The factor of 4 uncertainty is due to the uncertainty in beam end clamping conditions.

Our ring-and-beam shape was taken from pictures of Burns' rings [3]. Ours were laid out with radii from 25 to 50μm in 2.5μm increments, and from 50 to 170μm in 5μm increments, for a 6.8x range in R. Larger rings were not laid out due to die space limitations. All ring circumference widths were 20μm; all beam widths were 10μm. Discounting the factor of 4 in uncertainty in ε_r, and varying the thickness by a factor of 2, a range of 185x of tensile strain should be detectable. If the residual



(a) top view



(b) side view

Figure 3a: Ring and Beam Tensile Strain Diagnostic Structure

R(μm)	g(R)	R(μm)	g(R)
25	0.30722948	80	0.27286496
27	0.31433432	85	0.26172523
30	0.32289659	90	0.25069369
32	0.32724678	95	0.23987572
35	0.33189268	100	0.22934788
37	0.33385122	105	0.219164
40	0.33528448	110	0.20936008
42	0.33535185	115	0.19995815
45	0.33430365	120	0.19096949
47	0.33293573	125	0.18239706
50	0.33002935	130	0.17423764
52	0.32759957	135	0.16658331
55	0.32333424	140	0.15912284
57	0.32013753	145	0.15214263
60	0.31490436	150	0.14552752
65	0.30527369	155	0.13926139
70	0.29485602	160	0.13332765
75	0.28397114	165	0.12770957
		170	0.12239059

Figure 3b: g(R) Values for 20micron Ring Width and 10micron Beam Width

stress is compressive, the beams should be pulled tight and thus remain flat, although the rings themselves may buckle.

1.2 Two-Layer Structures: Cantilever Beams

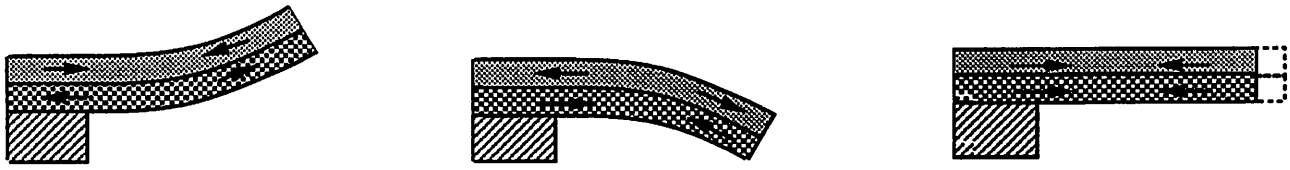
If two films with different residual strains are deposited onto a wafer, patterned into a cantilever beam, and freed from the substrate, the beam will curl to relieve the stress as shown in Fig. 4. If both films have the same residual strain, the composite beam will simply lengthen or contract while remaining flat. If the top film is more compressive, the beam will bend downward; conversely, if the top film is more tensile, it will curve up. Note that only the *difference* in residual strains matters here: for example, both films can be tensile, but if the top one is more tensile, the beam bends upward.

Because the bending moment due to the residual stresses is same at all points along the length of the beam, its curvature is constant, so the beam takes the shape of a circle of arc or a circle of radius ρ. The difference in residual strains can be calculated from this radius of curvature and film parameters. It does not suffer from the range-of-strain limitations that the single-layer test structures had: residual strain difference of any magnitude can be calculated provided ρ and material parameters can be measured.

Our cantilevers were designed to be fabricated alongside the single-layer structures from the same films. They are attached to a large pad which is raised from the substrate by a PSG pedestal. In addition to the composite cantilevers, cantilevers were made of each layer singly in order to look at self-curvature. Self-curvature is due to residual stress variation through the thickness of a single layer and must be accounted for in the residual stress calculations.

The cantilevers had to be wider than the diameter of a focused laser spot and spaced far enough from each other to avoid causing interference with the reflected laser beam (experiment to be described later). As a compromise between width and die space consumed, we chose 105μm-wide cantilevers with a 195μm beam-to-beam spacing. In order to make it possible to use different ρ measurement techniques, the cantilever lengths were varied logarithmically with lengths of 100, 200, 500, 1000, 2000, and 5000μm. The length variation was also intended to allow optical observation of downward vertical deflection: if the beams bend downward, all beams longer than a certain length will touch the substrate.

Square 5x5μm holes spaced 5μm apart (Fig. 5) were put in the cantilevers to reduce the undercutting distance during the etch step to free the structures (see Process Design). This does not affect the radius of curvature, because the bending moment across any width divided by the width of material resisting bending is constant. To demonstrate this, beams 20μm wide and 2000μm long were placed



(a) Top more tensile (b) Bottom more tensile (c) Both equally tensile

Figure 4. Two-layer cantilevers bending

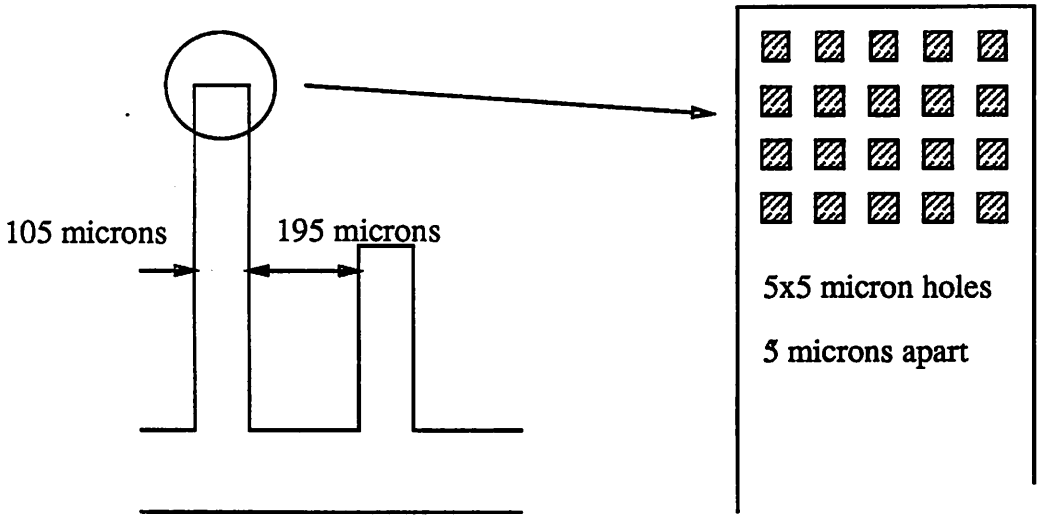


Figure 5. Close-up of cantilever etch holes

next to the rest of the cantilevers for comparison. Another set of cantilevers 100μm wide with the same 100 to 5000μm lengths was laid out but not used.

Figure 6 is our chip layout showing the doubly clamped beams, rings, and cantilevers made out of each single layer of film, and the two-layer cantilevers.

In the next sections, the process, methods of measuring the radius of curvature, and using the information to calculate residual stress are described.

2.0 Processing

2.1 Blanket-Deposited Films

Our "regular" process, in which all films are blanket-deposited, uses a three mask set to define test structures made from two different single layers (doubly clamped beams, etc.) and the composite cantilevers on the same die. The two films are each about 0.5 to 1μm thick. The different films used and their actual thicknesses are listed in the Results section.

The process starts with 4-inch (100) silicon wafers. About 3μm of PSG, which will later form the pedestals, is deposited and densified (Fig. 7a). Next, the first layer of thin film is deposited, patterned with Mask 1:PRO1, and etched, leaving behind large rectangles of Film 1 where one set of single layer and the two-layer structures will be (Fig. 7b). The etch can be plasma or wet, whichever is easiest for the films involved.

If Layer 1 is polysilicon or nitride, the wafer is dipped in HF immediately before the next deposition step to remove any native oxide that has grown (oxide between the two layers could cause delamination during the final PSG etch step). The second layer of thin film is then deposited over the entire wafer (Fig. 7c). Lithography using Mask 2:CANT, which contains the actual test structures, is performed. Layer 2 is plasma etched, then Layer 1 is plasma etched using the same photoresist, resulting in a self-aligned structure (Fig. 7d). Plasma etching is necessary for straight sidewalls, ensuring that the processed dimensions do not deviate much from the layout dimensions. A Layer 2 etch must be chosen that does not rapidly undercut Layer 1 when it reaches it. For example, the Technics-c nitride etcher etches polysilicon rapidly and fairly isotropically and is thus inappropriate). Likewise, the Layer 1 etch must not undercut Layer 2. The photoresist must be thick enough that some is left after two plasma etches; double and triple thick resist is used in our processing.

The final mask step, using Mask 3:PRO2, covers the entire wafer with photoresist except where the Layer 1-only structures are located. An etch that attacks only Layer 2 is used to remove the top

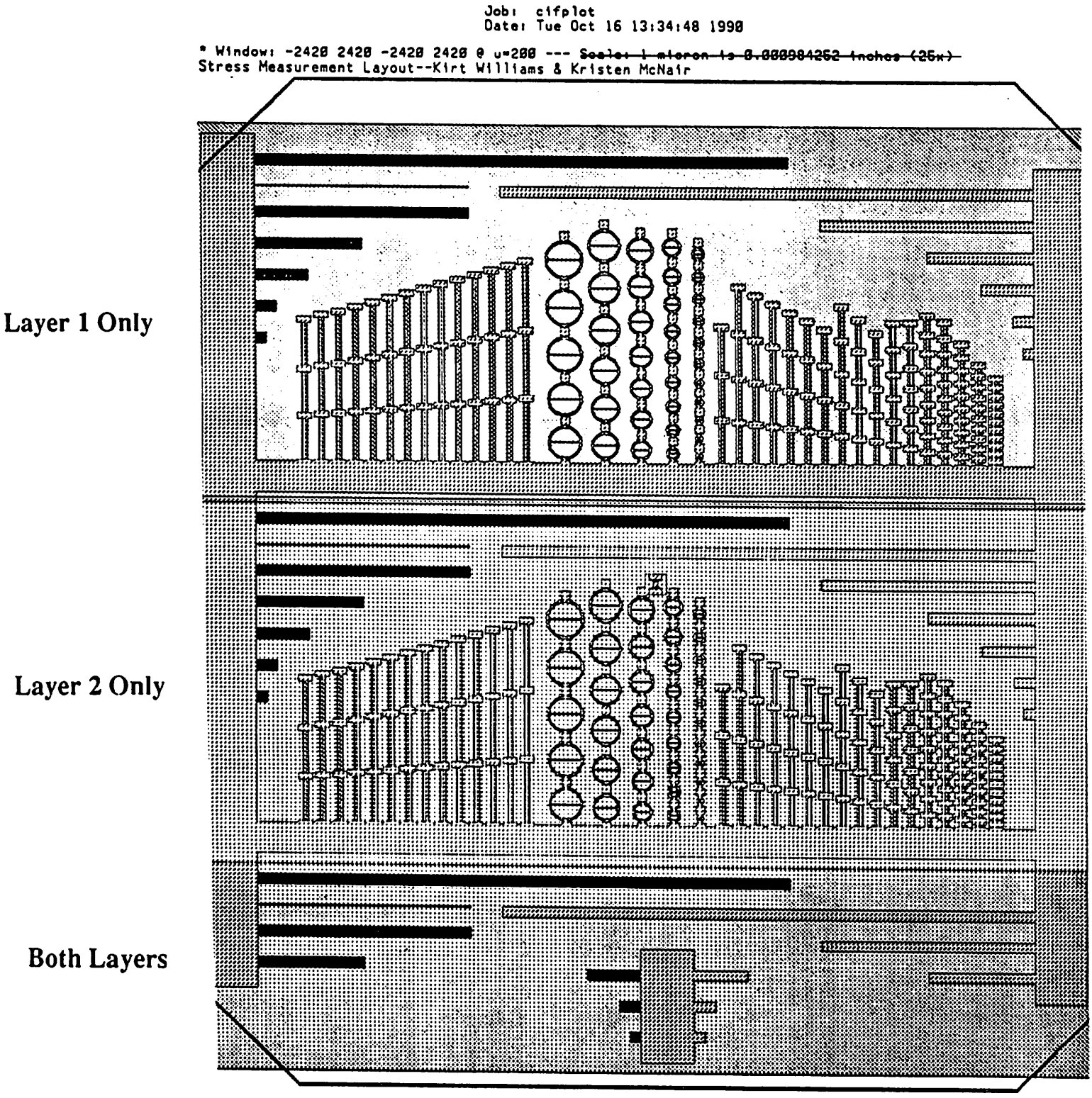


Figure 6: Mask Layout of Cantilevers and Strain Diagnostic Structures

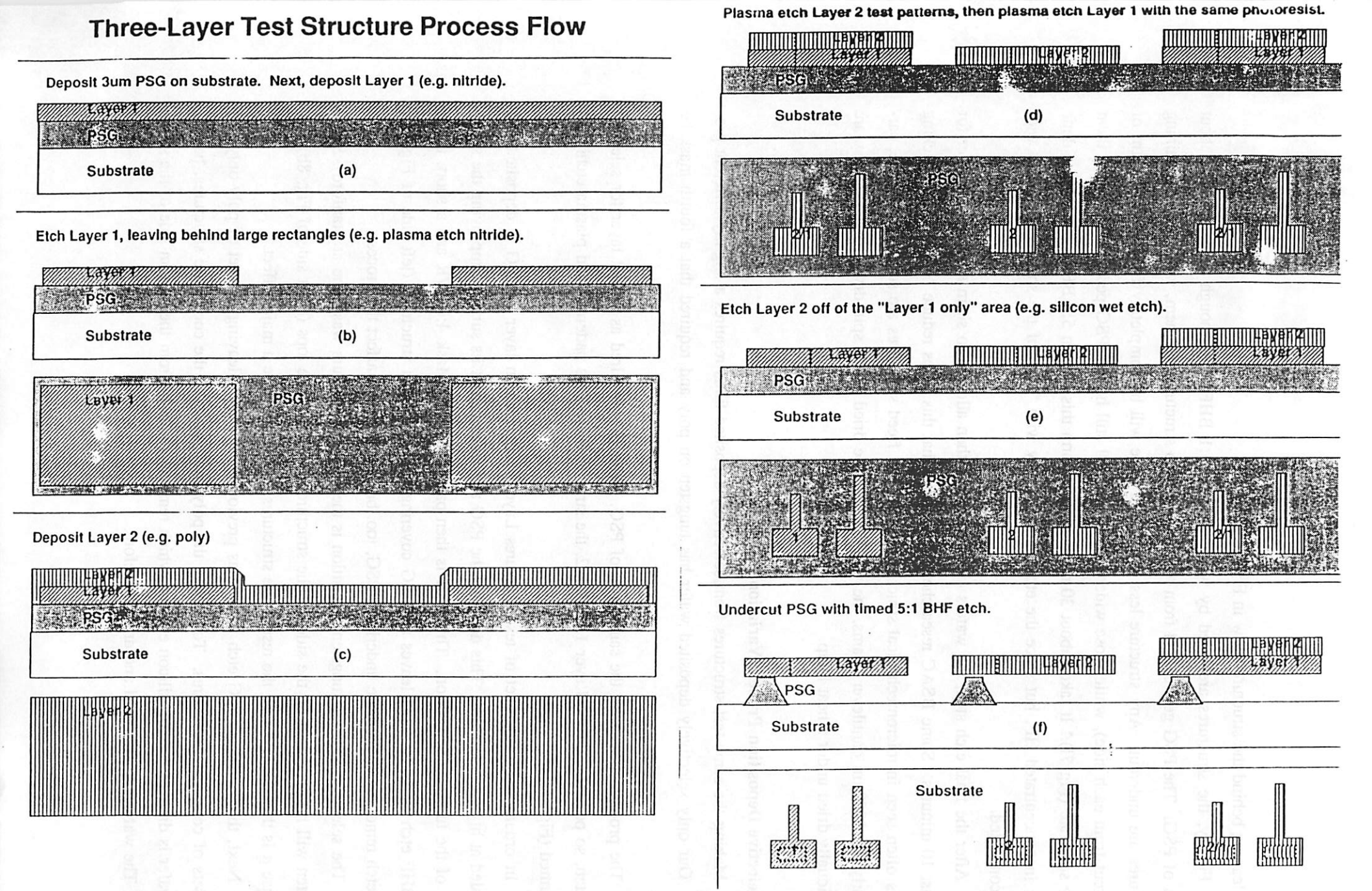


Fig. 7: Three-Layer Test Structure Process Flow

film, leaving behind the structure show in Fig. 7e.

Finally, the structures are freed by wet etching in 5:1 BHF long enough to go through about 13μm of PSG. The PSG gets etched from all areas with no structures on them, while areas containing structures are undercut. Any structure less than 20μm wide will be completely freed (about 10μm of undercut from each side), while those wider than 30μm will still have a PSG pedestal anchoring them to the substrate (Fig 7f). It takes about 30 minutes to perform this etch in 5:1 BHF. It could be done faster in concentrated HF, but since the etch rate will vary with "age" of the solution, it is not very well controlled.

After the final etch step, the wafers are rinsed well, then allowed to sit in stagnant DI water for at least 10 minutes. Some BSAC researchers have found that this helps reduce "stiction," the sticking that is often seen in micromechanical structures. Since the freed structures are somewhat fragile, particularly the 5000μm cantilever beams, the wafers cannot be dried in the spin dryers; instead, they are individually dried under a heat lamp.

2.2 Selective Deposition Process Variation

Making the same test structures from selectively deposited films requires a slightly different process. Our only selectively deposited wafer has tungsten on poly and required that a fourth mask be made.

The process starts with the same 3μm of PSG. Silicon is required as a seed layer for selective tungsten, so poly is used for Layer 1. Mask 2, the structures mask, is patterned and plasma etching is performed (Fig. 8a).

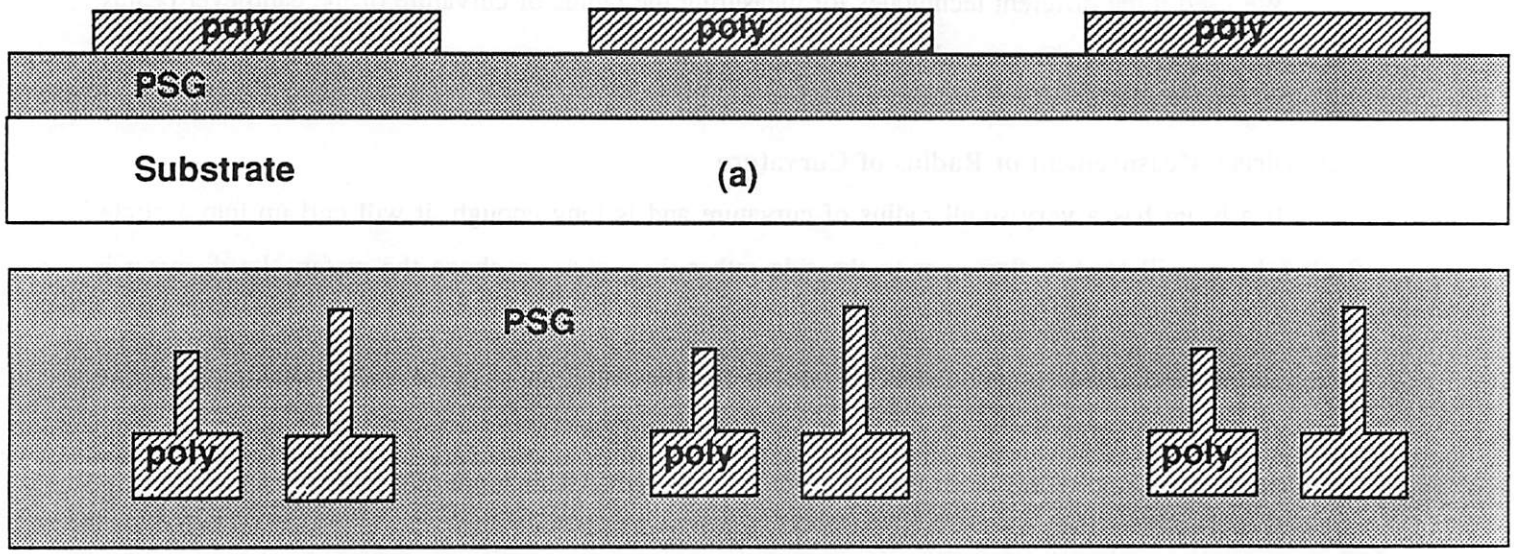
In order to keep one set of test structures Layer 1-only, a thin layer of PSG is deposited and annealed at high temperature. This densifies the PSG and smooths out its surface, improving the selectivity of the tungsten deposition. This PSG is then patterned using Mask 4:COVR and a short, timed 5:1 BHF etch is done. This leaves the PSG covering one set of test structures (left side of Fig 8b). This etch removes some of the underlying PSG, too, but this does not affect the process.

The selective LPCVD tungsten deposition is then performed (ours was done at Stanford). Some tungsten will be deposited on the sides of the structures as well as the tops (right side of Fig. 8b), but because it is thin compared to the rest of the structures, it will not have a major effect.

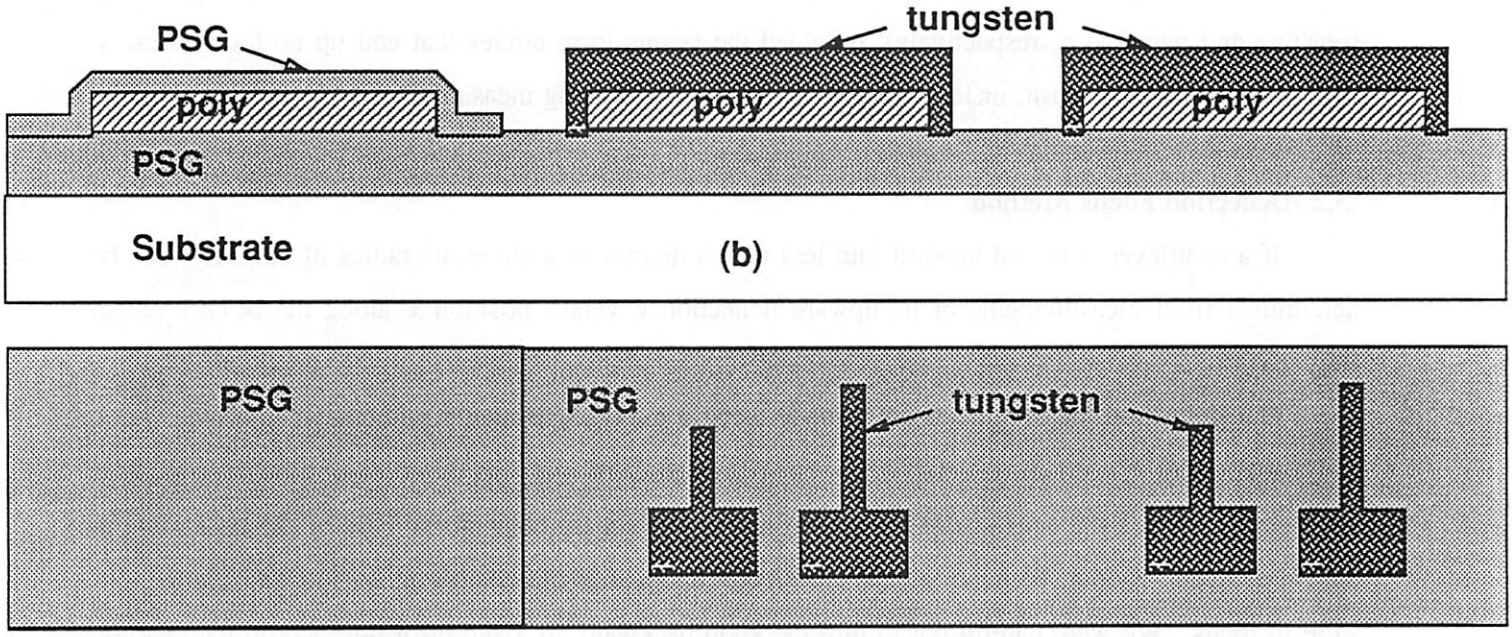
Next, the sacrificial PSG etch is done as previously described, leaving one set of poly-only and two sets of composite structures. To remove the poly from some of the composite structures, half of the wafer is dipped in a wet silicon etchant, which removes the poly from the bottom side of the structures. The wafer is then rinsed and dried as before.

Selective Deposition Process Variation

Pattern test structure on Layer 1 (poly)



Cover poly-only test structures with 2000A PSG to inhibit tungsten deposition
Selectively deposit tungsten



Do timed 5:1 BHF etch of PSG as usual.

Dip half of wafer in silicon etchant to remove poly from the underside of the tungsten.

Figure 8: Selective Deposition Process Variation

3.0 Measuring Radius of Curvature

We used three different techniques for measuring the radius of curvature of the cantilever beams: the direct, the deflection/focus, and the laser/angle methods.

3.1 Direct Measurement of Radius of Curvature

If a beam has a very small radius of curvature and is long enough, it will curl up into a circle. Such a beam will tend to flop over to the side rather than stick up above the wafer. Its diameter is then easily measured directly.

Our diameter measurement method was to put the wafer on a probe station and position two probe tips directly opposite each other above a circular beam (Fig. 9a). To ensure that we were measuring the curling of the beam due to residual stress rather than some other local phenomenon, we checked that other circles on the wafer had the same diameter. The probe tip spacing was then compared to structures on the wafer of known dimension to get a diameter reading (Fig. 9b).

This method is accurate from extremely small radii (about 10μm) up to about 1500μm, the radius of a 5000μm beam curled into a semicircle. It works for both upward and downward bending (positive and negative ρ, respectively), provided the beams form circles that end up on their sides. It is accurate to ±25 or 50μm, or less, depending on the radius being measured.

3.2 Deflection/Focus Method

If a cantilever is curled upward into less than a quarter of a circle, its radius of curvature can be determined from measurements of its upward deflection v versus position x along the beam's length (Fig. 10a).

Our deflection measurements were made using a microscope with a micrometer fine focus knob and a shallow (a few microns) depth of focus. A curved cantilever was observed in a microscope field large enough to see the entire beam. We focussed on the square holes at the base of the beam, noting the micrometer readout, then turned the fine focus knob until the beam's midpoint, and finally end, were in focus. We were careful not to turn the knob backward, to avoid error due play in the gearing. From the change in micrometer reading, deflections $v(x)$ at $x=0$, $x=\frac{1}{2}L$, *and* $x=L$ were determined. We used this data in two different ways to relate the $v(x)$ data to ρ.

The radius of curvature of any curve is related to the first and second derivatives at a point by

$$\frac{1}{\rho} = \frac{d^2v/dx^2}{[1+(dv/dx)^2]^{3/2}}. \qquad (3)$$

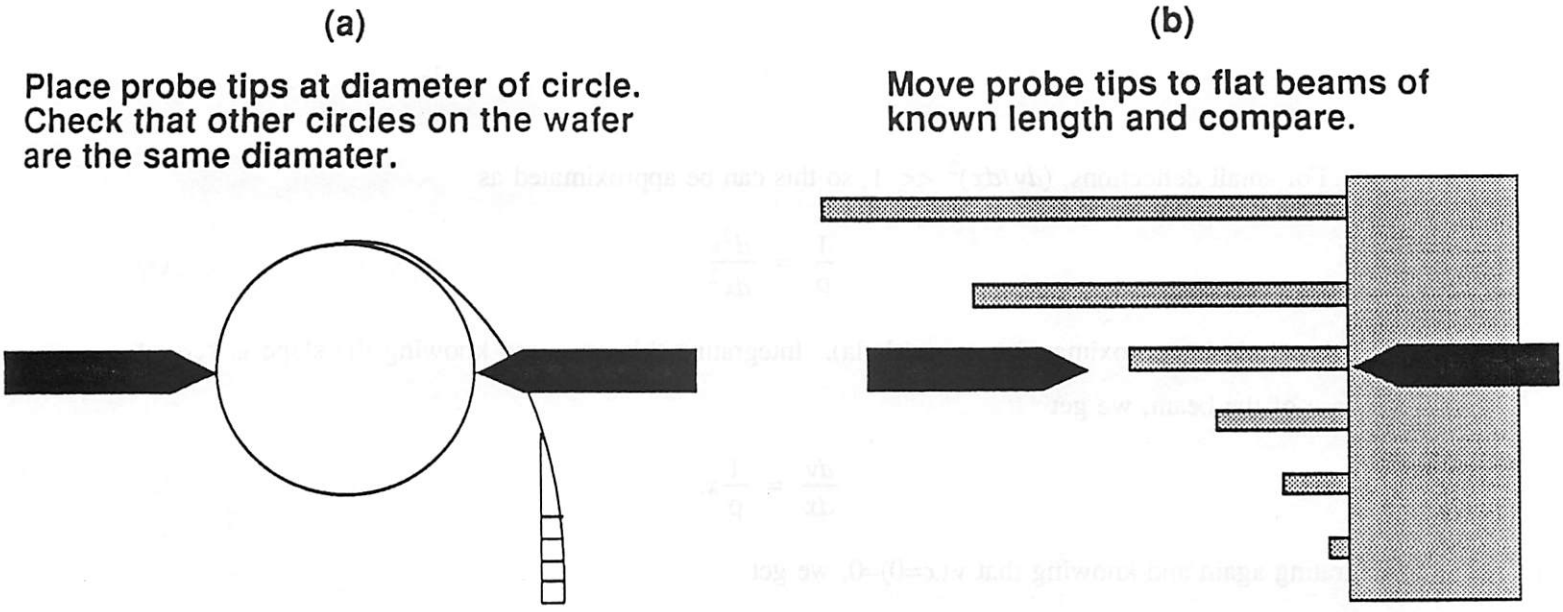
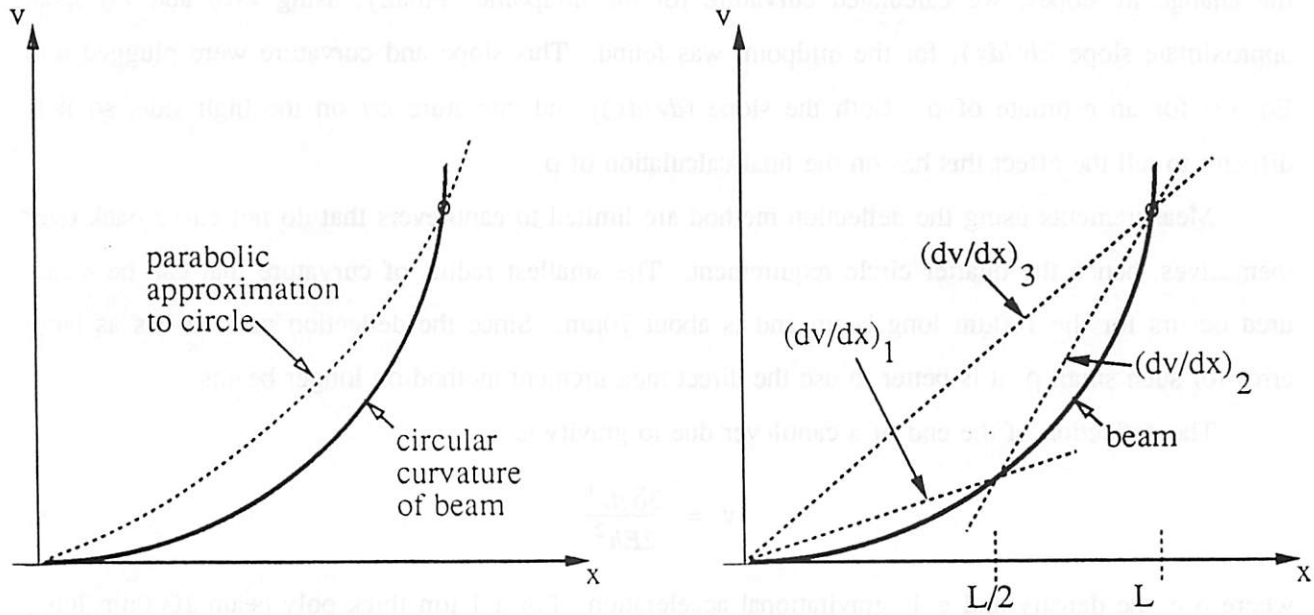


Figure 9: Direct Measurement of Long Beams Curled into Cylinders



(a) Side view of cantilever deflection (exaggerated)

(b) Slope approximations

Figure 10: Measuring Deflection of a Cantilever

For small deflections, $(dv/dx)^2 << 1$, so this can be approximated as

$$\frac{1}{\rho} = \frac{d^2v}{dx^2} \tag{4}$$

(i.e., the circle is approximated as a parabola). Integrating this once and knowing the slope is zero at the base of the beam, we get

$$\frac{dv}{dx} = \frac{1}{\rho}x. \tag{5}$$

Integrating again and knowing that $v(x=0)=0$, we get

$$v = \frac{1}{2\rho}x^2 \tag{6}$$

so

$$\rho = x^2/2v. \tag{7}$$

This method is commonly used in mechanics and is accurate when the slope is small enough to have an insignificant effect on ρ (for 1% error, the slope must be <0.08). If the slope does cause an error it results in an underestimation of ρ (that is, ρ is larger than the calculated value).

For larger deflections, we approximated Eq. (3) in a different way: We calculated approximate slopes between $x=0$ and $x=½L$ $(dv/dx)_1$, and between $x=½L$ and $x=L$ $(dv/dx)_2$ (Fig. 10b). From the change in slopes, we calculated curvature for the midpoint. Finally, using $v(0)$ and $v(L)$, an approximate slope $(dv/dx)_3$ for the midpoint was found. This slope and curvature were plugged into Eq. (3) for an estimate of ρ. Both the slope $(dv/dx)_3$ and curvature err on the high side, so it is difficult to tell the effect this has on the final calculation of ρ.

Measurements using the deflection method are limited to cantilevers that do not curve back over themselves, hence the quarter circle requirement. The smallest radius of curvature that can be measured occurs for the 100μm long beam and is about 70μm. Since the deflection method has as large error for such small ρ, it is better to use the direct measurement method on longer beams.

The deflection of the end of a cantilever due to gravity is

$$v = \frac{3\delta g L^4}{2Eh^2} \tag{8}$$

where δ is the density and g is gravitational acceleration. For a 1 μm thick poly beam 1000μm long, this is about 0.2μm. For L=2000μm it is 4μm, but for a 5000μm beam it is 140μm due to the L^4 term. Thus, for the effect of gravity to be neglected, L must be kept below about 2000μm.

The largest ρ that can be calculated is limited by detection of the upward deflection of the 2000μm beams. Under the appropriate microscope objective, deflections on the order of 10μm can be measured, corresponding to a ρ of 80cm.

3.3 Laser/Angle Method

The laser method was intended to measure large radii of curvature by determining the slope of a cantilever as a function of position. The approximation in Eq. (4) applies, so Eq. (5) is valid. Referring to Fig. 11, it is seen that the tangent of the angle θ at given point x is the same as the slope. Thus,

$$\tan\theta = \frac{dv}{dx} = \frac{1}{\rho}x. \tag{9}$$

$1/\rho$ can be determined from a measurement of two values of $\theta(x)$ from

$$\frac{1}{\rho} = \frac{\Delta(tan\,\theta)}{\Delta x}. \tag{10}$$

A measurement is made using the setup shown in Fig. 11 and Photograph 1. The laser is a battery-powered semiconductor laser pointer, with λ≈650nm (ruby red), spot size of a few millimeters, and partial spatial coherence over several meters. The laser is focused through the lens and goes through an N.D. filter. The focus is adjusted for the smallest spot at the wafer (depth of focus is a few millimeters). The iris is then closed to make an even smaller spot of about 90μm, narrow enough to fit on the cantilevers.

This spot is then reflected off a clean wafer up the Reticon array, a linear array of 1024 photodiodes with 25μm spacing. The Reticon array displays the intensity versus position on a personal computer monitor. to keep ambient light out, a red bandpass filter was placed over the array. To reduce the reflected light the laser hitting the photodiodes, a slit made of razor blades was placed over that. This slit was also intended to make the setup less sensitive to the reflected beam moving perpendicular to the usual direction of motion. Since the spot is larger that the slit is wide, the spot scanning across the slit should let about the same amount of light through.

When an actual measurement on a cantilever is made, it is aligned with micrometers on an x–y–ϕ stage. The laser spot can be seen through the microscope due to the square hole causing non-specular reflection. If the reflected beam saturates the detector (Fig. 12), a darker N.D. filter is installed; if the reflected waveform is too weak, a lighter one is used to get the largest non-saturating signal.

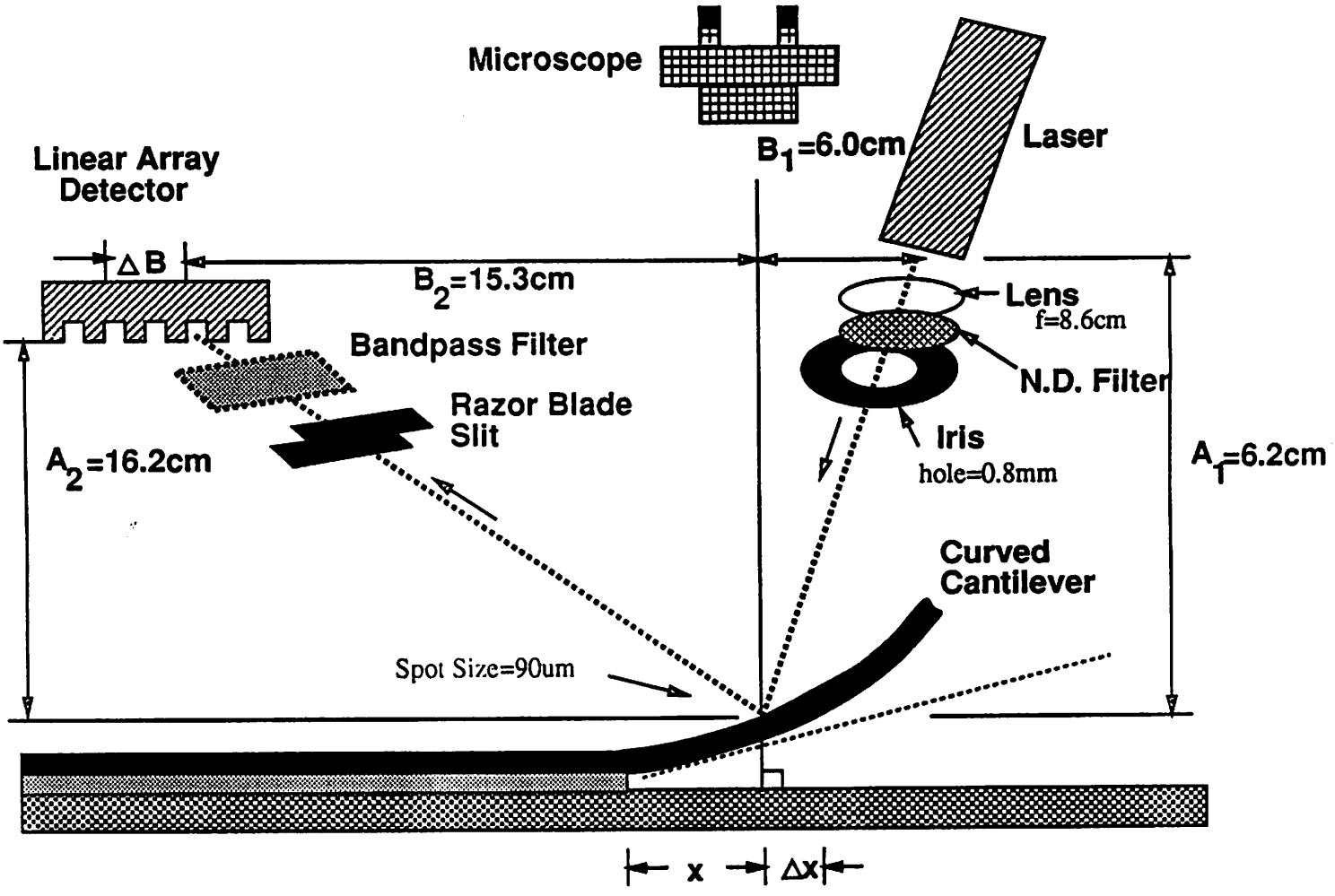


Figure 11: Curvature Measurement Setup

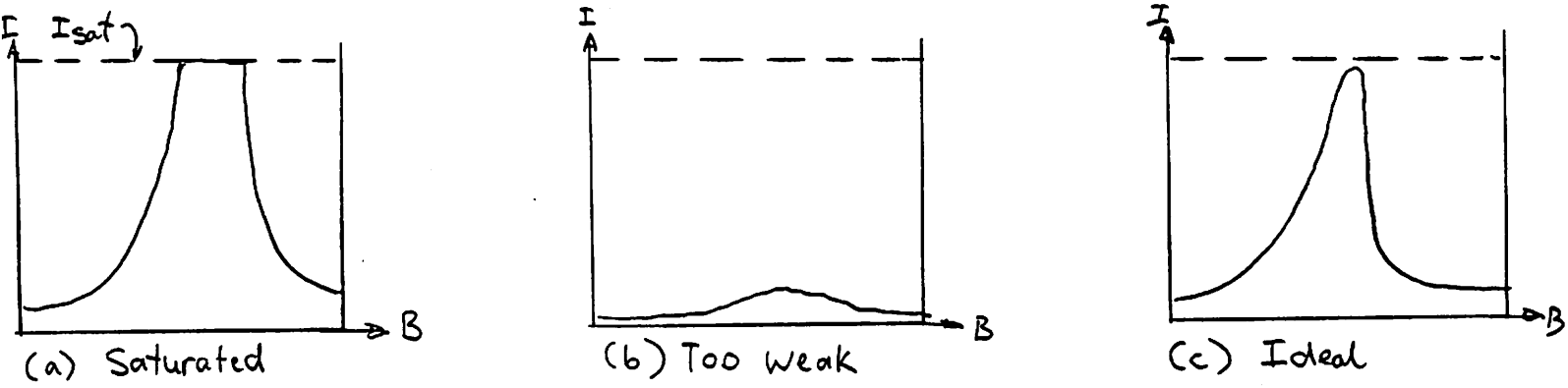


Fig. 12: Reticon Array Output

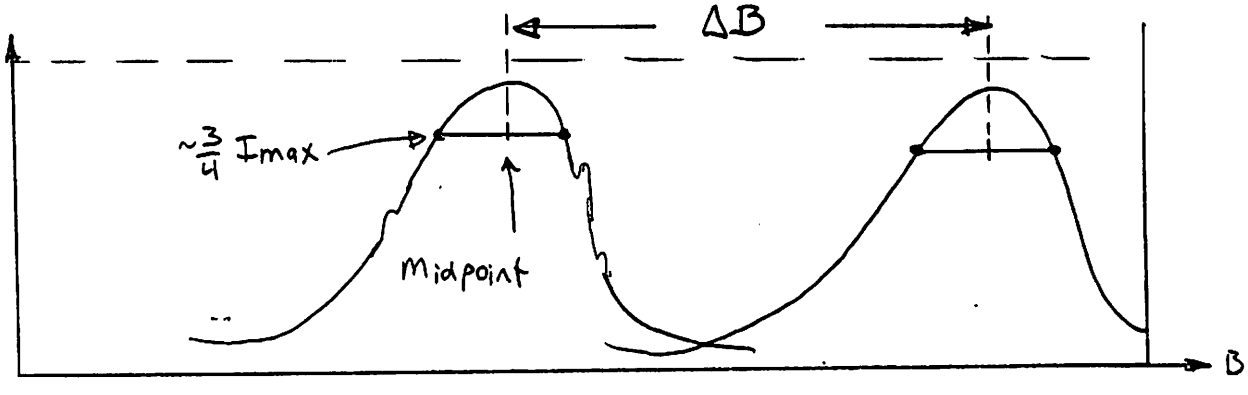


Fig. 13: Measuring Detected Beam Movement ΔB

The setup is now ready for a measurement. The laser spot is started at the base of the cantilever (x=0) and the position B of the reflected beam is noted as follows. Assuming the beam intensity has a Gaussian-type shape, its center is taken to be the midpoint of two points of about 3/4 maximum intensity (Fig. 13). The points selected are away from anomalies in the waveform. Finally, the cantilever is moved a known distance Δx by turning the micrometer stage and the spot movement ΔB is found.

Using the geometry, algebra, and approximations shown in Appendix A, we can determine the radius of curvature from:

$$\frac{1}{\rho} = \frac{\Delta B}{2A_2\Delta x}. \tag{11}$$

3.3.1 Limitations of the Laser Method

Ideally in the laser method, a narrow, Gaussian-shaped waveform would move a long distance ΔB for a small cantilever movement Δx. Several design tradeoffs affect this. The first is spot size. Ideally, the beam would emanate from the laser with a small diameter and no divergence, eliminating the need for the lens, but this would violate laws of diffraction. Since the lens is required, the converging beam diverges after reflection, so that the detected spot grows with distance from the wafer to the detector A_2 (Fig. 9). Thus, a long focal length and a short A_2 are desired for a small detected spot. We only had one lens available to us, so the focal length was set. This requirement for a short A_2 is countered by the need for a large spot deflection ΔB, which is also proportional to A_2.

In our setup, we were able to assuage this problem by using the iris to get the initial beam diameter down to about 0.8mm. This allowed the detector to be raised as high as the setup would tolerate. The spot at the detector should then be $0.8\text{mm}\cdot A_2/A_1 = 2\text{mm}$, a small fraction of the 25mm of detector area available. Using the small aperture of an iris can potentially cause problems with diffraction effects. Since the iris opening here is much larger than the laser wavelength, diffraction is not a problem.

Another limitation occurs when the cantilever curls up so much that it acts as a spherical mirror (Fig. 14). The beam will be focused through the mirror's focal point (approximately at the center of curvature), after which it will diverge and be very large by the time it reaches the detector. As the cantilever is moved, the spot will move, but it may not be detectable. Thus, focusing effects set a lower limit for the detectable radius of curvature. We estimate that a focal point about three times the wafer-to-detector distance is safe, which is about 70cm.

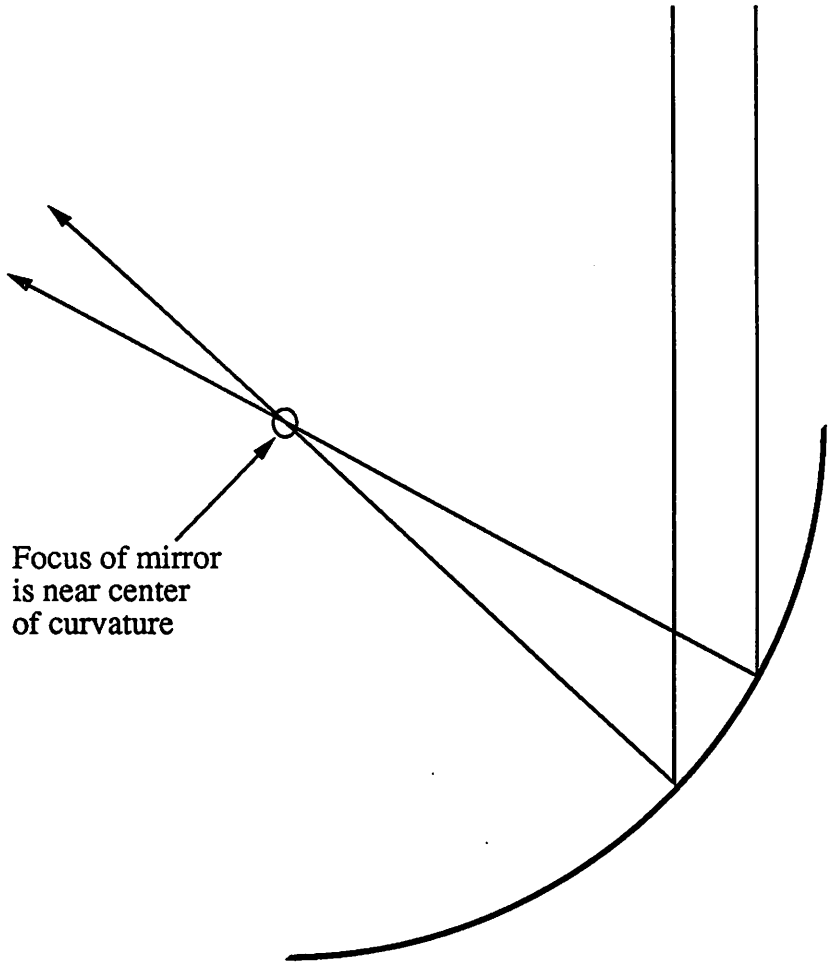


Figure 14. **Cantilever with a small radius of curvature acts like a spherical mirror**

The maximum detectable ρ is limited by detection of the spot movement ΔB. Noise alone can account for about three pixel of ΔB variation, so at *least* a 10-pixel (250μm) spot movement is desirable. This corresponds to a maximum measurable ρ (using a 2000μm beam) of 2.5 meters.

4.0 Calculation of Residual Stress for Cantilevers

4.1 Self Curvature

As shown in Appendix B, cantilever beams of a *single* film can curl up by themselves due to residual stress variation through their thickness. The stress gradient can be found from

$$\sigma' = E'/\rho. \tag{12}$$

This self bending represents a bending moment, which is related to ρ by

$$\frac{1}{\rho} = \frac{M}{E'I}. \tag{13}$$

It is seen that for larger M, the curvature increases (1/ρ decreases), while for a larger E' (stiffness) or I (moment of inertia; geometrical resistance to bending) it decreases.

The radius of curvature of a two-layer cantilever beam has been related to the residual stresses by Judy [9]. His equation does not take into account the effect of self bending, so that is done in Appendix C by subtracting off the bending moments due each film by itself to get an effective radius of curvature for bending due to the residual stresses alone. The resulting method and equations are as follows:

Measure ρ_1, ρ_2, and ρ_t (radii of Film 1, 2, and the composite)

Calculate $E_1'=E_1/(1-\nu_1)$ and $E_2'=E_2/(1-\nu_2)$ (effective Young's moduli)

Measure h_1 and h_2 (thicknesses of the bottom and top films)

Calculate $m=E_1'/E_2'$, $n=h_1/h_2$, $h=h_1+h_2$

$K_1=1+4mn+6mn^2+4mn^3+m^2n^4$

$K_2=[n(n+1)^2]^{-1}$, $K_3=(3m+K_1K_2)/6m$

The effective radius of curvature is

$$\rho_r = \left[\frac{1}{\rho_t} - \frac{mn+1}{K_1}\left[\frac{mn^3}{\rho_1} + \frac{1}{\rho_2}\right]\right]^{-1} \tag{14}$$

This is related to the residual stresses by

$$\sigma_2 - \frac{\sigma_1}{m} = \frac{K_3E_2'h}{\rho_r}. \tag{15}$$

It is seen that if $\sigma_2=\sigma_1/m$, or alternately, if $\varepsilon_1=\varepsilon_2$, that there is no bending. By selecting a stress-free bottom layer, this reduces to

$$\sigma_2 = \frac{K_3 E_2' h}{\rho_r}. \tag{16}$$

As usual, $\varepsilon_2=\sigma_2/E_2'$.

Temperature compensation is necessary to compare the residual strain of a free beam with one clamped at both ends to a rigid substrate. A free beam can change its strain (that is, its length) with temperature after being deposited as $\varepsilon=\alpha_f \Delta T$, where α_f is the coefficient of thermal expansion of the film and ΔT is the decrease in temperature from deposition conditions to ambient. A doubly clamped beam on the other hand will be strained by an amount $\alpha_{su}\Delta T$ by the substrate, but can relieve strain by $\alpha_f \Delta t$, so the total temperature contribution to the residual strain is the difference $(\alpha_{su}-\alpha_f)\Delta T$. Because most thin-film structures in sensor and actuator work are rigidly attached to the substrate, as the doubly clamped beams are, we will refer all residual strains to this condition. This is done for the cantilevers by *subtracting* the previous expression:

$$\Delta\varepsilon_{temp} = (\alpha_f-\alpha_{su})(T_{ambient}-T_{dep}). \tag{17}$$

For nitride deposited at 835°C, this is $245x10^{-6}$; for CVD tungsten deposited at 285°C, $\Delta\varepsilon_{temp}=-451x10^{-6}$. Single crystal silicon has an α that varies with direction, but that is neglected here. For our calculations we take α to be the same for poly and the substrate (not a bad approximation), thus polysilicon on silicon has no temperature strain effect. It is also noted that α is a moderate function of temperature, so this should be taken into account when precision is required.

VII. RESULTS AND INTERPRETATION

5.0 Process Variations

Our project involved 13 wafers. Wafers 1-12 started with about 4.3μm of PSG. During subsequent processing, most received various combinations of half or one micron films of polysilicon using the SDOPOLYG recipe or low-stress nitride using the SNITD.V recipe at 835 °C with a gas flow ratio of $SiCl_2H_2$:NH_3 = 64:16 sccm. Two others received aluminum sputtered in the CPA and CVD tungsten. Wafer 13 was donated to the project with CPA-sputtered tungsten over PSG (one layer only). Table 1 is a history of the films used on each wafer. It also lists failures, where applicable.

The thickness of the poly and nitride films were determined from dummy wafers measured on the Nanospec. The CVD tungsten thickness was reported by the processors at Stanford; they use a step height change measurement on the wafer itself. The sputtered aluminum and tungsten thicknesses were measured by the step height on a patterned dummy wafer.

5.1 Failure Mechanisms

As seen in Table 1, our structures on seven wafers were rendered unusable for six different reasons. Wafer 1 had a mask error that was spotted too late. Wafer 7 was broken when removed from the tylan furnace before exposing Mask 2. Since the automatic wafer handling system used in the wafer stepper requires whole wafers, this one had to be scrapped. Wafer 6 was chopped into quarters by the MTI automatic developer before developing before developing the final mask. It was salvaged by doing hand developing.

We didn't know whether 5:1 BHF would etch aluminum. We discovered that it will during the 30 minute sacrificial PSG etch step. The aluminum etching was not uniform, but rather looked very messy under inspection. Small chunks, presumably of Al, were left in the HF tank after the etch.

Of the over 200 structures of sputtered tungsten looked at on Wafer 13, more than 99% had visible cracks due to a high tensile residual stress (Photo 2). As would be expected, all cracks were initiated at points of high stress concentration. This tungsten appeared to be very brittle because of cracking without any evidence of plastic deformation and because when touched with a probe tip, pieces tended to chip off.

The fifth reason for structures not working was an inadequate plasma etch of the nitride forming the bottom layer of the structures on Wafers 9 and 10. Although the etch rate had been characterized and an overetch was done, a thin layer of nitride remained. Many etch steps were checked for completion by probing a conducting layer for an open or short circuit, but this could not be done here. Only

Table 1: History of Project Wafers

Wafer	Layer 1	Layer 2	All Structures Work?
1	.49μm poly	incomplete	no
2	.49μm poly	.41μm CVD W	yes
3	.49μm poly	.41μm nitride	yes
4	.49μm poly	.88μm nitride	yes
5	.49μm poly	incomplete	no
6	.49μm poly	.66μm aluminum	no
7	.82μm poly	.41μm nitride	no
8	.82μm poly	.88μm nitride	yes
9	.48μm nit	.47μm poly	no
10	.48μm nit	.86μm poly	no
11	.92μm nit	.97μm poly	yes
12	.92μm nit	.86μm poly	no
13	~1μm CPA W	none	no

Problems:

Wafer 6: wafer broken by MTI after last lithography; Al was destroyed by BHF etch

Wafer 7: wafer broken when removed from tylan furnace before mask2 patterning

Wafer 9: nitride not etched long enough for mask2; nitride structure not freed

Wafer 10: nitride not etched long enough for mask2, but a few dice were okay

Wafer 12: poor adhesion of nitride to PSG pedestal

Wafer 13: all structures cracked

when the PSG wet etch was performed was it discovered that the nitride blocked the BHF in the nitride-only and two-layer regions of each die, while the PSG was etched from the poly-only regions (refer to Figs. 7e and 7f). We could not salvage the wafers by etching the nitride more because the photoresist mask had already been removed. Even etching a little nitride from the entire wafer wouldn't have solved this problem, as repeating the BHF etch would have completely removed the PSG pedestals in the poly-only region.

After learning that the nitride etch could be a problem, we found that it could be checked by doing a 5 minute 5:1 BHF etch. If undercutting of the nitride structures was seen, the nitride had been cleared.

The sixth and final failure seen was poor adhesion of the nitride rings and beams to the PSG pedestals on Wafer 12. When pushed with a probe tip, these structures easily detached from their anchors while remaining intact. The PSG did not appear to be overetched. The cantilevers, with much larger-area pedestals, were unaffected. Wafer 11, which received almost identical processing, did not exhibit this problem, either. Its source is a mystery.

5.2 Freeing the Structures

At first glance almost all the doubly clamped beams longer than a certain length on all wafers *appeared* to be buckled, due to changes in color at the middle of the beam. This is demonstrated on Photo 3 for nitride beams on Wafer 3. It turned out that the green regions were where the middles of the beams were simply stuck down to the substrate after the rinse and dry steps, perhaps as droplets of water under the beams shrunk and pulled them down. The shortest beams could not be deflected far enough to get stuck. Upon gently touching the stuck beams with a probe tip, most sprang free and pulled taut, becoming the same uniform brown color as the other "unstuck" beams.

This phenomenon was also seen for the rings and cantilevers. Thus before measurements could be made for almost any structure, they had to be probed. The rings were freed as easily as the beams, although some were delicate and preferred to break. Most cantilevers took a lot of prying to get them free. Those that curled upward remained free after probing, while most that bent downward tended to stick down, their colors returning to the "stuck color." Some beams with significant downward curvature were able to prop themselves up and even curl into circles to remain free.

Most cantilevers were found to be incredibly resilient. The 5000μm ones could be bent in two and let go to spring back to their original positions, with no permanent effects observed.

5.3 Data and Interpretation

5.3.1 Single Layer Test Structures

The sacrificial PSG etch undercut all structures' pads about 10μm. To account for this, 20μm was added to the drawn lengths of the doubly clamped beams for an appropriate effective length. No adjustment was necessary for the rings; the effect on the long cantilevers in negligible.

The left side of Table 2 lists observed data for the single-layer test structures. It is seen that in most cases neither the doubly clamped beams nor the rings with beams buckled. Thus, the residual strain is less negative than a certain value that is calculated from the length of the longest beam used in Eq. 1, and is less positive than a different value calculated using the radius of the the largest ring in Eq. 2. This indicates that the residual strain is near zero (but can be either positive or negative), as one would expect from low-stress nitride and typical poly. The residual strain ranges and the corresponding residual stresses are tabulated in Table 3.

Wafers 9 through 12, for unknown reasons, had poly that was so compressive that all the doubly clamped beams buckled (Photo 3). The polysilicon on these wafers was deposited during the same run. This was accomplished by putting down 0.47μm on all four wafers, then removing two and rerunning the same deposition program. The minimum magnitudes of strain for the wafers vary by a factor of 3.8 due to the factor of 2 variation in poly thickness.

As expected from Howe and Muller [2], ripples were seen in places where the poly had been undercut, forming an overhang. The wavelength of the ripples could easily be measured, but the amplitude couldn't be accurately determined without the use of an SEM, so no strain calculations were made using their technique.

The CVD tungsten on Wafer 2 was so tensile that all the rings but the smallest two or three on each die had their beams buckled. Upon close inspection, it could be seen that those few had relieved their stresses by cracking instead. The smallest ring observed to have buckled set a lower limit for the tungsten's tensile strain, which was large enough to break all the doubly clamped beams (Photo 4).

The rings and beams on all the wafers (except 13, which had cracking) performed as expected for the residual strains that were seen in the processing we did. Unfortunately, there were no residual strains in the ranges the structures were designed for, so the intended behavior of beams buckling for lengths or radii longer than a critical value was not seen and proper operation could not be verified. This was somewhat surprising, given the orders of magnitude over which buckling/not buckling could have been seen.

Table 2: Observation under the Microscope

Wafer	Layer	Rmeas	Rcalc	Rest
2	poly	bends down slightly		<0
	W	160μm+/-25μm	>114μm	175+/-35μm
	both	150μm+/-25μm	>114μm	175+/-35μm
3	poly	too large	>7800μm	13,700+/-700μm
	nit	very little seen		~0
	both	too large	>2000μm	3180+/-160μm
4	poly	bends down		<0
	nit	bends down		<0
	both	bends down		
8	poly	too large		<0
	nit	very little seen	>9300μm	~0
	both	bends down	<0	
9	nit	bends down slightly		
	poly	bends down slightly		
	both	bends down slightly		
10	nit	bends down	<0	
	poly	bends down	<0	
	both	bends down	<0	
11	nit	bends down		<0
	poly	-825μm+/-50μm		bends down
	both	-400μm+/-25μm		bends down
12	nit	bends down slightly		<0
	poly	-1000μm+/-25μm		bends down
	both	-400μm+/-25μm		bends down

5.3.2 Cantilevers

Radius of curvature data for the cantilever beams is tabulated on the right-hand side of Table 2. It is seen that some beams bent down and stuck to the substrate even after being raised up by a probe; therefore, they could not be measured.

Some of the beams on Wafers 11 and 12 curled downward enough to form circles that laid on their sides (Photo 6). For a given structure, there circles had the same diameter across the entire wafer, within the limit of experimental error. The poly on Wafer 12 was curled enough sideways to form a measurable semicircle.

The long tungsten and composite beams on Wafer 2 curled into helixes (Photo 7), indicating a lateral as well as through thickness stress gradient in the tungsten.

Table 2 lists directly measured values of ρ where they were small enough for this technique. Next to that, two different ρ's from cantilever deflection are listed, where upward beam deflection made measurement possible. As discussed previously, the $\rho=x^2/2v$ calculation yields a low value for ρ; these numbers are therefore listed as such. The calculation using estimates of slope are also listed. They are estimated to have less than $\pm 30\mu m$ of error due to calculation regardless of ρ, and 5% error in measurement. Thus, this approximation gets better as ρ increases. For Wafer 2, the only wafer where direct and deflection measurements were possible, the results are self-consistent.

We had hoped to be able to estimate negative radii of curvature by seeing for which lengths downward-bending cantilevers touched the substrate. Using normal optical microscopy, we never saw this occur; however, most downward-deflecting beams stuck flat to the substrate, rather than just touching at their ends.

As shown in Appendix B, if a stress gradient exists in a film, its radius of curvature is independent of the thickness. This holds true (approximately) for the poly on Wafers 11 and 12, which have ρ=-825 and -1000μm.

Table 3 contains the temperature-compensated calculated values of residual stress and strain using the radius of curvature data in Eqs. (14) and (16). It is consistent with the data for the single-layer test structures. The 3.1GPa stress measured for the CVD tungsten is the same order of magnitude as the approximately 1.5GPa the people at Stanford expected (method unknown).

The strain change due to difference in coefficient of thermal expansion, as calculated with Eq. (17), is seen to have less than a 10% effect on the uncompensated cantilever strain values for these particular cases.

Table 3:

Residual Stresses and Strains Calculated from Different Methods

Wafer	Layer	Clamped Beams & Rings		Cantilevers(T corrected)		
		ε_r	σ_r(Pa)	ε_r	σ_r(Pa)	σ_r(Pa/μm)
2	poly	-2.05E-6<ε<1.24E-5	-4.1E5<σ<2.3E6	~0		<0
	CVD W	ε>1.03E-4	σ>5.9E7	5.4E-3	3.1E9	3.6E9
3	poly	-2.05E-6<ε<1.24E-5	-4.1E5<σ<2.3E6			~1.4E7
	nit	-1.46E-6<ε<8.87E-6	-5.2E5<σ<3.1E6			~0
4	poly	-2.05E-6<ε<1.24E-5	-4.1E5<σ<2.3E6			<0
	nit	-6.66E-6<ε<4.03E-5	-2.3E6<σ<1.4E7			<0
8	poly	-6.93E-6<ε<4.2E-5	-1.4E6<σ<8.2E6			<2.1E7
	nit	-6.66E-6<ε<4.03E-5	-2.3E6<σ<1.4E7			~0
9	poly	ε<-8.04E-4	σ<-1.6E8			<0
10	nit	-7.32E-6<ε<1.18E-5	-1.90E6<σ<3.07E6			
	poly	ε<-2.68E-3	σ<-4.02E8			<0
11	nit	-7.29E-6<ε<4.42E-5	-2.5E6<σ<1.5E7	~0	~0	<0
	poly	ε<-8.04E-4	σ<-1.6E8	-4.8E-3	-9.4E-4	-2.4E8
12	nit	poor adhesion to PSG		~0	~0	<0
	poly	ε<-2.68E-3	σ<-5.3E8	-5.3E-3	-1.0E-3	-2.0E8

$\Delta\varepsilon_{temp} = \Delta T(\alpha_f - \alpha_s) = (20 - T_{dep})(\alpha_f - \alpha_s)$

Nitride: E=259 GPa, υ=0.25, E'=E/(1-υ)=345 GPa, α=2.5E-6 K^{-1}

Polysilicon: E=150 GPa, υ=0.23, E'=E/(1-υ)=195 GPa, α=2.8E-6 K^{-1}

Tungsten: E=411 GPa, υ=0.28, E'=E/(1-υ)=571 GPa, α=4.5E-6 K^{-1}

5.3.3 The Laser Apparatus

No radius of curvature data from the laser apparatus was listed in Table 2 because none was taken that we felt at all comfortable with. The reflected beam never moved as much as anticipated from other measurements. It was usually hard to tell whether the beam even moved because the intensity profile changed shape irregularly. This is demonstrated in Photographs 8, 9, and 10. Photo 8 shows a pseudo-Gaussian waveform as a starting point. If it kept this shape as the cantilever was moved, it would be quite usable, but instead extra peaks often popped up and grew, completely obscuring the original waveform (Photo 9). A few hundred microns further down the beam, the waveform might almost disappear (Photo 10). Such problems occurred for every cantilever we tried to look at.

A further difficulty was poor repeatability. Moving a blank, slightly bowed strip of wafer back and forth 1mm, the waveform center only returned to within about 5 pixel of its original position, a ΔB error of 125 μm.

As mentioned in the design section, we do not think the incident beam had been diffracted before striking the cantilever. We never saw a pattern repeat itself spatially as might expected with a diffraction pattern.

Most of the beams we looked at had relatively large radii of curvature, so focusing effects probably were not occurring.

It is possible that the cantilevers, even though they appeared flat, had curling at the edges or local wrinkling, which would have scattered the light differently at each point along its length. The setup may also have been extremely sensitive to side-to-side tilt of the wafer, which could move the reflected beam out of the slit at the detector.

Even if the problems currently plaguing this apparatus could be found and corrected, it would still not be appropriate for casual use because its setup is time-consuming and tedious. Fortunately, all but the largest radii of curvature that were intended to be detected using this method can be measured more easily and accurately using the deflection/focus method.

VIII. CONCLUSION AND RECOMMENDATIONS

Our project was designed to measure residual stresses and strains in thin films using two different methods. The films we successfully looked at were polysilicon, low-stress silicon nitride, and selectively deposited LPCVD tungsten.

The first method was patterning free-standing doubly clamped beams (sensitive to compressive strain) and rings with beams across their diameters (sensitive to tensile strain). For a minimum magnitude of residual strain that depends on geometry, all beams in these structures longer than a critical length will buckle. We made an array of these structures, varying the beam lengths and film thicknesses to achieve sensitivity to several orders of magnitude of residual strain. Even with such a wide dynamic range, each film that we looked at either buckled all or none of the test structures, which only told us that the films had more or less strain than a certain value. It did show that the low-stress nitride indeed had a low stress, and that the normally low-stress poly could have a large compressive stress.

The second method was to make a two-layer cantilever beam, with one of the materials of a low or known residual strain and the second's to be determined. Differences in residual strain caused the beam to deflect into a circular arc. Measuring the radius of curvature of the beam using one of three methods, and taking into account bending of the beams on their own, residual stresses were calculated. This technique was not limited in range like the single-layer one was, allowing us to calculate a tensile residual stress for the tungsten of about 3GPa.

The simplest and most accurate of the three radius measurement methods is measuring long beams that have curled into circles directly. For larger radii of curvature, deflection measurements by focusing a microscope on different parts of a beam is appropriate. Together, these two methods cover a range of ρ = 10μm to 80cm. A third method using a laser beam to try to measure larger ρ's was unsuccessful. It is not recommended that this method be pursued.

The doubly clamped beam and ring test structures should be checked to see whether they really work as expected, with all beams above a certain size buckling and the others remaining flat. One way to do this for the low-stress films would be to generate a new mask with even larger test patterns. Another would be to try using thinner films. A simple method that could be tried on the existing nitride structures is heating and cooling the silicon wafer to take advantage of the $0.2x\,10^{-6}K^{-1}$ difference in thermal expansion coefficients. Heating or cooling the wafer by about 40°C should strain the nitride enough to cause some buckling.

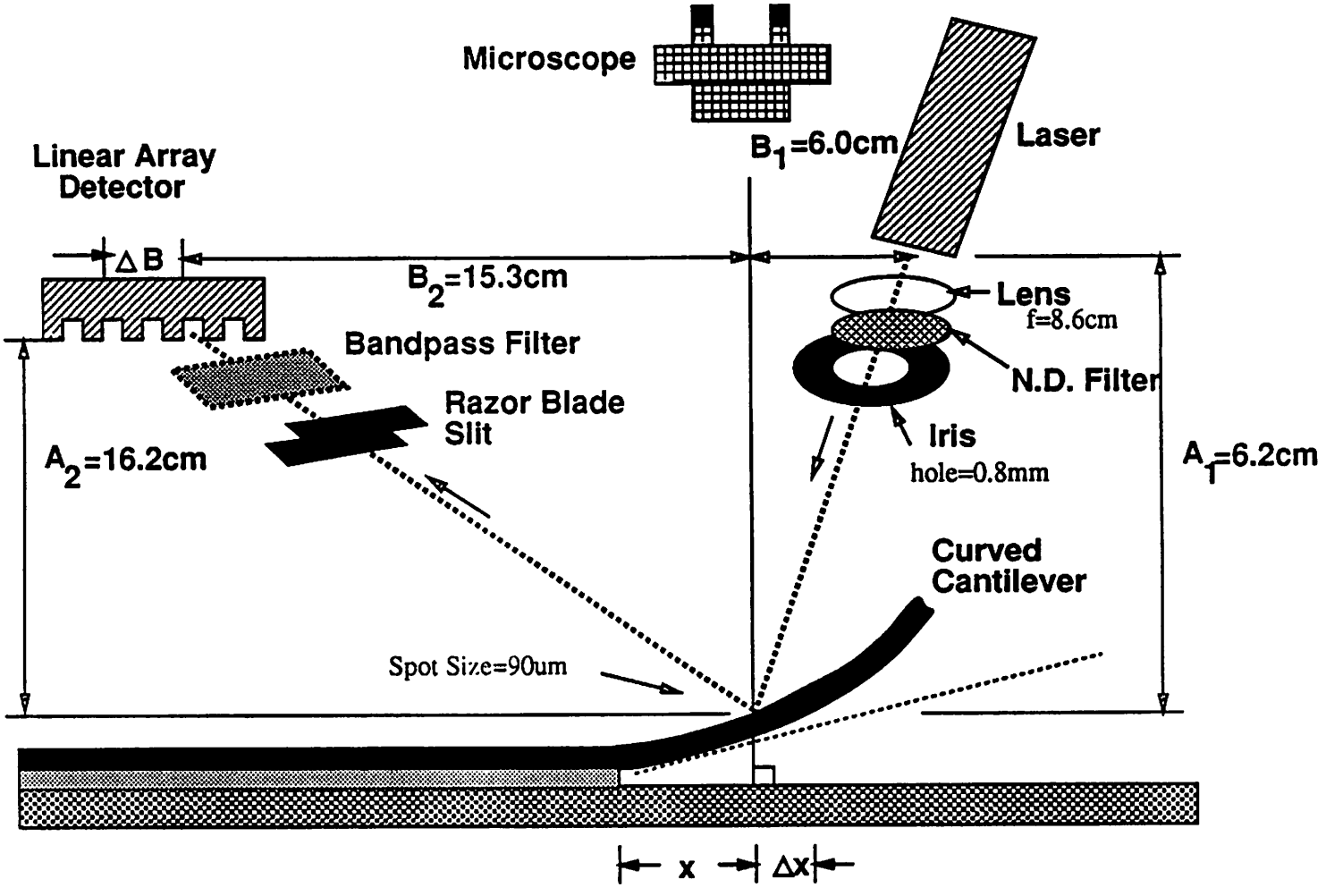
To get around the problem of the beams and rings sticking to the substrate, one might be tempted to make the PSG pedestals thicker. This wouldn't help these devices much: doubling the PSG thickness would increase the length of beams that don't stick by only $2^{1/3}$. It would also increase the undercutting during the BHF etch step, which would require all end rectangles to be made larger. It may be possible, however, to reduce sticking with an improved rinse or even a chemical treatment.

If other Microlab users desire to use some of our structures as diagnostic devices on their chips, we recommend first determining what magnitude of ε_r they want to look at, then, knowing the film thickness, selecting the length that will tell whether ε_r is greater or less than this value. If a little more accuracy is desired, a few different lengths should be used.

If other users want to use the cantilever method, both low-stress nitride and poly are good candidates for the "known" layer. The poly is easier to process because it can be plasma etched and easily wet etched.

Since single-layer cantilevers are excellent for looking at stress gradients in films, depositing films on different substrates and patterning cantilevers out of it could yield information on the film-substrate interface.

Appendix A: Determining Radius of Curvature Using the Laser Setup



From the geometry of the setup:

$B_1/A_1 = \tan\beta_1$, $B_2/A_2 = \tan\beta_2$

$2\alpha + \beta_1 + \beta_2 = 180° \Rightarrow \alpha = 90° - \frac{1}{2}\beta_1 - \frac{1}{2}\beta_2$

$\theta + \alpha + \beta_1 = 90° \Rightarrow \theta = 90° - \alpha - \beta_1$

$= 90° - (90° - \frac{1}{2}\beta_1 - \frac{1}{2}\beta_2) - \beta_1$

$= \frac{1}{2}(\beta_2 - \beta_1)$

so $\theta = \frac{1}{2}[\tan^{-1}(B_2/A_2) - \tan^{-1}(B_1/A_1)]$

Now to find ρ, $\Delta\tan\theta$ is required.

If we start at the flat root of the beam, $B_2/A_2 = B_1/A_1$.

Then $\theta_1 = 0$ and $\theta_2 = \frac{1}{2}[\tan^{-1}((B_2+\Delta B)/A_2) - \tan^{-1}(B_2/A_2)]$.

Since $\Delta B \ll B_2$ and ΔB is very small, $\theta_2 = \frac{1}{2}[\tan^{-1}\Delta B/A_2] = \frac{1}{2}\Delta B/A_2$,

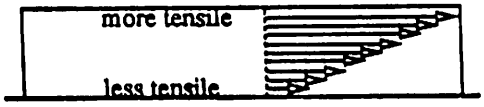
so $\Delta\tan\theta = \tan\theta_2 = \Delta B/A_2$.

Thus, $1/\rho = \Delta\tan\theta/\Delta x$

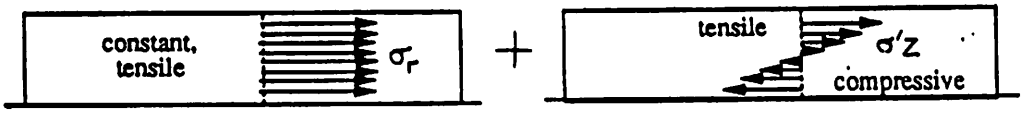
$1/\rho = \Delta B/2A_2\Delta x$.

Appendix B: Stress Gradient

- A film may have a uniform residual stress, or it may vary through the film's thickness. We'll assume this stress gradient is linear with gradient σ' in MPa/μm.

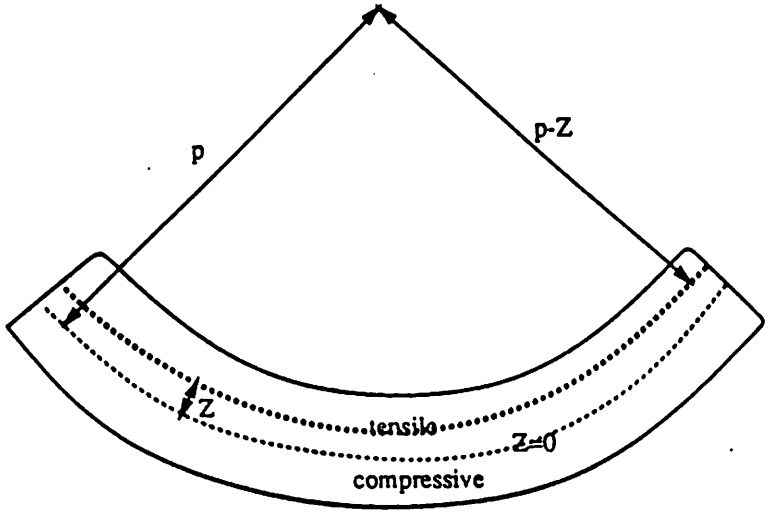


- This stress as a function of position can be divided into a constant residual stress σ_r and stress due to the gradient $\sigma' z$.



- Upon releasing a cantilever beam, the constant part of the residual stress is relieved expansion or contraction of the beam, leaving behind the gradient.

- This will cause the cantilever to bend upward if the film is tensile on top and compressive on the bottom.



- Calculating σ' :
  - strain = $\varepsilon(z) = \dfrac{x_1 - x_2(z)}{x_1} = \dfrac{\rho\sigma - (\rho - z)\sigma}{\rho\sigma} = \dfrac{z}{\rho}$
  - stress = $\sigma(z) = E'\varepsilon(z) = E'z/\rho$
  - Equating to $\sigma(z) = \sigma' z$,

$$\boxed{\sigma' = E'/\rho}$$

Appendix C

Derivation of Stress as a Function of Radius of Curvature

ρ = radius of curvature

E' = Effective Young's Modulus (Stiffness), $E'=E/(1-\upsilon)$ here

M = Bending moment

I = Moment of inertia (resistance to bending based on geometry)

Basic equation for a beam:

$$\frac{1}{\rho} = \frac{M}{EI}$$

It is seen that for higher M, the curvature increases ($1/\rho$ decreases) while for higher E or I, the curvature is smaller because the beams is harder to bend. For simplicity, E and I will be grouped together as (EI) here.

For a beam of a single material, $(EI)=Ebh^3/12$, where b is the width and h is the thickness. Thus, $(EI)_1 = E_1bh_1^3/12$, $(EI)_2 = E_2bh_2^3/12$. For a two-layer beam the equivalent EI is

$$(EI)_{eq} = \frac{K_1}{mn+1}\frac{E_2bh_2^3}{12,}$$

where $m = E_1/E_2$, $n = n_1/n_2$, $h = h_1+h_2$, $K=1+4mn+6mn^2+4mn^3+m^2n^4$ (from Roark [8], Judy [9])

Solving Eq. 13 for M, $M = (EI)/\rho$. Knowing the dimensions of the beams, and having measured ρ_1, ρ_2, and ρ_T (total, for the composite beam),

$$M_1 = (EI)_1/\rho_1,\ M_2 = (EI)_2/\rho_2 \text{ and } M_T = (EI)_{eq}/\rho_T.$$

Now the total binding moment M_T is due to the films alone, plus the interaction of the residual stresses (M_R):

$$M_T = M_1 + M_2 + M_R,$$

$$M_R = M_T - M_1 - M_2 = (EI)_{eq}/\rho_T - (EI)_1/\rho_1 - (EI)_2/\rho_2.$$

The effective radius of curvature due to interaction alone (that is, minus the effects of M_1 and M_2) is:

$$\frac{1}{\rho_R} = \frac{M_R}{(EI)_{eq}} = \frac{1}{\rho_T} - \frac{(EI)_1}{(EI)_{eq}}\frac{1}{\rho_1} - \frac{(EI)_2}{(EI)_{eq}}\frac{1}{\rho_2}$$

$$\frac{1}{\rho_2} = \frac{1}{\rho_T} - \frac{E_1bh_1^3/12}{\frac{K_1}{mn+1}\frac{E_2bh_2^3}{12}}\frac{1}{\rho_1} - \frac{E_2bh_2^3/12}{\frac{K_1}{mn+1}\frac{E_2bh_2^3}{12}}\frac{1}{\rho_2}$$

$$= \frac{1}{\rho_T} - \frac{mn+1}{K_1}\frac{E_1}{E_2}\left[\frac{h_1}{h_2}\right]^3\frac{1}{\rho_1} - \frac{mn+1}{K}\frac{1}{\rho_2} = \frac{1}{\rho_T} - \frac{mn+1}{K_1}\left[\frac{mn^3}{\rho_1}+\frac{1}{\rho_2}\right]$$

Now $1/\rho_R$ has been related to the residual stresses alone by Judy et. al. as

$$\frac{1}{\rho_R} = \frac{6(m\sigma_1 - \sigma_2)}{hE_2(3m+K_1K_2)}, \text{ when } K_2 = [n(n+1)^2]^{-1}$$

*can see if $\varepsilon_1 = \varepsilon_2$ => no curvature

Solving for σ_2, $\sigma_2 - \frac{\sigma_1}{m} = \frac{hE_2(3m+K_1K_2)}{6m\rho_R}$.

If σ_1 is small, σ_2 dominates, and

$$\sigma_2 = \frac{hE_2(3m+K_1K_2)}{6m\rho_R} = \frac{hE_2K_3}{\rho_R}, \; K_3 = \frac{3m+K_1K_2}{6m}.$$

Summary:

Measure ρ_1 (often $\to\infty$),ρ_2, ρ_T

Look up E_1', E_2'; $m = E_1'/E_2'$

Measure h_1, h_2; $n = h_1/h_2$, $h = h_1 + h_2$,
$K_1 = 1 + 4mn + 6mn^2 + 4mn^3 + m^2n^4$
$K_2 = [n(n+1)^2]^{-1}$
$K_3 = (3m + K_1K_2)/6m$

$$\rho_R = \left[\frac{1}{\rho_T} - \frac{mn+1}{K_1}\left[\frac{mn^3}{\rho_1} + \frac{1}{\rho_2}\right]\right]^{-1}$$

If $\sigma_2 \gg \sigma_1/m \to \varepsilon_2 \gg \sigma_1$,

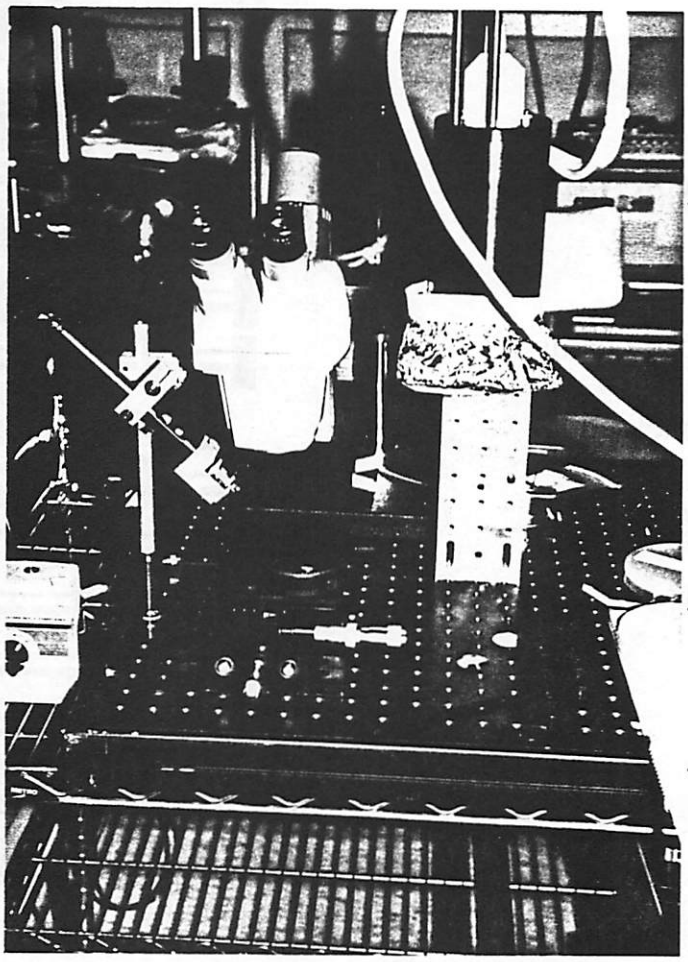
$$\sigma_2 = \frac{h}{\rho_R}E_2'K_3$$

Photo 1: Laser Setup

Microscope ------->

Laser ------->

Lens, ND Filter, and Iris ------->



<------- Reticon Array

<------- Wafer

<------- x-y-ϕ Stage

<------- Optical Bench

Photo 2:

Wafer 13, Sputtered Tungsten

Fine cracks start at corners, areas of high stress concentration.

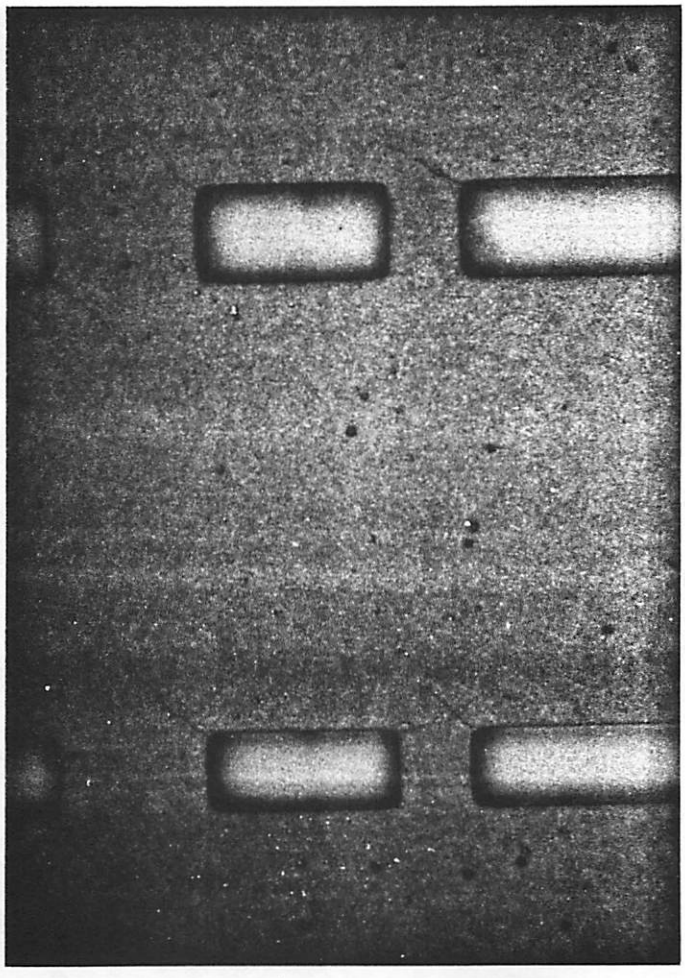


Photo 3:

Wafer 4, Silicon Nitride

Doubly clamped beams
appear to be buckled
but are simply stuck
to the substrate.

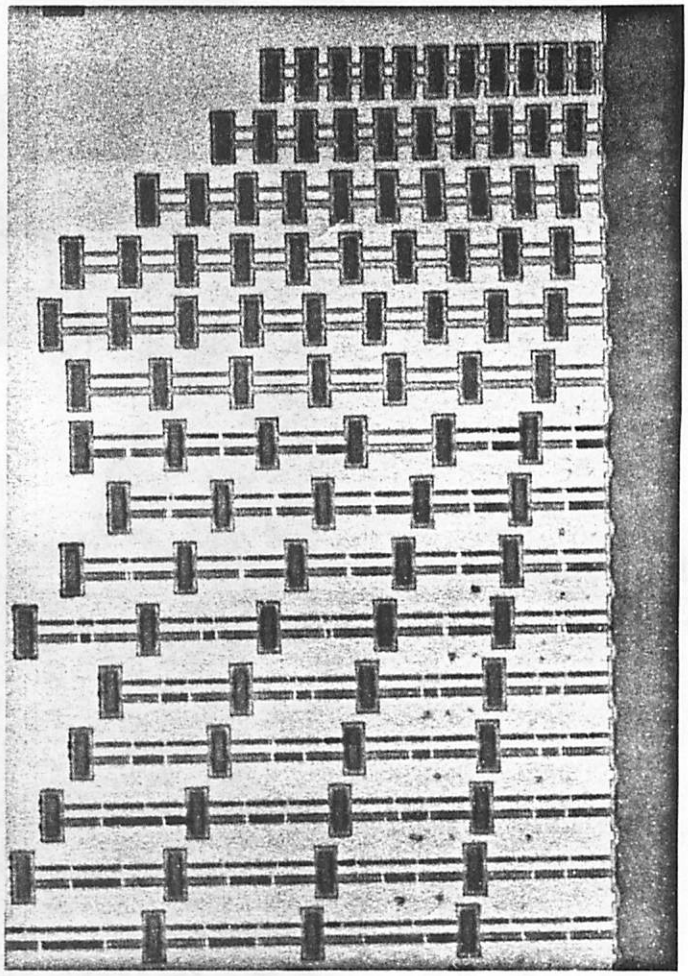


Photo 4:

Wafer 11, Polysilicon

Short doubly clamped
beams are buckled due
to high compressive
residual strain.

Also note fringes
in overhang ------->

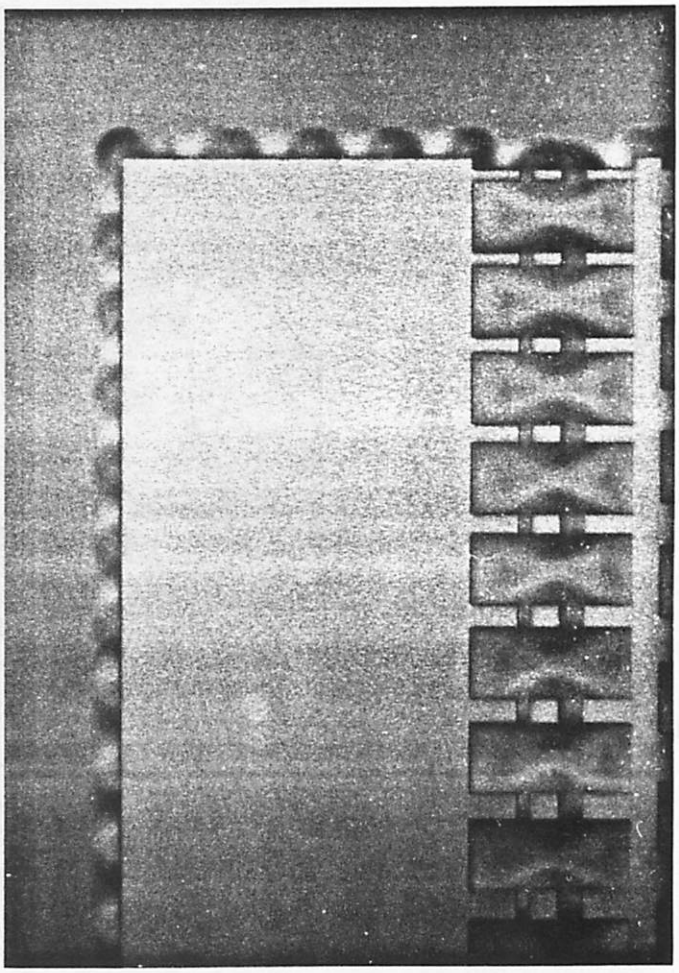


Photo 5:

Wafer 2, CVD Tungsten

Doubly clamped beams are broken due to high tensile residual stress.

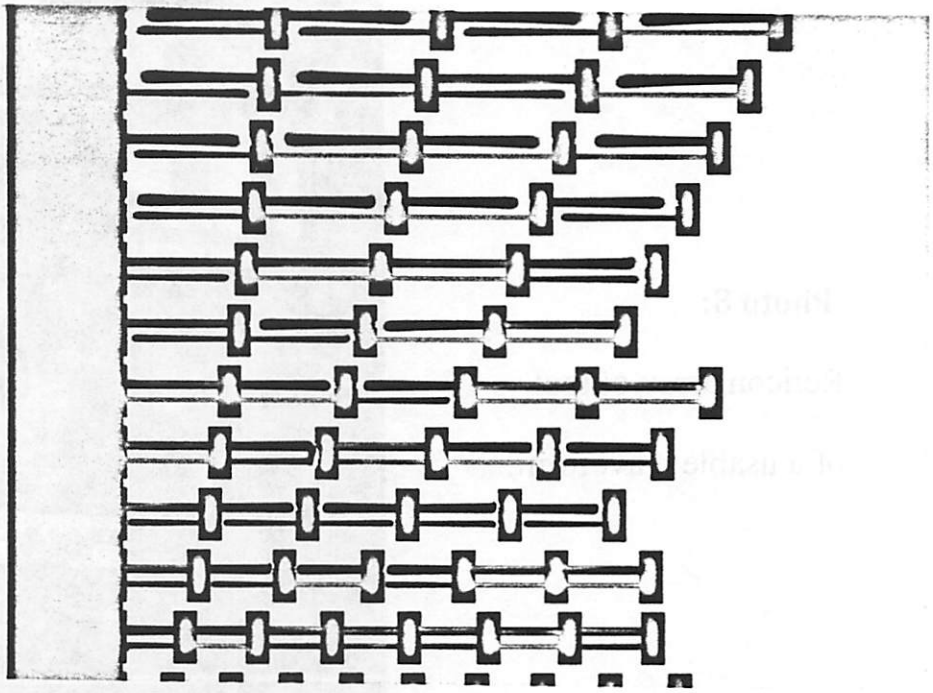


Photo 6:

Wafer 11, Polysilicon on Nitride

A 5000μm cantilever curves downward and to the side, ρ=-150μm.

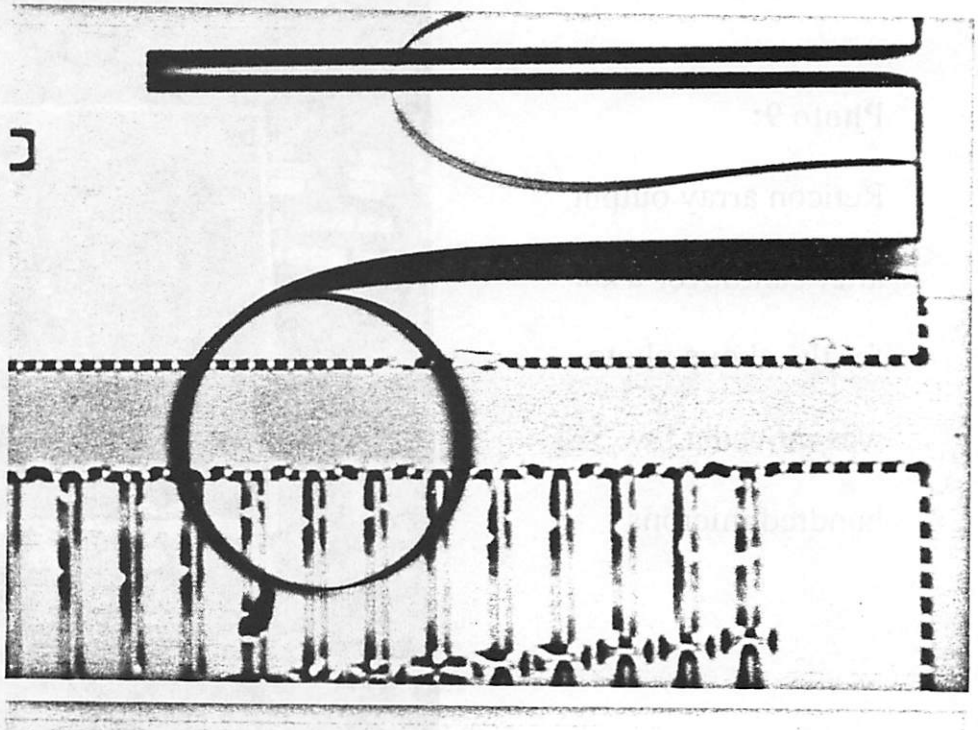


Photo 7:

Wafer 2, CVD Tungsten

5000μm cantilevers curl into helixes due to stress gradients through the thickness and across the width.

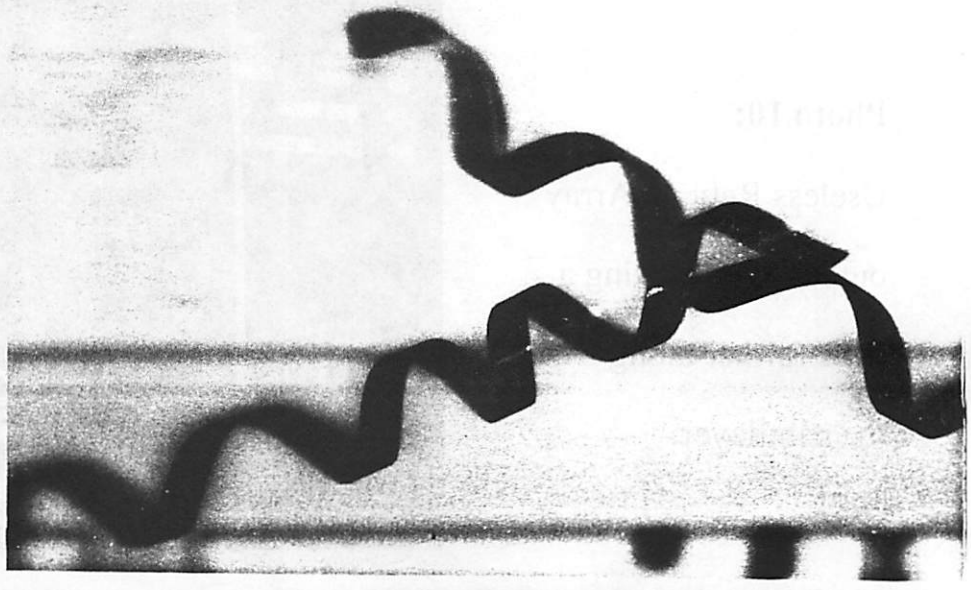


Photo 8:

Reticon array output of a usable waveform.

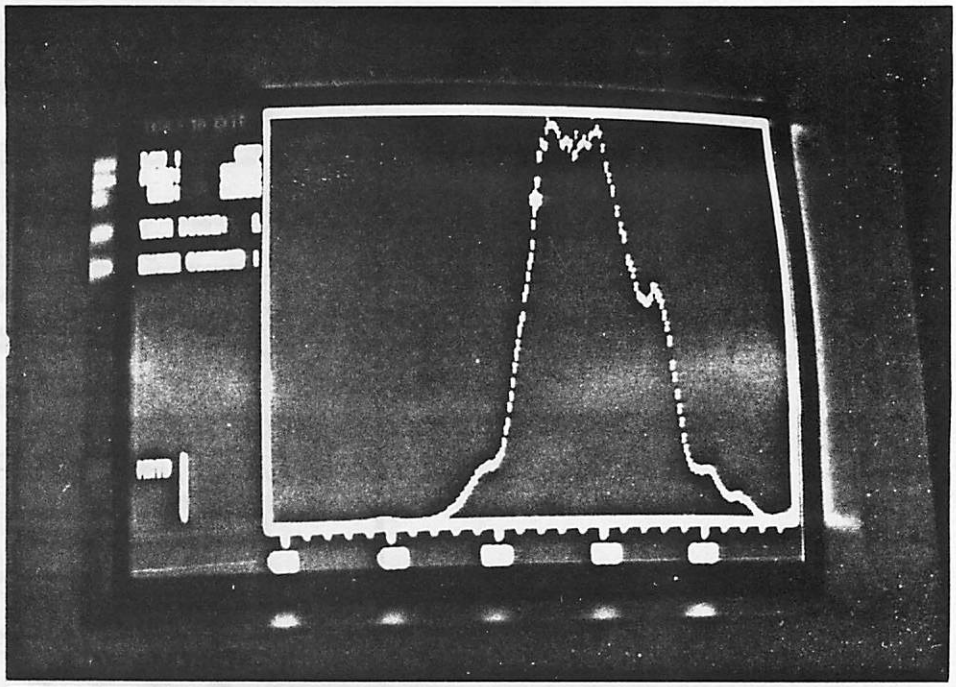


Photo 9:

Reticon array output after cantilever used for the above photo was moved a few hundred microns

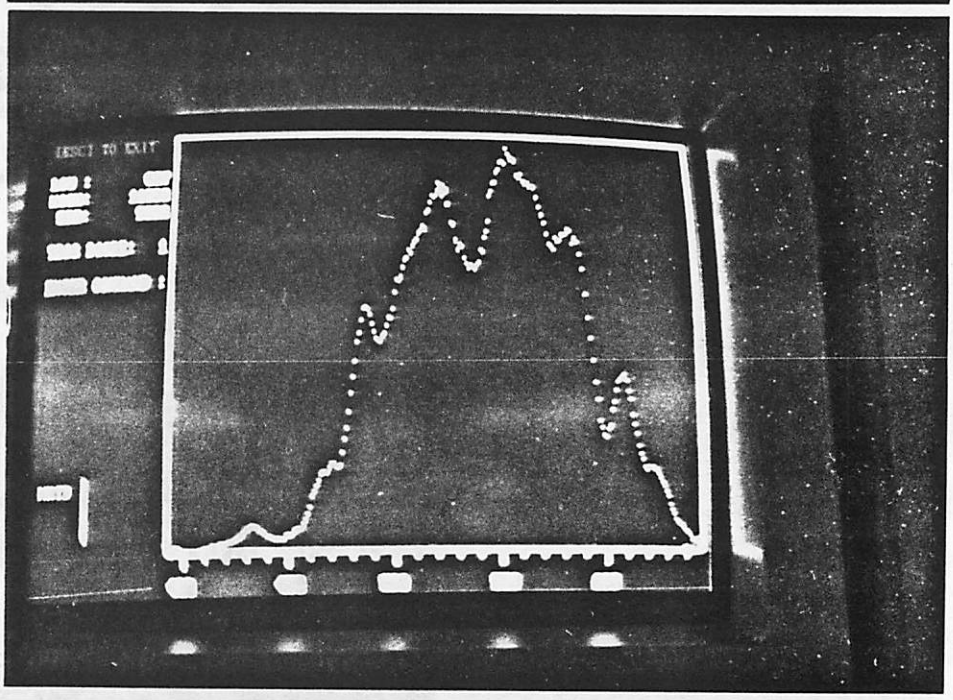


Photo 10:

Useless Reticon Array output after moving a little further along the cantilever.

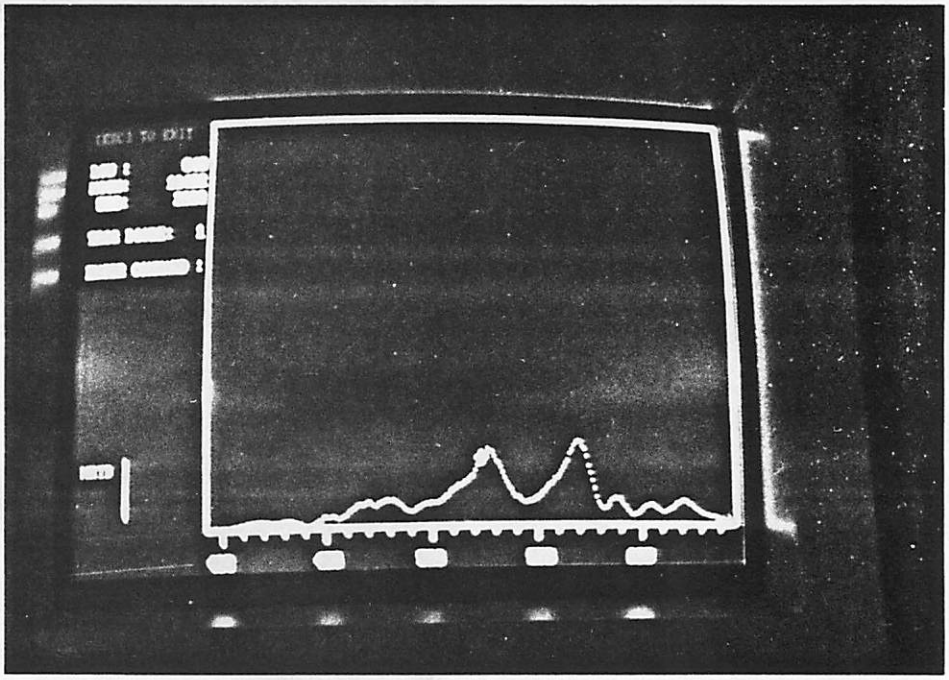
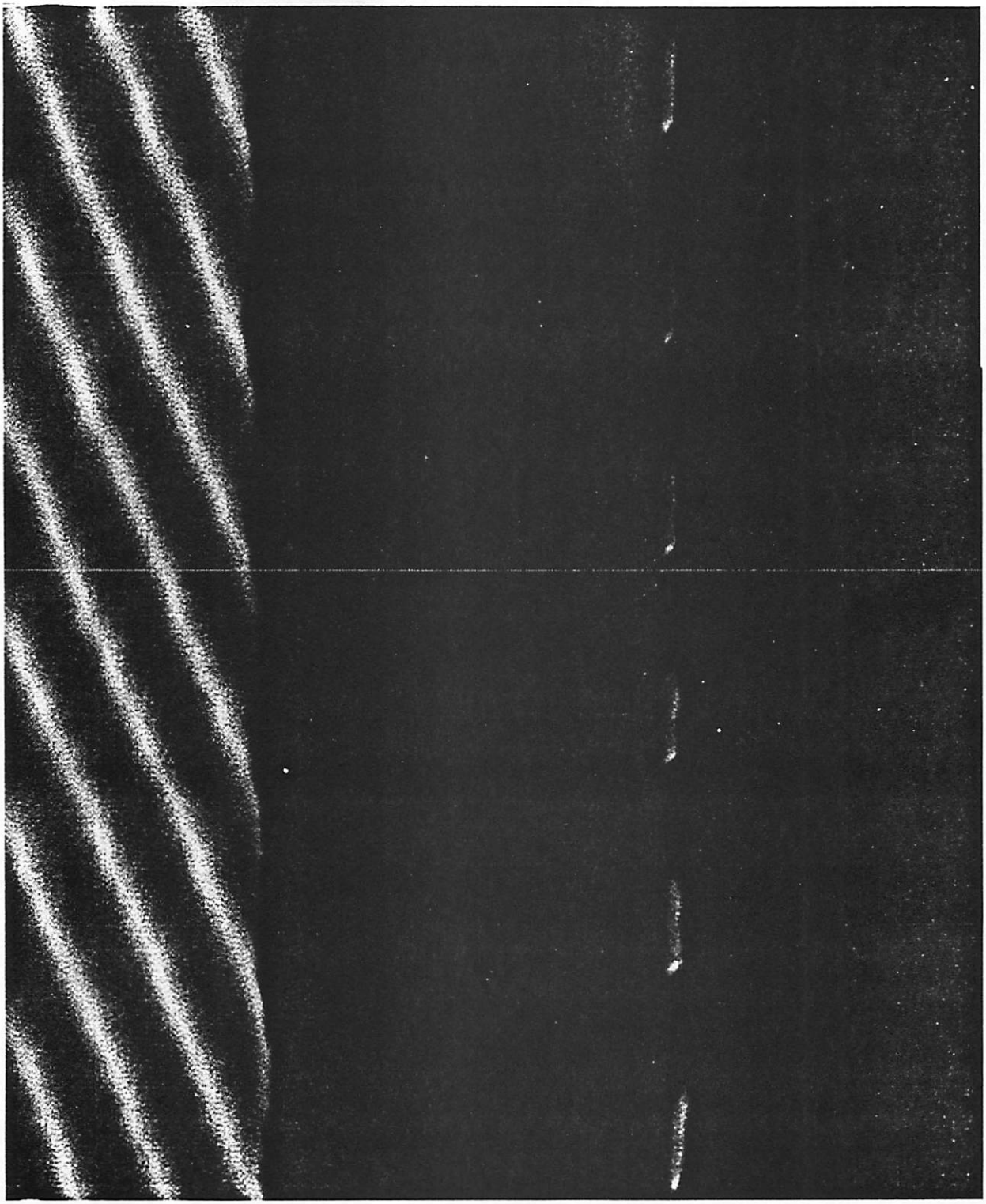
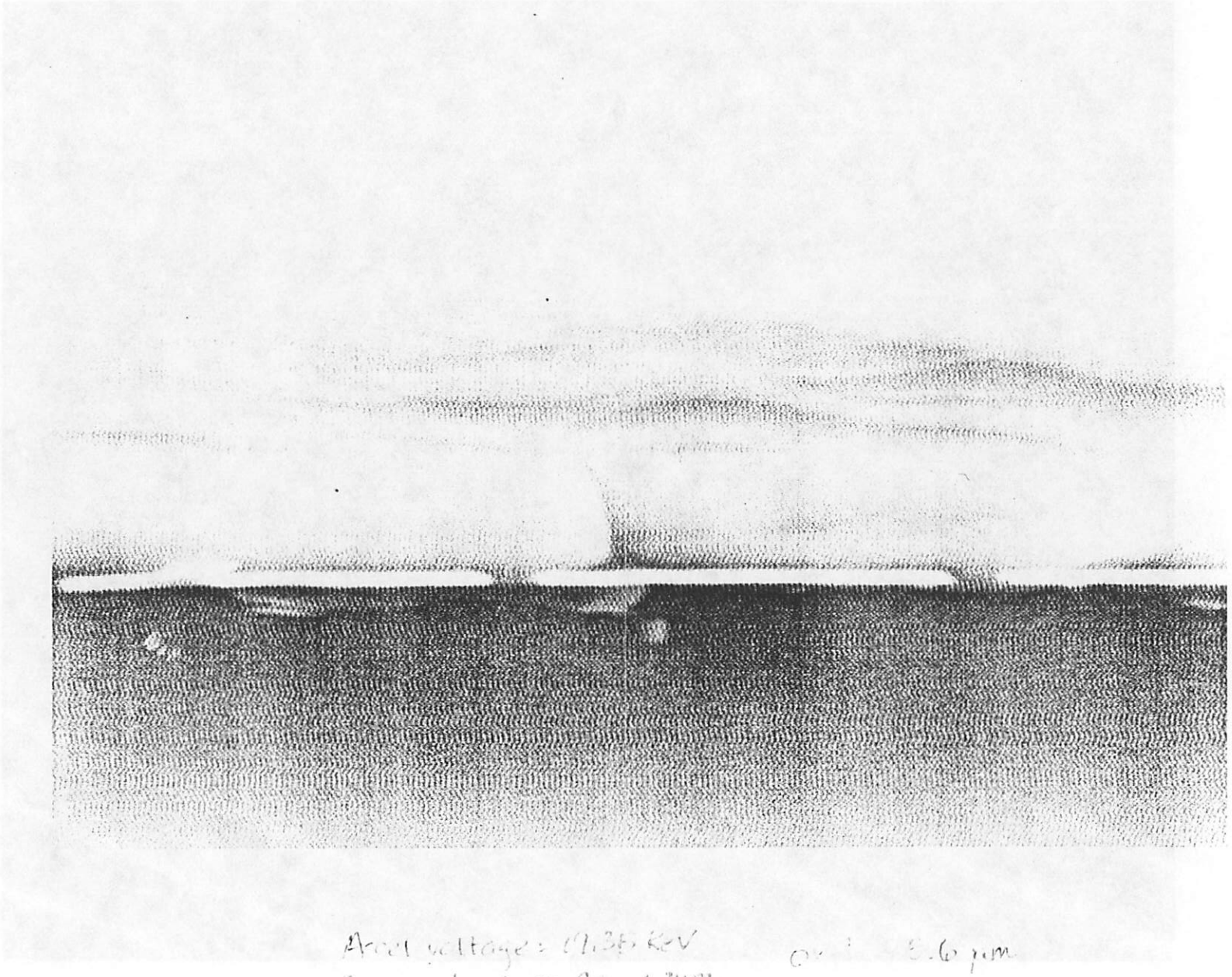


Fig. P1



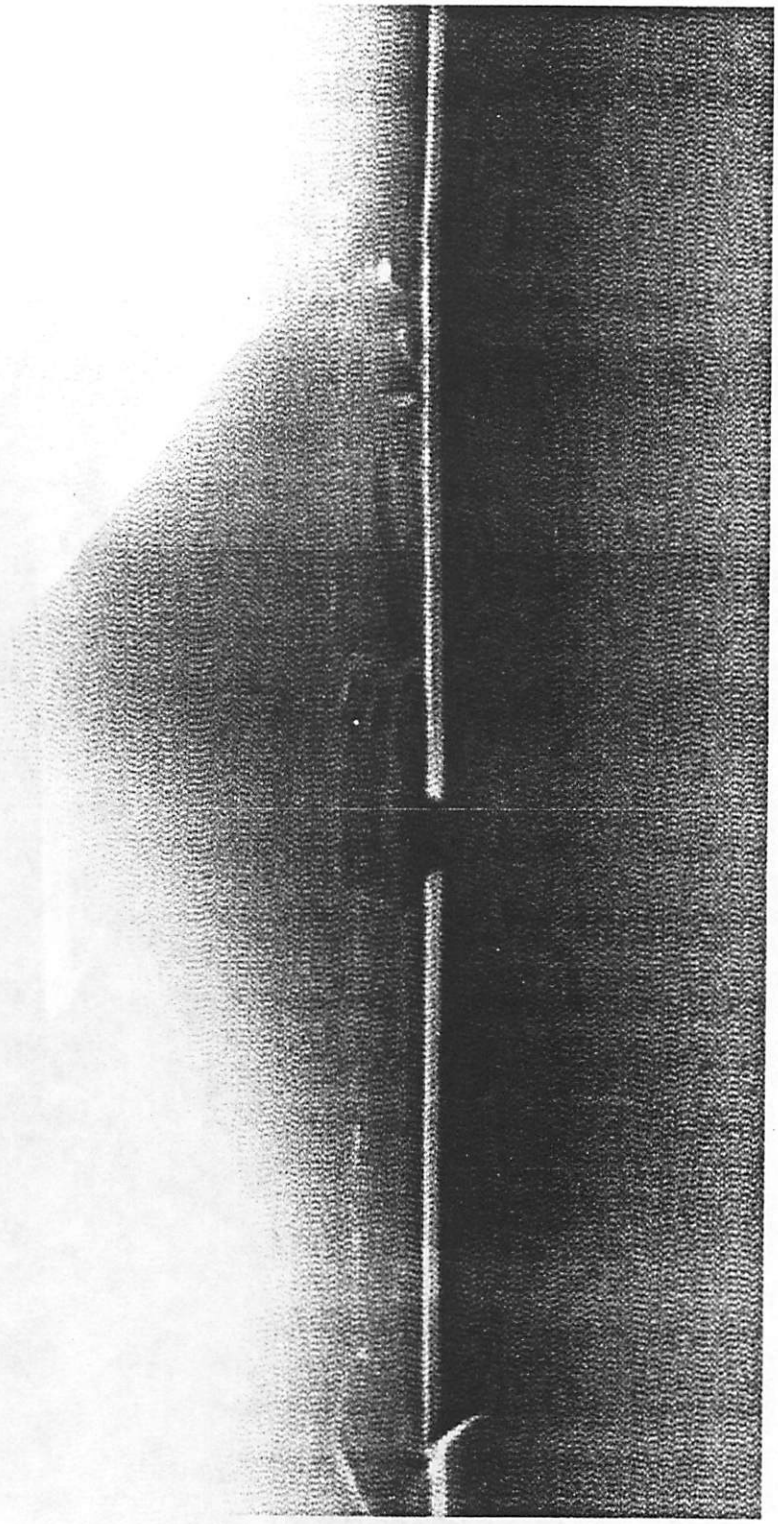


Accel voltage : 17.3 KeV
Focussed WD ≈ 23.6 mm

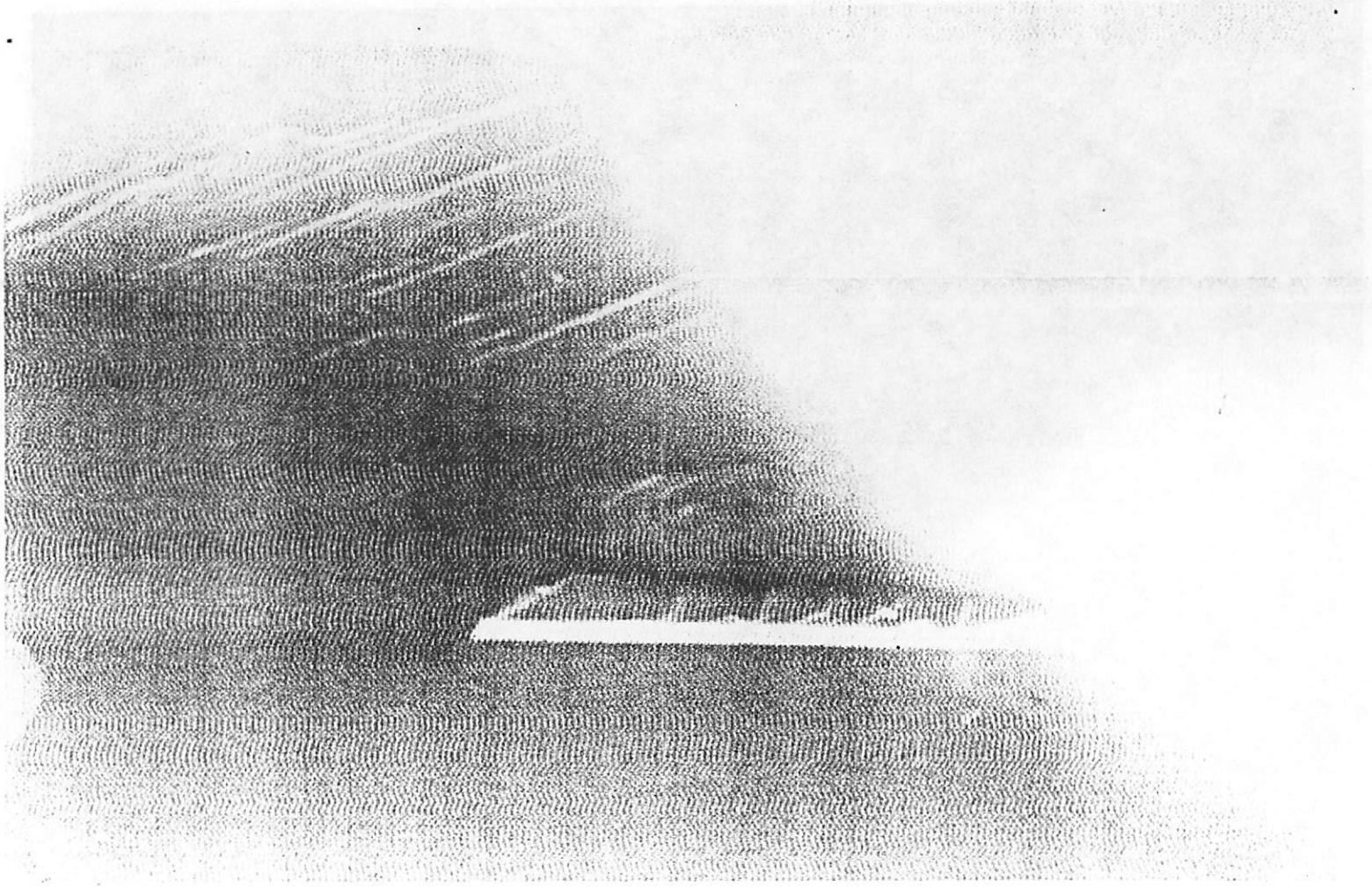
Mag = 7.11 Kx

Fig. P2

Fig. P3

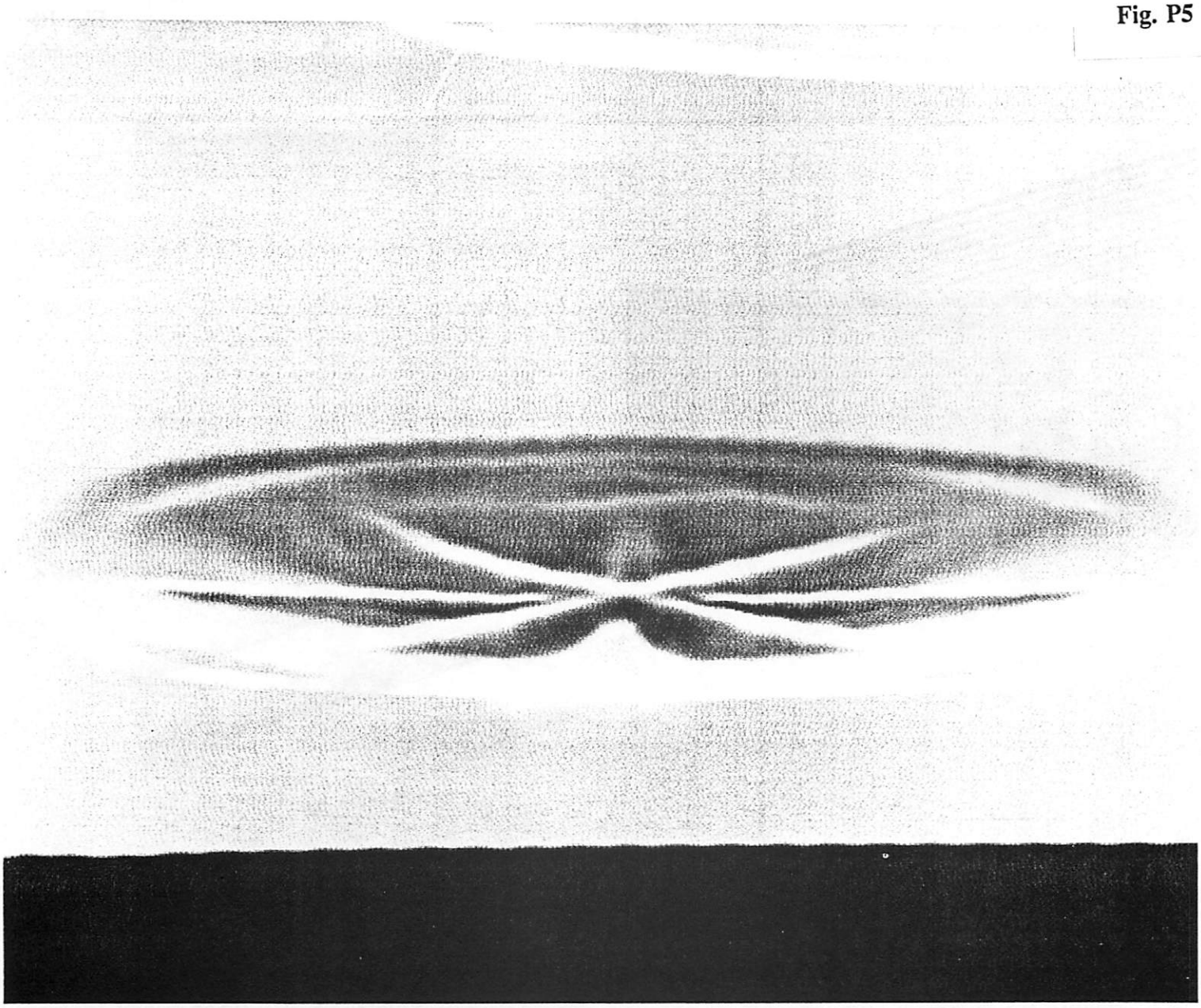


Mag = 9.13Kx
focussed WD = 25.7 mm



Mag = 9.13 Kx
E11
Epi thickness = 5.1 μm

Fig. P4

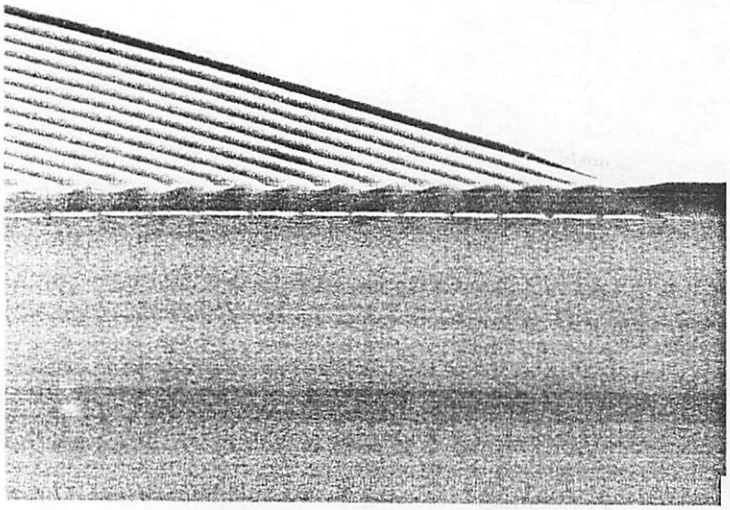


Mag: 8.38 kx

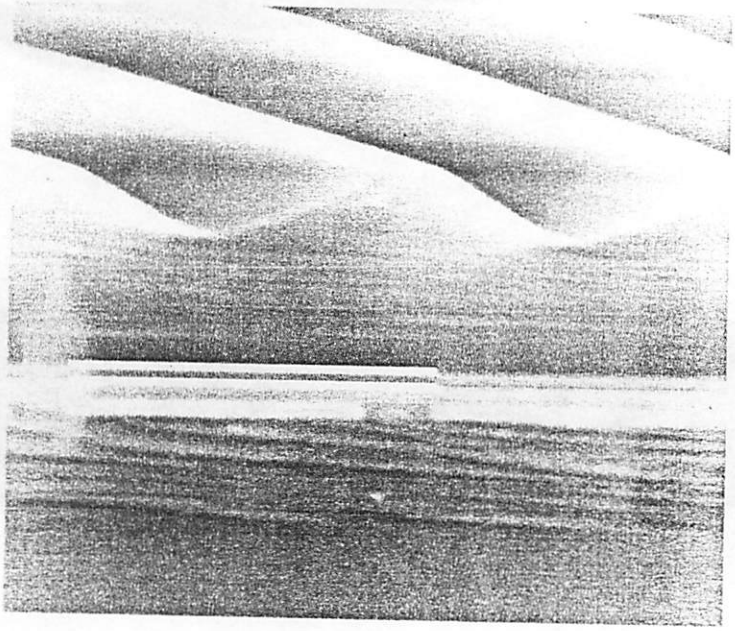
Mag = 100x

Fig. P6

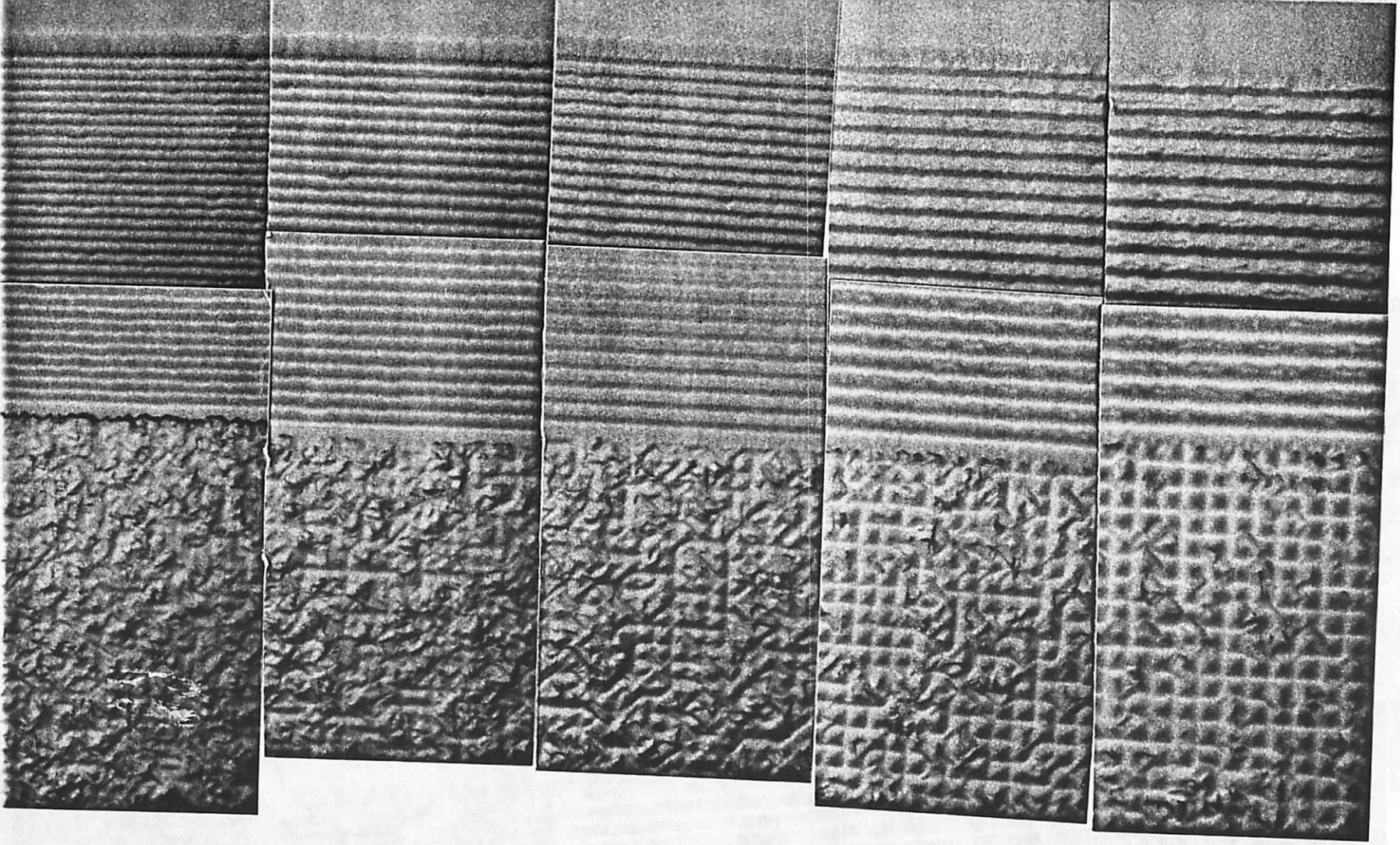
M = 1.78Kx



M = 13.16Kx



M = 9.8Kx



E12
Epi thickness =
5.2 μm

THICK EPI
(100) FLAT

P=2.0 W=0.5
S=1.5

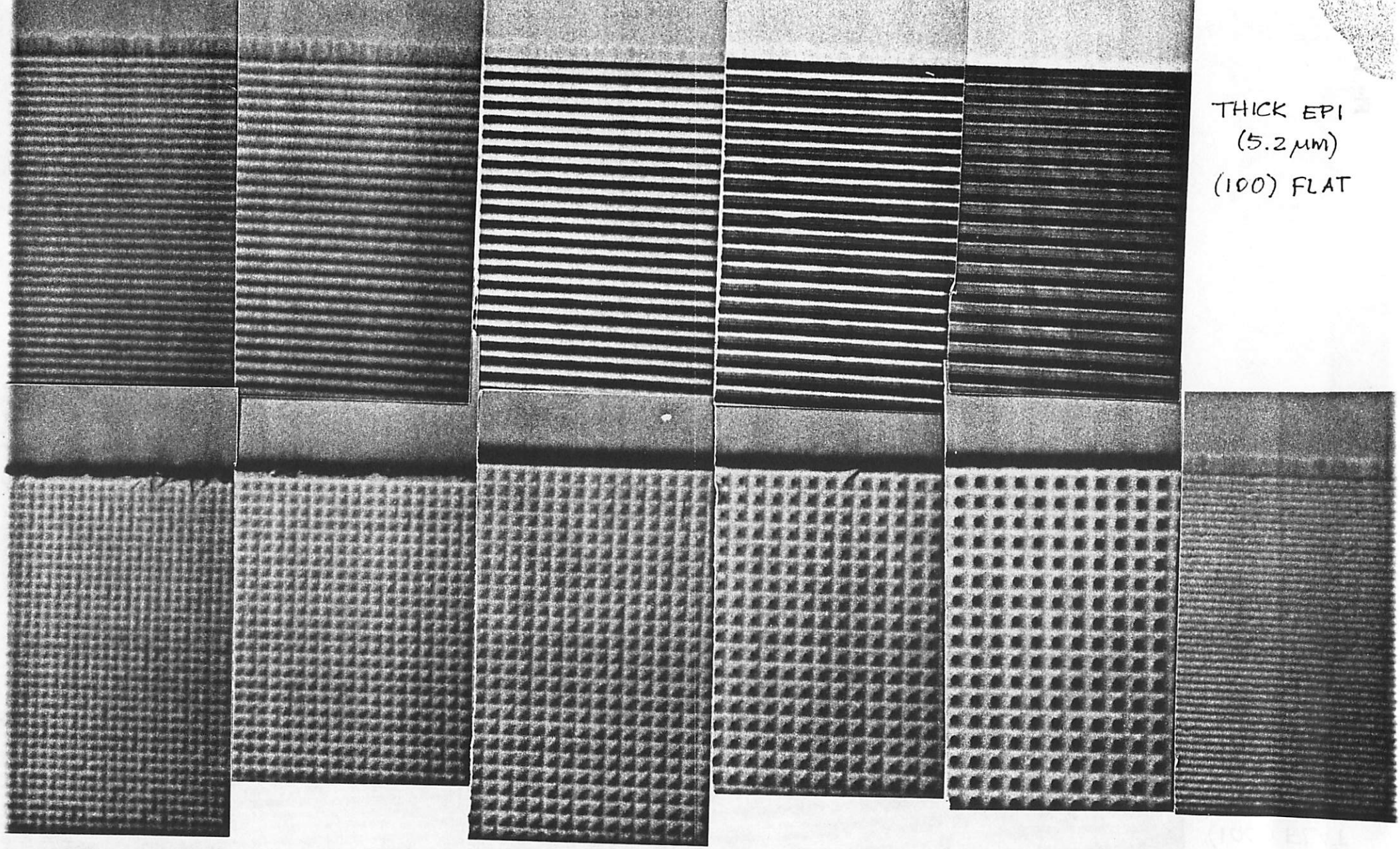
P=2.5
W=0.5
S=2.0

P=3.0
W=0.5
S=2.5

P=3.5
W=0.5
S=3.0

P=4.0
W=0.5
S=3.5

Fig. P7



P=2.0
S=1.0
W=1.0

P=2.5
S=1.0
W=1.5

P=3.0
S=1.0
W=2.0

P=3.5
S=1.0
W=2.5

P=4.0
S=1.0
W=3.0

P=1.5
S=1.0
W=0.5

Fig. P8

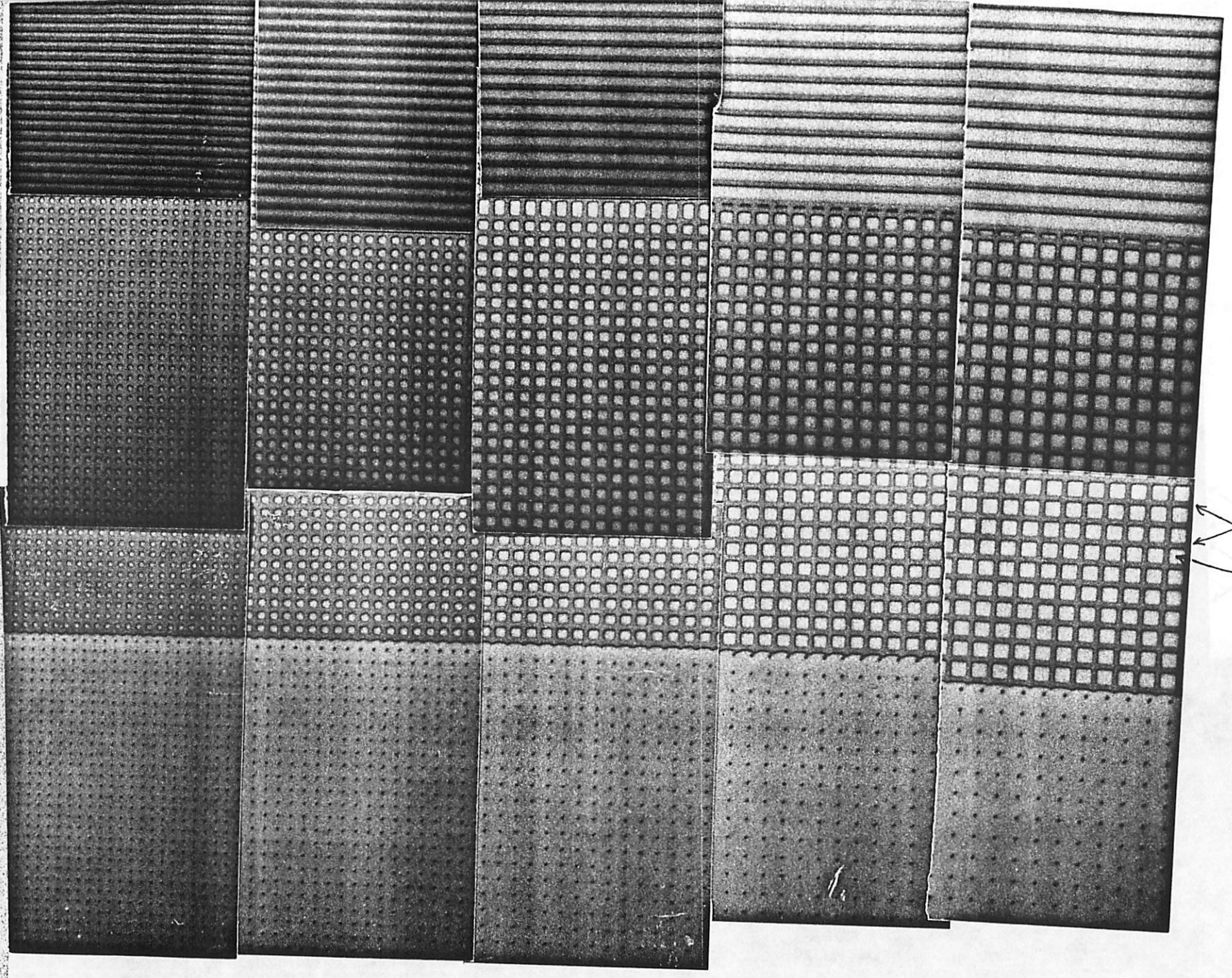
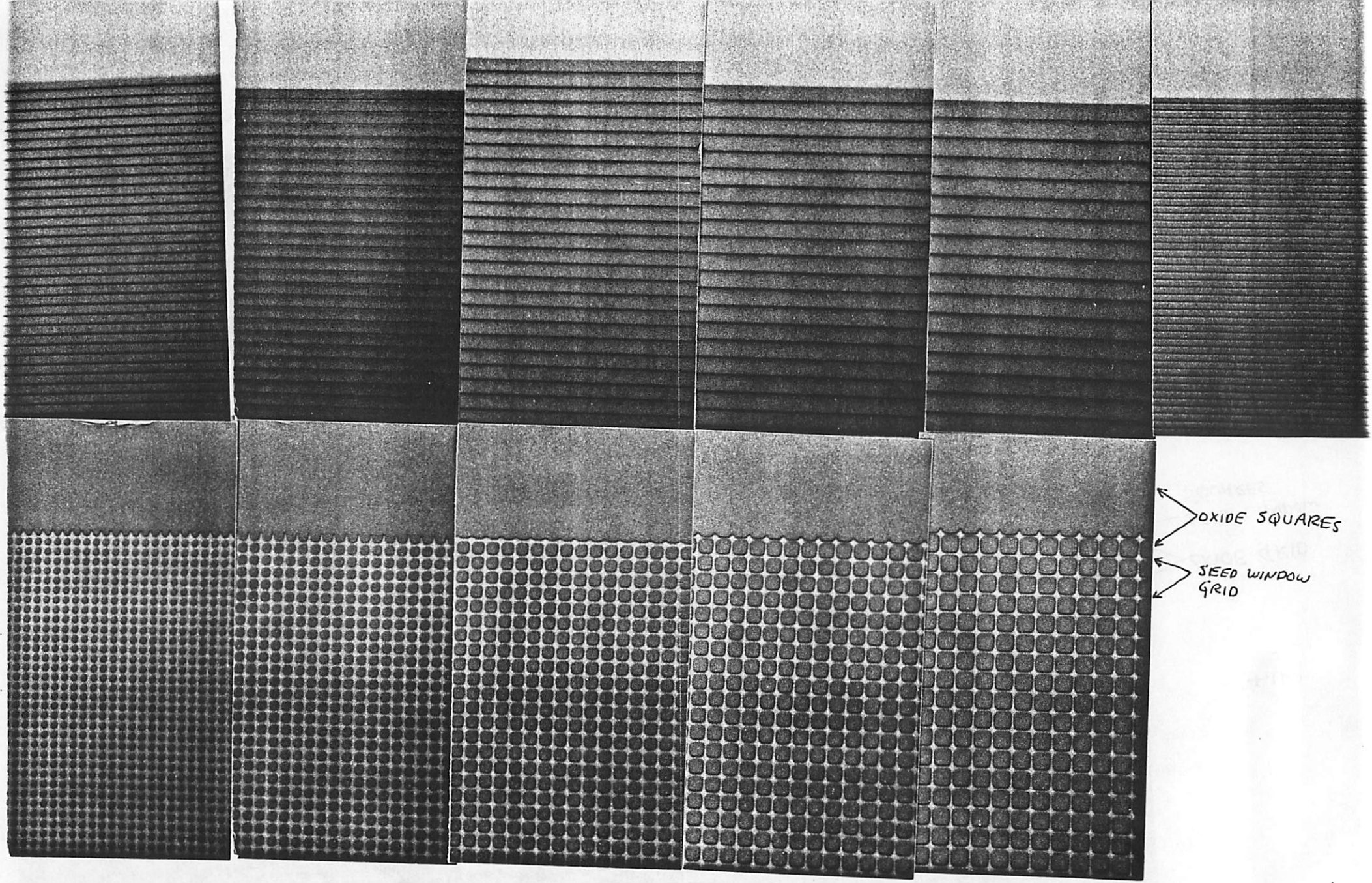

OXIDE SHAPES
LINES
GRIDS
OXIDE GRID
SEED WINDOW SQUARES
SQUARES
P=2.0 w=0.5 s=1.5
P=2.5 w=0.5 s=2.0
P=3.0 w=0.5 s=2.5
P=3.5 w=0.5 s=3.0
P=4.0 w=0.5 s=3.5
Mag=100X

Fig. P9



P=2.0
S=1.0
W=1.0

P=2.5
S=1.0
W=1.5

P=3.0
S=1.0
W=2.0

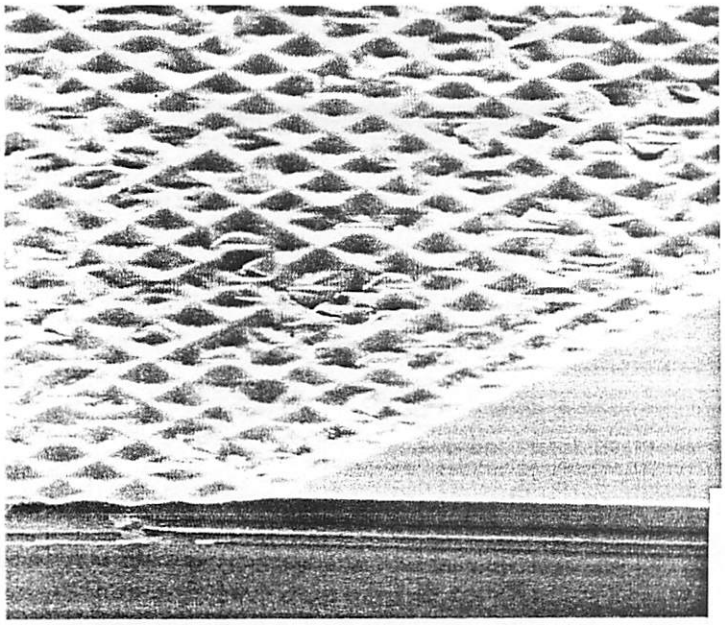
P=3.5
S=1.0
W=2.5

P=4.0
S=1.0
W=3.0

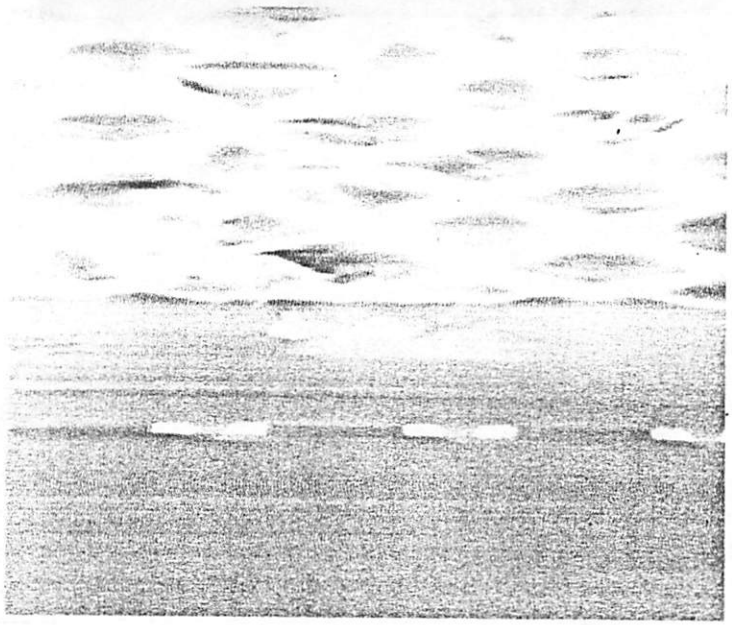
P=1.5
S=1.0
W=0.5

Fig. P10

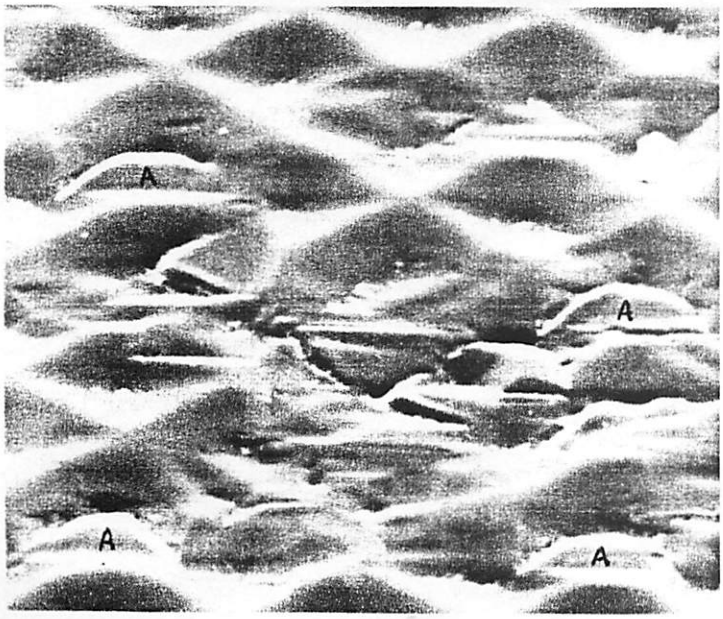
M = 2.48 Kx



M = 7.76 Kx



M = 7.83 Kx



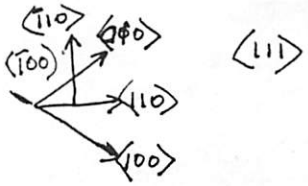
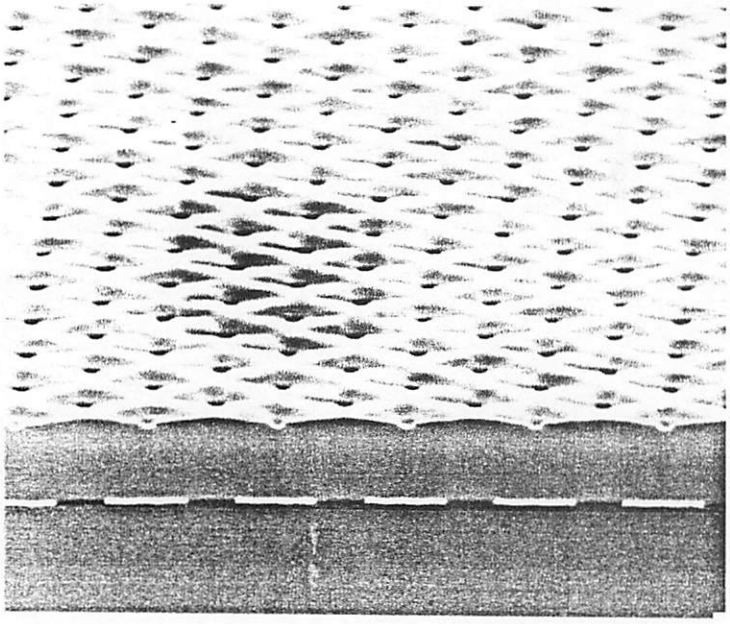
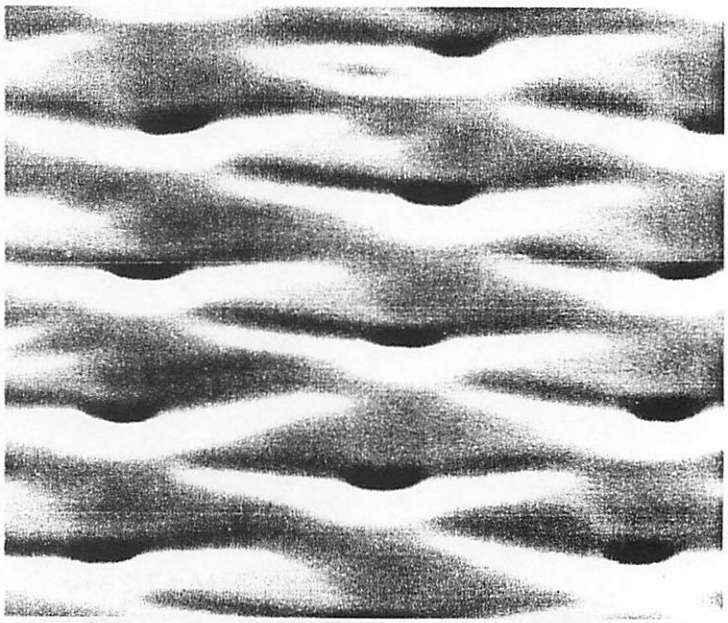


Fig. P12

M = 4.63 Kx



M=19.37 Kx



M= 25.9Kx

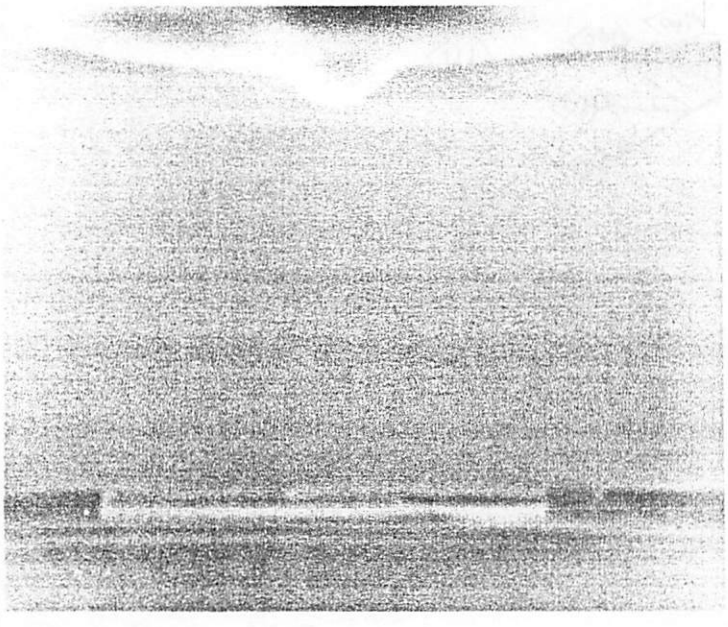
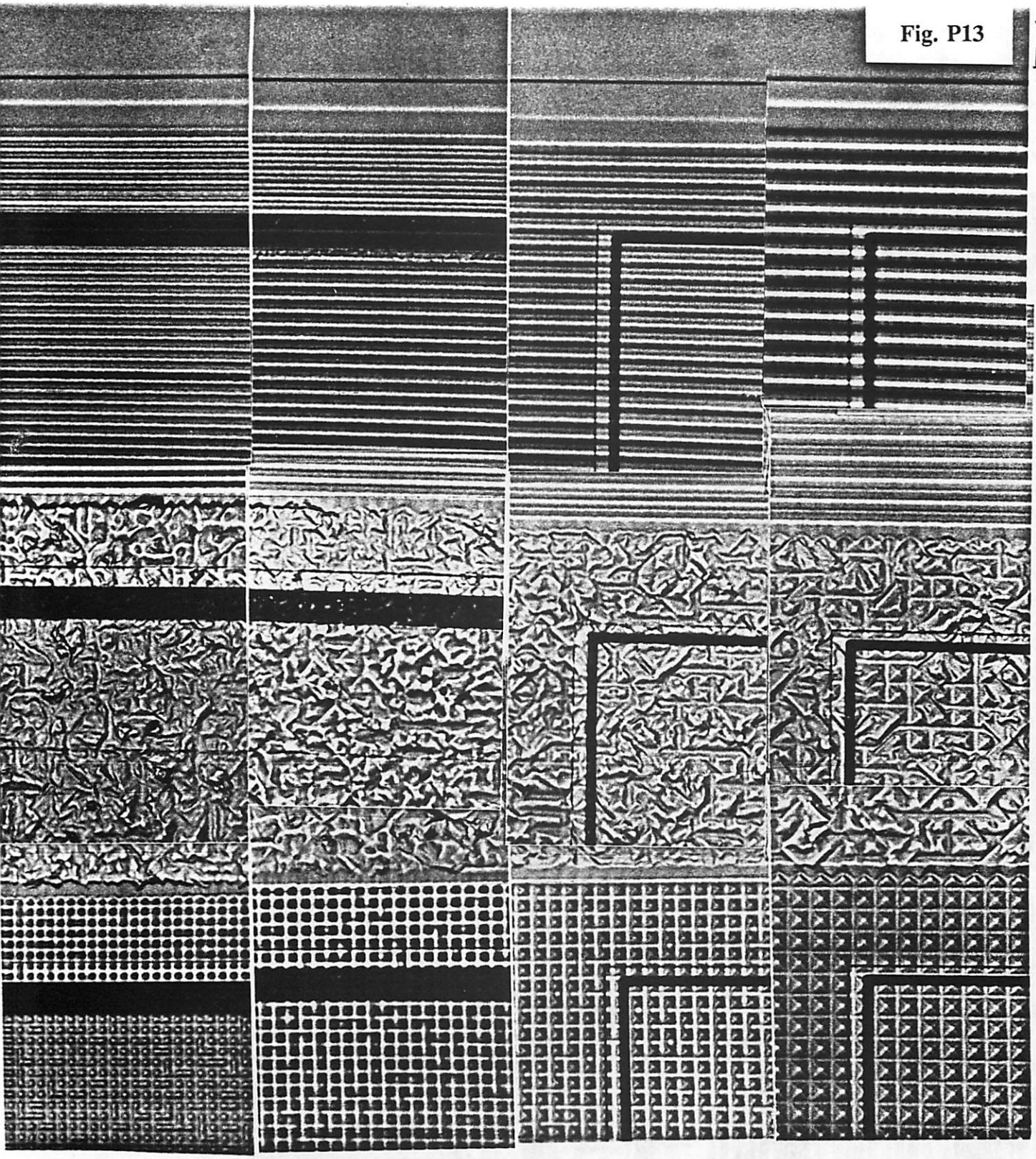


Fig. P13

E8

Epi = 33 μm
THIN EPI
(100) FLAT



LINES

GRIDS

SQUARES

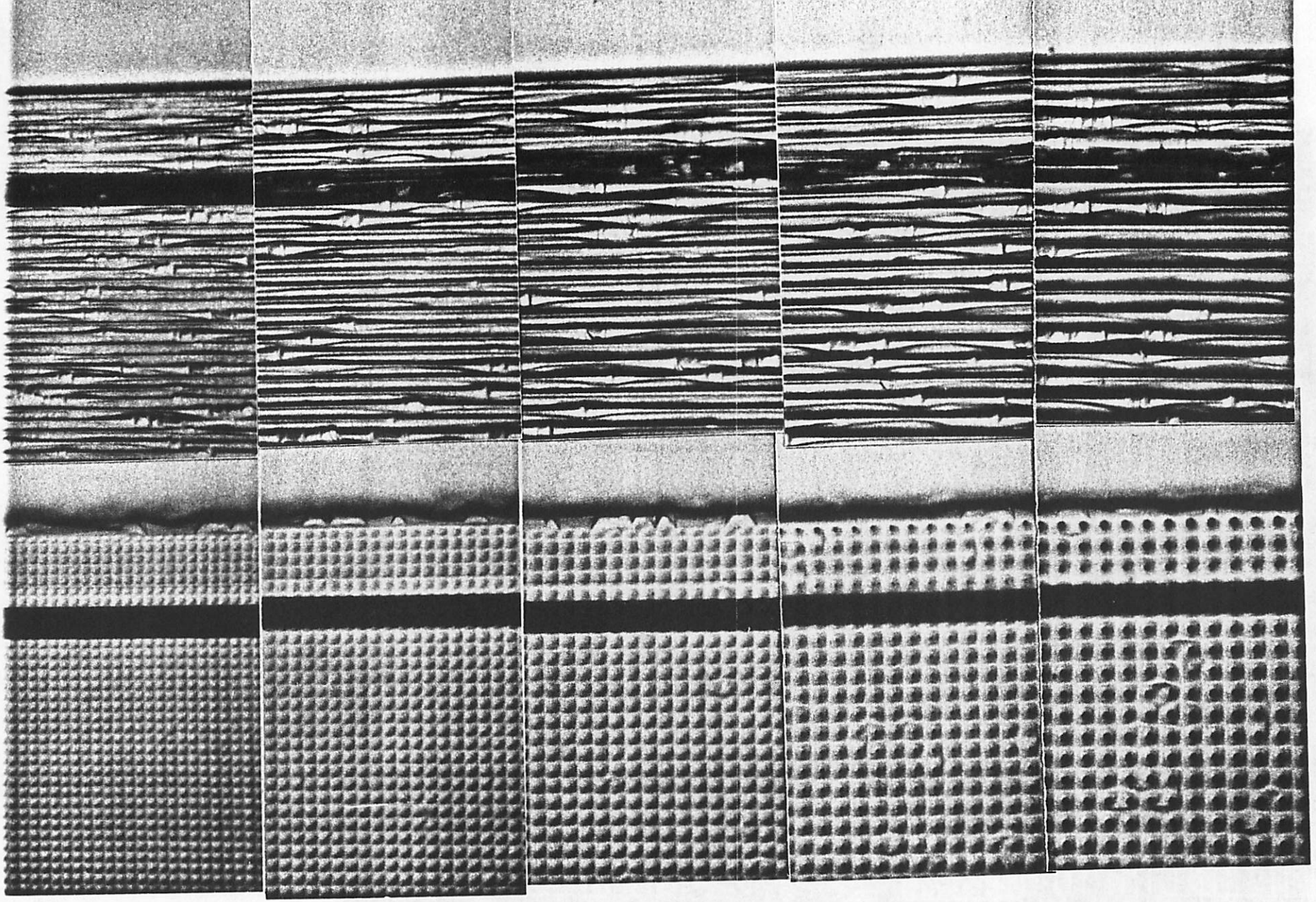
P = 2.0
W = 0.5
S = 1.5

P = 2.5
W = 0.5
S = 2.0

P = 3.0
W = 0.5
S = 2.5

P = 4.0
W = 0.5
S = 3.5

Mag = 100X



THIN D5
EPI (3.3μm)
(110) FLAT

LINES

SQUARES

P=1.0
S=1.0
W=1.0

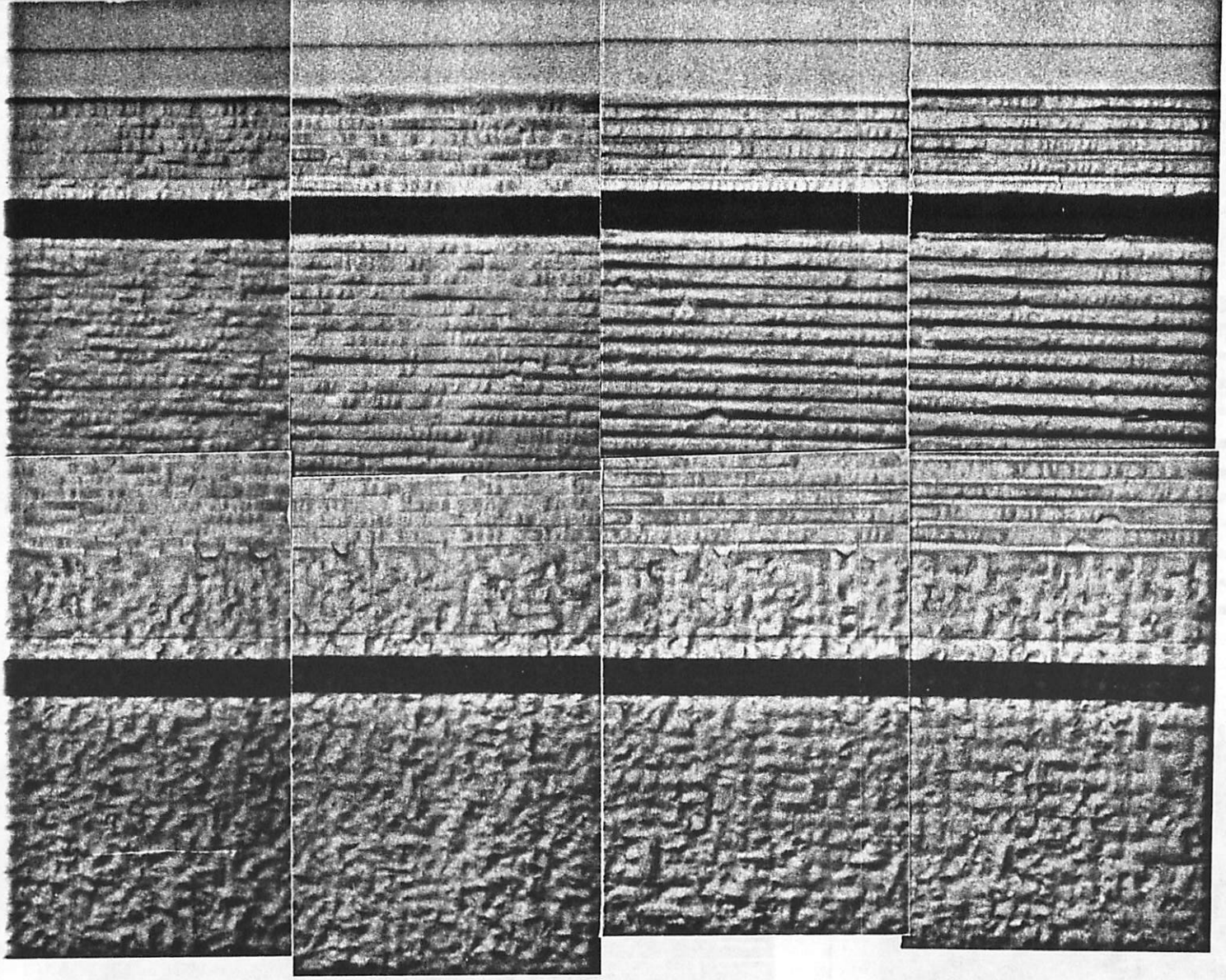
P=2.5
S=1.0
ω=1.5

P=3.0
S=1.0
ω=2.0

P=3.5
S=1.0
ω=2.5

P=4.0
S=1.0
ω=3.0

Fig. P14



D5

THIN EPI (3.3μ)

(110) FLAT

LINES

worse in smaller seeds

GRIDS

no visible pattern

P=2.0
W=0.5
S=1.5

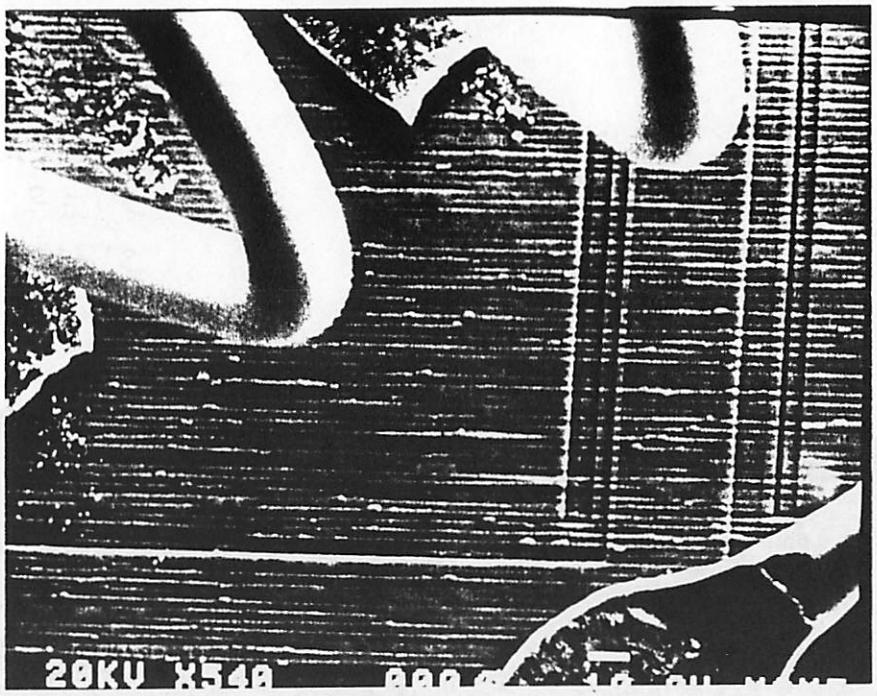
P=2.5
W=0.5
S=2.0

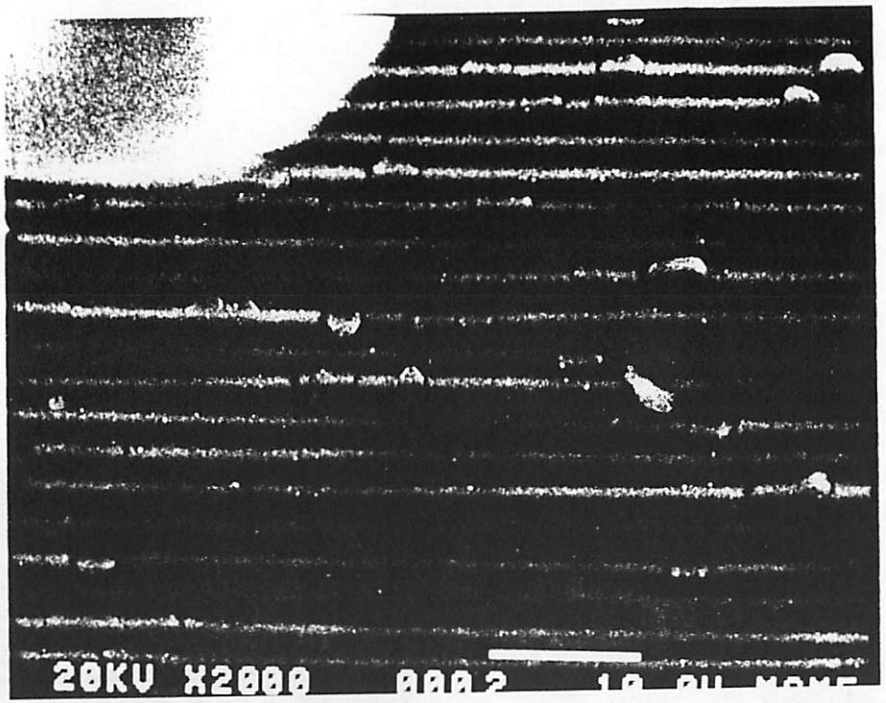
P=3.0
W=0.5
S=2.5

P=4.0
W=0.5
S=3.5

Fig. P15

MAG = 100λ





E8 (100) FLAT THIN EPI (3.3μm)

P=3.5 (OXIDE VARIATION)
W=2.5
S=1.0

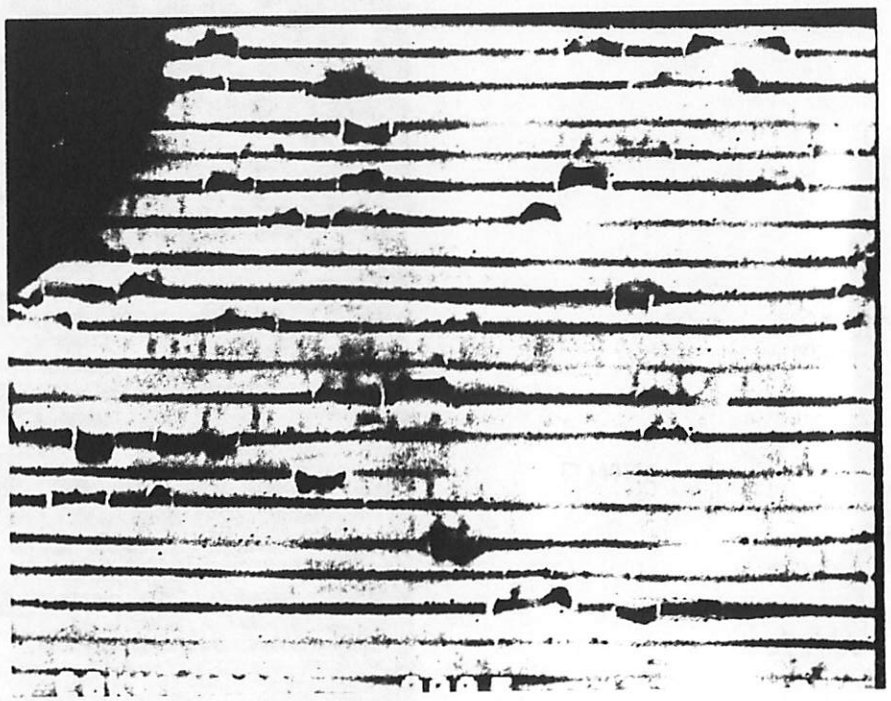
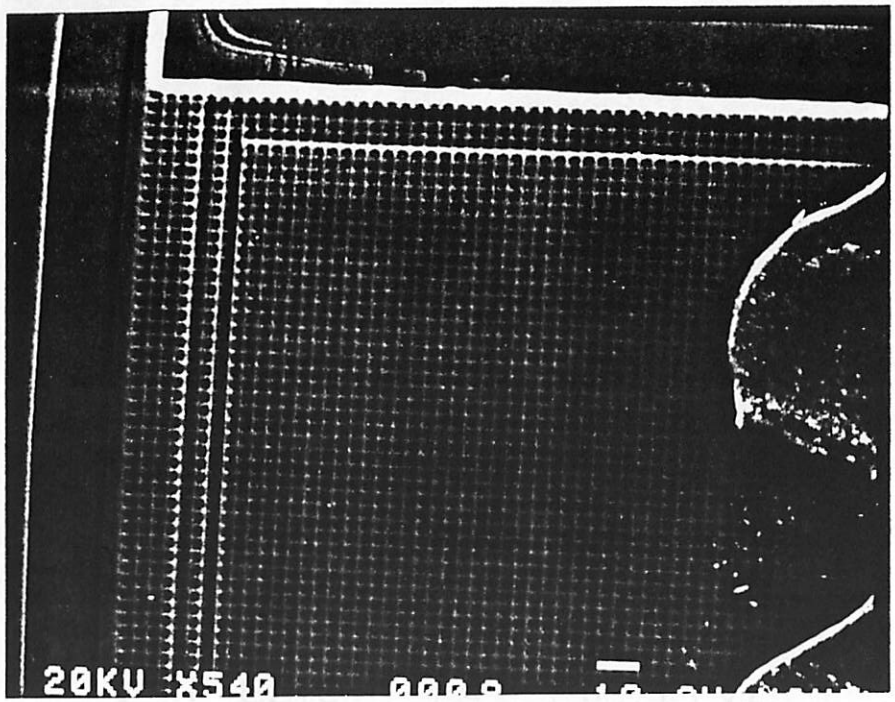
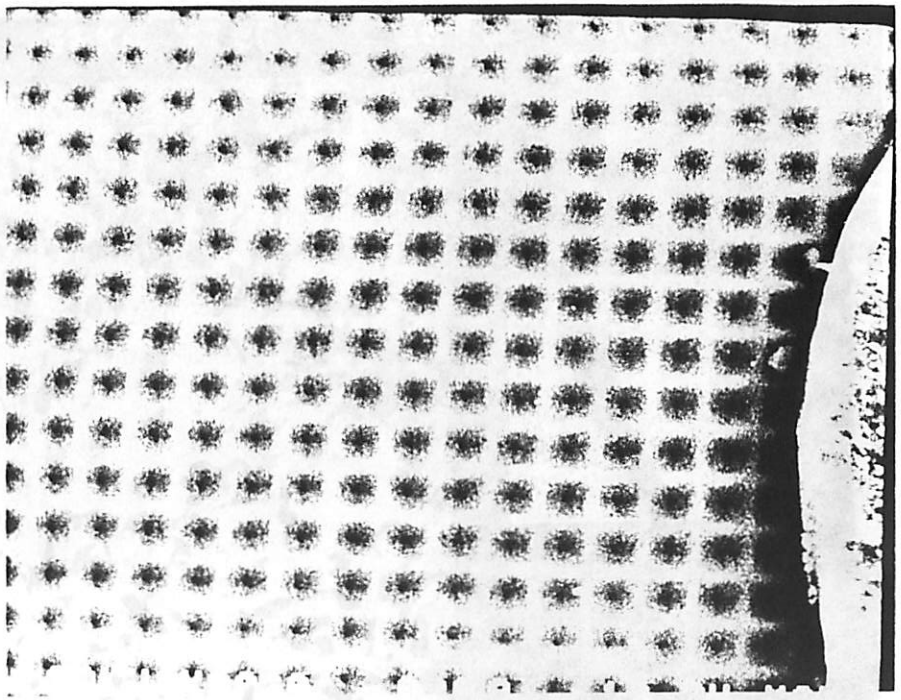


Fig. EBIC1





E8 (100) FLAT THIN EPI (3.3μm)

P=2.5
W=0.5
S=2.0
} SEED WINDOW EXPT.

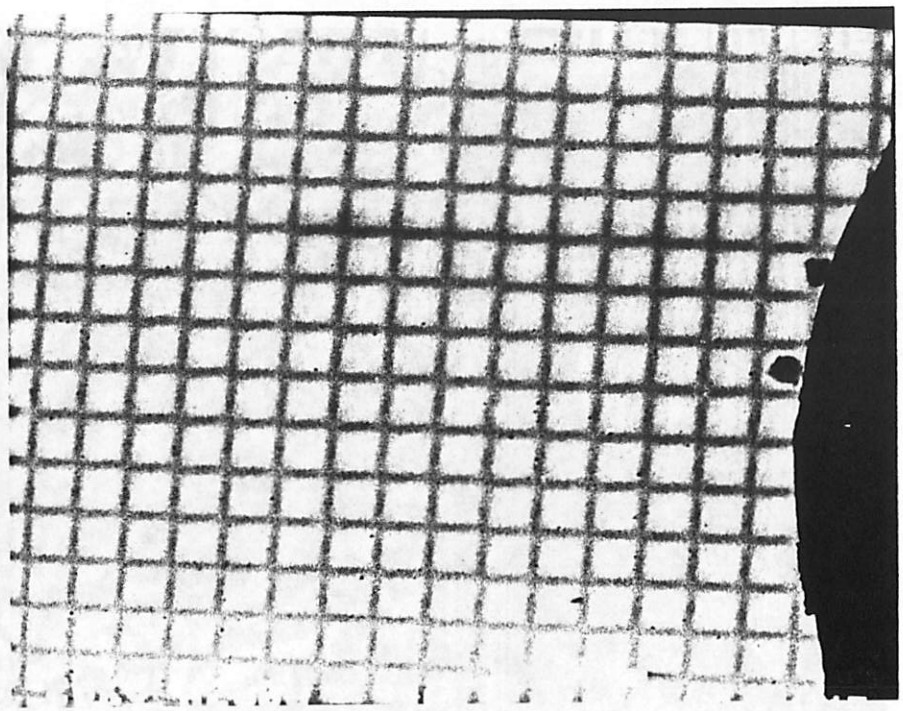
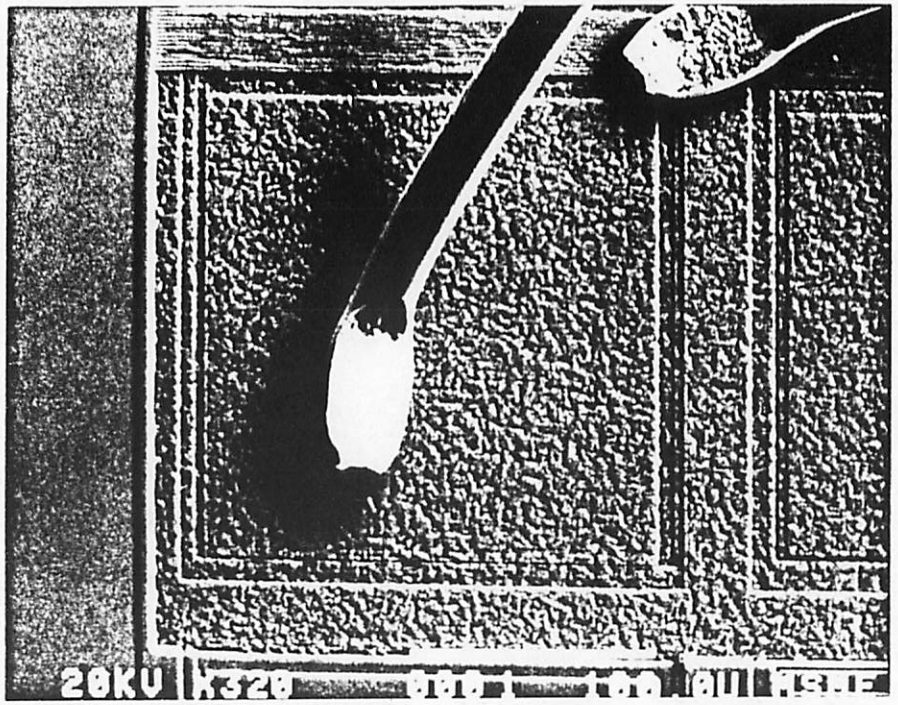
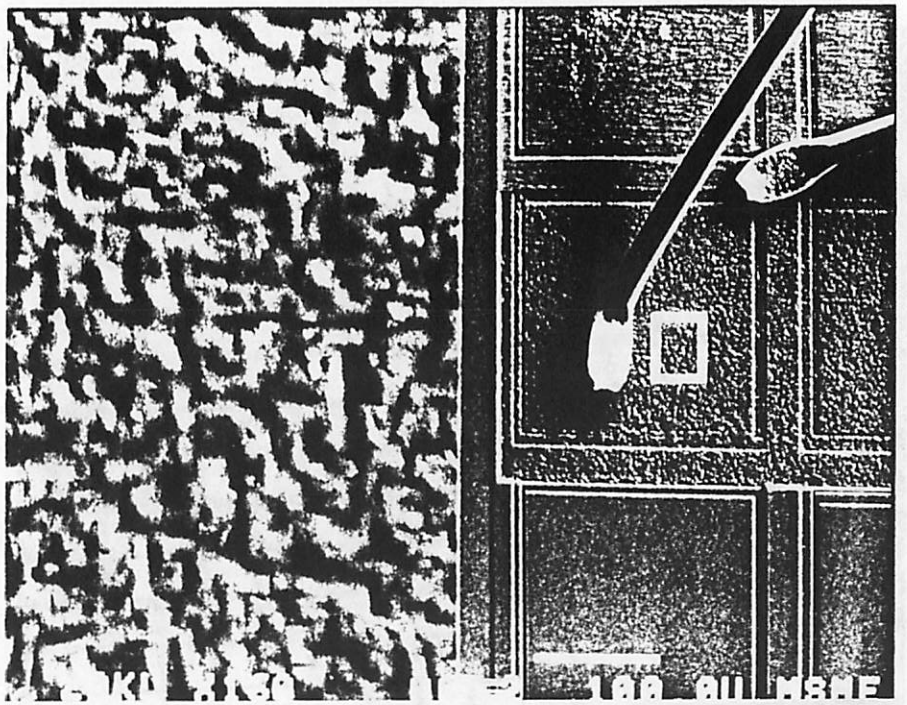


Fig. EBIC2



SE1



SEM EBIC
GRIDS
P= 2.0
W= 0.5
S= 1.5

E8 (100) FLAT
THIN EPI (3.3 μm)

SEED WINDOW EXPT.

EBIC

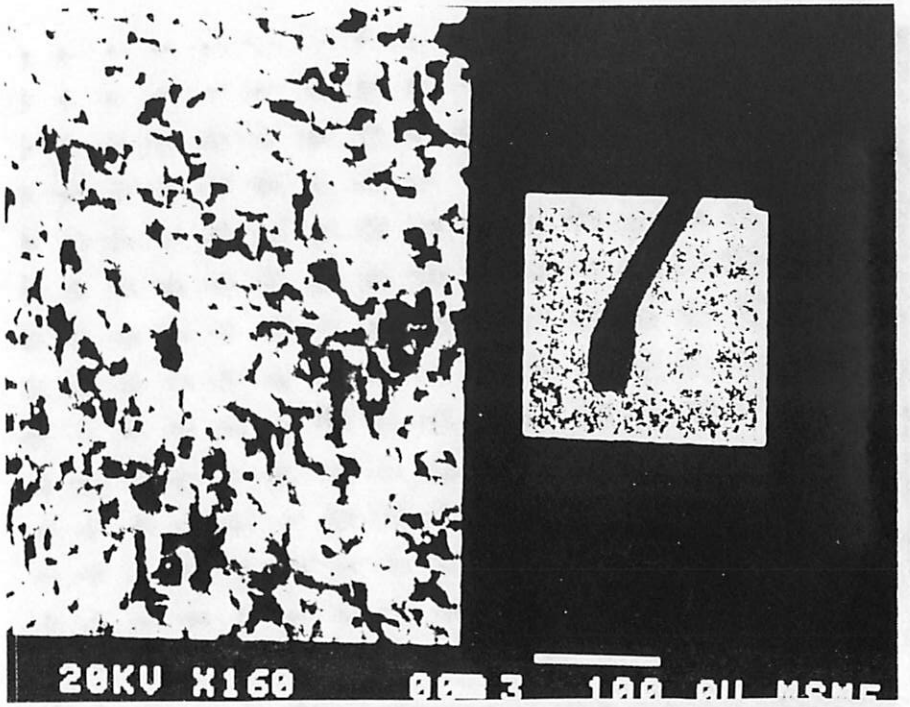


Fig. EBIC3