Monte Carlo simulation of surface kinetics during plasma enhanced chemical vapor deposition of SiO_2 using oxygen/tetraethoxysilane chemistry

Phillip J. Stout and Mark J. Kushner
University of Illinois, Department of Electrical and Computer Engineering, Urbana, Illinois 61801

(Received 6 October 1992; accepted 24 April 1993)

Tetraethoxysilane/oxygen (TEOS/O_2) chemistries are now used for low plasma enhanced chemical vapor deposition (PECVD) of SiO_2 when high conformality and low temperature are required. The surface processes leading to film growth, carbon elimination, and conformality are not well characterized. We have developed a model to investigate the surface kinetics of the PECVD of SiO_2 films using TEOS/O_2 chemistry. The model includes precursor adsorption and desorption, densification of the film, void formation, pyrolysis, plasma surface interactions, and geometric shadowing. Results for growth rate, fraction of carbon retained in the film, film conformality, and film roughness are discussed.

I. INTRODUCTION

The desire to deposit SiO_2 films for microelectronics fabrication having good conformality at low temperatures ($\approx$400 °C) has motivated a renewed interest in tetraethoxysilane [TEOS, $Si(OC_2H_5)_4$] as a source gas in plasma enhanced chemical vapor deposition (PECVD).[1,2] Depositions at a low temperature are important due to stringent thermal budgets dictated by use of shallow junctions, which rapidly diffuse into the substrate at higher temperatures. When fabricating interlayered dielectrics,[2] it is often necessary to fill submicron trenches without voids and in some cases form planar surfaces.[3] This requires highly conformal depositions. Insulation between metalized layers,[4] must retain high dielectric strength properties and conformality at low temperatures. Oxide films grown with PECVD TEOS chemistries have shown good conformality with moderate insulating properties when deposited at low temperatures,[5] and meet many of these criteria.

An alternate to PECVD TEOS for SiO_2 films is the plasma oxidation of silane (SiH_4). This process, however, sometimes suffers from poor insulating quality and poor conformality[6] due to the low surface mobility of adsorbed deposition precursors.[7] It is believed that PECVD TEOS provides better film properties than those obtained with silane chemistry due to a higher surface mobility of the deposition precursor.[8] The fact that TEOS is easier to handle and is safer than silane provides additional incentive for its use.

X-ray photoelectron spectroscopy (XPS) studies[9] have shown that the carbon content in PECVD oxide using TEOS/O_2 chemistries is low. This result might imply that the plasma oxidation of TEOS removes the majority of the carbon to form reactive precursors prior to incorporation into the bulk film. Other work[10] suggests that TEOS fragments, partially decomposed by the plasma or created by radicals impinging on adsorbed TEOS on the surface, are the deposition precursors. These adsorbed TEOS species are eventually incorporated into the oxide which then requires that carbon elimination occurs on the surface or in the near surface layers of the film. Results of infrared transmission spectroscopy suggest these precursors may include $(SiO)_2Si(OC_2H_5)_2$ and $(SiO)Si(OC_2H_5)_3$.[11] The surface reactions which create these species appear to be somewhat dependent on the surface OH concentration.

It has been observed that the deposition rate of PECVD oxide using TEOS/O_2 chemistry has a negative activation energy. This behavior is believed to result from a balance between adsorption of precursor fragments of TEOS, which in general has a negative activation energy, and incorporation of the precursor into the film which has a positive activation energy.[5] In the absence of re-emission processes, the high conformality of these films compared to other PECVD systems implies that once TEOS fragments are adsorbed they have a high degree of surface mobility. This supposition is consistent with isotopic labeling and step coverage studies of PECVD TEOS oxide deposition which support a deposition mechanism dominated by mobile surface species.[12]

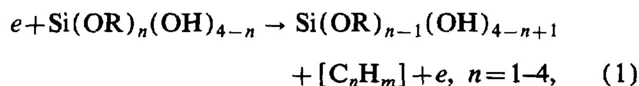
Models[13,14] for PECVD TEOS have recently been published. Raupp *et al.* have investigated the gas phase kinetics of O_2/TEOS plasmas.[15] Their results, which will be discussed below, suggest that production of atomic oxygen is a rate limiting process to the SiO_2 deposition. A surface profile model by Islamraja *et al.* for PECVD is based on a global surface reaction probability coupled with re-emission.[16,17] Their results for SiO_2 deposition using TEOS imply that re-emission is the dominant effect responsible for high conformality, while surface diffusion is not important. Monte Carlo models have been published to investigate step coverage during chemical vapor deposition (CVD) but are not specific to TEOS.[18,19] Since the details of surface kinetics have not been addressed by these works, we have developed a computer model for PECVD of SiO_2 using TEOS/O_2 chemistry with emphasis on surface processes. The purpose of this work is to phenomenogically investigate the contributions and requirements of surface kinetics to highly conformal films having low carbon content. The model will be described in Sec. II followed by a discussion of our results in Sec. III. Our concluding remarks are in Sec. IV.

 ©1993 American Vacuum Society

II. DESCRIPTION OF THE MODEL

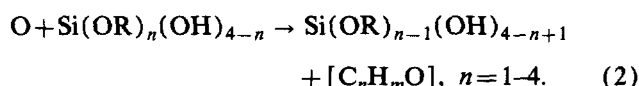
The model is a time-dependent Monte Carlo simulation of PECVD of SiO_2 in O_2/TEOS mixtures. The flow of the model will be briefly described. Deposition precursors, represented by Monte Carlo particles, are directed towards the growing film. These precursors adsorb on the surface, and diffuse from surface site to surface site until being incorporated into the film or desorbing from the surface. The film is represented by a discrete lattice to which the precursors bond, thereby allowing statistics on both morphology and stoichiometry to be gathered. When a precursor bonds to the lattice, becoming part of the film, it remains a chemically active species, reacting with surface resident species until it is buried beneath the surface. There are four surface processes and three bulk processes in the model. The surface processes are adsorption of deposition precursors, surface diffusion, desorption, and reactions with O atoms. The bulk processes are incorporation of surface resident precursors or densification of the film, void formation, and pyrolysis. All of the rate coefficients for these processes are temperature dependent. These dependencies and rate coefficients were derived by parameterizing the model using different reactions as a function of substrate temperature, and comparing our results to experimental values. This procedure does not provide unique values for rate coefficients but does generate an internally self-consistent reaction scheme which is useful for phenomenologically investigating TEOS kinetics. It should be noted that all of the temperature dependencies we discuss are only valid in the PECVD temperature regime (approximately 350–425 °C).

Precursors to the deposition are generated by plasma decomposition of TEOS, which allows the substrate to be held below pyrolytic temperatures. One set of precursors are believed to be $Si(OR)_n(OH)_{4-n}$ and $Si(OR)_nO_{4-n}$ where R represents C_2H_5. We will restrict our discussion to the former species. Two classes of reactions are hypothesized to occur in the plasma to generate these precursors in TEOS/O_2 mixtures. The first is direct electron impact dissociation, schematically represented by



$$e+Si(OR)_n(OH)_{4-n} \rightarrow Si(OR)_{n-1}(OH)_{4-n+1} + [C_nH_m]+e, \quad n=1\text{–}4, \qquad (1)$$

where the fragmentation products are $[C_nH_m]$. The second is oxidation of the R group



$$O+Si(OR)_n(OH)_{4-n} \rightarrow Si(OR)_{n-1}(OH)_{4-n+1} + [C_nH_mO], \quad n=1\text{–}4. \qquad (2)$$

In both reactions an OR group is removed from the silicon complex and is replaced by an OH group. In highly dilute TEOS/O_2 mixtures, we have assumed O oxidation reactions dominate over electron impact as the precursor source.[12] This supposition can be rationalized, in part, by the fact that most of the power deposited in the plasma goes into the O_2 to produce O atoms and excited O_2 species, as opposed to directly fragmenting the TEOS. Chemical reaction and excitation transfer of these species to the TEOS then generate deposition precursors. Deposition using O_2/TEOS mixtures can also be performed with remote plasmas where O_2, but not TEOS, flows through the discharge. These conditions produce films with similar properties to direct PECVD, thereby suggesting similar decomposition pathways. As described below, O reactions with TEOS fragments on the surface may also be an important method to eliminate carbon. Given the operational temperatures of a PECVD reactor, the rate of precursor generation is largely independent of the substrate temperature and depends largely on these gas phase processes.

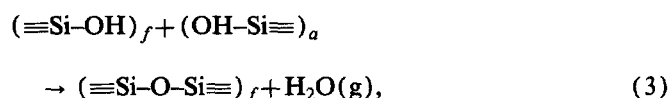
Although there are many other reactions which no doubt occur in the plasma, and which generate other species, the primary concern of this study is the surface kinetics of the deposition process. Therefore, we simply specify a flux of precursors to be incident on the surface which is consistent with the cited gas phase reaction scheme. In this respect, our reaction scheme allows fluxes of five gas phase species to be incident on the substrate; $Si(OH)_4$, $Si(OR)(OH)_3$, $Si(OR)_2(OH)_2$, $Si(OR)_3(OH)$, and O. We do not include the contribution of ions to the deposition, while noting that there is experimental evidence for ion-activated TEOS deposition.[15] $Si(OR)_4$ is not considered as a dynamic species in the model since our reaction mechanism does not distinguish between $Si(OR)_4$ fragmented in the plasma, and adsorbed $Si(OR)_4$ processed by O atoms on the surface. We note that recent modeling by Raupp *et al.*[15] proposes that O atom reactions with TEOS, either in the gas phase or on the surface, are the rate limiting deposition steps. This places an equal weight on gas phase and surface generation of TEOS fragments; and is conceptually consistent with the model described here.

The basis of our model is to treat the substrate as a three-dimensional network of sites. Each site can be occupied by a silicon atom bonded to an O, OH, or OR. For simplicity, the network is cubic; however, it is stoichiometrically amorphous by virtue of its random chemical composition. No attempt was made to account for the particulars of bond angles. The sites which a silicon complex may occupy are held in stacks or columns. The topmost layers of the network are fully resolved with respect to the site's position in the network, the type of species occupying that site and the neighboring species it is bonded to. As the film grows, a given Si atom "falls" to lower sites in the stack as other silicon complexes incorporate into the film on top of it. This process continues until it is buried below a specified number of layers. At this time the Si complex is no longer individually tracked. When buried beneath the near layer, the individual properties of the site are added to the bulk statistics of the film. The number of sites in the near layers was selected by performing many simulations while varying the depth of the near layer; and selecting a sufficiently thick layer so that the predicted properties of the film did not vary. This technique was used in modeling *a*-Si:H deposition, and is described in more detail in Ref. 20. We keep track of the stoichiometry of the film by recording the bonding partners of Si atoms in the film. The method allows macroscopically large films effectively having 10^6–10^7 lattice points to be simulated. True molecular dynamics simulations are usu-

ally restricted to a smaller number of atoms (100s to 1000s) due to their greater computational requirements.[21]

Given the chemical composition of the flux of precursors from the plasma, silicon in the film can be bonded to one of three species. It can bond to an O atom to form the SiO_2 network (denoted by B), to OH (denoted by O), or to an OR group (denoted by R). Given that one of the bonds must be to the amorphous SiO_2 network, and that a site can be empty, there are eleven possible types of sites (i.e., ten of the form BBOR, BOOO, etc.; and a void which denotes the absence of silicon at that site). A fully coordinated stoichiometric film would consist of only BBBB sites. Films containing OH or carbon will have BBBO and BBBR sites.

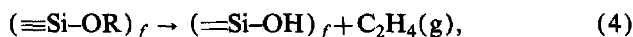
The primary reaction that propagates the SiO_2 network in our model is an interlinking step which integrates adsorbed TEOS fragments into the film and evolves water[12]



$$(\equiv\text{Si–OH})_f+(\text{OH–Si}\equiv)_a \rightarrow (\equiv\text{Si–O–Si}\equiv)_f+\text{H}_2\text{O(g)}, \quad (3)$$

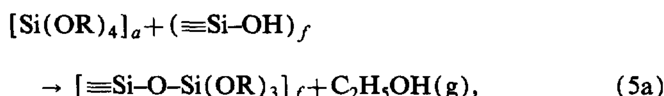
The subscripts f and a denote silicon incorporated into the film and adsorbed as a precursor, respectively. In this reaction the hydroxyl (OH) groups of adjacent silicon complexes interconnect to eliminate water, and two silicon atoms bond to a shared oxygen. In principle, the second (OH–Si≡) can be either an a or f species. With this mechanism, O atom "processing" of $Si(OR)_n$ is necessary to both remove carbon in the form of R, and prepare a ≡Si–OH bond. This processing can, in principle, occur in the gas phase or on the surface. The reverse reaction provides a mechanism for OH groups to integrate into the film through a reaction with adsorbed water.

The deposition precursors most likely contain substantial amounts of carbon since the gas phase oxidation of the TEOS is likely to be incomplete. Elimination of carbon from the film must therefore occur on the surface or in the near surface layers of the film. This may occur by oxidation or a pyrolytic reaction such as

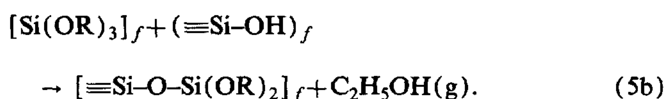


$$(\equiv\text{Si–OR})_f \rightarrow (\equiv\text{Si–OH})_f+\text{C}_2\text{H}_4\text{(g)}, \quad (4)$$

where an OR group of a Si complex eliminates C_2H_4, leaving behind a hydroxyl group. Since TEOS completely pyrolyzes at about 700 °C, this reaction likely plays a small role in the PECVD of oxide. It has been observed by Crowell *et al.*, however, that TEOS adsorbed on Si(100) will desorb C_2H_4 at temperatures as low as 230 °C, well within the operating temperatures of PECVD.[22] Crowell *et al.* also suggest that adsorbed TEOS or TEOS fragments can directly react with hydroxylated silicon surfaces, eliminating carbon in the process



$$[\text{Si(OR)}_4]_a+(\equiv\text{Si–OH})_f \rightarrow [\equiv\text{Si–O–Si(OR)}_3]_f+\text{C}_2\text{H}_5\text{OH(g)}, \quad (5a)$$



$$[\text{Si(OR)}_3]_f+(\equiv\text{Si–OH})_f \rightarrow [\equiv\text{Si–O–Si(OR)}_2]_f+\text{C}_2\text{H}_5\text{OH(g)}. \quad (5b)$$

The execution of the model will be briefly described followed by a more detailed description of the components of the model. Monte Carlo particles representing adsorbing deposition precursors are directed towards randomly selected locations on the substrate (typically 100×100 sites) at a frequency

$$f_a=\Phi A, \quad (6)$$

where Φ is the flux of adsorbing precursors and A is the area of the substrate. The site of the adsorption may be locally geometrically constrained, as described below. Once adsorbed, the precursor is randomly moved from site to site, simulating surface diffusion. As the deposition precursor moves over (or into) chemically active sites on the surface, it may interlink [Eq. (3)] thereby becoming a part of the film. The precursor is removed from the surface (that is, desorbed) if it has not interlinked after a specified time. While the precursor is on the surface, it may be chemically "attacked" by oxygen atoms from the plasma. This is accomplished by randomly directing O atoms onto the surface, and altering the chemical identity of the surface species it strikes.

Each of the processes in the model proceeds at a specified frequency. The mean time between occurrence of that event is, for example, $\Delta t_a=1/f_a$. The time step which is used to advance the global times for time step n is

$$\Delta t_n=\min(\tau_i+\Delta t_i)-t_{n-1}, \quad (7)$$

where t_{n-1} is the time of the last update and τ_i is the time of the last occurrence of process i. A particular process k is executed when $t_n=\tau_k+\Delta t_k$. In practice, the jump frequency of adsorbed precursors between surface sites is the largest of all frequencies, and therefore has the shortest mean time between events. This time therefore defines the increment by which the global clock is advanced. The meaningful parameter for process i is therefore f_i/f_j where f_j is the jump frequency between sites.

A. Precursor movement

Precursors on the surface are moved with a jump frequency of

$$f_j=1.7\times10^6\exp\left(-\frac{\Delta\epsilon_j}{RT}\right)\ \text{s}^{-1}, \quad (8)$$

where the activation energy is $\Delta\epsilon_j=1.9$ kJ/mol. This value compares favorably to activation energies for diffusion of adsorbates on metal which are typically a few kJ/mol.[23,24] The mean time between jumps at 450 K is $\approx 1\ \mu$s. (As mentioned above, the more meaningful parameter is f_i/f_j.) The temperature dependence of the jump frequency represents the increased probability that a precursor escapes the potential well in which it is adsorbed at higher substrate temperatures. This value was chosen to reproduce the observed temperature dependence for deposition.

Surface residence species are moved every Δt_j in one of eight directions (nearest and next nearest neighbors) to sites of equal or lower heights as determined by the choice of a random number $r=(0,1)$. Periodic boundary condi-

tions are used when simulating a flat surface. Reflective boundary conditions are used when simulating a macroscopic trench (see below). As the precursor moves on the surface, it may encounter sites of varying heights. If the site it moves over is the same height as the one it left, it may attempt to bond to the site it has moved over. If the adjoining site is higher, it will attempt to bond to the side of the "stack" it encounters. If the precursor moves into a hole, we assume that it will ultimately bond since there is no mechanism in our model to remove precursors from deep holes. If the precursor does not fall into a hole or does not bond, it is available to be moved again on its next jump.

The average residence time of the precursor on the surface before desorbing is a specified parameter. Long residence times allow a given precursor to sample a larger surface area, thereby reducing the effects of shadowing, either macroscopic (trenches) or microscopic (surface roughness). Longer residence times also increase the probability of incorporation into the lattice by allowing the precursor to visit more sites. The fact that the residence time of precursors may decrease with increasing temperature is a possible explanation for the observed negative activation energy for growth of SiO_2. We have chosen a desorption frequency of

$$f_d=3.3\times 10^5 \exp\left(\frac{-\Delta\epsilon_d}{kT}\right)\ \mathrm{s}^{-1}, \tag{9}$$

where $\Delta\epsilon_d=21.8$ kJ/mol with mean time between desorption of $\Delta t_d=1/f_d$. As discussed below, these values were chosen to yield a negative activation energy for the growth. This yields $f_d/f_j=10^{-3}$ at 450 K.

B. Interlinking or densification

The interlinking reaction which forms the SiO_2 network [Eq. (3)] is included in the model by checking adjacent surface sites at a specified frequency to see if they both have ≡Si–OH bondings. If so, then interlinking occurs with probability P_r and the site types are changed to reflect the reaction. [A random number $r=(0,1)$ is chosen. If $r\leqslant P_r$, then interlinking occurs.] The interlinking frequency is important with respect to the amount of OH in the film. If the rate of interlinking is not sufficiently high, a site could be buried before it eliminates hydrogen through bonding with an adjacent site having an OH group. A subtlety in implementing interlinking occurs if all of the local surface sites interlink with nearest neighbors. If so, then all of the OH surface sites are depleted. The result is that there are no ≡Si–OH configurations for incoming precursors to react with on the surface and integrate into the film. This effectively halts the growth of the film. As a result, when interlinking is implemented, the surface sites are always left with an OH (with the exception of BRRR). This action insures bonds are available for film growth.

The surface reaction probability P_r partly defines the effective mobility of the precursor on the surface. The ratio of $\Delta t_d/\Delta t_j$ determines the maximum possible number of sites a precursor will visit on the surface. If P_r is small, then many sites are sampled before incorporation. If P_r is large, then incorporation occurs near the landing site. Unless fixed for parametric purposes, the reaction probability was assumed to be

$$P_r=5\times 10^{-4} \exp\left(\frac{-\Delta\epsilon}{RT}\right), \tag{10}$$

where the activation energy $\Delta\epsilon=2.06$ kJ/mol, which yields $P_r=3\times 10^{-4}$ at 450 K. Due to its low activation energy, the reaction probability does not dramatically vary in the temperature range of interest to PECVD. This choice of activation energy was based on our observations that the temperature dependence of P_r must be weak in order for film growth to have a negative activation energy. This can be shown by defining an effective probability that a precursor incorporates into the film, P_i,

$$P_i=1-(1-P_r)^{\Delta t_d/\Delta t_j}. \tag{11}$$

The larger the P_i, the higher the rate of growth. As P_r increases, P_i also increases. As the temperature increases, however, the ratio $\Delta t_d/\Delta t_j$ decreases. That is, the time a precursor spends on the surface decreases with increasing temperature more quickly than the interval between moves, causing the exponent to decrease as temperature increases. Since $(1-P_r)<1$ and $\Delta t_d/\Delta t_j$ decreases as T increases, the incorporation probability goes down only if P_r is a weak function of temperature compared to the exponent.

C. Void formation

A gap in the amorphous SiO_2 network (i.e., no Si complex) is referred to as a void. Voids are formed in the same manner as interlinking. They result from interlinking between next nearest neighbors provided there is no occupied site in between. When a bond forms between next nearest neighbors, all open sites below are considered a void. Since there are no processes in the model which can change the status of a void, once a void is formed it becomes a permanent feature of the film. The mean time between attempting to bond with next nearest neighbors was chosen to be 50 times longer than bonding with nearest neighbors. This results in films having, at most, a few percent voids.

D. Oxygen surface chemistry

O atoms arrive on the surface at a rate determined by their flux from the plasma. The surface site an O atom strikes is picked at random, subject to the shadowing considerations discussed below. If the site has an (OR) group, an O insertion reaction is assumed to take place as would occur in the gas phase and the (OR) group is replaced by an OH group, thereby eliminating carbon from the film.

E. Pyrolysis

Pyrolysis is implemented by randomly choosing a site in the near layer, and, if the site has any (OR) groups, changing an (OR) group to an OH in accordance with reaction (4). As a method to remove carbon, this process differs from reactions with O atoms in that the pyrolysis reaction

can occur in the near layers of the film as well as on the surface. Also the rate of pyrolysis is temperature dependent whereas reactions with O atoms have no direct dependence on substrate temperature. We assumed that the frequency of the pyrolysis reaction is

$$f_p = 2.0\times10^{12}\exp\left(\frac{-\Delta\epsilon}{RT}\right)\ \mathrm{s}^{-1}, \tag{12}$$

where the activation energy $\Delta\epsilon = 184$ kJ/mol. As mentioned earlier, this process plays little part in the PECVD, though analogous processes, as discussed by Crowell *et al.*,[22] may be important at lower temperatures.

F. Geometric shadowing

When a precursor is incident on the surface, an accounting must be made for both microscopic and macroscopic shadowing. Since the precursors do not strike the substrate at normal incidence, a site on the surface could be shadowed by a stack next to it. This is accounted for in the model in the following fashion. First, an initial landing site for the precursor is randomly chosen. The differences in the heights of the eight stacks surrounding the selected landing site are then calculated. Any negative height differences are ignored (i.e., lower sites than the selected one) and the remaining values are used to obtain the average shadowing height. An acceptance angle is calculated based on this shadowing height. The acceptance angle is then used to obtain the probability that an incident radical lands on the initially selected stack and is not shadowed by one of the surrounding stacks. The choice of a random number determines the outcome of the event. If the decision is that the site is shadowed, the precursor is placed on a randomly chosen adjacent site which is higher.

Global shadowing effects, as caused by trenches, are accounted for in the following fashion. We assume that precursors arrive on the substrate isotropically. An acceptance angle for each site along the trench is computed and from these, a probability distribution is determined. The choice of a random number then determines the location along the surface that a precursor arrives.

III. PREDICTED FILM PROPERTIES

In this section, we will discuss results from our model for PECVD of SiO_2 in the context of growth rate, carbon content, and conformality. The results from our model, and the ensuing remarks should be considered a phenomenological discussion in which the importance of the cited processes are compared with respect to film properties. The results should be considered representative of low temperature PECVD where the deposition precursors have high surface mobility.

The base conditions for our study are a precursor mixture $Si(OH)_4/Si(OR)(OH)_3 = 0.9/0.1$, $\Phi = 10^{16}$ $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, $\Delta t_j = 10^{-6}$ s, $\Delta t_d = 10^{-3}$ s, and $P_r = 3\times10^{-4}$. In our nomenclature, Φ is the flux of radicals which directly adsorb upon striking the substrate. The total precursor flux is Φ/P_a, where P_a is the direct absorption probability.

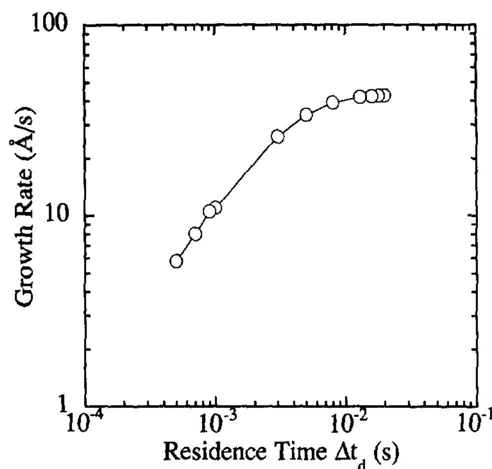


FIG. 1. Predicted growth rate of SiO_2 as a function of average residence time of adsorbed TEOS precursors on the substrate. The flux of adsorbing precursors is 10^{16} $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The growth rate saturates at large residence times when all adsorbing precursors are incorporated into the film.

A. Growth rate

The maximum growth rate is determined primarily by Φ, and the actual growth rate can be predicted by $P_i\Phi$. The growth rate can be increased for a given Φ by increasing the residence time of an adsorbed precursor (Δt_d) or increasing the reaction probability P_r. To demonstrate the dependence of growth rate on these parameters, we performed simulations as a function of Δt_d, P_r, and Φ. The growth rate increases nearly linearly with Δt_d as shown in Fig. 1, and is commensurate with experimental values.[12–16] Increasing the residence time for an adsorbed precursor increases the maximum number of interlinking attempts a precursor makes before being desorbed from the surface. This effectively increases the reactive sticking coefficient. The growth rate saturates at large Δt_d, at which time the integration probability P_i approaches unity and all adsorbed precursors integrate into the film.

The growth rate as a function reaction probability P_r is shown in Fig. 2. Increasing P_r for a given Δt_d increases the growth rate since the rate of incorporation of adsorbed precursors is larger. Shorter residence times are more sensitive to decreasing values of P_r since smaller values of P_r require that more sites be visited for a precursor to incorporate into the film. For a given set of parameters (Δt_d, P_r), the predicted growth rate is linear with increasing flux of adsorbing precursors as shown in Fig. 3. Growth rates are shown for two values of Δt_d, showing a larger rate for larger Δt_d. A weakness of the model is that radicals on the surface do not "see" each other. If reactions between adsorbed precursors were included, the deposition rate would predictably be less than linear with the adsorbing precursor flux.

The predicted growth rate as a function of temperature, shown in Fig. 4, has a negative activation energy of −0.11 eV. This value is commensurate with that observed by Pai and Chang[5] (−0.12 eV), and by Chin and van de Ven[2] (−0.19 eV). The predicted negative activation energy is a

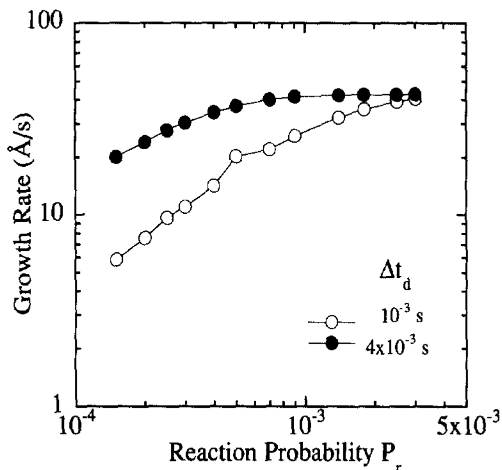


FIG. 2. Growth rate of SiO_2 as a function of the reaction probability for incorporation into the network for two average residence times Δt_d. The growth rate saturates with a smaller P_r at longer residence times since a larger number of sites are visited, thereby affording more opportunities to incorporate into the film.

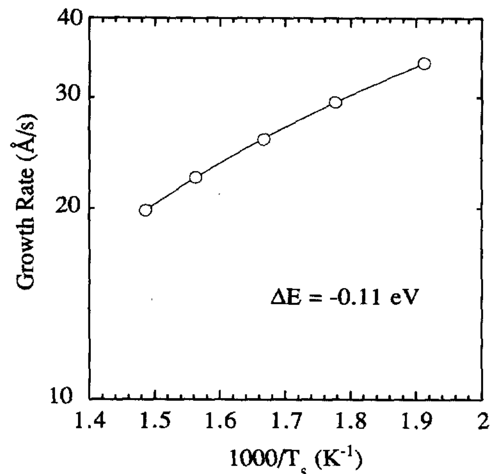


FIG. 4. Simulated growth of SiO_2 as a function of substrate temperature T_s. A mechanism in which the rate of desorption of TEOS precursors increases faster with temperature than the rate of incorporation produces a negative activation energy.

result of a decrease in residence time for adsorbed precursors with increasing temperature. This decreases growth rates as a result of there being fewer samplings of sites. This effect dominates over an increase in jump frequencies and reaction rates with increasing temperature processes which should increase growth rates. Other effects may also contribute to the apparent negative activation energy. Raupp *et al.*[15] have modeled SiO_2 growth from TEOS and propose that growth is rate limited by the availability of O atoms. Since the recombination of O radicals at the wall to form O_2 increases with temperature, fewer O radicals are available to react with TEOS adsorbed on the surface to form the film, and so the growth rate decreases.

B. Carbon content

The fraction of (OR) groups (C_2H_5O) in the film is a function of three parameters. The first is the type of precursor incident on the substrate [e.g., $Si(OR)_3(OH)$]. The second is the flux of oxygen which oxidizes (OR) groups on the surface. The third is pyrolysis which removes R groups from the near layers. The experimental observation is that the carbon content near the surface (topmost 10–20 Å) is 5%–10%.[12] At a depth below 30 Å, the carbon content is negligible. The R/Si fraction as a function of the O atom flux Φ_O incident on the substrate, is shown in Fig. 5. For low values of Φ_O, the R fraction is the same as that in the adsorbing precursor flux since, in the absence of oxidation, there is no mechanism to remove R from the adsorbed species. As Φ_O is increased to the same order of

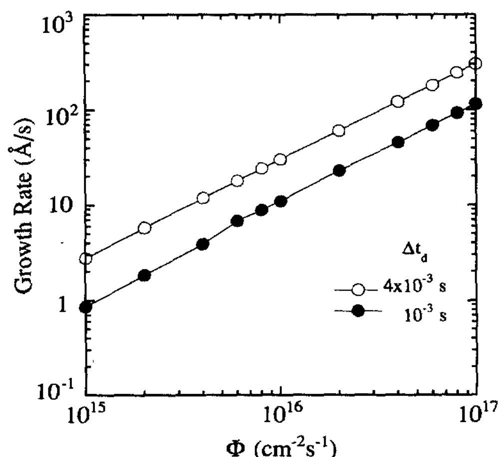


FIG. 3. Simulated growth rate of SiO_2 as a function of the flux of adsorbing TEOS precursors for two average residence times Δt_d. In our model, adsorbed TEOS fragments do not simultaneously compete for sites; and so the growth rate increases linearly with the adsorbed radical flux.

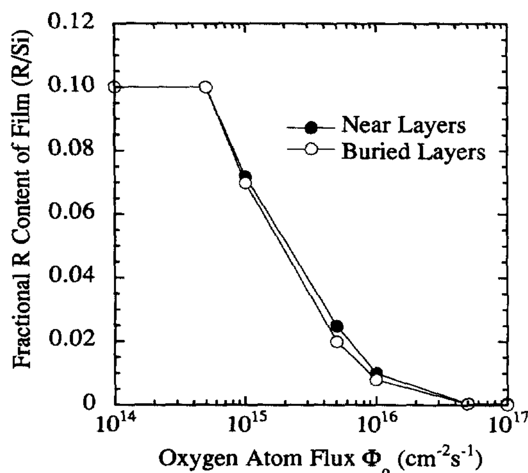


FIG. 5. Fraction of carbon groups ($R=C_2H_5O$) in the film as a function of oxygen atom flux incident on the film. In this parameterization, the only mechanism to remove carbon from the film once incident on the surface is by oxidation. Total removal therefore requires an oxygen atom flux commensurate with the precursor flux (10^{16} cm^{-2} s^{-1}).

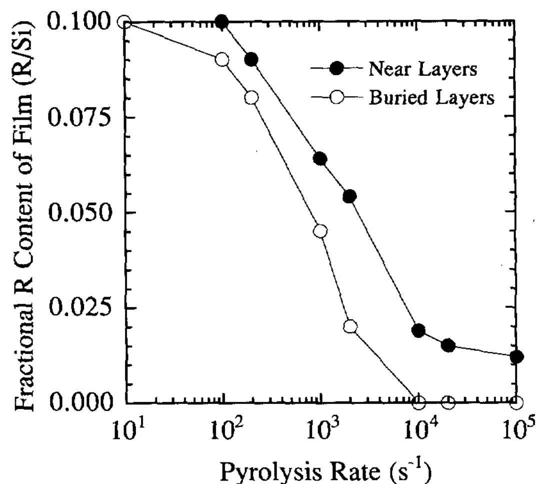


FIG. 6. Fraction of carbon groups ($R=C_2H_5O$) in the film as a function of the rate of pyrolysis or (other homogeneous process) which eliminates carbon. With this mechanism, the near layers retain more carbon than the buried layers as observed experimentally. The adsorbing precursor flux is 10^{16} cm^{-2} s^{-1}.

magnitude as the precursor flux (10^{16} cm^{-2} s^{-1}), the carbon content decreases. For $\Phi_O > 5 \times 10^{16}$ cm^{-2} s^{-1} carbon is essentially eliminated from the film. Note that the near layers and buried layers have nearly identical R content since the removal of carbon has no dependence on residence time and occurs only on the surface. These trends may support the conclusions of Raupp *et al.*[15] who propose that the O atom inventory ultimately controls the growth rate by producing deposition precursors on the surface saturated by TEOS. In our model we allow $Si(OR)(OH)_3$ to incorporate into the film, whereas in reality, the rate of incorporation of these species is most likely retarded compared to $Si(OH)_4$. Therefore, the oxidation of $\equiv$Si–(OR) on the surface by O atoms, which essentially activates that particular site for incorporation, may control film growth.

The R fraction as a function of the rate of pyrolysis is shown in Fig. 6. Carbon is nearly completely eliminated from the film at a frequency of 10^4 s^{-1}, corresponding to a substrate temperature of 760 °C. In these cases the near and buried layers have different R content. The carbon content of the buried layers is smaller than the near layers since the residence time of $\equiv$Si–(OR) in the film is longer, and therefore the opportunity to eliminate the carbon by pyrolysis is greater. Experimental measurements of carbon content shows higher values in the near layers than in the bulk.[12] These observations, and our results, suggest that at low temperatures the incorporation of $\equiv$Si–(OR) into the film is constrained [compared to $\equiv$Si–(OH)], until it is oxidized or activated by O atoms from the plasma. This results in a buildup of carbon in the near layers. This behavior mimics the removal of carbon from the near layers which might occur from a pyrolytic (or some other homogeneous) process once $\equiv$Si–(OR) is below the surface. Reactions of this type are suggested by the work of Crowell *et al.*[22]

C. Conformality

Conformality is a measure of how well a film mimics the microscopic topography of the substrate. In the case of trenches, good conformality is usually equated to having the film thickness on the side and bottom of the trench equal to that on the top surface. Under physical vapor deposition conditions, the growth rate at the bottom of the trench is higher than the walls due to geometric shadowing. If the desire is to completely fill a trench[25] to create a planar surface, the trench should fill with positive slopes from the bottom of the trench. In this way a large void will not be formed in the trench that might be uncovered in future planarization steps. (A trench with a positive slope has a film on the sides of the trenches which is thicker at the bottom of the trench than at the top.)

Recently, experimental and theoretical studies of the growth of SiO_2 from O_2/TEOS plasmas have been completed by Islamraja *et al.*[16] In their experimental work, microstructures were fabricated from which the deposition rates on surfaces exposed to the plasma, and surfaces shadowed from the plasma could be measured. Their model also used these structures; and predicted deposition rates based on a global sticking coefficient and re-emission of nonsticking radicals. They concluded that the highly conformal SiO_2 films which can be obtained in O_2/TEOS plasmas is dominantly a result of low reactive sticking coefficients and re-emission which allows deposition precursors to penetrate small structures and conform to step structures, and that surface diffusion played a small role in conformality.

In light of these results, we wanted to investigate the contribution of surface diffusion of adsorbed deposition precursors to conformality and roughness of the film in the context of a detailed surface kinetics model. If the residence time or incorporation time of precursors on the surface is relatively constant, precursors which sample a wider range of sites from its site of adsorption will produce both more conformal and microscopically smoother films. Therefore, one should expect that conditions which result in longer residence times or more rapid movement of the precursor on the surface should produce more conformal films. Regions of the film which do not receive large fluxes of precursors are therefore compensated by more mobile precursors arriving from distant sites. Trenches having large aspect ratios (height of trench/width) are particularly prone to having poor conformality since the direct precursor flux to the corners of the trench from the plasma is reduced by the solid angle open to the plasma.

We investigated conformality in small structures in the absence of re-emission (Figs. 7–11) using our model. Our test structure consisted of a periodic array of trenches whose dimensions are shown in Figs. 9 and 10. Aspect ratios are given by S/(2B), where half of the structure is typically 700–1500 sites long. Physically, these structures are only 0.2–0.5 μm in size. The predicted conformalities [expressed as ratios of the film thickness "side-to-top" (S/T) and "bottom-to-top" (B/T)] are shown in Fig. 7 as a function of aspect ratio. Results are shown for varying reaction probabilities and desorption times Δt_d. As ex-

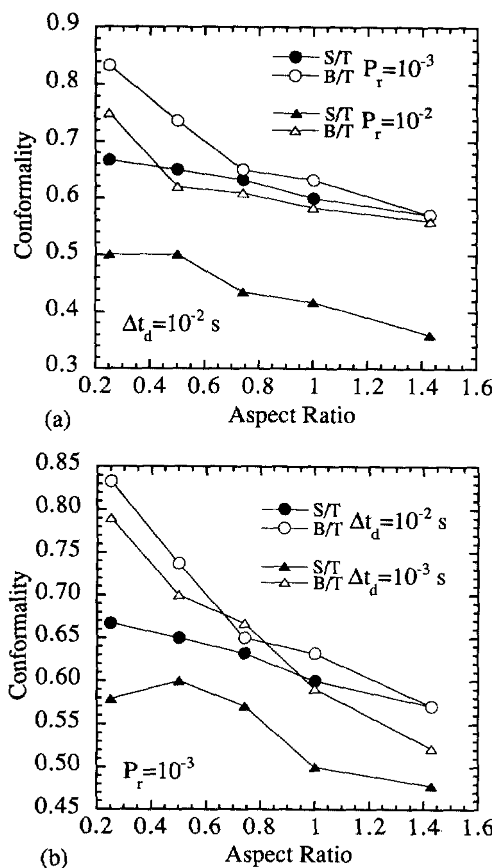


FIG. 7. Predicted conformality as a function of aspect ratio for (a) two values of the reaction probability P_r and (b) two values of the average residence time. The ratio of film thickness bottom-to-side (B/T) is generally greater than that for side-to-top (S/T), in agreement with experiments. Conformality decreases with increasing aspect ratio and increasing reaction probability, as precursors are incorporated into the film closer to their site of incidence.

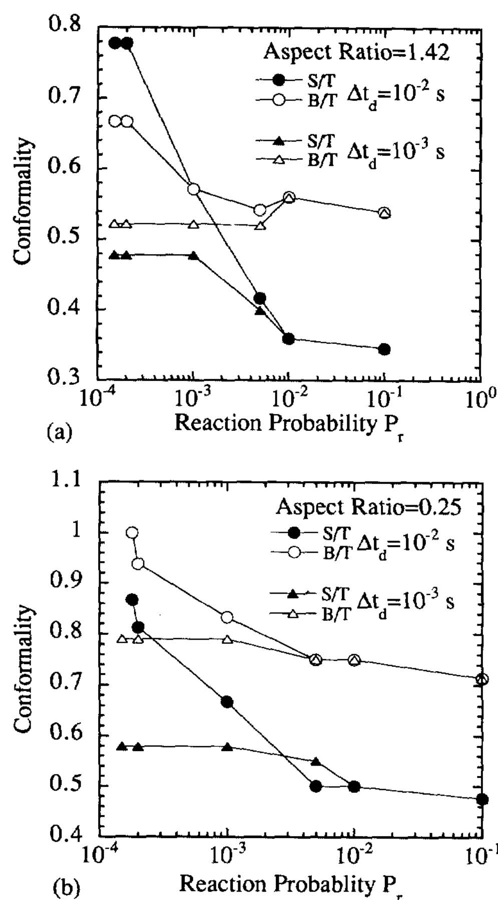


FIG. 8. Predicted conformality as a function of reaction probability for aspect ratios of (a) 1.42 and (b) 0.25. Results are shown for two values of the average residence time. The conformality decreases with increasing reaction probability, as precursors incorporate into the film near their landing site.

pected, conformality degrades with increasing aspect ratio. It also decreases with decreasing residence time and increasing reaction probability, as the area and number of sites sampled by a precursor decreases. Experimentally measured conformalities worsen with increasing aspect ratios, and typically have B/T > S/T (B/T $\approx$0.9 to 0.6 and S/T $\approx$0.6 to 0.2) and so our model generally agrees with these results.[12] The dependence on aspect ratio is less sensitive for conditions where the deposition is "directional" (B/T $\approx$1, S/T $\ll$1).[26] There are also examples (low frequency plasma[27] and downstream plasma reactor[5]) for which the sidewall thickness is greater than the bottom thickness.

A more direct demonstration of the effect of P_r on conformality is shown in Fig. 8 where B/T and S/T are shown for aspect ratios of 1.42 and 0.25. In both cases the conformality approaches unity as P_r is lowered. There is, however, also a limit (dependent on the residence time) to the degree to which small values of P_r can improve conformality. The shorter the residence time the higher the reaction probability at which conformality saturates at its high value. The ratio saturates when the reaction probability is low enough so that the mobility of the adsorbed precursors

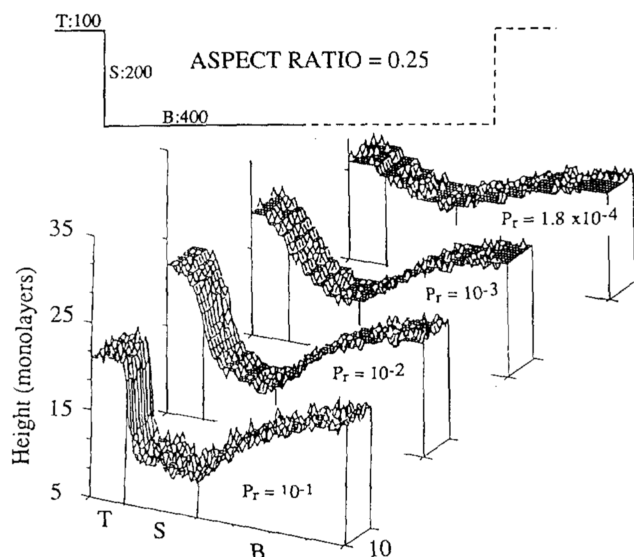


FIG. 9. Film thickness along a small trench having an aspect ratio of 0.25 as a function of reaction probability after adsorbing. The trench geometry is shown in the inset at the top of the figure. T, S, and B refer to top, side and bottom respectively. The numbers in the inset refer to the number of sites in the simulation. Note that as the conformality improves with decreasing reaction probability, the film smoothness also improves.

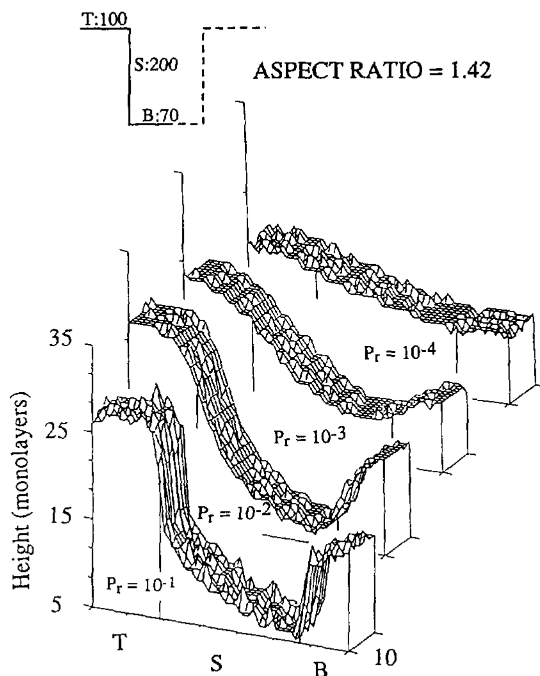


FIG. 10. Film thickness along a small trench having aspect ratio of 1.42 as a function of reaction probability after adsorbing. The conditions and notation are otherwise the same as in Fig. 9. Conformality and smooth films require a smaller reaction probability than for the smaller aspect ration trench.

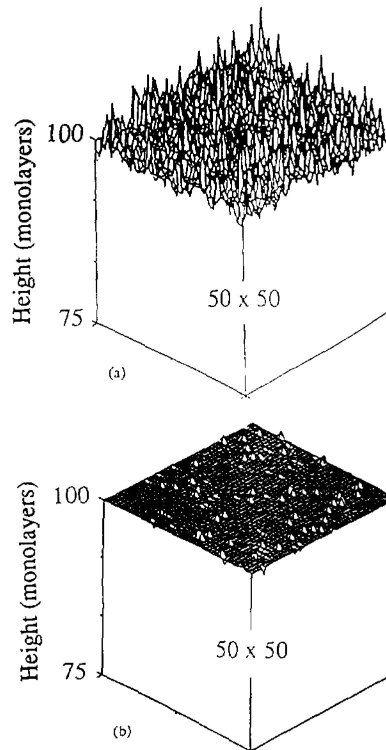


FIG. 11. Film smoothness as a function of reaction probability (a) 0.3 and (b) 0.001.

is limited by their residence time on the surface. In this scenario, all of the radicals have their maximum range of travel from their site of incidence on the surface and desorb before bonding.

The relative film thickness for trenches of aspect ratio 0.25 and 1.42 are shown in Figs. 9 and 10 for various values of P_r. The top, side, and bottom regions of the trench are labeled T, S, and B in the figures. In the model, these regions are distinguished from each other by the probability of a precursor arriving at a particular location. With a large reaction probability P_r, the film heights reflect the distribution of the arrival of precursors since the precursors incorporate near their site of incidence. The top and bottom of the trenches have thick films and the side of the trench, where shadowing occurs, has a thin film. As P_r is lowered this transition becomes more gradual and, in general, complete conformality is approached.

The microscopic smoothness of the film also improves with decreasing P_r, as shown in Fig. 11, where film morphologies are shown for $P_r=0.3$ and 0.001. The substrate was smooth before deposition and grown to an average height of 100 layers on a 50×50 site. Comparing these results to Figs. 9 and 11, we find that smooth films are obtained with a larger P_r than required to achieve conformality; a consequence of the shorter range involved.

IV. CONCLUDING REMARKS

A model for the surface kinetics during PECVD of SiO_2 using $TEOS/O_2$ chemistry has been developed, and used to investigate growth rates, carbon content, and conformality. The model which we have described is phenomenological, and many of the reaction rates have been estimated or derived. Its value stems from it being a self-consistent accounting of a fairly complex set of interdependent processes which can be used to examine those dependences. When the various estimated rate coefficients become experimentally or theoretically available, the infrastructure in this model exists to investigate their implications on film properties. In spite of the phenomenological nature of our model, certain trends have become evident, and tentative conclusions may be drawn.

Our results for carbon elimination from the film suggest that the observed experimental trends of decreasing carbon content with depth can be explained by one or a combination of two mechanisms. The first is a resident time dependent process, which mimics homogeneous pyrolysis, or an interlinking mechanism which discriminates against $\equiv$Si(OR) complexes. The former process requires low activation energy reactions as proposed by Crowell *et al.*[22] The latter mechanism is consistent with O atom reactions on the surface being responsible for the removal of carbon from the film. In that respect, O atoms would ultimately control the rate of incorporation, and film growth, thereby corroborating the conclusions of Raupp *et al.*[15] The observed negative activation energy for PECVD in O_2/TEOS plasmas can be reproduced by a mechanism in which an increasing rate of desorption of precursors with increasing substrate temperature dominates over any increase in rates of incorporation. Raupp *et al.*[15] suggest that the available O atoms on the surface also decrease with increasing temperature due to an increasing rate of recombination and

desorption, and the rate of growth decreases as a consequence. These two mechanisms, though different in their details, both rely on the depletion of precursors on the surface with increasing temperature.

The role of surface mobility in producing conformal films can explain experimental results provided that reaction probabilities are low enough and residence times long enough to extend the range of adsorbed precursors over macroscopic (many microns) distances. Low reaction probabilities, however, also result in low deposition rates. It is difficult to reconcile the experimentally observed deposition rates and conformal films based on only low reaction probabilities with long residence times (10s ms) prior to desorption. Re-emission, as proposed by Islamraja *et al.*,[16] therefore, appears to play an important role in obtaining conformality on large scales (many microns). Microscopically smooth films, however, are most likely explained by mobile surface precursors.

ACKNOWLEDGMENTS

This work was supported by IBM East Fishkill Facility, the National Science Foundation (ECS-91-09326) and the Semiconductor Research Corporation.

[1]P. H. Singer, Semicond. Int. **15**, 28 (1992).
[2]B. L. Chin and E. P. van de Ven, Solid State Technol. **31**, 119 (1988).
[3]A. C. Adams, Solid State Technol. **26**, 135 (1983).
[4]S. Rojas, A. Modelli, and W. S. Wu, J. Vac. Sci. Technol. B **8**, 1177 (1990).
[5]C. S. Pai and C.-P. Chang, J. Appl. Phys. **68**, 793 (1990).
[6]R. M. Levin and K. Evans-Lutterodt, J Vac. Sci. Technol. B **1**, 54 (1983).
[7]K. Law, J. Wong, C. Leung, J. Olsen, and D. Wang, Solid State Technol. **32**, 60 (1989).
[8]F. S. Becker, D. Pawlik, H. Azinger, and A. Spitzer, J. Vac. Sci. Technol. B **5**, 1555 (1987).
[9]S. Nguyen, D. Dobuzinsky, D. Harmon, R. Gleason, and S Fridman, J. Electrochem. Soc. **77**, 2209 (1990).
[10]C. P. Chang, C. S. Pai, and J. J. Hsieh, J. Appl. Phys. **67**, 2119 (1990).
[11]L. L. Tedder, Guangquan Lu, and J. E. Crowell, J. Appl. Phys. **69**, 7037 (1991).
[12]N. Selamoglu, J. A. Mucha, D. E. Ibbotson, and D. L. Flamm, J. Vac. Sci. Technol. B **7**, 1345 (1989).
[13]S. R. Kalidindi and S. B. Desu, J. Electrochem. Soc. **137**, 624 (1990).
[14]G. B. Raupp and T. S. Cale, Chem. Mater. **1**, 207 (1989).
[15]G. B. Raupp, T. S. Cale, and H. P. W. Hey, J. Vac. Sci. Technol. B **10**, 37 (1992).
[16]M. M. Islamraja, M. Capelli, J. McVittie, and K. Saraswat, J. Appl. Phys. **70**, 7137 (1991).
[17]L. Y. Cheng, J. P. McVittie, and K. C. Saraswat, *Proceedings of the Second International Conference on Ultra Large Scale Integration Scientific Technol.*, edited by C. M. Osburn and J. M. Andrews (The Electrochemical Society, Pennington, NJ, 1989), p. 506.
[18]M. Ikegawa and J. Kobayashi, J. Electrochem. Soc. **136**, 2982 (1989).
[19]M. J. Cooke and G. Harris, J. Vac. Sci. Technol. A **7**, 3217 (1989).
[20]M. J. McCaughey and M. J. Kushner, J. Appl. Phys. **65**, 186 (1989).
[21]See, for example, C. M. Gilmore and J. A. Sprague, Phys. Rev. B **44**, 8950 (1991).
[22]J. E. Crowell, L. L. Tedder, H.-C. Cho, F. M. Cascarano, and M. A. Logan, J. Vac. Sci. Technol. A **8**, 1863 (1990).
[23]R. Gomer, Rep. Prog. Phys. **53**, 917 (1990).
[24]R. B. Hall, T. H. Upton, and E. Herbolzheimer, J. Vac. Sci. Technol. B **5**, 1470 (1987).
[25]N. Selamoglu, C. N. Bredbenner, and T. A. Giniecki, J. Vac. Sci. Technol. B **9**, 2530 (1991).
[26]J. J. Hsieh, D. E. Ibbotson, J. A. Mucha, and D. L. Flamm, Mater. Res. Soc. Symp. Proc. **165**, 107 (1990).
[27]D. E. Ibbotson, J. J. Hsieh, D. L. Flamm, and J. A. Mucha, Proc. SPIE **1037**, 130 (1988).