Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Computational Physics 213 (2006) 591–612
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

A kinetic transport and reaction model and simulator for rarefied gas flow in the transition regime

Matthias K. Gobbert [a,*], Timothy S. Cale [b]

[a] Department of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA
[b] Focus Center – New York, Rensselaer: Interconnections for Hyperintegration, Isermann Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, CII 6015, 110 8th Street, Troy, NY 12180-3590, USA

Received 17 June 2005; received in revised form 24 August 2005; accepted 25 August 2005
Available online 14 October 2005

Abstract

We present a model for rarefied gas flows that are characterized by reactive species as minor constituents in a dominant inert carrier species. The kinetic transport and reaction model consists of a system of transient linear Boltzmann equations for the reactive species in the flow. This model applies to a wide range of transport regimes, including the transition regime in which both transport and collisions between molecules must be taken into account, characterized by Knudsen numbers on the order of unity. A numerical simulator based on a spectral Galerkin method in velocity space approximates each linear Boltzmann equation by a system of transient conservation laws in space and time with diagonal coefficient matrices, which are solved using the discontinuous Galerkin method. This deterministic solver gives direct access to the kinetic density that is the solution to the Boltzmann equation, as a function of position, velocity, and time. It is valuable to have direct access to the velocity dependence in order to analyze the underlying kinetic causes of macroscopic observeables. Using chemical vapor deposition as an important application example, the influence of process parameters is studied in two-dimensional reference studies and transient studies for a three-dimensional domain that represents structures seen during integrated circuit fabrication.

Keywords: Rarefied gas flow; Boltzmann transport equation; Spectral Galerkin method; Discontinuous Galerkin method; Chemical vapor deposition

1. Introduction

Many important manufacturing processes for integrated circuits involve the flow of gaseous reactants at pressures that range from very low to atmospheric [18]. Correspondingly, the mean free path λ (the average distance that a molecule travels before colliding with another molecule) ranges from less than 0.1 μm to over 100 μm. The typical size of the electronic components (called 'features' during integrated circuit processing) is

* Corresponding author. Tel.: +1 410 455 2404; fax: +1 410 455 1066.
E-mail addresses: gobbert@math.umbc.edu (M.K. Gobbert), calet@rpi.edu (T.S. Cale).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2005.08.026

now below 1 μm and the size of the chemical reactor, in which the gas flow takes place, can be on the order of a meter in one or more dimensions. Thus, models on a range of length scales L^* are of interest, each of which needs to be appropriately selected to be valid on its length scale.

The appropriate transport model at a given combination of pressure and length scale is determined by the Knudsen number Kn, defined as the ratio of the mean free path and the length scale of interest $\mathrm{Kn} := \lambda/L^*$, which is a dimensionless group obtained in the non-dimensionalization of the Boltzmann transport equation [18]: For very small Kn, the Boltzmann equation simplifies to the equations of continuum flow models. For very large Kn, the collision term in the Boltzmann equation goes to zero and collisionless or ballistic transport models are appropriate. For intermediate Kn, flow is in the transition regime and the Boltzmann equation is appropriate. Guidelines for deciding which flow regime should be modeled differ somewhat. It is safe to say that for $\mathrm{Kn} < 0.01$, continuum equations are appropriate, while for $\mathrm{Kn} > 100.0$, ballistic transport models are appropriate.

Our interest includes models on the micron- to millimeter-scale at a range of pressures, resulting in Knudsen numbers ranging across the wide spectrum from $\mathrm{Kn} = 0.01$ to $\mathrm{Kn} \to \infty$, with a particular focus on the transition regime with $\mathrm{Kn} \approx 1.0$. We have developed a kinetic transport and reaction model (KTRM) for multicomponent, reactive flows typical of those seen in integrated circuit fabrication equipment. The KTRM is represented by a system of linear Boltzmann equations for all n_s reactive species in dimensionless form

$$\frac{\partial f^{(i)}}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f^{(i)} = \frac{1}{\mathrm{Kn}} Q_i\big(f^{(i)}\big), \quad i = 1, \ldots, n_s, \tag{1}$$

with the linear collision operators

$$Q_i\big(f^{(i)}\big)(\mathbf{x}, \mathbf{v}, t) = \int_{\mathbb{R}^3} \sigma_i(\mathbf{v}, \mathbf{v}') \big[M_i(\mathbf{v}) f^{(i)}(\mathbf{x}, \mathbf{v}', t) - M_i(\mathbf{v}') f^{(i)}(\mathbf{x}, \mathbf{v}, t)\big] \mathrm{d}\mathbf{v}',$$

where $\sigma_i(\mathbf{v}, \mathbf{v}') = \sigma_i(\mathbf{v}', \mathbf{v}) > 0$ is a given collision frequency model and $M_i(\mathbf{v})$ denotes the Maxwellian density of species i. The left-hand side of (1) models the advective transport of molecules of species i (local coupling of spatial variations via the spatial derivatives $\nabla_{\mathbf{x}} f^{(i)}$), while the right-hand side models the effect of collisions (global coupling of all velocities in the integral operators Q_i). The unknown functions $f^{(i)}(\mathbf{x}, \mathbf{v}, t)$ in this kinetic model represent the (scaled) probability density, which we call kinetic density for short and to distinguish it clearly from other densities, that a molecule of species $i = 1, \ldots, n_s$ is at a position in $[\mathbf{x}, \mathbf{x} + \mathrm{d}\mathbf{x}]$ with a velocity in $[\mathbf{v}, \mathbf{v} + \mathrm{d}\mathbf{v}]$ at a time in $[t, t + \mathrm{d}t]$. Its values need to be determined at all points $\mathbf{x} \in \Omega \subset \mathbb{R}^3$ and for all velocity vectors $\mathbf{v} \in \mathbb{R}^3$ at all times $0 < t \leqslant t_{\mathrm{fin}}$. Models in both two and three dimensions are of interest for our applications, but we write all equations in three dimensions for clarity of presentation. Notice that while the equations in (1) appear decoupled, they actually remain coupled through the boundary conditions at the wafer surface that involve general models for the interaction of the gas phase reactive species with wafer surface.

The purpose of this paper is to provide a detailed derivation to the KTRM for multi-species collisional gas flow in a dominant carrier gas (Section 2), and to demonstrate its predictive capabilities (Section 4); the numerical method for the simulations is sketched in Section 3. This work extends the collision-less model that we introduced for the process of atomic layer deposition (ALD) in two dimensions [13–16] and three dimensions [26]. First, results for this extended model are contained in [12,25]. Other work, more focused on the numerical method, can be found in [17] and on the scalability of the parallel implementation in [12,17,25,27].

We accomplish the extension to collisional transport by combining modeling techniques from neutron transport (see, e.g., [8, Chapter IV]) that allow us to pose the KTRM as a system of linear Boltzmann equations with numerical techniques originally devised for the semiconductor Boltzmann equation in [22–24]. Two of the fundamental challenges of the Boltzmann transport equation are the quadratic non-linearity of its solution and the high dimensionality of the integral in the collision operator. As in the context of neutron transport, the fact that the reactive species are at least an order of magnitude less dense than the inert carrier gas gives rise to the linear Boltzmann equation as appropriate model in (1). The model is simpler in that the solution appears linearly in the collision operator which involves only integration over velocity space. The remaining fundamental challenge for numerical simulations of kinetic models is the high dimensionality of the phase space variables $(\mathbf{x}, \mathbf{v})$ that need to be discretized, and at every time step if transient studies are desired. Historically, Monte Carlo methods [3] were usually used to attack the problem, as deterministic methods

could not resolve the phase space adequately on available computers. Recent work [7,11] has demonstrated that efficient deterministic methods can be competitive with Monte Carlo methods. Besides the inherent advantage of avoiding stochastic variability, [7,11] also show that deterministic methods can give direct access to the kinetic density as a function of $\mathbf{v}$ at selected positions $\mathbf{x}$ and times t. For these same reasons, we chose a deterministic simulator for the KTRM.

Since the operating conditions of the applications under consideration give rise to Knudsen numbers in a range that includes the transition regime, a kinetic model is necessary. Simultaneously, since the KTRM can be formulated as a system of linear Boltzmann equations and the solutions are not expected to be too far from Maxwellian, it is appropriate to use a moment method approach following, for instance, the formulation in [22–24] for the semiconductor Boltzmann equation. But the numerical challenges are quite different in those applications: The spatial domain of a transistor channel is reasonably modeled as one-dimensional, and even with two-dimensional extensions of interest (e.g., [6]), the most fundamental numerical difficulty still lies in the coupling of the linear Boltzmann equation with the Poisson equation that is driven by an applied voltage whose values vary over several orders of magnitude. By contrast, we are materially interested in multi-species models, with coupling to reaction models through boundary conditions, and in developing a simulator for two- and three-dimensional problems, whose domains are irregular in shape. This explains our particular formulation [17] of the spectral Galerkin method in velocity space that gives a system of hyperbolic equations with diagonal system matrices for every Boltzmann equation in our model and makes dealing with higher dimensions systematic, and the use of a finite element method to discretize the physical domain Ω that is designed to discretize higher dimensional domains of any shape.

The remainder of this paper is organized as follows: Section 2 derives the KTRM in (1) in detail with its assumptions and our choices for the model and non-dimensionalization parameters. Section 3 provides a brief overview of the numerical method used in the simulator for the KTRM. Section 4 shows the ability of the KTRM to gain insight into phenomena in an application example, the simulation of initial deposition in chemical vapor deposition: First, Section 4.1 analyzes the behavior of the model as two physical parameters (the sticking factor γ_0 of the deposition and the Knudsen number Kn) are varied and demonstrates the meaning of the kinetic density $f^{(i)}(\mathbf{x},\mathbf{v},t)$ as a function of its velocity argument $\mathbf{v}$ at a specific position $\mathbf{x}$ and time t. Then, Section 4.2 shows results of the KTRM used to model chemical vapor deposition in a three-dimensional domain with a realistic irregular shape and analyzes the effect of different Knudsen numbers.

2. The model

2.1. The spatial domain

Fig. 1 shows two views of the spatial domain $\Omega \subset \mathbb{R}^3$ for a representative trench/via feature; more precisely, the plots show the solid wafer surface $\Gamma_w \subset \partial\Omega$ consisting of a 0.3 μm deep trench (which in practice can be very long) of width 0.4 μm (x_1-coordinates between 0.3 and 0.7), into which is etched another 0.3 μm deep via

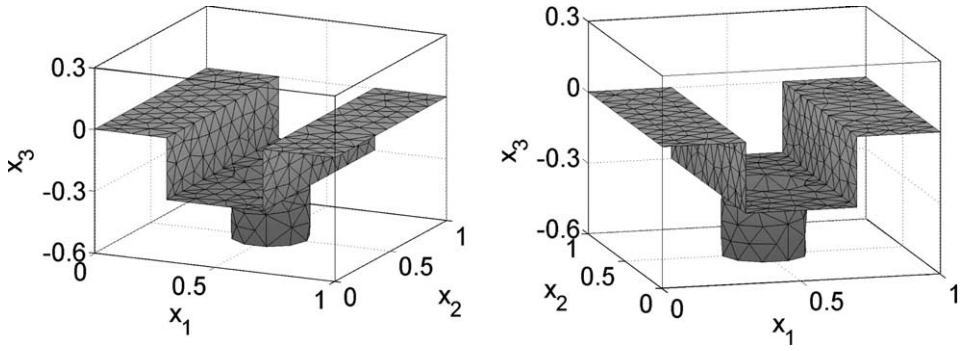

Fig. 1. Two views of the solid wafer surface of the three-dimensional trench/via feature. The domain Ω is the gaseous region above the wafer surface up to the interface to the bulk gas at $x_3 = 0.3$, the top of the plot box.

(round hole). The domain Ω for our model is the gaseous region above the solid wafer surface Γ_w up to the top Γ_t of the plot box at $x_3 = 0.3$ μm and bounded on the sides by the vertical parts of the plot box collectively denoted by Γ_s. The portions of the boundary are chosen mutually disjoint and such that $\partial\Omega = \Gamma_w \cup \Gamma_t \cup \Gamma_s$.

In our context of deposition processes in the manufacturing of computer chips, the trench and the via would have been etched in the previous production step. The purpose of the present manufacturing step is, for instance, to fill them with conducting material that will act as connection between the electronic components of the chip. Since this domain does not admit any symmetries that allow for a reduction to a two-dimensional problem, this example demonstrates the importance of being able to accommodate a three-dimensional domain with irregular shape.

2.2. The dimensional model equations

To better accommodate the chemical reactions vital to the applications of interest here, we use molar units, that is, the kinetic density $f^{(i)}(\mathbf{x},\mathbf{v},t)$ has units of mol/(cm^3 (cm/s)^3) such that $\int f^{(i)}(\mathbf{x},\mathbf{v},t)\,d\mathbf{v}$ defines the molar concentration $c_i(\mathbf{x},t)$ of species i in mol/cm^3. Using the molecular weight ω_i of species i with units g/mol, these variables can be related to the mass density $\rho_i = \omega_i c_i$ and to the mass-scaled kinetic density $\tilde{f}^{(i)} = \omega_i f^{(i)}$ such that $\rho_i = \int \tilde{f}^{(i)}(\mathbf{x},\mathbf{v},t)\,d\mathbf{v}$. Conventions of process engineering are the reason we do not use SI units but rather the cgs-system with base units cm for length, g for mass, and s for time.

Our derivation of the KTRM starts with the dimensional system of non-linear Boltzmann equations appropriate for a multicomponent system with $n_s + 1$ chemical species $i = 0,1,\ldots,n_s$

$$\frac{\partial f^{(i)}}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{x}} f^{(i)} = \sum_{j=0}^{n_s} Q_{ij}\big(f^{(i)},f^{(j)}\big) + \sum_{j=0}^{n_s}\sum_{k=0}^{n_s}\sum_{\ell=0}^{n_s} R_{ij}^{k\ell}\big(f^{(i)},f^{(j)},f^{(k)},f^{(\ell)}\big). \tag{2}$$

To distinguish the effects of purely elastic collisions from collisions that also involve a chemical reaction, the terms are written in additive form in (2); i.e., following, for instance, [9, p. 224] and the original references cited there, the collision kernels B_{ij} in the collision operators Q_{ij} model elastic collisions and the functions $W_{ij}^{k\ell}$ in the reaction operators $R_{ij}^{k\ell}$ reactive collisions. In more detail, the collision operator Q_{ij} models collisions between molecules of species i and j. It can be stated in many forms that differ in subtle details; we will follow the formulation used in [8, p. 64] (adjusted for our units)

$$Q_{ij}\big(f^{(i)},f^{(j)}\big) = \int_{\mathbb{R}^3}\int_0^{2\pi}\int_0^{\pi/2}\Big[f^{(i)'}f_*^{(j)'} - f^{(i)}f_*^{(j)}\Big]B_{ij}(\vartheta,V)\,d\vartheta\,d\varepsilon\,d\mathbf{v}_* \tag{3}$$

with the short-hand notations $f^{(i)'} = f^{(i)}(\mathbf{x},\mathbf{v}',t)$, $f_*^{(j)'} = f^{(j)}(\mathbf{x},\mathbf{v}_*',t)$, $f^{(i)} = f^{(i)}(\mathbf{x},\mathbf{v},t)$, and $f_*^{(j)} = f^{(j)}(\mathbf{x},\mathbf{v}_*,t)$. In turn, $\mathbf{v}'$ and $\mathbf{v}_*'$ are short-hand notations for the pre-collision velocities related to $\mathbf{v}$ and $\mathbf{v}_*$ via the conservation of momentum and energy

$$m_i\mathbf{v} + m_j\mathbf{v}_* = m_i\mathbf{v}' + m_j\mathbf{v}_*', \tag{4a}$$

$$m_i v^2 + m_j v_*^2 = m_i v'^2 + m_j v_*'^2, \tag{4b}$$

where we use the notations $v = |\mathbf{v}|$, $v_* = |\mathbf{v}_*|$, $v' = |\mathbf{v}'|$, and $v_*' = |\mathbf{v}_*'|$. In (3), $B_{ij}(\vartheta,V)$ in units of cm^3/(mol s) is the collision kernel scaled properly for our units. The function $B_{ij}(\vartheta,V)$ depends only on the magnitude of the relative velocity $V = |\mathbf{V}| = |\mathbf{v} - \mathbf{v}_*|$ and on the collision angle ϑ that is defined as the angle between $\mathbf{V}$ and $\mathbf{v} - \mathbf{v}'$.

For the form of the reaction operators, we follow [8,9,18] that propose a generalization from collisions to a class of chemical reactions in the gas phase by defining the reaction term $R_{ij}^{k\ell}$ as

$$R_{ij}^{k\ell} = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\Big[f^{(k)'}f_*^{(\ell)'} - f^{(i)}f_*^{(j)}\Big]W_{ij}^{k\ell}(\mathbf{v},\mathbf{v}_*,\mathbf{v}',\mathbf{v}_*')\,d\mathbf{v}_*\,d\mathbf{v}'\,d\mathbf{v}_*' \tag{5}$$

with the short-hand notations $f^{(k)'} = f^{(k)}(\mathbf{x},\mathbf{v}',t)$, $f_*^{(\ell)'} = f^{(\ell)}(\mathbf{x},\mathbf{v}_*',t)$, $f^{(i)} = f^{(i)}(\mathbf{x},\mathbf{v},t)$, and $f_*^{(j)} = f^{(j)}(\mathbf{x},\mathbf{v}_*,t)$, but where the velocities $\mathbf{v}$, $\mathbf{v}_*$, $\mathbf{v}'$, and $\mathbf{v}_*'$ are now taken as independent variables. Here, $W_{ij}^{k\ell}$ models a reaction in which molecules of species k and ℓ become molecules of species i and j.

To motivate the particular form of $R_{ij}^{k\ell}$ that models binary reactive collisions, notice that it is a generalization of the collision operator Q_{ij} that models binary elastic collisions: Prevent the molecules in a reactive collision in $R_{ij}^{k\ell}$ from changing species by fixing indices in $W_{ij}^{k\ell} = W_{ij}\delta_{ik}\delta_{j\ell}$ (with Kronecker deltas), thus dropping out the summations over k and ℓ, which makes the collision elastic. Then enforce momentum and energy conservation by introducing Dirac delta functions by setting [8, p. 65]

$$W_{ij}(\mathbf{v},\mathbf{v}_*,\mathbf{v}',\mathbf{v}'_*) = S_{ij}(\vartheta,V)\delta(m_i\mathbf{v}+m_j\mathbf{v}_* - m_i\mathbf{v}' - m_j\mathbf{v}'_*)\times\delta\big(m_iv^2+m_jv_*^2 - m_iv'^2 - m_jv_*'^2\big), \tag{6}$$

where the choice of the function

$$S_{ij}(\vartheta,V) = \frac{1}{2}\,\frac{B_{ij}(\vartheta,V)}{2V\cos\vartheta\sin\vartheta}\,\frac{m_i^3m_j^3}{m_{ij}^2} \tag{7}$$

with the reduced mass $m_{ij} := (m_im_j)/(m_i+m_j)$ of species i and j [8, p. 67] establishes the connection to a collision operator of the form (3), as follows: First, integrate out $\mathbf{v}'_*$ using the conservation of momentum, which introduces a factor $1/m_j^3$ from the delta function. Second, transform $\mathbf{v}'\in\mathbb{R}^3$ to $\boldsymbol{\lambda} := m_i(\mathbf{v}-\mathbf{v}')\in\mathbb{R}^3$, which gives a factor $1/m_i^3$ from the Jacobian of the transformation. Third, introduce spherical coordinates $(\lambda,\vartheta,\varepsilon)$ for $\boldsymbol{\lambda}$ with polar axis $\mathbf{V}$ and polar angle ϑ, which introduces the standard differential $\mathrm{d}\boldsymbol{\lambda} = \lambda^2\,\mathrm{d}\lambda\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varepsilon$ with $\lambda = |\boldsymbol{\lambda}|\in(0,\infty)$, $\vartheta\in(0,\pi)$, and $\varepsilon\in(0,2\pi)$. Finally, the integral with respect to λ is then evaluated using the delta function for energy conservation [8, p. 65], resulting in the factor $2m_{ij}^2V\cos\vartheta$. The integration over $\vartheta\in(0,\pi)$ and $\varepsilon\in(0,2\pi)$ amounts to an integration over the whole unit sphere; to get Q_{ij} with $\vartheta\in(0,\pi/2)$ as in (3), we have the additional factor 1/2 in S_{ij}; compare [8, pp. 57 and 65].

The applications in microelectronics manufacturing, for which we endeavor to develop a model, involve the flow of gases through chemical reactors. These processes typically use an inert carrier gas that is maintained at a constant (relatively) low pressure throughout, with the (expensive) reactive gases switched on only when desired. While at (relatively) low pressure, the carrier gas is still (at least) an order of magnitude denser than the reactive chemicals. We denote this special species as $i=0$ and the reactive species as $i=1,\ldots,n_s$. Since the inert carrier gas does not react with any other species, we have $R_{ij}^{k\ell}=0$ above, whenever *any* index i, j, k, or ℓ is 0, that is, the summations of $R_{ij}^{k\ell}$ in (2) can actually start at index 1 instead of 0 and are not present at all in the equation for $i=0$. Moreover, since species 0 is (at least) an order of magnitude denser than the other species $j=1,\ldots,n_s$, we can assume that $|Q_{i0}|\gg|Q_{ij}|$ as well as $|Q_{i0}|\gg|R_{ij}^{k\ell}|$ for $j,k,\ell=1,\ldots,n_s$ in all equations. These arguments decouple the equation for $i=0$ from those for $i=1,\ldots,n_s$ in (2). Notice that we also need the inertness of $i=0$ to ensure that it is not involved in the reactive boundary conditions at the wafer surface.

If we consider the flow on the scale of a feature, with its domain close to the wafer surface, we can additionally assume that the flow of the carrier gas is well established and not affected by the introduction of reactive species. Hence, we assume that species $i=0$ is in a spatially homogeneous steady-state during our simulations. Putting all information about $i=0$ together, $f^{(0)}$ satisfies the equation $Q_{00}(f^{(0)},f^{(0)})=0$. All solutions to this equation are in the form of Maxwellians, which we define in the following for all species for future reference after introducing reference quantities appropriate for our applications of interest.

Although we treat flow on small spatial scales in this paper, we retain a focus on coupling these small scale models with models for the reactor scale, which today can be on the scale of a meter in one or more direction. In general, the pressure, temperature, and concentrations of all species vary with both time and position in the reactor. Some processes in use today are operated in transient mode; e.g., the rapid pulses seen in a typical atomic layer deposition process [13–16]. A complete model for such a process would include spatial and temporal variations in the dependent variables of interest. For the purposes of this paper, we assume that conditions at the top of the small feature scale domain (much smaller than the chemical reactor) are spatially uniform. Solutions within a multiscale framework would be required to meaningfully improve upon this approach, for which the present model is intended to supply the small scale part.

In the definition of reference quantities, we therefore wish to only use information that is accessible macroscopically. Using the temperature T in K as one of the operating conditions of the chemical reactor, define the thermal reference speeds for all species $i=0,1,\ldots,n_s$ as

$$v_i^{\mathrm{ref}} = \sqrt{2\frac{k_{\mathrm{B}}}{m_i}T} = \sqrt{2\frac{R_{\mathrm{g}}}{\omega_i}T}, \tag{8}$$

where the universal Boltzmann constant $k_{\mathrm{B}} = 1.3807 \times 10^{-23}$ J/K is related to the universal gas constant $R_{\mathrm{g}} = 8.3145$ J/(K mol) by Avogadro's number $N_{\mathrm{A}} = 6.0221 \times 10^{23}$/mol via $R_{\mathrm{g}} = N_{\mathrm{A}} k_{\mathrm{B}}$. Useful units for R_{g} in our context are also $R_{\mathrm{g}} = 8.3145 \times 10^7$ (g cm^2)/(K mol s^2) = 62,400 (cm^3 Torr)/(K mol). This choice of reference speeds uses the so-called most likely speed for its physical significance. Then we can define Maxwellians for all species $i = 0,1,\ldots,n_{\mathrm{s}}$ by

$$M_i(\mathbf{v}) = \frac{1}{\left[\pi\left(v_i^{\mathrm{ref}}\right)^2\right]^{3/2}} \exp\left(-\frac{|\mathbf{v}-\mathbf{u}|^2}{\left(v_i^{\mathrm{ref}}\right)^2}\right), \tag{9}$$

a form chosen in units of 1/(cm/s)^3, such that $\int M_i \mathrm{d}\mathbf{v} = 1$ in dimensionless form. Here, $\mathbf{u}$ is the bulk velocity of the mixture. It is reasonable to assume that $\mathbf{u} = 0$ in a feature scale model, as we are considering at present. Using the total pressure T_{total} in Torr and the dimensionless mole fractions $0 \leqslant x_i \leqslant 1$ of all species in addition to the temperature T in K, we also define reference concentrations for $i = 0,1,\ldots,n_{\mathrm{s}}$ by

$$c_i^{\mathrm{ref}} = \frac{P_i}{R_{\mathrm{g}}T} = \frac{x_i P_{\mathrm{total}}}{R_{\mathrm{g}}T}, \tag{10}$$

where $P_i = x_i P_{\mathrm{total}}$ denotes the partial pressure of species i. Notice that the mole fractions satisfy $\sum x_i = 1$, hence $\sum P_i = P_{\mathrm{total}}$.

Then, the solution of $Q_{00}(f^{(0)},f^{(0)}) = 0$ is given by $f^{(0)}(\mathbf{x},\mathbf{v},t) = c_0 M_0(\mathbf{v})$ with a constant concentration $c_0 > 0$ in units of mol/cm^3. Putting all information for the reactive species $i = 1,\ldots,n_{\mathrm{s}}$ together, we arrive at the dimensional equations for the KTRM

$$\frac{\partial f^{(i)}}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{x}} f^{(i)} = Q_i\left(f^{(i)}\right), \quad i = 1,\ldots,n_{\mathrm{s}}, \tag{11}$$

where we recognize abstractly that the collision operator on the right-hand side $Q_i(f^{(i)}) := Q_{i0}(f^{(i)}, c_0 M_0)$ is a linear function of $f^{(i)}$, thus justifying its notation. To see this concretely, we write out and transform the linear collision operator $Q_i(f^{(i)})$ a number of times into a form more convenient both for modeling and for the design of the numerical method.

Using its definition $Q_i(f^{(i)}) = Q_{i0}(f^{(i)}, c_0 M_0)$, we have

$$Q_i\left(f^{(i)}\right) = \int_{\mathbb{R}^3}\int_0^{2\pi}\int_0^{\pi/2}\left[f^{(i)'} c_0 M_0(\mathbf{v}'_*) - f^{(i)} c_0 M_0(\mathbf{v}_*)\right] B_{i0}(\vartheta, V)\,\mathrm{d}\vartheta\,\mathrm{d}\varepsilon\,\mathrm{d}\mathbf{v}_*. \tag{12}$$

Using W_{i0} and S_{i0} from (6) and (7) with $j = 0$, we can use the same techniques that explained the connection of the collision operator Q_{ij} and reaction operator $R_{ij}^{k\ell}$ to introduce integrations over $\mathbf{v}'$ and $\mathbf{v}'_*$ to get

$$Q_i\left(f^{(i)}\right) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\left[f^{(i)'} c_0 M_0(\mathbf{v}'_*) - f^{(i)} c_0 M_0(\mathbf{v}_*)\right] W_{i0}\,\mathrm{d}\mathbf{v}_*\,\mathrm{d}\mathbf{v}'\,\mathrm{d}\mathbf{v}'_*. \tag{13}$$

Using the molar units of the collision kernel B_{i0} as cm^3/(mol s) from above, the units of S_{i0} are (cm^2 g^4)/mol. Since the units of a Dirac delta function are the reciprocal of the units of its argument and noting that the delta function of the momentum conservation has a vector-valued argument (a short-hand notation for the product of delta functions for each vector component), the units of W_{i0} are s^5/(mol cm^3).

Generalizing the notation in [8, Chapter IV] to multiple species, we introduce

$$K_i(\mathbf{v}' \to \mathbf{v}) := \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} c_0 M_0(\mathbf{v}'_*) W_{i0}(\mathbf{v},\mathbf{v}_*,\mathbf{v}',\mathbf{v}'_*)\,\mathrm{d}\mathbf{v}_*\,\mathrm{d}\mathbf{v}'_* \tag{14}$$

and

$$v_i(\mathbf{v}) := \int_{\mathbb{R}^3}\int_0^{2\pi}\int_0^{\pi/2} c_0 M_0(\mathbf{v}_*) B_{i0}(\vartheta, V)\,\mathrm{d}\vartheta\,\mathrm{d}\varepsilon\,\mathrm{d}\mathbf{v}_*, \tag{15}$$

and change the order of integrations in (13) and obtain

$$Q_i(f^{(i)}) = \int_{\mathbb{R}^3} K_i(\mathbf{v}' \to \mathbf{v}) f^{(i)}(\mathbf{x}, \mathbf{v}', t)\, d\mathbf{v}' - v_i(\mathbf{v}) f^{(i)}(\mathbf{x}, \mathbf{v}, t). \tag{16}$$

The units of K_i are s^2/cm^3 and of v_i are 1/s. Mass conservation requires that we have [8, p. 169]

$$v_i(\mathbf{v}) = \int_{\mathbb{R}^3} K_i(\mathbf{v} \to \mathbf{v}')\, d\mathbf{v}', \tag{17}$$

and the principle of detailed balance implies that [8, p. 170]

$$K_i(\mathbf{v}' \to \mathbf{v}) M_i(\mathbf{v}') = K_i(\mathbf{v} \to \mathbf{v}') M_i(\mathbf{v}). \tag{18}$$

To understand better why the term M_i would appear here, write out (18) explicitly. From the definition of $K_i(\mathbf{v}' \to \mathbf{v})$, obtain first

$$K_i(\mathbf{v} \to \mathbf{v}') = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} c_0 M_0(\mathbf{v}'_*) W_{i0}(\mathbf{v}', \mathbf{v}_*, \mathbf{v}, \mathbf{v}'_*)\, d\mathbf{v}_*\, d\mathbf{v}'_*. \tag{19}$$

The principle of detailed balance actually guarantees that $W_{ij}(\mathbf{v}, \mathbf{v}_*, \mathbf{v}', \mathbf{v}'_*) = W_{ij}(\mathbf{v}', \mathbf{v}'_*, \mathbf{v}, \mathbf{v}_*)$ [9, p. 217], which gives here

$$K_i(\mathbf{v} \to \mathbf{v}') = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} c_0 M_0(\mathbf{v}'_*) W_{i0}(\mathbf{v}, \mathbf{v}'_*, \mathbf{v}', \mathbf{v}_*)\, d\mathbf{v}_*\, d\mathbf{v}'_*. \tag{20}$$

Interchanging the labels of the dummy integration variables $\mathbf{v}_*$ and $\mathbf{v}'_*$ does not change the value of the double integral that reads now

$$K_i(\mathbf{v} \to \mathbf{v}') = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} c_0 M_0(\mathbf{v}_*) W_{i0}(\mathbf{v}, \mathbf{v}_*, \mathbf{v}', \mathbf{v}'_*)\, d\mathbf{v}_*\, d\mathbf{v}'_*, \tag{21}$$

where we have also interchanged the order of integration. Thus, all terms in $K_i(\mathbf{v}' \to \mathbf{v})$ and $K_i(\mathbf{v} \to \mathbf{v}')$ agree except the Maxwellians $M_0(\mathbf{v}'_*)$ and $M_0(\mathbf{v}_*)$. But we note that $M_i(\mathbf{v}')M_0(\mathbf{v}'_*) = M_i(\mathbf{v})M_0(\mathbf{v}_*)$ holds, since the M_j are Maxwellians and mass conservation holds, thus giving the detailed balance as stated in the form (18).

Taking advantage of the facts resulting from mass conservation and detailed balance, it is natural to introduce the following function:

$$\sigma_i(\mathbf{v}, \mathbf{v}') := \frac{K_i(\mathbf{v}' \to \mathbf{v})}{M_i(\mathbf{v})} \tag{22}$$

which is symmetric $\sigma_i(\mathbf{v}, \mathbf{v}') = \sigma_i(\mathbf{v}', \mathbf{v})$ and positive $\sigma_i(\mathbf{v}, \mathbf{v}') > 0$ and allows us to re-write the linear collision operator in the following form:

$$Q_i(f^{(i)}) = \int_{\mathbb{R}^3} \sigma_i(\mathbf{v}, \mathbf{v}') \left[M_i(\mathbf{v}) f^{(i)}(\mathbf{x}, \mathbf{v}', t) - M_i(\mathbf{v}') f^{(i)}(\mathbf{x}, \mathbf{v}, t) \right] d\mathbf{v}'. \tag{23}$$

In the case of a single reactive species ($n_s = 1$), this form of the linear collision operator agrees with one commonly used form in semiconductor device modeling and allows us to generalize numerical techniques that were originally introduced for the semiconductor Boltzmann equation [7,21,23] to the case of multiple species.

We note that σ_i has units of 1/s and is interpreted as a collision frequency. Hence, the simplest model chooses $\sigma_i = 1/\tau_i$ with the relaxation time τ_i of species $i = 1, \ldots, n_s$. This gives the relaxation time approximation with

$$Q_i(f^{(i)}) = -\frac{1}{\tau_i} \left[f^{(i)}(\mathbf{x}, \mathbf{v}, t) - M_i(\mathbf{v}) \int_{\mathbb{R}^3} f^{(i)}(\mathbf{x}, \mathbf{v}', t)\, d\mathbf{v}' \right], \tag{24}$$

where we used the scaling of the Maxwellian that guarantees $\int M_i\, d\mathbf{v} = 1$. We finally have to define the relaxation time τ_i that characterizes the time scale on which the molecules of species i reach their steady-state distribution (if one is permitted by the model conditions), which we choose to relate to the mean free path λ by the thermal reference speed v_i^{ref} via $\tau_i = \lambda / v_i^{\text{ref}}$. Notice that this relaxation time approximation constitutes a multi-species generalization from [22,24].

2.3. Dimensional boundary and initial conditions

Owing to the hyperbolic character of the Boltzmann equation, we have to supply the inflow velocity components characterized by $\mathbf{n}\cdot\mathbf{v}<0$ of the kinetic density $f^{(i)}(\mathbf{x},\mathbf{v},t)$ for $\mathbf{x}\in\partial\Omega$ for all reactive species $i=1,\ldots,n_s$ in (11), where $\mathbf{n}=\mathbf{n}(\mathbf{x})$ denotes the unit *outward* normal vector at $\mathbf{x}\in\partial\Omega$.

The boundary condition at the solid wafer surface Γ_w is the crucial part of the model that connects the gas flow with the deposition of solid material on the wafer surface. On the length and time scales of our interest, we neglect other effects (e.g., diffusion of molecules on the surface) and model the re-emission as purely Maxwellian in velocity space and thus choose

$$f^{(i)}(\mathbf{x},\mathbf{v},t)=\alpha_i(\mathbf{x},t)M_i(\mathbf{v}),\quad \mathbf{x}\in\Gamma_w,\quad \mathbf{n}\cdot\mathbf{v}<0. \tag{25}$$

We now enforce the conservation of molecules by requiring, for every reactive species, the flux into the gas domain from the surface to equal the flux of that species to the surface η_i plus the molar species generation rate (per area) r_i of that species due to chemical reactions. This can be written as

$$\int_{\mathbf{n}\cdot\mathbf{v}<0}|\mathbf{n}\cdot\mathbf{v}|f^{(i)}(\mathbf{x},\mathbf{v},t)\,d\mathbf{v}=\eta_i(\mathbf{x},t)+r_i(\mathbf{x},t), \tag{26}$$

where $\eta_i(\mathbf{x},t)=\int_{\mathbf{n}\cdot\mathbf{v}'>0}|\mathbf{n}\cdot\mathbf{v}'|f^{(i)}(\mathbf{x},\mathbf{v}',t)\,d\mathbf{v}'$ denotes the flux to the surface of species i. The reaction model that supplies the formulas for r_i, $i=1,\ldots,n_s$, can be any general, non-linear model, giving the KTRM the flexibility to use any reaction chemistry desired. Note that the flux η_i has units of mol/(s cm^2) because of our use of molar units for $f^{(i)}$, which agrees with the units that the species generation rate r_i has in process engineering; this explains our choice of molar units in the KTRM.

To obtain the final form of the boundary condition at Γ_w, insert the $f^{(i)}$ from (25) into the integral on the left-hand side of (26) to determine α_i. The boundary condition at the wafer surface then reads

$$f^{(i)}(\mathbf{x},\mathbf{v},t)=C_i[\eta_i(\mathbf{x},t)+r_i(\mathbf{x},t)]M_i(\mathbf{v}),\qquad \mathbf{x}\in\Gamma_w,\quad \mathbf{n}\cdot\mathbf{v}<0 \tag{27}$$

with the scaling factor $C_i=2\sqrt{\pi}/v_i^{\mathrm{ref}}$ in units of 1/(cm/s).

At the top of the domain Γ_t that constitutes the interface to the bulk gas phase of the chemical reactor, we assume that also the reactive chemicals are in Maxwellian form, and we use the boundary condition

$$f^{(i)}(\mathbf{x},\mathbf{v},t)=c_i^{\mathrm{top}}(\mathbf{x},t)M_i(\mathbf{v}),\qquad \mathbf{x}\in\Gamma_t,\quad \mathbf{n}\cdot\mathbf{v}<0. \tag{28}$$

Notice that the form of this boundary condition is not limited by the model but chosen because of its reasonableness for a micron-scale model; it essentially represents an infinite domain in the directions along the surface.

At the remaining parts of the boundary, the vertical sides of the domain Ω collectively labeled Γ_s, we use specular reflection

$$f^{(i)}(\mathbf{x},\mathbf{v},t)=f^{(i)}(\mathbf{x},\mathbf{v}',t),\qquad \mathbf{x}\in\Gamma_s,\quad \mathbf{n}\cdot\mathbf{v}<0 \tag{29}$$

with $\mathbf{v}=\mathbf{v}'-2(\mathbf{v}'\cdot\mathbf{n})\mathbf{n}$. This is a convenient and reasonable boundary condition for a micron-scale model.

Finally, we assume that the initial distributions of the reactive species are known; while this is not material to our model, we assume at present that it is in Maxwellian form given by

$$f^{(i)}(\mathbf{x},\mathbf{v},t)=c_i^{\mathrm{ini}}(\mathbf{x})M_i(\mathbf{v})\quad \text{at } t=0. \tag{30}$$

Specifically, we often choose $c_i^{\mathrm{ini}}\equiv 0$ to model the case, where the reactive chemicals are not present in the domain Ω at the beginning of the simulation.

2.4. Derivation of the dimensionless model

Up to this point, we have purposefully kept the equations in dimensional form in order to introduce the model parameters for our applications of interest with units. Now, we introduce reference quantities to non-dimensionalize the system of Boltzmann equations (11) with linear collision operators (23). We start by choosing the reference speed $v^*:=v_1^{\mathrm{ref}}$ from (8) and the reference concentration $c^*:=c_1^{\mathrm{ref}}$ from (10) based

on species $i = 1$. The transport on the left-hand side of (11) needs to be non-dimensionalized with respect to the typical scales for transport [1], which are the length scale of the domain L^* and correspondingly the time $t^* := L^*/v^*$. The collision operators (23) on the right-hand side of (11) need to be non-dimensionalized with respect to the typical scales for collisions, which are the mean free path λ and the corresponding time $\tau^* := \lambda/v^*$. For convenience, also introduce the notation $f^* := c^*/(v^*)^3$.

Using these reference quantities, we define the dimensionless independent variables $\hat{\mathbf{x}} := \mathbf{x}/L^*$, $\hat{\mathbf{v}} := \mathbf{v}/v^*$, and $\hat{t} := t/t^*$. To non-dimensionalize the dependent variable, introduce $\hat{f}^{(i)} := f^{(i)}/f^* = (v^*)^3 f^{(i)}/c^*$. The dimensionless Maxwellians are given by $\hat{M}_i := (v^*)^3 M_i = \exp(-|\hat{\mathbf{v}}|^2/(\hat{v}_i^{\mathrm{ref}})^2)/[\pi(\hat{v}_i^{\mathrm{ref}})^2]^{3/2}$ with the dimensionless thermal reference speeds $\hat{v}_i^{\mathrm{ref}} := v_i^{\mathrm{ref}}/v^*$. The collision frequencies σ_i in (23) have to be non-dimensionalized with respect to collisions by choosing $\hat{\sigma}_i := \sigma_i \tau^*$. For the relaxation time approximation in (24), we need to introduce the dimensionless relaxation times $\hat{\tau}_i := \tau_i/\tau^*$, which implies that $\hat{\tau}_i = v^*/v_i^{\mathrm{ref}} = 1/\hat{v}_i^{\mathrm{ref}}$, involving the dimensionless thermal reference speed $\hat{v}_i^{\mathrm{ref}}$. Using these definitions, (11) with (23) becomes

$$\frac{f^*}{t^*}\frac{\partial \hat{f}^{(i)}}{\partial \hat{t}} + \frac{f^* v^*}{L^*}\hat{\mathbf{v}} \cdot \nabla_{\hat{\mathbf{x}}} \hat{f}^{(i)} = \frac{(v^*)^3 f^*}{(v^*)^3 \tau^*} \int_{\mathbb{R}^3} \hat{\sigma}_i \left[\hat{M}_i \hat{f}^{(i)'} - \hat{M}_i' \hat{f}^{(i)}\right] \mathrm{d}\hat{\mathbf{v}}'. \tag{31}$$

Noting that $v^*/L^* = 1/t^*$ and simplifying results in

$$\frac{\partial \hat{f}^{(i)}}{\partial \hat{t}} + \hat{\mathbf{v}} \cdot \nabla_{\hat{\mathbf{x}}} \hat{f}^{(i)} = \frac{1}{\mathrm{Kn}} \int_{\mathbb{R}^3} \hat{\sigma}_i \left[\hat{M}_i \hat{f}^{(i)'} - \hat{M}_i' \hat{f}^{(i)}\right] \mathrm{d}\hat{\mathbf{v}}', \tag{32}$$

where the Knudsen number $\mathrm{Kn} := \lambda/L^* = \tau^*/t^*$ has been introduced. We recognize how the Knudsen number emerges naturally as the relevant dimensionless group that quantifies the relative importance of intermolecular and molecular-wall collisions. Notice that the present results agrees with the dimensionless system of linear Boltzmann equations in (1), with the hats dropped.

To transform the boundary and initial conditions to dimensionless form, we choose the reference flux $\eta^* := c^* v^*$ to define the dimensionless fluxes $\hat{\eta}_i := \eta_i/\eta^*$ as well as the dimensionless species generation rates $\hat{r}_i := r_i/\eta^*$. Introducing also the transformation $\hat{C}_i := v^* C_i$, the boundary and initial conditions have the same dimensionless form as the dimensional equations stated above.

We continue in dimensionless form and drop the hat notation for simplicity.

3. The numerical method

The numerical method for the system of Boltzmann equations (1) needs to discretize the 3-D spatial domain $\Omega \subset \mathbb{R}^3$ and the 3-D (unbounded) velocity space $\mathbb{R}^3$ for full 3-D/3-D simulations. The KTRM in (1) with its linear collision operator allows for the application of moment methods following [22–24]. In this approach, we start by discretizing in velocity space by approximating each $f^{(i)}(\mathbf{x}, \mathbf{v}, t)$ by an expansion $f_K^{(i)}(\mathbf{x}, \mathbf{v}, t) := \sum_{\ell=0}^{K-1} f_\ell^{(i)}(\mathbf{x}, t)\varphi_\ell(\mathbf{v})$. The classical basis functions in this approach are products of a Maxwellian and Hermite polynomials in each dimension. Each linear Boltzmann equation in (1) is discretized by inserting $f_K^{(i)}(\mathbf{x}, \mathbf{v}, t)$ for $f^{(i)}(\mathbf{x}, \mathbf{v}, t)$ and testing against the K basis functions $\varphi_k(\mathbf{v})$, resulting in a system of K transient linear hyperbolic transport equations

$$\frac{\partial F^{(i)}}{\partial t} + A^{(1)}\frac{\partial F^{(i)}}{\partial x_1} + A^{(2)}\frac{\partial F^{(i)}}{\partial x_2} + A^{(3)}\frac{\partial F^{(i)}}{\partial x_3} = \frac{1}{\mathrm{Kn}} B^{(i)} F^{(i)} \tag{33}$$

in space $\mathbf{x} = (x_1, x_2, x_3)^{\mathrm{T}}$ and time t for the vector of K coefficient functions $F^{(i)}(\mathbf{x}, t) := (f_0^{(i)}(\mathbf{x}, t), \ldots, f_{K-1}^{(i)}(\mathbf{x}, t))^{\mathrm{T}}$ for each reactive species $i = 1, \ldots, n_s$. Here, the $K \times K$ matrices $A^{(1)}$, $A^{(2)}$, $A^{(3)}$, and $B^{(i)}$ are constant owing to the linearity of (1). The use of the classical basis functions results in dense coefficient matrices $A^{(1)}$, $A^{(2)}$, and $A^{(3)}$ that can in turn be diagonalized simultaneously [23]. Based on this observation, we design collocation basis functions $\varphi_\ell(\mathbf{v})$ in the expansion $f_K^{(i)}(\mathbf{x}, \mathbf{v}, t)$ that give rise to *diagonal* coefficient matrices in (33) directly [17,25]. Recall that the solutions for all reactive species are coupled through the boundary condition at the wafer surface that is crucial for the applications under consideration. Our formulation has the advantage that boundary and initial conditions can be evaluated more directly without a transformation; this makes no difference in one dimension but the setup of simulations in higher dimensions becomes easier in practice.

Since the hyperbolic system in (33) can be transformed to the one based on classical basis functions [17,25], we can apply the analytic results in [23]. More specifically, under reasonable assumptions on the problem, there are available the stability result $\|f_K\|_{G(t)} \leqslant C$ with problem-dependent constant $C < \infty$ and the asymptotic convergence result $\|f - f_K\|_{G(t)} \to 0$ for all times t as the number of expansion coefficients $K \to \infty$, using the norm $\|f\|_{G(t)} := (\iint f^2/M(\mathbf{v})\,d\mathbf{v}\,d\mathbf{x})^{1/2}$ induced by properties of the linear Boltzmann equation [23]. Here, we are dropping the species index and quote the theory only for the single-species case. Extending the 1-D numerical results in [23,24] for the semiconductor Boltzmann equation, numerical demonstrations for representative simulations on 2-D and 3-D trench domains for a relevant range of Knudsen numbers in Table 1 confirm that convergence can be achieved using reasonable numbers of discrete velocities in each dimension, for which transient studies on adequate spatial meshes are feasible; see [17,25] for more details on the theoretical results and the stability and convergence studies. Therefore, the studies in this paper use $K = 16 \times 16 = 256$ for two-dimensional velocity space and $K = 4 \times 4 \times 4 = 64$ for three-dimensional velocity space.

The hyperbolic system (33) is now posed in a standard form as a system of partial differential equations on the spatial domain $\Omega \subset \mathbb{R}^3$ and in time t and amenable to solution by various methods. Since we are interested in discretizing spatial domains Ω with potentially irregular shapes, such as the one in Fig. 1, the discontinuous Galerkin method (DGM) is convenient. It was first in fact introduced in [19] for solving the neutron transport equation, an example of a scalar linear Boltzmann equation. More recently, the DGM in the code DG [20] has been used to solve systems with a *small number* of *non-linear* conservation laws (the Euler equations in two and three dimensions) [10,20]. We use this implementation of the DGM here by extending it to solve the system (33) with a *large number* of *linear* conservation laws. This implementation has both triangular and quadrilat-

Table 1
Stability and convergence studies for the velocity discretization as function of the number of expansion coefficients K, for selected Knudsen numbers

K	$\|f_K\|_{G(t)}$	$\|f\|_{G(t)} - \|f_K\|_{G(t)}$	$\|f - f_K\|_{G(t)}$
(a) 2-D trench domains			
Kn = 0.01			
4	0.47861	6.86e−03	8.73e−03
16	0.48252	2.95e−03	3.69e−03
64	0.48425	1.22e−03	1.51e−03
256	0.48508	3.90e−04	4.94e−04
Kn = 1.0			
4	0.52021	5.28e−03	9.31e−03
16	0.52299	2.51e−03	4.70e−03
64	0.52455	9.20e−04	4.63e−03
256	0.52522	2.50e−04	2.88e−03
Kn = ∞			
4	0.52306	6.20e−03	1.12e−02
16	0.52599	3.27e−03	8.78e−03
64	0.52805	1.21e−03	6.63e−03
256	0.52891	3.50e−04	4.34e−03
(b) 3-D trench domains			
Kn = 0.01			
8	0.61551	4.84e−03	5.89e−03
64	0.61848	1.87e−03	1.79e−03
Kn = 1.0			
8	0.65713	1.94e−03	8.43e−03
64	0.65869	3.80e−04	5.09e−03
Kn = ∞			
8	0.65956	3.58e−03	1.01e−02
64	0.66142	1.72e−03	6.75e−03

eral elements available in two and three dimensions. For the two-dimensional domain with regular shape in Section 4.1, we use quadrilateral elements. For the three-dimensional domain in Fig. 1 in Section 4.2, we use tetragonal elements. In both cases, we use linear finite elements, which are second-order accurate in their natural L^2-norm [2]. Currently, we use explicit, first-order accurate Euler time-stepping because of its memory efficiency and cheap cost per time step. Because the spatial and time discretizations are well known and we are using the established software package DG, we do not provide numerical demonstrations of their stability and convergence.

The degrees of freedom (DOF) of the finite element method are the values of the n_s species' coefficient functions $f_\ell^{(i)}(\mathbf{x},t)$ in the spectral Galerkin expansion at K discrete velocities on the vertices of each of the N_e elements of the DGM: Since quadrilateral elements in two dimensions and tetragonal elements in three dimensions both have four vertices per element, this happens to give the same formula for the number of degrees of freedom $\mathrm{DOF} = 4N_e n_s K$ for both 2-D/2-D and 3-D/3-D simulations here. This is the complexity of the computational problem that needs to be solved at every time step. For the two-dimensional studies in Section 4.1, we use a modest spatial mesh with $N_e = 16 \times 8 = 128$ elements; for the velocity discretization with $K = 256$, this implies $4 \times 128 \times 256 = 131{,}072$ degrees of freedom in a single-species simulation ($n_s = 1$) at every time step. For the three-dimensional studies in Section 4.2, the mesh of the domain in Fig. 1 uses $N_e = 7087$ elements; even in the case of a single-species model ($n_s = 1$) and using just $K = 4 \times 4 \times 4 = 64$ discrete velocities, the total DOF are $N = 1{,}814{,}272$ or nearly 2 million unknowns to be determined at every time step.

The size of problem at every time step motivates our interest in parallel computing for this problem. For the parallel computations on a distributed-memory cluster, the spatial domain Ω is partitioned in a pre-processing step, and the disjoint subdomains are distributed to separate parallel processes. The discontinuous Galerkin method for (33) needs the flux through the element faces. At the interface from one subdomain to the next, communications are required among those pairs of parallel processes that share a subdomain boundary. DG uses MPI for its parallel communications. If the number of DOF are large, parallel performance results show excellent efficiency for up to 64 processors on our cluster with high-performance Myrinet interconnect, while for more moderate number of DOF, such as in the two-dimensional studies here, the efficiency drops off for more than 16 processors [12,17,25]. Therefore, we typically use 16 processors in production runs.

4. Simulation results

To demonstrate the capabilities of the KTRM, we choose chemical vapor deposition (CVD) as an important example of a deposition process in microelectronics manufacturing. In this process, gaseous chemicals are supplied from the gas-phase interface at the top of the domain, for instance in Fig. 1, through the x_1–x_2-plane at $x_3 = 0.3$. The gaseous chemicals flow downwards throughout the domain Ω until they reach the solid wafer surface, where some of the molecules react to form a solid deposit. Deposition processes in or over sub-micron features on a wafer are critical to the fabrication of integrated circuits. Thin deposited films (a few nanometers thick) are used for several purposes in features such as that shown in Fig. 1. An important aspect of such thin films is that they need to coat the internal surface of the feature with uniform thickness. This is called high conformality or step-coverage. Thicker film deposits are used to fill features; e.g., with copper to form wires that electrically connect the devices to form logic, and get signals off the chip. The work here is very relevant to understanding the thin film deposition processes. What is really needed is an understanding of the relative rates of growth around the features, for the feature shape as given. For thicker depositions, the surface needs to be updated and the governing equations resolved. EVOLVE [5] and PLENTE [4] are two codes that perform physics-based simulations of such processes. Like all similar (topography) codes, they use pseudo-steady models for collisionless transport (for low pressures), and are not used to study the rapid transients in species fluxes that can be followed using the KTRM. For more details, see [4,5] and the references therein.

To focus on studying the interplay of transport and surface reactions, we use a single-species model with $n_s = 1$ reactive species. Hence, we drop the species superscripts in $f^{(1)}$ as well as subscripts in c_1, M_1, η_1, r_1, σ_1, and τ_1 in the following. Recall that we are solving the dimensionless problem (1) here and all quantities are dimensionless (with hats dropped), except explicitly re-dimensionalized quantities such as the domain Ω and time t. For clarity, we consider chemical reaction rates that can be considered linear in the species flux;

i.e., we use 'sticking factor' based chemical reaction rates that take the reaction rate R as proportional to the flux to the surface: $R(\mathbf{x},t) = \gamma_0 \eta(\mathbf{x},t)$ for $\mathbf{x} \in \Gamma_w$; here, the proportionality constant $0 \leqslant \gamma_0 \leqslant 1$ denotes the sticking factor that represents the fraction of molecules that are modeled to deposit at (''stick to'') the wafer surface. The species generation rate in the relevant boundary condition is then given by $r(\mathbf{x},t) = -R(\mathbf{x},t)$. This reaction model yields a boundary condition for the inflowing components of f, proportional to the fraction of molecules *not* 'sticking' $(1-\gamma_0)$, proportional to the flux to the surface η, and with Maxwellian re-emission of molecules.

We focus on how the flow behaves when starting from no gas present throughout Ω, modeled by initial condition $f \equiv 0$ at $t = 0$. The boundary condition at the top of the domain Γ_t, modeling the interface to the bulk gas, is $f(\mathbf{x},\mathbf{v},t) = c^{\text{top}} M(\mathbf{v})$ with $c^{\text{top}} \equiv 1$ for the inflowing velocity components $\mathbf{n}\cdot\mathbf{v} < 0$. The collision operator uses a relaxation time approximation by choosing $\sigma(\mathbf{v},\mathbf{v}') \equiv 1/\tau$ with the (dimensionless) relaxation time $\tau = 1.0$ for a single-species model [22,24].

4.1. Reference studies for a flat wafer surface

First, we present reference studies to analyze the interplay between sticking factor γ_0 and the pressure regime; the latter is equivalently determined by selecting the Knudsen number Kn, which controls the relative importance of transport and collisions in (1). A small Kn corresponds to higher pressure, while a large value corresponds to low pressure. For instance, the choice of $\text{Kn} = 0.01$ gives flow in the near-hydrodynamic regime and $\text{Kn} = 0.1$ denotes flow in the transition regime, while $\text{Kn} \to \infty$ leads to a cancellation of the collision terms on the right-hand side of (1) and models free molecular flow in the Knudsen regime; see [18, p. 50]. To focus only on the interplay of the parameters γ_0 and Kn, we choose a flat wafer, so as to suppress any effect that the geometry of the domain has on the flow. For simplicity and easier depiction of the results, we use a two-dimensional domain. Concretely, the domain is chosen as $\Omega := (-0.25, 0.25) \times (0.0, 0.25)$ in units of μm; since we choose the reference length as 1 μm, the dimensionless domain is given by the same numerical values. Here, the first component x_1 of $\mathbf{x} \in \Omega$ counts along the wafer surface, and the second component x_2 orthogonal to it, with $x_2 = 0.0$ denoting the flat wafer surface.

The transient studies were run sufficiently long in each case to approximate steady-state, which takes longer for smaller Knudsen number than for large values. In Figs. 2–4, these steady-state results are shown for 12 cases, namely in four rows for the sticking factors $\gamma_0 = 1.0$, 0.1, 0.01, and 0.0 and in three columns for the Knudsen numbers $\text{Kn} = 0.01$, 0.1, and ∞. Three quantities are plotted for each steady-state result:

- Fig. 2: the dimensionless concentration $0 \leqslant c(\mathbf{x},t) \leqslant 1$ as functions of $\mathbf{x} \in \Omega$.
- Fig. 3: the kinetic density $f(\mathbf{x},\mathbf{v},t) \geqslant 0$ as functions of $\mathbf{v} \in \mathbb{R}^2$ at point $\mathbf{x} = (0.0625, 0.0625)$.
- Fig. 4: the saturation of the kinetic density $0 \leqslant f(\mathbf{x},\mathbf{v},t)/M(\mathbf{v}) \leqslant 1$ as functions of $\mathbf{v} \in \mathbb{R}^2$ at point $\mathbf{x} = (0.0625, 0.0625)$.

Each individual plot is a mesh plot, in which the quantity is plotted in the vertical direction over the two-dimensional domain of the independent variable, being $\mathbf{x} \in \Omega$ for $c(\mathbf{x},t)$ and $\mathbf{v} \in \mathbb{R}^2$ for $f(\mathbf{x},\mathbf{v},t)$ and $f(\mathbf{x},\mathbf{v},t)/M(\mathbf{v})$. The fact that the Maxwellian $M(\mathbf{v})$ is the steady-state limit for the solution $f(\mathbf{x},\mathbf{v},t)$ for the model with the boundary conditions considered here and in the absence of reactions motivates the definition of the kinetic saturation as $f(\mathbf{x},\mathbf{v},t)/M(\mathbf{v})$. Since additionally the values of f are never larger than M in our simulations, we have the bound for the saturation $f/M \leqslant 1$ and for the (dimensionless) concentration $c = \int f \,\mathrm{d}\mathbf{v} \leqslant \int M \,\mathrm{d}\mathbf{v} = 1$. The spatial point $\mathbf{x}$ at which f and f/M are plotted is chosen $x_2 = 0.0625$ μm above the wafer surface. Its x_1-coordinate is chosen for convenience, as the results are identical at all values of x_1. Notice that the first and last choice of γ_0 are particularly useful for demonstration purposes: $\gamma_0 = 1.0$ means that all molecules deposit on the wafer surface, meaning that *no* molecules re-emit from the surface; this is seen in some widely used processes, and so has practical relevance. And $\gamma_0 = 0.0$ implies that no deposition takes place and *all* molecules re-emit from the surface, with Maxwellian velocity distribution.

We start by interpreting the *upper right corner* for the steady-state results in Figs. 2–4 for the demonstration case of $\gamma_0 = 1.0$ and $\text{Kn} = \infty$. This choice of Knudsen number means that no collisions take place among the molecules. Additionally, the choice of sticking factor implies that no molecules re-emit from the surface.

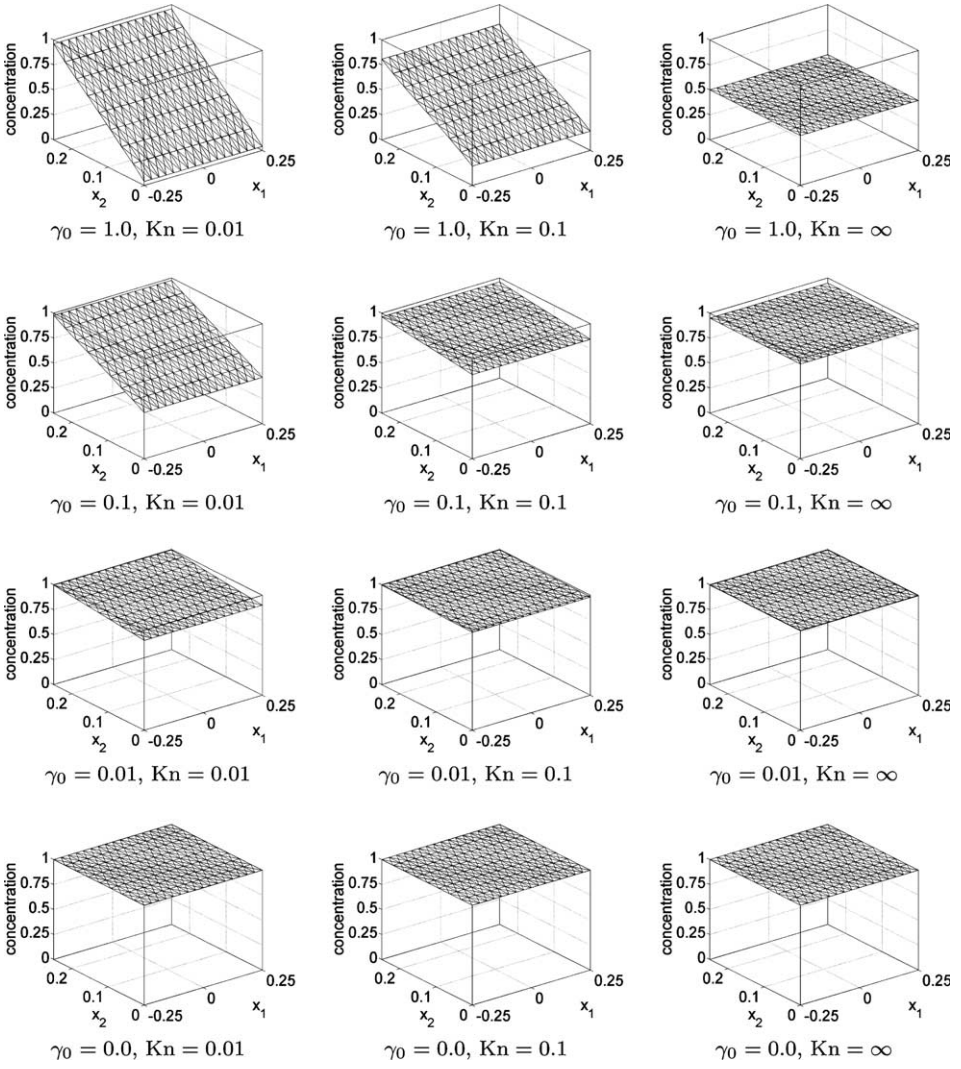

Fig. 2. Concentration $c(\mathbf{x}, t)$, for selected Knudsen numbers Kn and sticking factors γ_0, as function of $\mathbf{x}$ in the two-dimensional domain with flat wafer surface at steady-state.

Starting from the initial condition of zero throughout the domain, molecules flow in from the boundary at the top of the domain $x_2 = 0.25$. Since there are no collisions, the kinetic density f from the boundary condition is translated through the spatial domain convectively until the wafer surface at $x_2 = 0.0$ is reached, where all molecules get consumed. This results in the spatially uniform steady-state concentration in Fig. 2 for this case. Notice that the boundary condition at the top of the domain feeds only downward components ($v_2 < 0$) of the kinetic density. Hence, also only those components are non-zero at steady-state, since neither collisions nor re-emission from the wafer surface contribute to the other components with $v_2 > 0$. This is born out most clearly in Figs. 3 and 4 which confirm the zero value for the $v_2 > 0$ components of f for the chosen point $\mathbf{x}$. For $v_2 < 0$, Fig. 3 only shows that $f > 0$, but due to widely varying scales of different components, their exact values cannot be assessed; this is the value of the saturation plots in Fig. 4, which in fact shows that all components of f for $v_2 < 0$ are fully saturated at their respective maximum values. In turn, since the components of f are zero for $v_2 > 0$ and fully saturated for $v_2 < 0$ *at all points* $\mathbf{x} \in \Omega$, we can explain the value of 0.5 for the concentration c throughout the domain for this case of $\gamma_0 = 1.0$ and $\mathrm{Kn} = \infty$, seen in Fig. 2.

We continue by interpreting and comparing the results in the *right column* for the steady-state results in Figs. 2–4 that all have $\mathrm{Kn} = \infty$ and contrast the cases of different sticking factors γ_0. The case $\gamma_0 = 1.0$ showed

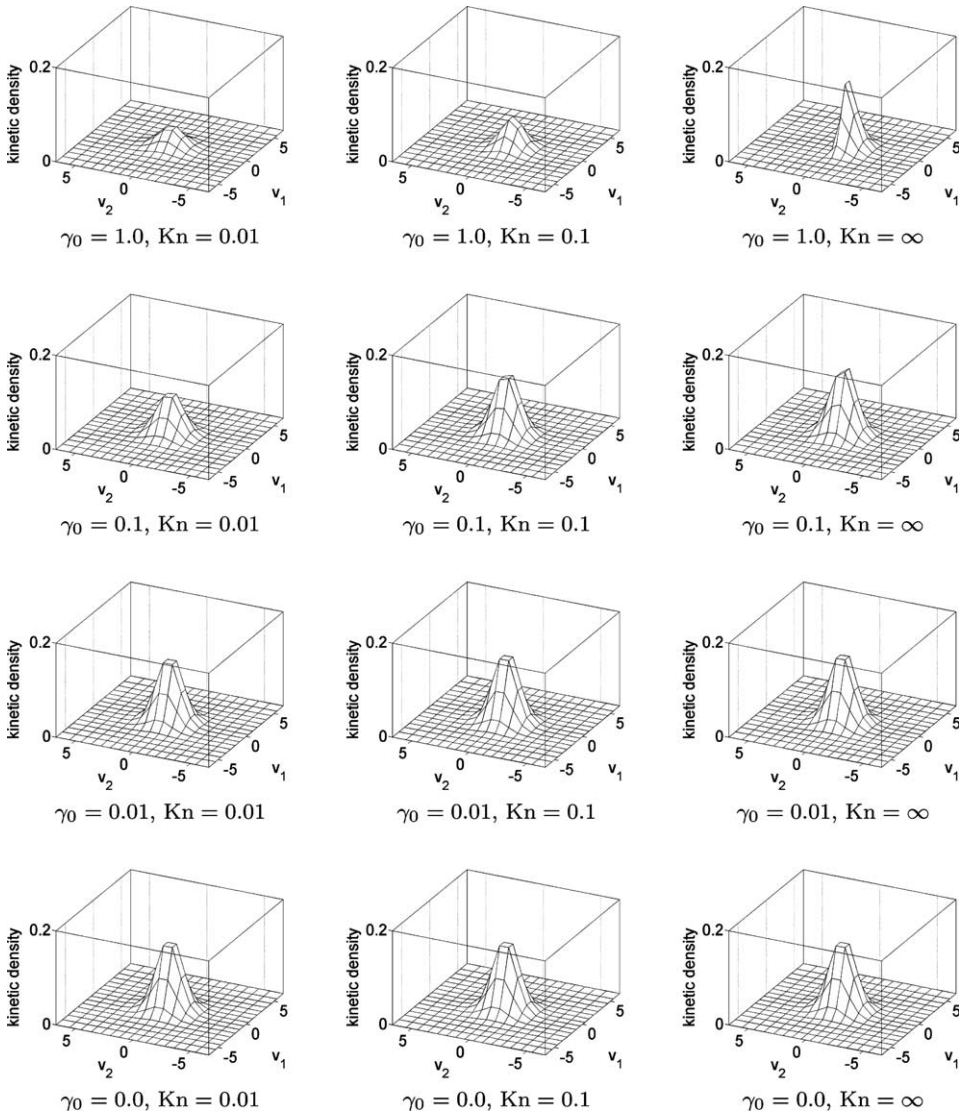

Fig. 3. Kinetic density $f(\mathbf{x},\mathbf{v},t)$, for selected Knudsen numbers Kn and sticking factors γ_0, as function of $\mathbf{v}\in\mathbb{R}^2$ at $\mathbf{x}=(0.0625,0.0625)$ in the two-dimensional domain with flat wafer surface at steady-state.

that the $v_2>0$ components of f remain zero at steady-state, as no molecules re-emit from the surface. Compared to that, the re-emitting molecules for $\gamma_0<1.0$ generate non-zero components of f for $v_2>0$, as can be seen in Fig. 3. To see the values of the f components more precisely, consider the saturation plots in Fig. 4 that bear out that the downward components for $v_2<0$, unaffected by any collisions for $\mathrm{Kn}=\infty$, still attain full saturation at steady-state, while the values of the upward components for $v_2>0$ are proportional to $1-\gamma_0$ (most clearly visible for $\gamma_0=0.1$ in the second row). The results for f and f/M are for a particular point $\mathbf{x}$, but since there are no collision effects that flow attains the same steady-state value for all $\mathbf{x}\in\Omega$ after sufficient time. Therefore, the concentrations c in each plot in the right column of Fig. 2 are spatially uniform. Their values are larger, the more molecules are re-emitted from the wafer surface, as born out by the plots. The extreme case is for both concentrations and saturation to reach the value 1.0 uniformly throughout their domains in the case where all molecules are re-emitted from the wafer surface for $\gamma_0=0.0$ in the last row.

Next, consider the *first row* for the steady-state results in Figs. 2–4 now, which share the sticking factor $\gamma_0=1.0$ and contrast the effect of collisions controlled by the Knudsen number. For the collisionless flow in the upper right corner the upward components of f remained zero, since there was neither collisions nor re-emis-

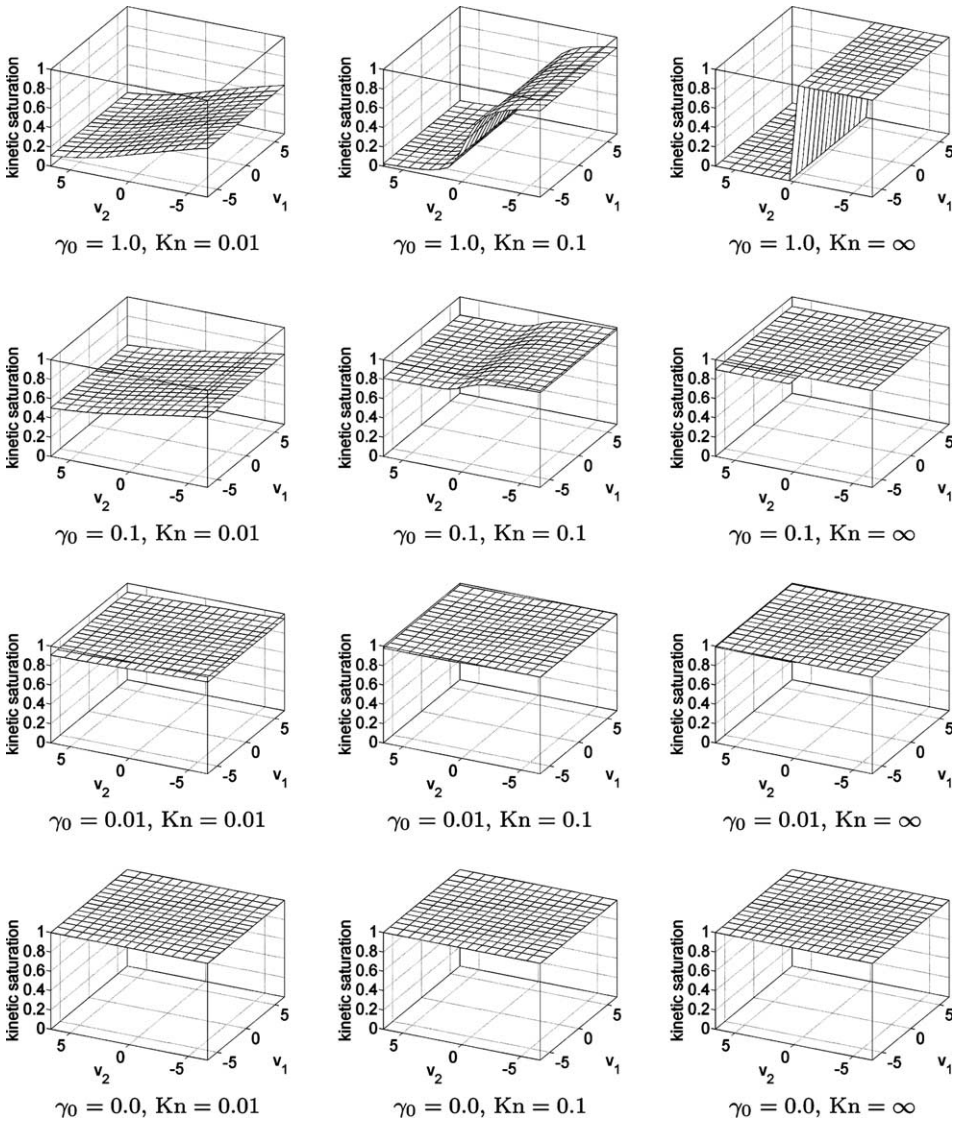

Fig. 4. Saturation of kinetic density $f(\mathbf{x},\mathbf{v},t)/M(\mathbf{v})$, for selected Knudsen numbers Kn and sticking factors γ_0, as function of $\mathbf{v} \in \mathbb{R}^2$ at $\mathbf{x} = (0.0625, 0.0625)$ in the two-dimensional domain with flat wafer surface at steady-state.

sion from the wafer surface to contribute to them. Here, now, $\gamma_0 = 1.0$ implies that there are still no molecules re-emitting from the wafer surface, but as Kn gets smaller, collisions become progressively more dominant and we see non-zero f components for $v_2 > 0$ in Fig. 3. The saturation plots in Fig. 4 clearly show the smoothing out that results from progressively more collisions among the molecules, as more and more molecules collide to attain upward velocities as Kn decreases. To understand the concentration plots in Fig. 2, recall that all molecules are consumed at the wafer surface at $x_2 = 0.0$, hence the concentration must decrease from the top of the domain to the wafer surface. This explains the spatial dependence seen in the plots. Moreover, since the upward components of f are solely fed by the collisions (as no molecules re-emit from the surface), this effect depends on the location of the point $\mathbf{x}$, with more contribution to all components of f the closer to the inflow at the top and the more collisions (smaller Kn). This leads in turn to larger values of c at those locations, with the opposite effect the closer to the wafer surface at which molecules are consumed.

Consider now the *last row* of steady-state results in Figs. 2–4, which share the sticking factor $\gamma_0 = 0.0$. For collisionless transport with $\mathrm{Kn} = \infty$, a spatially uniform steady-state with full saturation of f was attained at

all points $\mathbf{x} \in \Omega$, as the molecules flow from the top of the domain at $x_2 = 0.25$ convectively without collisions to the wafer surface at $x_0 = 0.0$, where all of them get re-emitted. After sufficiently long time, this leads to the saturation also of the upward components of f everywhere in Ω and a uniform concentration of 1.0. For $\mathrm{Kn} < \infty$ in the last row, collisions smooth out the kinetic density f from the start, leading to smaller numbers of molecules reaching the wafer surface than for $\mathrm{Kn} = \infty$ at a fixed time. But after sufficiently long time, and the smaller the Kn the longer time it takes, sufficiently many molecules have reached the wafer surface and have been re-emitted from there to saturate all components of f, and we see that the case of $\gamma_0 = 0.0$ admits a uniform Maxwellian steady-state solution. This is exhibited in the plots of $c = 1.0$ in Fig. 2 and full saturation in Fig. 4.

It remains to discuss the *left and center columns* for the *second and third rows* in Figs. 2–4: In those cases, we have a combination of the effects of collisions and increasing re-emission of molecules from the wafer surface. Comparing the concentration plots in Fig. 2 for fixed Knudsen number shows the higher steady-state values attainable throughout the domain if less molecules are consumed. This reaches the extreme case of all molecules being re-emitted for $\gamma_0 = 0.0$ in the last row, where at steady-state the concentration attains its maximum value of 1.0 throughout the domain Ω. Fig. 3 bears out that all components of f are non-zero for collisional regimes with $\mathrm{Kn} < \infty$, while Fig. 4 shows in more detail that collisions result in smoothing the kinetic density as Kn decreases; compare the plots in the second row for $\gamma_0 = 0.1$. This tends to the extreme case, where no molecules are consumed at the surface and then smoothed out by collisions, resulting in full saturation at steady-state independent of the Knudsen number.

In summary, the concentration results in Fig. 2 show that the steady-state concentrations form a decreasing function from the top of the domain to the wafer surface, with the value at the wafer surface depending on the fraction of molecules that re-emit. But as γ_0 gets closer to 1.0 and a smaller fraction of molecules are re-emitted from the wafer surface, we see the effect of the transport regime characterized by the Knudsen number: the larger Kn, the less interaction there is between the different velocity components in f, such that some of the components remain zero in the extreme case of $\mathrm{Kn} = \infty$, thus limiting the concentration values attainable at steady-state. This behavior can only be explained in detail by considering the kinetic density f as a function of its velocity arguments in Figs. 3 and 4. This demonstrates the value of being able to plot the kinetic density as a function of its velocity arguments, because it provides more insight into the structure of the solution beyond the macroscopic concentration as a function of space.

4.2. Application results for a three-dimensional irregular wafer surface

The results in Figs. 5–7 are designed to bring out the capabilities of the KTRM to simulate the *transient* behavior of flow in a general domain Ω with irregular shape, such as the example shown in Fig. 1.

We continue to use the single-species ($n_s = 1$) sticking factor-based CVD model from the previous section and contrast the results for three different Knudsen numbers $\mathrm{Kn} = 0.01$, 0.1, and 1.0. The studies shown use a sticking factor of $\gamma_0 = 0.01$, that is, 99% of all molecules re-emit from the surface. The results show the following quantities:

- Fig. 5: the dimensionless concentration $0 \leqslant c(\mathbf{x}, t) \leqslant 1$ as functions of $\mathbf{x} \in \Omega$.
- Fig. 6: the kinetic density $f(\mathbf{x}, \mathbf{v}, t) \geqslant 0$ as functions of $\mathbf{v} \in \mathbb{R}^3$ at the mouth of the trench at $\mathbf{x} = (0.5, 0.5, 0.0)$.
- Fig. 7: the saturation of the kinetic density $0 \leqslant f(\mathbf{x}, \mathbf{v}, t)/M(\mathbf{v}) \leqslant 1$ for $(v_1, v_2) = (0, 0)$ as functions of $v_3 \in \mathbb{R}$ at mouth of the trench at $\mathbf{x} = (0.5, 0.5, 0.0)$ and at mouth of the via at $\mathbf{x} = (0.5, 0.5, -0.3)$.

For these transient studies, the figures show the results for the case of each Kn column-wise, indicated at the top of the column, with each column showing the quantity at four different points in time, in re-dimensionalized units at $t = 5$, 10, 15, and 20 ns. These values are based on a reference time $t^* = 5$ ns corresponding to a reference length $L^* = 1$ μm and reference speed $v^* = 2.0 \times 10^4$ cm/s; this speed is appropriate for relatively high temperature $T = 500$ K and a relatively heavy reactant with molecular weight ω of about 200 g/mol.

It is challenging to display three-dimensional results in a meaningful and elucidating way and different choices are necessary for each quantity plotted. For each plot in Fig. 5, we choose to use a slice plot, which

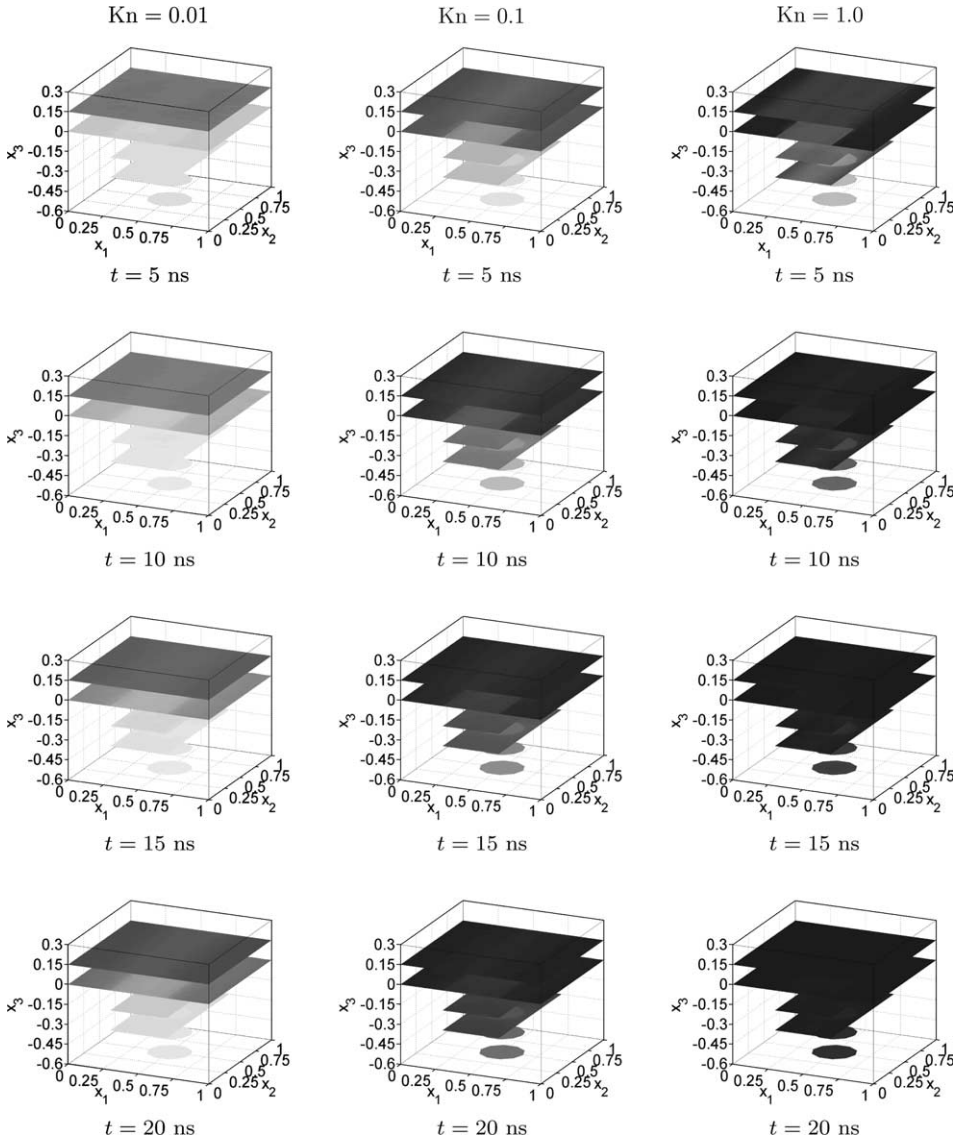

Fig. 5. Slice plots of the dimensionless concentration $c(\mathbf{x}, t)$, for selected Knudsen numbers Kn at times $t = 5$, 10, 15, 20 ns, as function of $\mathbf{x}$ in the three-dimensional trench/via feature with slices at heights $x_3 = -0.60$, -0.45, -0.30, -0.15, 0.00, 0.15. Grayscale from light $\iff$ $c = 0$ to dark $\iff$ $c = 1$.

encodes the values of $0 \leqslant c(\mathbf{x}, t) \leqslant 1$ for $\mathbf{x} \in \Omega$ on slices through Ω in grayscale from light color for $c = 0$ to dark color for $c = 1$; simultaneously, a slice plot still gives an indication of the shape of the domain Ω by the shapes of all slices taken together. We choose to show slice plots at 6 horizontal levels with heights of $x_3 = -0.60$, -0.45, -0.30, -0.15, 0.00, 0.15. The two top slices lie in the gaseous area of Ω above and at the opening of the trench. The two middle layers lie inside the trench, as seen by the shape of the slices. The two bottom layers cut through the via below the trench, indicated by their shape as disks.

In the process considered here, gaseous chemicals are supplied from the gas-phase interface (at $x_3 = 0.3$ in Fig. 1) and flow downwards throughout the domain Ω until they reach the solid wafer surface (the surface plotted in Fig. 1 with flat parts at height $x_3 = 0.0$), where a fraction of molecules react to form a solid deposit. Considering first the *right column* in Fig. 5 at time $t = 5$ ns, the top-most slice at $x_3 = 0.15$ is mostly dark-colored, indicating that a relatively high concentration of molecules have reached this level from the inflow at the top of the domain. The slice at $x_3 = 0.0$ shows that the concentration at the flat parts of the wafer

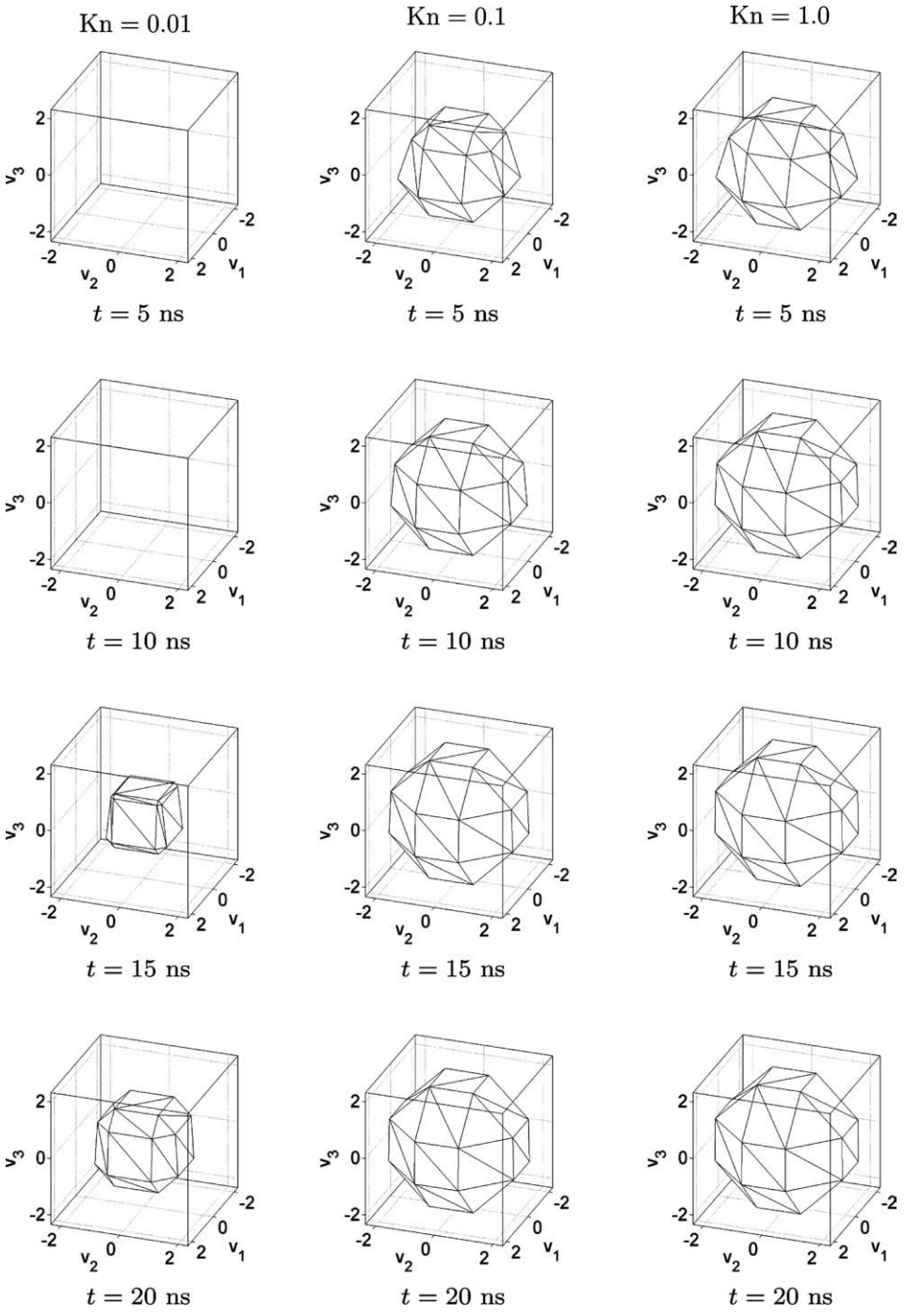

Fig. 6. Isosurface plots of the kinetic density $f(\mathbf{x},\mathbf{v},t)$, for selected Knudsen numbers Kn at times $t = 5$, 10, 15, 20 ns, as function of velocity $\mathbf{v} \in \mathbb{R}^3$ at the mouth of the trench at $\mathbf{x} = (0.5, 0.5, 0.0)$ in the three-dimensional trench/via feature. Isosurface level at $f(\mathbf{x},\mathbf{v},t) = 0.005$.

surface has reached relatively high values, as well, due to reflections from that surface, while the lighter color above the mouth of the trench (x_1-coordinates between 0.3 and 0.7) is explained by the ongoing flow of molecules into the trench. At the slice for $x_3 = -0.3$, we observe the same phenomenon where the concentration has reached a higher value in the flat areas of the trench bottom as compared to the opening into the via (round hole) below. Finally, not many molecules have reached the via bottom, yet, indicated by the light color there.

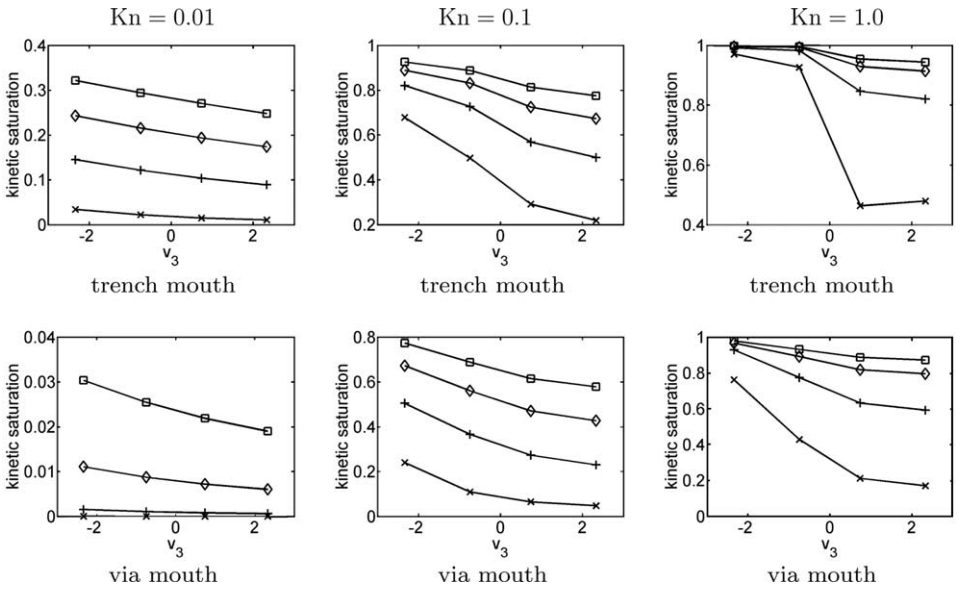

Fig. 7. Line plots of the saturation of the kinetic density $f(\mathbf{x},\mathbf{v},t)/M(\mathbf{v})$, for selected Knudsen numbers Kn, for $(v_1,v_2)=(0,0)$ as function of $v_3\in\mathbb{R}$; at the mouth of the trench at position $\mathbf{x}=(0.5,0.5,0.0)$ and at the mouth of the via at position $\mathbf{x}=(0.5,0.5,-0.3)$ in the three-dimensional trench/via feature; at times: ×, 5 ns; +, 10 ns; ◇, 15 ns; □, 20 ns. Notice the different scales of the vertical axes.

Comparing the three plots in the *first row* in Fig. 5 for Kn = 0.01, 0.1, and 1.0 at $t=5$ ns with each other, the ones for the smaller Kn have generally lighter color indicating a slower fill of the feature with gas. The smaller Knudsen numbers imply more collisions among molecules with the carrier gas, leading to a less directional flow than for the larger Knudsen number shown. Since the bulk direction of the flow is downward because of the supply at the top with downward velocity, the feature fills faster with molecules in the case of high Kn.

The plots in *each column* in Fig. 5 show how the fill of the entire domain with gaseous molecules continues over time. The results for the concentration $c(\mathbf{x},t)$ show that steady-state distribution of concentration is reached faster for the larger Knudsen number, which is realistic. Steady-state is not reached in 20 ns for the Knudsen numbers Kn = 0.01 and 0.1.

In Fig. 6, each plot shows the kinetic density $f(\mathbf{x},\mathbf{v},t)$ plotted as functions of the three-dimensional $\mathbf{v}\in\mathbb{R}^3$ for a fixed point $\mathbf{x}$. We show isosurface plots that indicate the surface of a body, along which f has a fixed value, here $f=0.005$. The choice of isosurface plots is motivated by the observation that a Maxwellian distribution has symmetric values about the origin $\mathbf{v}=(0,0,0)$ (the center of each plot) and hence the isosurface for a Maxwellian would look like a ball in three dimensions, up to the resolution used to discretize velocity space.

In Fig. 6, we start by considering the *right column* for Kn = 1.0. The isosurface plots as function of $\mathbf{v}\in\mathbb{R}^3$ are evaluated at position $\mathbf{x}=(0.5,0.5,0.0)$, which is located at the height of the mouth of the trench at $x_3=0.0$ and in the center of the trench and via structure with respect to the x_1- and x_2-coordinates. At $t=5$ ns, the isosurface plot has a skewed shape with more extent at the bottom, indicating higher values in f for velocity components with $v_3<0$. Over time, the f values for the velocity components with $v_3>0$ continue to grow due to both re-emission from the wafer surface and collisions, until at $t=20$ ns, the isosurface of f appears to approach the shape of a ball. The *center column* for Kn = 0.1 shows a similar evolution. Comparing the plots at equal times reveals that the approach to equilibrium is slower for the smaller Knudsen number; notice that even at $t=20$ ns, the sides of the apparent ball are not as near-vertical for Kn = 0.1 as those for Kn = 1.0. The results for Kn = 0.01 in the *left column* in Fig. 6 show the same evolution at yet slower pace; the apparently empty plots for the first and second point in time reveal that none of the components in f has reached a value above the cut-off of $f=0.005$ yet. Recall that these plots show f evaluated at the fixed position $\mathbf{x}$ at the mouth of the trench, and the previous figure showed that the concentration had not yet reached significant values at this position at these points in time.

In Fig. 7 for the saturation of the kinetic density $f(\mathbf{x},\mathbf{v},t)/M(\mathbf{v})$, we build on the realization that the kinetic density exhibits its main variation with respect to the third velocity component v_3, when evaluated at a point $\mathbf{x}$ that lies on the central axis of the domain in Fig. 1. To bring out this effect, we fix now $(v_1, v_2) = (0, 0)$ and plot the kinetic saturation f/M as a function of v_3 only. The columns in Fig. 7 fix the Knudsen numbers Kn = 0.01, 0.1, and 1.0, respectively. The first row shows the saturation again at the same spatial point $\mathbf{x} = (0.5, 0.5, 0.0)$ at the mouth of the trench and in the center with respect to x_1 and x_2 as in the previous figure. But we take advantage of the compact display of information possible for f/M as a function of v_3 by also showing it at the point $\mathbf{x} = (0.5, 0.5, -0.3)$ that lies at the mouth of the via and directly below the former point. Each individual plot in Fig. 7 contains four lines that plot f/M at the times $t = 5$, 10, 15, 20 ns, distinguished by markers as indicated in the caption. Notice that the scales of the vertical axes are different in each plot, with both the minimum and the maximum varying widely. The v_3 values on the horizontal axis vary from $v_3 = -3$ on the left to $v_3 = +3$ on the right in the conventional increasing direction; this means that f/M values on the left end are for downward velocities and for upward velocities on the right end.

We first observe that the increasing nature of the kinetic density f as a function of time is confirmed by the fact that the lines for latter times lie above lines for earlier times in every plot in Fig. 7. Looking at the *first row*, which plots f/M at the same position $\mathbf{x}$ as f in Fig. 6, we can confirm that the $v_3 < 0$ components of f are not just larger than the isosurface level value of $f = 0.005$ used in that figure, but they have in fact reached near-saturation for Kn = 1.0 and values above 0.9 for Kn = 0.1; the relative closeness of the lines for $t = 15$ and 20 ns indicates that the steady-state is being approached for these cases. For Kn = 0.01, all lines in the plot are still far apart from each other, indicating a continued increase of f/M over time beyond $t = 20$ ns. Notice how the lines of the saturation change more smoothly as function of v_3 for the more collisional cases of the smaller Knudsen numbers.

The *second row* in Fig. 7 shows the kinetic density at the point $\mathbf{x} = (0.5, 0.5, -0.3)$ further down inside the feature, at the mouth of the via. Since the chemicals are fed from the top of the spatial domain, it takes longer for them to reach a point deeper inside the feature. Hence, all values of f/M lag behind those for corresponding times in the first row. Only for Kn = 1.0, values of above 0.9 have been reached at all. The transition of the saturation values from $v_3 < 0$ to $v_3 > 0$ tends to be smoother in all cases in the second row than in the first, owing to the fact that this position is closer to the wafer surface, from which molecules with velocity components in all directions are re-emitted; recall the small value of the sticking factor $\gamma_0 = 0.01$ used here, hence 99% of all molecules re-emit from the wafer surface.

The results validate the effectiveness of the model and its numerical method for this application in three dimensions. The plots of the kinetic density provide here additional insight into the approach to steady-state and into the behavior observed for a macroscopic quantity such as the molar concentration.

5. Conclusions

In Section 2, we summarize the basis and formulation of the KTRM, which is appropriate to low pressure reactive flows in the presence of a dominant inert carrier gas. Such conditions are commonly found in the processing steps used to fabricate integrated circuits. In Section 3, we then summarize the design of the numerical method and its parallel implementation used to solve the KTRM. Two example are presented in Section 4, both using a simple sticking factor based reaction model for chemical vapor deposition: (1) reference studies for a flat wafer surface to analyze in detail the effect of the Knudsen number and species reactivity (sticking factor) on concentration and the kinetic density in Section 4.1; (2) transient studies for a three-dimensional trench/via feature that is commonly found in integrated circuit fabrication for several Knudsen numbers in Section 4.2. The KTRM provides reasonable solutions which can guide process development by increased understanding of low pressure transport and reaction.

Acknowledgments

The hardware used in the computational studies was partially supported by the SCREMS Grant DMS-0215373 from the US National Science Foundation with additional support from the University of Maryland,

Baltimore County. See www.math.umbc.edu/~gobbert/kali for more information on the machine and the projects using it. Prof. Gobbert also wishes to thank the Institute for Mathematics and its Applications (IMA) at the University of Minnesota for its hospitality during Fall 2004. The IMA is supported by funds provided by the US National Science Foundation. Prof. Cale acknowledges support from MARCO, DARPA, and NYSTAR through the Interconnect Focus Center. We also thank Max O. Bloomfield for supplying the original mesh of the trench/via structure.

References

[1] C. Bardos, F. Golse, C.D. Levermore, Fluid dynamics limits of kinetic equations II: convergence proofs for the Boltzmann equation, Commun. Pure Appl. Math. 46 (1993) 667–754.
[2] C.E. Baumann, J.T. Oden, An adaptive-order discontinuous Galerkin method for the solution of the Euler equations of gas dynamics, Int. J. Numer. Meth. Eng. 47 (2000) 61–73.
[3] G.A. Bird, Molecular Gas Dynamics, Oxford University Press, Oxford, 1976.
[4] M.O. Bloomfield, T.S. Cale, Formation and evolution of grain structure in thin films, Micro. Eng. 76 (1–4) (2004) 195–204.
[5] T.S. Cale, M.O. Bloomfield, D.F. Richards, K.E. Jansen, M.K. Gobbert, Integrated multiscale process simulation, Comput. Mater. Sci. 23 (2002) 3–14.
[6] J.A. Carrillo, I.M. Gamba, A. Majorana, C.-W. Shu, A direct solver for 2D non-stationary Boltzmann–Poisson systems for semiconductor devices: a MESFET simulation by WENO-Boltzmann schemes, J. Comput. Electron. 2 (2003) 375–380.
[7] J.A. Carrillo, I.M. Gamba, A. Majorana, C.-W. Shu, A WENO-solver for the transients of Boltzmann–Poisson system for semiconductor devices: performance and comparisons with Monte Carlo methods, J. Comput. Phys. 184 (2003) 498–525.
[8] C. Cercignani, The Boltzmann Equation and Its Applications, Applied Mathematical Sciences, vol. 67, Springer, Berlin, 1988.
[9] C. Cercignani, Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2000.
[10] 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, Berlin, 2000.
[11] E. Fatemi, F. Odeh, Upwind finite difference solution of Boltzmann equation applied to electron transport in semiconductor devices, J. Comput. Phys. 108 (1993) 209–217.
[12] M.K. Gobbert, M.L. Breitenbach, T.S. Cale, Cluster computing for transient simulations of the linear Boltzmann equation on irregular three-dimensional domains, in: V.S. Sunderam, G.D. van Albada, P.M.A. Sloot, J.J. Dongarra (Eds.), Computational Science—ICCS 2005, Lecture Notes in Computer Science, vol. 3516, Springer, Berlin, 2005, pp. 41–48.
[13] M.K. Gobbert, T.S. Cale, A feature scale transport and reaction model for atomic layer deposition, in: M.T. Swihart, M.D. Allendorf, M. Meyyappan (Eds.), Fundamental Gas-Phase and Surface Chemistry of Vapor-Phase Deposition II, vol. 2001-13, The Electrochemical Society Proceedings Series, 2001, pp. 316–323.
[14] M.K. Gobbert, V. Prasad, T.S. Cale, Modeling and simulation of atomic layer deposition at the feature scale, J. Vac. Sci. Technol. B 20 (3) (2002) 1031–1043.
[15] M.K. Gobbert, V. Prasad, T.S. Cale, Predictive modeling of atomic layer deposition on the feature scale, Thin Solid Films 410 (2002) 129–141.
[16] M.K. Gobbert, S.G. Webster, T.S. Cale, Transient adsorption and desorption in micrometer scale features, J. Electrochem. Soc. 149 (8) (2002) G461–G473.
[17] M.K. Gobbert, S.G. Webster, T.S. Cale, A Galerkin method for the simulation of the transient 2-D/2-D and 3-D/3-D linear Boltzmann equation, submitted (2005).
[18] A. Kersch, W.J. Morokoff, Transport Simulation in Microelectronics, Progress in Numerical Simulation for Microelectronics, vol. 3, Birkhäuser Verlag, Basel, 1995.
[19] W.H. Reed, T.R. Hill, Triangular mesh methods for the neutron transport equation, Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.
[20] J.-F. Remacle, J.E. Flaherty, M.S. Shephard, An adaptive discontinuous Galerkin technique with an orthogonal basis applied to compressible flow problems, SIAM Rev. 45 (1) (2003) 53–72.
[21] C. Ringhofer, Computational methods for semiclassical and quantum transport in semiconductor devices, Acta Numer. 6 (1997) 485–521.
[22] C. Ringhofer, Space–time discretization of series expansion methods for the Boltzmann transport equation, SIAM J. Numer. Anal. 38 (2) (2000) 442–465.
[23] C. Ringhofer, C. Schmeiser, A. Zwirchmayr, Moment methods for the semiconductor Boltzmann equation on bounded position domains, SIAM J. Numer. Anal. 39 (3) (2001) 1078–1095.
[24] C. Schmeiser, A. Zwirchmayr, Convergence of moment methods for linear kinetic equations, SIAM J. Numer. Anal. 36 (1) (1998) 74–88.
[25] S.G. Webster, Stability and convergence of a spectral Galerkin method for the linear Boltzmann equation, Ph.D. thesis, University of Maryland, Baltimore County, 2004.

[26] S.G. Webster, M.K. Gobbert, T.S. Cale, Transient 3-D/3-D transport and reactant-wafer interactions: Adsorption and desorption, in: P. Timans, E. Gusev, F. Roozeboom, M. Ozturk, D.L. Kwong (Eds.), Rapid Thermal and Other Short-Time Processing Technologies III, vol. 2002-11, The Electrochemical Society Proceedings Series, 2002, pp. 81–88.

[27] S.G. Webster, M.K. Gobbert, J.-F. Remacle, T.S. Cale, Parallel numerical solution of the Boltzmann equation for atomic layer deposition, in: B. Monien, R. Feldmann (Eds.), Euro-Par 2002 Parallel Processing, Lecture Notes in Computer Science, vol. 2400, Springer, Berlin, 2002, pp. 452–456.