PHYSICAL REVIEW B VOLUME 62, NUMBER 24 15 DECEMBER 2000-II

Reaction-diffusion model for thermal growth of silicon nitride films on Si

R. M. C. de Almeida and I. J. R. Baumvol
Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, 91501-970, Porto Alegre RS, Brazil
(Received 13 October 2000)

Thermal growth of ultrathin silicon nitride films on Si in NH_3 is modeled as a dynamic system governed by reaction-diffusion equations. Solution of the model yields profiles of the involved species consistent with experimental observations of a stoichiometric silicon nitride layer in the near-surface region and a subnitride layer of comparable thickness in the near-interface region. Self-limited growth kinetics are also obtained from the model equations in good agreement with experimental results, owing to a diffusion barrier layer to the nitridant species formed in the near-surface region by stoichiometric silicon nitride.

Silicon nitride is one of the major materials for thin and ultrathin dielectrics in microelectronics. They are commonly used as diffusion barriers to silicon oxidant species, as well as to impurities that can contaminate the dielectric/Si interface and the active semiconductor region in metal-oxide-semiconductor (MOS) structures used in present and future Si-based devices.[1] More recently, ultrathin (1 nm thick and less) thermal silicon nitride films grown on Si(001) (c-Si) in ammonia (NH_3) are being applied as the best barriers to silicon oxidation[2–5] in advanced, high-dielectric constant (high-K) oxides like Ta_2O_5, TiO_2, Al_2O_3, $ZrSiO_4$, and barium strontium titanate deposited on Si, which are alternatives to SiO_2. The thickness and consequently the efficacy of any intermediate diffusion barrier film is, however, severely limited by the degradation that it causes on the magnitude of the overall dielectric constant.[3] This last factor renders silicon nitride films thermally grown in NH_3 specially attractive since: (i) for any practical temperature, the growth kinetics is essentially self-controlled;[6–11] (ii) they are effective barriers even at thickness as low as one or two atomic layers;[12] (iii) the dielectric constant of silicon nitride (7.8) itself is more than twice that of SiO_2 (3.8). The aim of the present work is to model the kinetics of thermal growth of silicon nitride films on Si in NH_3 as a dynamic system governed by reaction-diffusion equations and comparing the model kinetics with those experimentally determined. Owing to its increasing technological importance, molecular dynamics simulations of ultrathin Si_3N_4 films deposited on Si and of resulting Si_3N_4/Si interfaces have been recently reported.[13,14] However, to the best of our knowledge, there is not in the literature a model capable of describing the observed thermal growth kinetics of silicon nitride films on Si in NH_3.

Kinetics of thermal growth of silicon nitride films on Si in NH_3 were determined,[6–11] showing very high growth rates at initial stages, followed by self-limited growth to a maximum thickness of 6 nm even for processing times as long as 18 h at 1200 °C. Previous work discussed thermal nitride growth within the framework of the Deal and Grove model for the thermal oxidation of silicon in O_2,[11] assuming that the nitriding species were nitrogenous radicals diffusing through the stoichiometric and uniform nitride film to react with Si at the nitride-Si interface. This model was criticized[8] since the extracted diffusion lengths (around 0.4 nm) imply annihilation of the nitridant species inside the nitride film, in contradiction with the assumption made of a stoichiometric and uniform growing silicon nitride film. A completely different approach[7,15] assumed that the nitridation of silicon proceeds mainly by migration of Si cations, under the action of an electric field out to the surface of the growing nitride film, being the reaction between NH_3 and Si occurring at the surface the rate-determining factor. However, since Si was shown to be immobile during thermal nitridation in NH_3 (Refs. 16–19) this approach must be disregarded.

Surface analysis and adsorption/desorption spectroscopy[16,20–24] showed that at initial stages thermal nitridation of Si proceeds by dissociative adsorption of NH_3, followed by reaction of dissociation products like NH_2 and H with the substrate. At the end of the initial fast growth regime, silicon nitride films with thickness around 1 nm are formed. Analysis of the near-interface region by core-level photoemission spectroscopy[15,16] revealed the presence of substoichiometric compounds (subnitrides Si_3N_x, $x \leqslant 4$), similarly to those existing in the near-interface regions of thermal silicon oxide films on Si. The thickness of the transition subnitride region can be estimated from the various core level spectroscopy studies as being approximately 0.5 to 1 nm. The surface composition of thermal silicon nitride films is stoichiometric Si_3N_4 (Si^{4+} state).[16] Depth profiling of Si and N revealed that the films are not homogeneous.[9,25] A stoichiometry corresponding to Si_3N_4 only exists in a narrow layer (1 nm or less) below the surface. Deeper into the films N concentration decreases and reactive Si concentration increases continuously until reaching the nitride/Si interface, confirming the subnitrides (silicon-excess) region.

Thus, experimental evidence indicates that (i) nitrogenous species are reactive and mobile during thermal growth of silicon nitride films on Si in NH_3. Si, on the other hand, is reactive and immobile, although very short range injection of Si from the substrate into the growing nitride film cannot be ruled out, even if it has not been observed directly and (ii) thermal nitridation of Si in NH_3 begins with the dissociation of the gas molecule into smaller fragments, like NH_2 and H, which react with silicon generating silicon nitride. Film growth proceeds by the migration of NH_2 towards the nitride/Si interface to react with Si, either as an interstitial defect reacting with the nitride network or as a substitutional defect driven by diffusion. Since the silicon nitride/Si transi-

tion region containing subnitrides corresponds usually to more than half of the film thickness, the mobile nitrogenous species react substantially within the subnitride region, promoting growth and releasing H. Only a minor fraction of the reaction takes place at the interface, nitriding bare Si. Reaction of ammonia fragments within the substoichiometric transition region leads to the formation of a stoichiometric Si_3N_4 layer, which is stable only in the near-surface region, since the subnitride region is always necessary to accommodate stress due to the transition from bulk *c*-Si to stoichiometric, amorphous Si_3N_4.[12]

Based on experimental evidence, thermal growth of silicon nitride films on Si is modeled as a reaction-diffusion phenomenon in one spatial dimension. Solution of the model equations yields concentration versus depth of the involved species and growth kinetics. Initially a sharp *c*-Si surface at $x=0$ is considered, which is exposed to a nitridant atmosphere (NH_3) at constant temperature and pressure. NH_3 is dissociatively adsorbed, reacting with Si and forming silicon nitride. Film growth proceeds by diffusion and reaction of the nitridant species, mainly NH_2, through silicon nitride and Si. The observed self-limited growth kinetics may be explained by diffusion barriers. One assumes here that the nitridant species diffuses through substoichiometric (subnitride) layers as well as through Si, having zero diffusivity in stoichiometric Si_3N_4. Since Si_3N_4 does not react with NH_3, after establishment of a stoichiometric silicon nitride layer in near-surface regions nitridation may still occur by reaction within the subjacent subnitride layer, but film growth will eventually stop when the previously adsorbed nitridant species have been consumed.

Model equations are built-up to describe the time evolution of relative concentrations of silicon and nitridant species, defined as

$$\rho_{Si}(x,t)=\frac{C_{\mathrm{Si}}(x,t)}{C_{\mathrm{Si}}^{\mathrm{bulk}}},$$

$$\rho_{N}(x,t)=\frac{C_{\mathrm{N}}(x,t)}{C_{\mathrm{Si}}^{\mathrm{bulk}}}, \tag{1}$$

where $C_{\mathrm{Si}}(x,t)$ and $C_{\mathrm{N}}(x,t)$ are, respectively, the concentration of reactive Si atoms and of nitridant species molecules at depth x and time t. $C_{\mathrm{Si}}^{\mathrm{bulk}}$ is the number of silicon atoms in bulk *c*-Si. Both relative concentrations are hence adimensional.[26]

Silicon is immobile and, owing to conservation of Si species which are only transformed from reactive to nitrided by reaction with nitridants species, silicon nitride concentration is taken as $\rho_{\mathrm{SiN}}(x,t)=1-\rho_{\mathrm{Si}}(x,t)$. Here $0\leqslant\rho_{\mathrm{Si}}(x,t)\leqslant 1$, implying coexistence of different nitridation states of Si, thus modeling the observed sub-nitride region leading to a nonabrupt interface between stoichiometric Si_3N_4 and the *c*-Si substrate. The evolution equations for the two relative concentrations are

$$\frac{\partial\rho_{\mathrm{N}}(x,t)}{\partial t}=D\rho_{\mathrm{Si}}(x,t)\frac{\partial^2\rho_{\mathrm{N}}(x,t)}{\partial x^2}-D\rho_{\mathrm{N}}(x,t)\frac{\partial^2\rho_{\mathrm{Si}}(x,t)}{\partial x^2}$$

$$-k\rho_{\mathrm{Si}}(x,t)\rho_{\mathrm{N}}(x,t)\frac{\partial\rho_{\mathrm{Si}}(x,t)}{\partial t}$$

$$=-k\rho_{\mathrm{Si}}(x,t),\rho_{\mathrm{N}}(x,t), \tag{2}$$

where k is a reaction rate and D is the diffusivity of the nitridant species in Si. The origin of the two second time derivatives can be understood by expressing reaction-diffusion of the nitridant species by finite differences:

$$\rho_{\mathrm{N}}(x,t+\Delta t)-\rho_{\mathrm{N}}(x,t)$$

$$=D\frac{\Delta t}{(\Delta x)^2}\{\rho_{\mathrm{Si}}(x,t)[\rho_{\mathrm{N}}(x-\Delta x,t)+\rho_{\mathrm{N}}(x+\Delta x,t)]$$

$$-\rho_{\mathrm{N}}(x,t)[\rho_{\mathrm{Si}}(x-\Delta x,t)+\rho_{\mathrm{Si}}(x+\Delta x,t)]\}$$

$$-k\Delta t\rho_{\mathrm{Si}}(x,t)\rho_{\mathrm{N}}(x,t), \tag{3}$$

that is, the diffusing nitridant species cannot ''jump'' into regions where there is fully nitrided Si, where $\rho_{\mathrm{SiN}}(x,t)=1-\rho_{\mathrm{Si}}(x,t)=1$, whereas the jumping probability into regions containing partially nitrided Si is modulated by the concentration of these species. The above expression may be completed to yield the two finite difference versions of the second derivatives in Eq. (2). This restriction upon nitridant species diffusion through stoichiometric silicon nitride is the origin of the difference between kinetics of silicon nitridation and silicon oxidation.[26]

The initial and boundary conditions, stating constant NH_3 pressure at the surface, P_N, and an initially bare silicon surface are written as

$$\rho_{\mathrm{N}}(0,t)=P_{\mathrm{N}},$$

$$\rho_{\mathrm{Si}}(x,0)=1. \tag{4}$$

One can adopt natural, adimensional units for length and time[26] given respectively by $\sqrt{D/k}$ and $1/k$, leading to an adimensional form of Eq. (2), which have one single parameter, namely P_N.

Model equations are solved using finite differences methods, that is, we rewrite Eq. (2) for discrete time and length (depth) and successively iterate in time. Figure 1(a) shows snapshots of the Si profile at increasing nitridation times. Since the profiles in Fig. 1 were calculated for fixed temperature and pressure, the depth scale could be reconverted into real, dimensional depth by using the corresponding $\sqrt{D/k}$ value.[26] Figure 1(a) shows that surface reactive Si is initially consumed, being transformed into silicon nitride and, as time proceeds, the reaction front advances into the *c*-Si matrix. The amount of nonfully nitrided silicon in the near-surface region decreases fast, preventing progressively incorporation of nitridant species and consequently slowing down film growth. For the two longest nitridations (time intervals ratio of 10^3) the relative widths of the near-surface, stoichiometric Si_3N_4 layer and of the near-interface layer presenting gradual transition from stoichiometric Si_3N_4 to substrate *c*-Si are comparable, modeling well experimental observations.[9,25] Figure 1(b) shows snapshots of the nitridant species profiles

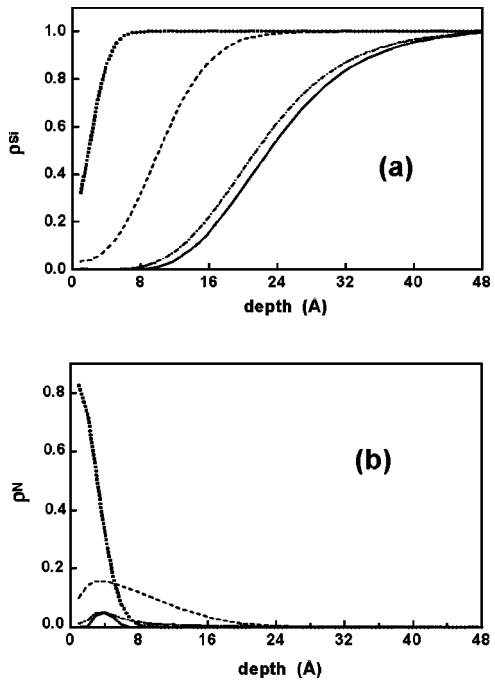

FIG. 1. (a) Snapshots of Si profiles for increasing—from left to right—nitridation times; (b) the corresponding nitridant species profiles for each nitridation time considered in (a).

for the same nitridation time intervals as Fig. 1(a). Initially, the profile is typical of a source of a certain adsorbed species, the nitridant species, that is diffusing and being consumed by reaction. The source concentration decreases with time, owing to the diffusion barrier represented by the Si_3N_4 layer just formed in the near-surface region. Concomitantly, the nitridant species profile also decreases, as it is being consumed by reaction with silicon in the reactive, sub-nitride region and with substrate-Si.

Nitride growth kinetics can be obtained by integrating $\rho_{SiN}(x,t)=1-\rho_{Si}(x,t)$ in all space.[26] In Fig. 2 experimental kinetics,[9] measured from early stages of very fast growth up to saturation into a self-limited regime are well fitted by model calculated kinetics. As both experimental kinetics were obtained at the same temperature we used the same

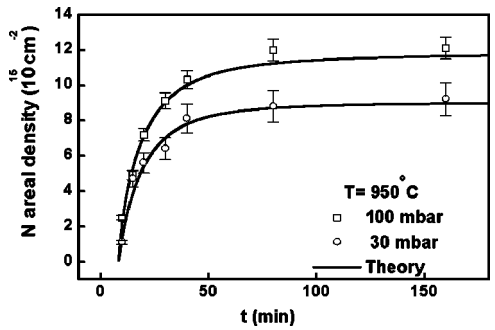

FIG. 2. Experimental kinetics of thermal growth of silicon nitride films on Si in NH_3 from Ref. 9 and calculated kinetics obtained by integration of the model equations. The two calculated kinetic curves (solid lines) used the same values of D and k and a ratio between NH_3 pressures (0.30) identical to the experimental one. Approximate thickness can be obtained by the equivalence relation 1×10^{15}N cm$^{-2}=0.18$ nm of silicon nitride.

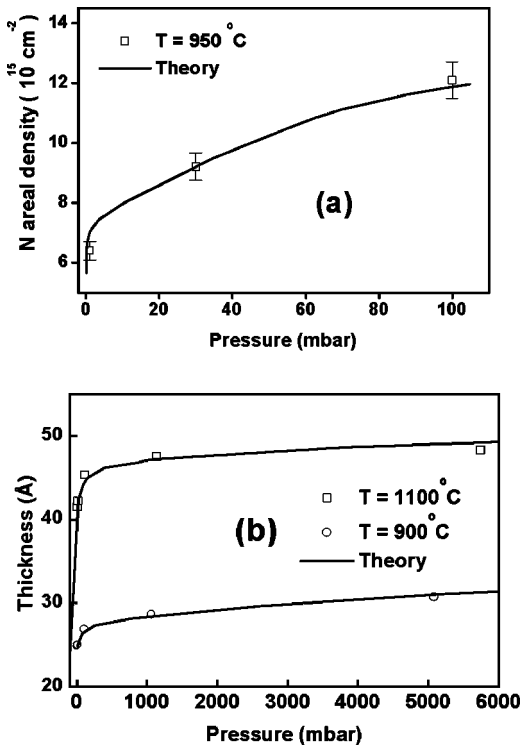

FIG. 3. Calculated film thickness at saturation as function of pressure (solid lines) and experimental data obtained from (a) Fig. 2 complemented by kinetics at very low pressure (1 mbar) of ^{15}N-enriched NH_3, measured by nuclear reaction analysis (12) and (b) Ref. 7.

natural units, namely same D and k for the calculations, while the ratio between the values of P_N for both calculated curves were identical (0.30) to that of the experimental ones.

Figure 3 shows experimental results of film thickness in the saturation, self-limited regime versus NH_3 pressure and theoretical curves obtained by iterating the evolution equations for very long times. The data taken from Fig. 2 above were complemented by kinetics at very low pressure (1 mbar) of ^{15}N-enriched NH_3, measured by nuclear reaction analysis.[12] For each nitridation temperature, convenient natural length and time units were adopted in the calculation by fixing D and k, whereas P_N was varied proportional to the experimental values. The behavior depicted in these figures can be understood as follows: The determinant quantity for the final film thickness is the total amount of nitridant species that is adsorbed and transported to reactive sites, integrated over the whole nitridation time, and not only the rate at which the adsorption takes place. In the lower pressures range, an increase in NH_3 pressure at sample surface produces a significant increase in film thickness because more nitridant species are adsorbed per unit time and more nitridation of Si occurs. However, in the higher pressure range, it can happen that nitridation occurs so fast that a stoichiometric Si_3N_4 layer quickly develops in the near-surface region, and consequently a diffusion barrier is built up earlier. That is, although the adsorption rate may be higher in this pressure range, it happens during a shorter time interval, leading to smaller total amounts of adsorbed nitridant species. Thus in the higher pressures range this effect of de-

creasing the time interval at which adsorption occurs grows in importance and the increase in total thickness with pressure is slower.

The present model is a simple approach to a phenomenon that may present other aspects which are not being taken into account. For example, diffusivity of nitridant species through stoichiometric Si_3N_4 is certainly much smaller than through sub-nitrides or pure silicon, but it may be different from zero. Also, recombination and/or decay of the nitridant species into inert species have been evidenced experimentally,[27,28] acting as sinks for nitridant species inside the sample. Nevertheless, this first approach is powerful enough to provide an adequate description of concentration profiles, kinetics, and pressure dependence of the final thickness of thermally grown silicon nitride films on Si. Further developments of the model incorporating some of these aspects are in progress.

ACKNOWLEDGMENT

This work has been partially supported by Brazilian Agencies CNPq, CAPES, and FAPERGS.

[1] D. A. Buchanan, IBM J. Res. Dev. **43**, 245 (1999).
[2] T. P. Ma, in *The Physics and Chemistry of SiO_2 and of the Si-SiO_2 Interface- 4*, edited by H. Massoud *et al.* (The Electrochemical Society, Pennington, NJ, 2000), Vol. 2000-2, p. 19.
[3] C. Chaneliere *et al.*, Mater. Sci. Eng. **R22**, 269 (1998).
[4] M. C. Gilmer *et al.*, in *Ultrathin SiO_2 and High-k Materials for ULSI Gate Dielectrics*, edited by H. R. Huft *et al.* (Mater. Res. Soc., Warrendale, 1999), Vol. 567, p. 323.
[5] G. B. Aler *et al.*, in *Ultrathin SiO_2 and High-k Materials for ULSI Gate Dielectrics*, edited by H. R. Huft *et al.* (Mater. Res. Soc., Warrendale, 1999), Vol. 567, p. 391.
[6] S. P. Murarka, C. C. Chang, and A. C. Adams, J. Electrochem. Soc. **126**, 996 (1979).
[7] Y. Hayafuji and K. Kajiwara, J. Electrochem. Soc. **129**, 2102 (1982).
[8] M. M. Moslehi and K. C. Saraswat, IEEE Trans. Electron Devices **ED-32**, 105 (1985).
[9] I. J. R. Baumvol *et al.*, J. Electrochem. Soc. **142**, 1205 (1995).
[10] J.-J. Ganem *et al.*, Nucl. Instrum. Methods Phys. Res. B **64**, 778C (1992).
[11] Wu, C. King, M. Lee, and C. Chen, J. Electrochem. Soc. **129**, 1559 (1982).
[12] I. J. R. Baumvol, Surf. Sci. Rep. **36**, 1 (1999).
[13] A. Omeltchenko *et al.*, Phys. Rev. Lett. **84**, 318 (2000).
[14] M. E. Bachlechner *et al.*, Phys. Rev. Lett. **84**, 322 (2000).
[15] C. H. F. Peden *et al.*, Phys. Rev. B **47**, 15 622 (1993).
[16] G. Dufour *et al.*, Surf. Sci. **304**, 33 (1994).
[17] F. Rochet *et al.*, Surf. Sci. **320**, 371 (1994).
[18] C. Maillot *et al.*, Appl. Surf. Sci. **26**, 326 (1986).
[19] I. J. R. Baumvol *et al.*, Nucl. Instrum. Methods Phys. Res. B **B118**, 499 (1996).
[20] F. Bozso and Ph. Avouris, Phys. Rev. Lett. **57**, 1185 (1986).
[21] A. Glachant, D. Saidi, and D. F. Delord, Surf. Sci. **168**, 672 (1986).
[22] J. L. Bischoff *et al.*, Surf. Sci. **251/252**, 170 (1991).
[23] C. U. S. Larsson and A. S. Flodstrom, Surf. Sci. **241**, 353 (1991); **271**, 349 (1992).
[24] S. M. Cherif, J.-P. Lacharme, and C. A. Sebenne, Surf. Sci. **251/252**, 737 (1991); **262**, 33 (1992); **274**, 258 (1992).
[25] S. W. Sun, P. J. Tobin, J. Weiheimer, and E. Reed, J. Electrochem. Soc. **134**, 1799 (1984).
[26] R. M. C. de Almeida, S. Gonçalves, I. J. R. Baumvol, and F. C. Stedile, Phys. Rev. B **61**, 12 992 (2000).
[27] P. A. Taylor *et al.*, Surf. Sci. **215**, L286 (1989).
[28] P. J. Chen, M. L. Colaianni, and J. T. Yates, Jr., Surf. Sci. Lett. **274**, L605 (1992).