

## Two deterministic approaches to topography evolution

Timothy S. Cale<sup>a,b</sup>, Max O. Bloomfield<sup>a,b,\*</sup>, Matthias K. Gobbert<sup>c</sup>

<sup>a</sup> Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, United States

<sup>b</sup> Process Evolution, United States

<sup>c</sup> Department of Mathematics and Statistics, University of Maryland, Baltimore County, United States

Available online 3 May 2007

### Abstract

We discuss two deterministic approaches to model free-molecular transport and reaction inside features on patterned wafers during thermal deposition processes. One approach that has been widely used to model chemical vapor deposition processes, which are largely steady state at the equipment scale, is based upon the ballistic transport and reaction model (BTRM). The major computational burden in BTRM based codes is usually computing the geometry-defined matrix of transmission probabilities; each element of which is used to determine what fraction of material leaving one point on the surface goes to each other point. This computation scales quadratically in the size of the discretizations, but is done in parallel. The second approach to modeling transport and reaction in the free-molecular flow regime is based on the kinetic transport and reaction model (KTRM). The KTRM starts from the Boltzmann equation and is particularly appropriate for processes operated under transient conditions at the reactor scale; *e.g.*, atomic layer deposition. The KTRM is computationally expensive; models in three spatial dimensions require the discretization of the three spatial dimensions, three velocity dimensions, and time for transient studies. The spatial mesh scales with the third power of the system size, for the same resolution. The KTRM is implemented in parallel.

© 2007 Elsevier B.V. All rights reserved.

**Keywords:** Ballistic transport; PVD; CVD; Simulation; Boltzmann transport equation

### 1. Introduction

The semiconductor industry makes extensive use of processes in which films are deposited on patterned wafers that are exposed to multi-component reactive gas sources; *e.g.*, chemical vapor deposition (CVD), and atomic layer deposition (ALD). Being able to predict and even engineer the shapes and characteristics of the deposited films is highly desirable in order to design and integrate these process steps to form reliable ICs. Both academic and commercial ‘topography simulation’ packages [1–3], which give answers, or at least guidance, about the shapes and perhaps compositions of the resulting films on a feature scale, are available for features that can be represented in 2D as well as those that require 3D representations.

While several groups have contributed to the literature of topography evolution, this paper focuses on two deterministic

approaches used by us to determine film profile evolution in features that need to be represented in 3D. The ratio of the mean free path of the molecules to the characteristic length scales of the features (the Knudsen number or  $Kn$ ) is significantly larger than 1 for the processes. These systems are said to be in the free-molecular, or ballistic, transport regime. After a molecule enters a feature, the probability it will collide with another gas phase molecule before making contact with the feature surface is small. In such cases, the chemical reactions that need to be considered are those between gas phase species and the surface. We first discuss the ballistic transport and reaction model (BTRM) that has been widely used to model deposition processes that are run under largely steady-state conditions (at the equipment scale) [2–4]. We then discuss the Boltzmann equation-based kinetic transport and reaction model (KTRM) that is appropriate to processes that are run under transient conditions at the equipment scale; *e.g.*, ALD. The KTRM and BTRM have been used to model different types of processes, for reasons discussed below, though there is some overlap of applicability. Although they are beyond the scope of the discussion here, other deterministic approaches have been

\* Corresponding author. Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, United States.

*E-mail address:* [bloomm2@rpi.edu](mailto:bloomm2@rpi.edu) (M.O. Bloomfield).

used to address similar problems, including but not limited to Green's function methods [5], in addition to the large number of Monte Carlo based methods that have been used [6–8] among many].

Although we do not deal in this paper with the actual movement of the surfaces due to deposition [9,10], it is important to clarify how fluxes are determined as the feature changes shape. In brief, for both the BTRM and the KTRM the surface is assumed to evolve slowly relative to the redistribution of fluxes [11]. This 'pseudo-steady-state' assumption essentially means that the fluxes of all of the species in the problem are computed assuming a fixed feature surface. After the species fluxes and film deposition rates are determined as functions of position, as discussed below, the surface is moved assuming the growth rates are constant for a small time interval. After the surface has been moved, the species fluxes and deposition rates along the surface are re-computed.

We deal with thermal deposition processes to limit our discussions to cosine, or diffuse, sources of species and re-emission from surfaces [4], though much of the discussion is valid for other 'species flux distributions' either as-is or with little modification. Much of the discussion is valid for other deposition and etch processes, with more general flux distributions. While the chemistries considered here are simple, in order to focus on species transport, the ability to implement chemistry models that are fairly complicated, even mechanistic, is an important aspect of 'topography simulators' [3].

## 2. The ballistic transport and reaction model

One well-studied approach to species transport and reaction in features at low pressures is the ballistic transport and reaction model (BTRM). The BTRM was reported by Cale et al. [12–14], and has been applied to feature-scale simulations many times over the years (see Refs. [2,3] and references therein); some aspects are reviewed below. Other views of transport and reaction at low pressures can be found in Ref. [2].

For many deposition processes, the BTRM may be expressed as a system of integral equations that are species balances along  $\partial R$ , the feature surface, given here for diffuse re-emission:

$$\eta_i^{a,t}(\mathbf{x}) = \eta_i^{a,l}(\mathbf{x}) + \int_{\partial R} q(\mathbf{x}, \mathbf{x}') [\eta_i^{a,t}(\mathbf{x}') + R_i(\eta_i^{a,t}, \xi, T)] d\mathbf{x}' \quad \text{for } \mathbf{x}', \mathbf{x} \in \partial R \quad (1)$$

Here,  $\eta_i^{a,l}(\mathbf{x})$  represents the flux of species  $i$  arriving at point  $\mathbf{x}$  from the source, and not by re-emission from other surfaces.  $\eta_i^{a,t}$  represents the total flux of the  $i$ -th species arriving at  $\mathbf{x}$ , which is often dominated by fluxes from other points on the surface; *i.e.* re-emission.  $R_i$  is the reaction rate expression that represents the generation rate of species  $i$  at point  $\mathbf{x}$  (molecules/time/area, just like fluxes), based on the local temperatures  $T$ , the local fluxes of all species  $\eta_i^{a,t}$ , and the local surface coverages  $\xi$  of reactive species on the surface. The term in brackets within the integrand is the flux of the  $i$ -th species being re-emitted from point  $\mathbf{x}'$ . The quantity  $q(\mathbf{x}, \mathbf{x}')$  is the differential transmission probability for molecules leaving position  $\mathbf{x}'$  to

arrive at position  $\mathbf{x}$ , after traveling in a straight line. It includes both the possibility of intervening solids (shadowing) and the relative orientations of the surface at both locations,  $\mathbf{x}$  and  $\mathbf{x}'$  [4].

The 'transmission probability matrix'  $\mathbf{q}$  contains the geometric information about a particular system that is needed to solve the transport part of the BTRM for a given feature geometry. After discretizing the surface using  $N$  triangles, the transmission probabilities are stored in an  $N$  by  $N$  matrix ( $\mathbf{q}$ ). All BTRM solvers compute this matrix, either implicitly or explicitly (perhaps convolved with another quantity). The size of the task of determining  $\mathbf{q}$  increases by the square of the size of the discretizations ( $N$ ), and usually constitutes the bulk of the computation workload for solving the BTRM; certainly for problems involving large or complicated structures and reasonable chemistries.

To determine the elements of  $\mathbf{q}$  'directly', one can construct lines or rays between all pairs of triangles, most commonly between their centroids. If two triangles are facing each other and can see each other, *i.e.*, the line between them does not hit another triangle, the transmission probability can be determined from the geometry in Fig. 1. The transmission probability between  $\mathbf{x}'$  and  $\mathbf{x}$  is:

$$q(\mathbf{x}, \mathbf{x}') = - \frac{\cos\Omega \cos\Omega'}{||\mathbf{s}||^2} \quad (2)$$

where the terms are defined in Fig. 1. Note that Eq. (2) is all the information that is needed to compute the transport for 'cosine' or 'diffuse' re-emission, which is a commonly used assumption for re-emission of species with reasonable (thermal) energies. A single line drawn between the centroids of two triangles can be used to determine if the view is obstructed (visibility is 0), or not (visibility is 1). Visibility and transmission probability determinations are symmetric, *i.e.*,  $q_{ij}=q_{ji}$ , which decreases the size of the computation. Although beyond the scope of this paper, non-diffuse re-emission models (*e.g.*, specular reflection) can be dealt with through careful alternate formulations of  $\mathbf{q}$  when computing the integral in Eq. (1).

A Monte Carlo method can also be used to determine the transmission probabilities. A straightforward MC method is to

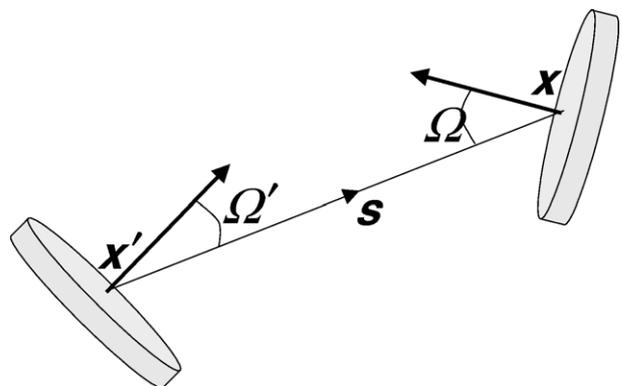


Fig. 1. Schematic of the quantities used to compute  $q(\mathbf{x}, \mathbf{x}')$ .

emit large numbers of rays from the  $i$ -th triangle, and then following or ‘tracing’ each one to see which triangle it intersects first. By emitting rays at angles sampled from a distribution, an entire row of the  $q$  matrix associated with triangle  $i$  may be determined; that is, the fraction of the total number of rays emitted that hits each other triangle (or leaves the feature). The main downside to using this method is that the number of rays that must be cast to determine a row in  $q$  can be quite large in order to get good statistics in each column (receiving triangle). Although good statistics can be quickly achieved on nearby or large triangles, smaller triangles or triangles that are far from the emitting triangle receive rays at much lower rates, and so the quality of the statistics associated with these triangle pairs increases more slowly. This becomes a problem because neighboring triangles can be assigned significantly different values in the same row, resulting in a ‘noisy’ simulation and artificial roughness in the film growth rates, particularly for non-linear reaction kinetics [15].

It is easy to demonstrate that the direct method discussed above is more efficient and usually more accurate than the MC method discussed. Though there are different ways to implement the direct method to determine  $q$ , the approach avoids the noise associated with poor statistics for entries in  $q$  with low values when using the MC method. Example comparisons of the approaches for simple structures are given in Ref. [16]. Both approaches to determining  $q$  can be implemented in parallel.

Once the transmission probability matrix is determined and the incoming distribution of species fluxes from the source gas is known (thereby determining  $\eta_i^{a,t}$  in Eq. (1)), the system of equations may be solved. There is not a closed form solution for the general case. However, the system usually responds well to fixed point iteration methods, particularly when  $R$  is linear in the fluxes  $\eta^{a,t}$ . The reaction rates determined as part of this solution can then be used to compute local deposition rates and film compositions along the surface. Examples of the application of the BTRM can be found in several places [4 and references therein].

### 3. The kinetic transport and reaction model

The KTRM is a deterministic model developed by Gobbert et al. [17,18], and is based on the Boltzmann transport equation [19,20]. The model applies to both collisional and collisionless

(free-molecular transport) conditions. For collisional systems, the KTRM is limited to systems in which there is a gas phase species that is dominant; *e.g.*, a carrier gas. The KTRM is really intended for systems with  $Kn$  of about 1 or higher. The quantity solved for when the KTRM is used is the kinetic density of each species  $f_i(\mathbf{x}, \mathbf{v}, t)$ , which is the probability of finding a molecule of species  $i$  at spatial position  $\mathbf{x}$  (in  $d\mathbf{x}$ ), that has velocity  $\mathbf{v}$  (in  $d\mathbf{v}$ ), at time  $t$ . Thus, in general one needs to consider three spatial dimensions, three velocity directions, and time for transient studies. So, rather than only the three spatial dimensions explicitly considered in the BTRM, the KTRM explicitly considers six or seven dimensions.

The Boltzmann transport equation for the kinetic density  $f_i$  of species  $i$ , in non-dimensional form is

$$\frac{\partial f_i}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_i = \frac{1}{Kn} Q_i(f) \quad (3)$$

where the subscript  $\mathbf{x}$  indicates that the gradient is over spatial variables, and  $Q_i$  is a collision operator. Gobbert et al. [17,18] use a spectral Galerkin method in velocity space and the discontinuous Galerkin finite element method in physical space to compute kinetic density functions of species, and a finite difference method to integrate it in time. This time variable is not necessarily process time, but can be rather just a variable used to arrive at the solution for a fixed geometry; *i.e.*, when the usual pseudo-steady-state assumption discussed above is used.

The KTRM is applied in a fundamentally different way than the BTRM, though both are ‘driven’ by the fluxes of species into the feature from the source volume. The BTRM requires ‘surface discretizations’, *e.g.*, triangles that represent a surface in 3D. In contrast, the KTRM requires a ‘volume discretizations’, *e.g.*, tetrahedral meshes of the spatial domains of interest. Additionally, a mesh or grid in velocity space must be established. Boundary conditions are applied both to represent the species concentrations and incoming fluxes from the reactor volume, as well as to represent the effects of collisions with the surface and any resulting chemical reactions. With these discretizations and boundary conditions in place, the kinetic densities  $f_i$  may be calculated as functions of velocity, at each position in the domain; by integrating the equations to steady state. The KTRM has been implemented in parallel [21].

To demonstrate the amount of information obtained in the solutions of the KTRM, Fig. 2 shows representative results of a

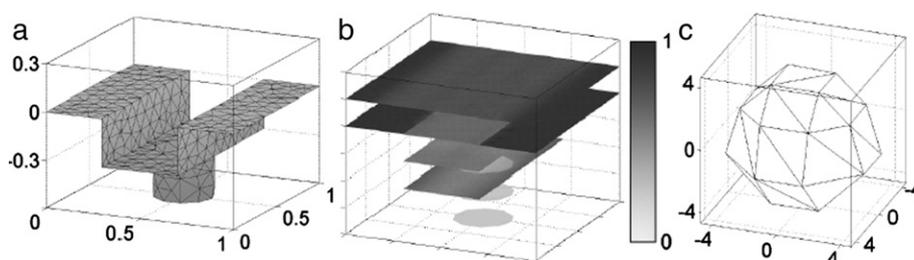


Fig. 2. (a) A representative dual damascene trench. Units in micrometers. (b) Slice plots of the dimensionless concentration of a precursor species, relative to its dimensionless concentration in the source ( $Kn=1.0$ , sticking factor 0.01) 5 ns after the concentration front reaches the wafer surface. (c) Kinetic density isosurface for  $f=5 \times 10^{-3}$  in velocity space. Units in  $10^4$  cm/s.

transient study of a single-species chemical vapor deposition model with  $Kn=1$  (in the transition regime) and a sticking factor of 0.01. Fig. 2a shows a surface discretization of an idealized ‘dual damascene’ structure (showing triangular faces from the volume mesh on the solid wafer surface), into which the species is flowing from the top of the three-dimensional domain. Fig. 2b shows a slice plot of the number density in the feature 5 ns after the species arrives at the wafer surface; assuming the feature had no molecules of this species in it initially. Fig. 2c shows the isosurface in velocity space of the kinetic density  $f$  where  $f=5 \times 10^{-3}$ , at a spatial point that is level with the top of the trench and centered over the via, at the same time of 5 ns. This indicates how much information is embedded in the set of  $f_s$  for the system; for each species there is a 3D probability density function in velocity space for each point in physical space, for each time. Summarizing results becomes a challenge. From the  $f_s$ , many quantities of interest may be calculated directly, including the number densities and fluxes to points on the surface. The reaction rates at the feature surface, and the film growth rates, may then be calculated as functions of position from these intermediate quantities.

#### 4. Discussion

There is clearly more information contained in the result of a KTRM simulation than in a corresponding result of a BTRM simulation. However, KTRM simulations require significantly more complex mathematics [22], require more computer power, and have far worse scaling behavior. In 3D, the number of volume entities typically used to resolve a geometry increases as the cube of the size of the system, faster than the number of surface entities used to describe it for a BTRM calculation, which increases quadratically. Thus, in problems for which the assumptions of the BTRM hold, it seems like the clear choice of approach.

The KTRM has the advantage that it starts with the Boltzmann transport equation, which employs only the most basic of assumptions [18,20] about the underlying gas and conditions. It is thus applicable to a wider variety of systems than the BTRM. There are two problem types for which the KTRM seems to be the right choice. When the Knudsen number is no longer significantly greater than 1, the assumptions of the BTRM are violated, but the KTRM remains valid for certain processes — though the difficulty expands considerably, since collisions have to be considered [18]. The second problem type, and one of interest to the microelectronics community, is that in which transients in gas flow and pressure are important, such as in ALD. ALD involves quick changes in partial pressures of component gases, which must propagate from the reactor inlets to the entirety of the feature-scale structures, in the presence of a perhaps chemically active surface. Gobbert et al. [17,23,24], and Webster et al. [21] performed KTRM calculations with direct application to ALD systems and were able to calculate the number density of precursor molecules as a function of position inside microelectronic features (such as that in Fig. 2) during ALD pulses. One important result of those simulations is that the time for species to enter features and the fluxes to stabilize takes

only order 100 ns. This supports the use of the pseudo-steady-state assumption, even when considering ALD systems [21].

The KTRM continues to undergo development. For example, the speed at which the kinetic density responds to changes in boundary conditions is very quick. The presence of these small time scales makes for expensive calculations when simulating the response of a system that has additional, longer time scales, such as those set by reaction kinetics or reactor set-point changes. This issue can be addressed through the use of implicit time integrators, adaptive time stepping; or steady-state approaches when applicable.

#### 5. Conclusions

Although different methods exist for simulating thermal low-pressure vapor deposition processes, we feel that deterministic methods continue to be good choices for microelectronics applications. The BTRM is a standard way to address feature-scale simulations in which the mean free path is significantly greater than the size of features. One deterministic approach for finding the transmission probability matrix  $q$ , an essential quantity in formulating the BTRM, employs a ray-tracing technique similar to those used in Monte Carlo approaches. However, it applies an analytical expression for pair-wise transmission probabilities and avoids the statistical sampling noise found when using MC methods. Depending upon the chemistry of the process under consideration, the technique can be expected to be at least a factor of 10 faster than the direct sampling MC technique discussed for systems with thousands of surface triangles. The BTRM scales as the square of the size of the discretization for systems in which calculation of  $q$  dominates.

The KTRM approach is also deterministic, and although more computationally expensive for the same problem, offers capabilities that the BTRM does not have. The primary dependent variables are the kinetic densities of each species in the problem, from which species fluxes and reaction rates can be computed. More generally, it can also simulate transient and transition regime gas phase behavior. The KTRM scales with the size of both the spatial mesh and the velocity mesh. To maintain a particular resolution, the spatial mesh tends to scale as the third power of the system size, although the velocity mesh is unlikely to require scaling. Further development of the KTRM has the potential to make it more efficient and reduce its hefty computational cost.

#### References

- [1] S. Abdollahi-Alibeik, J. Zheng, J.P. McVittie, K.C. Saraswat, C.T. Gabriel, S.C. Abraham, *J. Vac. Sci. Technol., B, Microelectron. Nanometer Struct. Process. Meas. Phenom.* 19 (2001) 179.
- [2] T.P. Merchant, M.K. Gobbert, T.S. Cale, L.J. Borucki, *Thin Solid Films* 365 (2000) 368.
- [3] S. Rosnagel, A. Ulman (Eds.), *Modeling of Film Deposition for Microelectronics Applications*, Thin Films, vol. 22, Academic Press, San Diego, CA, 1996.
- [4] T.S. Cale, T.P. Merchant, L.J. Borucki, A.H. Labun, *Thin Solid Films* 365 (2000) 152.
- [5] P. Vogl, U. Hansen, V. Fiorentini, *Comput. Mater. Sci.* 24 (2002) 54.

- [6] S.T. Rodgers, K.F. Jensen, *J. Appl. Phys.* 83 (1998) 524.
- [7] L. Houlet, A. Rhallabi, G. Turban, *J. Vac. Sci. Technol., A, Vac. Surf. Films* 17 (1999) 2598.
- [8] V. Sukharev, *Vacuum* 65 (2002) 281.
- [9] S. Hamaguchi, in: S. Rossnagel, A. Ulman (Eds.), *Math Methods for Thin Film Deposition Simulations, Modeling of Film Deposition for Microelectronic Applications*, vol. 22, Academic Press, San Diego, CA, 1996, p. 81.
- [10] T.S. Cale, V. Mahadev, Z. Tang, G. Rajagopalan, L.J. Borucki, in: F. Williams (Ed.), *Plasma Processing of Semiconductors*, France, Chateau de Bonas, 1996, p. 109.
- [11] T.S. Cale, V. Mahadev, in: S. Rossnagel, A. Ulman (Eds.), *Modeling of Film Deposition for Microelectronic Applications, Modeling of Film Deposition for Microelectronics Applications*, vol. 22, Academic Press, San Diego, CA, 1996, p. 175.
- [12] T.S. Cale, G.B. Raupp, *J. Vac. Sci. Technol., B, Microelectron. Nanometer Struct. Process. Meas. Phenom.* 8 (1990) 649.
- [13] T.S. Cale, *J. Vac. Sci. Technol., B, Microelectron. Nanometer Struct. Process. Meas. Phenom.* 9 (1991) 2551.
- [14] T.S. Cale, G.B. Raupp, T.H. Gandy, *J. Vac. Sci. Technol., A, Vac. Surf. Films* 10 (1992) 1128.
- [15] M.O. Bloomfield, T.S. Cale, *Microelectron. Eng.* 76 (2004) 195.
- [16] M.O. Bloomfield, "A computational framework for modeling grain structure development in three dimensions," Doctoral Thesis, Howard P. Isermann Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, 2007, Troy, NY, USA.
- [17] M.K. Gobbert, S.G. Webster, T.S. Cale, *J. Electrochem. Soc.* 149 (2002) 461.
- [18] M. Gobbert, T.S. Cale, *J. Comput. Phys.* 213 (2006) 591.
- [19] A. Kersch, W.J. Morokoff, *Progress in Numerical Simulation for Microelectronics*, vol. 3, Birkhauser Verlag, Basel, 1995.
- [20] G.A. Bird, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, 2nd ed. Oxford University Press, Oxford, 1994.
- [21] S.G. Webster, M.K. Gobbert, J.-F. Remacle, T.S. Cale, *Parallel numerical solution of the Boltzmann Equation for atomic layer deposition*, *Lecture Notes in Computer Science*, vol. 2400, Springer-Verlag, 2002, p. 452.
- [22] M.K. Gobbert, S.G. Webster, T.S. Cale, *J. Sci. Comput.* 30 (2) (2007) 237.
- [23] M.K. Gobbert, V. Prasad, T.S. Cale, *Thin Solid Films* 410 (2002) 129.
- [24] M.K. Gobbert, V. Prasad, T.S. Cale, *J. Vac. Sci. Technol., B, Microelectron. Nanometer Struct. Process. Meas. Phenom.* 20 (2002) 1031.