ELSEVIER

Sensors and Actuators A 89 (2001) 142–151

SENSORS AND ACTUATORS A PHYSICAL

www.elsevier.nl/locate/sna

In situ characterization of CMOS post-process micromachining

Brett Warneke*, Kristofer S.J. Pister

Berkeley Sensor and Actuator Center, University of California at Berkley, 497 Cory Hall Berkeley, CA 94720, USA

Abstract

We have developed and demonstrated a new methodology for in situ monitoring and characterization of CMOS post-process micromachining utilizing integrated circuits and micromachine test-structures. In our demonstration, the circuits provide automated readout of N-well resistors surrounding each of the 140 test pit structures at up to 14,000 samples per second per device during the post-process silicon etch and, thus, also provide etch progress and end point determination without extra analytical equipment. We use this technique to examine the effect of pit size, surrounding thin film layers, and topology in a 2 μm CMOS process with a XeF_2 post-process step, although the technique and results are of use to EDP, TMAH, and plasma post-processing.

Keywords: MEMS; CMOS post-processing; In situ characterization; Micromachining; CMOS integrated circuits; XeF_2

1. Introduction

Many MEMS devices utilize specialized processes and equipment and so require the MEMS designer to have access to a fabrication facility. This causes the process design to be tightly coupled to the design of the entire device. In contrast, integrated circuit processes have seen widespread use due to foundry services and looser coupling between circuit design and process design. To provide such capabilities to MEMS designers, CMOS post-processing was developed [1] to allow those without a fab to do micromachining with a single maskless post-processing step on standard foundry CMOS, thus requiring very little extra equipment. This method relies on the ability to stack the contact, via, and overglass cut layers (Fig. 1) to yield a 'pit' of exposed silicon when the chip returns from the foundry, which can then be sacrificially etched by Si etchants such as xenon difluoride (XeF_2) [2], tetramethylammonium hydroxide (TMAH), and ethylenediamine-pyrocatechol (EDP) [3] with the oxide acting as the mask and structural layer. This technique has been used to create a wide variety of devices including accelerometers [4], microwave power sensors [3], higher-Q spiral inductors [5], heart-cell transducers [6], thermal sensors [7], and thermally isolated circuits [8]. Although this method does not work in submicron processes that use tungsten plugs in the vias, it is still useful in the older, less expensive processes.

One problem with this approach is that it violates the design rules in most such CMOS processes, so the results are not guaranteed. Marshall et al. [9] at the National Institute of Standards and Technology (NIST) did work to develop a set of design rules for this method in the Orbit 2 μm process available through the MOSIS service.[1] At the time that this characterization was taking place, the process technicians at Orbit were discovering that the vias and overglass cuts were not clearing, so they performed overetches and over-exposures until they did clear. Because of the success of these runs, NIST published their design rules. However, the circuit designers that used the Orbit process began to complain that their pads and vias were being overetched, so Orbit stopped performing the extra steps that ensured that the pits cleared. With the process now changed, the new design rules were no longer valid. To make matters worse, the processing of the pits was no longer consistent from run to run [10].

The primary issue in properly fabricating the pits is that the resulting abrupt topography can prevent proper development of the photoresist that ends up in them during subsequent processing steps, leaving SiO_2 and Al residue in the corners of the pits (Fig. 2), which can moderate etching or even prevent it when the pit is filled. The minimum size pit that will not be plugged depends on several factors, including local topography, the surrounding layers, and orientation. It varies from run to run due to changes in die placement on the wafer and process variations. To quantify the effects of these factors, a set of 140 test pits was designed, along with a readout circuit to ease measurement. Tea et al. [3] previously described the use of *n*-diffusion resistors to precisely control the extent of the

* Corresponding author.

[1] http://www.mosis.org.

0924-4247/01/$ – see front matter
PII: S0924-4247(00)00543-4

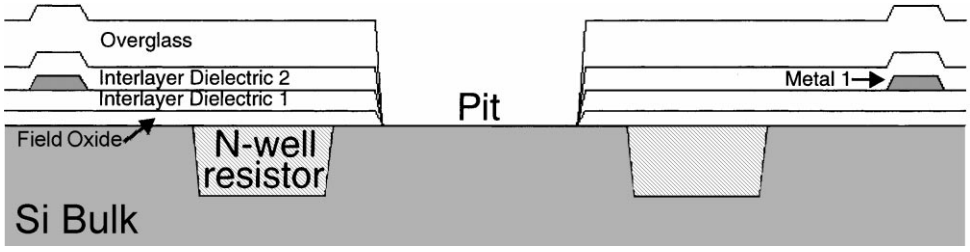

Fig. 1. Cross-section of the basic pit cell in Fig. 3 showing the N-well resistor used to detect etch progression. The cross-sections in this paper were made with L-Edit 7.3 in the Tanner MEMS-Pro 1.5[2] suite and have exaggerated vertical scales for clarity.

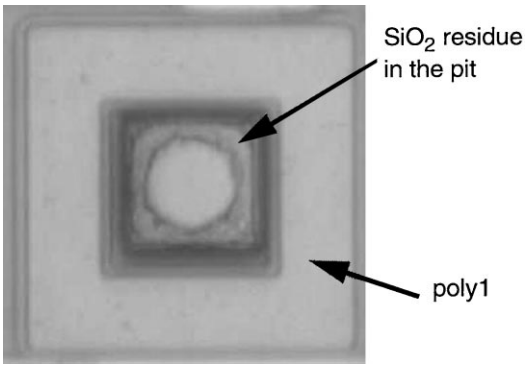

Fig. 2. Photomicrograph of a square pit surrounded by poly 1 showing the problematic pit residue.

XeF_2 Si etch by measuring the resistance during the etch to determine the location of the etch front, then modifying the etch rate parameters accordingly. We have extended this technique by utilizing the CMOS circuit capability to monitor the etch at many more etch fronts throughout the post-process, and to provide characterization of the preceding CMOS fabrication process.

2. Pit design

The basic pit cell (Fig. 3) consists of a pit surrounded by an N-well resistor that is progressively destroyed as the etch progresses. The resistance is nominally 14.4 kΩ, but as the etch front moves out from the pit and gradually destroys the resistor, the resistance will increase until the resistor has been totally destroyed. If the etch is allowed to proceed long enough, the row select transistor will also be destroyed, but without a detrimental effect on the circuit's performance. While this cell is designed for isotropic etchants such as XeF_2, it would only need some modification to work with anisotropic etchants such as EDP, TMAH, and plasmas. The primary change would be that the N-well resistor would be inside the pit instead of around it so that the vertical etch would destroy the device. However, one might choose just to do a post-etch characterization of the pits, rather than an in situ run, since operating circuits is more difficult in these etchants.

To determine the effect of several key features of the pits, a 14 × 10 array of pit cells was created. The right-hand half of the array varies the size of the pit from 5 to 30 μm in 5 μm increments along each row. Up each column, the stack of layers surrounding the pit varies from just the oxide films (Fig. 1) to various combinations of the dielectrics and poly 1 (P1), poly 2 (P2), metal 1 (M1), and metal 2 (M2)(Fig. 4) in order to determine how increasing the thickness of the region surrounding the pit at various points in the process affects the resulting amount of exposed silicon.

In an attempt to make the pits more manufacturable, the left-hand half of the array provides a slope to the edge of the pits by adjusting the hole size on each layer such that the overglass cut is 15 μm across and holes in underlying layers are progressively larger (Figs. 5–7). Thus, the topographical changes are more gradual and the film being removed is always on a nominally flat region, factors which were expected to decrease the minimum size hole that can be patterned. This approach is often used in dicing streets. Across each row the differential hole size from layer to layer is increased from 2 to 8 μm in 2 μm steps. Within each column the stack of layers surrounding the pit is adjusted similarly to the right-hand half of the array.

3. Circuit design

With such a large array of test structures and the desire to have a technique that could scale up to wafer level, a method of reducing the number of connections was necessary. Since the post-processing is performed on fully functional CMOS circuits and allows electrical feedthoughs, it was realized that on-chip circuits could be put to work automatically measuring the test pits, and in so doing reduce the pad count. This capability greatly reduces the amount of work necessary to examine the 140 pits and facilitates the use of this structure in monitoring run to run variations, just as one would automatically monitor electrical device variations. Furthermore, the wasted pad area and number of connections would be small enough that the circuit and array could be put in the dicing streets of a wafer and bondwires placed around the wafer for in situ measurement.

[2] http://www.tanner.com/eda/products/memspro/over.htm.

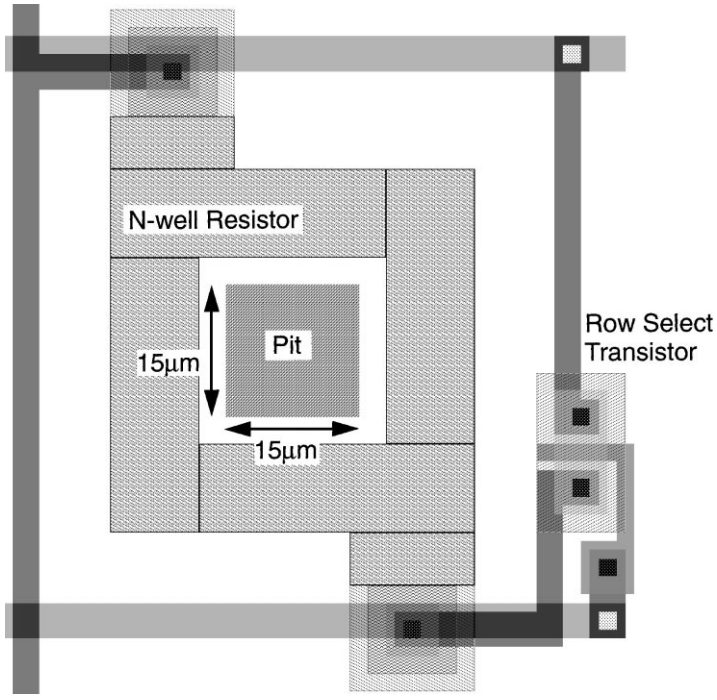

Fig. 3. Layout of a the basic CMOS pit cell showing a 15 µm × 15 µm pit surrounded by the N-well resistor, the row select transistor, and the access wiring.

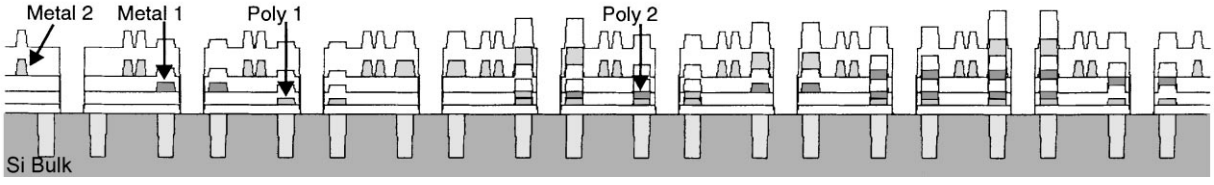

Fig. 4. Cross-section through a column of 15 µm × 15 µm pits showing the variation in layers surrounding the pits. The metal 2 row select and output wires are also visible between each pit.

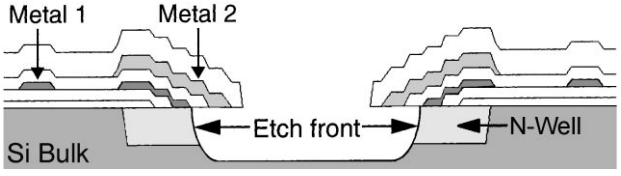

Fig. 5. This pit is surrounded by metals 1 and 2, yielding increased topography, and utilizes a gradual pit cut to ease the processing requirements. This approach is often used in dicing streets. This figure also portrays the pit mid-way through an isotropic etch showing the destruction of the N-well resistor.

The goal when designing the circuit was to minimize the number of pads, but still maintain an ability to know exactly which pit was being read from. The readout circuitry requires five electrical feedthroughs on the etch chamber: Vdd, GND, clock, reset, and output; the reset signal allows us to set the circuit in a known state then clock through it in a predetermined sequence. On chip, the clock signal is first converted to a non-overlapping two phase clock, as the circuit utilizes the pseudo-static logic style of flip-flops. The core of the circuit consists of two circular pointers, the

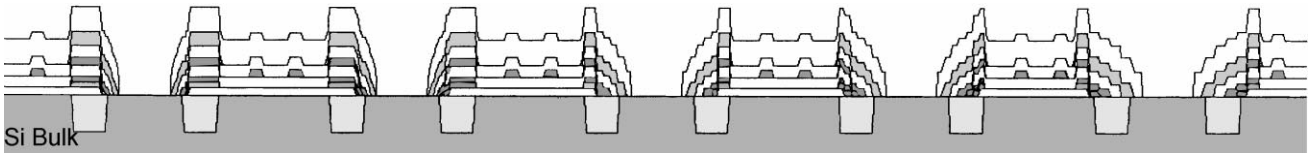

Fig. 6. Cross-section through part of a row of pits showing the variation in the slope to the pit as in Fig. 5. These pits are surrounded by both polysilicon and both metallization layers. The metal 1 column select and output wires are also visible between each pit.

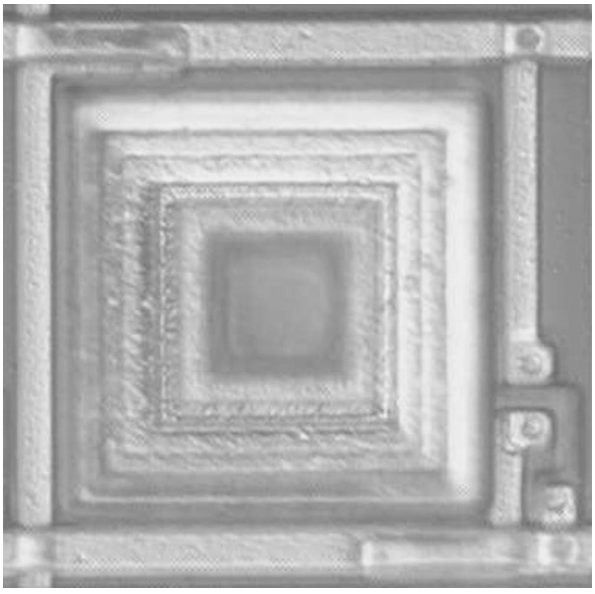

Fig. 7. Phase-contrast photomicrograph of the rightmost pit in Fig. 6 showing the stepping down into the pit.

first of which drives each column in turn, while the second drives each successive row as the column pointer passes from the last to first column (Fig. 8). Each column pointer cell has a digital buffer that acts as a current source driving

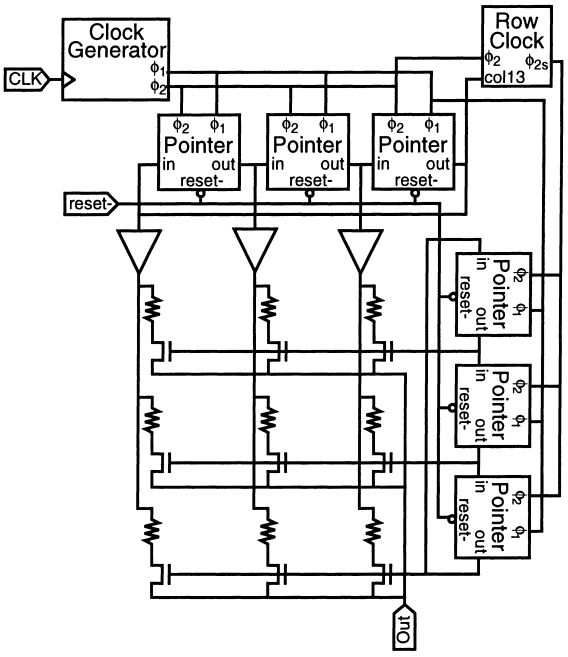

Fig. 8. Circuit schematic for a partial 3 × 3 pit array. The circuit contains two circular pointers that repetitively scan the entire array according to the clock. Each pit in the array is surrounded by an N-well resistor that is destroyed during etching. The circuit outputs a current proportional to the resistance of the selected pit's resistor and operates between dc and 2 MHz, but typical operation would be from 100 Hz to 1 kHz.

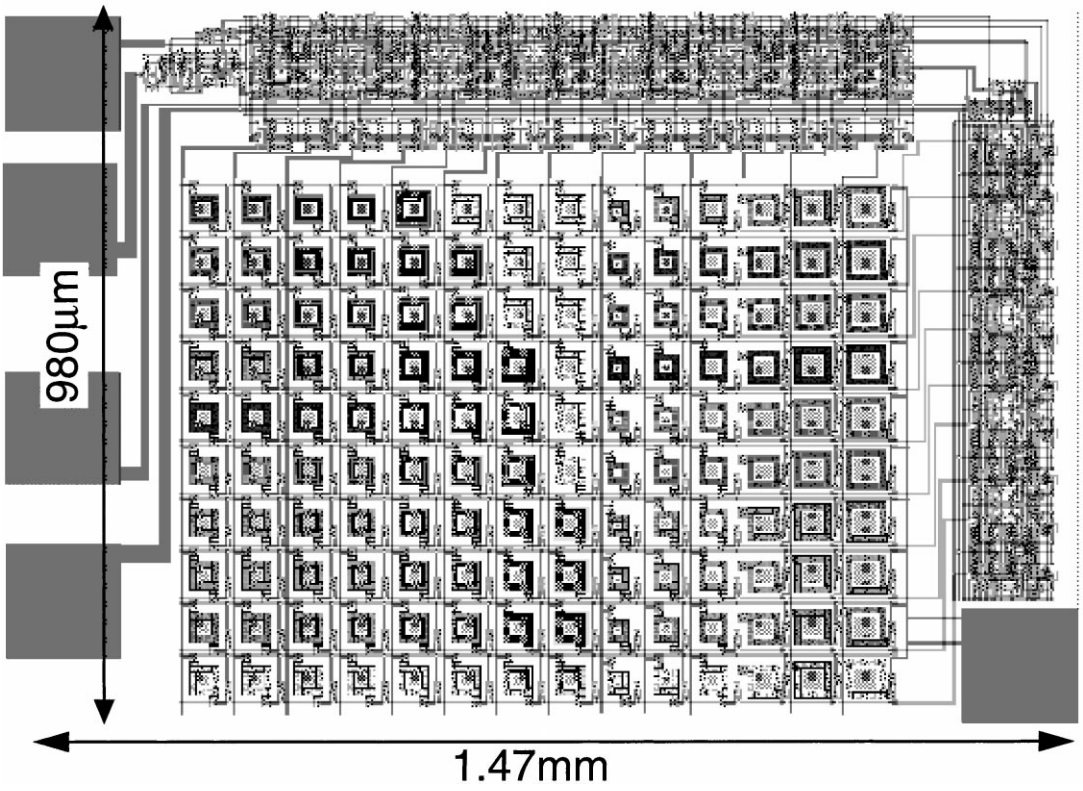

Fig. 9. 1.47 mm × 0.98 mm layout of the full circuit and pit array for the Orbit 2 μm process. Only five connections are necessary: GND, Vdd, reset, clock, and out.

the N-well resistors in the column. The row pointer cells then control *n*-channel pass transistors connected to each resistor of the row so that the current from the column driver only flows through the resistor in the currently selected row. The source of all the pass transistors are connected so that there is a single output from the array. In this manner, each element of the 14 × 10 array is driven in succession, yielding an output current that is dependent on the resistance of the N-well resistor surrounding the pit. As the etch front progresses through the resistor, the resistance increases until the element is open-circuited, resulting in zero output current.

The system has been fabricated through the MOSIS service in the Orbit 2 μm CMOS process (Fig. 9), which has two polysilicon layers and two metal layers. The entire circuit, including bond pads, is 1.47 mm × 0.98 mm. The die was epoxied to a ceramic package, wirebonded, and inserted into a breadboard that provided connections to the electrical feedthroughs on the XeF_2 etching chamber. Feedthroughs are relatively easy to implement in XeF_2 systems: in the past we have used commercial vacuum chamber BNC feedthroughs or drilled holes in the acrylic lid to feed wire through, then sealed the holes with vacuum epoxy. For the current work, a hole was milled in a 0.25 in. aluminum plate used as the lid for the etch chamber. DB25 connectors were sealed onto either side of the plate with vacuum epoxy to provide quick connections to the device in the chamber and the test equipment. An aluminum lid was used instead of the normal acrylic lid that allows visual inspection of the etch progress. An opaque lid is necessary to prevent photogenerated carriers that would affect circuit operation. This system can also be run in situ in other post-process etchants

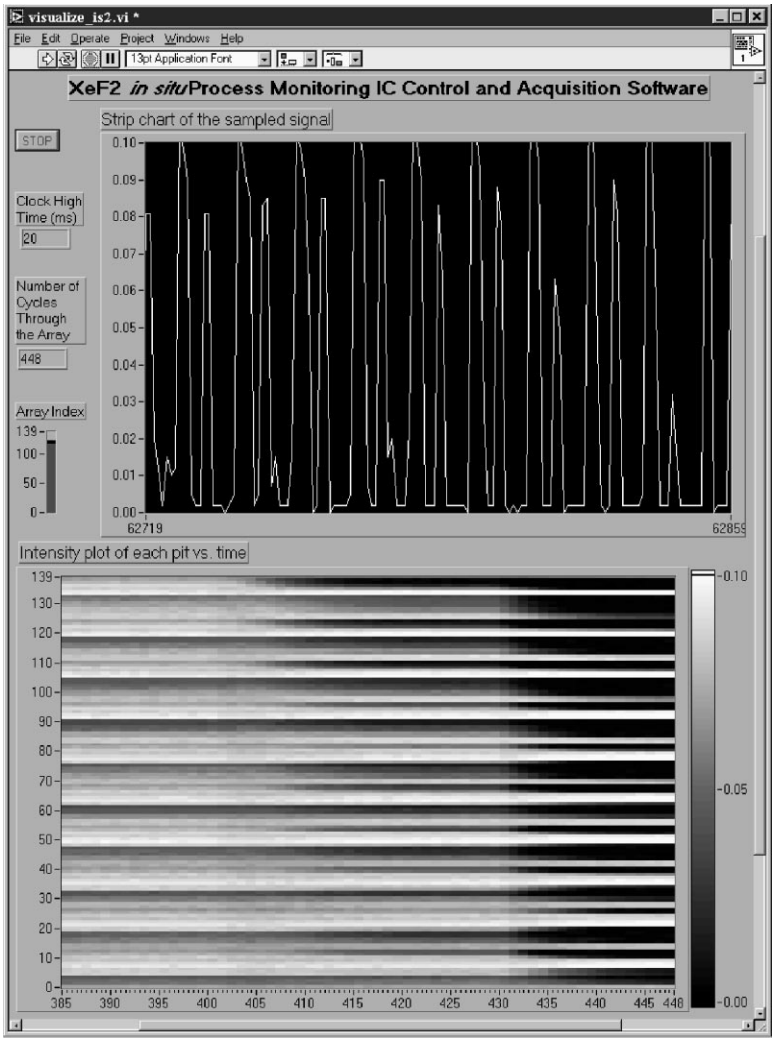

Fig. 10. Screen capture of the LabVIEW virtual instrument that generates the clock and reset signals to drive the circuit and reads the resulting output as the array is scanned. The top chart is a strip chart representation of the data coming directly out of the circuit for one entire scan of the array. The bottom chart is an intensity chart where the data from the top chart has been demultiplexed into individual pits such that the *y*-axis represents the pit number, the *x*-axis represents time (the number of the cycle through the array), and intensity is the relative value at the output. This capture was taken towards the end of the etch when a good number of pits have been etched through, giving a zero output. The bottom chart shows the pit signals gradually fading out, allowing rapid visualization of the large amount of data present being gathered. Periodic bands are apparent from the 5 and 10 μm columns in which no pits etched.

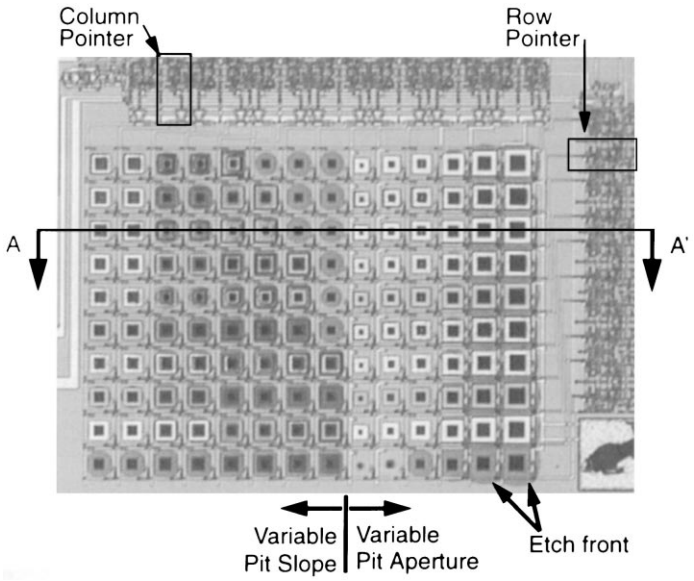

Fig. 11. Photomicrograph of the circuit after the post-process XeF_2 etch. Note the different amounts of undercut for each type of pit. The section used in Fig. 12 is indicated.

such as TMAH, although special considerations must be made for the feed-through wires, such as using Teflon™-coated wires [8].

The circuit has been measured to operate between dc (with non-50% duty cycle) and 2 MHz, providing up to 14,290 scans of the array per second. XeF_2 is a fast etchant at up to 15 μm/min [2], so to watch the etch progress through a 10 μm wide resistor, we would want 1–10 scans of the array per second. Thus, the circuit would typically operate between 100 Hz and 1 kHz.

A computer with a data acquisition card and LabVIEW[3] generates the input signals for the circuit and measures the voltage dropped across a 500 Ω resistor, allowing ready correlation between the pit currently being measured and its output. The LabVIEW virtual instrument we created (Fig. 10) shows the circuit output in a strip chart mode and on an intensity chart where the *y*-axis is the element of the array, the *x*-axis is time, and the color of the point corresponds to the output value. The latter chart provides a clear, succinct picture of the status of the etch for each pit and thus can also be used as an end-point determination. A second virtual instrument reads the data files generated by the first virtual instrument, reformats the data into the 14×10 array, and generates an animated intensity graph, allowing one to see the progress of the etch across the array, quickly visualize how the variations in the parameters affect the etch time, and filter out crosstalk between pits.

4. Measurements

A photomicrograph of the fabricated circuit after etching (Fig. 11) clearly shows the difference in etch progression among the various elements of the array, while a scanning electron micrograph (Fig. 12) of a die cleaved after etching provides an even more dramatic illustration. Fig. 13 shows the output wave-forms for the first row of the array at several times during the etch. The first cycle is before the etch has started, showing small variations in the output current due to differences in the N-well resistance. In the right side of the array, these variations are primarily due to the change in size of the resistor with pit size, while in the left side of the array, the variations are due to differing sizes of contact and via cuts whose overetches reduce the thickness of portions of the N-well and thus increase the resistance. The later cycles show the output changing as the resistors are etched, until after 565 cycles many of them are removed. Each cycle lasted 2–3 s.

Fig. 14 shows the waveforms from each of the 15 μm pits. The pits with no extra layers, poly 1, and metal 1 etched, while the pit with poly 1 and poly 2 started to etch before the experiment ended. The late beginning and slow rate of the etch for this pit indicates that the etch gas found a tiny hole to penetrate the residue, which on two other arrays it did not find. Similarly, given enough time some of the other pits may have etched, but they would also have low etch rates and low yield making them unsuitable for use. In this experiment the etch was terminated just before the faster etching pits began to attack their neighbors' resistors and contaminate the measurements. The steps in the falling waveforms are due to the pulsed nature of the etch [2]; 1.5 t of gas is released into the etch chamber for 30 s, the chamber is evacuated, and the cycle repeats. The rise in the data after some of the pits have begun etching is due to the fact that the columns are not driven by real switchable current sources that float when turned off, but by inverters. When the column is ''off'', it actually has an NFET current sink, so the current

[3] http://www.ni.com/labview.

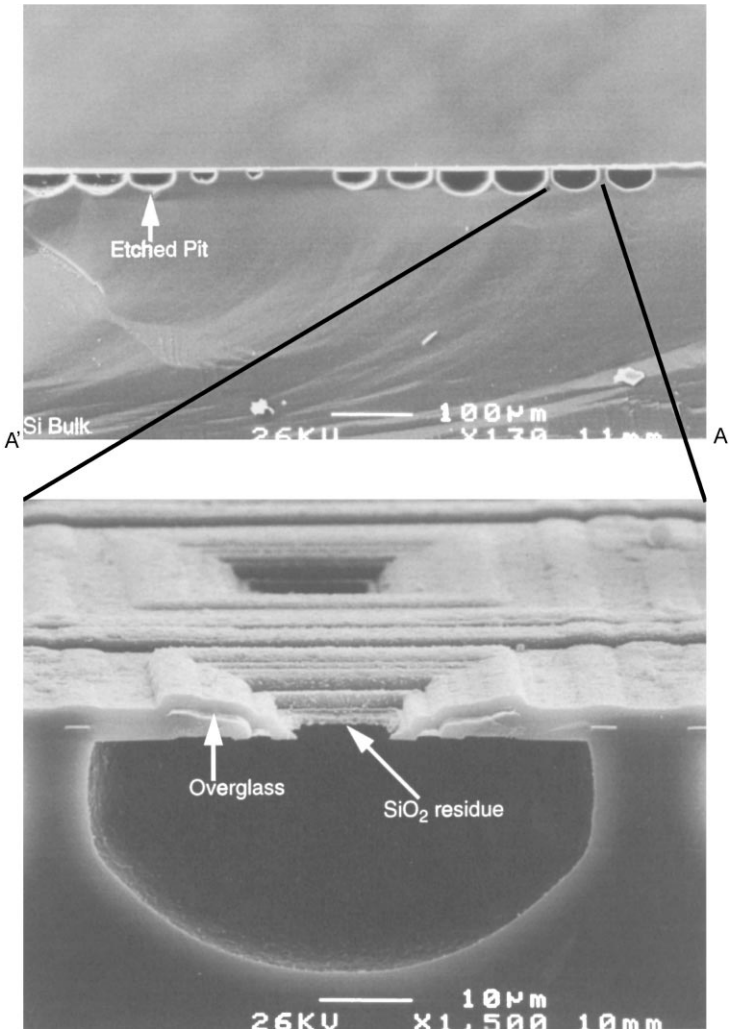

Fig. 12. Scanning electron micrographs (SEM) of the section indicated in Fig. 11 showing the variations in etch fronts and the unetched pits. The bottom SEM is a close-up of the fourth pit from the left in Fig. 11 and has a 4 μm stepped pit slope with poly 1, poly 2, and metal 2 surrounding the pit.

flowing through the ''on'' column resistor is split between the output branch and the other resistors in the same row since their pass transistors are also conducting. As the resistors are destroyed, less current is shared from an ''on'' column with them, so more current flows through the output branch (Fig. 15). The output resistor is about 30 times less than the nominal resistance around the pits, but there are up to 13 of those resistors, so at most a third of the current can be flowing through the ''off'' columns. If more accuracy is desired, such as to perform detailed etch system characterization, more ideal current sources that float when turned off would need to be utilized instead.

Table 1 summarizes the data gathered from the right side of three different arrays, one on die 1 and two on die 2. The relative etch times are calculated such that zero indicates the time that the first pit in each array etches its resistor through. We see that the array on die 1 did not have as clean of pits as the arrays on die 2, demonstrating the significant amount of die to die variation in these features; meanwhile, array 2b usually etched before array 2a. In addition, 10 identical, dielectric-only, 15 μm pits distributed throughout the array showed relative etch times between 27 and 67 on array 2a and between 24 and 35 on array 2b, while only one of the 10 etched at all on array 1b. Despite this significant amount of non-uniformity, some general trends can still be observed for this run. First of all, pits with no more than one extra layer can be made relatively reliably down to 20 μm, while for more layers 25 or 30 μm pits should be used. Secondly, metal 2 by itself does not yield as clean of pits as other films by themselves. If poly 1 and poly 2 are in the stack, however, then metal 2 does not make a significant difference over metal 1.

Data gathered from the left side of the three arrays is summarized in Table 2. The first three rows duplicate the 15 μm pit data from Table 1 as a reference point. For each step size, there is a version where the contact cut is actually the same size as the gate oxide (active) region to see if stepping is necessary so early in the process. The data shows that stepping the contact inside the active region helps some, but for step sizes above 4 μm, it does not make much

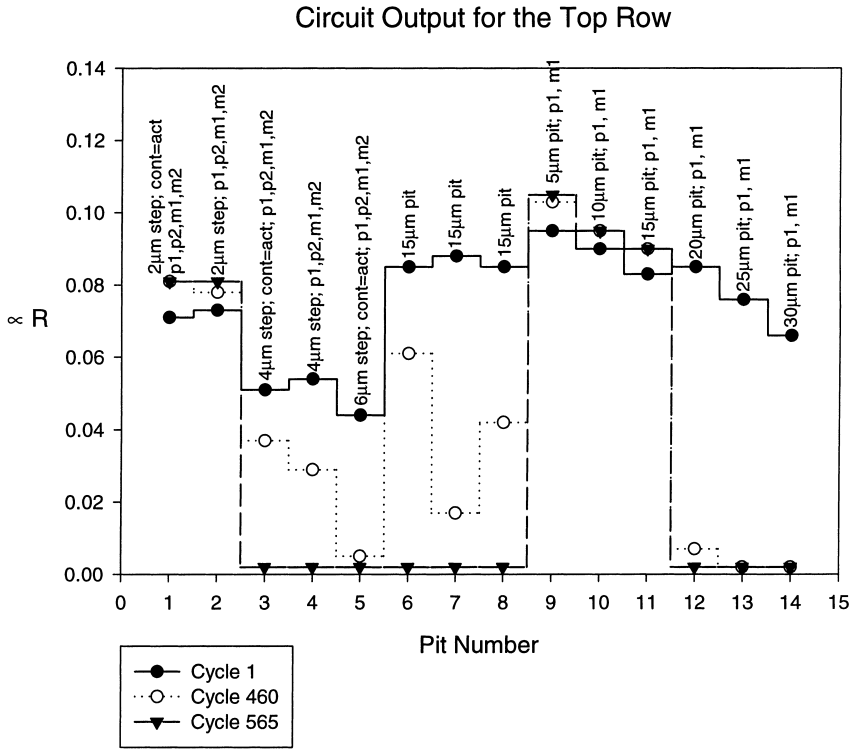

Fig. 13. Output waveforms for the first row of the test pit array at several times during the etch. Cycle 1 is before the etch, showing small variations due to differences in the N-well resistance. Cycle 460 shows the output part way through the experiment showing that the 25 and 30 μm pits have been totally etched, while the 20 μm pit, some of the 15 μm pits, and the 4 and 6 μm stepped pits are partially etched. At cycle 565 all of these pits have finished etching. Each cycle lasted 2–3 s.

Table 1
Relative etch times for various pit sizes in three arrays

Pit size (μm)	Array	Layers besides dielectrics									
		None	M1	P1	M2	P1, P2	P1, M1	M1, M2	P1, P2, M1	P1, P2, M2	P1, P2, M1, M2
30	1b	0	25	25	25	78	115	60	88	82	180[a]
	2a	0	5	5	15	7	12	27	24	24	32
	2b	0	2	2	10	9	11	12	14	22	27
25	1b	7	14	8	28	44	–[b]	63	63	34	–
	2a	2	5	5	22	12	5	30	25	27	40
	2b	0	5	5	15	9	20	12	21	21	28
20	1b	15	33	33	47	–	–	–	–	180[a]	–
	2a	7	17	17	27	30	30	50	107	42	107
	2b	4	10	10	23	20	27	29	60	37	72
15	1b	113	–	–	–	–	–	–	–	–	–
	2a	27	92	97	–	–	–	–	–	–	–
	2b	34	55	48	–	132	–	–	–	–	–
10	1b	–	–	–	–	–	–	–	–	–	–
	2a	382[a]	–	–	–	–	–	–	–	–	–
	2b	–	–	–	–	–	–	–	–	–	–
5	1b	–	–	–	–	–	–	–	–	–	–
	2a	882[a]	–	–	–	–	–	–	–	–	–
	2b	–	–	–	–	–	–	–	–	–	–

[a] Projected value; did not finish etching during the experiment.
[b] Did not etch.

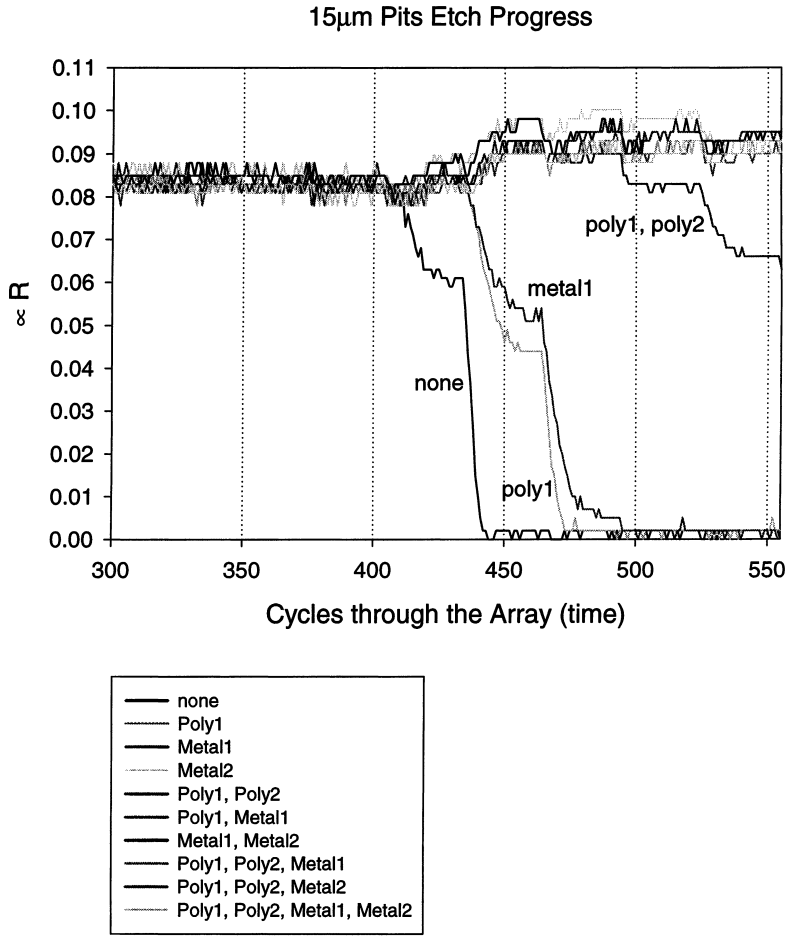

Fig. 14. Demultiplexed data for the column of 15 µm pits with one curve per pit, showing that the pits with no extra layers, poly 1, and metal 1 etched, the pit with poly 1 and poly 2 started to etch, and the other pits did not etch. For pit sizes larger than 15 µm, all the pits etched, while for pits smaller than 15 µm, only the pits with no extra layers etched. The steps in the falling waveforms are due to the pulsed nature of the etch.

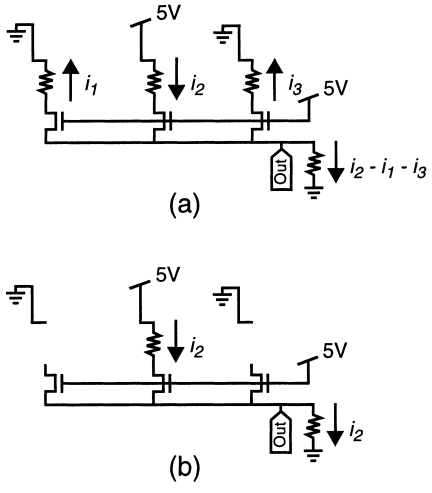

Fig. 15. Schematics of a three column row to illustrate why the data in Fig. 14 rises after some pits have etched: (a) before any pits have etched, the current flowing through the selected column is split between the output and the other columns in the row since the row transistors are all on; (b) after some resistors have been destroyed, more of the current can flow into the output, increasing the voltage dropped across the output resistor.

difference. Another deviation from the proposed stepping profile is that when poly is in the stack, the contact is actually larger than the hole in the poly, so there is a reverse step in the stack at that point. The primary observation from Table 2 is that reducing the slope into the pit does improve the yield on the pits, particularly for 4 µm step sizes and above.

5. Conclusion

We have demonstrated a new methodology for in situ monitoring and characterization of the foundry CMOS used for post-process CMOS micromachining and the post-process etching itself. The on chip circuits allow a large number of structures to be measured automatically as the etch is proceeding and provide end point determination. This technique allows design guidelines to be developed as to the smallest pit that will be open for a given set of layers surrounding the pit; a 20 µm pit was approximately the minimum for the Orbit 2 µm run evaluated in this study. In addition, new styles of pits that reduce the abrupt

Table 2
Relative etch times for 15 μm pits with various slopes and surrounding layers in three arrays

Step size (μm)/deviation	Array	Layers besides dielectrics									
		None	M1	P1	M2	P1, P2	P1, M1	M1, M2	P1, P2, M1	P1, P2, M2	P1, P2, M1, M2
0	1b	113	–[a]	–	–	–	–	–	–	–	–
	2a	27	92	–	–	–	–	–	–	–	–
	2b	34	55	48	–	132	–	–	–	–	–
2 (contact = active)	1b	–	–	–	–	–	–	–	–	–	–
	2a	36	72	59	87	134	–	–	–	–	–
	2b	35	79	57	68	67	–	–	–	–	–
2	1b	–	–	–	–	–	–	–	–	–	–
	2a	28	64	53	69	102	–	–	–	–	–
	2b	32	63	50	63	60	–	362	–	–	–
4 (contact = active)	1b	31	100	35	125	42	–	–	–	–	–
	2a	21	24	21	24	21	–	40	25	33	52
	2b	22	24	22	25	22	–	36	25	32	38
4	1b	30	80	66	83	43	133	–	108	–	–
	2a	21	24	21	24	23	29	57	32	42	52
	2b	23	26	23	25	24	27	37	25	35	35
6 (contact = active)	1b	13	18	14	40	18	92	27	33	34	–
	2a	19	21	20	20	19	29	19	18	19	30
	2b	20	20	20	21	19	26	20	19	19	25
6	1b	15	20	14	–	20	20	–	31	33	
	2a	21	20	20	21	21	21	21	19	21	
	2b	21	19	19	21	19	19	20	19	19	
8 (contact = active)	1b	5	6	7	10	8	15	28			
	2a	20	18	18	18	19	18	18			
	2b	20	18	19	18	18	19	18			
8	1b	11	16	7	30						
	2a	19	18	18	18						
	2b	20	19	19	19						

[a] Did not etch.

topographies have been shown to increase the yield of small pits. Besides its use in process characterization, this concept could be used to provide automatic end-point detection wherein the circuits on a chip or wafer would tell the etch system exactly when to terminate the etch.

Acknowledgements

Support for this work was provided by the Howard Hughes Doctoral Fellowship. Special thanks to Josie Kung and Ron Wilson for sample preparation and SEM operation.

References

[1] M. Parameswaran, H.P. Baltes, L. Ristic, A.C. Dhaded, A.M. Robinson, A new approach for the fabrication of micromechanical structures, Sens. Actuators 19 (1989) 289–307.

[2] P.B. Chu, J.T. Chen, R. Yeh, G. Lin, J.C.P. Huang, B.A. Warneke, K.S.J. Pister, Controlled pulse-etching with xenon difluoride, in: Proceedings of the International Conference on Solid-State Sensors and Actuators — TRANSDUCERS'97, Chicago, USA, 16–19 June 1997, pp. 665–668.

[3] N. Tea, V. Milanovic, C. Zincke, J. Suehle, M. Gaitan, M. Zaghloul, J. Geist, Hybrid post-processing etching for CMOS-compatible MEMS, IEEE J. Microelectromech. Syst. 6 (1997) 363–372.

[4] E.J.J. Kruglick, B.A. Warneke, K.S.J. Pister, CMOS 3-axis accelerometers with integrated amplifier, in: Proceedings of the IEEE 11th Annual International Workshop on Microelectromechanical Systems (MEMS'98), Heidelberg, Germany, 25–29 January 1998, pp. 631–636.

[5] A. Rofougaran, J.Y.-C. Chang, M. Rofougaran, A.A. Abidi, A 1 GHz CMOS RF front-end IC for a direct-conversion wireless receiver, IEEE J. Solid-State Circuits 31 (1996) 880–889.

[6] G. Lin, R.E. Palmer, K.S.J. Pister, K.P. Roos, Single heart cell force measured in standard CMOS, in: Proceedings of the International Conference on Solid-State Sensors and Actuators — TRANSDUCERS'97, Chicago, USA, 16–19 June 1997, pp. 199–200.

[7] H. Baltes, O. Paul, O. Brand, Micromachined thermally based CMOS microsensors, in: Proceedings of the IEEE, Vol. 86, 1998, pp. 1660–1678

[8] E.H. Klaassen, R.J. Reay, C. Storment, G.T.A. Kovacs, Micromachined thermally isolated circuits, Sen. Actuators A 58 (1997) 43–50.

[9] J. Marshall, M. Gaitan, M. Zaghloul, D. Novotny, V. Tyree, J.-I. Pi, C. Piñá, W. Hansford, Realizing suspended structures on chips fabricated by CMOS foundry processes through the MOSIS service, NIS-TIR 5402, US Department of Commerce, June 1994.

[10] Personal communication with J. Marshall.