Aspect ratio dependent etching lag reduction in deep silicon etch processes

S. L. Lai,[a] D. Johnson, and R. Westerman
Unaxis USA, Inc., Saint Petersburg, Florida 33716

(Received 12 October 2005; accepted 9 January 2006; published 22 June 2006)

Microelectromechanical system (MEMS) device fabrication often involves three dimensional structures with high aspect ratios. Moreover, MEMS designs require structures with different dimensions and aspect ratios to coexist on a single microchip. There is a well-documented aspect ratio dependent etching (ARDE) effect in deep silicon etching processes. For features with different dimensions etched simultaneously, the ARDE effect causes bigger features to be etched at faster rates. In practice, ARDE effect has many undesired complications to MEMS device fabrication. This article presents a physical model to describe the time division multiplex (TDM) plasma etch processes and thereafter the experimental results on ARDE lag reduction. The model breaks individual plasma etch cycles in the TDM plasma etch processes into polymer deposition, polymer removal, and spontaneous silicon etching stages. With the insights gained from the model and control over the passivation and etch steps, it has been demonstrated that ARDE lag can be controlled effectively. Experiments have shown that a normal ARDE lag can be changed to an inverse ARDE lag. Under optimized conditions, the ARDE lag is reduced to below 2%–3% for trenches with widths ranging from 2.5 to 100 μm, while maintaining good etch profile in trenches with different dimensions. Such results are achieved at etch rates exceeding 2 μm/min. © *2006 American Vacuum Society.* [DOI: 10.1116/1.2172944]

I. INTRODUCTION

In recent years, the time division multiplex (TDM) plasma etch processes, in which alternating deposition and etching steps are used cyclically, have been developed and extensively applied in manufacturing silicon based microelectromechanical system (MEMS) devices.[1,2] Such structures frequently have depths ranging from tens to hundreds of micrometers. The advantages of the TDM plasma etch processes over the conventional single-step plasma etch processes include high etch rates ($\geqslant$10 μm/min), high etch selectivity to mask materials (e.g., $\geqslant$200:1), good anisotropy control, smooth sidewall (e.g., $\leqslant$20–50 nm), and other etch performance parameters.

There is a well-documented aspect ratio dependent etching (ARDE) phenomenon in deep silicon etching using the TDM plasma etch processes, as well as in single-step plasma etch processes.[3] For instance, one would observe that the silicon etch rate decreases as the depth or aspect ratio (AR, defined as feature depth divided by feature width) increases. ARDE is a highly complex phenomenon and many mechanisms are proposed to explain it. In general, many factors contribute prominently to ARDE: (i) ion flux loss at the bottom of the etched structure[3–7] and (ii) reactive neutral species depletion due to neutral shadowing[8] and Knudsen transport.[9,10] When using the TDM plasma etch processes to fabricate three dimensional structures, the ARDE effect can be manifested in two ways. First, as an etch process proceeds, the aspect ratio of a specific feature, such as a trench or via, increases with increasing etch time and the etch rate decreases over time. Second, when features of different dimensions are present on the same substrate and etched simultaneously the bigger features are etched at faster rates than the smaller features, resulting in ARDE lag. An example of ARDE lag is shown in the cross-sectional scanning electron microscopy (SEM) image in Fig. 1(a). The silicon trenches are etched with a Unaxis inductively coupled plasma (ICP) etcher operated at an ICP power level of 1200 W and a rf bias power at 18 W. The chamber pressure in the deposition step is maintained at 20 m Torr, while the pressure in the etch step is maintained at 70 m Torr. The deposition and etch step times are 2.0 and 6.0 s, respectively. While the 100-μm-wide trench is etched to a depth of 130 μm, the 2.5-μm-wide trench is only etched to a depth of 62 μm. The ARDE lag, sometimes called the reactive ion etching (RIE) lag, is measured as trench depths are normalized to that of a 100-μm-wide trench, and the result is given in Fig. 1(b). In this case, an ARDE lag over 50% is exhibited. Indeed, the ARDE effect presents undesirable complications to MEMS device fabrication. When structures with various lateral dimensions coexist and are etched at the same time, the resulted vertical dimensions are different, which may be incompatible with device design requirements. Even for a single structure, as etching progresses, the vertical etch rate is not constant, which may represent a challenge to good process control.

Many researchers have conducted experimental investigations aiming to reduce ARDE lag. For example, in single-step etch processes, Doh *et al.* described the reduction of RIE lag at increased bias voltage and increased bias frequency in an electron cyclotron resonance (ECR) plasma etching system.[11] Lill *et al.* reported reduced RIE lag at high pressure

[a]Electronic mail: shouliang.lai@unaxis.com

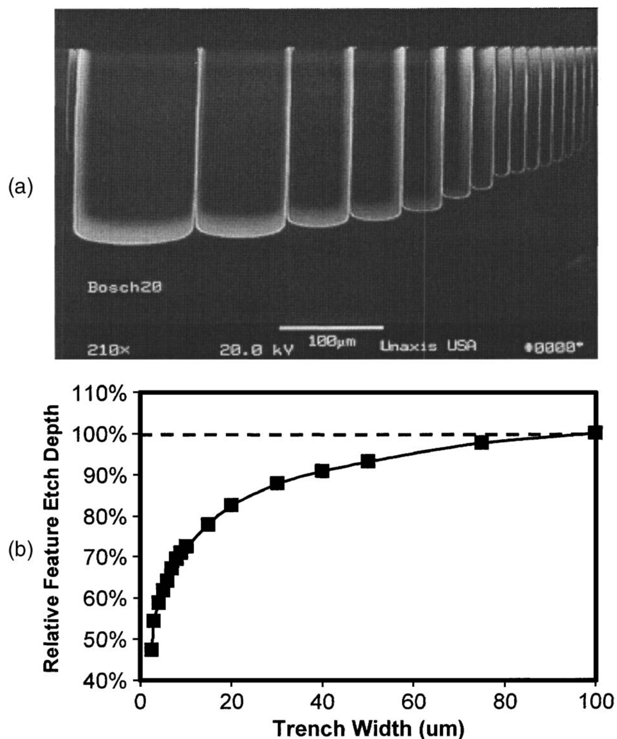

FIG. 1. (a) ARDE lag exhibited in typical TDM plasma etch process. The final etch depths in wider trenches are larger. (b) A plot of trench depth normalized to that of the 100-μm-wide trench. An ~50% ARDE lag is observed in the 2.5-μm-wide trench.

and medium cathode temperature when etching polysilicon with SF_6 plasma in an ICP system.[12] Tsujimoto *et al.* studied a method in which gas residence time is reduced to mitigate RIE lag when polysilicon is etched in a Cl_2 plasma.[13] In TDM plasma etch processes, one widely used method is to employ an etch stop layer. In silicon-on-insulator (SOI) and silicon-on-glass (SOG) wafers, a buried oxide layer is used to stop silicon etching to compensate for the ARDE lag. However, two disadvantages emerge when SOI or SOG wafers are used: (1) notching occurs at the silicon/oxide interface,[14] and (2) SOI and SOG wafers are generally more expensive than silicon wafers. Chung and Lu disclosed that by raising the process pressure in both the etch and deposition steps the ARDE lag can be reduced.[15] Rickard *et al.* discovered that ARDE is minimized through shorter etch time, low pressure, low rf bias power, and increased deposition time,[16] while Ayon *et al.* reported the minimization of ARDE for TDM processes by using high SF_6 flow rates.[17]

As demonstrated above, most previous research in the ARDE lag reduction in plasma etch processes has focused on the empirical aspects. In particular, for the ARDE lag reduction in the TDM plasma etch processes the experimental findings are often contradictory. In this work, we first present a model for the TDM plasma etch processes based on the principles of the process schemes and thereafter use the experimental evidence to construct a different approach to reduce the ARDE lag. Experimental demonstrations in controlling and reducing the ARDE lag will also be presented.

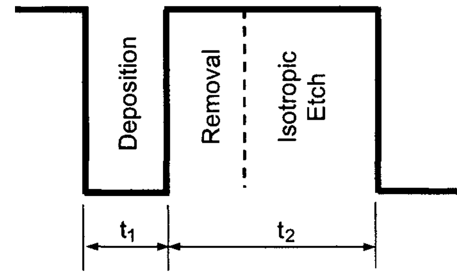

FIG. 2. Three-stage model for TDM plasma etching processes. In a complete deposition/etch cycle, the deposition step lasts for a period of time t_1, and the etch step for t_2. The etch step is divided into a polymer removal stage and a spontaneous isotropic silicon etching stage. a_1, a_2, and a_3 are polymer deposition rate, polymer removal rate, and isotropic silicon etching rate, respectively.

II. THREE-STAGE MODEL

The principal steps in TDM plasma etch processes were described in detail previously.[18] The TDM processes consist of multiple steps and a set of steps (e.g., etch and deposition) is called a cycle. In a complete deposition/etch cycle, the deposition step time and the etch step time are denoted t_1 and t_2, respectively, as shown in Fig. 2.[19] During the deposition step, deposition precursor gas such as C_4F_8 is dissociated by the plasma to form the ion and radical species, which undergo polymerization reactions to form a polymeric layer, probably $(CF_2\text{-})_n$, on the surface of silicon structures. In the etch step, SF_6 gas is used to promote spontaneous silicon etching. The etch step essentially consists of two different stages. In the first stage, the polymer already deposited at the bottom of the structures is preferentially removed. This process is facilitated by directional ion bombardment. In the second stage, which starts after the polymer deposited at the bottom surface is completely removed, F radicals cause spontaneous and isotropic Si etching by adsorption followed by formation and desorption of volatile products such as SiF_4. Therefore, a complete deposition/etch cycle can be divided into three distinctive but sequential physical stages: polymer deposition stage, polymer removal stage, and isotropic Si etching stage.

The deposition stage can be characterized by the polymer deposition rate. Since the polymer formed at a feature's bottom surface is the most relevant, the polymer deposition rate at the bottom surface is denoted a_1. Then the thickness of polymer layer deposited at the bottom surface can be expressed as $a_1 \cdot t_1$. In the polymer removal stage, if the polymer removal rate from the bottom surface is denoted a_2, the time required to completely remove the polymer layer will be $a_1 \cdot t_1 / a_2$. Thus the remaining time for the spontaneous and isotropic etching of silicon can be expressed as

$$t = t_2 - \frac{a_1 \cdot t_1}{a_2}. \qquad (1)$$

The isotropic Si etching stage is characterized by the silicon etch rate, denoted a_3. The etch depth accomplished during

this last stage will represent the overall silicon etch depth during a complete deposition/etch cycle. As such, the overall silicon etch depth, L, and the overall silicon etch rate, R, achieved during the complete deposition/etch cycle can be, respectively, expressed as

$$L = a_3 \cdot t = a_3 \cdot \left(t_2 - \frac{a_1 \cdot t_1}{a_2} \right), \tag{2}$$

$$R = a_3 \cdot t \cdot \frac{1}{t_1 + t_2} = a_3 \cdot \left(t_2 - \frac{a_1 \cdot t_1}{a_2} \right) \cdot \frac{1}{t_1 + t_2}. \tag{3}$$

It should be pointed out that assumptions are made in this model. These assumptions include that (1) polymer is uniformly deposited at the horizontal bottom surface, (2) polymer removal at the horizontal bottom surface also occurs in a uniform manner, and (3) vertical sidewalls have no other influence on polymer deposition, polymer removal, or isotropic Si etching at the horizontal bottom surface. The last assumption particularly implies that the previously exposed vertical sidewalls are always under the protection of polymer passivation and thus no lateral etching of Si occurs at the sidewalls. In other words, this model assumes an overall anisotropic etch profile. The aforementioned ARDE effects should exhibit in all three stages, namely, the polymer deposition stage, the polymer removal stage, and the isotropic Si etching stage. More specifically, such ARDE effects should be exclusively reflected in the polymer deposition rate, a_1, polymer removal rate, a_2, and spontaneous Si etching rate, a_3.

In a typical TDM plasma etch process, the polymer deposition step uses C_4F_8 as a gas precursor. The deposition process can be performed with or without applying rf bias. In the case with no applied rf bias, there is minimal ion bombardment of the substrate. Figure 3(a) shows experimental measurements of the polymer deposition rate as a function of feature aspect ratio. Over the range of plasma pressure tested, the deposition rate measured at the bottom of the silicon trench decreases with increasing aspect ratio. The strong AR dependence is attributed to the transport of radicals into the trenches.

In contrast to the polymer deposition process, an Ar/SF_6 based polymer passivation removal process is primarily ion driven, though in some cases the addition of O_2 gas during the etch step will drive the process toward an ion-assisted chemical mechanism. Figure 3(b) shows the experimental results on polymer removal rate at the trench bottom over a range of aspect ratios. As demonstrated by the measurement results, polymer removal rate is relatively weakly aspect ratio dependent under the process conditions, with the polymer removal being slightly faster in the lower aspect ratio features. Under normal TDM plasma etch conditions, the mean free path of ions is close to ~1 mm. Thus, when ions are driven to the trench bottom by a dc electrical field of a few hundreds of volts, the loss of ion flux due to collision and scattering at feature walls is minimal and not affected by aspect ratio as significantly.

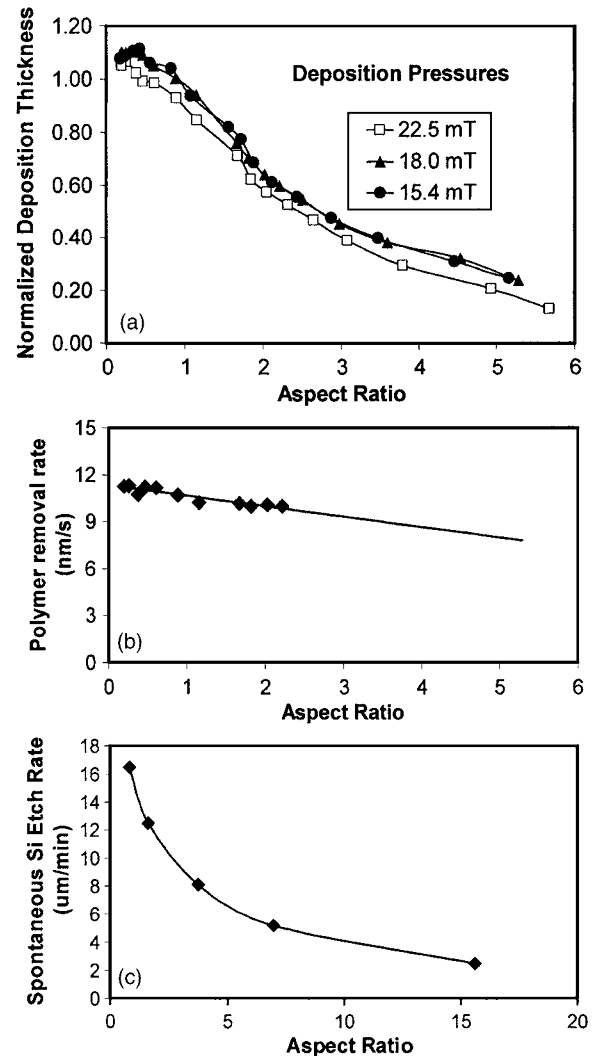

FIG. 3. Experimental measurements on (a) normalized thickness of polymer deposited at trench bottom (normalized to the thickness of polymer deposited on the horizontal mask surface), (b) polymer removal rate, and (c) spontaneous silicon etching rate as functions of aspect ratio.

The spontaneous Si etching process is more chemical in nature. Typically in a TDM plasma etch process, the rf bias in this step is set to relatively low values that enable acceptable feature profiles while maximizing the etch selectivity to mask materials. In this stage, however, the transport of F radicals and etch products into and out of deep features will again play a more important role in etching efficiency. Figure 3(c) shows the spontaneous silicon etch rate as a function of aspect ratio. The silicon etch rate is a strong function of aspect ratio, with the features of smaller aspect ratios having significantly higher etching rates.

Conceptually, one should be able to control ARDE lag with the insights gained from the aspect ratio dependence of a_1, a_2, and a_3. Figure 4 depicts two features with different widths. Let us assume that both features have the same depth before a new TDM cycle starts [Fig. 4(a)]. By the end of the deposition step, the deposited polymer film in the wider feature is thicker than in the narrower feature due to the higher polymer deposition rate in lower aspect ratio feature [Fig.

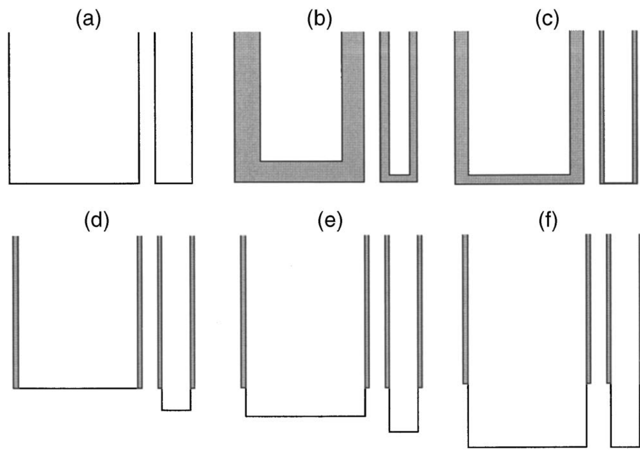

FIG. 4. Schematic illustration on ARDE lag control in two features with different sizes. Inverse (points D and E) and zero (point F) ARDE lags can be obtained by terminating the process cycle at proper points of time.

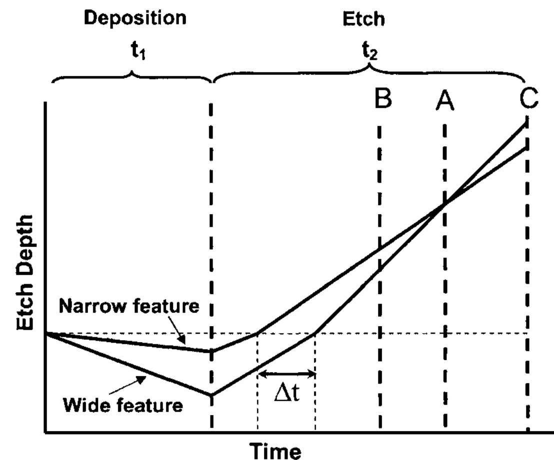

FIG. 5. Graphic illustrations on controlling ARDE lag by controlling the process step time.

4(b)]. Then the first portion of the etch step serves to remove the polymer passivation at the feature bottom before the isotropic silicon etching can take place. Since the polymer removal rate is nearly independent of aspect ratio and the narrow feature starts with a thinner polymer layer, the narrow feature will clear the polymer layer first [Fig. 4(c)]. This allows the silicon etching to proceed in this narrow feature while the remaining polymer in the wider feature continues to clear [Fig. 4(d)]. Essentially the narrow feature gets a "head start" in silicon etching. By the time the polymer in the wider feature has cleared, the narrow feature has etched some depth into the silicon. Once the polymer has cleared from in wider feature, it begins to be etched at a higher rate than the narrow feature [Fig. 4(e)]. During the remaining time period both the wide and narrow features etch further into the silicon. The narrow feature is deeper due to the increased isotropic etch time, but the wider feature is etched at a faster rate. At a certain point of time, the wide and narrow features have reached the same depth [Fig. 4(f)]. If the TDM cycle ends at this point of time, perfect ARDE lag elimination is achieved during this complete TDM cycle. Additional iterations of TDM etch cycles will also result in zero ARDE lag, as long as every TDM cycle is so maintained that ARDE lag is eliminated during the cycle.

This illustration is recaptured graphically in Fig. 5. Here, the polymer deposition rate, polymer removal rate, and spontaneous silicon etching rate are represented by the slopes of the solid straight lines. A steeper slope represents a higher rate. The case of complete ARDE lag elimination is now described as the etch step stops at point A. Two other instances can be also visualized using Fig. 5. If a TDM cycle ends at point B which is earlier in time than point A, "inverse" ARDE lag will result. A narrow feature will be etched deeper than a wide feature. If a TDM cycle ends at a later point C, "normal" ARDE lag will result, as observed so often in practice. In this case, a narrow feature will be etched to a shallower depth than a wide feature. Also based on the model, it is possible to adjust the TDM process to compensate for the different starting depths to achieve zero ARDE lag as the etch proceeds, if the features start at different depths. This compensation may take one or more TDM cycles to accomplish.

The elimination of ARDE lag can be further examined by defining a time ratio χ. Letting $\chi=t_2/(t_1+t_2)$, one can rewrite Eq. (3) as

$$R=\left(a_3+\frac{a_1\cdot a_3}{a_2}\right)\cdot\chi-\frac{a_1\cdot a_3}{a_2}. \quad (4)$$

In Eq. (4), a_1,a_2, and a_3 are all dependent on aspect ratios if process parameters, such as rf source power, rf bias power, gas flow rates, chamber pressure, and temperature are fixed. Thus, one can adjust χ so as to maintain the etch rate R at a constant value as plasma etching proceeds and aspect ratio increases. In other words, ARDE in a specific feature across time in a TDM plasma etch process can be eliminated.

The model for TDM plasma etch process thus provides tremendous insights with regard to controlling the ARDE lag in deep silicon etching processes. In constructing a TDM recipe, there are many practical alternatives among possible choices to adjust the process to reduce or eliminate ARDE lag, such as to (a) adjust the recipe etch time, (b) adjust the recipe deposition time, (c) adjust the polymer removal rate, (d) adjust the polymer deposition rate, and (e) adjust the spontaneous Si etching rate. The critical element is to offset the different efficiencies in polymer deposition and spontaneous silicon etching by manipulating step times or process rates. Other technical means, such as real-time monitoring of deposition thickness and etch depth[20] and continuous adjustment of process parameters,[21] can also be employed to assist effective control over ARDE lag.

III. EXPERIMENTAL RESULTS

Experiments have been conducted to validate the above model on the Unaxis DSE™ III platforms. The plasma etch-

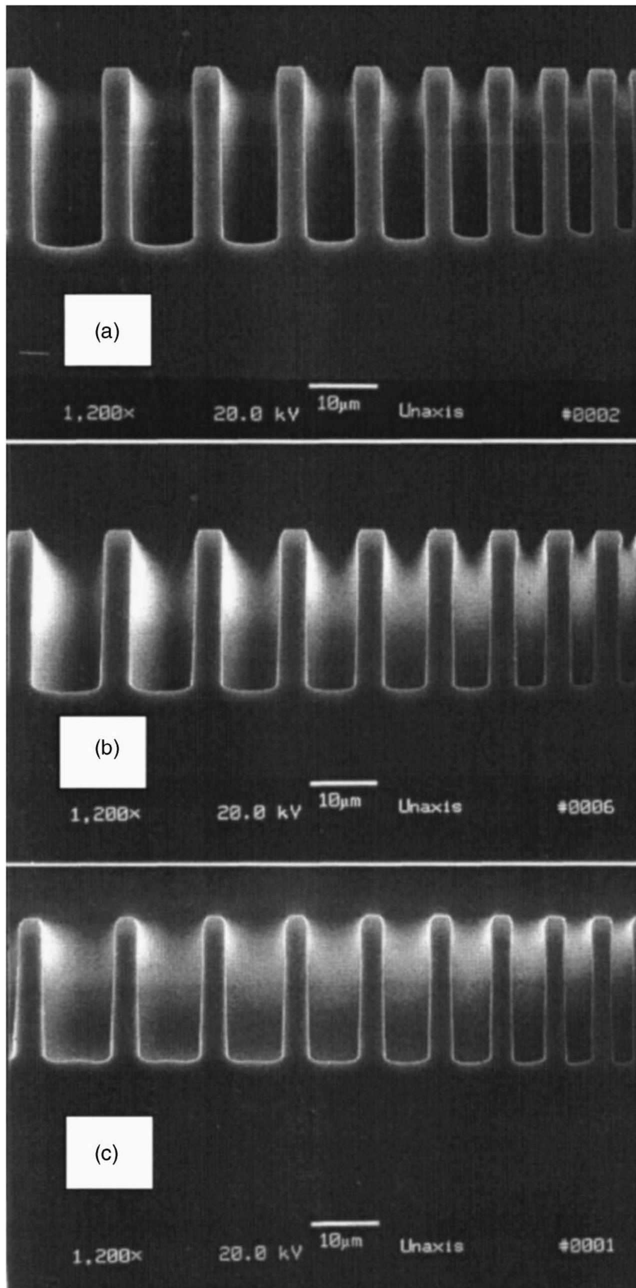

FIG. 6. Controlling the ARDE lag. For trenches with width ranging from 2.5 to 10 μm under different process conditions: (a) normal ARDE lag, ~10%, (b) elimination of ARDE lag (<2%), and (c) inverse ARDE lag (−5%).

ing of silicon is carried out by generating high-density plasma using an ICP source operated at 2 MHz. A rf bias is applied independently to the silicon substrate to control the ion energy during the etch process. The etched substrate is either electrostatically or mechanically clamped to an electrode with helium backside cooling. SF_6 and C_4F_8 are the gases used in the alternating etching/deposition cycles in the deep silicon etching (DSE) processes. By controlling plasma etching parameters, elimination and even the reversing of ARDE lag have been accomplished, as demonstrated in Fig. 6. In Fig. 6(a), the normal ARDE lag is observed but reduced to below 10% in trenches with trench widths ranging from 2.5 to 10 μm. In Fig. 6(b), the ARDE lag is eliminated. Examination of trenches etched under the same optimized conditions shows that the ARDE lag is controlled within 2%–3% across a wide range of trench width, i.e., 2.5–100 μm. In Fig. 6(c), the inverse ARDE lag is observed. The 2.5-μm-wide trench is etched 5% deeper than the 10-μm-wide trench. In all three instances shown, the TDM plasma etch rates exceed 2 μm/min.

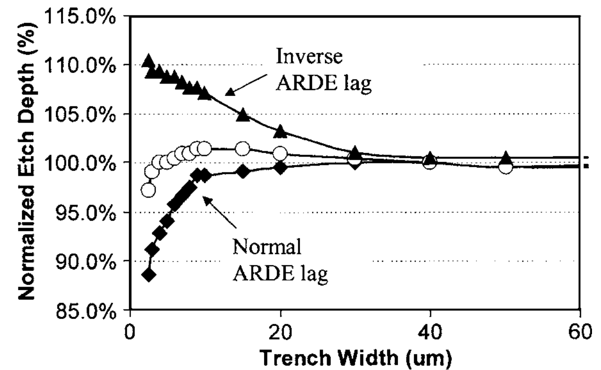

FIG. 7. Plot showing ARDE lags as a function of silicon trench width.

Figure 7 shows a plot of ARDE lag as a function of trench width. In the three experiments, normal, minimal, and inverse ARDE lag are measured. Over the range of trench width from 5 to 100 μm, ARDE lag are all controlled well below ~10%. When compared with the ~50% ARDE lag shown in Fig. 1, this is indeed a significant technological improvement.

IV. CONCLUSIONS

Today, the most widely employed TDM deep silicon etching processes use F-based chemistries for deposition and etching steps. Our three-stage model for such TDM plasma etch processeses focuses on a complete deposition/etching cycle and divides it into three consecutive stages, i.e., polymer deposition, polymer removal, and spontaneous silicon etching. These three stages are characterized by the respective polymer deposition rate, polymer removal rate, and silicon etching rate, all of which can be directly measured from experiments and controlled by processing parameters. While the ARDE phenomenon in TDM deep silicon etching processes is very complex in nature, significant insights into reducing and eliminating ARDE lag can be obtained from this simple model. The experimental results have been available to validate the physical model and demonstrate the effective control over ARDE lag in TDM plasma etch processes. Under optimal conditions, ARDE lag below 2%–3% has been readily achieved.

[1]K. Suzuki, K. Ninomiya, S. Nishimatsu, S. Okudaira, and O. Okada, U.S. Patent No. 4,579,623.
[2]F. Laermer and A. Schilp, U.S. Patent No. 5,498,312; U.S. Patent No. 5,501,893.
[3]R. A. Gottscho, C. W. Jurgensen, and D. J. Vitkavage, J. Vac. Sci. Technol. B **10**, 2133 (1992).
[4]C. W. Jurgensen, A. E. Novembre, and E. S. G. Shaqfeh, Proc. SPIE **94**, 1262 (1990).
[5]M. A. Blauw, T. Zijlstra, R. A. Bakker, and E. van der Drift, J. Vac. Sci. Technol. B **18**, 3453 (2000).
[6]S. G. Ingram, J. Appl. Phys. **68**, 500 (1990).
[7]J. C. Arnold and H. H. Sawin, J. Appl. Phys. **70**, 5314 (1991).
[8]K. P. Giapis, G. R. Scheller, R. A. Gottscho, W. S. Hobson, and Y. H. Lee, Appl. Phys. Lett. **57**, 983 (1990).
[9]S. Dushman and J. M. Lafferty, *Scientific Foundations of Vacuum Technology* (Wiley, New York, 1962), p. 94.
[10]W. Coburn and H. F. Winter, Appl. Phys. Lett. **55**, 2730 (1989).
[11]H. H. Doh, C. K. Yeon, and K. W. Whang, J. Vac. Sci. Technol. A **15**, 664 (1997).
[12]Th. Lill, M. Gribergen, and D. Mui, J. Vac. Sci. Technol. B **19**, 2123 (2001).
[13]K. Tsujimoto, T. Kumihashi, N. Kofuji, and S. Tachi, J. Vac. Sci. Technol. A **12**, 1209 (1994).
[14]T. Nozawa, T. Kinoshita, T. Nishizuka, A. Narai, T. Inoue, and A. Nakaue, Jpn. J. Appl. Phys., Part 1 **34**, 2107 (1995).
[15]C. K. Chung and H. C. Lu, U.S. Patent No. 6,900,136.
[16]A. Rickard and M. McNie, Proc. SPIE **4407**, 78 (2001).
[17]A. A. Ayón, R. A. Braff, R. Bayt, H. H. Sawin, and M. A. Schmidt, J. Electrochem. Soc. **146**, 339 (1999).
[18]S. L. Lai, D. Johnson, R. J. Westerman, J. Nolan, D. Purser, and M. Devre, Proc. SPIE **4979**, 43 (2003).
[19]S. L. Lai, R. J. Westerman, D. Johnson, and J. Nolan, Proc. SPIE **5342**, 94 (2004).
[20]D. Johnson, Invited Talk, *SPIE Conference on Micromachining and Microfabrication Process Technology VIII*, San Jose, CA, January 2004 (unpublished).
[21]M. Teixeira, M. Devre, W. Dawson, and D. Johnson, U.S. Patent No. 6,417,013.