Modeling and simulation of atomic layer deposition at the feature scale

Matthias K. Gobbert
Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland 21250

Vinay Prasad[a] and Timothy S. Cale
Focus Center—New York, Rensselaer: Interconnections for Gigascale Integration, Rensselaer Polytechnic Institute, CII 6015, Troy, New York 12180-3590

(Received 4 September 2001; accepted 1 April 2002)

We present a transient Boltzmann equation based transport and reaction model for atomic layer deposition (ALD) at the feature scale. The transport model has no adjustable parameters. In this article, we focus on the reaction step and the postreaction purge steps of ALD. The heterogeneous chemistry model consists of reversible adsorption of a reactant on a single site, and irreversible reaction of a second gaseous reactant with the adsorbed reactant. We conduct studies on the effect of the kinetic rate parameter associated with the reaction. We provide results for number densities of gaseous species, fluxes to the surface of the feature, and surface coverage of the adsorbing reactant as functions of time. For reasonable reaction rate parameter values, the time scale for gas transport is much smaller than that for reaction and desorption. For these cases, an analytic expression for the time evolution of the surface coverage of the adsorbing reactant provides a good approximation to the solution obtained from the transport and reaction model. The results show that fractional coverage of the adsorbing reactant reduces significantly in the reaction step due to reaction with the gaseous reactant and desorption. Larger values of the reaction rate parameter lead to larger reductions in the fractional coverage during the reaction step. For smaller values of the reaction rate parameter, the decrease in coverage is dominated by desorption. The surface coverage of the adsorbing reactant also decreases during purge steps, due to desorption. [DOI: 10.1116/1.1481754]

I. INTRODUCTION

Atomic layer deposition (ALD) has been shown to provide excellent film thickness uniformity over severe topography, in addition to the ability to control film thickness.[1–4] The uniformity of deposition and control over thickness can be attributed to the self-limiting mechanism of film growth, with the film being (ideally) deposited one monolayer at a time. Because of this, ALD is a viable technique to deposit ultrathin films for diffusion barriers and conformal films in high aspect ratio features found in modern integrated circuit fabrication.[5]

ALD involves pulsing reactant gases over a substrate in series, with purges of an inert gas being employed 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, with surface sites for adsorption of **A** being made available as the reaction proceeds. The reactor is then purged again, and the next ALD cycle is started with a fresh pulse of **A**. The duration of each pulse of the ALD cycle is adjusted to provide a high rate of deposition while maintaining the properties of conformality and control over the rate of deposition that make ALD attractive.

A complete ALD model would account for the transients caused by the sequence of reactant pulses and purges, at all spatial scales in the reactor. In this article, the flux from the reactor volume to each feature is assumed to be constant in time, except for step changes that represent idealized changes from one step in the ALD cycle to another. Note that most of the relevant literature on feature scale modeling in ALD consists of descriptions of the surface processes without a gas phase transport model.[6–9] The dominant approach to feature scale transport and reaction analyses was developed to model topography evolution during conventional steady-state deposition and etch processes. These models are pseudosteady; i.e., the local surface reaction rates are computed assuming fluxes are constant in time[10,11] and are not appropriate for ALD.

We use a Boltzmann equation based gas transport model along with heterogeneous chemical reaction mechanisms[12] to study ALD. This formulation allows us to study the transients of both gas phase transport and surface reaction that are inherent to ALD on the feature scale. For discussions of how to integrate reactor scale and feature scale models, see Refs. 13–16. For extensions to transient integrated multiscale modeling, see Ref. 17. For a discussion of transients that are present on the time scale of processes, see Ref. 18.

The following section describes the transport and reaction models used. After briefly explaining the numerical method

[a]Author to whom correspondence should be addressed; electronic mail: prasav@rpi.edu

in the following section, we report our simulation results for the reaction and postreaction stages of the ALD process. Finally, we present results on surface coverage of the adsorbing reactant over a complete ALD cycle.

II. MODEL

A. Domain

The domain Ω of the feature scale model includes the interior area of one feature and a small part of the gas domain above and around the feature mouth; a schematic of a two-dimensional domain chosen as a cross section of a typical feature is shown in Fig. 1(a). The differential equation needs to be accompanied by boundary conditions along $\partial\Omega$, which is comprised of three parts with different boundary conditions: $\partial\Omega=\Gamma_w\cup\Gamma_t\cup\Gamma_s$. Here, Γ_w denotes the portion of the boundary along the solid wafer surface, Γ_t is the top of the domain that forms the interface to the bulk of the reactor, and Γ_s denotes the union of the portions of the boundary on the sides of the domain.

B. Differential equations in the gaseous domain

The flow of a rarefied gas is described by the Boltzmann equation[19–21] for each gaseous species:

$$\frac{\partial f^{(i)}}{\partial t}+v\cdot\nabla_x f^{(i)}=\sum_{j=0}^{2} Q_{ij}(f^{(i)},f^{(j)}),\quad i=0,1,2. \tag{1}$$

The unknown variables are the density distribution functions $f^{(i)}(x,v,t)$, i.e., the number of molecules of species i at position $x=(x_1,x_2,x_3)\in\Omega\subset\mathbb{R}^3$ with velocity $v=(v_1,v_2,v_3)\in\mathbb{R}^3$ at time $t\geqslant 0$. The $f^{(i)}$ have to be determined for all points x in the domain $\Omega\subset\mathbb{R}^3$ and for all possible velocities $v\in\mathbb{R}^3$. The distribution functions are scaled such that

$$c_i(x,t):=\int_{\mathbb{R}^3} f^{(i)}(x,v,t)\,dv,\quad i=0,1,2 \tag{2}$$

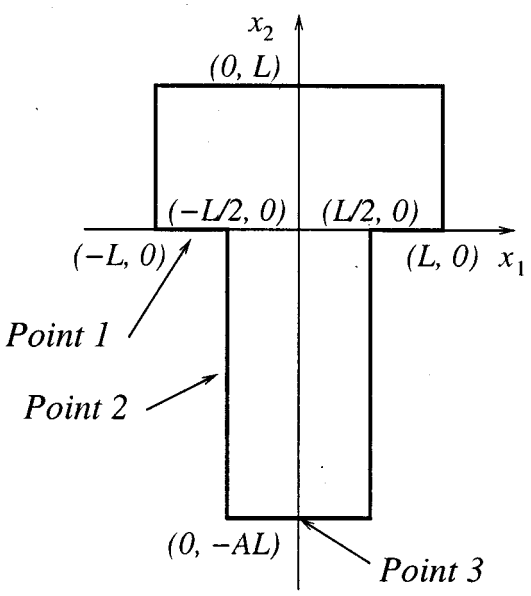
gives the molar concentration of species i at $x\in\Omega$ at time t. As written, Eq. (1) is appropriate for ALD with one inert background species **I** with index $i=0$ and reactive species **A** and **B** with indices $i=1$ and $i=2$, respectively. The left-hand side describes the transport of species i. The right-hand side describes the effect of collisions among molecules of all species, in which the collision operators Q_{ij} model the collisions between molecules of species i and j. The following paragraphs show how we treat collisional transport of reactive species in a background gas. This derivation is important to arrive at the appropriate dimensionless formulation of the Boltzmann equation for free molecular flow.

Assuming that the reactive species $j=1,2$ are at least an order of magnitude less concentrated than the background gas 0, it can be shown that it is justified to keep only the collision operators Q_{i0} and neglect Q_{i1} and Q_{i2} in every equation $i=0,1,2$. If we also assume that the background gas is uniformly distributed in space ($\nabla_x f^{(0)}=0$), at equilibrium ($\partial f^{(0)}/\partial t=0$), and inert (does not react with the species $j=1,2$), then the equation for $f^{(0)}$ is decoupled from the remaining ones for the reactive species and consists in fact of $Q_{00}(f^{(0)},f^{(0)})=0$ only, which has as a solution a Maxwellian:[19,20]

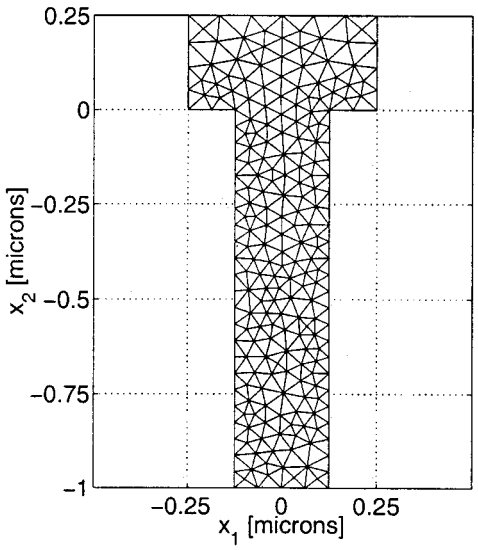
$$f^{(0)}(x,v,t)=M_0^{\mathrm{ref}}(v):=\frac{c_0^{\mathrm{ref}}}{[2\pi(v_0^\infty)^2]^{3/2}}\exp\left(-\frac{|v|^2}{2(v_0^\infty)^2}\right), \tag{3}$$

where c_0^{ref} and v_0^∞ denote a reference concentration and the thermodynamic average speed, respectively.

Using the explicit solution for the background species, we solve the linear Boltzmann equation for the reactive species



(a)



(b)

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 $A=4$.

TABLE I. Physical constants, operating conditions, and species reference quantities.

Physical constants	
Universal gas constant	$R_g = 8.3145$ J/(K mol) $= 62\,400$ (cm^3 Torr)/(K mol)
Universal Boltzmann constant	$k_B = 1.3807\times10^{-23}$ J/K
Avogadro's number	$N_A = 6.0221\times10^{23}$/mol
Operating conditions	
Ambient temperature	$T = 500$ K
Total pressure	$P_{\text{total}} = 1$ Torr
Reference quantities for reactive species A ($i=1$)	
Mole fraction	$x_1 = 0.10$
Partial pressure	$P_1 = 0.10$ Torr
Reference concentration	$c_1^{\text{ref}} = 3.2\times10^{-9}$ mol/cm^3
Molecular weight	$\omega_1 = 104$ g/mol
Thermal average speed	$v_1^{\infty} = 2.0\times10^4$ cm/s
Reference quantities for reactive species B ($i=2$)	
Mole fraction	$x_2 = 0.05$
Partial pressure	$P_2 = 0.05$ Torr
Reference concentration	$c_2^{\text{ref}} = 1.6\times10^{-9}$ mol/cm^3
Molecular weight	$\omega_2 = 128$ g/mol
Thermal average speed	$v_2^{\infty} = 1.8\times10^4$ cm/s

$$\frac{\partial f^{(i)}}{\partial t} + v\cdot\nabla_x f^{(i)} = Q_i(f^{(i)}), \quad i=1,2 \tag{4}$$

with the linear collision operator $Q_i(f^{(i)}) := Q_{i0}(f^{(i)}, M_0^{\text{ref}})$. Notice that while these equations are decoupled, the solutions for the reacting species are coupled to each other through the boundary condition at the wafer surface that models the surface reactions; decoupling from the background gas relies materially on the assumption that it is an inert gas.

Define reference Maxwellians also for the reactive species by

$$M_i^{\text{ref}}(v) = \frac{c_i^{\text{ref}}}{[2\pi(v_i^{\infty})^2]^{3/2}} \exp\left(-\frac{|v|^2}{2(v_i^{\infty})^2}\right), \quad i=1,2, \tag{5}$$

where c_i^{ref} and v_i^{∞} denote again reference concentrations and the thermodynamic average speeds for the species. The reference concentrations are chosen from the ideal gas law as

$$c_i^{\text{ref}} = \frac{P_i}{R_g T}, \quad i=1,2, \tag{6}$$

where the partial pressure of species i is given by $P_i = x_i P_{\text{total}}$ based on the chosen reference mole fraction x_i, and R_g denotes the universal gas constant; see Table I. The temperature T in this article is the constant and spatially uniform temperature in Table I. The thermal average speeds, which are used in the Maxwellians, are given by

$$v_i^{\infty} = \sqrt{R_i T} = \sqrt{\frac{k_B}{m_i}T} = \sqrt{\frac{R_g}{\omega_i}T}, \quad i=1,2, \tag{7}$$

TABLE II. Reference quantities.

For gaseous species	
Reference concentration	$c^* := c_1^{\text{ref}} = 3.2\times10^{-9}$ mol/cm^3
Reference speed	$v^* := v_1^{\infty} = 2.0\times10^4$ cm/s
For transport	
Referernce length	$L^* = 1\ \mu\text{m} = 10^{-4}$ cm
Reference time for transport	$t^* = L^*/v^* = 5\times10^{-9}$ s $= 5$ ns
For collisions	
Mean free path	$\lambda = 100\ \mu\text{m} = 10^{-2}$ cm
Reference time for collisions	$\tau^* = \lambda/v^* = 5\times10^{-7}$ s
For reactions	
(Formal) reference flux	$\eta^* := c^* v^* = 6.4\times10^{-5}$ mol/(s cm^2)
Total concentration of surface sites	$S_T = 10^{-9}$ mol/cm^2

based on the molecular weights ω_i. The universal gas constant R_g and the universal Boltzmann constant k_B are related through Avogadro's number N_A by $R_g = N_A k_B$; see Table I. Writing $(v_i^{\infty})^2 = R_i T$ results in another common representation of the Maxwellians

$$M_i^{\text{ref}}(v) = \frac{c_i^{\text{ref}}}{[2\pi R_i T]^{3/2}} \exp\left(-\frac{|v|^2}{2R_i T}\right). \tag{8}$$

Note that the Maxwellians are designed to have the same units as the density functions $f^{(i)}$.

The reference quantities for the nondimensionalization procedure are listed in Table II. The reference concentration c^* and reference speed v^* are chosen equal to the corresponding quantities for the first reactive species. After defining the reference length appropriate for the domain size as $L^* = 1\ \mu$m, we obtain on the one hand the reference time for transport as $t^* = L^*/v^*$. The mean free path λ is about 100 μm at the operating conditions listed in Table I and determines, on the other hand, a reference time for collisions (the mean collision time) by $\tau^* = \lambda/v^*$. The ratio of those times or lengths is equal to the Knudsen number $\text{Kn} = \lambda/L^* = \tau^*/t^*$.

TABLE III. Dimensionless variables.

Time	$\hat{t} = \dfrac{t}{t^*}$
Lengths	$\hat{x} = \dfrac{x}{L^*}$
Velocities	$\hat{v} = \dfrac{v}{v^*}, \quad \hat{v}_i^{\infty} = \dfrac{v_i^{\infty}}{v^*}$
Concentrations	$\hat{c}_i = \dfrac{c_i}{c^*}, \quad \hat{c}_i^{\text{ref}} = \dfrac{c_i^{\text{ref}}}{c^*}, \quad \hat{c}_i^{\text{top}} = \dfrac{c_i^{\text{top}}}{c^*}, \quad \hat{c}_i^{\text{ini}} = \dfrac{c_i^{\text{ini}}}{c^*}$
Density distributions	$\hat{f}^{(i)} = \dfrac{(v^*)^3}{c^*} f^{(i)}, \quad \hat{f}_i^{\text{top}} = \dfrac{(v^*)^3}{c^*} f_i^{\text{top}}, \quad \hat{f}_i^{\text{ini}} = \dfrac{(v^*)^3}{c^*} f_i^{\text{ini}}$
Maxwellians	$\hat{M}_i^{\text{ref}} = \dfrac{(v^*)^3}{c^*} M_i^{\text{ref}}$
Collision operators	$\hat{Q}_{ij} = \dfrac{(v^*)^3\tau^*}{c^*} Q_{ij}, \quad \hat{Q}_i \dfrac{(v^*)^3\tau^*}{c^*} Q_i$
Fluxes, reaction rates	$\hat{\eta}_i = \dfrac{\eta_i}{\eta^*}, \quad \hat{R}_k = \dfrac{R_k}{\eta^*}$
Fractional surface coverage	$\vartheta_A = \dfrac{S_A}{S_T}$

TABLE IV. Dimensionless groups.

For species **A** ($i=1$)	
Dimensionless reference concentration	$\hat{c}_1^{\text{ref}}=1.0$
Dimensionless reference speed	$\hat{v}_1^{\infty}=1.0$
For species **B** ($i=2$)	
Dimensionless reference concentration	$\hat{c}_2^{\text{ref}}=0.5$
Dimensionless reference speed	$\hat{v}_2^{\infty}=0.9$
For transport and collisions	
Knudsen number	$\text{Kn}:=\lambda/L^*=100$
For reactions	
Reaction coefficients for reaction 1	$\gamma_1^f=S_T k_1^f, \quad \gamma_1^b=\frac{S_T}{\eta^*}k_1^b$
Reaction coefficient for reaction 2	$\gamma_2^f=S_T k_2^f$
Prefactor	$\alpha_p=\frac{\eta^* t^*}{S_T}=0.32\times 10^{-3}$

The choices of dimensionless variables are listed in Table III. They result in the dimensionless Maxwellians

$$\hat{M}_i^{\text{ref}}(\hat{v})=\frac{\hat{c}_i^{\text{ref}}}{[2\pi(\hat{v}_i^{\infty})^2]^{3/2}}\exp\left(-\frac{|\hat{v}|^2}{2(\hat{v}_i^{\infty})^2}\right), \quad i=1,2, \tag{9}$$

where the dimensionless groups $\hat{c}_i^{\text{ref}}$ and $\hat{v}_i^{\infty}$ are included in Table IV. The dimensionless Boltzmann equation is obtained by introducing the dimensionless variables listed in Table III. Notice that the left-hand side is nondimensionalized with respect to transport, while the right-hand side is nondimensionalized with respect to collisions. This results in the Knudsen number appearing in the dimensionless Boltzmann equations for the reactive species

$$\frac{\partial \hat{f}^{(i)}}{\partial \hat{t}}+\hat{v}\cdot\nabla_{\hat{x}}\hat{f}^{(i)}=\frac{1}{\text{Kn}}\hat{Q}_i(\hat{f}^{(i)}), \quad i=1,2. \tag{10}$$

Since the Knudsen number Kn for gaseous flow on the feature scale is large, $\text{Kn}\gg 1$ (see Table IV), we obtain the equations of free molecular flow

$$\frac{\partial \hat{f}^{(i)}}{\partial \hat{t}}+\hat{v}\cdot\nabla_{\hat{x}}\hat{f}^{(i)}=0, \quad i=1,2. \tag{11}$$

Recall again that the reacting species will be coupled through the surface reacting entering through the boundary conditions at the wafer surface.

C. Surface reaction model

The surface chemistry model consists of reversible adsorption of **A** on a single site, and irreversible reaction of **B** with the adsorbed **A** (Ref. 12)

$$\begin{aligned}&\mathbf{A}+v\rightleftharpoons\mathbf{A}_v,\\ &\mathbf{A}_v+\mathbf{B}\rightarrow v+(*),\end{aligned} \tag{12}$$

where $\mathbf{A}_v$ is adsorbed **A**, v stands for a surface site available for adsorption, and $(*)$ is the nonadsorbing gaseous product. Notice that surface sites available for adsorption are produced by the reaction of **B** with adsorbed **A**.

The total molar concentration of surface sites available for deposition is denoted by S_T, see Table II. If S_A denotes the concentration of adsorbed molecules of **A**, the difference S_T-S_A is the concentration of vacant sites, and the reaction rates can be written as

$$\begin{aligned}R_1&=k_1^f(S_T-S_A)\eta_1-k_1^b S_A,\\ R_2&=k_2^f S_A\eta_2,\end{aligned} \tag{13}$$

where η_i denotes the flux of species i to the surface, which is related to the distribution function of Eq. (1) by

$$\begin{aligned}&\eta_i(x,t)=\int_{\nu\cdot v'>0}|\nu\cdot v'|f^{(i)}(x,v',t)dv',\\ &i=1,2, \quad x\in\Gamma_w.\end{aligned} \tag{14}$$

Here, Γ_w denotes the points at the wafer surface and $\nu\equiv\nu(x)$ is the unit outward normal vector at $x\in\Gamma_w$. Notice that the integral is over all velocities pointing out of the domain due to the condition $\nu\cdot v'>0$. The evolution of the concentration of sites occupied by **A** at every point x at the wafer surface Γ_w is given by

$$\frac{dS_A(x,t)}{dt}=R_1(x,t)-R_2(x,t), \quad x\in\Gamma_w. \tag{15}$$

Notice that this model assumes that there is no significant movement of molecules along the surface.

If we nondimensionalize the reaction rates with respect to the reference flux η^* and introduce the fractional surface coverage $\vartheta_A=S_A/S_T\in[0,1]$, we obtain the dimensionless reaction rates

$$\begin{aligned}\hat{R}_1&=\gamma_1^f(1-\vartheta_A)\hat{\eta}_1-\gamma_1^b\vartheta_A,\\ \hat{R}_2&=\gamma_2^f\vartheta_A\hat{\eta}_2\end{aligned} \tag{16}$$

with the dimensionless coefficients given in Table IV. Making the differential equation for S_A dimensionless, we obtain

$$\frac{d\vartheta_A(\hat{x},\hat{t})}{d\hat{t}}=\alpha_p[\hat{R}_1(\hat{x},\hat{t})-\hat{R}_2(\hat{x},\hat{t})], \quad \hat{x}\in\Gamma_w, \tag{17}$$

with the prefactor $\alpha_p=(\eta^* t^*)/S_T$. This differential equation is supplied with an initial condition that represents the fractional coverage ϑ_A^{ini} at the initial time, which is assumed known.

Remark: It is in general impossible to find a closed-form solution $\vartheta_A(\hat{t})$ to the differential equation [Eq. (17)], because the coefficients involving $\hat{\eta}_1$ and $\hat{\eta}_2$ are not constant. But if these fluxes are constant, then Eq. (17) 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

$$\frac{d\vartheta_A(\hat{t})}{d\hat{t}}=-\alpha_p b\,\vartheta_A(\hat{t})+\alpha_p\gamma_1^f\hat{\eta}_1, \quad \vartheta_A(0)=\vartheta_A^{\text{ini}} \tag{18}$$

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

$$\vartheta_A(\hat{t})=\vartheta_A^{\infty}(1-e^{-\alpha_p b\hat{t}})+\vartheta_A^{\text{ini}}e^{-\alpha_p b\hat{t}} \tag{19}$$

with the equilibrium limit

$$\vartheta_A^\infty = \frac{\gamma_1^f \hat{\eta}_1}{\gamma_1^f \hat{\eta}_1 + \gamma_1^b + \gamma_2^f \hat{\eta}_2}, \tag{20}$$

provided that $\hat{\eta}_1$ and $\hat{\eta}_2$ are constant. Clearly, $\vartheta_A(\hat{t}) \to \vartheta_A^\infty$ as $\hat{t} \to \infty$, hence the name for the constant ϑ_A^∞. Note that this assumes that the species fluxes from the source above the wafer are constant.

D. Boundary conditions for the Boltzmann equation

At the wafer surface Γ_w, we use the boundary conditions

$$\begin{aligned} f^{(1)}(x,v,t) &= [\eta_1(x,t) - R_1(x,t)] C_1(x) M_1^{\text{ref}}(v), \\ &\quad \nu \cdot v < 0, \quad x \in \Gamma_w, \\ f^{(2)}(x,v,t) &= [\eta_2(x,t) - R_2(x,t)] C_2(x) M_2^{\text{ref}}(v), \\ &\quad \nu \cdot v < 0, \quad x \in \Gamma_w, \end{aligned} \tag{21}$$

where η_i is the flux of species i to the surface and R_k is the reaction rate of reaction k. The boundary condition assumes diffusive emission of molecules, i.e., with the same velocity distribution as the reference Maxwellian.[19,21] In the absence of reactions ($R_k = 0$), the inflowing part of $f^{(i)}$ is then proportional to the flux to the surface η_i, because all molecules are being re-emitted. In the presence of reactions thought, the rates of re-emission differ from the incoming flux by the reaction rates, which could have either sign.

The factors C_i are chosen as

$$C_i(x) = \left(\int_{\nu \cdot v < 0} |\nu \cdot v| M_i^{\text{ref}}(v) dv \right)^{-1}, \quad i = 1,2, \tag{22}$$

to guarantee mass conservation in the absence of reactions, i.e., we require that influx equal to outflux for each species for $R_k = 0$:

$$\int_{\nu \cdot v < 0} |\nu \cdot v| f^{(i)}(x,v,t) dv = \int_{\nu \cdot v > 0} |\nu \cdot v| f^{(i)}(x,v,t) dv. \tag{23}$$

Notice that $C_i(x)$ depends on the position $x \in \Gamma_w$ via the unit outward normal vector $\nu(x)$.

Using the reference flux η^*, formally chosen as $\eta^* = c^* v^*$ (see Table II), the dimensionless boundary conditions attain the same form as the dimensional ones as

$$\begin{aligned} \hat{f}^{(1)}(\hat{x},\hat{v},\hat{t}) &= [\hat{\eta}_1(\hat{x},\hat{t}) - \hat{R}_1(\hat{x},\hat{t})] \hat{C}_1(\hat{x}) \hat{M}_1^{\text{ref}}(\hat{v}), \\ &\quad \nu \cdot \hat{v} < 0, \quad \hat{x} \in \Gamma_w, \\ \hat{f}^{(2)}(\hat{x},\hat{v},\hat{t}) &= [\hat{\eta}_2(\hat{x},\hat{t}) - \hat{R}_2(\hat{x},\hat{t})] \hat{C}_2(\hat{x}) \hat{M}_2^{\text{ref}}(\hat{v}), \\ &\quad \nu \cdot \hat{v} < 0, \quad \hat{x} \in \Gamma_w, \end{aligned} \tag{24}$$

with the dimensionless fluxes to the surface

$$\hat{\eta}_i(\hat{x},\hat{t}) = \int_{\nu \cdot \hat{v}' > 0} |\nu \cdot \hat{v}'| \hat{f}^{(i)}(\hat{x},\hat{v}',\hat{t}) d\hat{v}', \quad i = 1,2 \tag{25}$$

and with

$$\hat{C}_i(\hat{x}) = \left(\int_{\nu \cdot \hat{v} < 0} |\nu \cdot \hat{v}| \hat{M}_i^{\text{ref}}(\hat{v}) d\hat{v} \right)^{-1}, \quad i = 1,2. \tag{26}$$

The top of the domain of the feature scale model Γ_t forms the interface to the bulk of the gas domain in the reactor, and we assume that the distribution of $f^{(i)}$ is known there. More precisely, we assume that the inflow has a Maxwellian velocity distribution, hence

$$\begin{aligned} f^{(i)}(x,v,t) &= f_i^{\text{top}} := \frac{c_i^{\text{top}}}{[2\pi (v_i^\infty)^2]^{3/2}} \exp\left(-\frac{|v|^2}{2(v_i^\infty)^2} \right), \\ &\quad i = 1,2, \quad \nu \cdot v < 0, \ x \in \Gamma_t. \end{aligned} \tag{27}$$

Using the dimensionless variables in Table III results in the dimensionless boundary condition

$$\begin{aligned} \hat{f}^{(i)}(\hat{x},\hat{v},\hat{t}) &= \hat{f}_i^{\text{top}} = \frac{\hat{c}_i^{\text{top}}}{[2\pi (\hat{v}_i^\infty)^2]^{3/2}} \exp\left(-\frac{|\hat{v}|^2}{2(\hat{v}_i^\infty)^2} \right), \\ &\quad i = 1,2, \quad \nu \cdot \hat{v} < 0, \quad \hat{x} \in \Gamma_t. \end{aligned} \tag{28}$$

On the sides of the domain Γ_s, which are perpendicular to the mean wafer surface, we use specular reflection for the boundary condition to simulate an infinite domain. This condition can immediately be stated in dimensionless form as

$$\hat{f}^{(i)}(\hat{x},\hat{v},\hat{t}) = \hat{f}^{(i)}(\hat{x},\hat{v}',\hat{t}), \quad i = 1,2, \quad \nu \cdot \hat{v} < 0, \quad \hat{x} \in \Gamma_s \tag{29}$$

with

$$\hat{v}' = \hat{v} - 2\nu(\nu \cdot \hat{v}). \tag{30}$$

Finally, we assume that the initial distribution of gas is given by

$$\begin{aligned} f^{(i)}(x,v,t) &= f_i^{\text{ini}} := \frac{c_i^{\text{ini}}}{[2\pi (v_i^\infty)^2]^{3/2}} \exp\left(-\frac{|v|^2}{2(v_i^\infty)^2} \right), \\ &\quad i = 1,2, \quad x \in \Omega, \quad t = 0 \end{aligned} \tag{31}$$

with a Maxwellian velocity distribution; in particular, the choice of $c_i^{\text{ini}} = 0$ results in no gas of species i in the domain initially. The dimensionless initial condition is then

$$\begin{aligned} \hat{f}^{(i)}(\hat{x},\hat{v},\hat{t}) &= \hat{f}_i^{\text{ini}} = \frac{c_i^{\text{ini}}}{[2\pi (\hat{v}_i^\infty)^2]^{3/2}} \exp\left(-\frac{|\hat{v}|^2}{2(\hat{v}_i^\infty)^2} \right), \\ &\quad i = 1,2, \quad \hat{x} \in \Omega, \quad \hat{t} = 0. \end{aligned} \tag{32}$$

Remark: If the reference concentrations in the reference Maxwellians are chosen as $c_i^{\text{ref}} = c_i^{\text{top}}$ and there are no surface reactions, the reference Maxwellians will be the exact equilibrium solution of the model by construction.

III. NUMERICAL

This article reports on numerical results obtained in two dimensions, and the numerics are stated in two-dimensional form here; the generalization to three dimensions is straightforward, but considerably more computationally intense. To simplify notation, the hats (^) used to indicate dimensionless variables are omitted in this section.

The solutions $f^{(i)}(x,v,t)$ to the kinetic equation [Eq. (11)] together with the boundary conditions [Eqs. (24), (28), and (29)] and the initial condition [Eq. (32)] depend on $x \in \Omega \subset \mathbb{R}^2$, $v \in \mathbb{R}^2$, and $t \geq 0$. We approach the problem by expanding the unknowns $f^{(i)}$ for the reactive species in velocity space:

$$f^{(i)}(x,v,t) = \sum_{k=0}^{K-1} f_k^{(i)}(x,t)\varphi_k(v), \quad i=1,2, \tag{33}$$

where the $\varphi_k(v)$, $k=0,1,...,K-1$, form an orthogonal set of basis functions in velocity space with respect to some inner product $\langle \cdot,\cdot \rangle_C$, namely, $\langle \varphi_k, \varphi_k \rangle_C = q_k \neq 0$ for all k and $\langle \varphi_k, \varphi_l \rangle_C = 0$ for all $k \neq l$.

To obtain an equivalent system of equations for the vectors of coefficient functions

$$F^{(i)}(x,t) = \begin{pmatrix} f_0^{(i)}(x,t) \\ \vdots \\ f_{K-1}^{(i)}(x,t) \end{pmatrix}, \quad i=1,2, \tag{34}$$

the expansions in Eq. (33) are inserted into Eq. (11) and the equations tested against φ_k in scalar product $\langle \cdot,\cdot \rangle_C$. This Galerkin approach yields the systems of linear hyperbolic equations

$$\frac{\partial F^{(i)}}{\partial t} + A^{(1)} \frac{\partial F^{(i)}}{\partial x_1} + A^{(2)} \frac{\partial F^{(i)}}{\partial x_2} = 0, \quad i=1,2 \tag{35}$$

with matrices $A^{(1)}, A^{(2)} \in \mathbb{R}^{K \times K}$. Following ideas in Ref. 22, it is possible to make a judicious choice of basis functions to make these matrices diagonal. First mathematical results based on this approach can be found in Refs. 23 and 24. Notice again that while the systems of differential equations appear decoupled, their solutions are still related through the boundary condition at the wafer surface.

This system of linear hyperbolic equations is now posed in a standard form amenable for numerical computations. However, due to its large size, the irregular structure of the domain, and the requirement to compute for long times, it still poses a formidable challenge. It is solved using the discontinuous Galerkin method implemented in the code DG,[25] which is well-suited to the task. See Ref. 26 for more detailed information on the numerical method.

The demonstration results presented in this article are computed using four discrete velocities in each x_1 and x_2 direction; hence, there are $K=16$ equations for $i=1,2$ hence 32 equations total. For the results of the adsorption and post adsorption purge steps presented in Ref. 27, some results were checked against discretizations using six discrete velocities in each direction, and good agreement was found for each of these comparisons. The spatial domain was meshed rather coarsely to save on computation time. The mesh for the domain is shown in Fig. 1(b). As shown below, the coarse mesh and the value of K are sufficient to show that the time scale for transport is much faster than the time scale for adsorption and reaction for reasonable chemistries. In turn, this leads to a significantly simpler model that has the analytic solution of Eq. (19) for the surface fractions.

IV. RESULTS

In this section, we present simulation results for the reaction and postreaction purge steps of ALD, and show that Eq. (19) for the surface coverage of **A** is accurate for the conditions studied. We also show a predicted time evolution for surface coverage of **A** over one complete ALD cycle. For all the simulation results reported here, we choose the geometry displayed in Fig. 1(a) with the aspect ratio $A=4$. The feature mouth is assumed to be $L=0.25$ μm wide. Figure 1(b) shows the mesh used in the simulations. Results on the adsorption and postadsorption purge steps for the same geometry are reported in Ref. 27; the following subsections provide results for the reaction and the postreaction purge steps. We also provide results for surface coverage over one complete ALD cycle, which are computed using the analytical approximation of Eq. (19). The model is given by the dimensionless equations detailed in previous sections. Some parameter values are listed in Table IV. Additionally, we need to specify the (dimensionless) reaction parameters in Eq. (16). In Ref. 27, numerical simulations on the adsorption and desorption step were conducted for the values $\gamma_1^f = 0.0, 1.0, 10^{-2}$, and 10^{-4}. In the majority of the cases, the desorption coefficient was chosen relative to the adsorption coefficient as $\gamma_1^b = \gamma_1^f/100$. In this work, we choose the values $\gamma_1^f = 10^{-2}$ and $\gamma_1^b = 10^{-4}$, and provide simulation results for reaction rate coefficient values of $\gamma_2^f = 10^{-2}$, 10^{-4}, and 10^{-6}. We justify these values for dimensionless reaction rate constants in the subsection on the reaction step.

To complete the model, we need to choose the initial condition for the (dimensionless) gas concentration throughout the domain $\hat{c}_i^{\text{ini}}$ and for the fractional coverage ϑ_A^{ini}, as well as the coefficient in the boundary condition at the top of the domain $\hat{c}_i^{\text{top}}$. The values for $\hat{c}_i^{\text{ini}}$ and $\hat{c}_i^{\text{top}}$, which are used to specify the initial condition for the flux distributions of **A** and **B**, are different for each step of the pulse cycle and are specified in the following subsections. To generate appropriate initial conditions for surface coverage and number densities, we assume an ALD pulse cycle with adsorption and reaction steps of 450 ms duration each, and purges that are 50 ms long.

In order to analyze the behavior of the fluxes $\hat{\eta}_i$ to the surface and of the fractional surface coverage ϑ_A over time, three significant points on the wafer surface are chosen as shown in Fig. 1(a). Point 1 is located on the flat area of the wafer surface at $(-0.75L,0)$, point 2 is located halfway down the trench and has x_2 coordinate $-0.5AL$, and point 3 is located at the bottom of the feature at $(-AL,0)$. The fluxes $\hat{\eta}_i$ and the fractional coverage ϑ_A are presented as time evolutions at these representative locations. The choice of these locations is consistent with the locations used in Ref. 27.

A. Reaction step

At the beginning of the reaction step, we assume there are no molecules of species **A** present in the domain, hence $\hat{c}_i^{\text{ini}} = 0.0$. As shown in the section on the postreaction purge step, it is a good approximation to assume that the number densities of reactant species go to zero during purge steps.

No **A** is fed from the top, so $\hat{c}_1^{\text{top}}=0.0$. No molecules of species **B** are present initially in the domain, so $\hat{c}_2^{\text{ini}}=0.0$. We assume that **B** is fed into the domain during the reaction step with a Maxwellian velocity distribution and the full reference concentration, hence $\hat{c}_2^{\text{top}}=\hat{c}_2^{\text{ref}}=0.5$; see Table IV. For values of $\gamma_1^f=10^{-2}$ and $\gamma_1^b=10^{-4}$ and the specified durations of the adsorption and postadsorption purge steps, the initial condition for the fractional coverage of species **A** for the reaction step is chosen to be 0.71 and is assumed to be spatially uniform. This value is computed as described in Ref. 27. The initial condition for the postreaction purge depends on the values obtained from simulations of the reaction step for the different values of γ_2^f chosen. We fix $\gamma_1^f=10^{-2}$, $\gamma_1^b=10^{-4}$, and study three different values of the dimensionless reaction rate, $\gamma_2^f=10^{-2}$, 10^{-4}, and 10^{-6}.

Figure 2 shows plots of the dimensionless number density of **B** for $\gamma_2^f=10^{-2}$ at times $t=10.0$, 40.0, and 80.0 ns, respectively. From the plots, we see that the dimensionless number density of **B** reaches a uniform constant value at all points in the domain in the first 80.0 ns of the reaction step. The number density of **B** at the top of the feature reaches the steady-state value within the first 10.0 ns; however, the number densities at points inside the feature are below this value. This is partly because some **B** is being consumed as it reacts with adsorbed **A** on the feature surface, and partly due to the inherent time required to reach the bottom of the feature. The results for number density of **B** for $\gamma_2^f=10^{-4}$ and 10^{-6} are similar to the results for $\gamma_2^f=10^{-2}$, and the plots for those cases are not displayed. In all cases, the feature fills up with species **B** by the end of 80.0 ns of the reaction step.

Figure 3 shows plots of the dimensionless flux of species **A** to the surface at the three observation points on the feature surface for the three values of γ_2^f under consideration. This flux of **A** is produced by desorption during the reaction step. The flux shows an initial increase to a small value and decays with time. This increase is due to the initial condition for the number density of **A** (zero). There is a small amount of **A** present in the feature due to desorption from the surface, which continues through the reaction step. The flux decays with time much faster for the value of $\gamma_2^f=10^{-2}$ than for the lower values of the dimensionless reaction rate parameter. This is due to the larger rate of reaction for $\gamma_2^f=10^{-2}$.

Figure 4 shows the corresponding plots of the dimensionless flux of **B** to the surface with time. The flux at all three observation points reaches the same steady-state value $\hat{\eta}_2=0.207$ in no more than 30 ms for all values of γ_2^f. The initial transient for $\hat{\eta}_2$ at Point 1 is so short that it is not visible on the scale in the plot; it occurs in a period of less than 10.0 ns. For low values of γ_2^f, the feature floods with reactant, and an essentially spatially uniform flux is established; the results are essentially the same for $\gamma_2^f=10^{-4}$ and $\gamma_2^f=10^{-6}$. Reaction proceeds as a result of this flux. The flux at the interior points of the feature takes longer to reach the steady-state value for $\gamma_2^f=10^{-2}$ than for the other two values of the reaction rate parameter; this is due to more **B** being consumed by reaction when it first contacts the surface than in the case with lower coefficients.

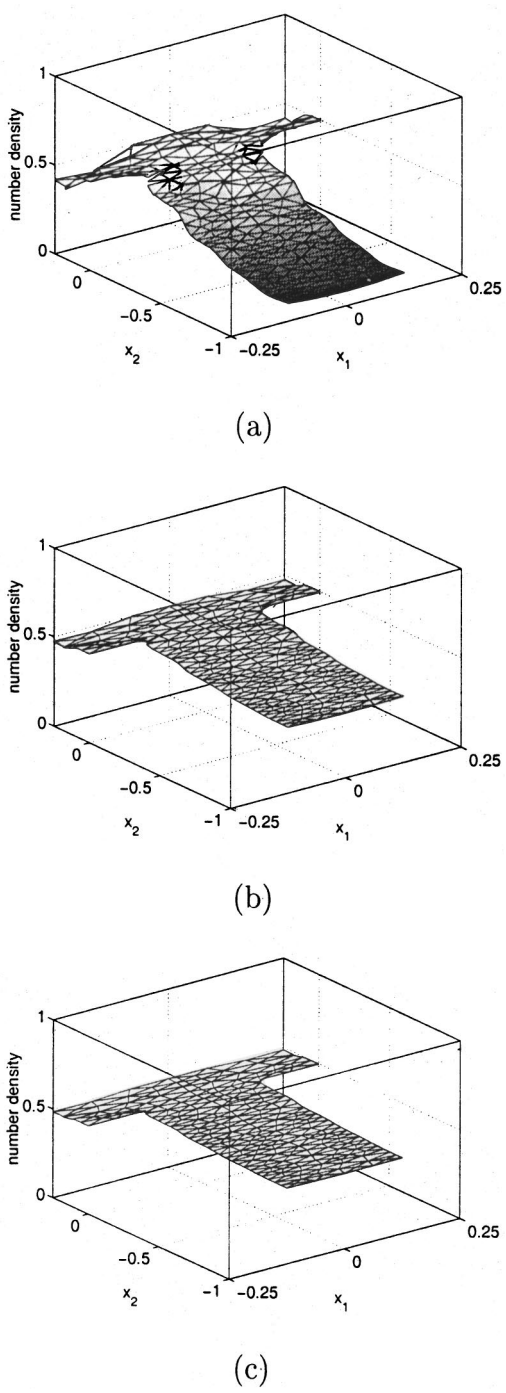


FIG. 2. Reaction step: dimensionless number density of species **B** for a feature with aspect ratio $A=4$ for $\gamma_2^f=10^{-2}$ at times (a) 10.0 ns, (b) 40.0 ns, and (c) 80.0 ns. Note the different scales on the x_1 and the x_2 axes.

The fractional coverage of **A** during the reaction step at the three observation points is shown in Fig. 5. For $\gamma_2^f=10^{-2}$, the coverage at all three points reduces almost to zero in this time due to consumption by reaction and desorption. The fractional coverage at the two interior observation points decreases at slower rates than the coverage in the flat area of the wafer. Figures 5(b) and 5(c) show the coverage

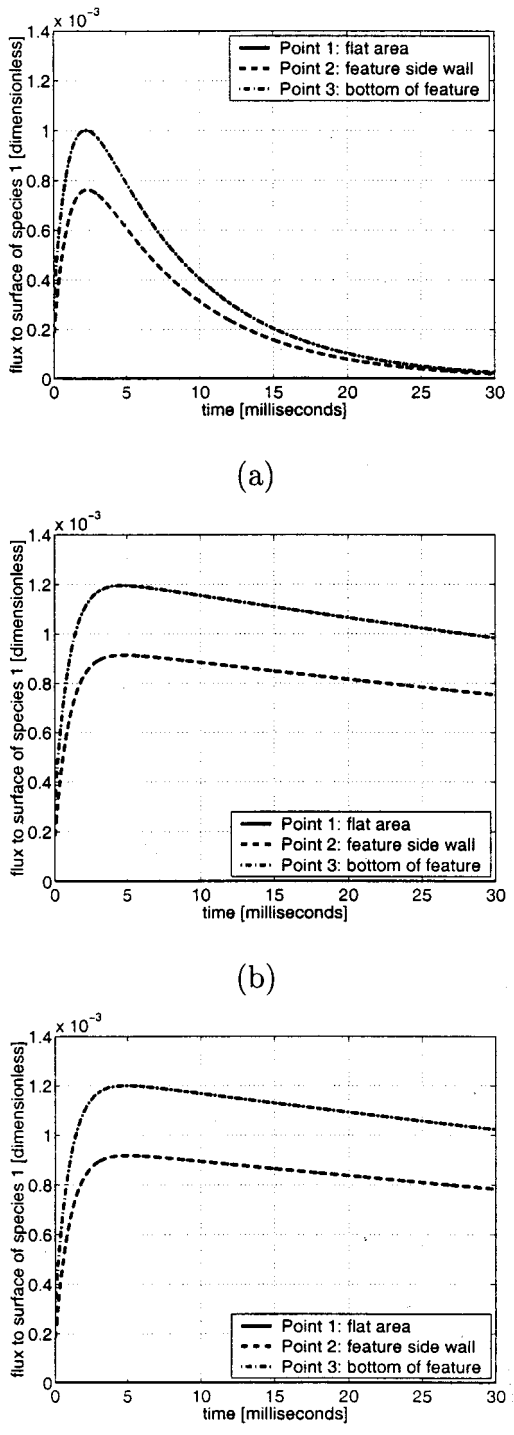


FIG. 3. Reaction step: dimensionless flux to the surface of species **A** vs time at the three observation points [see Fig. 1(a)] for (a) $\gamma_2^f = 10^{-2}$, (b) $\gamma_2^f = 10^{-4}$, and (c) $\gamma_2^f = 10^{-6}$. Notice the scales of the vertical axes.

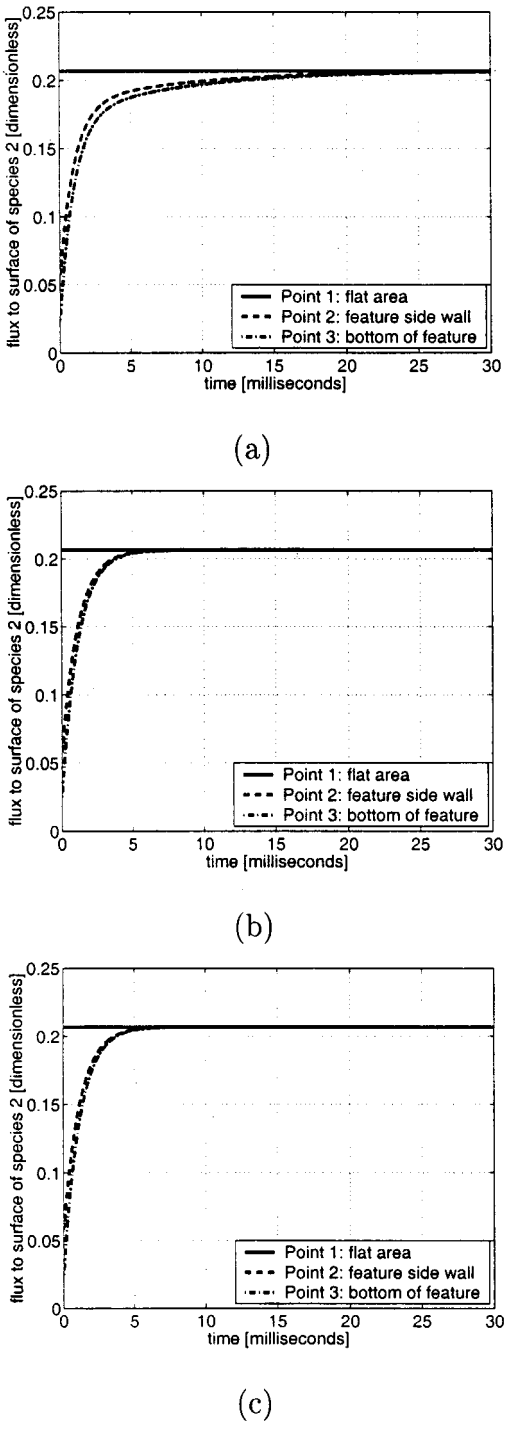


FIG. 4. Reaction step: dimensionless flux to the surface of species **B** vs time at the three observation points [see Fig. 1(a)] for (a) $\gamma_2^f = 10^{-2}$, (b) $\gamma_2^f = 10^{-4}$, and (c) $\gamma_2^f = 10^{-6}$.

for $\gamma_2^f = 10^{-2}$ and 10^{-4}, respectively. The reduction in coverage is much slower at these lower values of the reaction rate parameter. The results shown in Figs. 5(b) and 5(c) are similar because the decrease in coverage is dominated by desorption of **A**, as explained in the following paragraph.

The simulation results shown in Figs. 3 and 4 indicate that we can use the approximations $\hat{\eta}_1 = 0.0$ and $\hat{\eta}_2 = 0.207$ in Eq. (19) to obtain an analytic representation for the fractional coverage. The solid diamonds in Fig. 5 represent the predictions from the analytic expression, and are in excellent agreement with the simulation results at the top of the feature, where the flux reaches the steady-state value the earliest. The

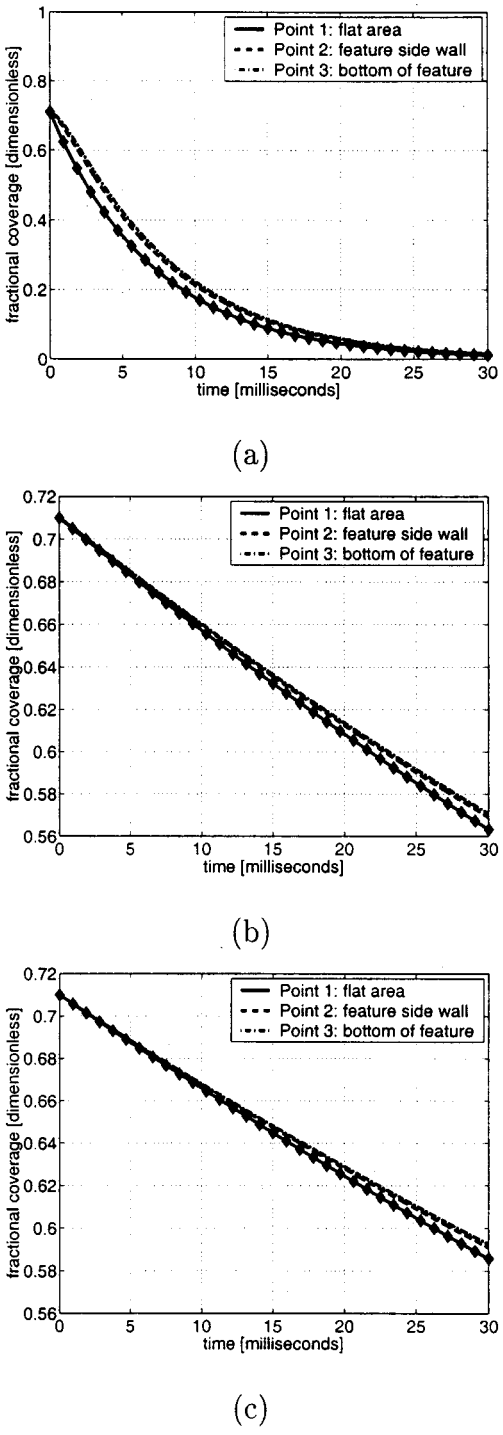


FIG. 5. Reaction step: fractional coverage vs time at the three observation points [see Fig. 1(a)] for (a) $\gamma_2^f = 10^{-2}$, (b) $\gamma_2^f = 10^{-4}$, and (c) $\gamma_2^f = 10^{-6}$. The solid diamonds show the analytical solution given by Eq. (19). Notice the different scales on the vertical axes.

analytic solution also explains the very small difference in the coverage between Figs. 5(b) and 5(c) for $\gamma_2^f = 10^{-2}$ and 10^{-4}, respectively. The parameter in Eq. (19) that determines the rate of decrease in the coverage is $b = \gamma_1^f \hat{\eta}_1 + \gamma_1^b + \gamma_2^f \hat{\eta}_2$. For the flux values in both cases, b is dominated by $\gamma_1^b = 10^{-4}$ for both values of γ_2^f. This again highlights how important it is to account for desorption during the reaction step.

To justify the parameter values chosen in our simulation studies, we perform an analysis on the deposition rates for each value of γ_2^f we have considered. This is done by obtaining estimates of the maximum deposition rate at the start of the reaction step. The deposition rate r is governed by the reaction rate R_2. Using Table III and Eq. (16), we find

$$R_2 = \eta^* \hat{R}_2 = \eta^* \gamma_2^f \vartheta_A \hat{\eta}_2 . \tag{36}$$

For our choice of initial condition for the reaction step, ϑ_A is 0.71 at the start and can be set to unity in an order of magnitude analysis. Using the value for η^* from Table II and $\hat{\eta}_2 = 0.207$ obtained from the simulation results

$$R_2 = 1.3 \times 10^{-5} \gamma_2^f \text{ mol/(s cm}^2\text{)} \tag{37}$$

with the dimensionless reaction coefficient γ_2^f still present. This molar reaction rate can be converted to a deposition rate by multiplication with an appropriate molar volume. If a molar volume of $V = 10$ cm^3/mol is assumed, the deposition rate r is

$$r = V R_2 \approx 10^{-4} \gamma_2^f \text{ cm/s}. \tag{38}$$

Thus, the maximum deposition rates are of the order of 10, 0.1, and 0.001 nm/s for dimensionless reaction rates of $\gamma_2^f = 10^{-2}$, 10^{-4}, and 10^{-6}, respectively. Deposition rates in constant rate processes such as chemical vapor deposition are at the higher end of the range considered in our simulations, and we conclude that appropriate values of γ_2^f have been chosen for the analysis in this study.

B. Postreaction purge step

At the beginning of the purge step after reaction, the domain is filled with gaseous molecules of **B**, hence $\hat{c}_2^{\text{ini}} = \hat{c}_2^{\text{ref}} = 0.5$; see Table IV. But no more gas is fed from the top, therefore $\hat{c}_2^{\text{top}} = 0.0$. The corresponding values for **A** are chosen to be $\hat{c}_1^{\text{ini}} = 0.0$ and $\hat{c}_1^{\text{top}} = 0.0$. This means that there is no **A** initially present in the domain or being fed into the domain at the top.

The initial condition for the fractional coverage ϑ_A for the postreaction purge is assumed to be the value of the coverage at the end of the reaction step. Using the length of 450 ms for the reaction step and the observed value of $\hat{\eta}_2$, the analytic solution given by Eq. (19) is used to predict the appropriate values. For $\gamma_2^f = 10^{-2}$, the fractional coverage at the end of the reaction step is zero. The corresponding values of coverage for $\gamma_2^f = 10^{-4}$ and 10^{-6} are 0.022 and 0.040, respectively. The coverage at the end of the reaction step for all cases is essentially spatially uniform, and this spatially uniform initial condition is used for the coverage in the postreaction purge.

Plots of the dimensionless number density of **B** for $\gamma_2^f = 10^{-2}$ are presented in Fig. 6 for 10.0, 40.0, and 80.0 ns into the postreaction purge step. Plots for $\gamma_2^f = 10^{-4}$ and 10^{-6} are similar and are not shown. In all cases, the number density of **B** at the top of the feature reduces almost to zero

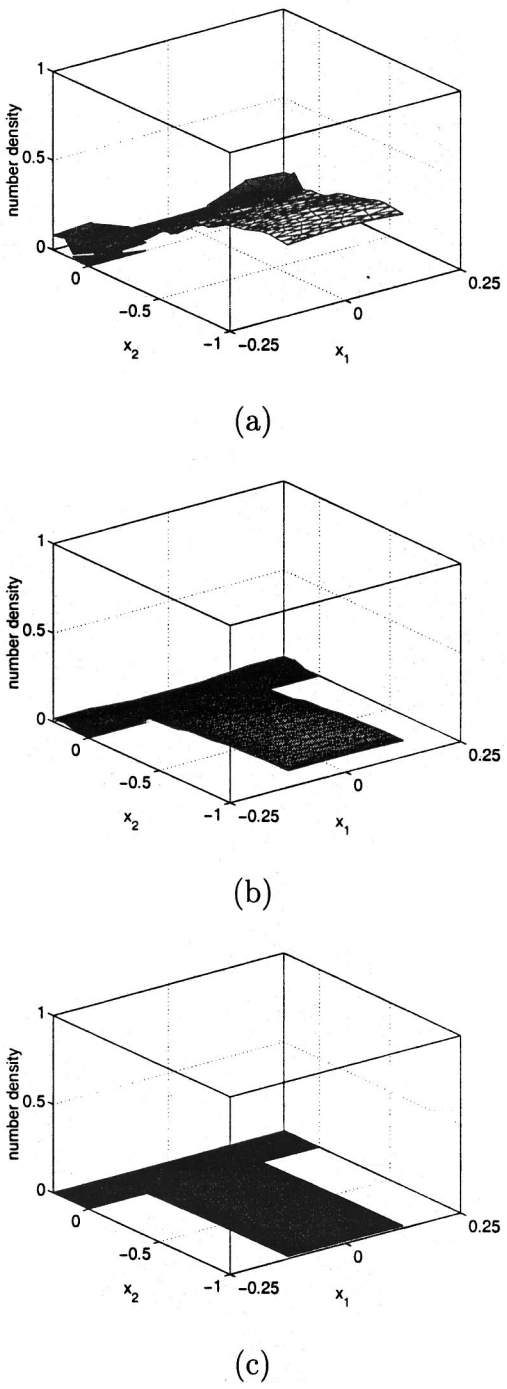


FIG. 6. Postreaction purge step: dimensionless number density of species **B** for a feature with aspect ratio $A=4$ for $\gamma_2^f=10^{-2}$ at times (a) 10.0 ns, (b) 40.0 ns, and (c) 80.0 ns. Note the different scales on the x_1 and the x_2 axes.

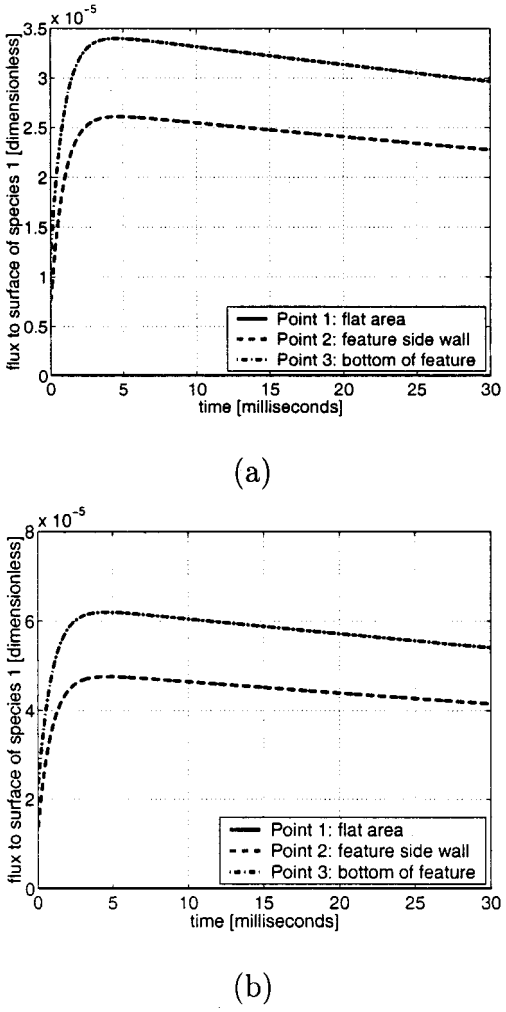


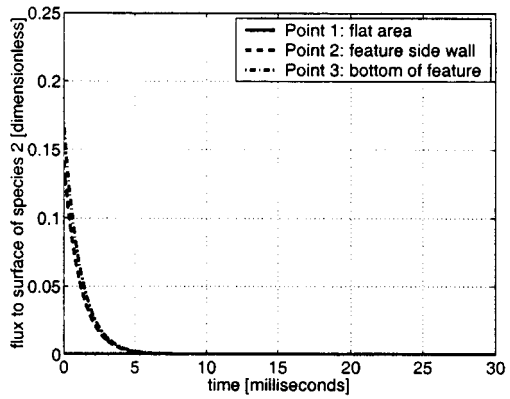
FIG. 7. Postreaction purge step: dimensionless flux to the surface of species **A** vs time at the three observation points [see Fig. 1(a)] for (a) $\gamma_2^f=10^{-4}$ and (b) $\gamma_2^f=10^{-6}$. Notice the scales of the vertical axes.

within the first 10.0 ns, but it takes close to 80.0 ns for **B** to be removed from the feature volume. This result justifies the initial condition used for the number density of species **A** in the reaction step.

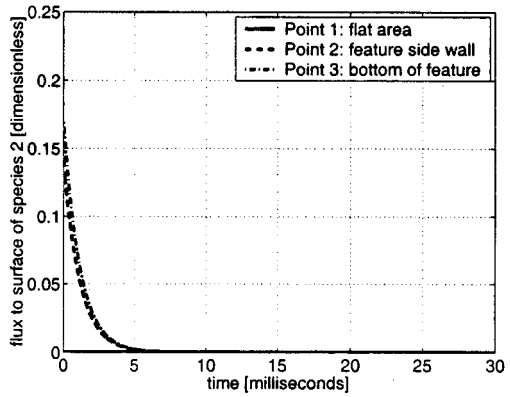
Figures 7(a) and 7(b) show the dimensionless flux of **A** for $\gamma_2^f=10^{-4}$ and 10^{-6}, respectively. In both cases, the flux of **A** decays after an initial increase. The flux values for $\gamma_2^f=10^{-4}$ are lower than the flux values for $\gamma_2^f=10^{-6}$, because the level of adsorbed **A** available for desorption is lower. The initial increase is a consequence of the choice of a zero initial condition for **A**. Although this initial condition is an approximation, the number densities are several orders of magnitude lower than the flux values for $\gamma_2^f=10^{-6}$, because the level of adsorbed **A** available for desorption is lower. The initial increase is a consequence of the choice of a zero initial condition for **A**. Although this initial condition is an approximation, the number densities are several orders of magnitude lower than those of **B**. The flux of **A** for $\gamma_2^f=10^{-2}$ is zero, because no **A** is present in the system, and is not shown.

Figure 8 shows the corresponding plots of the dimensionless flux of **B**. The dimensionless flux of **B** decays to zero in approximately 6 ms for all values of γ_2^f. The results for the flux of **B** are similar for all three values of γ_2^f, largely because **B** does not adsorb.

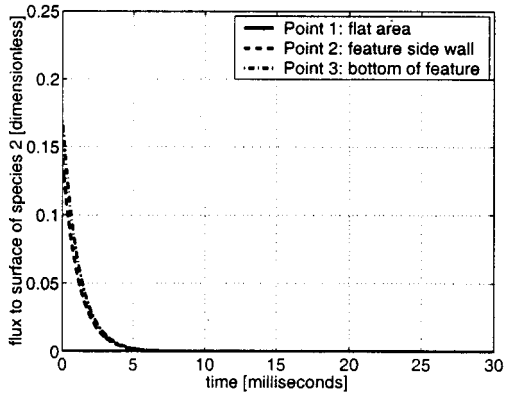
Figures 9(a) and 9(b) show results for the fractional coverage of species **A** for $\gamma_2^f=10^{-4}$ and 10^{-6}, respectively. The results for $\gamma_2^f=10^{-2}$ are not presented, since the coverage at the end of the reaction step was zero, and no **A** is present in the system. The diamonds in the plots show the predictions



(a)

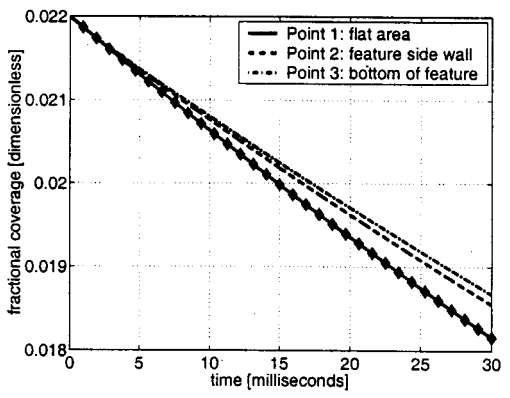


(b)

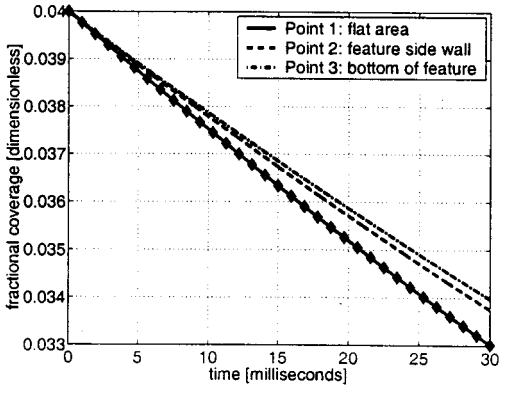


(c)

FIG. 8. Postreaction purge step: dimensionless flux to the surface of species **B** vs time at the three observation points [see Fig. 1(a)] for (a) $\gamma_2^f = 10^{-2}$, (b) $\gamma_2^f = 10^{-4}$, and (c) $\gamma_2^f = 10^{-6}$.



(a)



(b)

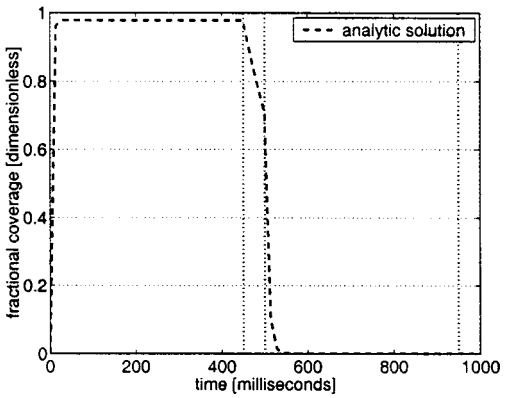
FIG. 9. Postreaction purge step: fractional coverage vs time at the three observation points [see Fig. 1(a)] for (a) $\gamma_2^f = 10^{-4}$ and (b) $\gamma_2^f = 10^{-6}$. The solid diamonds show the analytical solution given by Eq. (19). Notice the different scales on the vertical axes.

of the analytic solution given by Eq. (19) with the dimensionless fluxes of **A** and **B** being set to zero. The analytic expression provides good estimates of the reduction in coverage with time for the observation point in the flat area of the feature. The coverage at the points in the interior of the feature show a lag from the coverage in the flat area.

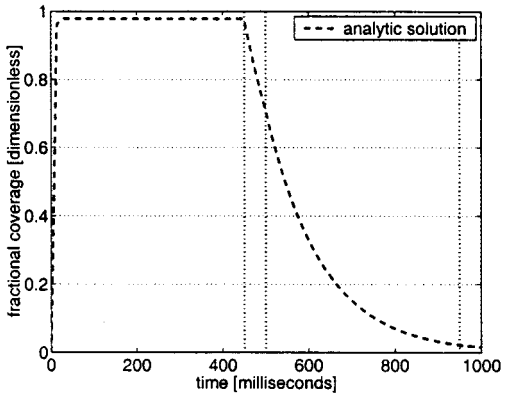
From the plots shown up to 30 ms and from the plots of number density and flux for the postreaction purge, we conclude that the purge step can be shortened. This assumes that changes on the reactor scale could be accomplished at the rates required. The values of coverage are low, and desorption further reduces the coverage. The fluxes and number densities indicate that the feature is free of **B** within the first 6 ms of the postreaction purge, and the rest of the purge time only serves to reduce the fractional coverage of **A** within the feature.

C. Analytic model of one ALD cycle

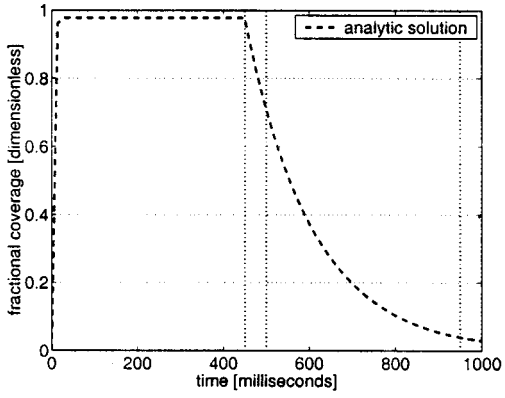
The comparisons between simulation results and the analytic solution for the fractional coverage ϑ_A demonstrate that the analytic solution can be used as a predictor for many choices of reaction coefficients. Therefore, we complete the information given by the simulations for all four steps of one ALD cycle (in Ref. 27 and above) by using the analytic solution to plot the evolution of ϑ_A for one full ALD cycle in Fig. 10. The vertical lines mark the beginning and end of the steps: adsorption for $0 \leqslant t \leqslant 450$ ms, postadsorption purge for $450 \leqslant t \leqslant 500$ ms, reaction for $500 \leqslant t \leqslant 950$ ms, and postreaction purge for $950 \leqslant t \leqslant 1000$ ms.



(a)



(b)



(c)

FIG. 10. Prediction of fractional coverage vs processing time during one full ALD cycle using the analytic solution given by Eq. (19) for (a) $\gamma_2^f = 10^{-2}$, (b) $\gamma_2^f = 10^{-4}$, and (c) $\gamma_2^f = 10^{-6}$. Adsorption step: $0 \leqslant t \leqslant 450$ ms, postadsorption purge step: $450 \leqslant t \leqslant 500$ ms, reaction step: $500 \leqslant t \leqslant 950$ ms, postreaction purge step: $950 \leqslant t \leqslant 1000$ ms.

As all three plots show, we chose coefficients for the adsorption and desorption of species **A** that result in equilibrium coverage at the end of the adsorption step, and a sizable decrease of coverage during the postadsorption purge step. Figure 10 shows that for the largest value of γ_2^f considered, the reaction of gaseous **B** with the adsorbed molecules of **A** proceeds rapidly. This causes the rapid decrease of fractional coverage of **A** during the reaction step. As Figs. 10(b) and 10(c) show for the smaller values of γ_2^f considered, the decrease in fractional coverage is slow by comparison but still reaches very low levels at the end of the reaction step (450 ms). However, as the similarity of Figs. 10(b) and 10(c) demonstrates, the decrease of coverage is dominated by desorption of adsorbed molecules of **A** from the wafer surface and not by reaction of gaseous molecules of **B** with adsorbed molecules of **A**. This means that the speed of the decrease of coverage in itself does not indicate a desired performance of the process: deposition occurs if the coverage of **A** decreases due to reaction with **B**. This may be interpreted to mean that the choice of chemistry should be made with the relative size of desorption rate of **A** and reaction rate of **A** and **B** in mind, the best being a slow desorption and fast reaction rate.

V. CONCLUSIONS

A model and a numerical method capable of simulating the transient behavior during the various stages of atomic layer deposition were presented. The reaction and postreaction purge steps of ALD were studied. The transport in the gas phase was modeled by the Boltzmann equation. The heterogeneous chemistry model included reversible adsorption of a reactant on a single site, and irreversible reaction of a second gaseous reactant with the adsorbed reactant. Parametric studies were conducted on the effect of the reaction rate constant. Number densities of gaseous species, fluxes to the surface of the feature, and the surface coverage of the adsorbing reactant were computed as functions of time. The results provide information on the interplay between gas phase transport and surface reaction mechanisms during ALD. The gas transport in the feature volume is relatively fast, and the surface mechanisms are rate-controlling for the parameter values studied. For these cases, the flux and the number density within the feature volume are spatially uniform and at steady-state for almost the entire time of each pulse in an ALD cycle. An analytic expression provides a good approximation to the solution for surface coverage of the adsorbing reactant obtained from the transport and reaction model. The results show that the fractional coverage of the adsorbing reactant reduces significantly in the reaction step due to reaction with the gaseous reactant. Larger values of the reaction rate parameter lead to larger reductions in the fractional coverage during the reaction step. For smaller values of the reaction rate parameter, the decrease in coverage is dominated by desorption. The surface coverage of the adsorbing reactant also decreases during the reaction and purge steps, due to desorption.

ACKNOWLEDGMENTS

Professor Gobbert acknowledges support from the NSF under Grant No. DMS-9805547. The RPI authors acknowledge support from MARCO, DARPA, and NYSTAR. We are also thankful to C. Ringhofer and J.-F. Remacle, without whose help this work would not have been possible.

[1]M. Ritala and M. Leskela, Nanotechnology **10**, 19 (1999).
[2]M. Leskela and M. Ritala, J. Phys. IV **5**, 937 (1995).
[3]T. Suntola, Thin Solid Films **216**, 84 (1992).
[4]H. Siimon and J. Aarik, J. Phys. D **30**, 1725 (1997).
[5]J. Klaus, S. Ferro, and S. George, Thin Solid Films **360**, 145 (2000).
[6]J.-W. Lim, J.-S. Park, and S.-W. Kang, J. Appl. Phys. **87**, 4632 (2000).
[7]J.-W. Lim, H.-S. Park, and S.-W. Kang, J. Appl. Phys. **88**, 6327 (2000).
[8]J.-W. Lim, H.-S. Park, and S.-W. Kang, J. Electrochem. Soc. **148**, C403 (2000).
[9]H.-S. Park, J.-S. Min, J.-W. Lim, and S.-W. Kang, Appl. Surf. Sci. **158**, 81 (2000).
[10]T. S. Cale, T. P. Merchant, L. J. Borucki, and A. H. Labun, Thin Solid Films **365**, 152 (2000).
[11]T. S. Cale and V. Mahadev, *Thin Films* (Academic, New York, 1996), Vol. 22, p. 175.
[12]R. I. Masel, *Principles of Adsorption and Reaction on Solid Surfaces* (Wiley–Interscience, New York, 1996).
[13]M. K. Gobbert, C. A. Ringhofer, and T. S. Cale, J. Electrochem. Soc. **143**, 2624 (1996).
[14]M. K. Gobbert, T. P. Merchant, L. J. Borucki, and T. S. Cale, J. Electrochem. Soc. **144**, 3945 (1997).
[15]M. K. Gobbert, T. S. Cale, and C. A. Ringhofer, VLSI Design **6**, 399 (1998).
[16]S. T. Rodgers and K. F. Jensen, J. Appl. Phys. **83**, 524 (1989).
[17]T. P. Merchant, M. K. Gobbert, T. S. Cale, and L. J. Borucki, Thin Solid Films **365**, 368 (2000).
[18]T. S. Cale, T. P. Merchant, and L. J. Borucki, in *Advanced Metallization Conference in 1998*, edited by G. S. Sandhu, H. Koerner, M. Murakami, Y. Yasuda, and N. Kobayashi (Materials Research Society, Pittsburgh, 1999), pp. 737–741.
[19]C. Cercignani, *The Boltzmann Equation and Its Applications*, Vol. 67, Applied Mathematical Sciences (Springer, New York, 1988).
[20]C. Cercignani, *Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations*, Cambridge Texts in Applied Mathematics (Cambridge University, Cambridge, 2000).
[21]G. N. Patterson, *Introduction to the Kinetic Theory of Gas Flows* (University of Toronto, Toronto, 1971).
[22]C. Ringhofer, Acta Numerica **6**, 485 (1997).
[23]M. K. Gobbert and C. Ringhofer, *IMA Volumes in Mathematics and its Applications*, edited by I. Gamba, D. Levermore, and C. Ringhofer (Springer, New York, to be published).
[24]M. K. Gobbert and C. Ringhofer (to be published).
[25]J.-F. Remacle, J. Flaherty, and M. Shephard (to be published).
[26]M. K. Gobbert, J.-F. Remacle, and T. S. Cale (to be published).
[27]M. K. Gobbert, S. G. Webster, and T. S. Cale, J. Electrochem. Soc. (to be published).