

Predictive modeling of atomic layer deposition on the feature scale

Matthias K. Gobbert^a, Vinay Prasad^{b,*}, Timothy S. Cale^b

^aDepartment of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA

^bFocus Center—New York, Rensselaer: Interconnections for Gigascale Integration, Rensselaer Polytechnic Institute, CII 6015, 110 8th Street, Troy, NY 12180-3590, USA

Received 26 October 2001; received in revised form 21 February 2002; accepted 22 February 2002

Abstract

A feature scale simulator for atomic layer deposition (ALD) is presented that combines a Boltzmann equation transport model with chemistry models. A simple but instructive chemistry is considered; one reactant species adsorbs onto the surface, and a second reactant reacts with it from the gas phase (Eley–Rideal). This work includes potential desorption of the adsorbed species during purge steps, which may or may not play a role in any given ALD system. Three sets (cases) of rate parameters are chosen to compare chemical rates with transport rates. The duration of the ALD pulses and the geometry of the representative feature are the same for each case. Simulation results are presented for all four steps in one ALD cycle, adsorption, post-adsorption purge, reaction, and post-reaction purge. The results are extended to multiple ALD cycles, and the monolayers per cycle are estimated. We highlight the potential trade-off between pulse durations and deposition rate (wafer throughput); e.g. the time penalty required to increase the amount adsorbed during the adsorption step. The simulation methodology we present can be used to determine the pulse durations that maximize throughput for a given chemistry and chemical rate parameters. One overall observation is that transport is fast relative to chemical reactions, for reasonable kinetic parameters. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Deposition process; Monolayers; Boltzmann equation; Atomic layer deposition

1. Introduction

Atomic Layer Deposition (ALD) involves pulsing reactant gases over a substrate in series, with purges of an inert gas between reactant pulses. Typically, a gaseous species (A) is fed into the reactor, perhaps in a carrier gas, and adsorbs on the surface in the first step of the ALD cycle. The reactor is then purged with the inert gas, and a second gaseous reactant (B) is pulsed into the reactor. The adsorbed A reacts with B to deposit a layer of film on the substrate, or on previously deposited film. Surface sites for adsorption of A are regenerated as the reaction proceeds. The reactor is then purged again, and the next ALD cycle is started with a pulse of A. The duration of each pulse of the ALD cycle is adjusted to increase the rate of deposition, while maintaining the properties of conformality and uniformity that make ALD attractive.

A goal of process optimization is to achieve deposition as fast as possible while meeting film property constraints. The deposition rate can be expressed in several ways, but they all measure wafer throughput. One informative way to express deposition rate is ‘monolayers of deposited film per cycle’ divided by the ‘period of one ALD cycle’. In order to increase the overall deposition rate, one can shorten the duration of each ALD cycle or increase the fraction of a monolayer deposited per cycle. It has been observed that if steps in each cycle are too short, the fraction of monolayer deposited per cycle decreases [1,2]. However, more cycles can be run in a given process time, if each is shorter, and that may make up for the reduced effectiveness of one cycle. A trade-off can exist between the number of monolayers deposited per cycle and the duration of a cycle, even assuming no degradation in across wafer or intrafeature uniformity.

This paper uses a reasonable chemistry and reasonable chemical rate parameters (adsorption, desorption and

*Corresponding author. Tel.: +1-518-276-2814.

E-mail address: prasav@rpi.edu (V. Prasad).

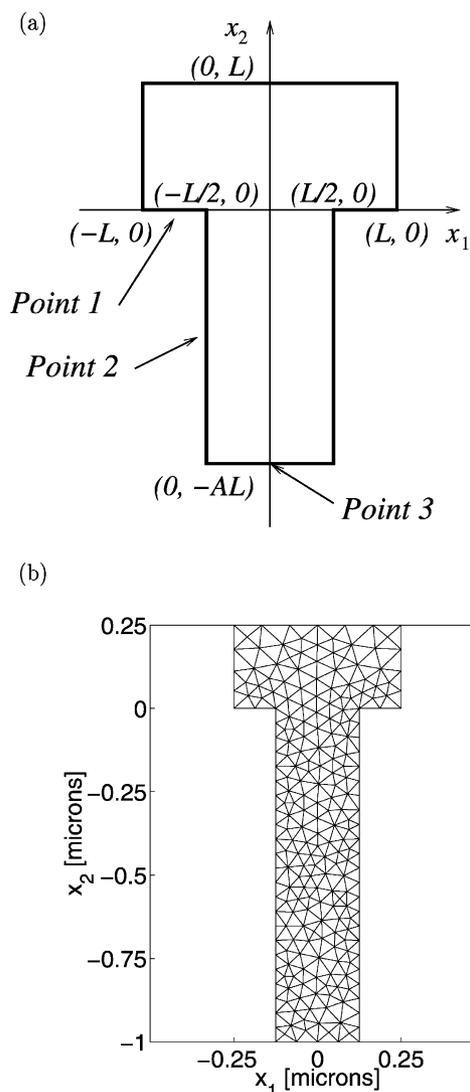


Fig. 1. (a) Schematic of a two-dimensional domain defining length L and aspect ratio A . (b) Numerical mesh for the feature with $L=0.25$ μm and aspect ratio $A=4$.

surface reaction) to demonstrate how simulation can be used to evaluate the trade-off between pulse durations and monolayers per cycle deposited. For this discussion, we consider one monolayer deposited per cycle in one-second equivalent to, for instance, half a monolayer per cycle in half a second.

We focus on the feature scale; i.e. we consider transport and reaction in and near a feature that might be found in a multilevel metallization process flow during the fabrication of integrated circuits (ICs). Transport and reaction during ALD are inherently transient, and it is not clear that the widely used approach to feature scale modeling of low-pressure deposition processes is valid. In that approach, species fluxes are assumed to be constant over small, but reasonable time scales [3–6]. In a previous paper, we dealt with transient

adsorption and desorption that might occur during the adsorption and purge steps in an ALD cycle [7]. We showed that, depending upon the adsorption and desorption rates, it may well be reasonable to assume that the fluxes are not only constant in time, but are also spatially uniform along the feature's surface.

In order to maintain our focus on the feature scale, we idealize the changes at the reactor scale; changes in flows and concentrations occur such that changes directly over the feature can be considered to occur as step functions. Thus, we consider the case in which all fluxes from the reactor volume to the feature mouth are constant in time, except for step changes; e.g. from zero to some specified value. Though our results depend upon this assumption quantitatively, it does not impact the main messages in the paper. To go further, we would need to integrate a reactor scale model with our feature scale model, as has been done for LPCVD [8–10] and ECD [11,12]. Reactor scale models that describe the formation of a traveling wave of the reactant are found in Siimon and Aarik [6,13,14].

The following Model section introduces the transport and reaction models used in our simulations, with a particular emphasis on the surface reaction model and its coefficients. The Results and discussion section details the rationale for the three chemistries chosen, then explains the simulation results, and concludes by discussing trade-offs.

2. Model

The domain of the feature scale model comprises the inside of one feature and a small region of the gas-phase above the feature mouth; Fig. 1a shows the domain under consideration. We focus on a feature of width $L=0.25$ μm and aspect ratio $A=4$, giving a feature depth of 1.0 μm . The appropriate transport model for representative operating conditions is given by a system of Boltzmann equations [15–17]. On the feature scale, the mean-free path of the molecules is on the order of 100 μm , and the Knudsen number satisfies $\text{Kn} \approx 100 \gg 1$. Hence, we are in the free molecular flow regime, and the transport model consists of a linear Boltzmann equation for each reactive species with a zero right-hand side. Using a Galerkin ansatz in velocity space, each Boltzmann equation is converted to a system of linear hyperbolic conservation laws in the time and spatial variables. The discontinuous Galerkin (DG) method [18] is used to solve the resulting system on the mesh shown in Fig. 1b.

In order to aid in analysis and the presentation of results, we non-dimensionalize the Boltzmann transport equations, the number density of the reacting species, the fluxes of the species to the surface of the feature, the reaction rates of the surface reactions, and the surface coverage of the adsorbing species. The Boltzmann trans-

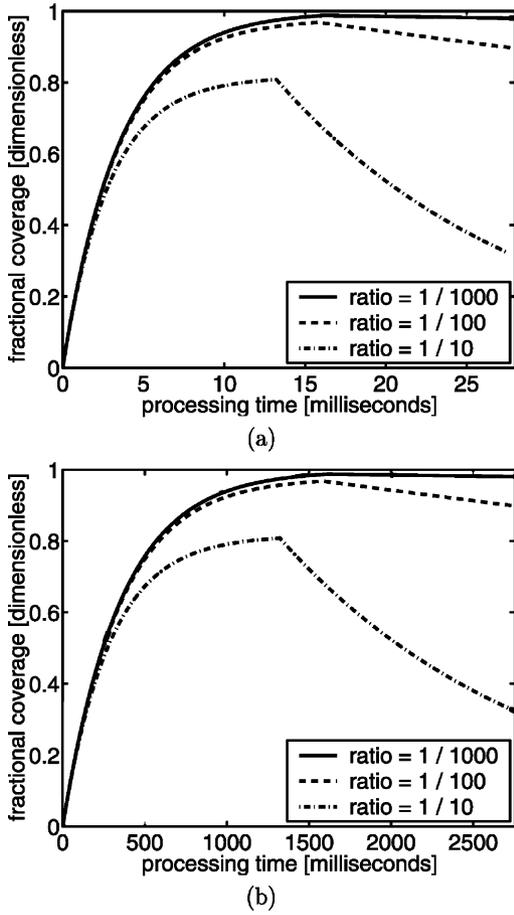


Fig. 2. Fractional coverage vs. process time for adsorption and purge together for various ratios of γ_1^b/γ_1^f with (a) $\gamma_1^f = 10^{-2}$, (b) $\gamma_1^f = 10^{-4}$. Process time for adsorption is depicted up to 99% of equilibrium coverage. Notice the different scales on the time axes.

port equations are non-dimensionalized using a reference concentration and reference speed based on reactant A. Using the ideal gas law, the reference concentration is chosen as

$$c^{\text{ref}} = \frac{P_A}{R_g T}, \quad (1)$$

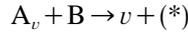
where P_A is the partial pressure of species A based on a specified mole fraction, and R_g denotes the universal gas constant. The reference speed is chosen as the thermal average speed of species A, and is given by

$$v^{\text{ref}} = \sqrt{\frac{R_g T}{\omega_A}} \quad (2)$$

where ω_A is the molecular weight of species A. We assume that the number density of species B in the reactor during the reaction pulse is half that of A during the adsorption pulse, and its thermal (average) speed is 0.9 times that of A. The fluxes of A and B are non-

dimensionalized with respect to a reference flux, which is chosen to be the product of the reference concentration and reference speed used to non-dimensionalize the Boltzmann transport equations.

The surface chemistry model consists of reversible adsorption of A on a single site, and irreversible reaction of B with the adsorbed A [19], viz



where A_v is adsorbed A, v stands for a surface site

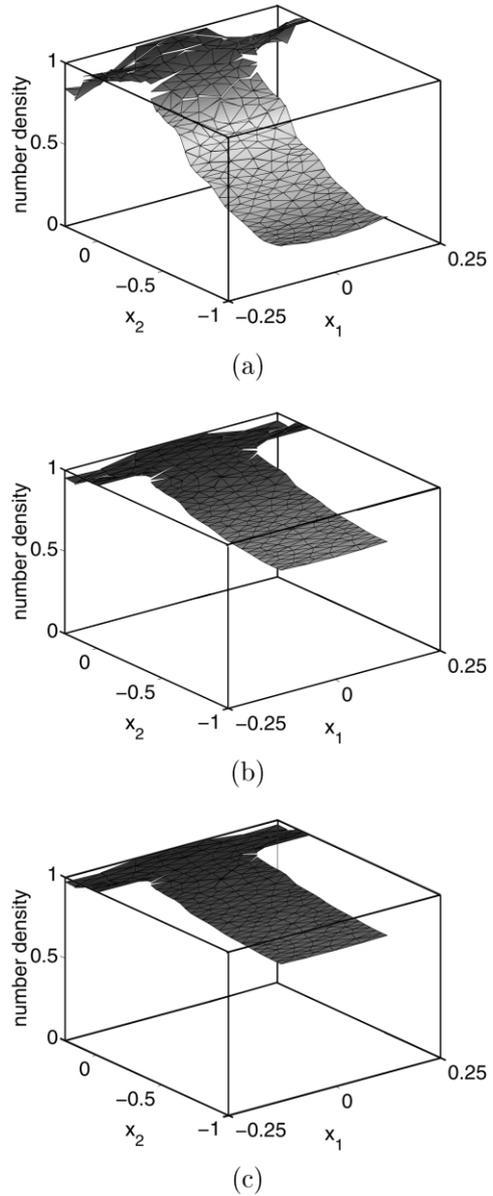


Fig. 3. Case 1, adsorption step: dimensionless number density of species A for a feature with aspect ratio $A=4$ at times (a) 10.0 ns; (b) 40.0 ns; (c) 80.0 ns. Note the different scales on the x_1 - and the x_2 -axes.

available for adsorption, and (*) is the non-adsorbing gaseous product. The first reaction models the adsorption and desorption of molecules of species A at the wafer surface. The second reaction (Eley–Rideal) models the reaction of gaseous molecules of species B with adsorbed molecules of species A at the wafer surface, which is assumed to be irreversible. Notice that surface sites available for adsorption are produced by the reaction of B with adsorbed A.

Let $0 \leq \vartheta_A(x, t) \leq 1$ denote the fraction of the wafer surface at position x and time t that is occupied by adsorbed molecules of reactive species A. If $\hat{\eta}_1(x, t)$ and $\hat{\eta}_2(x, t)$ denote the dimensionless fluxes of species A and B to the surface, respectively, the dimensionless reaction rates for the surface reactions are modeled by

$$\hat{R}_1 = \gamma_1^f(1 - \vartheta_A)\hat{\eta}_1 - \gamma_1^b\vartheta_A, \quad (4)$$

$$\hat{R}_2 = \gamma_2^f\vartheta_A\hat{\eta}_2$$

with the dimensionless coefficients γ_1^f , γ_1^b and γ_2^f . The coefficients γ_1^f and γ_1^b control the adsorption and desorption of molecules of A, respectively, and γ_2^f controls the reaction of B with adsorbed A.

The evolution of the fractional surface coverage $\vartheta_A(x, t)$ is modeled by the differential equation

$$\frac{d\vartheta_A(\hat{x}, \hat{t})}{d\hat{t}} = \alpha_p(\hat{R}_1(\hat{x}, \hat{t}) - \hat{R}_2(\hat{x}, \hat{t})), \quad \hat{x} \in \Gamma_w, \quad (5)$$

at every point \hat{x} on the wafer surface Γ_w , where α_p denotes a constant pre-factor from the non-dimensionalization procedure. This differential equation is supplied with an initial condition that represents the initial fractional coverage ϑ_A^{ini} at the beginning of the process A step.

Remark: It is in general impossible to find a closed-form solution $\vartheta_A(\hat{t})$ to the differential equation in Eq. (5), because the coefficients involving $\hat{\eta}_1$ and $\hat{\eta}_2$ are not constant. But if these fluxes are constant, then Eq. (5) becomes a first-order linear ordinary differential equation with constant coefficients and can be solved analytically. Specifically, at each point on the feature surface, we have the problem

$$\frac{d\vartheta_A(\hat{t})}{d\hat{t}} = -\alpha_p b \vartheta_A(\hat{t}) + \alpha_p \gamma_1^f \hat{\eta}_1, \quad \vartheta_A(0) = \vartheta_A^{\text{ini}} \quad (6)$$

with $\alpha_p = (\eta^* t^*)/S_T$ and $b = \gamma_1^f \hat{\eta}_1 + \gamma_1^b + \gamma_2^f \hat{\eta}_2$, which has the solution

$$\vartheta_A(\hat{t}) = \vartheta_A^\infty (1 - e^{-\alpha_p b \hat{t}}) + \vartheta_A^{\text{ini}} e^{-\alpha_p b \hat{t}} \quad (7)$$

with the equilibrium (long time) limit

$$\vartheta_A^\infty = \frac{\gamma_1^f \hat{\eta}_1}{\gamma_1^f \hat{\eta}_1 + \gamma_1^b + \gamma_2^f \hat{\eta}_2} \quad (8)$$

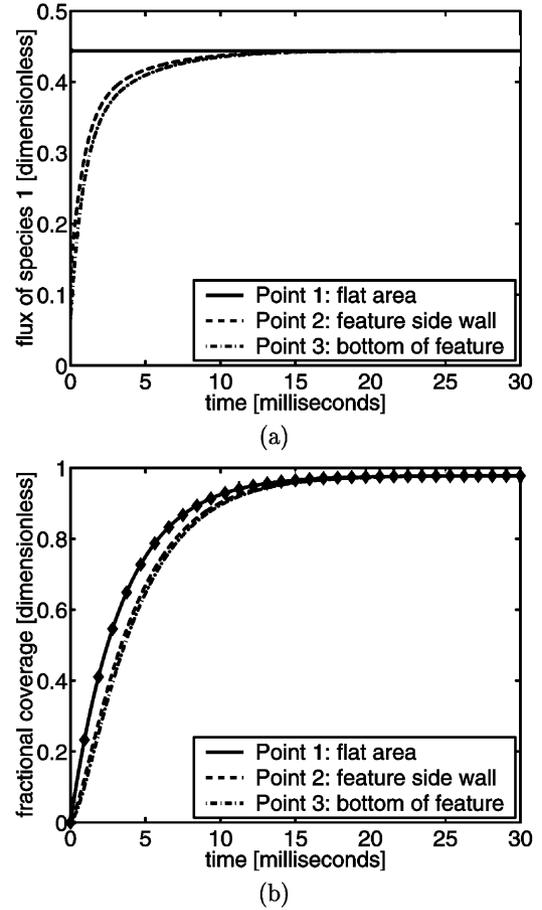


Fig. 4. Case 1, adsorption step: (a) dimensionless flux to the surface of species A; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

3. Results

3.1. Case studies

Studies of parameter values in the surface reaction model presented above are presented in Gobbert et al. [7]. In general, those studies show that transport is fast compared to typical pulse durations. For example, after introducing a species (reactant pulse starts), the fluxes to the surface attain steady state values in time frames on the order of tens of milliseconds. Because of that, the fluxes are approximately constant and uniform during adsorption and reaction process steps and the analytic solution given by Eq. (7) can predict the fractional coverage of adsorbed molecules of species A very well for long times. Hence, we use the analytic solution as a guide for which reaction coefficients to consider for further study.

Case studies for the adsorption and post-adsorption purge steps [7] show that the ratio of the adsorption coefficient γ_1^f to the desorption coefficient γ_1^b is crucial

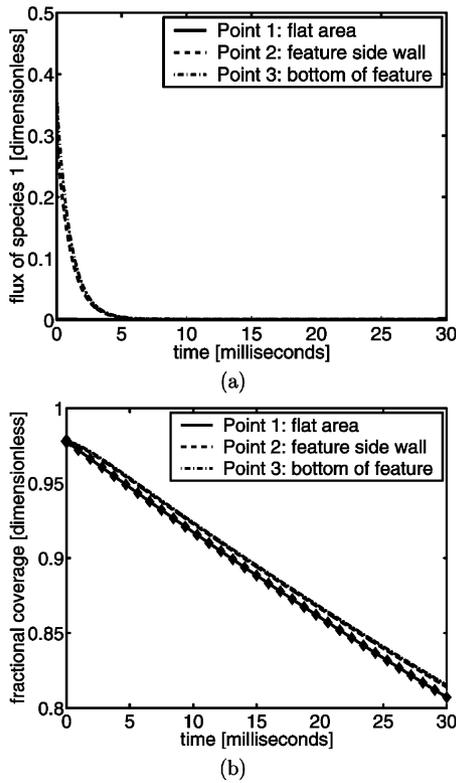


Fig. 5. Case 1, post-adsorption purge step: (a) dimensionless flux to the surface of species A; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

to the behavior during the adsorption-purge sequence. Fig. 2 summarizes this dependence on the ratio of coefficients: the larger the adsorption coefficient, the faster the coverage increases, but desorption needs to be considered as well. The purge step can lead to a significant decrease in coverage if the ratio γ_1^b/γ_1^f is larger than 1/100 for any value of γ_1^f [7]. If this ratio is smaller than 1/100, e.g. 1/1000, the decrease of coverage during the purge is negligible for any value of γ_1^f and the length of the purge step does not matter as much. We consider an interesting case by fixing $\gamma_1^b = \gamma_1^f/100$.

Case studies for the reaction and post-reaction purge steps show that the ratio between the reaction coefficient γ_2^f and desorption coefficient γ_1^b is crucial. If they are equal, desorption of A will dominate during the reaction step, and the decrease in coverage of adsorbed A is due more to desorption than to reaction between A and B. Thus, the desired deposition of the product of A–B will not take place at the levels expected based upon adsorption rates. If the reaction coefficient is larger than the desorption coefficient, then reaction will dominate and more deposition will take place, see Gobbert et al. [20].

Based upon the above summary of previous results

on this simple but illustrative chemistry, we consider the following three cases:

- Case 1: $\gamma_1^f = 10^{-2}$, $\gamma_1^b = 10^{-4}$, $\gamma_2^f = 10^{-2}$,
- Case 2: $\gamma_1^f = 10^{-4}$, $\gamma_1^b = 10^{-6}$, $\gamma_2^f = 10^{-2}$,
- Case 3: $\gamma_1^f = 10^{-4}$, $\gamma_1^b = 10^{-6}$, $\gamma_2^f = 10^{-4}$,

Note that $\gamma_1^b = \gamma_1^f/100$ and $\gamma_1^b \ll \gamma_2^f$ in all cases. We assume that values of the adsorption and reaction coefficients closer to unity than 10^{-2} are unlikely in ALD practice; therefore, we consider the two values $\gamma_1^f = 10^{-2}$ and 10^{-4} . The coefficient γ_1^b is then fixed by $\gamma_1^b = \gamma_1^f/100$. And the remaining coefficient γ_2^f is chosen to obey the restriction $\gamma_1^b \ll \gamma_2^f$.

For comparison purposes, we fix the duration of each

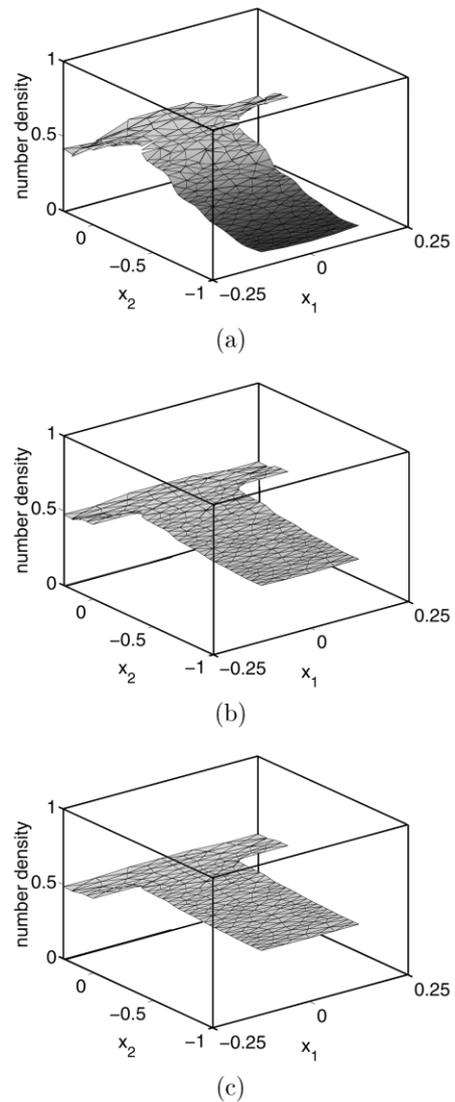


Fig. 6. Case 1, reaction step: dimensionless number density of species B for a feature with aspect ratio $A=4$ at times (a) 10.0 ns; (b) 40.0 ns; (c) 80.0 ns. Note the different scales on the x_1 - and the x_2 -axes.

step of one ALD cycle for all cases. The adsorption step and the reaction step each last 450 ms, the post-adsorption purge step and the post-reaction purge step each last 50 ms. Shorter purge times are preferable, and we demonstrated that transport is fast [7]. Times much shorter than 50 ms would suffice to evacuate reactant from features on a wafer surface, as has been shown in Gobbert et al. [7] for the post-adsorption purge. But purge times are limited by the capabilities of the equipment to switch, which we make no attempt to model in this paper. We therefore, fix the purge time at 50 ms.

For the simulation results presented, the initial values of the number densities (A and B), fluxes (A and B), and surface coverage of A for each process step are taken to be the values at the end of the previous step. They are assumed to be spatially uniform [7], though this can be relaxed. We present results for number densities, fluxes, and surface coverage of A in dimensionless form. Based on the non-dimensionalization procedure described in the Model section, the dimensionless number density of A lies between 0 and 1, and the dimensionless number density of B lies between 0

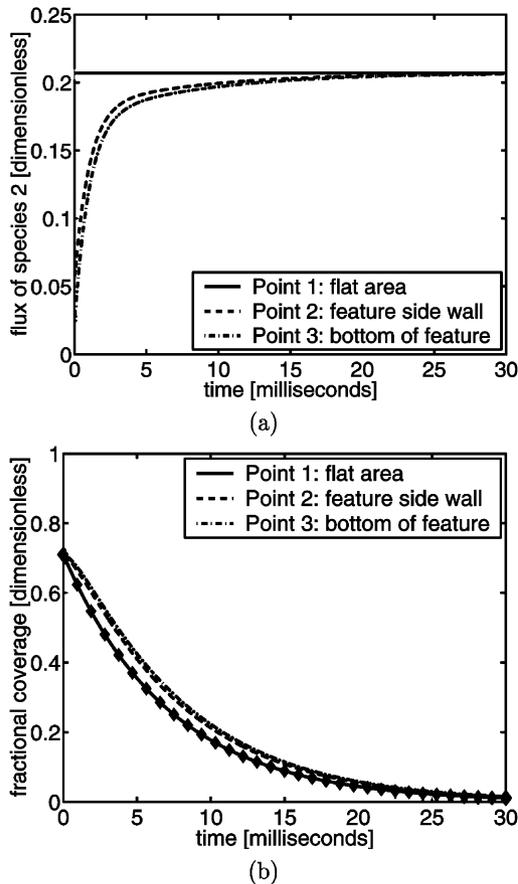


Fig. 7. Case 1, reaction step: (a) dimensionless flux to the surface of species; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

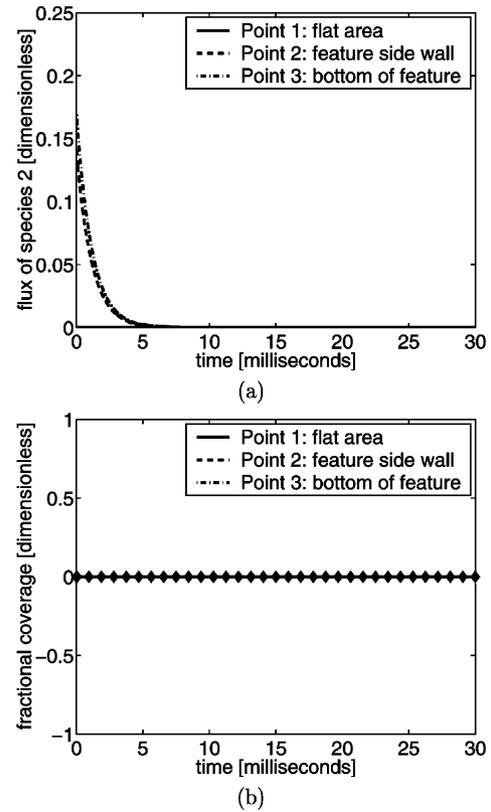


Fig. 8. Case 1, post-reaction purge step: (a) dimensionless flux to the surface of species B, (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

and 0.5. The dimensionless surface coverage of A lies between 0 and 1.

3.2. Simulation results for Case 1

In Case 1, the adsorption coefficient is $\gamma_1^f = 10^{-2}$, the desorption coefficient is $\gamma_1^b = 10^{-4}$, and the reaction coefficient is $\gamma_2^f = 10^{-2}$. Among the three cases, Case 1 represents the one with the fastest adsorption. Simulation results for this case are shown in Figs. 3–8.

Fig. 3 shows the increase in the (dimensionless) number density of A throughout the feature, with the feature filling from the top. It is completely filled within approximately 100 ns; i.e. transport is much faster than the pulse duration. In Gobbert et al. [7], an analytic solution was used to predict that the fractional coverage ϑ_A reaches 99% of its steady state value in approximately 1.6 ms. This is confirmed by Fig. 4a,b in which both the flux of A to the surface and the fractional coverage attain their steady state values of 0.444 and 0.98, respectively, much more quickly than the pulse duration. These observations confirm that Case 1 has fast adsorption.

Fig. 5 shows the results for the post-adsorption purge step of the ALD cycle. The flux of A goes to zero

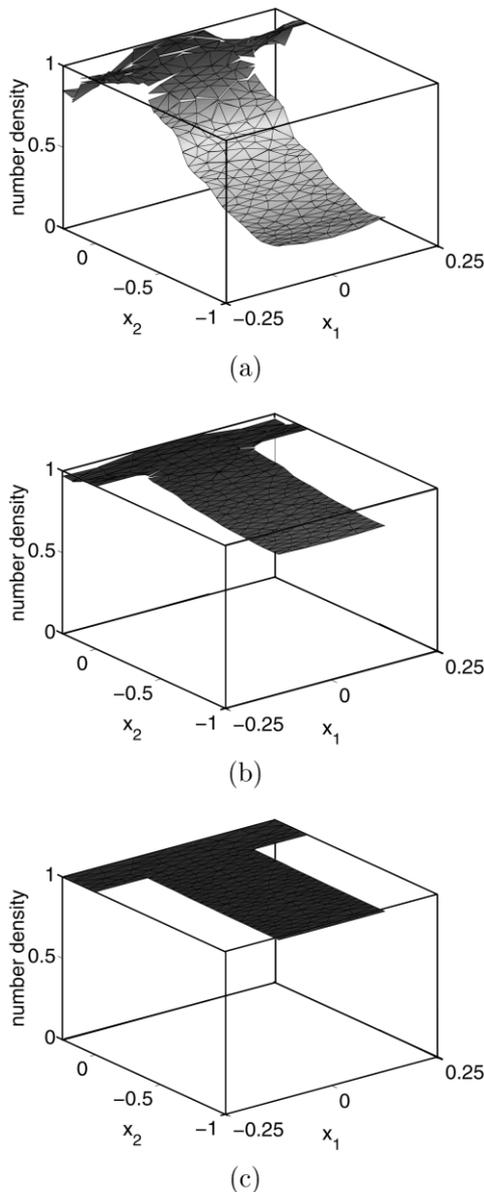


Fig. 9. Case 2, adsorption step: dimensionless number density of species A for a feature with aspect ratio $A=4$ at times (a) 10.0 ns, (b) 40.0 ns, (c) 80.0 ns. Note the different scales on the x_1 - and the x_2 -axes.

quickly, and the fractional coverage of A decreases essentially linearly. Notice that this decrease during the purge results in a significant loss of adsorbed molecules of A at the surface. The purge time should be as short as possible, because desorption of A during this step decreases the amount of film that potentially can be deposited in the following reaction step. As discussed above, we assume for our simulations that the purge time is 50 ms. Using the analytic solution to predict the value of the coverage at the end of the purge step yields $\Phi_A=0.71$. This is significantly less than the value of 0.98 at the start of the purge. Not shown are the results for the number density of A throughout the feature; the

gaseous species A is evacuated from the domain within 100 ns.

For the reaction step, we obtain the simulation results in Figs. 6 and 7. Fig. 6 demonstrates that the feature fills completely with reactive gas B during the first 100 ns of the reaction step. The dimensionless number density depicted in Fig. 6 increases only to approximately 0.5, because the number density of species A was used to non-dimensionalize the equation. Fig. 7a shows that the flux of B to the surface goes to its steady state value within tens of milliseconds. The corresponding flux of A is solely the result of continued desorption of A from the wafer surface. It remains at negligible levels after the previous purge step and is not shown. Fig. 7b shows the decrease in fractional coverage during the reaction step.

Notice that the simulation time of 30 ms is long enough for the coverage to decrease to zero; i.e. faster than the 450 ms duration of the reaction step. This is a consequence of the value $\gamma_2^f=10^{-2}$ of the reaction coefficient. Note that the decrease of coverage is caused both by desorption of adsorbed molecules of A and by reaction of gaseous B with adsorbed A. But due to our choice of $\gamma_1^b \ll \gamma_2^f$, the contribution of desorption is

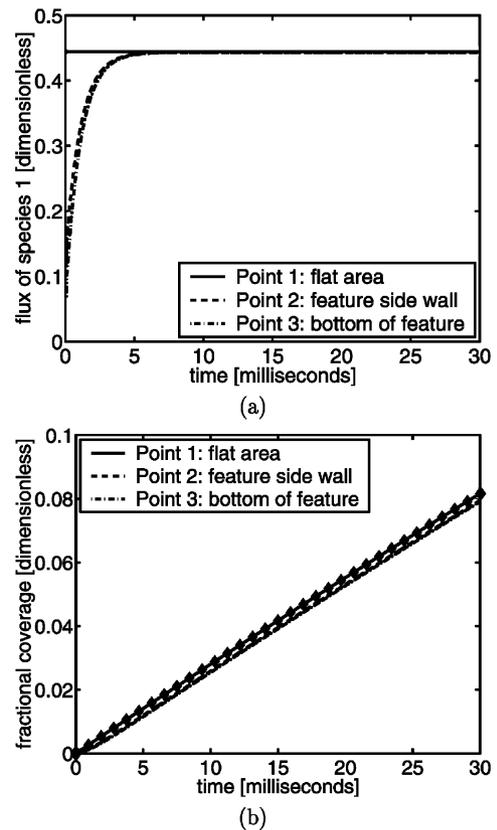


Fig. 10. Case 2, adsorption step: (a) dimensionless flux to the surface of species A; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

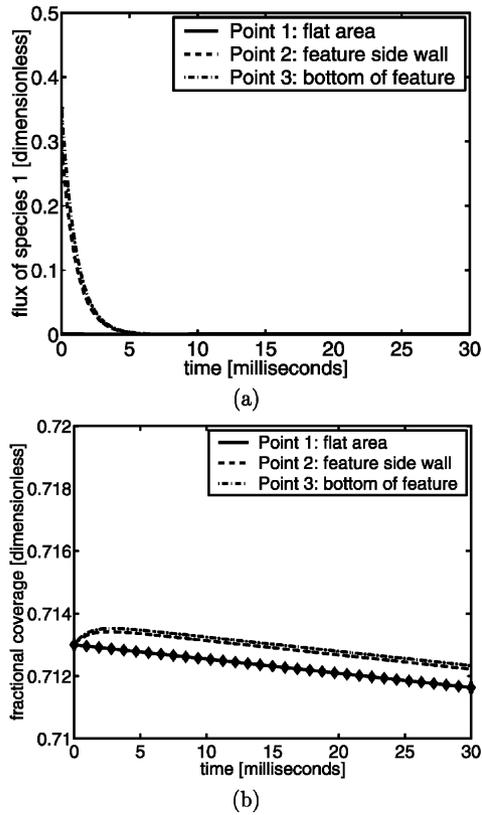


Fig. 11. Case 2, post-adsorption purge step: (a) dimensionless flux to the surface of species A; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

much less significant than the contribution of the reactions; see the previous subsection.

Fig. 8 shows how rapidly the flux of B to the surface decreases during the postreaction purge step. This reflects the quick evacuation of gaseous B from the domain. For completeness, Fig. 8b shows the evolution of the fractional coverage of A during the purge step; it remains at negligible levels throughout, as expected, since all available adsorbed A was consumed in the previous reaction step.

3.3. Simulation results for Case 2

In Case 2, the adsorption coefficient is $\gamma_1^f = 10^{-2}$, the desorption coefficient is $\gamma_1^b = 10^{-6}$, and the reaction coefficient is $\gamma_2^f = 10^{-2}$. Simulation results for this case are shown in Figs. 9–14. The choice of the adsorption coefficient $\gamma_1^f = 10^{-4}$ reflects the observation that the corresponding choice for Case 1 may be higher than values expected in practice. The value of γ_1^b is fixed as $\gamma_1^b = \gamma_1^f/100$. The reaction coefficient has the same value as in Case 1.

Fig. 9 confirms that the feature fills with A during the adsorption step as fast as in Case 1. While Fig. 10a,

shows that the flux of A to the surface tends to its steady state value as fast as in Case 1, Fig. 10b, demonstrates that the fractional coverage does not increase as fast as before due to the lower value for γ_1^f . Based on a prediction using the analytic solution, we see that the coverage only increases to 0.71 during the 450 ms of the adsorption step; this is short of the steady state value of 0.98.

Fig. 11 shows the results for the flux of A and the coverage during the post-adsorption purge step. As in Case 1, the flux decreases to zero within 10 ms. Using the spatially uniform value of 0.71 as initial condition for the coverage ϑ_A Fig. 11b, shows the decrease of

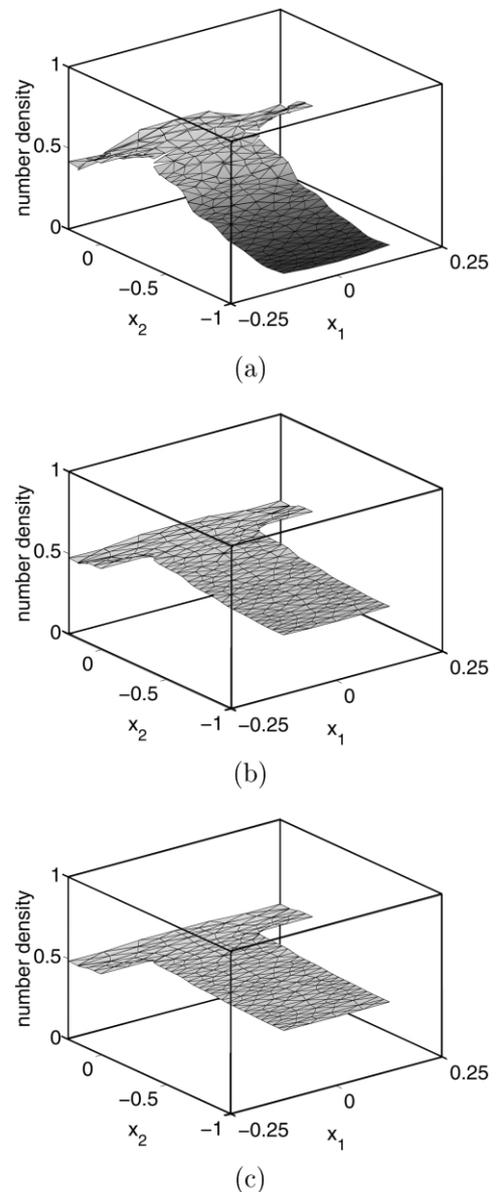


Fig. 12. Case 2, reaction step: dimensionless number density of species B for a feature with aspect ratio $A=4$ at times (a) 10.0 ns, (b) 40.0 ns, (c) 80.0 ns. Note the different scales on the x_1 - and the x_2 -axes.

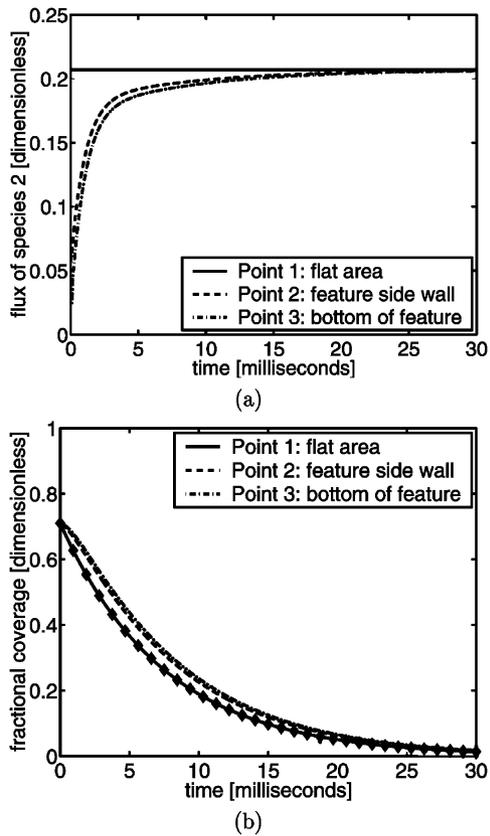


Fig. 13. Case 2, reaction step: (a) dimensionless flux to the surface of species B; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

adsorbed A during this purge. The initial increase of fractional coverage at points 2 and 3, which are located inside the feature, is due to continued desorption of A from the wafer surface that re-adsorb at those points. No corresponding increase is observed at point 1, because no points of the wafer surface are visible to it. Notice however that the decrease of coverage at each point as well as the absolute difference of coverage between the three observation points is small compared to their initial value of 0.71.

To predict the value of the surface coverage after 50 ms of post-adsorption purge, we use the analytic solution (as in Gobbert et al. [7]) to find the predicted value of 0.71. Note that this is the same value as was found at the beginning of the reaction step in Case 1. This demonstrates that despite the differences in parameter values, two chemistries can yield the same value of ϑ_A at this crucial stage of the cycle. If the relative size of adsorption and desorption coefficients are assumed to be constant, as we do here, any advantage of higher coverage at the end of the adsorption step might be obtained at the price of greater loss of coverage during the following purge step. Of course, step durations play a critical role as well, which we have fixed here.

Figs. 12 and 13 show the results for the reaction step in Case 2. Fig. 12 shows that the feature fills with molecules of species B as quickly and to the same level as for Case 1 (Fig. 6). Fig. 13a demonstrates that the flux of B to the surface tends to its steady state value as quickly as in Case 1 (Fig. 7a). Also, Fig. 13b shows that the coverage has essentially the same behavior as the coverage for Case 1 (Fig. 7b), however, the values are slightly higher than for Case 1; this reflects the fact that the reaction coefficients γ_2^f agree, but the desorption coefficient γ_1^b for Case 2 is two orders of magnitude smaller than for Case 1.

Fig. 14a,b show the flux of B and the fractional coverage results for the postreaction purge step for Case 2. They agree with the corresponding results for Case 1 (Fig. 8).

3.4. Simulation results for Case 3

In Case 3, the adsorption coefficient is $\gamma_1^f = 10^{-4}$, the desorption coefficient is $\gamma_1^b = 10^{-6}$, both as in Case 2, but the reaction coefficient is $\gamma_2^f = 10^{-4}$. This case combines the lower value of the adsorption coefficient γ_1^f with a reaction coefficient γ_2^f of equal value. Since the adsorption and desorption coefficients are the same

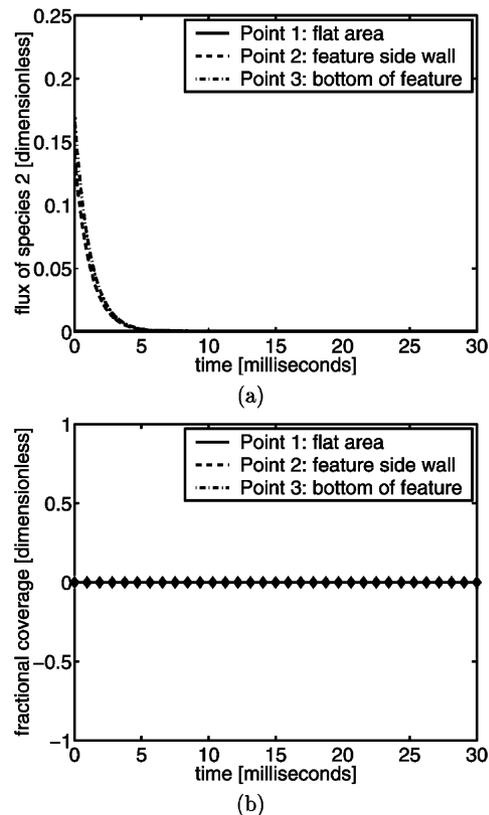


Fig. 14. Case 2, post-reaction purge step: (a) dimensionless flux to the surface of species B; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

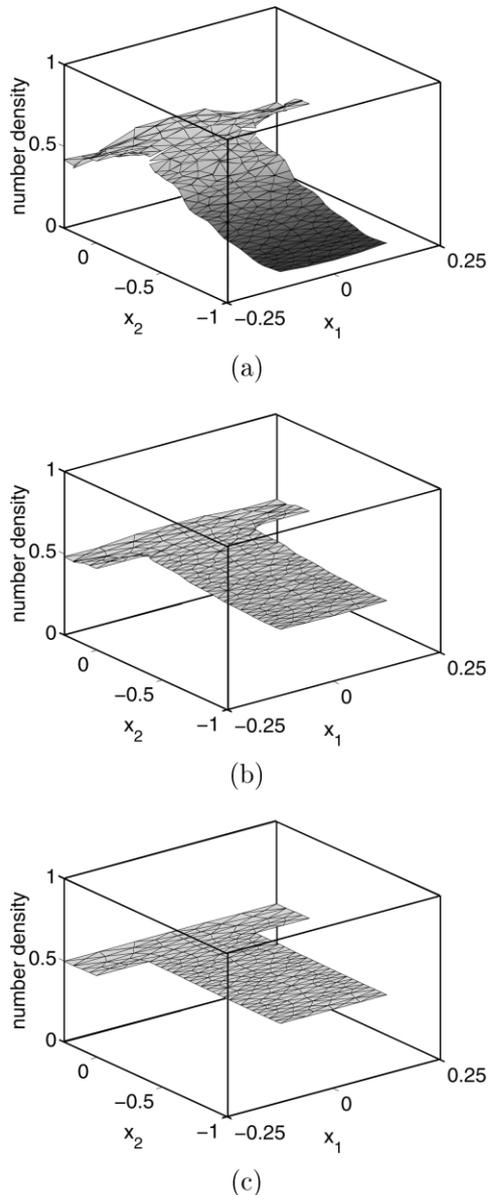


Fig. 15. Case 3, reaction step: dimensionless number density of species B for a feature with aspect ratio $A=4$ at times (a) 10.0 ns, (b) 40.0 ns, (c) 80.0 ns. Note the different scales on the vertical axes.

as in Case 2, the results for the adsorption step and the post-adsorption purge step are the same as in Case 2 (Figs. 9–11). Figs. 15–17 summarize the results for the reaction step and the post-reaction purge step.

As in both previous cases (Fig. 6 and Fig. 12), the plots of the number density of B in Fig. 15 show that the feature fills quickly with gaseous species B. As a result, the flux of B to the surface increases to its steady state value within 10 ms, as shown in Fig. 16a. However, due to the lower reaction coefficient γ_2^f than in Cases 1 and 2, fewer molecules of adsorbed A at the wafer surface are consumed in reactions with B. Hence, the decrease in fractional coverage in Fig. 16b is smaller

than in Cases 1 and 2 (Fig. 7b and Fig. 13b). Recall that the reaction coefficient γ_2^f is larger than the desorption coefficient γ_1^b ; this implies that the dominant reason for loss of coverage in Fig. 16b is reaction. This is important, because we want to emphasize the fact that there are two ways to lose adsorbed A; desorption of A and reaction. Reaction with B is desired, as it results in deposition.

Fig. 17 summarizes the results for the post-reaction purge. The flux of B tends to zero quickly, as shown in Fig. 17a. But the coverage in Fig. 17b starts at a positive value (contrary to Fig. 8b and Fig. 14b for Cases 1 and 2), because the previous reaction step has not depleted all molecules of adsorbed A on the wafer surface. Since the desorption coefficient γ_1^b is smaller than in Case 1, there is not a significant reduction in coverage during the process time for this purge step. Notice that the values of ϑ_A in Fig. 17b are approximately the same, and the coverage is essentially spatially uniform for the chemistry under consideration.

3.5. Comparison of the case studies

The three previous subsections detailed simulation results for the Cases 1, 2, and 3 separately. In this

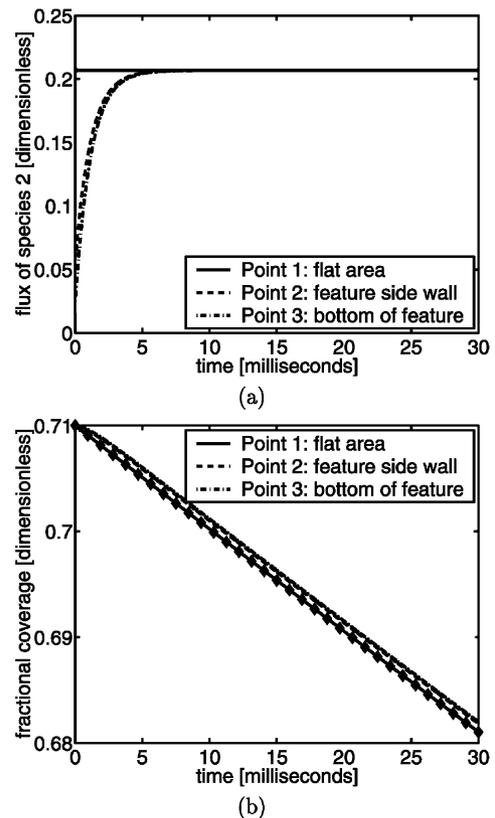


Fig. 16. Case 3, reaction step: (a) dimensionless flux to the surface of species B; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the vertical axes.

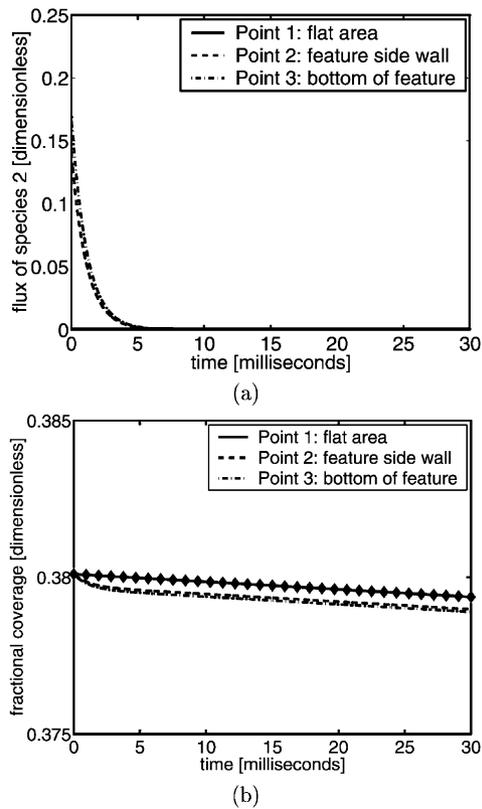


Fig. 17. Case 3, post-reaction purge step: (a) dimensionless flux to the surface of species B; (b) fractional coverage. The solid diamonds in (b) show the analytic solution. Note the different scales on the x_1 - and the x_2 -axes.

section, we compare and contrast the three cases. In particular, we wish to arrive at an estimate for the fraction of a monolayer deposited in one ALD cycle for each of the three chemistries considered. This is set in the context of possible changes of pulse durations for the four steps of one ALD cycle. One quantity on which to base this estimate is the fractional surface coverage ϑ_A . Therefore, Figs. 18–20 show the evolution of the fractional coverage ϑ_A vs. process time, subplots (a) for one ALD cycle and subplots (b) for four ALD cycles. In each case, the vertical lines denote the switch times between process steps. Simulation results are shown as thick solid lines and only cover the first 30 ms of each process step for which simulation is used. But Gobbert et al. [7] show that the analytic solution in Eq. (7) using the steady state flux values observed in the simulations is a good predictor of the fractional coverage throughout the entire process step. Hence, the analytic solution is used to complete the pictures given by the simulation results; it is also used to extend the results to four cycles in the subplots (b). In all cases, the final value of ϑ_A in one step is used as initial value in the following step.

Adsorbed molecules of A can be removed from the surface due to desorption or due to reaction with B. The

number of monolayers deposited in one cycle is obtained by integrating the deposition reaction rate from the beginning to the end of the reaction step. For Case 1, we find the fraction of a monolayer deposited in a cycle to be 0.68, the value for Case 2 is 0.71, and the value for Case 3 is 0.31. For Cases 1 and 2, these estimates remain the same for the cycles following the first one. For Case 3, the fraction of monolayer deposited per cycle increases to 0.37 in subsequent cycles. This is due to the larger fractional coverage of A present at the start of the adsorption step in the later cycles. The coverage

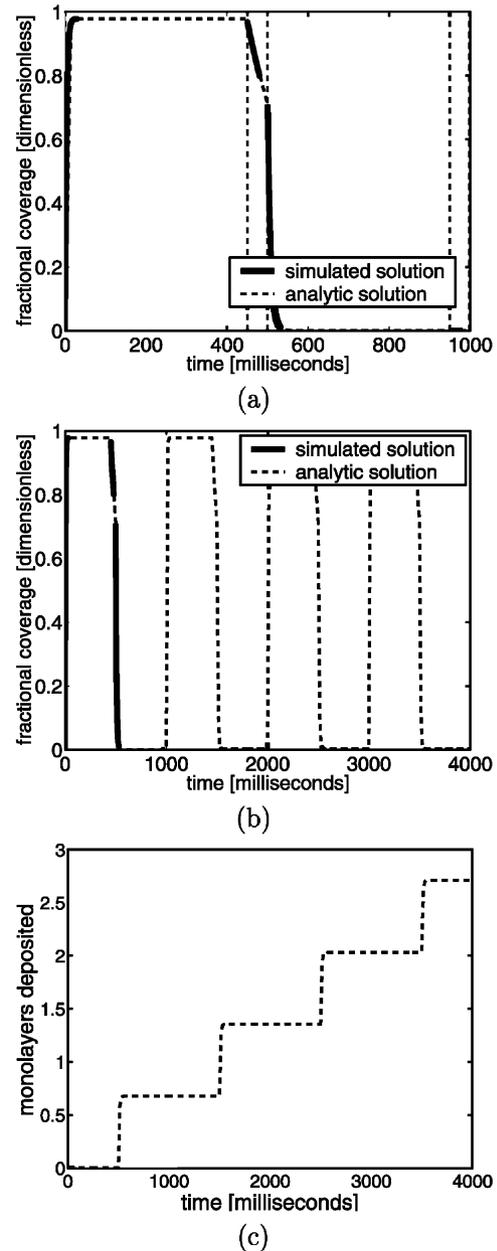


Fig. 18. Case 1: (a) fractional surface coverage for one ALD cycle; (b) fractional surface coverage for four ALD cycles; (c) monolayers deposited for four ALD cycles. The thick solid line segments show simulation results, the dashed lines plot the analytic solution.

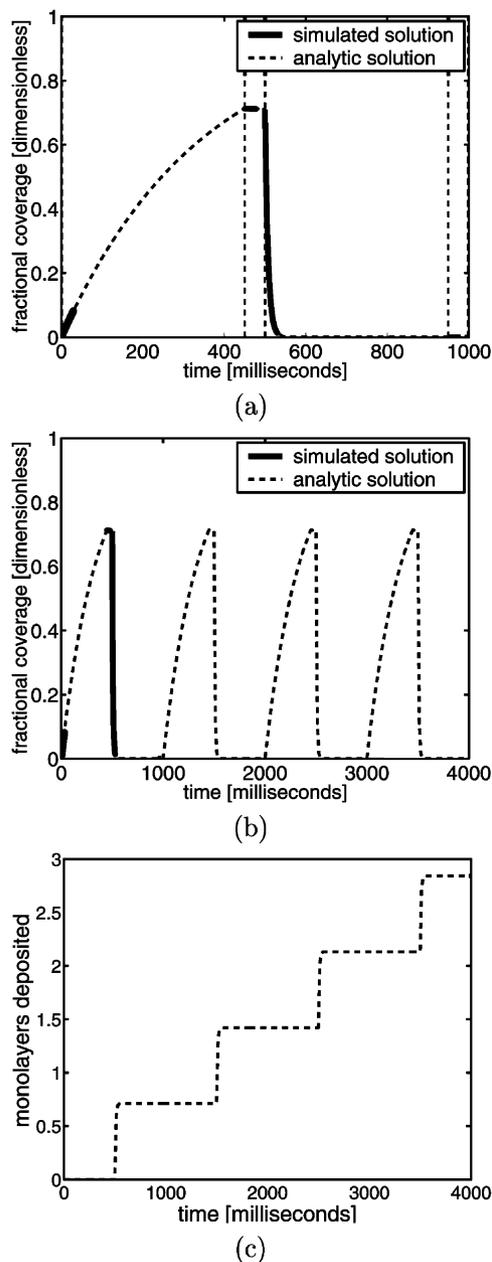


Fig. 19. Case 2: (a) fractional surface coverage for one ALD cycle; (b) fractional surface coverage for four ALD cycles; (c) monolayers deposited for four ALD cycles. The thick solid line segments show simulation results, the dashed lines plot the analytic solution.

of A did not reach its saturation value at the end of the adsorption step in the first cycle. To provide a picture of deposition vs. process time, Fig. 18c, 19c, and 20c show predictions for monolayers deposited using the simulation and analytically determined surface coverages in Fig. 18b, 19b, and 20b, respectively.

The first two cases provide higher fractions of a monolayer deposited per ALD cycle than the third case. Additionally, both Cases 1 and 2 allow for a reduction of the process time during the reaction step to 100 ms

without loss of deposition effectiveness. Case 1 additionally allows for a reduction in time for the adsorption step. These conclusions demonstrate that on balance the best chemistry is one with high adsorption as well as high reaction rates. However, in experimental studies, fractions of the order of or less than a third of a monolayer have been observed to be deposited in one ALD cycle [1,21,22]. The values we have used for the rate coefficients in Cases 1 and 2 might be higher than is realized in those experiments. In that case, Case 3

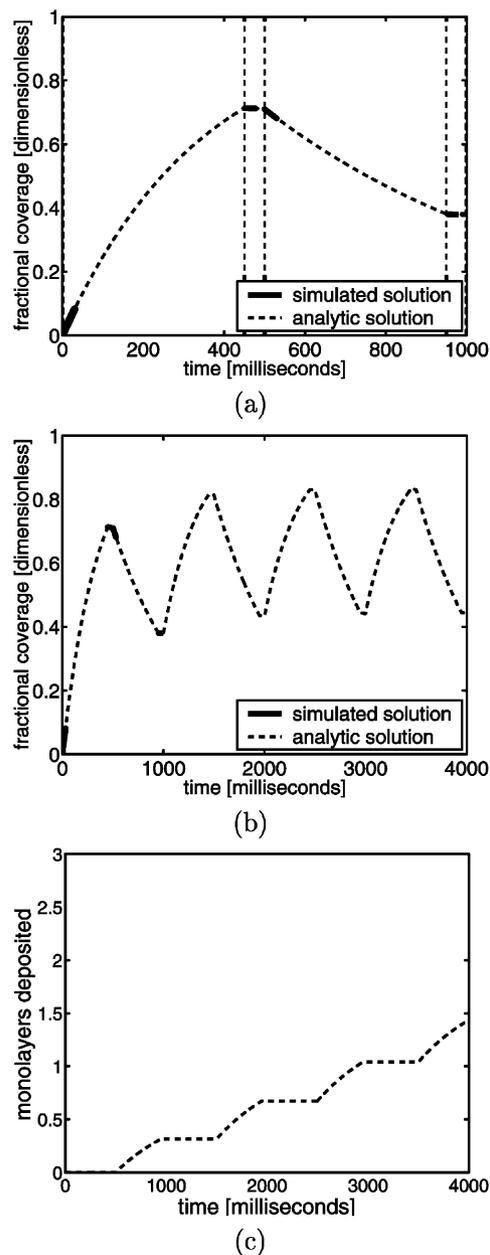


Fig. 20. Case 3: (a) fractional surface coverage for one ALD cycle; (b) fractional surface coverage for four ALD cycles; (c) monolayers deposited for four ALD cycles. The thick solid line segments show simulation results, the dashed lines plot the analytic solution.

may apply and the process times cannot be reduced without losing crucial adsorption and reaction times. Of course, this depends on the actual chemistry. Other considerations, such as steric hinderance, may also decrease the fraction of a monolayer deposited per ALD cycle.

4. Conclusions

A feature scale simulator for atomic layer deposition (ALD) is presented that combines a Boltzmann equation transport model with chemistry models. A simple but instructive chemistry is considered; one reactant species adsorbs onto the surface, and a second reactant reacts with it from the gas phase (Eley–Rideal). This work includes potential desorption of the adsorbed species during purge steps, which may or may not play a role in any given ALD system. Three sets (cases) of rate parameters are chosen to compare chemical rates with transport rates. The duration of the ALD pulses and the geometry of the representative feature are the same for each case. The particular choices of the adsorption, desorption, and reaction coefficients in the chemistry model are justified by prior detailed studies for the steps in the ALD cycle [7]. Simulation results are presented for all four steps in one ALD cycle; adsorption, post-adsorption purge, reaction and post-reaction purge. The results are extended to multiple ALD cycles, and the monolayers per cycle are estimated. We highlight a trade-off between pulse durations and deposition rate (wafer throughput) e.g. the increase in time required to increase the amount adsorbed during the adsorption step. The simulation methodology we present can be used to determine the pulse durations that maximize throughput for a given chemistry and chemical rate parameters.

Acknowledgments

Prof. Gobbert acknowledges support by the NSF under grant DMS-9805547. The RPI authors acknowledge support from MARCO, DARPA and NYSTAR.

References

- [1] M. Ritala, M. Leskela, E. Rauhala, J. Jokinen, *J. Electrochem. Soc.* 145 (1998) 2914.
- [2] M. Ritala, M. Leskela, E. Rauhala, *Chem. Mater.* 6 (1994) 556.
- [3] M. Ritala, M. Leskela, *Nanotechnology* 10 (1999) 19.
- [4] M. Leskela, M. Ritala, *J. de Physique, Colloque C5 (5)* (1995) 937.
- [5] T. Suntola, *Thin Solid Films* 216 (1992) 84.
- [6] H. Siimon, J. Aarik, *J. Phys. D: Appl. Phys.* 30 (1997) 1725.
- [7] M.K. Gobbert, S.G. Webster, T.S. Cale, *J. Electrochem. Soc.* (in press).
- [8] M.K. Gobbert, C.A. Ringhofer, T.S. Cale, *J. Electrochem. Soc.* 143 (1996) 2624.
- [9] T.P. Merchant, M.K. Gobbert, T.S. Cale, L.J. Borucki, *Thin Solid Films* 365 (2000) 368.
- [10] S.T. Rodgers, K.F. Jensen, *J. Appl. Phys.* 83 (1998) 524.
- [11] T.S. Cale, M.O. Bloomfield, D.F. Richards, K.E. Jansen, J.A. Tichy, M.K. Gobbert, In: I. Gamba, D. Levermore, C. Ringhofer (Eds.), *IMA Volumes in Mathematics and its Applications*, 2002 (in press).
- [12] M. O. Bloomfield, K.E. Jansen, T.S. Cale, In: G.S. Mathad, M. Yang, M. Engelhardt, H.S. Rathore, B.C. Baker, R.L. Opila (Eds.), *Thin Film Materials, Processes, and Reliability in Microelectronics*, PV 2001-24 (in press).
- [13] H. Siimon, J. Aarik, *J. de Physique, Colloque C5 (5)* (1995) 245.
- [14] J. Aarik, H. Siimon, *Appl. Surf. Sci.* 81 (1994) 281.
- [15] C. Cercignani, *The Boltzmann Equation and Its Applications*, Applied Mathematical Sciences, 67., Springer-Verlag, 1988.
- [16] C. Cercignani, *Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2000.
- [17] G.N. Patterson, *Introduction to the Kinetic Theory of Gas Flows*, University of Toronto Press, 1971.
- [18] B. Cockburn, G.E. Karniadakis, C.-W. Shu (Eds.), *Discontinuous Galerkin Methods: Theory, Computation and Applications*, Lecture Notes in Computational Science and Engineering, vol. 11, Springer-Verlag, 2000.
- [19] R.I. Masel, *Principles of Adsorption and Reaction on Solid Surfaces*, Wiley Interscience, 1996.
- [20] M.K. Gobbert, V. Prasad, T.S. Cale, In T. Wade (Ed.), *Proceedings of the Eighteenth International VLSI Multilevel Interconnection Conference (VMIC)*, Santa Clara, CA, November 27–30, 2001, p. 413.
- [21] M. Ritala, M. Leskela, E. Rauhala, J. Jokinen, *J. Electrochem. Soc.* 142 (1995) 2731.
- [22] M. Mashita, M. Sasaki, Y. Kawakyu, H. Ishikawa, *J. Crystal Growth* 131 (1993) 61.