

Parallel Numerical Solution of the Boltzmann Equation for Atomic Layer Deposition

Samuel G. Webster¹, Matthias K. Gobbert¹,
Jean-François Remacle², and Timothy S. Cale³

¹ Department of Mathematics and Statistics,
University of Maryland, Baltimore County,
1000 Hilltop Circle, Baltimore, MD 21250, U.S.A.

² Scientific Computing Research Center, Rensselaer Polytechnic Institute,
110 8th Street, Troy, NY 12180-3590, U.S.A.

³ Focus Center — New York, Rensselaer: Interconnections for Gigascale Integration,
Rensselaer Polytechnic Institute, CII 6015,
110 8th Street, Troy, NY 12180-3590, U.S.A.

Abstract. Atomic Layer Deposition is one step in the industrial manufacturing of semiconductor chips. It is mathematically modeled by the Boltzmann equation of gas dynamics. Using an expansion in velocity space, the Boltzmann equation is converted to a system of linear hyperbolic equations. The discontinuous Galerkin method is used to solve this system. The speedup becomes near-perfect for the most complex two-dimensional cases. This demonstrates that the code allows for efficient parallel computation of long-time studies, in particular for the three-dimensional model.

1 Introduction

Atomic Layer Deposition (ALD) provides excellent film thickness uniformity in high aspect ratio features found in modern integrated circuit fabrication. In an ideal ALD process, the deposition of solid material on the substrate is accomplished one atomic or monolayer at a time, in a self-limiting fashion which allows for complete control of film thickness. The ALD process is appropriately modeled by a fully transient, Boltzmann equation based transport and reaction model [1,4,6].

The flow of the reactive gases inside an individual feature of typical size less than 1 μm on the feature scale is described by the Boltzmann equation [1], stated here in dimensionless form as

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = \frac{1}{\text{Kn}} Q(f, f). \quad (1)$$

The unknown variable is the density distribution function $f(x, v, t)$, that gives the scaled probability density that a molecule is at position $x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2$ with velocity $v = (v_1, v_2) \in \mathbb{R}^2$ at time $t \geq 0$. The velocity integral of $f(x, v, t)$ gives the dimensionless number density of the reactive species

$c(x, t) = \int f(x, v, t) dv$. The left-hand side of (1) describes the convective transport of the gaseous species while the right-hand side of the Boltzmann equation models the effect of collisions among molecules. For feature scale models the Knudsen number Kn is large and hence the transport is free molecular flow. Mathematically, this corresponds to a special case of (1) with zero right-hand side. The stated model is two-dimensional, a generalization to three dimensions is straightforward and first results are presented in [9].

Initial coarse and fine meshes of the domain Ω for the feature scale model are shown in Fig. 1. The fine mesh contains approximately twice as many elements as the coarse mesh. The meshes are slightly graded from top to bottom with a higher resolution near the wafer surface.

We model the inflow at the top of the domain ($x_2 = 0.25$) by prescribing a Maxwellian velocity distribution. We assume specular reflection on the sides of the domain ($x_1 = -0.25$ and $x_1 = +0.25$). Along the remainder of the boundary, which represents the wafer surface of the feature, a reaction model is used to describe the adsorption of molecules to the surface and diffusive emission describes the re-emission of molecules from the surface [1,4,6]. Initially, no molecules of the reactive species are present in the domain.

2 The Numerical Method

To numerically solve (1) with the given boundary conditions and initial condition, the unknown f for the reactive species is expanded in velocity space $f(x, v, t) = \sum_{k=0}^{K-1} f_k(x, t)\varphi_k(v)$, where the $\varphi_k(v)$, $k = 0, 1, \dots, K - 1$, form an orthogonal set of basis functions in velocity space with respect to some inner product $\langle \cdot, \cdot \rangle_C$. Using a Galerkin ansatz and choosing the basis functions judiciously, the linear Boltzmann equation (1) is converted to a system of linear hyperbolic equations

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

where $F(x, t) = (f_0(x, t), \dots, f_{K-1}(x, t))^T$ is the vector of coefficient functions. $A^{(1)}$ and $A^{(2)}$ are $K \times K$ diagonal matrices with components $A^{(\ell)} = \text{diag}(A_{kk}^{(\ell)})$ ($\ell = 1, 2$) [5]. Therefore, each equation for component function $f_k(x, t)$

$$\frac{\partial f_k}{\partial t} + a_k \cdot \nabla_x f_k = 0 \tag{3}$$

is a hyperbolic equation with constant velocity vector $a_k = (A_{kk}^{(1)}, A_{kk}^{(2)})^T$ given by the diagonal elements of $A^{(1)}$ and $A^{(2)}$. Note that the equations remain coupled through the reaction boundary condition at the wafer surface [6].

This system is then solved using the discontinuous Galerkin method (DGM) [2]. In the implementation in the code DG [7], we choose to use a discontinuous L^2 -orthogonal basis in space and an explicit time-discretization (Euler’s method). This leads to a diagonal mass matrix so that no system of equations

has to be solved. The degrees of freedom are the values of the K solution components $f_k(x, t)$ on all three vertices of each of the N_e triangles. Hence, the complexity of the computational problem is given by $3KN_e$; it is proportional both to the system size K and to the number of elements N_e .

The domain is partitioned in a pre-processing step, and the disjoint subdomains are distributed to separate parallel processors. The code uses local mesh refinement and coarsening and dynamic load-balancing using the Zoltan library as load balancer [3] and the graph partitioning software ParMETIS [8].

3 Results

Numerical studies were conducted for three different velocity discretizations. The demonstration results presented below were computed using four and eight discrete velocities in each spatial direction, respectively; hence, there are $K = 16$ and $K = 64$ equations, respectively. In each case, simulations were run for the two different initial meshes of Fig. 1. The solutions are presented in [4,6].

The studies were performed on a 8-processor cluster of four dual Linux PCs with 1000 MHz Pentium III processors with 256 KB L1 cache and 1 GB of RAM per node. The nodes are connected by 100 Mbit commodity cables on a dedicated network, forming a Beowulf cluster. Files are served centrally from one of the nodes using a SCSI harddrive.

Figure 2 shows observed speedup for up to eight processes for the various numerical studies conducted; the speedup measures the improvement in wall-

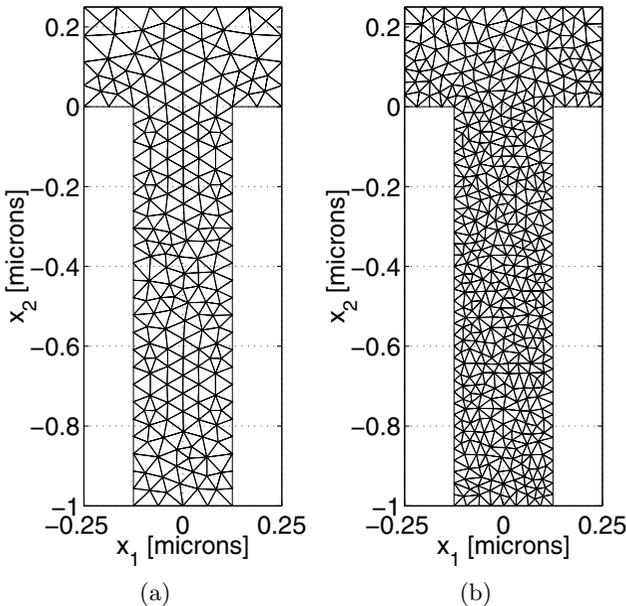


Fig. 1. (a) Coarse initial mesh, (b) fine initial mesh.

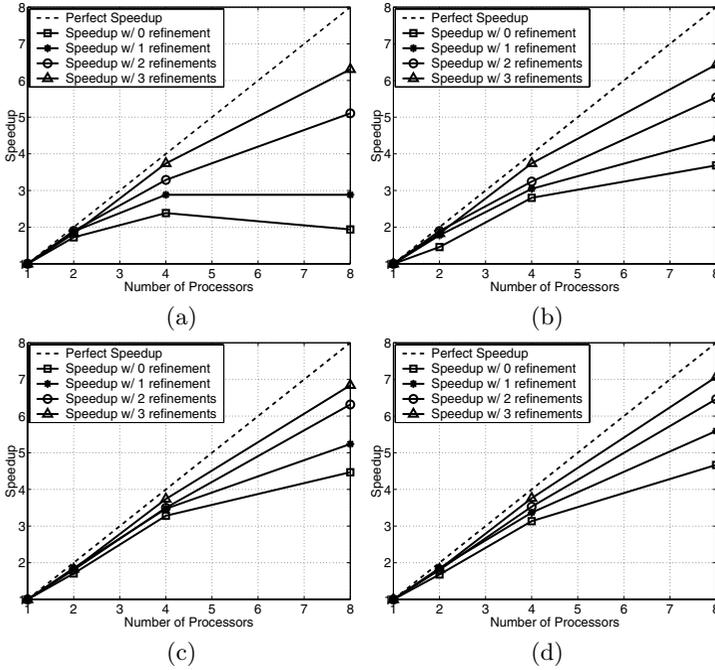


Fig. 2. Observed speedup for (a) coarse mesh / $K = 16$, (b) fine mesh / $K = 16$, (c) coarse mesh / $K = 64$, (d) fine mesh / $K = 64$.

clock time of the parallel code using p processes over the serial version of the code. The first row of plots in the figure corresponds to four discrete velocities ($K = 16$), and the second row corresponds to eight discrete velocities ($K = 64$). The left-hand column and right-hand column of Fig. 2 correspond to the coarse initial mesh and fine initial mesh of Figs. 1(a) and (b), respectively. Figure 2(a) compares the speedup for different levels of refinement of the initial coarse mesh with $K = 16$. Observe the decay in speedup without refinement due to the small number of degrees of freedom per process. Thus, as the maximum allowable refinement level increases and, consequently, the number of degrees of freedom increases, the speedup improves. Figures 2(a) and (b) demonstrate speedup for $K = 16$ for the two initial meshes. The finer mesh contains approximately twice as many elements as the coarse mesh; hence, the number of degrees of freedom increases by a factor of two. A comparison of the respective mesh refinement levels between the two plots shows that speedup improves because the degrees of freedom for the finer mesh is larger than for the coarse mesh. Figures 2(a) and (c) display speedup for the coarse mesh for the two studies $K = 16$ and $K = 64$. The finer velocity discretization in Fig. 2(c) introduces additional degrees of freedom which again improves speedup. Figure 2(d) combines the effect of the fine initial mesh and the finer velocity discretization. Observe that this is the most complex numerical study and thus possesses the best speedup.

4 Conclusions

It is demonstrated that the observed speedup improves with increasing levels of complexity of the underlying numerical problem. The studies were conducted using a two-dimensional model up to final times that are small compared to the time scales used for the process in industrial practice. The requirement to compute for long times, coupled with desired accuracy necessitates the use of an optimal parallel algorithm. While the demonstrated speedups are already extremely useful to conduct studies using the two-dimensional model, they become crucial in cases, when a three-dimensional model has to be used.

Acknowledgments

The authors acknowledge the support from the University of Maryland, Baltimore County for the computational hardware used for this study. Prof. Cale acknowledges support from MARCO, DARPA, and NYSTAR through the Interconnect Focus Center.

References

1. C. Cercignani. *The Boltzmann Equation and Its Applications*, volume 67 of *Applied Mathematical Sciences*. Springer-Verlag, 1988.
2. B. Cockburn, G. E. Karniadakis, and C.-W. Shu, editors. *Discontinuous Galerkin Methods: Theory, Computation and Applications*, volume 11 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, 2000.
3. K. Devine, B. Hendrickson, E. Boman, M. St.John, and C. Vaughan. *Zoltan: A Dynamic Load-Balancing Library for Parallel Applications; User's Guide*. Technical report, Sandia National Laboratories Tech. Rep. SAND99-1377, 1999.
4. M. K. Gobbert and T. S. Cale. A feature scale transport and reaction model for atomic layer deposition. In M. T. Swihart, M. D. Allendorf, and M. Meyyappan, editors, *Fundamental Gas-Phase and Surface Chemistry of Vapor-Phase Deposition II*, volume 2001-13, pages 316–323. The Electrochemical Society Proceedings Series, 2001.
5. M. K. Gobbert, J.-F. Remacle, and T. S. Cale. A spectral Galerkin ansatz for the deterministic solution of the Boltzmann equation on irregular domains. In preparation.
6. M. K. Gobbert, S. G. Webster, and T. S. Cale. Transient adsorption and desorption in micron scale features. *J. Electrochem. Soc.*, in press.
7. J.-F. Remacle, J. Flaherty, and M. Shephard. An adaptive discontinuous Galerkin technique with an orthogonal basis applied to Rayleigh-Taylor flow instabilities. *SIAM J. Sci. Comput.*, accepted.
8. K. Schloegel, G. Karypis, and V. Kumar. Multilevel diffusion algorithms for repartitioning of adaptive meshes. *Journal of Parallel and Distributed Computing*, 47:109–124, 1997.
9. S. G. Webster, M. K. Gobbert, and T. S. Cale. Transient 3-D/3-D transport and reactant-wafer interactions: Adsorption and desorption. *The Electrochemical Society Proceedings Series*, accepted.