Multiscale Modeling of Chemical Vapor Deposition and Plasma Etching

by
Seth Thomas Rodgers

B.S., Chemical Engineering
University of California, San Diego, 1994

M.S.C.E.P., Chemical Engineering
Massachusetts Institute of Technology, 1996

Submitted to the Department of Chemical Engineering
In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemical Engineering

at the

Massachusetts Institute of Technology

February 2000

Author..
Department of Chemical Engineering

Certified
By...
Klavs F. Jensen
Lamont Du Pont Professor
Department of Chemical Engineering
Thesis Supervisor

Accepted
by...
Robert E. Cohen
St. Laurent Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Multiscale Modeling of Chemical Vapor Deposition and Plasma Etching

by
Seth Thomas Rodgers

Submitted to the Department of Chemical Engineering on
December 22, 1999 in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Chemical Engineering

ABSTRACT

In this work, a framework and a set of modeling tools capable of describing systems with key processes occurring on widely separated length and time scales has been developed. The major focus of this work is linking atomistic and continuum descriptions of gas phase transport. This problem is of considerable practical interest, as most etching and CVD processes are run at low pressures ~ 1 torr or less. Under these conditions, the continuum diffusion models used to describe flow and transport in a typical reactor will fail below scales of a few hundred microns, and thus are not useful in describing transport in and around microscale topography. This is a serious limitation, as such topography is present in most microelectronic devices.

Two methods for linking discrete particle (or feature scale) and continuum models of precursor transport are presented. The discrete and continuum models are coupled by boundary conditions at their mutual interface (just above any reactive surface with microscale detail)

The first approach employs an effective reactivity function ε,, which is computed through a hybrid probabilistic-deterministic MC method. ε can be interpreted as a representation of the average fate of molecules entering the feature scale domain from the macroscopic model. An example of tungsten CVD over a substrate with surface topography typical of modern microelectronic devices is presented.

A second, deterministic technique was also developed as an improvement on the Monte Carlo approach. The deterministic method uses the matrix of transmission probabilities, or shape kernel, to summarize all microscale events in a fashion consistent with a continuum macroscopic model. The deterministic linking algorithm is over 1,000 times faster than the previously presented MC method. The speed advantage enables simulation of detailed chemistry. Plasma etching presents a very similar multiscale problem and a strategy for linked plasma etching simulations is presented.

Finally, a study of ionized physical vapor deposition of aluminum is presented as an example of atomistic-continuum linking. Molecular dynamics simulations are used to represent atomistic events. The Molecular Dynamics results are summarized in a manner that allows the combination of atomistic information with a continuum (level –set) model for evolution of the deposited metal film.

Thesis Supervisor : Dr. Klavs F. Jensen
Title : Professor of Chemical Engineering

Acknowledgements

Many thanks to my research advisor, Klavs, for his guidance during my stay at MIT, and for his open mindedness in allowing me the freedom to try lots of ideas (some successful others less so!). Working with Sadasivan Shankar at Intel was a great experience, I was lucky to find such a bright collaborator with just the right combination of enthusiasm and skepticism. Thanks as well to Peter Ventzek for introducing me to plasmas and for his support and encouragement.

Table of Contents

Chapter 1

Research Motivation and Background

The objective of this work is to develop a framework and a set of modeling tools capable of describing systems with key processes occurring on widely separated length and time scales. In general, at each length and time scale a particular set of simplifying assumptions and modeling techniques have proven useful. Most modeling work to date has been focused on a 'single scale'. While these single-scale efforts have been very successful, it has been shown that they can be even more powerful when combined in a self-consistent fashion to take a holistic, or 'multiscale' perspective on the systems of interest.

This type of multiscale analysis is of interest in microelectronics manufacturing, particularly in examining chemical vapor deposition (CVD) and plasma etch processes. In these processes, human operators attempt to manipulate macroscopic parameters such as temperature, reactor design, feed flow rates and pressures to enhance the performance and yield of circuit elements as small as 0.25 micron. Thus, the links between macroscopic and microscopic processes are of critical importance as circuit dimensions continue to decrease.

1.1 Multiscale Simulation in Microelectronics Processing

Chemical vapor deposition and plasma etching are two of the most versatile unit operations commonly used in the manufacture of microelectronic devices. Successful design and optimization of a such process requires first, selection of precursor chemistry.

Next, a 'process window' , in terms of a range of temperatures, pressures, flow rates. etc. must be found to maximize the economic potential of the process.

Process design and optimization is often a daunting task, as CVD and etching are complex processes with physical and chemical phenomena occurring on widely separated length and time scales. A summary of the length scales that may be important in CVD is presented in Fig 1-1, as well as some of the physical phenomena and modeling techniques used at each scale.

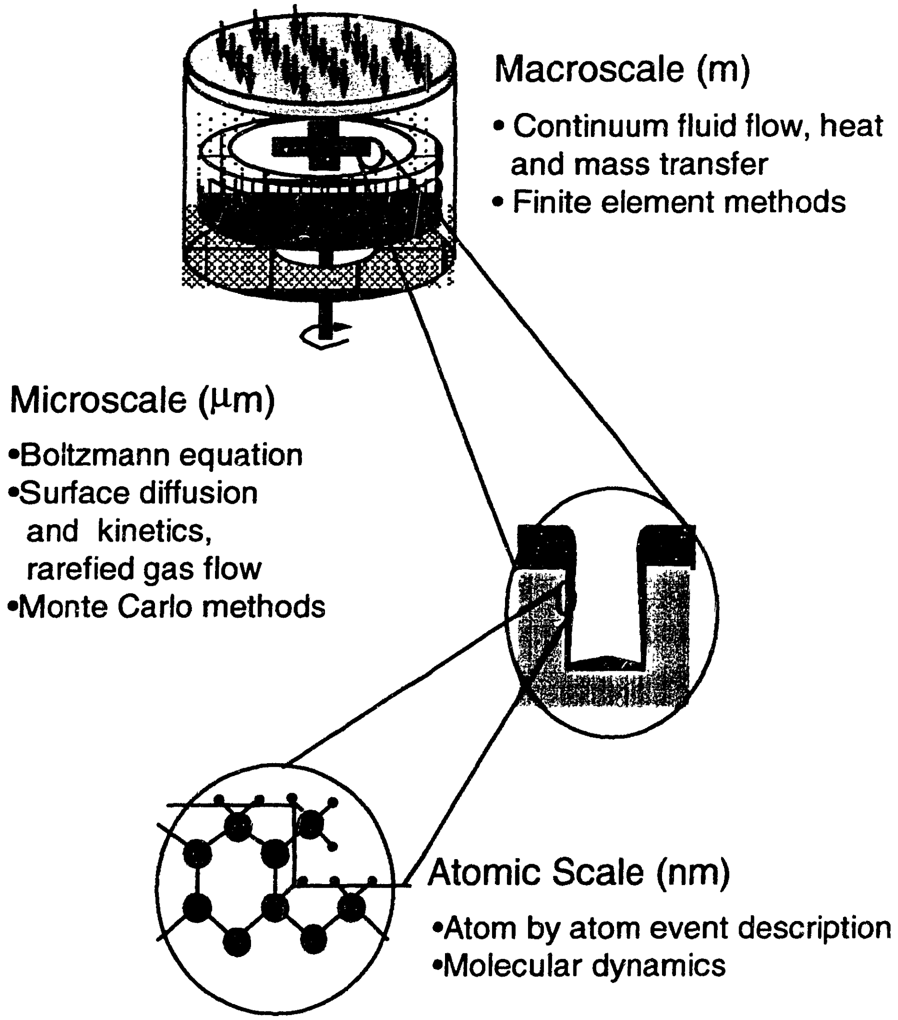


Figure 1- 1 Schematic of CVD length scales. This work fill focus on linking feature scale and reactor scale, as well as feature scale and atomic scale.

In principle, fundamental molecular models could be developed for the entire system. Starting at the atomic scale, and simulating each event in the system, atom by atom, would produce the distibutions solving the boltzmann equation (microscale). In turn averaging the boltzmann distributions would yield solutions at the macrosclae. The computational requirements of simulating a real system in this fashion would rapidly become prohibitive. Macroscopic models can often be solved with less computational effort, but will fail as length scales decrease. This is because the system will no longer be well characterized by its 'average' behavior and the underlying assumptions of the macroscopic approach will break down.

At each of the length scales depicted above, reactor[1, 2], feature,[3, 4] and atomic,[5, 6] simulation efforts have made important contributions. There is thus a need to develop an efficient, multiscale linking approach that combines the cost-effective macroscopic perspective with the accuracy of a detailed microscale simulation.[7]

For plasma processing, a very clear discussion of the multiscale problem is given by Economou.[8] Although the important phenomena are somewhat different, the length scale groupings and modeling techniques are largely the same as those important in CVD.

From a modeling perspective, once the key physical phenomena have been selected, two main reasons exist for introducing divisions that create distinct groups of scales. The more fundamental of these is a division made on physical grounds, where the underlying assumptions used to formulate a model at one length scale become invalid as scales change. The major thrust of this thesis is toward models that are consistent with the important physical assumptions made at each length scale in the overall domain. In each chapter of this thesis, continuum and atomistic perspectives are linked in as self-consistent fashion.

The second, and no less compelling, reason for grouping scales is convenience. For example, consider an atmospheric pressure CVD process where a thin film is to be deposited selectively in features of a few hundred microns. Here, the underlying assumptions of a continuum transport model are valid all the way to the feature scale, but actually refining a finite element or finite difference mesh to the point where such detail may be resolved would be very expensive.

1.2 Linking Reactor and Feature Scale

The parameter governing the grouping of length scales in reactive gas processing systems is the Knudesn number, *Kn* which represents the ratio of the gas phase mean free path, λ, to the length, *L*, over which flow conditions change. When *Kn* is small, molecules collide many times while their environment changes, and the gas phase may be regarded as being in collisional equilibrium, and transport through such a flow is well described as a combination of ordinary and thermal diffusion. Conversely, large *Kn* may imply the absence of collisional equilibrium, and an atomistic perspective must be employed to correctly address the problem of species transport.

In most low- pressure chemical vapor deposition (LPCVD) processes, λ is typically between 10mm and 100mm. Thus, the assumptions used in making a model of the reactor will be incorrect if applied on the feature scale. However when an LPCVD process is evaluated, feature scale results are of prime importance, a multiscale simulation approach is needed to reconcile the 'split' between tunable macroscopic process parameters, and microscopic results. In plasma processing, lower pressures are often used, but the same problem exists.

Examples taken from LPCVD are used as model problems in chapters 2 and 3 of this thesis. Chapter 4 exploits the similarities between LPCVD and plasma processing, to extend the multiscale CVD approach to plasma etching. In each case, a continuum view of species transport in the reactor at large must be linked in a self-consistent manner with the atomistic perspective that is required on the micron scale of typical of the features of microelectronic device circuitry. A detailed statement of this problem is given at the beginning of chapter 2., but the figure 1.2 highlights the key interactions between the macroscopic and feature scale perspectives.

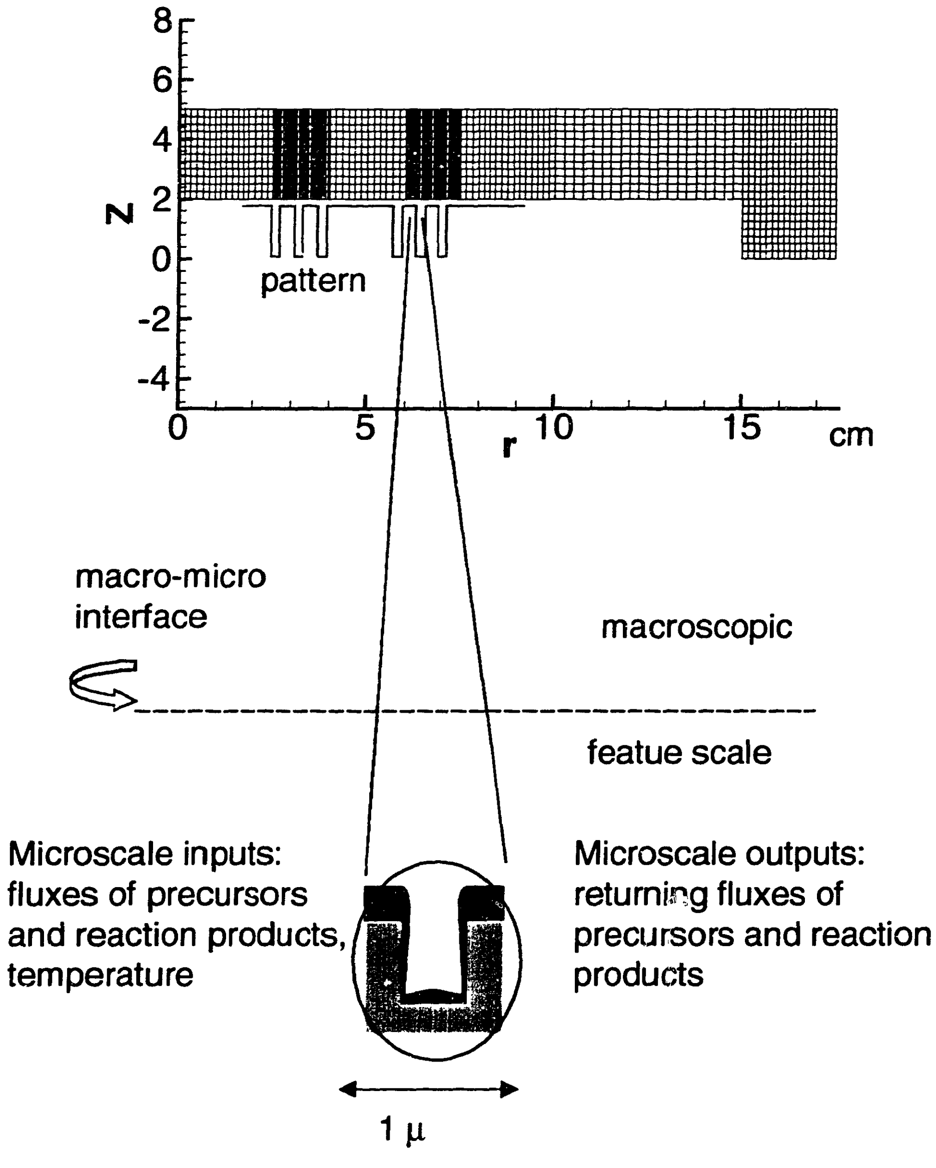


Figure 1- 2 Feature and rector environments are linked, as the reactor and substrate constantly exchange material. Feature scale models require a boundary condition at their interface with the macroscopic environment, reactor scale models will need a boundary condition at the substrate surface that reflects the presence of microscale topography.

From the feature scale perspective, one question immediately presents itself. How much flux of each species is arriving at the surface? Clearly, feature scale simulation cannot be undertaken independently of reactor scale modeling unless a boundary condition is measured or guessed.

Several investigators have recognized this split and attempted to link continuum and atomistic perspectives with a sequential model approach. In this approach, macroscopic models are solved information is passed on to successively smaller scales. Hasper *et al.*[9] have employed this approach for tungsten deposition from both WF_6/H_2 and WF_6/SiH_4. After solving the macroscopic conservation equations to find conditions at the substrate surface, they performed feature scale calculations to determine step coverage. A multiscale model for CVD of diamond spanning from reactor to atomic scale has been described by Srolovitz *et al.*,[10] were reactor scale solutions were fed to atomic and microscale models of the evolving surface.

In plasma processing, reactor scale results have also been used to determine a set of species fluxes as an input to a feature scale model.[11][Economou, 1988 #92] This approach can provide insights into the effects of reactor scale conditions on feature scale processes. The sequential model approach is well suited to systems such as PVD, or CVD of very reactive precursors, as few molecules return to the reactor chamber after visiting the substrate surface and no information is fed back to the gas phase from the surface.

However, the sequential model approach is not always adequate. Many CVD and plasma processes involve constant exchange of material between substrate and reactor chamber, thus allowing a strong interaction between the two domains. In these situations, a bi-directional linking approach is required to describe the collective effects of microscale events on the macroscale which could affect behavior on the microscale.

1.3 Bi-Directional Linking

In cases where reactor scale and feature scale processes interact, obtaining the correct inputs for feature scale simulation amounts to finding a consistent set of conditions at the border between macroscopic and microscopic models. In a nutshell, it is not possible to correctly model events in one scale domain without treating events in the other.

This multiscale resolution is difficult, since in many CVD operations, deposition occurs over integrated circuit (IC) device layouts that are heterogeneous patchworks on length scales ranging from centimeters to microns. Each region may present a somewhat different reactive surface area to film precursors. If reaction rates at the surface are high enough compared to diffusion through the gas phase, concentration gradients, *i.e.*, loading effects, will appear at the edges of pattern fields. In silicon device manufacture, transport of precursors into clusters of sub-micron features must be considered. In addition, heterogeneity at larger scales may be present *e.g.* areas where device features are dense may border sparse areas with few features. The same is true in plasma etching, where 'masked' areas may be adjacent to dense field of features, again, resulting in significant concentration gradients.

Efforts to simultaneously resolve both reactor and feature scale environments in CVD, using a mesoscale model as an intermediate, have been undertaken by Cale *et al.*[12, 13] In their work, the macroscale problem is solved with the FEM, and mesoscale models are created by local refinement of the finite element mesh in the vicinity of clusters of identical features. The effects of feature scale topography are treated with an averaging scheme based on asymptotic analysis. Difficulty with this approach may be encountered as on an actual substrate, the deposition surface may be heterogeneous on several sub-continuum length scales, with variation in both the type of surface, and geometry of features present. Averaging schemes would have to be redone at each scale of detail present. Moreover, arguments based on reactive area are not sufficient to completely describe the effects of feature scale heterogeneity, as microscale effects that restrict access of precursors to the growth interface, such as 'shadowing,' must be considered.

In this work an integrated approach is developed to link continuum and discrete particle transport models. The method is very flexible, and works with either MC or ballistic integral feature scale models. A superposition strategy is used to build a reactor scale flux boundary condition from individual feature scale simulations. Level set[14] and MC based profile simulators were used and linking demonstrated for tungsten and aluminum CVD as well as chlorine etch of polysilicon.

1.4 Atomic Scale information at the Feature scale

A sequential model linking approach is also demonstrated, this is a common situation when atomic simulations are used to provide inputs to models at larger scales. Interactions between the substrate surface and material arriving from the gas phase can be treated with various degrees of detail. One common assumption is that arriving particles come to thermal equilibrium before being re-emitted from the substrate surface.

This equilibration has an important consequence: once a particle arrives at the surface, it is always re-emitted diffusely, and thus it looses all 'memory' of where it came from, and its subsequent behavior is not correlated with its past. While this is true in many systems, it is not a good assumption for the very energetic ions and atoms that arrive at the surface in etching and PVD processes. In this case, an obvious solution is to simulate every particle with molecular dynamics and grow the film atom by atom. If realistic surface interactions are used, this method is too expensive by two or three orders of magnitude to simulate film growth over micron size structures. An alternative would be to let each arriving particle represent a 'bucket' of atoms. While this would be faster, it is much less rigorous, furthermore, simulating enough 'buckets' is also rather slow. Instead, we develop an efficient and accurate scheme for employing (and summarizing) molecular dynamics simulations within the framework of a level set model of the deposited film. We present a case study of ionized PVD aluminum film growth. The method is able to capture the influence of atomistic phenomena on the morphology of the deposited film. Furthermore, the model is fast enough to treat cases where the growing film is subjected to simultaneous argon bombardment, which defeat an atom by atom approach completely.

1. Kleijn, C., in *Computational Modeling in Semiconductor Processing*, M. Meyappan, Editor. 1995, Artech House: Norwood, Ma. p. 97.
2. Jensen, K.F., *Transport Phenomena in Epitaxy Systems*, in *Handbook of Crystal Growth*, H. DTJ, Editor. 1994, Elsevier: Amsterdam. p. 543.
3. IslamRaja, M.M., *et al.*, *J. Appl. Phys*, 1991. **70**(11): p. 7137-7140.
4. Cale, T.S., G.B. Raupp, and T.H. Gandy, *J. Appl. Phys.*, 1990. **68**: p. 3645.
5. Ruggerone, P., A. Kley, and M. Scheffler, *Progress in Surface Science*, 1997. **54**(3-4): p. 331-340.
6. Gilmer, G.H., *et al.*, *Materials Science and Engineering B*, 1996. **37**: p. 1-7.
7. Jensen, K.F., S.T. Rodgers, and R. Venkataramani, *Current Opinion in Solid State and Materials Science*, 1998. **3**: p. 562.
8. Economou, D.J., *Plasma Engineering*, in *Advances in Electrochemical Science and Engineering*, R.C.Alkire and D.Kolb, Editors. 1999, Wiley: New York. p. 241.
9. Hasper, A., *et al.*, *J. Electrochem Soc.*, 1991. **138**: p. 1728.
10. Sorolovitz, D.J., *et al.*, *JOM*, 1997. **49**: p. 42.
11. Hoekstra, R.J., M.J. Graperhaus, and M.J. Kushner, *J. Vac. Sci. Technol. A*, 1997. **15**: p. 1913.
12. Gobbert, M.K., C.A. Ringhofer, and T.S. Cale, *J. Electrochem. Soc*, 1996. **143**: p. 2624-2631.
13. Gobbert, M., *et al.*, *J. Electrochem. Soc.*, 1997. **144**: p. 3945-3951.

14. Adalsteinsson, D. and J.A. Sethian, *Journal of Computational Physics*, 1995. **120**: p. 128-144.

Chapter 2

Monte Carlo Methods for Multiscale Modeling of CVD

In the broadest sense, there are two types of Monte Carlo simulations. The first is a probabilistic simulation, the second a deterministic one.[1] An example of a probabilistic problem might be estimation of demand for telephone service given some idea of individual calling habits. On a particular day, each person might have a particular chance of using the phone for a given length of time. Each person might be at home or not at home, or a person's line might be already busy when a call arrives. While these situations are easy to describe on an individual basis, understanding the collective behavior of the system is a much more difficult proposition. A Monte Carlo simulation, using information about the individual behaviors as inputs, can easily solve this type of problem. Because MC is so well suited to building a picture of collective behavior from individual actions it is a natural choice as a method for multiscale analysis.

Deterministic MC calculations are somewhat different. Here, a problem may be written down as set of equations describing a deterministic process, but with a solution that is hard to achieve using standard analytical techniques. Multi-dimensional integrals in difficult geometries, for example, are often performed by MC as they often defeat any other technique.[2] The Boltzamnn equation can be solved by deterministic MC, as can its simpler analog, the transport equation. In fact, most of the more sophisticated MC methods were designed for use in exactly this capacity, and are a very powerful way of handling any transport problem.

In this work, MC methods are used from both the deterministic and probabilitic point of view. The goal of these simulations is two-fold, first, to determine the final shape of films deposited in microscale features. Secondly MC simulation is used to determine the 'average fate' of a molecule visiting a patterned surface an thus the effect of the pattterns on the macroscopic solution. Frequently, enough insight was gained from the probabilistic approach that deterministic model was later designed based on the probabilistic results. In general, replacing probabilistic models with a deterministic approach produces a faster, more accurate simulation and is done wherever possible.

2.1 Multiscale CVD Problem Statement

At the macroscopic level, continuum descriptions of fluid flow, heat transfer and mass transfer apply. For the dilute gas systems common in CVD, only the transport of chemical species is affected by microscale heterogeneity, and is thus the only problem to require multiscale resolution. The macroscopic species balance is.

$$\frac{\partial}{\partial t}(\rho Y) + \nabla \bullet (\rho\, \hat{u} Y) = -\nabla \bullet \mathbf{j} + R \qquad (2\text{-}1)$$

Here, Y represents the mass fraction of a species in the gas phase, $\mathbf{j}$ is the diffusion flux, and R is the generation term in the gas phase. The flow velocity, $\hat{u}$, and the temperature are determined by FEM solutions of the momentum and energy conservation equations. The density, ρ, is found from the ideal gas law.

Microscale heterogeneity affects the macroscopic solution through the boundary conditions imposed at reactive surfaces. In general, a mass balance may be written for each species at each point microscale point, x, on the surface of the substrate as Eq. 2-2.

$$r^{s}(x) = j(x) \qquad (2\text{-}2)$$

where r^{s} represents the local surface reaction rate and j the net local reagent flux our of the adjacent gas phase. However, when heterogeneity is present at length scales below λ, the expression in Eq. 2-2 cannot be readily combined with a continuum model of chemical species transport. This is because the gas is no longer assumed to be in collisional equilibrium, a diffusion coefficients cannot be defined. This is the case for the LPCVD process in figure 2.1

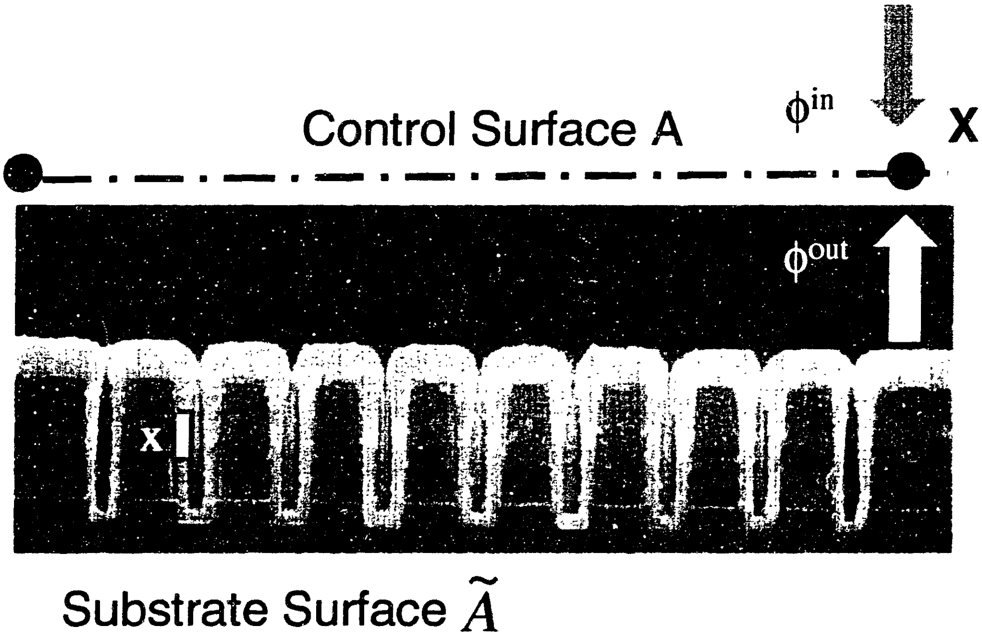


Figure 2. 1 Continuum methods – such as finite element techniques are not suitable for modeling transport and reaction in and around microscale topography. Multiscale techniques summarize events at the microscale in a fashion compatible with a continuum, finite element model.

Introducing a control surface, Fig. 2-1 having surface area A just above the substrate $\tilde{A}$, divides the macroscopic model of the reactor chamber from the microscale environments on the surface. In order to capture the effect of the many microscale features on the macroscale, events occurring at all microscopic points x, must be realted to conditions at all points, X, in the macroscopic domain. We can now write a mass balance on the macroscopic side, Eq. 2-3:

$$j\ (X) = \phi^{in}(X) - \phi^{out}(X) \qquad (2\text{-}3)$$

Here, $\phi(X)$ represents the scalar flux density of molecules crossing the control surface, and $j(X)$ represents the net flux density at point X. In systems displaying chemistry that can be described with a 'sticking coefficient' approach, another mass balance at the control surface, Eq.2-4, is possible

$$j(X) = \phi^{in}(X)P^{effective}(\mathrm{X}) \tag{2-4}$$

$P^{effective}$ is the effective sticking coefficient, which is the chance that a molecule leaving the macroscopic domain at X will be incorporated into the growing film before returning to the macroscopic. This may differ from the intrinsic, surface kinetic sticking coefficient as a molecule might make repeated encounters with the surface while in the microscale domain. For convenience we separate $P^{effective}(X)$ into the intrinsic sticking coefficient, P_o and a function $\varepsilon(X)$ which includes the effects of everything else occurring at the microscale. We call $\varepsilon(X)$ the 'effective reactivity'.

$$P^{effective}(X) = \varepsilon(X)P_o \tag{2-5}$$

From the macroscopic point of view, Eqs. 2-3 to 2-5 have served to exchange the problem of precursor transport to surface $\tilde{A}$ for that of transport to A, and the flux to the control surface, A, now becomes the boundary condition on the macroscopic finite element calculation: Eq. 2-6.

$$D\nabla C(X) = j(X) \tag{2-6}$$

In order for the boundary condition represented by Eq. 2-6 to be useful, variations in flux must be gradual enough for a continuum view of diffusion to be valid. That is, the flux map must include information from all scales where continuum transport models do not apply, but must itself be compatible with a continuum approach. Once reaction and transport inside microscale topography have been accounted for, a macroscopic description of microloading can be pursued. In general, $\varepsilon(X)$ will reflect a convolution of surface geometry, surface kinetics, and microscale transport resistance through variations in $j(X)$, all might be functions of position and time. The following sections illustrate the computation of effective reactivity to create microscale flux maps. Note that some details of the calculation are most easily expressed in deterministic form and are treated in chapter 3.

2.2 Monte Carlo Methods for Feature Scale Transport Problems

Assuming sufficient computational resources are available, probabilistic MC models may be formulated in a very straight froward fashion as direct or near-direct analogs to the physical processes of interest. While this generally results in models that are expensive relative to other strategies, considerable physical detail may be easily included. MC methods are thus well suited to systems that are difficult to describe with deterministic sets of equations, but whose physical nature is well known. We began our work with a probabilistic MC approach.

Monte Carlo models have been used for some time to address discrete particle transport problems, especially rarefied gas flow and neutron transport.[3, 4] Within the context of CVD models, MC methods have been employed to examine precursor transport inside microelectronic device features.[5, 6] More recent efforts have employed MC models to simultaneously examine feature scale transport and the microstructure of deposited thin films.[7] While MC methods have seen widespread acceptance, many problems do not require all the detail available in a MC calculation. In those cases, MC models are usually a poor choice if the desired calculation can be preformed by other means.

In this work, a Monte Carlo model of CVD over a non-planar substrate was constructed. A 2-dimensional cross section of an infinite feature was taken as the simulation domain. While an actual CVD process may involve any number of active species and chemical reactions at the surface, simplifications were required to keep simulation times short enough to be useful. This model considers collisionless transport of a single, limiting reagent inside a long trench. Under the assumption of collisionless flow, particles do not interact and trajectories are generated independently of each other. The collisionless flow assumption is justified upon comparison of typical gas phase mean free paths in low- pressure CVD (tens of microns at 1 torr) with feature dimensions typical of modern microelectronic devices (less than one micron).

Simulating each atom,[7], is too detailed for our purposes and would lead to excessive computational requirements for film growth on the scale of microns as opposed to angstroms. Several authors have addressed this issue by introducing 'test particles' representing many atoms each.[8, 9] As long as enough trajectories are simulated, the statistics will be smooth enough to determine the flux profile inside the feature.

A schematic of the simulation domain is shown in Fig. 2.2.

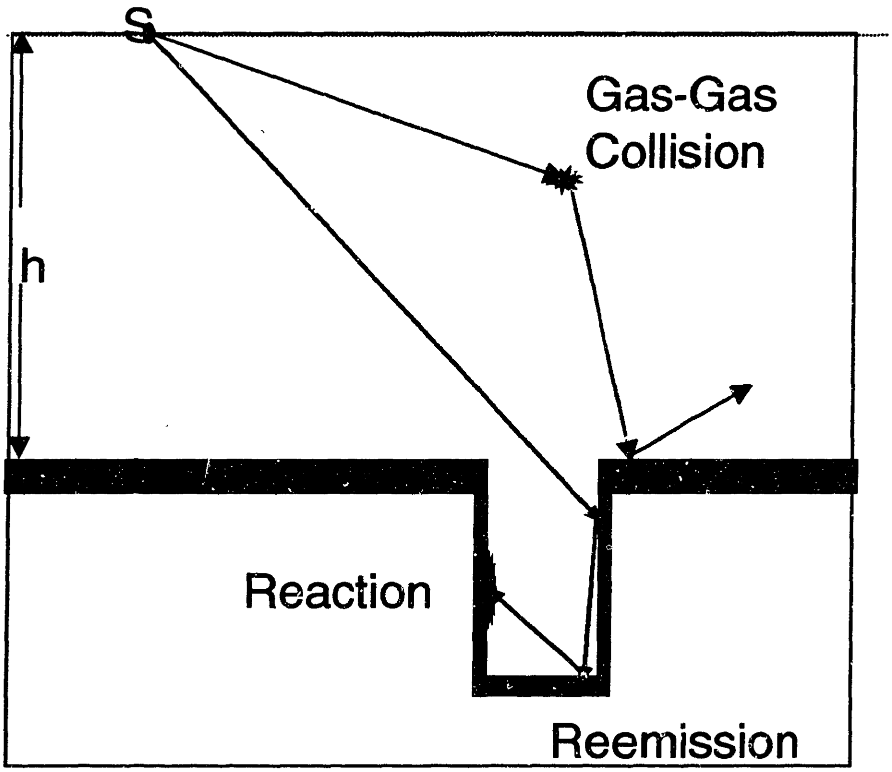


Figure 2. 2 Monte Carlo feature scale model. Moleculaes a re generated at a random position on the source line and followed until they impact the surface or leave the simulation domain. Boundary conditions are periodic.

Precursor molecules are generated at the source plane (line in a 2-D simulation). Their initial position, S, in the source plane is chosen randomly. The initial velocity of the particles is chosen from a Maxwellian distribution that may include a drift velocity depending on macroscopic conditions. Particles are then translated along their initial trajectories until they impact the substrate surface. Upon impact, the particle may be incorporated into the growing film with probability, P_r or re-emitted with probability $1-P_r$. If a molecule reaches the edge of the simulation domain it is reflected (periodic boundary conditions), while if it reaches the 'top' after interacting with the surface, it is assumed to escape. The simulation approach described in the formulation of the trench model may be readily applied in 2 dimensions to an axisymmetric via, or in 3 dimensions.

2.3 Effective Reactivity Concept and Calculation of Effective Reactivity at the Feature Scale

It is easy to guess that the details of deposition of a thin film on a silicon wafer covered with microscopic trenches are different than deposition over a flat, clean wafer. For one thing, the trenched wafer will present a greater surface area for reaction. Some precursors might be more or less able to diffuse into the microscale structures and reach the growing film. Although a few tiny features would not affect the macroscopic behavior of the reactor very much, modern device layouts have as many as several million features, which collectively may exert a significant effect. We will use a probabilistic MC calculation to determine the effect of a single feature and demonstrate how to extrapolate this to clusters or groups of features. The effect of a microscale feature on the macroscopic environment is a function of the number and nature of encounters a molecule has with the feature upon macroscopic environment is determined by the number and nature of encounters that a molecule has with the surface of the feature. Recalling the discussion at the beginning of the chapter, MC methods are well suited to determine the 'average fate' of molecules entering the feature sale model given their individual behavior. We will introduce molecules from the macroscopic domain and reco~d statistics on which react and which do not. recording the results as the 'effective reactivity of the surface. More formally, the calculation of the effective reactivity function at a point *S* above the substrate for a few test trajectories is described in Fig. 2.3.

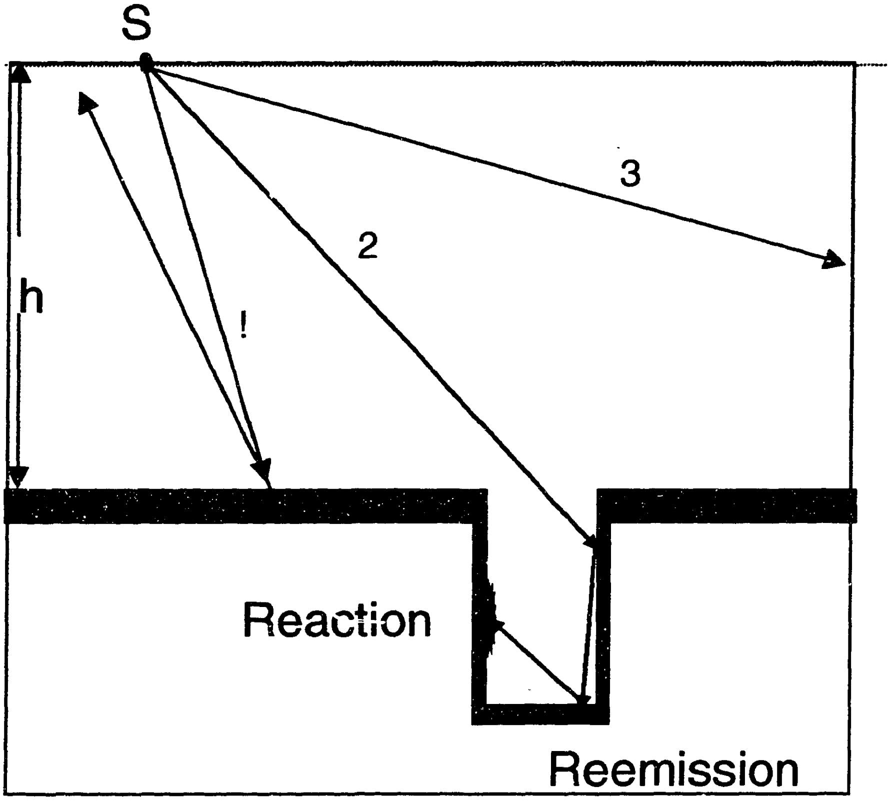


Figure 2. 3 The net flux into a feature or group of features from the macroscopic environment at point X can be calculated by considering the average fate of entering molecules

A large number of test trajectories is required to get an accurate estimate of the effective reactivity. Consider trajectories 1, 2, and 3 in fig 2.3. Each represents a very different fate. Particle 1 does not encounter the feature at all, is diffusely reflected and returns to the reactor chamber. Particle 2 is trapped by the feature, reacting after multiple surface encounters, while particle 3 makes no surface encounter at all. None of them individually is a very good description of what the feature 'looks like' when viewed from point S on the macroscopic side, However, after simulating many trajectories, we can build up a distribution (and an average !) of what the molecules entering the feature scale model will do.

More formally, statistics are recorded for some number, β, of test trajectories initiated at S. The history of each test trajectory yields an estimate of $\varepsilon(S,t)$, the effective reactivity of the surface with respect to a molecule entering the microscopic domain at position S. Each simulated trajectory, α, will consist of k steps (collisions). The locations of these collisions are stored as the x_i and define the test particle's history, Eq.2.7.

$$\alpha = (x_1, x_2 ... x_k) \tag{2.7}$$

Trajectories are assumed to terminate, either by surface reaction or escape to the macroscopic environment, within some finite k. ε can be expressed as a random variable on the space of all possible trajectories α. The expected value of ε, $E(\varepsilon)$, is given in Eq. 2.8.

$$E(\varepsilon) = \sum_{\alpha} P(\alpha)\varepsilon(\alpha) \tag{2.8}$$

This expectation value of ε represents the average environment seen by particles entering from the macroscopic gas phase.

Each collision may contribute differently to a history and thus, ε. For example, gas phase collisions with inert molecules may change a trajectory's direction, while surface interactions might influence both direction and reactive history. For a given trajectory $\alpha, \varepsilon(\alpha)$ is given by the following expression: Eq. 2.9

$$\varepsilon(\alpha) = \frac{\sum_{i=1}^{k} P_r(x_i)}{P_o} \tag{2.9}$$

Throughout the course of the trajectory, if a collision occurs, the local probability of reaction, $P_r(x_i)$ is added to the history. At each collision a random number f is compared to $P_r(x_i)$, if it is $<Pr(x_I,)$, the particle is said to have reacted and a new trajectory begins. A reference reaction probability, P_o, is obtained by comparison with a flat surface under the same macroscopic conditions.

To simplify subsequent arguments, deposition kinetics are taken to be first order decomposition of a single precursor, resulting in constant reaction probability over the substrate surface. Thus, all surface encounters are equally productive. Under these conditions, ε can be given a geometric interpretation, it is the effective area of the growth interface exposed to reactive precursors, after accounting for microscale transport resistance, or the average number of collisions a molecule makes with the surface in the feature scale domain. If the system of interest does not exhibit simple chemistry, due to flux-dependent reaction probabilities, or the presence of multiple reactions, reaction probability might be expressed in the style of Cooke and Harris,[5] but the computational cost will be large.

For a collisionless gas phase, the initial velocity of each test particle determines the point of impact on the substrate surface. Since only particles that enter surface features may experience multiple collisions, only these histories need be simulated. The estimator for ε above can now be re-written: as eq. 2.10

$$E(\varepsilon) = 1 + \frac{\beta^*}{\beta}\left[\left(\sum_{\alpha} P^*(\alpha)\gamma(\alpha)\right) - 1\right] \tag{2.10}$$

Where P has been changed to P^* to denote simulation only of the fraction of trajectories from a Maxwellian source, (β^*/β), that intersect the mouth of a feature, and thus might undergo multiple surface collisions. The number of surface collisions on a trajectory, α is $\gamma(\alpha)$. The term in round brackets is then the average number of collisions a particle entering the feature is expected to suffer. The advantages of this formulation are two-fold. For collisionless flows, the view factor term is easily computed analytically for most cases, but can be done as a MC integral, if analytic methods prove intractable. Because the fraction of molecules entering the feature depends only on feature geometry and source position, the feature scale transport problem is uncoupled from the surface chemistry. Secondly, if re-emission is assumed to be independent of incident trajectory, *e.g.*, cosine re-emission, the re-emission term depends only on feature geometry and reaction kinetics. Figures 2-4 through 2-6 illustrate effective reactivity maps for both the trench and a circular via. Note that estimates of the reactivity improve as the number of test trajectories is increased. Improvement with sample size is better than standard MC because only the 'relevant' trajectories (those entering the feature) were simulated.

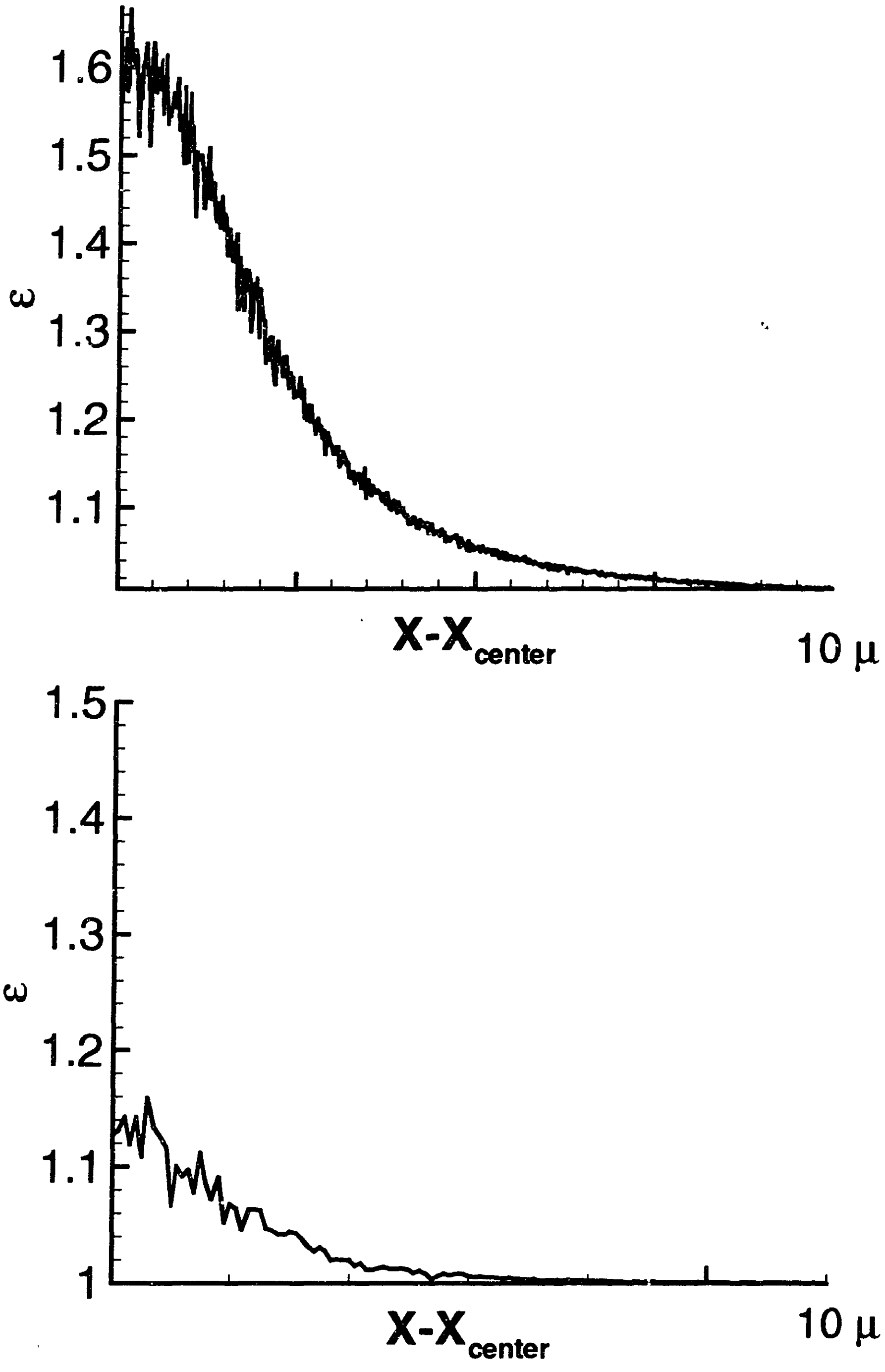


Figure 2. 4Effective flux map for trench (top) and via(bottom). 100 particles were used at each test point, note the large variance, or scatter, in the flux estimate. Aspect ratio of each feature was 3.0 and the sticking coefficient was 0.1

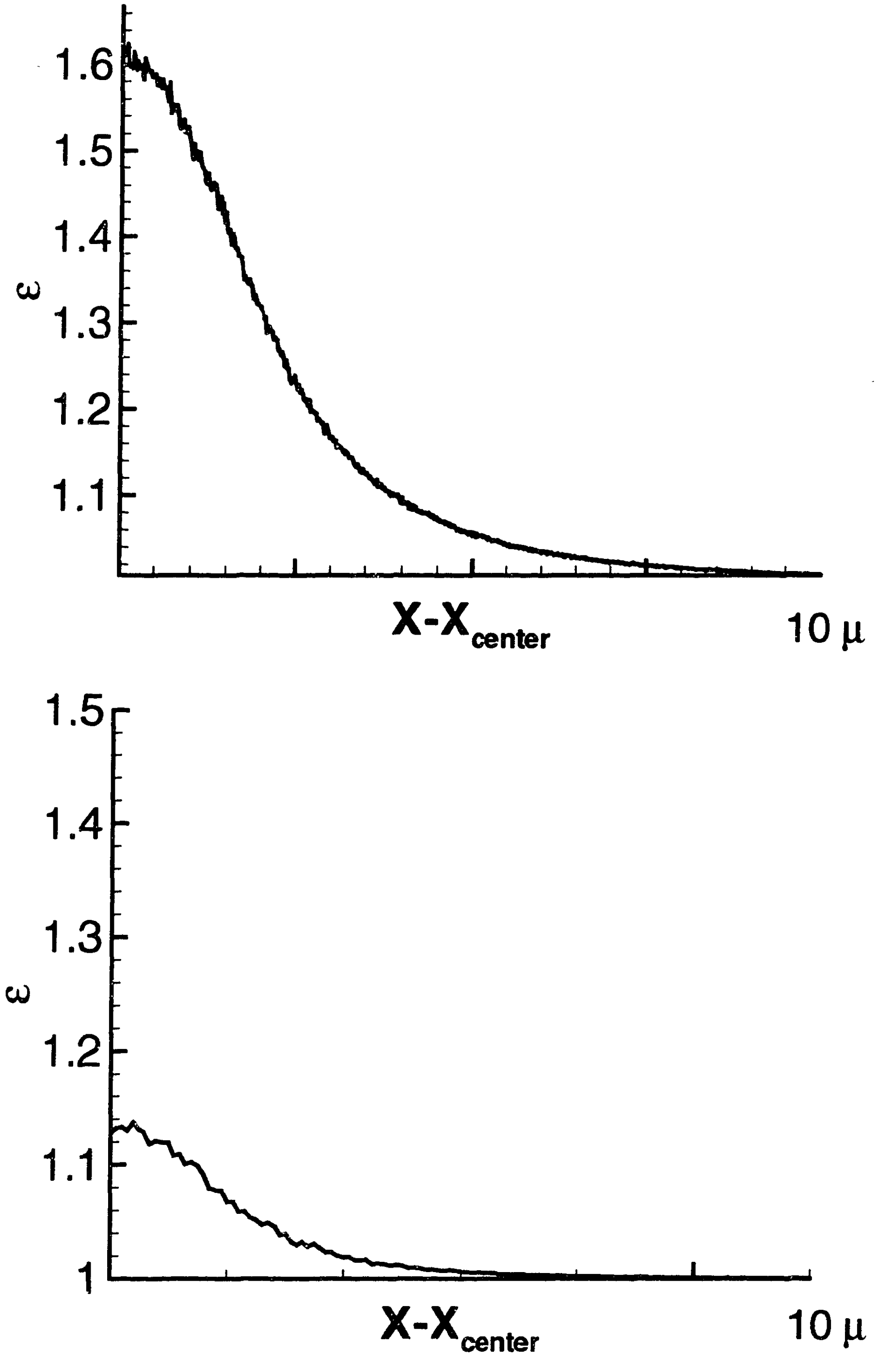


Figure 2. 5Effective flux map for trench (top) and via(bottom). 1000 particles were used at each test point. Aspect ratio of each feature was 3.0 and the sticking coefficient was 0.1

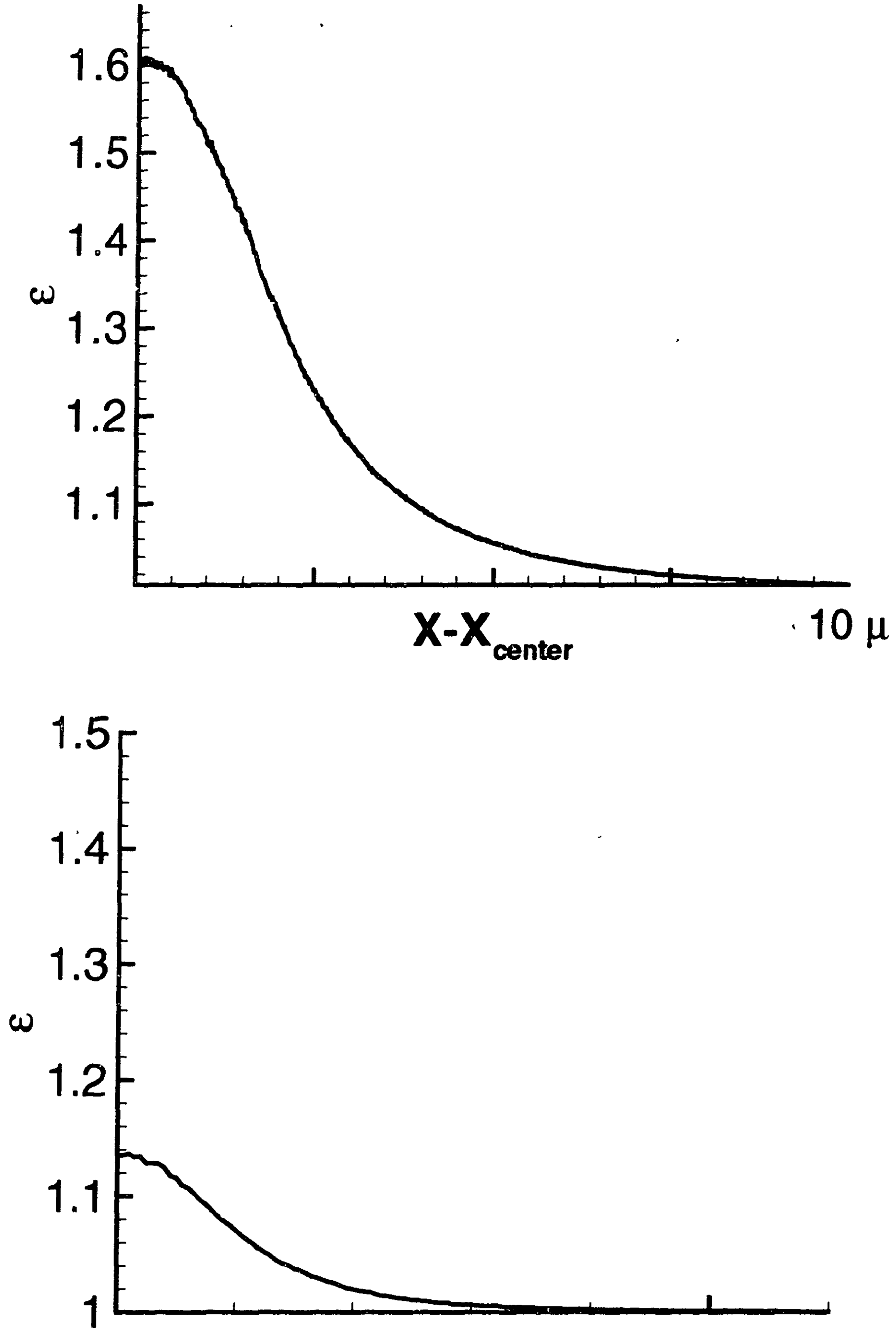


Figure 2. 6Effective flux map for trench (top) and via(bottom). 1000 particles were used at each test point. Aspect ratio of each feature was 3.0 and the sticking coefficient was 0.1

2.4 The Moelcular trajectory as a Markov Process

A very important simplification can be made for processes such as thermal CVD where precursor molecules come to equilibrium with the surface before being re-emitted. Such molecules are said to be diffusely reflected, with a distribution of speeds determined by the surface temperature and a velocity distribution that is a cosine around the local surface normal. Thus, the following relation holds Eq. 2-11 that is, that upon making a collision with a wall, the molecules lose all connection with their previous trajectory. This lack of memory is known as the Markov property.[10]

$$P(x_i|x_{i-1}) = P(x_i|x_{i-1}, x_{i-2} \ldots x_1) \qquad (2\text{-}11)$$

Furthermore, this implies Eq. 2-12

$$P(x_1, x_2, x_3 \ldots x_n) = P(x_2|x_1)P(x_3|x_2)\ldots P(x_n|x_{n-1}) \qquad (2\text{-}12)$$

This 'memory less' property means that the trajectory of any molecule in the feature scale domain may be described as a 'random walk', and the probability of any trajectory imaginable may be computed by multiplying together the 'one-step' transmission probabilities. The one step transmission probabilities could be pre-computed before the rest of the simulation by launching test particles from each possible x_i in the system and taking statistics on the location of their next visit. This approach is employed to evaluate the summation term in Eq. 2-10, and results in a large speed up over simple MC.

In the event that the transmission probabilities can be determined analytically, a several thousand fold speed-up is possible, this will be detailed in the next section. Finally, within the MC framework, in cases where angular distributions for the re-emitted molecules are known *a priori*, as in thermal CVD, low discrepancy Sobol sequences can be employed to considerable advantage for variance reduction.[2].

2.5 From feature scale to macroscale

The effective reactivity of a surface with several features may be computed with the following approach, exemplified for a via of aspect ratio 3.

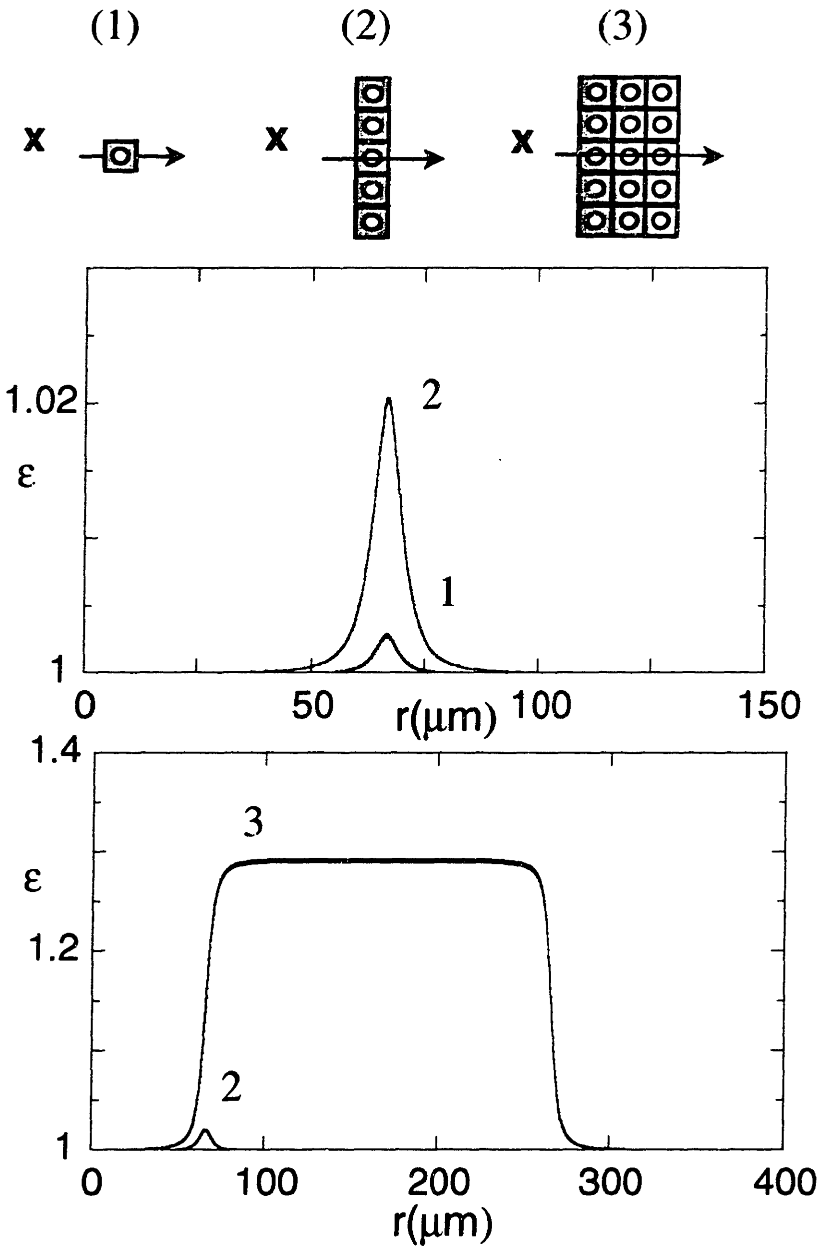


Figure 2. 7 Construction of a die-scale reactivity map for circular vias. Individual results are superimposed to make clusters which may then be superimposed to map at any length scale.

Beginning with a single feature, an effective reactivity map is found by generating 10^3-10^4 test particles from several hundred map points. If the reference value of ε, ε_o is subtracted from this map, the difference represents the marginal contribution of the feature to the effective reactivity. Once individual feature scale ε functions have been generated for each feature geometry present, they can be superimposed, serving as 'building blocks' to map the effective reactivity of heterogeneous fields of features to the reactor scale. This superposition is possible because no collisions are assumed to take place under the macro-micro interface. Thus, if a molecule is emitted from the surface, it will return to the macroscopic domain before visiting the substrate again. Molecules leaving one feature cannot visit another feature before returning to the macroscopic environment, and no mechanism for 'cross talk' between features exists aside from the macroscopic model.

We now present some scaling arguments for the size of the cross-talk effect. One million test trajectories were simulated for particles leaving a substrate in thermal equilibrium with the surrounding gas. The particles were followed until they reached the control surface above the emitting substrate or were scattered back to the substrate surface *Kn* denotes the ratio of mean free path to control surface height. At the end of the trajectory the particles were put in bins according to the number of collisions they had suffered on their trip.

This problem has been studied within the context of physical vapor deposition (PVD) where considerable attention is directed to developing models of weakly collisional transport.[11-13] While these studies have all addressed the issue of 'back diffusion', or 'inverse diffusion', where atoms leaving a source undergo collisions with the background gas phase and are re-directed back to the source, most of the emphasis was on determining the angular and energy distribution of material arriving at the surface from the sputter chamber.

These simulations have suggested fewer collisions and less back-scattering[12] than are found here in table 2.1. There are two likely reasons for this, first, sputtered particles in PVD systems are emitted with much more kinetic energy than the thermalized particles simulated here, and so have a correspondingly longer mean free path. Second, a hard sphere potential was used in simulating molecular collisions, which will overstate the role of collisions relative to simulations using the 'softer' potentials like variable soft sphere (VSS).

Table 2. 1 Monte Carlo simulation of collisions under the macro-micro interface.

Collisions	Kn=100	Kn=10	Kn=5	Kn=3	Kn=2	Kn=1.0
0	996841	967266	934146	890475	833677	285635
1	3159	31472	61160	95135	136714	345871
2	0	1262	3911	10826	19447	169555
3	0	0	734	2751	6017	83098
4	0	0	49	500	2275	37765
5	0	0	0	125	975	20041
5+	0	0	0	188	895	57535
total transmitted	998916	987937	974869	956404	929157	669537
total reversed	1083	12063	25131	43596	70843	330463

Consideration of the above results suggests that as long as the control surface is not more that 0.33 λ above the substrate, errors due to neglecting cross-talk will be small, and the superposition approach will not introduce artifacts into the results.

The resulting reactivity maps must meet two criteria. First, the MC results must exhibit variance, or "scatter," small enough to allow superposition. Eq. 2.13 was taken as an estimate of the error, δ, in ε.

$$\delta = \sqrt{\frac{E(\varepsilon^2) - (E(\varepsilon))^2}{\beta}} \qquad (2.13)$$

Sufficiently small errors were achieved at most map points using less than $2*10^3$ test particles. Second, as mentioned previously, gradients in ε must be resolvable on a finite element mesh to allow formulation of a flux boundary condition on the continuum model as:
where j is the net flux leaving the macroscpoic model at X, ϕ^{in} is the total flux leaving the reactor scale domain and P_o is the reference reaction probability

For collisionless flows, gradients in ε are dictated by the starting height, h, of the test trajectories. The source plane should be located at the smallest h, where gradients in ε become resolvable upon superposition (any closer and the continuum equations would be invalid) while a source farther away would result in simulation of a large number of gas phase collisions, and thus excessive computational effort. In this work, the source plane height was taken to be at 0.33λ. Use of a ballistic transport integral to determine (β^*/β) shows that gradients in ε will have a scaling length of about $20h$ or just less than 7λ. While this is very near the limit of a continuum formulation,[14] it should be noted that collisionless flows present the steepest gradients possible, and as such represent a 'worst case' picture of loading effects. If gas-gas collisions are allowed, gradients in ε will be smoothed, facilitating representation of ε on a finite element mesh.

Given ε to complete the boundary condition in Eq. 2.4, the macroscale model can now be solved with FEM to provide concentration and temperature profiles to serve as imputs for the microscale simulation. Iteration between the two domains will produce mutually consistent micro- and macroscopic environments, allowing tracking of feature scale morphology and its effects on macroscopic quantities of interest throughout the course of deposition. The resulting algorithm for consistent multiscale simulations is shown in Fig. 2.8

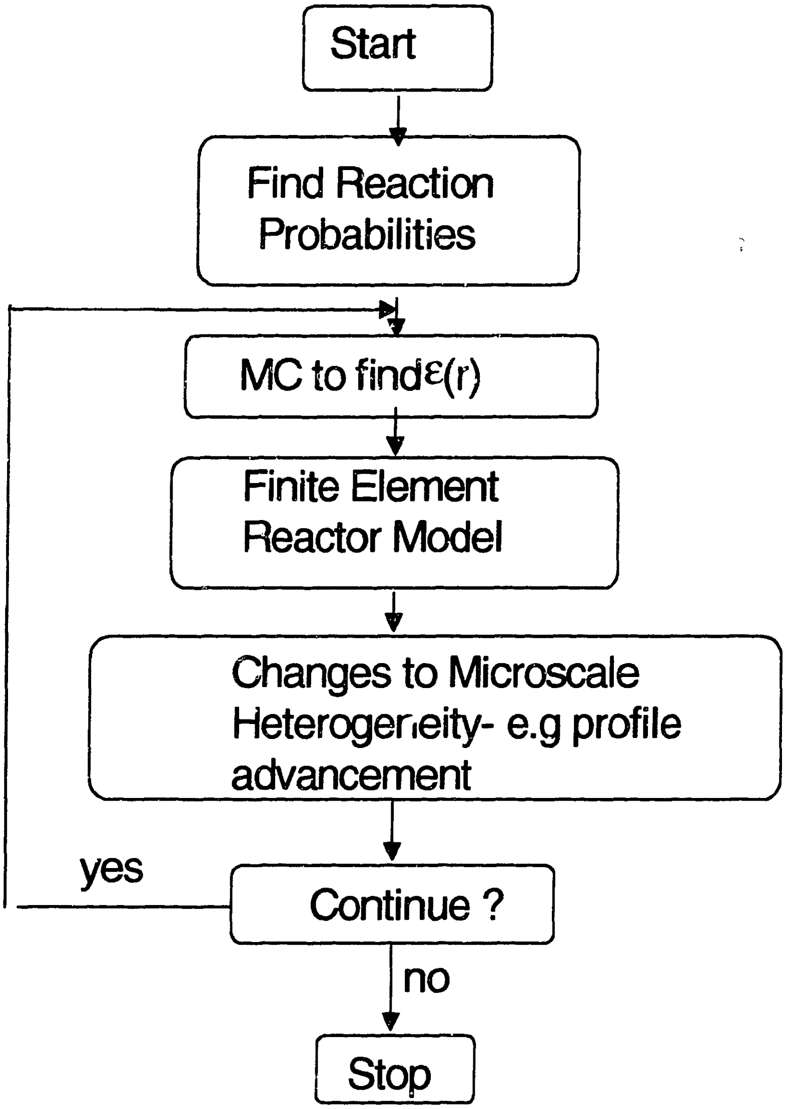


Figure 2. 8 Flowchart for full linked simulation using Monte Carlo calculations for feature scale models. Feature scale and reactor scale models are matched in an iterative fashion.

In the case study of tungsten CVD presented in section 2.6, the majority of simulation time was spent in calculations at the feature scale. Approximately 30 minutes was required for profile evolution at each time step, for each trench and 150 minutes per via monitored. Computing ε functions required about 1 hour for each feature geometry present. Computations were executed on an IBM workstation capable of 22 Mflops, and could be parallelized to run considerably faster than the figures reported here.

2.6 Example Problem

2.6.1 Problem Description

Although deposition of tungsten from hydrogen-reduced tungsten hexaflouride has been chosen as a case study the methods presented here are suitable for any CVD problem requiring linked treatment of transport at the continuum and discrete particle levels. We have chosen CVD of tungsten because it is relatively familiar, in particular because information on typical deposition conditions is available in the literature.

The operating conditions for this example were chosen to be typical for CVD tungsten with a few modifications. Temperature was taken to be 750K, this is at the high end of the range of temperatures commonly used, resulting in poorer than usual film conformality. This choice was made to illustrate the importance of microscale transport resistance. Total pressure was 1 torr, the mole fraction of WF_6 at the inlet was taken as 0.005, with the remaining carrier gas 90% H_2 and 10% Ar. As the goal of this work was to develop a linking methodology, and not to explore tungsten chemistry, we chose a simplifed kinetic model for WF_6/H_2 deposition from the considerable body of work available. Deposition was modeled with a single heterogeneous decomposition reaction, Eq. 2.14 and the rate of reaction was calculated with an expression (Eq. 2.15) of the form proposed by Charra and Cale.[15]

$$WF_6 + 3H_2 \rightarrow 6HF + W \tag{2.14}$$

$$R_s = 2.0\exp\left(-\frac{8300}{T}\right)\frac{P_{H_2}^{0.5}P_{WF_6}}{1+1000P_{WF_6}} \tag{2.15}$$

The chosen partial pressure of WF_6 is somewhat lower than that normally used, as we wished to minimize the error in treating the kinetics as first order at the feature scale.

A single wafer axisymmetric reactor, consistent with current CVD reactor designs, is chosen as the macroscale system. The substrate is a 200 mm diameter wafer, with 2 annular fields of features chosen to provide microscale heterogeneity typical of microelectronic device features.

Field 1 is comprised of 0.33 µm wide trenches spaced 0.33 µm apart. This field covers the substrate from r = 17 mm to 33 mm. The first half of field 1 consists of aspect ratio 3 trenches, the second half consists of aspect ratio 2 trenches. This geometry highlights the effects of microscale transport resistance, *i.e.*, when reaction probabilities are high, there is little difference between deep and shallow trenches. Both appear as 'reagent sinks' when viewed from the macroscopic environment.

Field 2, ranging from h= 50 mm to h = 70 mm, is composed of aspect ratio 3 trenches interspersed with lines of aspect ratio 3 vias. The layout of this field illustrates calculation of ε over features of varying geometry; in particular, the combination of fully 3-dimensional effective reactivity maps over circular vias and with the 2-dimensional maps obtained over long trenches.

2.6.2 Results and Discussion

The results of linked model calculations for WF_6 profiles at the start of deposition are shown in Fig. 2.9. Curve A is the solution obtained for an unpatterned; (homogeneous) substrate, and is shown for reference. Curve B represents the WF_6 concentration profile at the start of deposition. Not only do patterned regions deplete reagents significantly faster than flat areas, the growth rate over the entire wafer surface is depressed, as observed in the work of Kleijn *et al.* for silane reduced tungsten.[16]

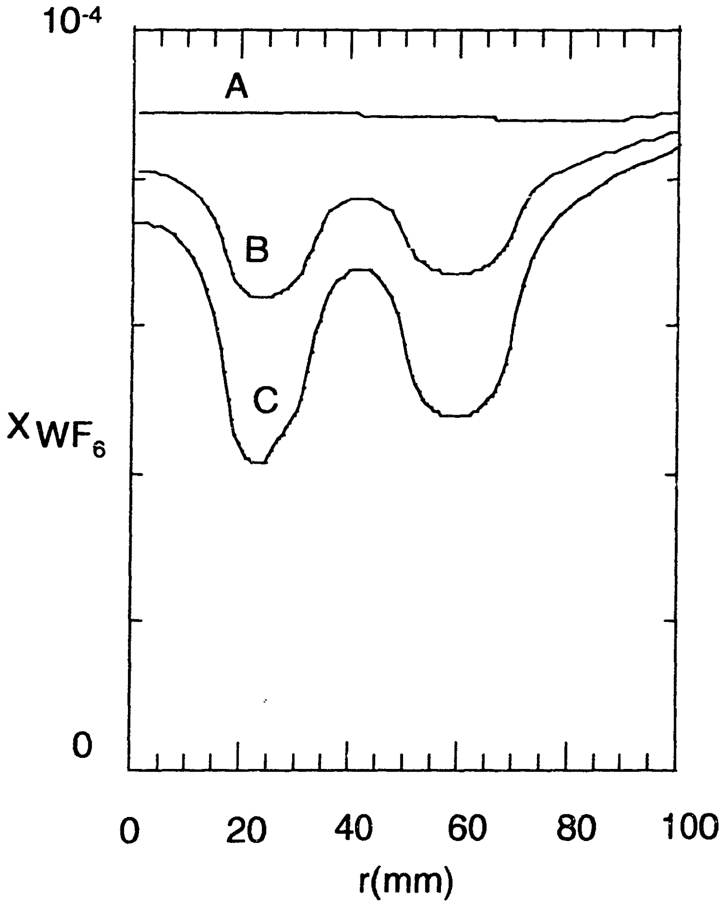


Figure 2. 9 Computed WF_6 profiles at the start of deposition. (A) over flat wafer, (B) over patterned wafer as predicted with effective reactivity calculations, (C) over patterned wafer with microscale transport resistance neglected – i.e. employing a simple 'exposed area' concept.

Film uniformity suffers considerably with the introduction of microscale heterogeneity. Maximum growth rate differences are approximately 25% for the patterned substrate B, compared to less that 1% for the flat substrate represented by curve A. Furthermore, considerable growth rate variation (about 15%) is seen from the center to edge of feature fields. The importance of microscale transport resistance is apparent when comparing curve B and curve C. Curve C derives from a 'simple perimeter' concept to account for the additional surface area offered by device features. By neglecting microscale transport resistance, precursor consumption and loading effects are over-estimated. Another example of microscale transport resistance is shown in the shape of curve B. Note that the first field is nearly symmetric about its midpoint, $r = 25$ mm, in spite of the inner half trenches being composed of aspect ratio 3, while the outer trenches are aspect ratio 2.

As deposition progresses, features are filled with deposited films, and the effective area presented to the macroscopic environment changes. The time evolution of WF_6 concentration profiles appears in Fig. 2.10. The circular vias in the outer field begin to close rapidly, causing the outer field to quickly resemble a flat substrate.

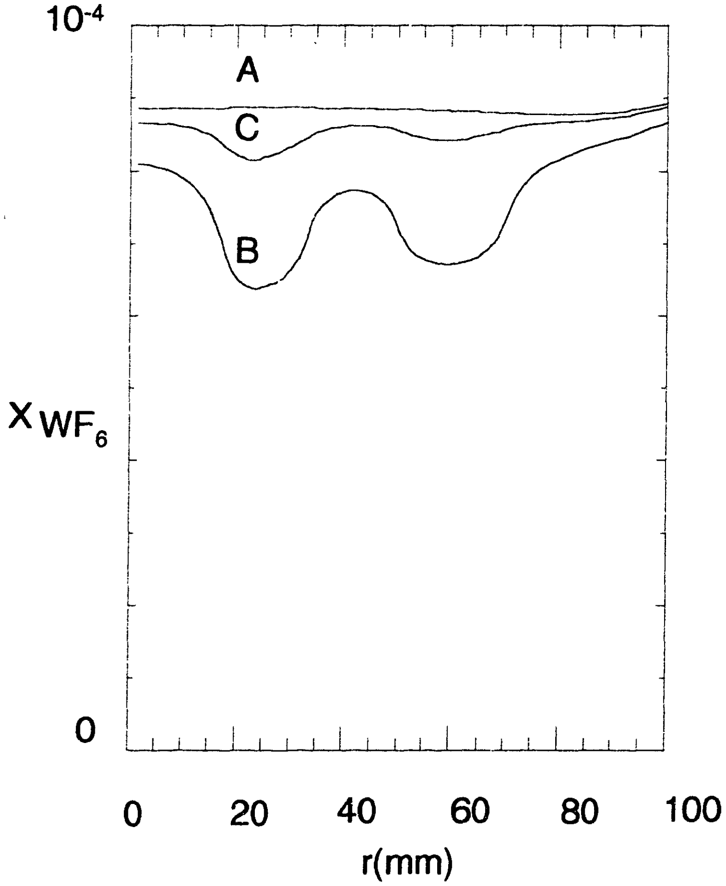


Figure 2. 10 Time evolution of computed WF6 profiles (A) over flat wafer, identical to that at the end of deposition, (B) over patterned surface at the beginning of deposition. (C) over patterned surface at the end of deposition.

Examination of feature profiles from field 1 at the end of deposition shows that very little flux actually penetrates to the bottoms of the aspect ratio 2 or 3 features. The features behave as 'sinks,' and thus the extra area offered by the deeper features is not 'seen' by the macroscopic environment (see Fig. 2.11). Thus, the effect of these different geometries is actually quite similar from a macroscopic perspective.

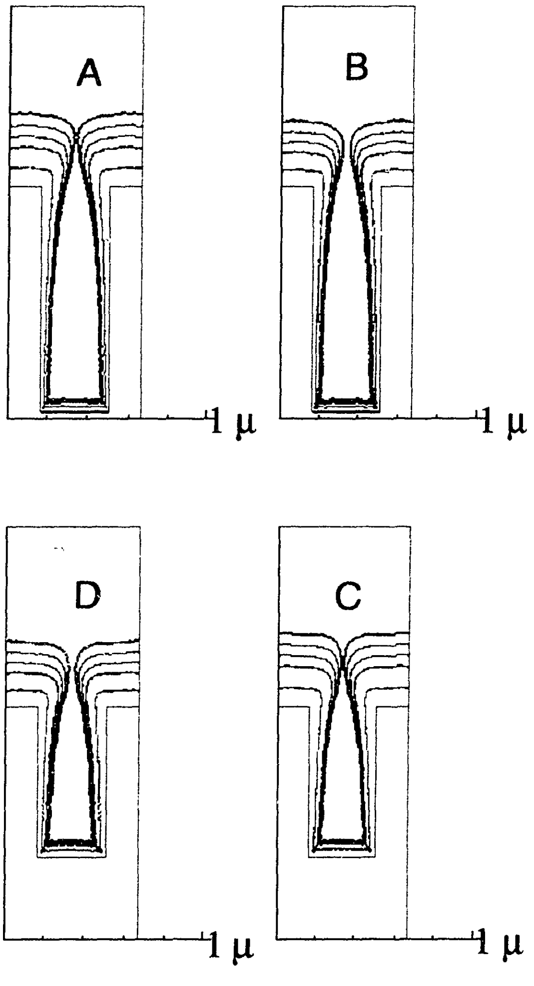


Figure 2. 11 **Calculated feature profiles at the end of deposition (inner field). Film thickness at field edges is roughly 115 % of that at the field center. Clockwise from upper left: (A) aspect ratio 3 trench at r=17 mm (B) aspect ratio 3 trench at r=25 mm (C) aspect ratio 2 trench at r=25 mm (D) aspect ratio 2 trench at r=33 mm.**

Post-deposition profiles from field 2 also show a significant difference-roughly 10% from center ($r = 60$ mm) to the edge ($r = 50$ mm). Shadowing is particularly pronounced in the vias, which received almost no bottom or sidewall deposition.

(see Fig. 2.12).

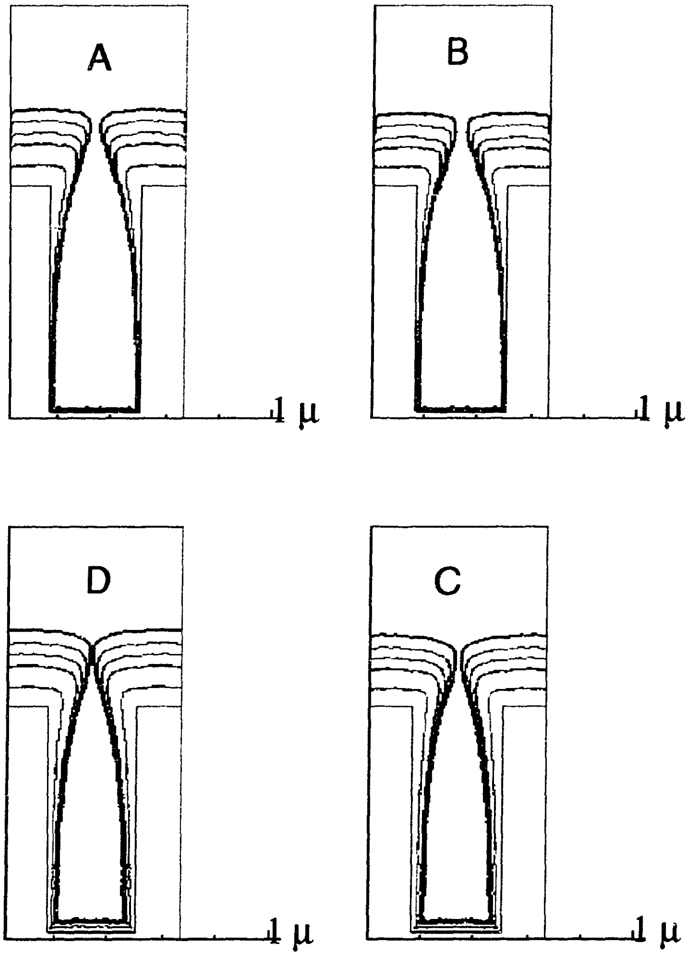


Figure 2. 12 Calculated feature profiles at the end of deposition (outer field). Film thickness at field edges is roughly 110 % of field center. Clockwise from upper left: (A) aspect ratio 3 via at r=50 mm (B) aspect ratio 3 via at r=60 mm (C) aspect ratio 3 trench at r= 50 mm (D) aspect ratio 3 trench at r=60 mm

2.7 Conclusions

A probabilistic, Monte Carlo method for linking continuum and discrete particle transport models has been developed. The resulting linked model can be used to investigate interactions between reactor scale and feature scale processes in low pressure CVD. We observed that reactor-feature scale interactions affect film conformality and growth rate on both the micro scale and over the wafer as a whole, thus exerting a large influence on the quality of finished devices. These interactions will become even more important if wafer sizes and pattern densities continue to increase according to current trends!

The linking algorithm presented in this chapter employs an effective reactivity function ε, which is computed through a hybrid probabilistic-deterministic MC method. E can be interpreted as a representation of the average fate of molecules entering the feature scale domain from the macroscopic model.

While submicron trenches and vias were taken as typical structures for this work, a map of ε can be calculated for any feature suitable for MC simulation. Calculations at the feature scale include microscopic transport resistance, which cannot be treated by continuum models because of the particulate nature of the gas phase at the micron length scale. Under conditions of free molecular flow, or nearly free-molecular flow, at the feature scale, ε calculations over individual features can be superimposed to yield maps over macroscopic fields of features. In this fashion, the effective reactivity of an entire substrate can be mapped. The effective reactivity concept presented here can be applied to other CVD processes where feature scale heterogeneity is important at the macroscale, such as selective growth. Treatment of collisional flows is also possible within the framework presented, but collisionless calculations are valuable as a 'worst case' picture of loading effects, in the absence of smoothing of the ε functions resulting from gas phase collisions.

In spite of these successes, the Monte Carlo method is not always optimal from a practical point of view - the biggest problem quite simply is speed. While adequate performance was achieved in this example by selection of simple chemistry and aggressive optimization of the MC method, the method is probably not affordable for difficult, multi species chemistry, particularly if reactive sticking coefficients are not given *a priori* to describe precursor-surface interactions. Such flux-dependent calculations are possible, however, using the deterministic methods presented in the next section.

1. Kalos, M.H. and P.A. Whitlock, *Monte Carlo Methods*. 1986: John Wiley & Sons.
2. Press, W.H., *et al.*, *Numerical Recipes*. 2nd ed. 1992, Cambidge: Cambridge University Press. 963.
3. Bird, G.A., *Astronaut. Aeronaut.*, 1981. **74**: p. 239.
4. Spanier, J. and E.M. Gelbard, *Monte Carlo Principles and Neutron Transport Problems*. 1964, Reading, Mass: Addison-Wesley.
5. Cooke, M.J. and G. Harris, *J. Vac. Sci. Technol. A*, 1989. **7**(6): p. 3217-3221.
6. Coronell, D.G. and K.F. Jensen, *J. Computer-Aided Materials Design*, 1993. **1**(1): p. 3-26.
7. Gilmer, G.H., *et al.*, *Materials Science and Engineering B*, 1996. **37**: p. 1-7.
8. Smy, T., K.L. Westra, and M.J. Brett, *IEEE Transactions on Electron. Devices*, 1990. **37**(3): p. 591.
9. Wulu, H.C., K.C. Saraswat, and J.P. McVittie, *J. Electrochem. Soc.*, 1991. **138**(6).
10. Fishman, G.S., *Monte Carlo Concepts, Algorithms and Applications*. Springer Series in Operations Research, ed. P. Glynn. 1995, New York: Springer-Verlag. 693.
11. Kersch, A., W. Morkoff, and C. Werner, *J. Appl. Phys.*, 1994. **75**(4): p. 2279.
12. G.M.Turner, *et al.*, *J. Appl. Phys.*, 1989. **65**(9): p. 3671.
13. Myers, A.M. and J.R. Doyle, *J. Appl. Phys.*, 1992. **72**(7): p. 3064.
14. Bird, G.A., *Molecular Gas Dynamics an dthe Direct Simulation of Gas Flows*. Oxford Engineering Science Series. Vol. 42. 1995, Oxford: Clarendon.
15. Cale, T.S. and M.B. Charra, *Thin Solid Films*, 1990. **220**: p. 19.
16. Kleijn, C., in *Computational Modeling in Semiconductor Processing*, M. Meyappan, Editor. 1995, Artech House: Norwood, Ma. p. 97.

Chapter 3
A Deterministic Method for Multiscale Modeling of Chemical Vapor Deposition

In this chapter, a determinstic method is presented for calculating effective reactivity at the feature scale. The method is several thousand fold faster than the MC methods presented previously, and enables treatment of multi-component chemistry. Although the linking method can be used with any feature scale model, a level set profile simulator is constructed to calculate the morphology or ‘step coverage’ of deposited films.

The biggest ‘plus’ of the deterministic formulation is its speed advantage over MC. A probabilistic MC simulation converges as the square root of the number of trials, thus if a calculation is to be accurate within a fractional error of 10^{-3}, at least a million test trajectories will be needed. In order to simulate surface chemistry, the flux of every species at every point of the feature scale model must be estimated with at least this level of accuracy. While a deterministic MC approach is also possible for such a particle transport problem,[1] it would be slower than the deterministic scheme presented here.

3.1 Detailed Chemistry at the Feature Scale

In this work, a ballistic integral transport model is combined with a detailed treatment of surface kinetics. Construction of the ballistic transport models has been described in detail previously[2, 3], but the key points will be highlighted here to clarify discussion of our linking method. The surface is discretized into elements, x_i, with material reaching element x_i directly from the deposition chamber, or after being re-emitted by another surface element, x_j of the feature.

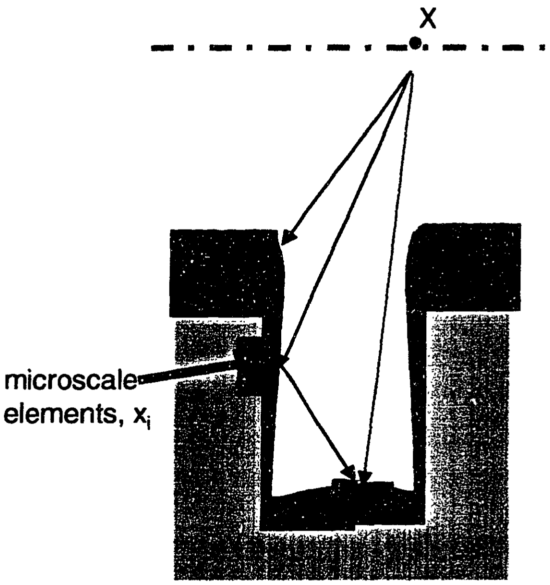


Figure 3- 1 The evolving film front is discretized into a set of segments, xi. Flux is assumed uniform over each segment. Linking to the macroscopic is accomplished by solving the feature scale problem to determine the flux at each segment. Then the transmission probability is determined to points X in the macroscopic domain.

A mass balance, Eq 3.1, on element *xi* gives:

$$\phi^{in}(x_i)A_i = \phi^{direct}(x_i)A_i + \sum_{x_j \neq x_i} \phi^{out}(x_j)A_j Q(x_j \rightarrow x_i) \tag{3.1}$$

Where $\phi^{in}(x_i)$ and $\phi^{out}(x_i)$, represent the incident and emitted flux densities at element x_1, and A_i is the area of element x_i. Q is the matrix of transmission probabilities, or 'shape kernel' with $Q(x_i \rightarrow x_j)$ representing the probability that a molecule emitted from x_i will make its next encounter with the surface at x_j. Because of the 'memory less' Markov property introduced in the last chapter, Q may be computed using an approach similar to the computation of view factors in radiative heat transfer problems. If the ϕ^{out} terms are a linear function of incident flux, Eq. 3.1 can be applied to each surface segment and the resulting set of linear equations may be solved directly. However, in many cases of practical interest, film growth reactions are described by Langmuir-

Hinschelwood (see table 1), or power law models where reaction rates are not linear functions of incident species fluxes. Thus, the emitted fluxes are known only implicitly, and an iterative approach must be employed. Equations 3.2 through 3.4 describe an iterative solution scheme.

$$\phi_0^{in}(x_i) = \phi^{direct}(x_i) \tag{3.2}$$

$$\phi_n^{out}(x_i) = \phi_n^{in}(x_i) - G\left[\phi_n^{in}(x_i), x_i\right] \tag{3.3}$$

$$\phi_{n+1}^{in}(x_i) = \phi^{direct}(x_i) + \sum_{x_j \neq x_i} \phi_n^{out}(x_j) Q(x_j \to x_i) \tag{3.4}$$

The flux arriving directly from the deposition chamber is taken as an initial guess, Eq.3.2. Subsequently, each iteration consists of two steps. Given the incident fluxes, compute the populations of adsorbed species and vacant sites on the surface, as well as reaction rates, $G(\phi^{in}(x),x)$, at each element. In this work, *G* represents a Langmuir-Hinschelwood description of surface kinetics, but in principle, any surface kinetic model can be used. The mass balance at each element is closed by calculating the amount of material to be emitted at the next iteration, (Eq. 3.3). Next, the new fluxes arriving at each surface element, are determined considering both the flux arriving directly from the reactor chamber, as well as emission by other surface elements (Eq. 3.4). Iterations are continued until the flux profile in converged. A simple power law model is presented to illustrate the effect of varying fluxes on sticking coefficients. Expressions like Eq. 3.5 have been used by Schmitz and Hasper,[4] as well as by Kuijlaars, Kleijn, and van den Akker.[5] to model CVD of tungsten from WF_6 in H_2.

$$rate = k[WF_6]^0 [H_2]^{0.5} \tag{3.5}$$

Writing a sticking coefficient for WF_6 as the reaction rate divided by the incident thermal flux Eq. 3.6, shows that the sticking coefficient will be inversely proportional to the incident flux and that using a single, constant value would be incorrect.

$$P^{rxn} = \frac{4k[WF_6]^0[H_2]^{0.5}}{[WF_6]\bar{c}} \tag{3.6}$$

Under mass-transfer limited conditions, precursor fluxes in the interior of a feature may be very different than those to the surfaces exposed directly to the reactor chamber. For a tungsten deposition mechanism like Eq. 3.5, this means the apparent sticking coefficient cannot be written a priori, and must be solved for with the iterative technique presented above. (Fig 3-2)

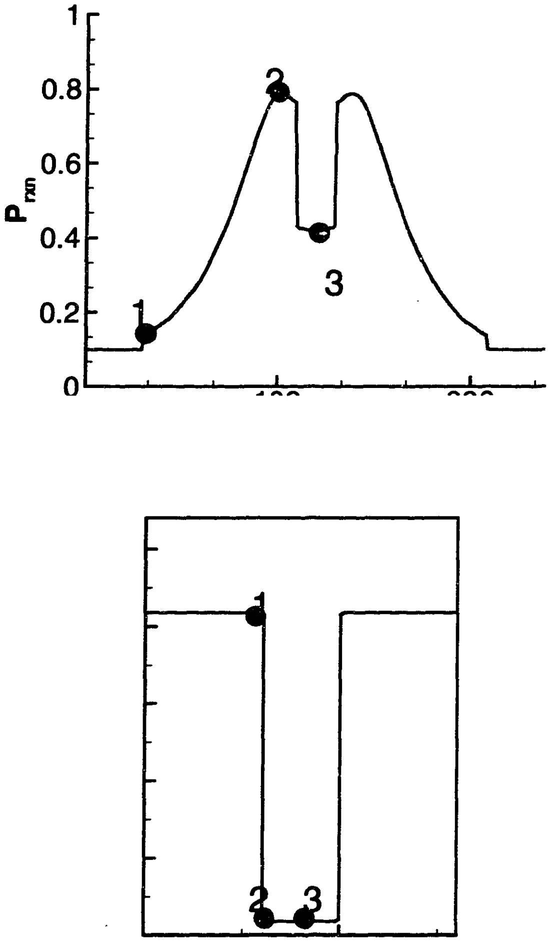


Figure 3- 2 Sticking coefficient as a function of position for power law model of tungsten CVD. Note surfaces starved for flux such as interior corners are relatively 'stickier' due to lower populations of adsorbates.

3.2 Alternative Formulation of the Multiscale CVD Problem

Given a detailed surface chemistry problem, it is convenient to work directly with conservation laws for the chemical species passing into and out of the microscale domain. Again write a mass balance at the control surface:

$$j\,(X) = \phi^{in}(X) - \phi^{out}(X) \tag{3.6}$$

Here, $\phi(X)$ represents the scalar flux density of molecules crossing the control surface, and $j(X)$ represents the net flux density at point X. The relationship between material emitted from all points on the heterogeneous substrate, x_j, and the flux entering the macroscopic domain at X is found by mass balance at the control surface, Eq.3.7.

$$\phi^{out}(X) = \sum_{\tilde{x}_j} \phi^{out}(x_j) A_j q(x_j \to X) \tag{3.7}$$

The transport kernel q, gives the flux density at point X on the control surface for a given emission from point x_j on the substrate surface. That is, $q(x_i \to X)dX$ is the probability a molecule emitted from point x_j on the substrate will next arrive at the control surface within dX of X. For ballistic integral calculations, q may be calculated in the same manner as Q in Eq. 3.1. If MC methods are used, q may still be evaluated[6] although the computation will be more costly. Once reaction and transport inside microscale topography have been accounted for, a macroscopic description of microloading, as a continuum gas phase-continuum surface interaction can be pursued. Thus, a multiscale problem is reduced to a familiar, single scale problem.

Scaling analysis at the control surface shows that lateral concentration gradients, and thus loading effects at X are proportional to the ratio of the apparent reaction rate to the diffusion rate. A surface Damkoehler number, *Da,* representing the ratio of reaction rates to diffusion rates, Eq. 3.8 has been used as a dimensionless metric of loading effects in continuum systems.

$$Da = \frac{kL}{D} \tag{3.8}$$

Where L is the width of the patterend area. All the effects of microscale heterogeneity are found in k, is the appearant surface reaction rate constant in (m/s), or in terms of fluxes, Eq. 3.9

$$k(X) = \frac{j(X)}{\phi_{in}(X)} \sqrt{\frac{RT}{2\pi M_w}} \tag{3.9}$$

Where R is the gas constant, T the temperature, and M_w the specie smolecular weight. In general, k will reflect a convolution of surface geometry, surface kinetics, and microscale transport resistance through variations in $j(X)$, all might be functions of position and time. The following section illustrates the computation of microscale flux maps, as well as the smoothing performed by the kernel, q, enabling combination of microscale information with continuum models.

3.3 Deterministic Calculation of Feature Scale Flux Maps

From the microscopic side, accounting for feature scale topography begins with determination of the fluxes emitted from each surface of the feature. The reason for this 'bottom up' approach is that fluxes of precursors, and thus reaction rates may differ considerably within a feature due to microscale transport resistance *i.e.* shadowing. The flux profiles in figure 3-3 illustrate the effects of microscale transport resistance.

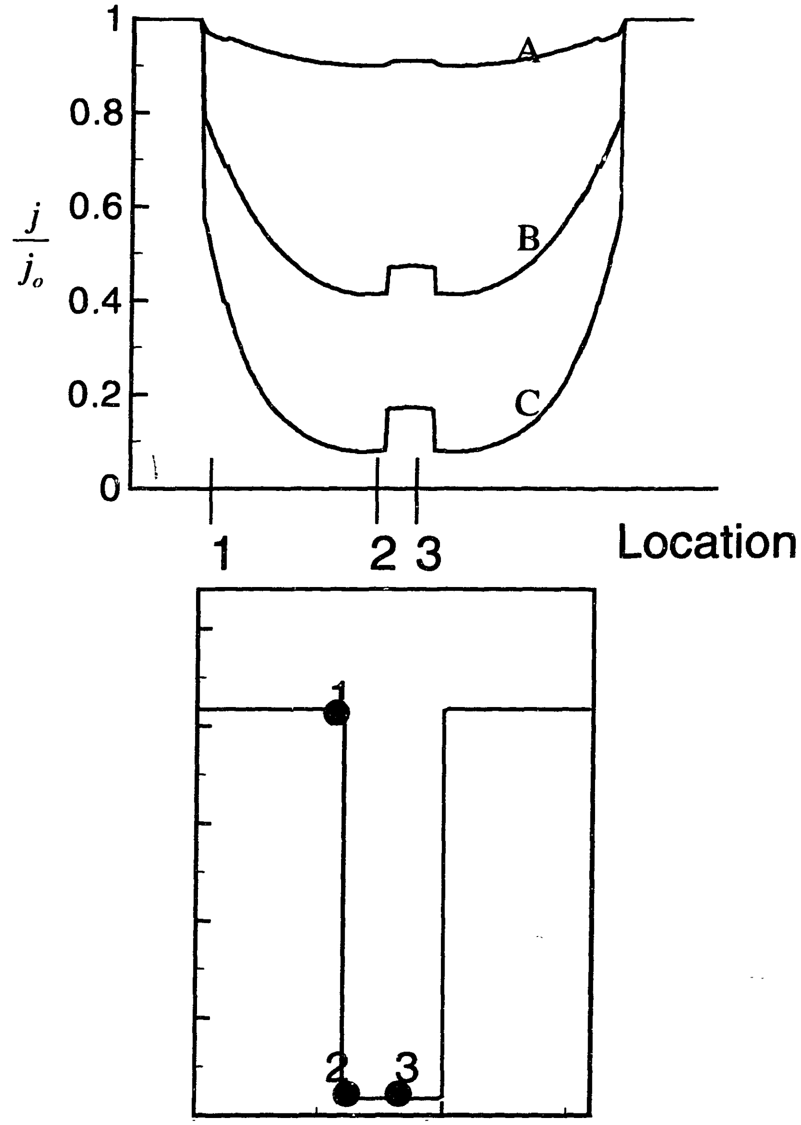


Figure 3- 3 Microscale transport resistance is seen in flux profiles for sticking coefficients 0.01 (curve A), 0.1 (curve B) and 0.5 (curve C).

As reactivity increases, less re-emission is possible, and the trench interior becomes less active, given the reduced supply of reactive species

Once the emitted fluxes are available, the contribution of this feature to the net flux at any point on the control surface is determined using Eq. 3-6. To illustrate the procedure, two example features are presented, (Fig. 3-4).

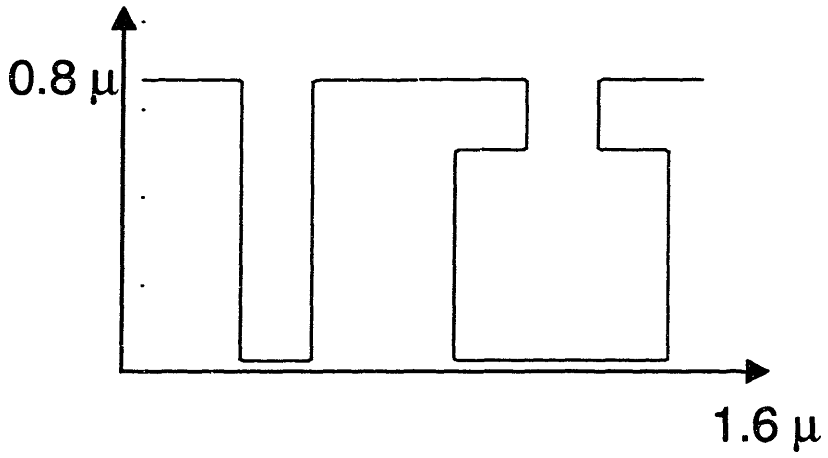


Figure 3- 4 Paired trench and overhang for demonstration of superposition technique

The trench and overhang features shown are arranged into 2 clusters. Each cluster contains two identical trenches interspersed with two identical overhangs, *i.e.* figure 3-4 is repeated twice in each cluster.

To simplify the discussion, sticking coefficient kinetics will be used. Panels A, B, and C of figure 3-5 show flux maps for sticking coefficients of 0.01, 0.1 and 0.5 respectively. While all flux maps show peaks over the additional surface area offered by the microscale topography, the effects vary according to the microscopic kinetics. Note that both surface area and microscale transport resistance will influence the net flux to the surface.

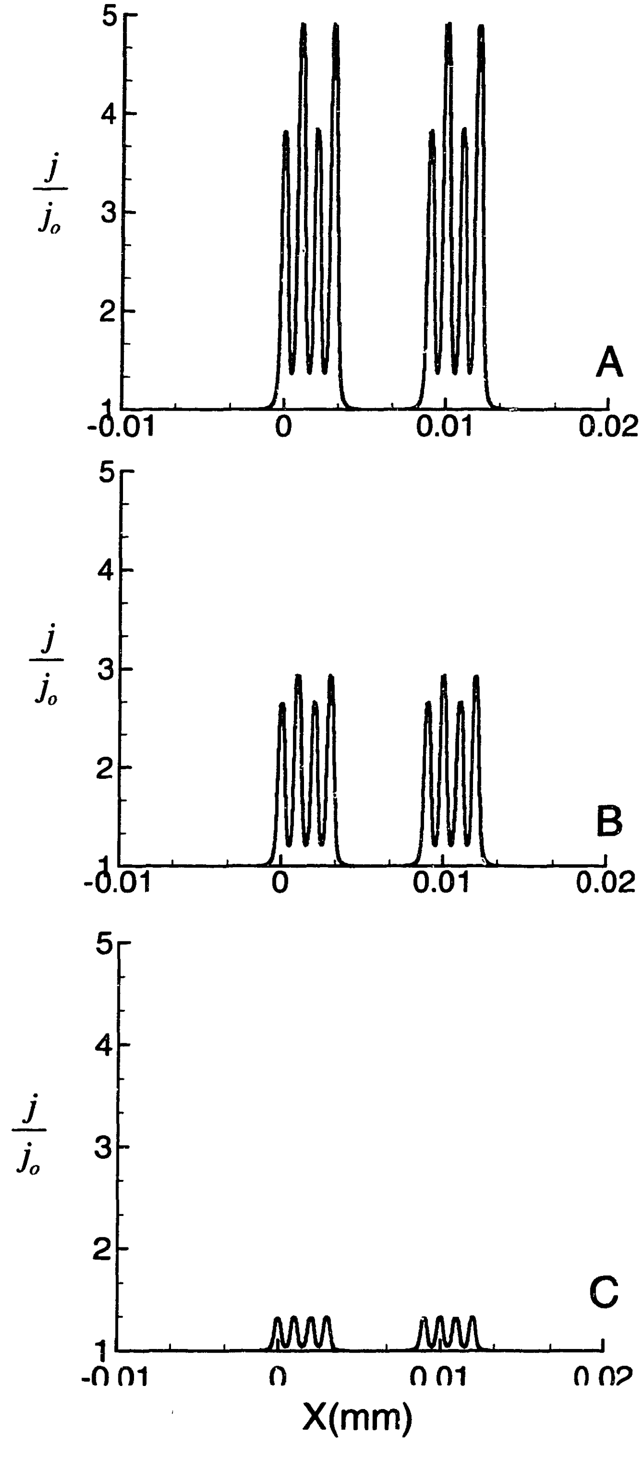


Figure 3- 5 Net flux to a patterned surface for sticking coefficients 0.01 (curve A), 0.1 (curve B) and 0.5 (curve C). Test features are located at x =1 μ, 3 μ, 9 μ, and 11 μ (trenches) and 2 μ, 4 μ, 10 μ and 12 μ (overhangs)

The effect of additional surface area is seen most clearly in panel A of figure. 3-5, when the high peaks in the flux map over the overhang structures are compared with the smaller ones over the trenches. However, the observed macroscopic flux map is also influenced by microscale transport resistance. figure 3-5 shows that the additional area provided by the microscale topography exerts a much smaller influence as the sticking coefficient is increased, as much of the reactive area is starved for precursor as seen previously (Fig. 3-3). By the time the sticking coefficient has increased to 0.5, panel C figure 3-5, the peaks in the flux map over the trench are not very different from those over the overhang. This is because when viewed from the macroscopic perspective, no material is returned after visiting the surface, and the features appear as 'sinks'.

Although the location of the control surface will not affect the total amount of flux, it does influence the shape of the flux map. The height of the control surface influences the 'view factors' or exchange probabilities that make up the transport kernel, q. As the control surface is raised above the substrate, molecules arriving at point X can see, and sample, conditions over an increasing area of the substrate surface. As a result, q maps fluxes over an increasing domain to X, leading to an 'averaging' over the substrate that is proportional to control surface height.

Note that as control surface is raised, individual feature maps rapidly coalesce, and gradients at the edges of the clusters begin to flatten. (Fig. 3-6)

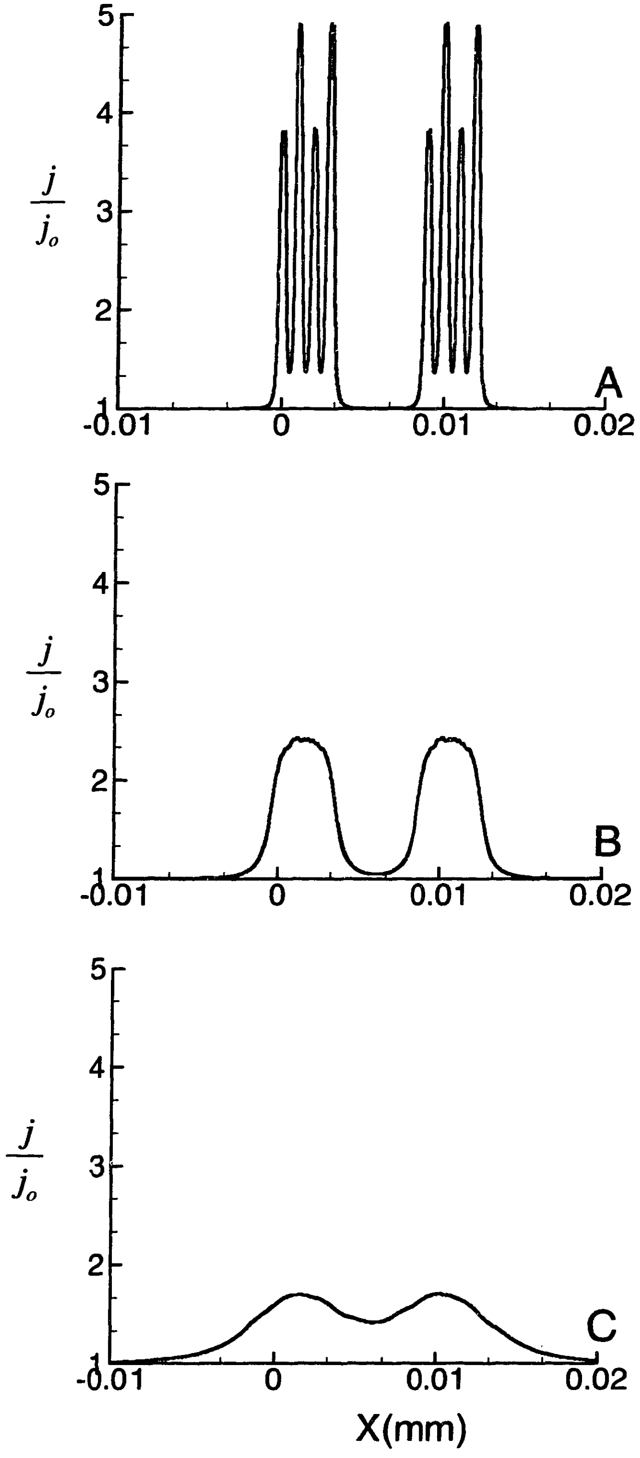


Figure 3- 6 Net flux to a patterned surface with variation of control surface height. Sticking coefficient 0.01 Test features are located at x =1 μ, 3 μ, 9 μ, and 11 μ (trenches) and 2 μ, 4 μ, 10 μ and 12 μ (overhangs)

The flux mapping method presented may be extended to any number of features in a striaght forward fashion. Recalling that the shape kernel, for a single feature will be a matrix of 'transmission probabilities', with element *Qij* representing the chance that a particle emitted from segment *i* of the trench shown in fig 3.7 will next encounter segment *j* of the trench, before escaping to the macroscopic environment. The stencil of a shape kernel for the single trench in Figure 3.7 is shown below.

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Q_{23} & Q_{24} & 0 \\ 0 & Q_{32} & 0 & Q_{34} & 0 \\ 0 & Q_{42} & Q_{43} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

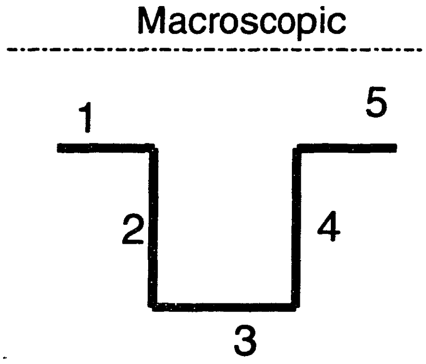
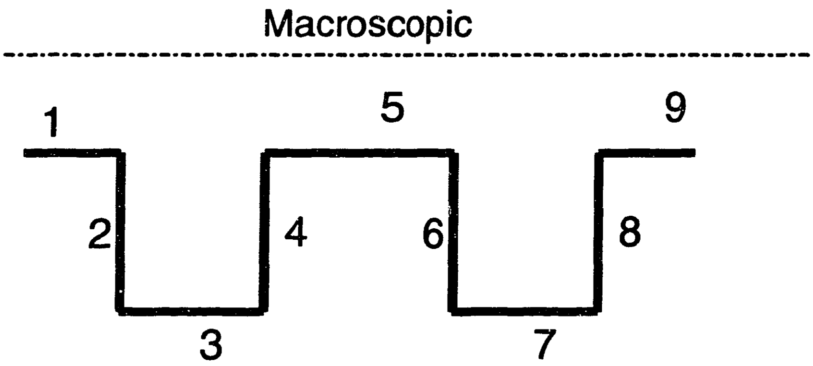


Figure 3- 7 Stencil of shape kernel fro a single trench. Interior segments interact and exchange material with each other, while exterior segments see only the macroscopic environemt

Note that while the interior segments of the trench see each other and exchange material, they are bounded on each end by segments that see only the macroscopic environment. Thus, as long as line of sight transport is assumed, a pair of features would produce a shape kernel with a pair of independent, non- interacting blocks, the double shape kernel, D, is shown below, with the D's representing the only non-zero entries. Fig 3.8



$$D = \begin{bmatrix} & D_{23} & D_{24} & & & & \\ D_{32} & & D_{34} & & & & \\ D_{42} & D_{43} & & & & & \\ & & & & & & \\ & & & & & D_{67} & D_{68} \\ & & & & D_{76} & & D_{78} \\ & & & & D_{86} & D_{87} & \end{bmatrix}$$

Figure 3- 8 Stencil of the shape kernel of a double trench, composed of non-interacting single features. The entire feature scale problem can be solved on feature at a time.

In principle, a shape kernel matrix could be written for the entire surface, with each piece of topography represented, by adding additional blocks to the D matrix. This would be a large task, as a typical integrated circuit layout contains several million features, each of which might be need to be represented by a hundred or so line segments. If 3-dimensional feature scale simulations are desired, the computational requirements would be still greater. Fortunately, this effort is not needed. Because the line segments features do not interact, (this is clear after looking at the D matrix !) each small feature scale problem may be solved individually, and their results superimposed to yield a map of the species uptake and generation over the substrate surface.

This is an appealing solution as long as the collisionless assumption remains valid. CVD processing at pressures greater than a few torr present a more challenging problem, in that the collisionless assumption is not valid at the feature scale. Under these conditions, transport is no longer line-of-of sight, and coupling terms arise, filling the off diagonal blocks in D to create a dense matrix. The bandwith of this dense matrix is determined by the height of the control surface.Ideally, the control surface should be located just where gradients in the flux map become resolvable on an FEM mesh, while avoiding costly Monte Carlo treatment of gas phase collisions.

3.4 A few Words About Level Set Methods

In order to take full advantage of the faster linking possible with a ballistic integral approach, a level set profile simulator was constructed. This section is not intended to be a complete discussion of level set methods given the abundant literature on the subject,[7] instead it is a description of how level set methods were implemented in this work, combined with some discussion on the pitfalls of the method. Level set methods were chosen for their flexibility, implementation is almost identical for deposition or etching.

The central idea of the level set approach involves viewing the evolving front as the zero level contour of a higher dimensional function, $\Psi(x,t)$.

The derivation given in[8] results in Eq. 3.10 which relates $\Psi(x,t)$ to the speed $F(x,t)$ of the growing front in the outward normal direction.

$$\Psi_t + F|\nabla\Psi| = 0 \qquad (3.10)$$

$$\Psi(x,t=0) \quad \text{given} \qquad (3.11)$$

We begin by considering the initial position of the front, a closed curve $\Gamma(t=0)$, we want to create an initial condition for Ψ that satisfies the property below:

$$\Gamma(t=0)=\left[x\middle|\Psi(x,t=0)=0\right] \tag{3.12}$$

A natural choice to create $\Psi(x,t)$ is to assign each point in the domain a value equal to its distance from the initial front if it is outside the closed curve and the opposite if it is inside. This is done at the start of simulation.

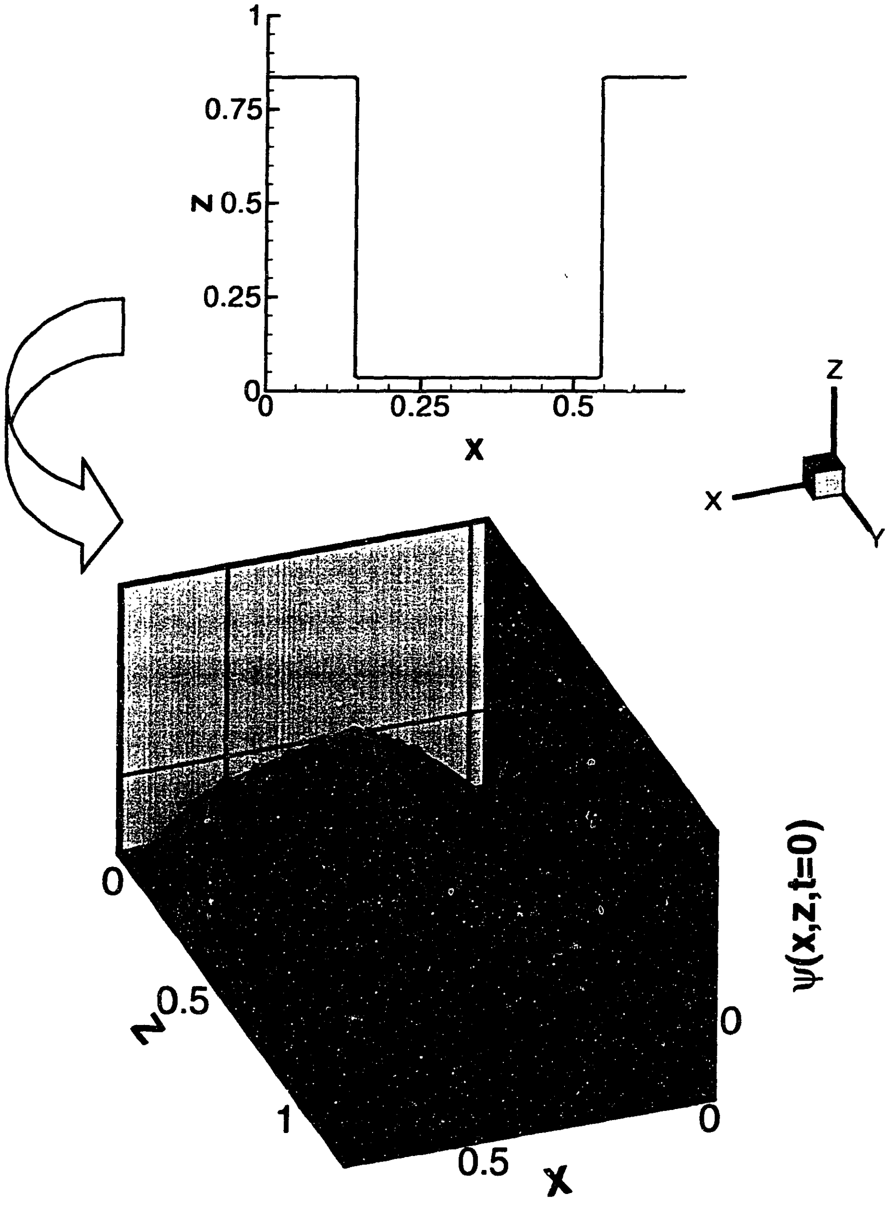


Figure 3- 9 Initialization of level set profile simulation. The desired initial profile is drawn (top), and the level set function, ψ is defined initially as the distance form the starting profile.

Then, the speed normal to the surface, F, (this is the growth rate for CVD), is determined using the ballistic integral transport model and the detailed surface kinetics mode described earlier. Of course in etching, F will be negative, but every other aspect of implementation is exactly the same.

One of the subtleties of applying the level set method to this problem is that although the speed function F, has meaning only on the film growth interface, it must be extended to all points where $\Psi(x,t)$ is defined in order for $\Psi(x,t)$ to be updated in time. This is known as the *extension problem* there are several ways of treating this problem described by Sethian.[8] The most intuitive approach is to assign each point in the domain the speed of the closest point on the zero-level curve, this is the approach taken here, and although somewhat time consuming, no problems were encountered.

Once the speed of each level curve in the domain has been found, a finite difference approach can be used to advanced the function $\Psi(x,t)$ *to* $\Psi(x,t+dt)$ if a suitable approximation to the gradient term can be found. The key issues in approximation of the gradient are related to 'upwinding'. This is not surprising, as Eq. 3.10 is essentially a hyperbolic partial differential equation. Fine detail, such as micro-trenching in plasma etching is difficult to resolve without adaptive meshing. At the end of the simulation, the zero level curve is extracted, which represents the original growth front, advanced in time.

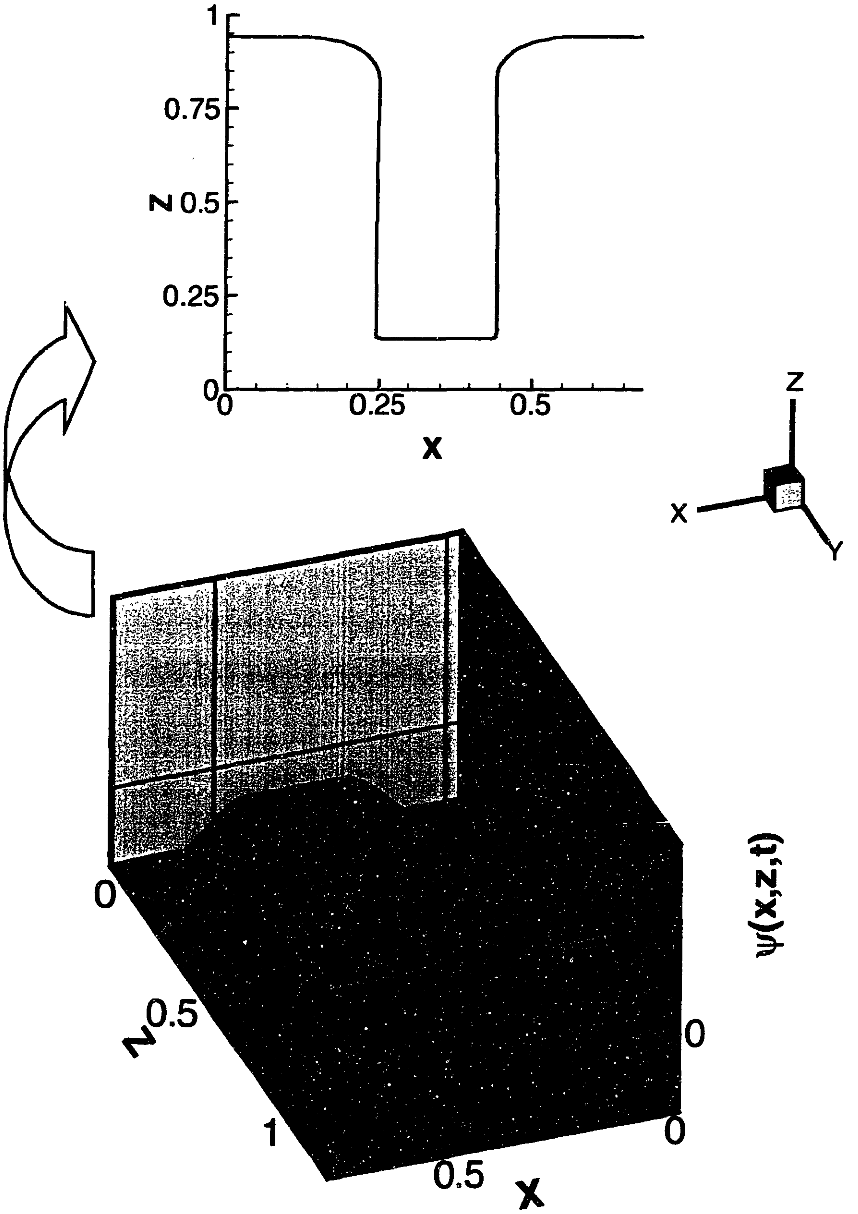


Figure 3- 10 Recovery of finished profile. At the end of of level set simulation, the zero contour line is extracted, representing the final shape of the front.

3.5 Putting it All Together

Beginning with an initial guess, *e.g* a flat substrate, the macroscopic model is called to provide a set of incident fluxes for the microscale simulations. Individual feature scale calculations are performed every few mean free paths along the substrate surface for each type of feature present. The resulting feature scale maps are then combined to make a map of the outgoing flux over the entire wafer surface. This flux is then used as a boundary condition on the macroscopic model, and a new iteration is begun. When the fluxes have been matched, the level set profile advancement model is invoked. As deposition progresses, features will fill, altering their geometry, and thus their flux uptake. Therefore after a small deposition, time the micro and macroscopic conditions are re-matched. A strategy for a fully linked simulation is illustrated in figure 3-9.

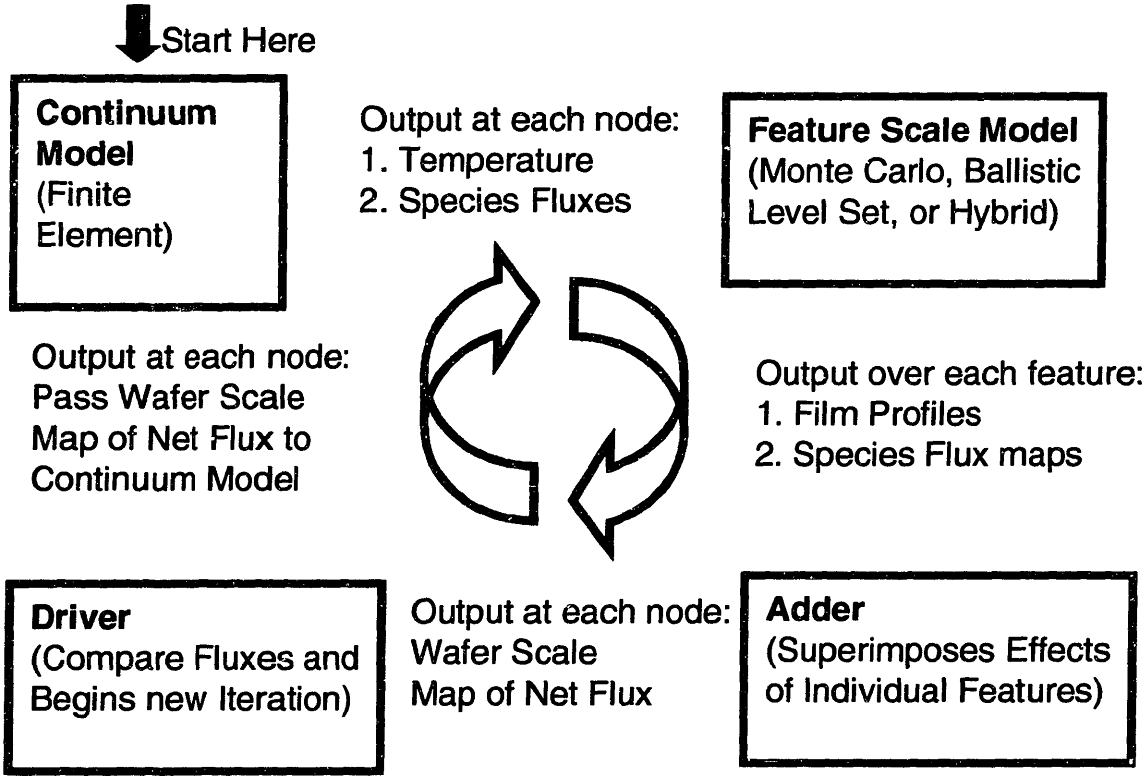


Figure 3- 11 Diagram of the iterative flux matching scheme. After fluxes are matched, the level set profile simulator is invoked to advance the surface.

At low pressures, ~ 1 torr computational burdens are not excessive. Ballsitic integral models of species transport execute every rapidly, and construction of the flux map requires only an additional line of sight calculation. Simulation of detailed surface chemistry has proven to be the most time-consuming part of the overall simulation. The examples presented below required approximately 16 hours on a 60 Mflops HP workstation for a complete, time dependent simulation, using 12 representative features, and matching micro and macroscopic fluxes 10 times throughout the deposition. However, less than an 20 minutes were required to obtain a self-consistent initial flux profile for examination of microloading effects.

3.6 Example: Aluminum CVD

Model Chemistry

CVD of aluminum from dimethylaluminum hydride, DMAH, has been chosen for this case study, as it is a promising precursor for aluminum interconnect applications.[9] DMAH exhibits complex chemistry on the substrate surface as well as in the gas phase. As the goal of this work was to develop a general linking methodology and not to explore aluminum chemistry, we consider deposition in the low temperature regime (450 K) where surface reactions are the rate-determining step. Furthermore, carbon incorporation does not take place at this temperature, and so is not considered.[10]

A simplified aluminum CVD mechanism was chosen[11], consisting of the reaction set shown in table 1.

Table 3- 1 Reaction mechanism for blanket aluminum CVD from dimethylaluminum hydride

No.	Reaction	Rate Law
1	$AlHMe_2 + S(D) = AlMe_2(D) + \frac{1}{2} H_2$	$k1[AlHMe_2][S(D)]$
2	$AlMe_2(D) + 2S(Al) = 2CH_3(Al) + l(Al) + S(D)$	$k2[AlMe_2][S(AL)]^2$
3	$AlMe_2(D) + CH_3(Al) = AlMe_3 + S(D) + S(Al)$	$k3[AlMe_2(D)][CH_3(Al)]$
4	$AlMe_3 + S(D) + S(Al) = AlMe_2(D) + CH_3(Al)$	$k4[AlMe_3][S(D)][S(Al)]$

The growing aluminum surface is assumed to offer 2 types of adsorption sites. Aluminum adsorbates are assumed to occupy only 'defect' sites on the growing film, while methyl adsorbates may populate the Al surface at large. Hydrogen desorbs readily and does not participate in surface reactions. Methyl is removed from the surface as trimethylaluminum, (TMA). TMA can also act as a growth inhibitor by competing for surface sites with arriving DMAH.[11],[10] The net flux of the byproduct TMA will be 2/3 of the DMAH consumption rate, consistent with the overall growth reaction presented in [12].

$$3DMAH_{(g)} \Rightarrow 2TMA_{(g)} + Al_{(s)} + 3/2H_{2(g)} \tag{14}$$

3.7 Deposition System

A single wafer axisymmetric reactor was chosen as the macroscale system. Blanket deposition was carried out over a 200 mm diameter wafer. Total pressure was taken to be 2 torr, and the inlet mole fraction of DMAH in argon was taken to be 10^{-4}. Operating conditions were chosen to achieve optimal macroscopic uniformity over the wafer, and thus highlight microscale effects. Two annular fields of features were introduced to provide microscale heterogeneity. The first field, positioned at 0.025 m < r < 0.04 m and was comprised of trenches 0.2m wide and 0.8m deep. The second field, located over the interval 0.06 m < r < 0.075 m, consisted of overhang structures 0.2μ wide at the mouth and 0.8μ deep. The individual features were spaced 0.8μ apart. Although each field contains exactly the same number of features, the overhang structures present larger surface area from reaction, and taken collectively, exert a larger effect on the macroscopic solution than do the trenches.

3.8 Linked Simulation Results

Fig. 3-11 shows calculated concentrations at the substrate surface for DMAH and TMA at the beginning, midpoint, and end of deposition. The concentration profile at the end of deposition is identical to that for a flat wafer. Conditions at the wafer surface differ significantly from the inlet conditions, and reflect the influence of feature scale topography.

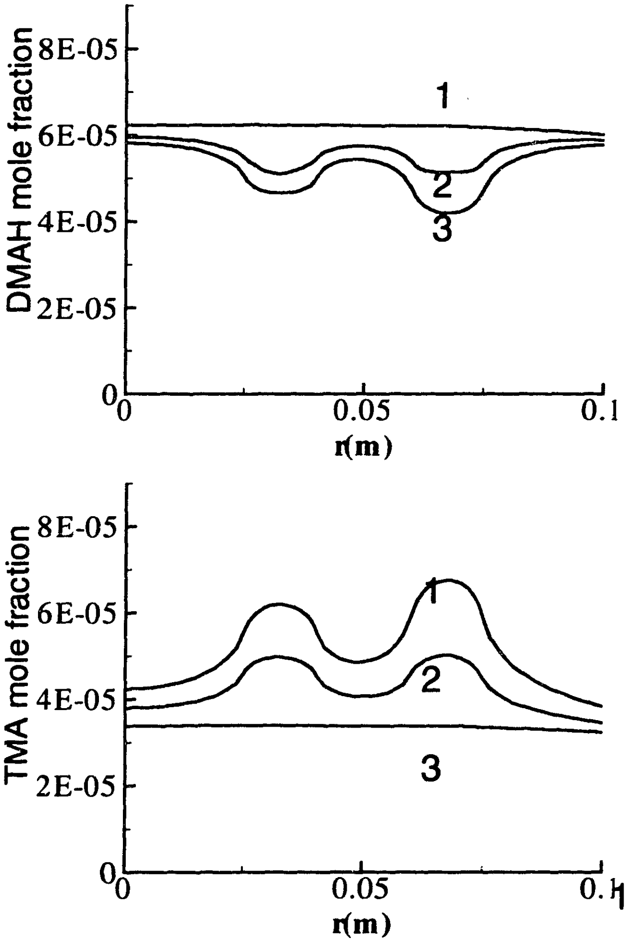


Figure 3- 12 Concentration profiles for DMAH and TMA at the start (curve 1), midpoint, (curve 2) and end (curve 3) of deposition.

Consideration of Fig. 3-10 shows that the actual conditions at the surface would have to be measured or guessed if a 'linked' multiscale simulation strategy were not used. Curve 1 demonstrates that at the start of deposition, patterned areas adversely affect film uniformity by depleting reagents faster than the flat areas. Concentration profiles over the patterned surface vary by approximately 25%, while the maximum variation over the flat surface is 4%. In addition, growth rate over the entire substrate is depressed, in agreement with the observations of Kondoh and Ohta for DMAH CVD.[9] As deposition progresses, features fill, and loading effects become less severe, curve 2, as patterned areas begin to resemble a flat wafer, curve 3.

Within the patterned areas, features at the edges of patterned areas are exposed to a greater concentration of precursors than features at the center of the cluster. Figure 3-11 reflects the cumulative effects of microloading.

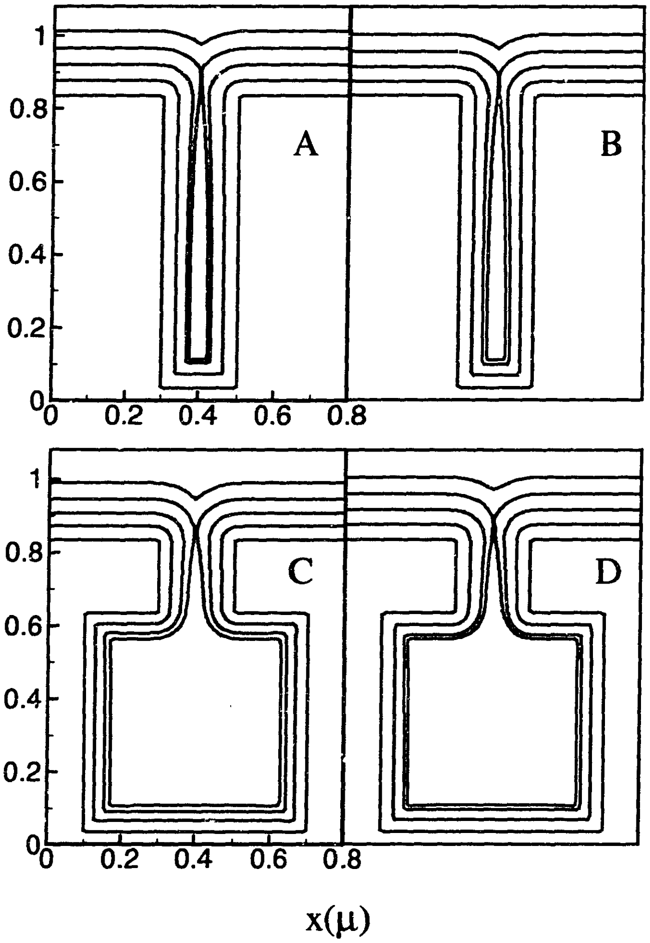


Figure 3- 13 Feature scale simulation results. Film thickness at r=0.025m is 106% of that at the field center r= 0.0325m, while film thickness at r=0.075m is 111 % of that at the field center, r=0.0675m.

Features from the center (panels 3-11B and 3-11C) and edge (panels 3-11A and 3-11D) of each field are shown. While the film morphology of features at the edge differ from those at the center, reflecting slightly more flux, differences are small; consistent with the experiments of Kondoh and Ohta.[9]

Operating pressures were seen to influence the degree of microloading observed. As pressure is decreased, diffusion coefficients increase while apparent reaction rates fall. Concentration gradients are seen to decrease, (Fig. 3-12).

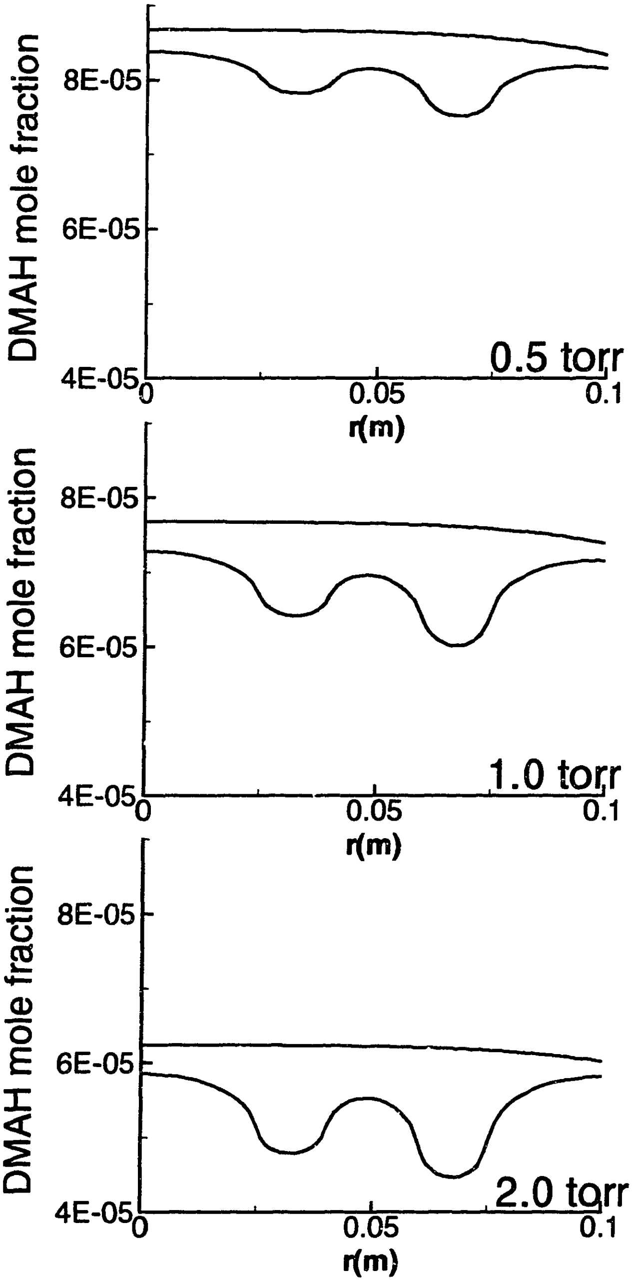


Figure 3- 14 Effect of pressure microloading. Reactor scale DMAH concentration profiles at the start of deposition are shown for patterned and unpatterned substrates at pressures of 500mTorr,1.0 Torr and 2.0 Torr.

While the maximum non-uniformity at 2 torr is 25%, the maximum at 1 torr is 17 %, and 9% at 0.5 torr. This trend is consistent with Eq. 11 and Eq.12, as well as results seen in higher pressure CVD systems exhibiting continuum-continuum interactions [9],[13].

3.9 Conclusions

A deterministic method for linking continuum and discrete particle transport models has been developed. The resulting linked model can be used to investigate interactions between reactor scale and feature scale processes in low pressure CVD. Such interactions affect film conformality and growth rate on both the micro scale and over the wafer as a whole, thus exerting a large influence on the quality of finished devices.

The linking algorithm in this work is over 1,000 times faster than the previously presented MC method. The speed advantage enables simulation of detailed chemistry. While submicron trenches and overhangs were taken as typical structures for this work, a flux map can be calculated for any feature. Calculations at the feature scale include microscopic transport resistance, which cannot be treated by continuum models because of the particulate nature of the gas phase at the micron length scale. Under conditions of free molecular flow at the feature scale, flux maps over individual features can be superimposed to yield maps of precursor and by-product fluxes over clusters and dies of features. In this fashion, the effective reactivity of an entire substrate can be mapped at moderate computational expense. The effective reactivity concept presented here is equated to a macroscopic Damkoehler number, and can be applied to other CVD processes where feature scale heterogeneity is important such as selective growth. Treatment of collisional flows by Monte Carlo is also possible within the framework presented, but collisionless calculations are valuable as a 'worst case' picture of loading effects, without the benefit of smoothing of the flux maps resulting from gas phase collisions.

1. Fishman, G.S., *Monte Carlo Concepts, Algorithms and Applications*. Springer Series in Operations Research, ed. P. Glynn. 1995, New York: Springer-Verlag. 693.
2. T.S.Cale and G.B.Raupp, *J. Vac. Sci. Technol. B*, 1990. **8**(4): p. 649-655.
3. IslamRaja, M.M., *et al.*, *J. Appl. Phys*, 1991. **70**(11): p. 7137-7140.
4. Schmitz, J.E.J. and A. Hasper, *J. Electrochem. Soc.*, 1993. **140**(7): p. 2112.
5. Kuijaars, K.J., C.R. Kleijn, and H.E.A.v.d. Akker, *Thin Solid Films*, 1995. **270**: p. 456.
6. Maynard, C.W., *Nuclear Science and Engineering*, 1961. **10**: p. 97-101.
7. Adalsteinsson, D. and J.A. Sethian, *Journal of Computational Physics*, 1995. **120**: p. 128-144.
8. Sethian, J.A., *Level Set Methods*. Cambridge Mononographs on Applied and Computational Mathematics, ed. P.G. Cialrlet, *et al.* 1996, Cambridge: Cambridge University Press. 218.
9. Kondoh, E. and T. Ohta, *J. Vac. Sci. Technol. A*, 1995. **13**(6): p. 2863-2871.
10. Littau, K.A., *et al.*, *Microelectronic Engineering*, 1997. **33**: p. 101-111.
11. Willis, B.G., *Complementary Computational Chemistry and Surface Science Experiments of Reaction Pathways in Aluminum Chemical Vapor Deposition*, in *Chemical Engineering*. 1999, Massachusetts Institute of Technology: Cambridge, Ma. p. 214.
12. Willis, B.G. *Growth Chemistry of Dimethyl Aluminum Hydride*. in *Advanced Metallization and Interconnect Systems for ULSI Applications*. 1997. Boston. Ma: Mater. Res. Soc.
13. Fujii, T. and M. Ekawa, *J. Appl. Phys.*, 1995. **78**(9): p. 5373-5386.

Chapter 4

Multiscale Models for Plasma Etching

The goal of this work was to produce a linked level set reactor model for plasma etching. This work builds on the linking methods previously developed for chemical vapor deposition processes. A model of this type can describe the interactions between the macroscopic plasma environment and observed feature profiles, and would be useful in addressing etch phenomena such as microloading, and aspect ratio dependent etching (ARDE), in addition to interaction between etch products and the plasma discharge.

A schematic of an inductively coupled plasma etch process is presented in figure 4.1. Strong electric fields generated by current - carrying coils create a zone of partially ionized gas (plasma). Plasmas are very important in materials processing, first because a wide range of 'cold' chemistries is possible in a plasma environment relative to a high temperature process. The second feature of materials processing plasmas, or glow discharges is the formation of a sheath: when a surface is first exposed to a plasma, it will receive a much higher flux of light, fast moving electrons than heavy, slow-moving ions. As a result the surface will become negatively charged, and a large local electric filed develops. This field repels electrons and negative ions while pulling in positive ions. The region in the vicinity of a surface where a large electric field is present is known as a sheath. Because the ions are accelerated by an electric field that is very large compared to any fields in the bulk plasma they arrive with a narrow distribution of incident angles and energies.

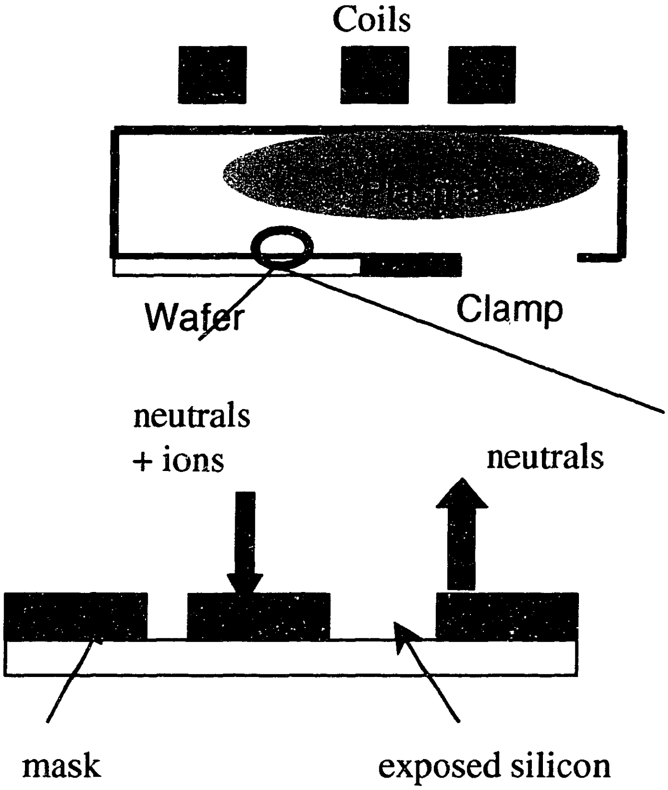


Figure 4. 1 The plasma etching process presentes a multisclae linking challenge similar to that of CVD. Both ions and neutrals flow outward from the plasma zone, but only the neutrals are fed back. Thus, neutral transport is likely to be important in microloading

4.1 Reactor Scale Model

A two-dimensional axisymmetric, inductively coupled plasma (ICP) reactor was chosen for this study. The model was developed and provided to us by Peter Ventzek of Motorola incorporated the model is based on earlier research by Ventzek et al, and is not representative of the more advanced models used by Motorola. Since the goal of this work was to highlight interactions between phenomena at the feature scale (microns) and the reactor scale (meters) a relatively simple, steady state model with small computational requirements was used to represent reactor scale phenomena. More sophisticated plasma models incorporating improvements in ion and electron dynamics have been developed.[1] Although approximate, this model can describe:

1) Feedstock dissociation
2) Sticking at surfaces
3) Multiple reaction, multi-species chemistry

The model is an extension of Paranjpe's work.[2]A schematic of the model is shown in fig 4.2.

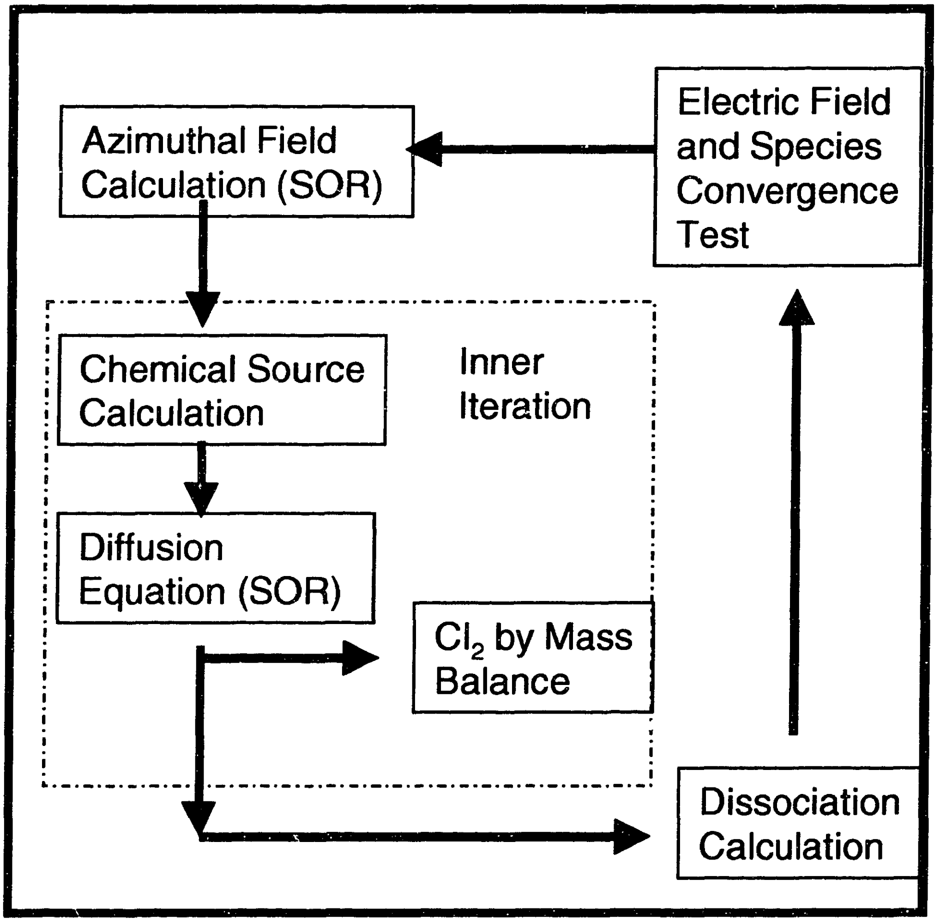


Figure 4. 2 Schematic of solution method for the plasma equipment model.

At the beginning of the simulation, the flow field in the reactor is determined by solving a potential flow model. After the flow field is calculated, the electromagnetic fields were obtained by solving Maxwell's equations within the reactor. In the style of Paranjpe, we solve Eq. 4.1 by successive over relaxation (SOR) to determine the azimuthal component of the magnetic vector potential. The induced electric field in the plasma given by Eq. 4.2

$$\nabla^2\mathbf{A} - \mu\varepsilon\frac{\partial^2\mathbf{A}}{\partial t^2} = \mu\sigma\mathbf{E} \tag{4.1}$$

$$\mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t} \tag{4.2}$$

The ionization rate is then calculated as a function of the local reduced electric field, E/N that then gives species production rates. Next, we solve an ambipolar diffusion equation for each species of charged particle Eq. 4.3, and convective-diffusion Eq. 4.4 equations for each neutral species.

$$\frac{\partial n_i}{\partial t} = D_i\nabla^2 n_i + S_i \tag{4.3}$$

$$\frac{\partial}{\partial t}(n_i) + \nabla \bullet (\hat{u}n_i) = -\nabla \bullet j_i + S_i \tag{4.4}$$

The source functions are updated in keeping with the new concentration profiles, and iterations continued until concentrations are converged, this is the inner iteration in Fig. 4.2. At this point, the azimuthal electric field is recalculated, the rate coefficients re-determined, and another pass is made trough the inner loop. The outer loop is stopped when all of the fields and species densities are converged at every spatial location.

Sticking and generation at surfaces is treated with a mass balance on the finite difference mesh adjacent to the substrate, Eq. 4.5, which is invoked in the inner iteration at each pass through the SOR scheme.

$$D_i \frac{(n_i^{wall+1} - n_i^{wall})}{\Delta X}\frac{A^{wall+1}}{V} - n_i^{wall}\bar{c}\frac{A^{wall}}{4V}P_{rxn} = S_i \tag{4.5}$$

A final consistency check, Eq. 4.6 is required due to inclusion of flux boundary conditions in the SOR scheme.

$$\underset{V}{\oiiint} S_i = \underset{S}{\oint} j_i \bullet \mathbf{dS} \tag{4.6}$$

Eq 4.6 is evaluated to verify overall mass balance on the reactor model. If there is an imbalance, all wall densities are multiplied by the ratio of the generation rate to the loss rate, and another pass through the SOR is made. The plasma chemistry model was made to converge for an inductively coupled chlorine discharge, over reactive surfaces representing both a bare and masked silicon wafer. Operating pressure was 10 mtorr, power was 1900 watts and inlet chlorine flowrate was set to 5 sccm. The predicted molecular chlorine density under these conditions is shown in Fig 4.3.

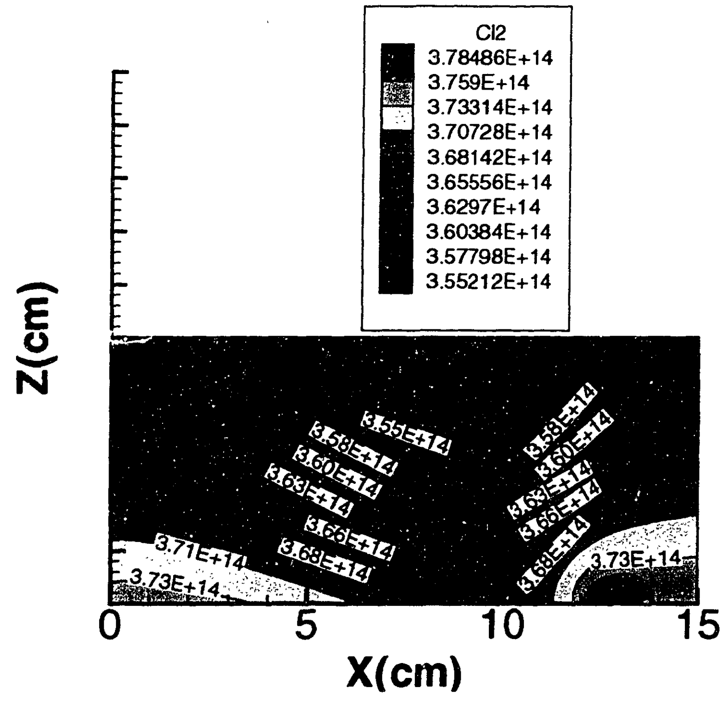


Figure 4. 3 Molecular chlorine concentration contours calculated by ICP model. Highest dissociation is seen in the toroidal region consistent with inductive power dissipation. Highest concentrations are seen near the substrate and walls due to recombination of Cl radiacals

Although the sticking coefficient approach and finite difference scheme presented are not well suited to treating generation of products, (it would involve the introduction of negative sticking coefficients) $SiCl_4$ production rates can be obtained by mass balance at the start of etching. Predicted $SiCl_4$ density with boundary conditions for an un-patterned wafer at the substrate surface are shown in figure 4.4. as expected, the silane production rate and concentration are largest at the edge of the substrate, which receives a slightly larger ion flux than the center.

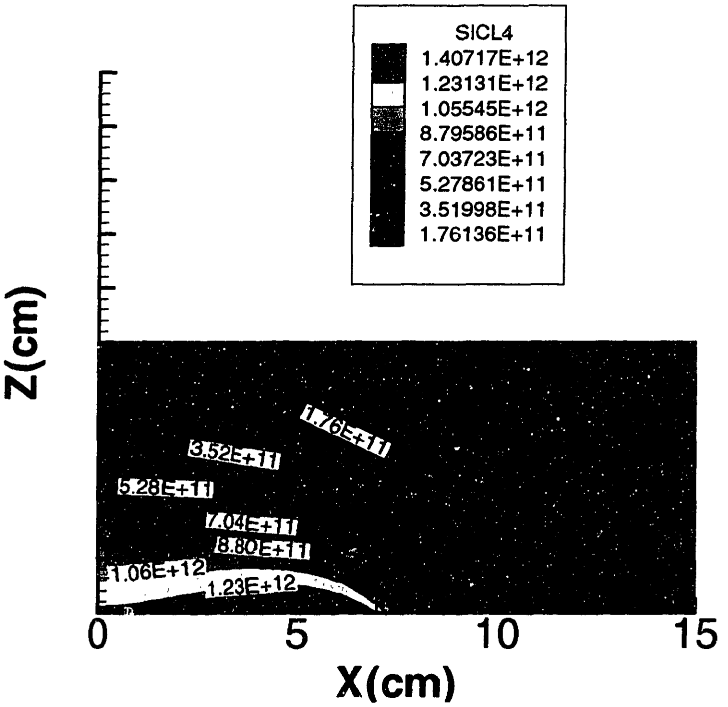


Figure 4. 4 SiCl4 concentration profiles over bare (un-patterned) silicon wafer, showing expected maximum in product concentration at the wafer surface.

The un-patterned wafer boundary condition simulated above is not the most accurate depiction of an actual etching process. Just like CVD, masked, and thus non-reactive, areas will exist adjacent to areas that react with and thus deplete etchants, as well as producing etch byproducts like $SiCl_4$. In this case, the etchants (chlorine atoms) are present in large excess, and so are not depleted appreciably by etching. The concentration profile of $SiCl_4$, however is markedly different as shown in figure 4.5, with a distinct maximum over the unmasked area from X = 2 to 4 cm.

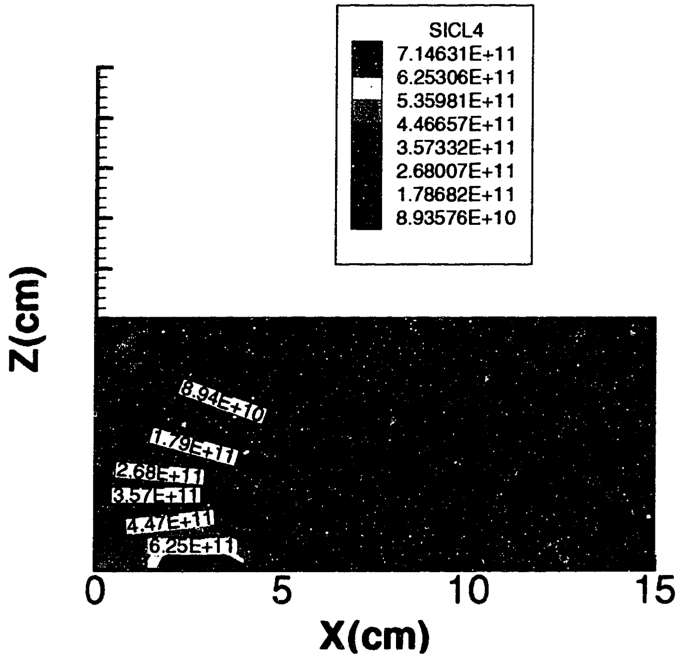


Figure 4. 5 SiCl4 concentration profiles over masked wafer, with active region between X=2cm and X=4cm SiCl4 flux is highly non-uniform over the wafer surface.

4.2 Feature Scale Etch Model

At the microscale, both thermalized neutrals and energetic ions play important roles in etching, and must be accounted for. As the subject of this work is polysilicon etching, surface charging effects are small, and the resulting distortion of ion trajectories is ignored. Thus, line of sight transport of both ions and neutrals is assumed. Etch profile models have been created using ballistic integral [3] [4],[5] and MC approaches.[6, 7] Given that we will not consider surface charging, a ballistic integral approach is taken here.

We begin with the ions. Due to the narrow range of energies possible in our system, we assume incident ions to be mono-energetic. We take the ions as arriving with the angular distribution Eq. 4.7 presented by Gottscho[8], which has been used successfully in profile simulation.

$$f(\xi_s) \approx \exp(\frac{-\xi_s^{\,2}}{2\sigma^2}) \tag{4.7}$$

$$\sigma \approx \sqrt{\frac{kT_i}{2eV_s}} \tag{4.8}$$

Where $f(\xi_s)d\xi_s$ is the probability that an ion entering the feature scale domain will have a trajectory at an angle within $d\xi_s$ of the ray leaving the source plane at angle ξ_s to the normal. If the total ion flux intensity is denoted I, the direct ion flux having trajectories within $d\xi_s$ of ξ_s is $If(\xi_s)d\xi_s$ and the arriving at point x in the feature directly from the plasma is given by Eq. 4.9

$$\Phi_{ion}(x,\xi) = \cos(\xi)If(g(\xi)) \tag{4.9}$$

Here, ξ represents the angle of the incident ion with respect to the local surface normal, and $g(\xi)$ is the angle that the ions' trajectory makes with the source plane normal. Upon impacting the surface, ions may then participate in etching reactions or be reflected. For the profiles presented in this chapter, reflection of energetic particles was not considered, in general, ballistic integral models neglect re-emission of energetic particles,[3, 5] as resolving a full angular and energy spectrum in a 2-d feature is a demanding task, leading to large computation times. However, a method for incorporation of specular reflections is presented in the next chapter. Once the incident ion flux density is known, neutral flux is determined by a line of sight calculation. Neutral transport is treated with the same line of sight model presented in Ch. 3.

The surface is discretized into elements, x_i, with material reaching element x_i directly from the deposition chamber, or after being re-emitted by another surface element, x_j of the feature. A mass balance on element xi gives:

$$\phi^{in}(x_i)A_i = \phi^{direct}(x_i)A_i + \sum_{x_j \neq x_i} \phi^{out}(x_j)A_jQ(x_j \rightarrow x_i) \tag{4.10}$$

Where $\phi^{in}(x_i)$ and $\phi^{out}(x_i)$, represent the incident and emitted neutral flux densities at element x_I, and A_i is the area of element x_i. We employ a Langmuir-Hinschelwood description of surface kinetics described in the next section, and solve the feature scale model with the iterative approach described previously.

4.3 Surface Chemistry Model

The underlying physical phenomena in chlorine plasma etching are quite complex,[9, 10] but a few assumptions can simply the problem considerably. Perhaps the most important assumptions involve which species to consider. Reactive plasmas generate a large variety of ions, neutrals and radicals, if etch by-products are included, the chemistry becomes even more complex. Because we wish to focus on the interactions between feature and reactor scale, a surface reaction set including some treatment of etch byproducts was required. A detailed model of feature scale surface chemistry is taken from the work of Chang, Mahorowala, and Sawin.[7]Our goal was to combine this mechanism with the plasma chemistry scheme of Lee. Graves, and Lieberman.[11] Some difficulty was encountered in that not all the species in plasma chemistry kinetic scheme were considered in the formulation of the surface models. Due to scarcity of data, we make some assumptions in reconciling the surface and plasma kinetic models. The first is that lattice Si is removed from the surface only as $SiCl_4$. The unsaturated silane, $SiCl_3$ is the dominant product of interaction of $SiCl_4$ with the plasma, assumed to re-deposit with the sticking coefficient measured by Chang for $SiCl_2$. $SiCl_4$ is considered inert with respect to the surface. Finally, Cl_2^+ was assumed to interact with the surface in the same fashion as Cl^+. While these assumptions are fairly severe, our intent is to illustrate the coupling of feature scale and reactor scale etching models, rather than develop an all-encompassing plasma-surface chemistry model. The resulting surface reaction set is shown below.

Table 4. 1 Feature Scale Etch Model. Reaction rates are functions of incident angle and energy, although mono-energetic ions are assumed

No	Reaction	Rate Law
1	$Cl + {*}(Si) = Cl(Si)$	$S[\phi_{Cl}][{*}Si]$
2	$Cl^+ + {*}(Si) = Cl(Si)$	$\int_{-\pi/2}^{\pi/2} C(\theta)\Phi_{Cl^+}(\theta)d\theta\,[{*}Si]$
3	$Cl_2^+ + 2{*}(Si) = 2Cl(Si)$	$\int_{-\pi/2}^{\pi/2} C(\theta)\Phi_{Cl_2^+}(\theta)d\theta\,[{*}Si]$
4	$SiCl_3 + {*}(Si) = 3Cl(Si) + Si(s)$	$\delta\,[\phi_{SiCl3}][{*}Si]$

5	$Si(s) + 4Cl(Si) = SiCl_4 + 4^*(Si)$	$[Cl(Si)]\,\beta \int_{-\pi/2}^{\pi/2} C(\theta)\{\Phi_{Cl_2^+}(\theta) + \Phi_{Cl^+}(\theta)\}d\theta$

The parameters $S, c(6)$, and β were obtained from the experiments of Chang, Mahorowala and Sawin.[7] S represents the sticking coefficient of Cl on Si under ion bombardment. $c(6)$ is the chance an incident ion will be interact with the surface instead of being reflected. It is a strong function of the incident angle, and the measurements made for Cl^+ impacting silicon are used here. No data were available for Cl_2^+, and the same angular dependence is assumed. Redeposition of etching by-product $SiCl_3$ is described with the sticking coefficient measured for $SiCl_2$ on a silicon surface under ion bombardment. Finally, β represents the 'etching yield' of an incident ion, and is a strong function of energy. Given the parameters above, a steady state approximation is made to determine the extent of surface chlorination, and the etch rate and 'outbound' species fluxes. The fluxes of neutral species are then solved for iteratively to determine the final etch rate. After the etch rate calculation, we invoke the level set profile simulator, etching for a small time interval and then re-computing the species fluxes. In this manner, the etch problem differed significantly than the CVD problems presented earlier.

Because of the strong angular dependence of the reactions in table 4-1, the etch rate was very sensitive to the orientation of the front, and the etch rate had to be re-computed for even very small movements of the front. Specifically, for the level set etching model a speed calculation was required every other time step, while for CVD a speed calculation every 25th time step would produce acceptable results.

4.4 Feature Scale Etch Model Results

Test cases were run with the etch profile simulator to examine the dependence of the final shape on ion to neutral flux ratio. Flux maps of etching by-products $SiCl_4$ and $SiCl_3$ are also presented. Level set etching models have also been presented by Sethian,[12] and Hwang,[5] but with no treatment of chemistry (Sethian) or simplified treatment (Hwang). Several investigators have noted the effect of ion to neutral flux ratio on etching rates[13, 14], as well as the shape of etch profiles.[4, 5]

Figures 4.6 and 4.7 depict etch profiles for ion to neutral flux ratios of 1 and 0.01 respectively to facilitate comparison with the results described by Hwang[5] and Tuda.[4] Test features were etched to a depth twice their width, and so appear as the same depth, in spite of very different real etch rates. In each case, ion flux was set to 10^{16} ions/cm^2 per second for both Cl^+ and Cl_2^+. Incident ion energy was chosen to be 50 eV.

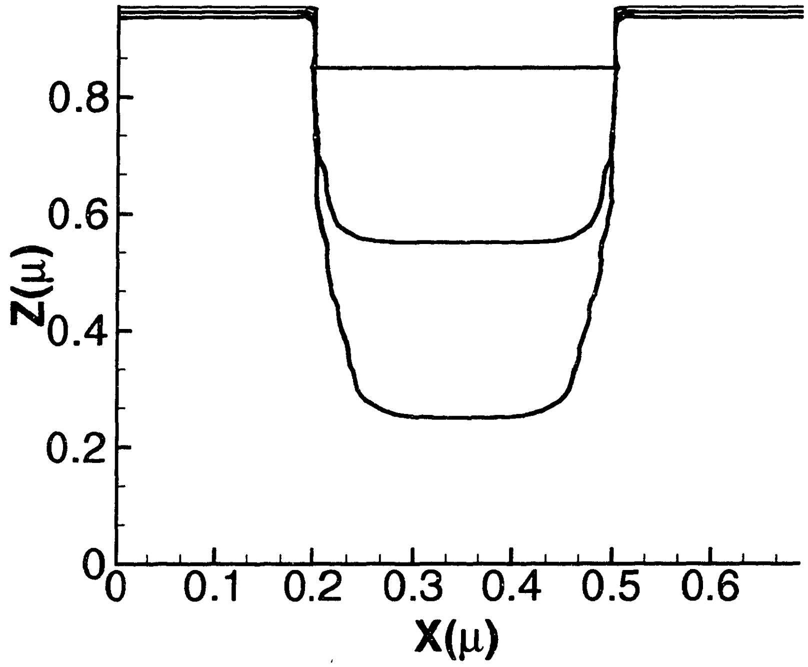


Figure 4. 6 Etch profile for ion to neutral flux ratio 1. 0. Lines are shown denoting the starting point, mid-point (aspect ratio=1) and end point (aspect ratio =2) of the etch. The profile bottom maintains a nearly constant shape, as anisotropic ions are responsible for both chlorination of the surface and removal of volatile etch products.

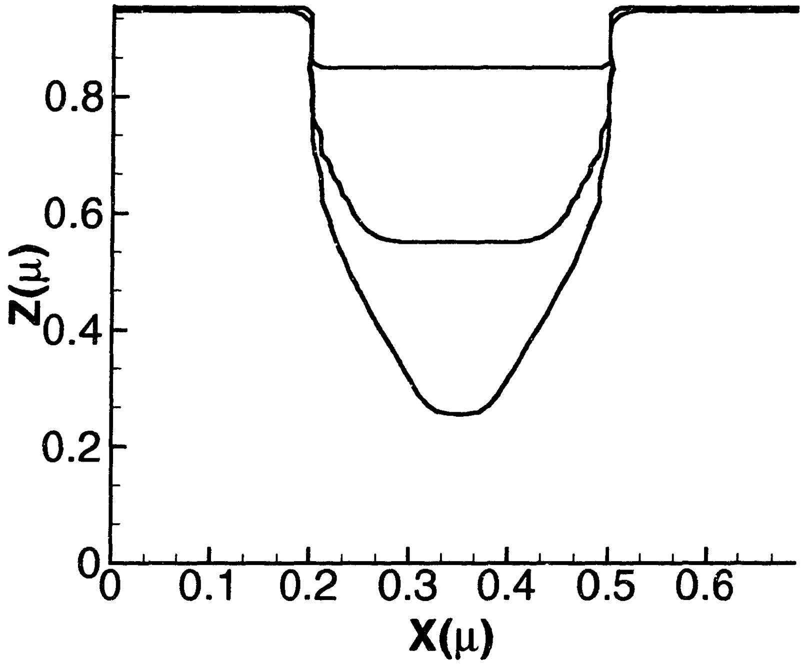


Figure 4. 7 Etch profile for ion to neutral flux ratio 0.01 Lines are shown denoting the starting point, mid-point (aspect ratio=1) and end point (aspect ratio =2) of the etch. As the aspect ratio increases, the bottom becomes sharply curved, due to the contribution of chlorine atom flux.

The changes in profile shape agree well with those described previously, and are explained by considering the synergy between the ion and neutral fluxes. Recalling that although the ions arrive in a highly directional fashion, neutral particles do not, arriving with the same isotropic flux distribution encountered in CVD processes. Thus , the ion flux will be nearly constant on the top surfaces , and along the base of the feature, with very little flux lost to the sidewalls. This in contrast to the neutral flux, which will be increasingly lost to the sidewalls as the feature deepens. The neutral flux is smallest at the corners where the sidewall meets the feature bottom. As a result, even in the presence of ion flux, the etch rate is retarded relative to that at the feature center due to differing degrees of surface chlorination. Thus, if cases where neutrals play an important role the character of the etch can change as the aspect ratios of etched features increase. Flux maps for use in a multiscale plasma model are presented in figures 4.8 and 4.9.

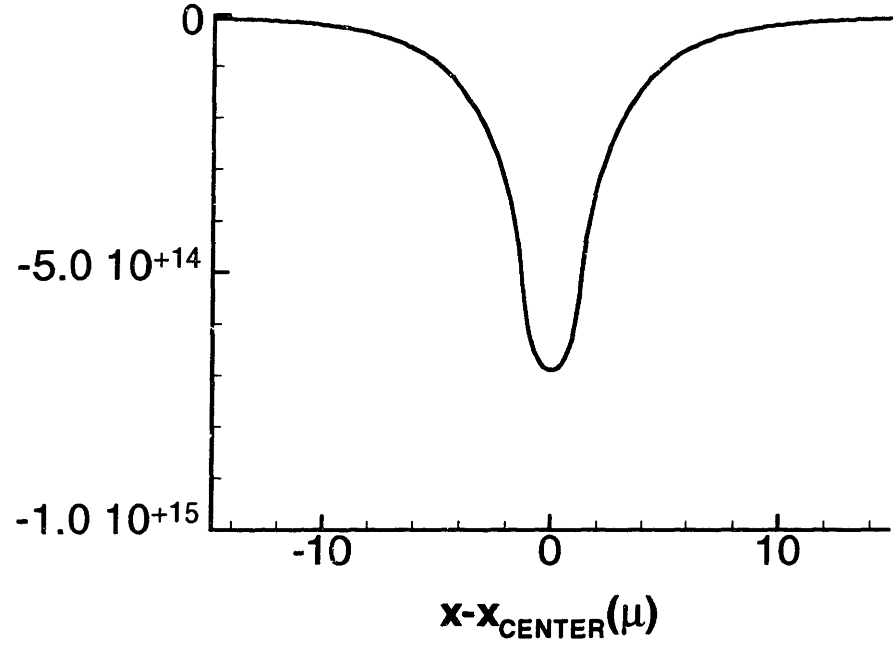


Figure 4. 8 Flux map showing incoming $SiCl_4$ flux over a 0.2m feature. Masked areas are assumed non-reactive. Negative values correspond to flux leaving the feature scale domain.

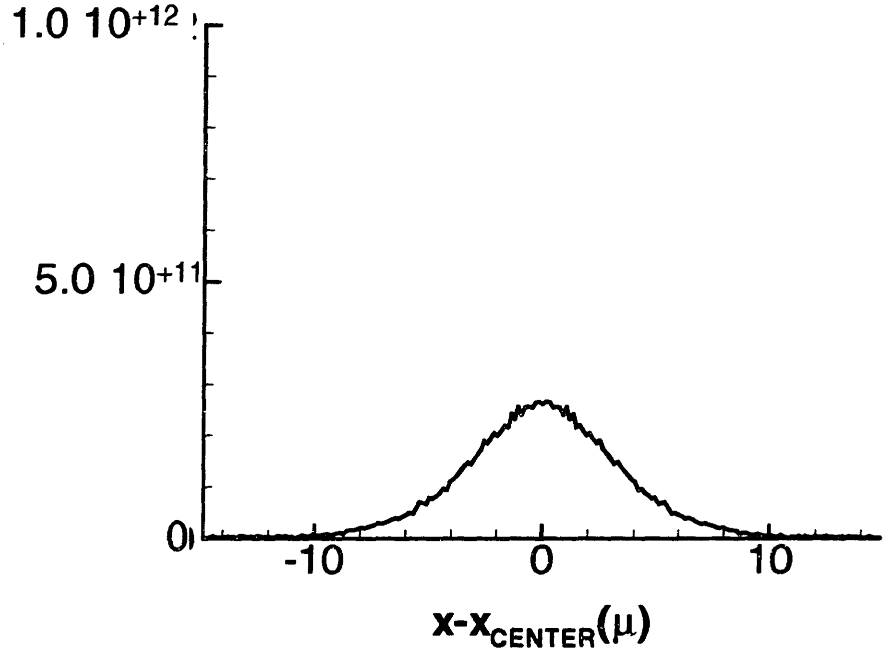


Figure 4. 9 Flux map showing incoming $SiCl_3$ flux over a 0.2m feature. Masked areas are assumed non-reactive. SiCl3 arrives from the reactor chamber and deposits on exposed silicon surfaces.

4.5 Prospects for Linking

The plasma equipment model feels the effects of the feature scale through the boundary condition at the wafer surface, represented by equation 4.6. In chapters 2 and 3, the effects of the feature scale were calculated by computing the flux from the substrate surface and formulating a flux boundary condition on the macroscopic model. When the feature scale linking model is invoked, the boundary condition in Eq. 4.6 is modified to yield Eq. 4.11.

$$D_i \frac{(n_i^{wall+1} - n_i^{Wall})}{\Delta X} \frac{A^{wall+1}}{V} - j_{FeatureModel} \frac{A^{wall}}{4V} = S_i \qquad (4.11)$$

However, problems were encountered when attempting to enforce this equation and

Eq. 4.6 simultaneously. The model equations 4.1 through 4.5 serve to determine the 'shape' of the flux distribution arriving at the walls. This flux profile is then scaled until it is consistent with the volumetric production rate in Eq 4.6. As long as this solution procedure is employed the boundary condition in Eq 4.11 cannot be used because the feature scale result cannot be rescaled, if it were, the macroscopic solution would no longer be consistent with the feature scale results.

The inability to rescale the FSM results to force consistency with the overall mass balance is quite problematic, for if $j_{Feature\ Model}$ is significantly different than the solution over a flat wafer, then an SOR technique will not converge the plasma model. Under-relaxation was also attempted. We suspect that the 'correction' between surface and reactor models was too large to converge without a more sophisticated approach to finite differencing, and a rigorous treatment of species generation at surfaces. As mentioned previously, further problem with the use of Eq. 4.11 as a boundary condition is that it introduces negative sticking coefficients for species with positive generation rates. In general, such a model will not converge easily. While a positive sticking coefficient implies " the more you have the more you consume " a negative sticking coefficient implies "the more you have the more you make", clearly this is not likely to be a well-behaved model.

If a transient model, or a finite element plasma model were available, a continuation approach could be attempted where the correction due to pattern effects could be added into the solution a little at a time, thus damping errors. In summary a transient finite difference or finite element plasma model, with rigorous treatment of product generation at reactive surfaces is required to complete a full, linked, simulations.

4.6 Etching conclusions

The original intent of this etching work was to demonstrate a linked feature scale and reactor scale etching model. The task proved more difficult than originally anticipated due to the strong interaction between gas phase and surface processes in a plasma environment. Surface products not only travel about the reactor, making repeated visits to the substrate, they also interact with plasma chemistry and may even influence the nature of the plasma discharge. Each time the plasma and feature models are matched, the solution procedure begins by converging the equipment model with a surface boundary condition that is not consistent with conditions in the discharge. If convergence of this first step fails, a new guess cannot be generated. With a finite element model a simple solution to this problem could be tried, the feature scale fluxes can be added in a little at a time, in a continuation-type approach. For finite difference techniques a remedy is probably possible but not readily apparent. Thus, a finite element, as opposed to finite difference approach would have been useful in the reactor model due to a more natural incorporation of flux boundary conditions, and the possibility of a continuation strategy. A level set etch profile simulator was created, and combined a ballistic integral approach to neutral transport with a simple treatment of ion transport. A complete model for ion transport, resolving the full angle and energy spectrum is presented in the next section.

1. Ventzek, P.L.G., R.J. Hoekstra, and M.J. Kushner, *J. Vac. Sci. Technol. B*, 1994. **12**(1): p. 461.

2. Paranjpe, A.P., *J. Vac. Sci. Technol. A*, 1994. **12**(4): p. 1221.

3. Levinson, J.A. and E.S.G. Shaqfen, *J. Vac. Sci. Technol. A*, 1997. **15**(4): p. 1902.

4. Tuda, M., K. Nishikawa, and K. Ono, *J. Appl. Phys*, 1997. **81**: p. 960.

5. Hwang, H.H., T.R. Govindan, and M. Meyappan, *J. Electrochem. Soc.*, 1999. **146**(5): p. 1889.

6. Hoekstra, R.J., M.J. Graperhaus, and M.J. Kushner, *J. Vac. Sci. Technol. A*, 1997. **15**: p. 1913.

7. Chang, J.P., A.P. Mahorowala, and H.H. Sawin, *J. Vac. Sci. Technol. A*, 1998. **16**(1): p. 217.

8. Gottscho, R.A., *J. Vac. Sci. Technol. B*, 1993. **11**(5): p. 1884.

9. Rossen, R.A. and H.H. Sawin, *J. Vac. Sci. Technol. A*, 1997. **5**(4): p. 1595.

10. Barone, M.E. and D.B. Graves, *J. Appl. Phys.*, 1995. **78**(11): p. 6605.

11. Lee, C., D.B. Graves, and M.L. Liebermann, *Plasma Chemistry and Plasma Processing*, 1996. **16**(1): p. 99.

12. Sethian, J.A., *Level Set Methods*. Cambridge Mononographs on Applied and Computational Mathematics, ed. P.G. Cialrlet, *et al.* 1996, Cambridge: Cambridge University Press. 218.

13. Dane, D. and T.D. Mantei, *Appl. Phys. Lett.*, 1994. **65**(4): p. 478.

14. Coburn, J.W., *J. Vac. Sci.Technol B*, 1994. **12**(3): p. 1385.

Chapter 5

Physical Vapor Deposition: Linking Feature Scale and Atomic Scale

In this work, a general method for modelling ionized physical vapor deposition (IPVD) is developed. The method links the atomic scale and feature scale. The model presented here includes atomistic detail with a molecular dynamics based growth rate model. This information is combined with a line of sight transport model within the feature. The resulting deposition model can account for arbitrary angular and energetic dependence of surface reaction rates. All effects of reemission and re-deposition are included. Finally, a level set scheme is combined with the reaction and transport model and used to follow the shape of the film in time. The resulting scheme is highly accurate and efficient, hundreds of times faster than Monte Carlo. We choose Al deposition with and without simultaneous argon bombardment as an example problem, as well as a conventional PVD example the method performs well in each case.

5.1 Ionized Physical Vapor Deposition (IPVD)

Ionized physical vapor deposition (IPVD) has received much attention as a method for depositing material at the bottom and on the side-walls of high aspect ratios proposed for sub 0.25 micron integration[1] IPVD is based on in-flight ionization of atoms sputtered from a target. The metal atoms knocked out of the target by argon ions experience ionization as they pass through a high-density plasma before reaching the substrate. The electric field at the biased substrate tends to collimate this metal ion flux. From a practical point of view, IPVD has two intrinsic advantages over conventional

---

*This work done at MIT in collaboration with visiting scholar Uwe Hansen sputtering. First, if the substrate is biased negatively, then positive ions will be accelerated toward the substrate causing incident metal ions arriving at near normal incidence. Second the arriving energy of the metal depositing species is controlled by adjusting the bias voltage. IPVD improves the morphology of deposited films by reducing the buildup of overhanging metal deposit at the mouth of the structure, and be re-sputtering material from the feature bottom to the sidewalls. Due to the growing technological demands for the sputter process an understanding of the underlying key processes at the atomic level is required.[2] Atomistic simulations are playing an increasingly prominent role in materials science, offering a microscopic physical view that cannot be readily obtained from experiment alone. Hence there has been increased effort [3-5] to model highly non-thermal deposition techniques using the molecular dynamics (MD) approach.[6, 7]. The strengths of this technique are that it allows study of atomic trajectories and thus an atomistic view of the deposition.

5.2 Linking atomistic and feature scale

Using molecular dynamics data in growth modeling is an approach clearly superior to simplified continuum models which usually do not include atomic level information at all and employ energy and angular independent rate constants. In principle the reaction rates obtained from the MD or *ab-initio* calculations can be employed in an atomistic Monte Carlo (MC) model where the motion of each individual atom is traced during the course of the simulation.[8, 9]. Although completely correct, this approach is computationally quite intensive, and thus inconvenient if the final goal is to predict metal film topographies at length scales of microns as opposed to angstoms. In many cases of practical interest, growth proceeds under simultaneous rare-gas bombardment, and the number of deposited particles and the number of particles etched away (re sputtered) from the surface are of similar size. In such a case the overall growth rate is very low, but within atomistic MC, the trajectories of all atoms still have to be traced, making this approach too expensive to be useful for systems with very low growth rates. In this work a new and computationally highly efficient scheme to model the growth of thin films is presented. This approach enables us to employ data from atomistic simulations (i.e. angular and energy dependent surface reactions rates) within the level-set method. Furthermore all effects of reemissions are included so that we arrive at a model that provides a great speed advantage compared to the Monte Carlo methods while retaining the accuracy of an atomistic approach, and easily resolving problems with low growth rates.

5.3 Reaction Rates from Molecular Dynamics

Uwe Hansen performed classical molecular dynamics simulation[6, 7] using an embedded atom type[10, 11] interaction potential for the Al-Al interaction and a repulsive Born-Mayer[12] type term for the Ar-Al interaction. The details of the molecular dynamics calculations are described in previous work.[9, 13, 14] Initially, the incident Al atom is placed outside the interaction range of the surface. Its initial kinetic energy is set in the range of 0 to 150 eV and its starting angle off the surface normal in the range 0^{o} to 85^{o}, which corresponds to typical ionized physical vapor deposition conditions. The trajectories of the incident atom, and of any other atom which may be etched away from the surface upon impact are monitored. Analyzing a large number of trajectories per incident energy and angle, we collected a statistically significant sample of well-defined adsorption, reflection and etching events. The relative probability of the corresponding process is calculated as the ratio of the number of events of each kind to the total number. For the reflected particles and the particles sputtered away from the surface we also recorded the angular distributions after the impact. This distributions comprise the function *G*, which is used to summarize surface events. In the following we briefly discuss the energy and angular dependence of the three key surface reactions, namely adsorption, reflection and sputtering.

5.3.1 Adsorption of Al on Al(111)

In the low energy regime (typical for molecular beam epitaxy) all Al atoms are adsorbed independent of their impact angle, since their initial kinetic energy does not suffice to leave the strongly attractive surface adsorption well. To be more precise, an impinging particle needs after the interaction with the surface atoms a momentum parallel to the surface normal that enables it to surmount the surface adsorption well (for an Al adatom the well depths are 3.10 eV on Al(111), 3.77 eV on Al(100) and 3.89 eV on Al(110)).[15] This picture changes if the incident kinetic energy is increased and a strongly energy and angular dependent sticking probability is found. The dependence of the sticking coefficient, $S(E,\alpha)$, on the energy, E and off-normal angle alpha for energies between 0 and 150 eV and incident angles in the range between 0 and 90° according to our MD calculations is depicted in Fig 5.1.

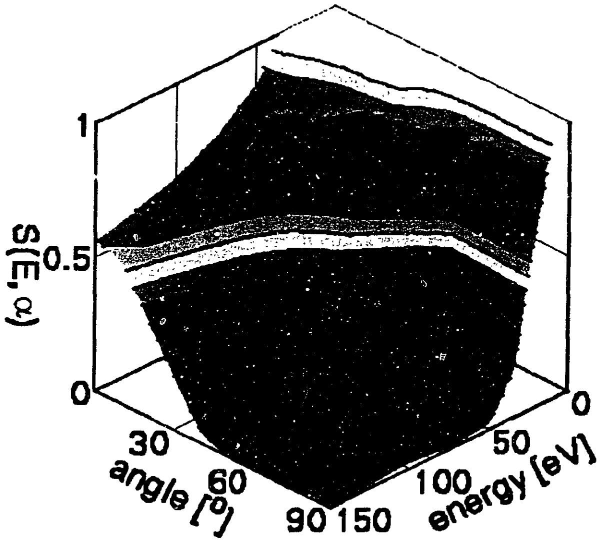


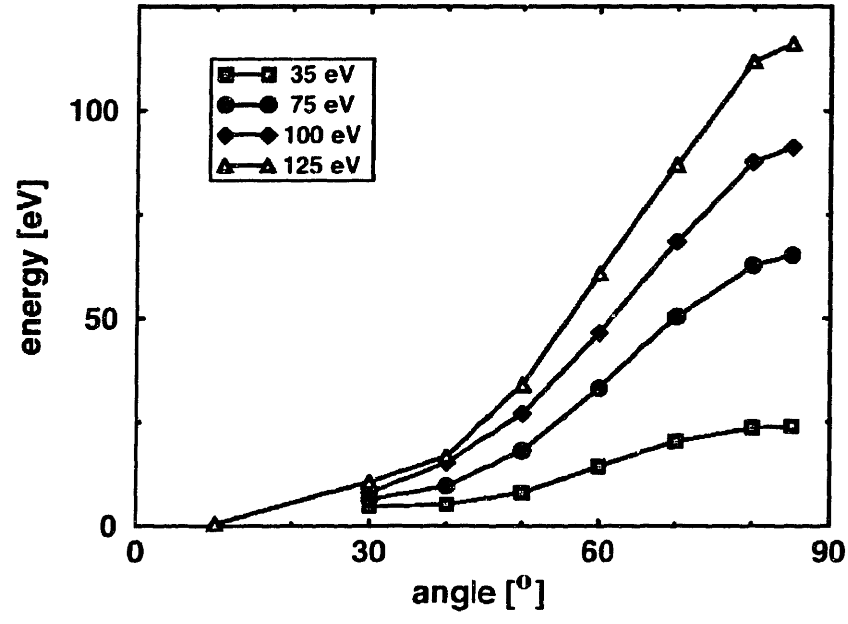
Figure 5. 1 Calculated sticking probabilities for hyper-thermal Al atoms impinging on an Al(111) surface as a function of kinetic energy and incident off-normal angle, a. The surface temperature was set to 450 K. Contour lines mark increments of 0.1 in sticking probability.

For high energies and large off-normal angles particles are reflected so that the adsorption probability drops to zero. In the intermediate angular range sputtering events [see Sec. 5.5.3] are the competing process that reduce the overall adsorption probability. From this contour plot is should be evident that using an angular and energy independent sticking probability is a very crude approximation for non-thermal deposition conditions.

5.3.2 Reflection of Al at Al(111)

We now turn our attention to the reflection events. With increasing angle we find a transition from diffuse to specular reflection events [for a detailed analysis see also [9]], as at large off-normal angles, the particle has a large momentum component parallel to the surface and a small momentum component perpendicular to the surface. During the collision mainly the perpendicular component is changed and the parallel component is conserved, such that the particle can escape the attractive adsorption well. This picture is reversed for the case of near normal incidence. Very few reflections occur since the particle makes a collision where its transfers most of its momentum to the surface. This is further confirmed by analyzing the particles kinetic energy after undergoing a reflection event. For a statistically significant sample we averaged the kinetic energy of particles being reflected from the surface. For incident energies in the range between 35 and 125 eV and off-normal angles between 10° and 85°, Fig.5.2 shows the mean kinetic energy for particles being reflected from the surface as a function of angle.

Figure 5. 2 Calculated mean kinetic energy for hyper-thermal Al atoms being reflected from the



Al(111) surface as a function of incident off-normal angle and energy. The open squares, circles, diamonds and triangles represent incident energies of 35, 75,100, and 125 eV respectively. The surface temperature was 450K.

The open squares, circles, diamonds and triangles correspond the energies of 35, 75, 100 and 125 eV respectively. For near normal incidence the momentum transfer towards the surface is most efficient, while for near grazing incidence the average energy loss is about 10 eV independent of the initial kinetic energy. For the use in our thin film model we did not only average the kinetic energies after the reflection event, we also calculated the variance in kinetic energy. This width in the kinetic energies later on enters the distribution functions for the reemitted fluxes.

5.3.3 Sputtering of Al from Al(111)

Under non-thermal deposition conditions, the key physical processes are adsorption, reflection and sputtering, i.e. the removal of surface atoms and adsorbates by the impact of energetic particles. The probability for particles being etched away due to the impact of an energetic particle according to our MD calculations is shown in Fig.5.3

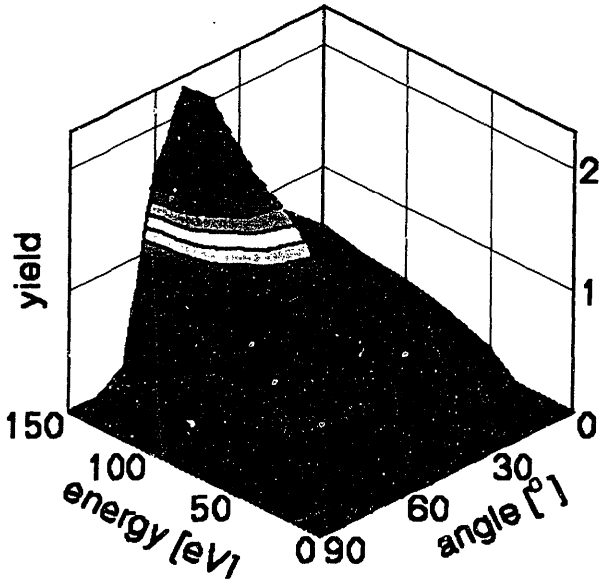


Figure 5. 3 Calculated sputter yield for hyper-thrmal Al atoms impinging on an Al(111) surface as a function of kinetic energy and incident off-normal angle, a. The surface temperature was 450K, with contour lines marking increments of 0.2 in the yield

Note that in contrast to the adsorption and reflection events, the etch probability does not show a monotone dependence on the parameters of incidence, *i.e.* energy and angle. More interestingly it reaches a maximum for angles around 50° and decreases as near grazing angles are approached. For energies smaller than 25 eV no sputter events are observed. But increasing the energy we find considerable changes in the sputter probability. Even at small angle the etching probability is non-vanishing. For small deviations form the normal incidence, the etch rate initially rises, since the probability of a surface atom to gain momentum directed away from the surface increases when the incident atom arrives at an oblique angle at the surface. At large off-normal angles the latter probability drops because of the competing specular reflection events. Analyzing the energies of the particles sputtered away from the surface we find, that they have mainly energies less than 15 eV such that they are most likely adsorbed in the course of their next interaction with the growing film front. Furthermore is should be pointed out that there exists a preferred direction for the sputtered particles, which depends on the underlying crystal structure and surface orientation. During the sputter process not only single particles are etched away from the surface so that the yield *i.e.*, the number of particles leaving the surface divided by the number of particles impinging on the surface is obtained by multiplying the sputter probability by the actual multiplicity for the event (multiplicity=n mean that upon impact of a single particle n+1 particles are sputtered from the surface).

5.3.4 Reaction rates for Ar on Al(111)

From experiments it is known that Ar is not incorporated into the film during growth. So for Ar, we only have to consider the reflection and etching probabilities since no adsorption takes place. We find that for energies below 25 eV hardly any argon atom can etch away Al atoms from the surface. As in the case of Al the yield curves for Al display a distinct maximum for off-normal angles around 45°. Due to the similar mass of

Al and Ar, the reflection probabilities and the energy loss during reflection events is similar for both of the atoms and will therefore not be discussed in further detail here.

5.3.5 Application of MD Results to Micron Scale Thin Film Growth

In summary, there exists complex balance between the different competing surface reactions. Only with an accurate description of their probabilities does reliable and predictive thin film growth modeling become feasible. Before the MD results can be used in a simulation of film morphology, they must be summarized in a fashion that is readily compatible with a species transport model. Given an incident vector flux, $\Phi^{in}(E^{in},\theta^{in})$, at a particular energy E^{in} and angle, θ^{in}, a function, $G\ (E^{in},\theta^{in} \rightarrow E^{out},\theta^{out})$ is used to relate the incident and emitted (outbound) fluxes in a manner consistent with the pre-tabulated MD results. Thus, $G(E,\theta \rightarrow E^{'},\theta^{'})$ summarizes all the surface events and gives the re-emitted flux, $\Phi^{out}(E^{'},\theta^{'})$ after operating on $\Phi^{in}(E\ ,\theta\ )$. Operation of G for the data presented in the precious section is depicted in Fig.5.4

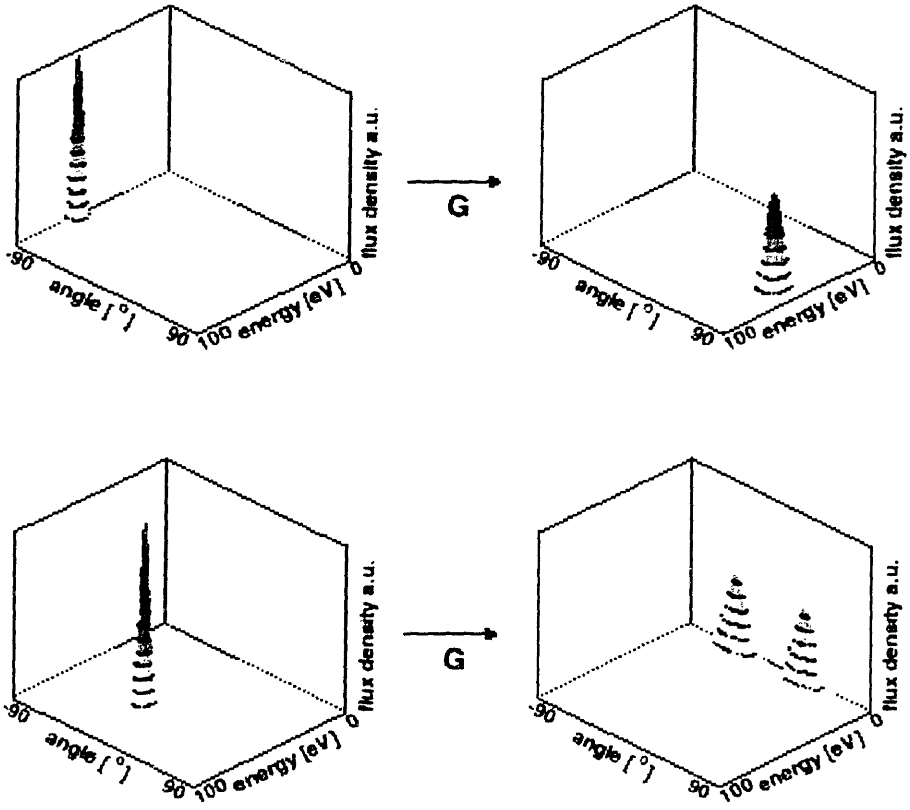


Figure 5.4 Illustration of the atomistic summary function, G that maps the incoming vector flux density to the re-emitted density. In the top left, a narrow distribution with energy of 60 ev arrives with incident angle of +85°. The emitted flux is shown at right The lower panels depict an incident distribution of 60 ev and +10° the re emitted flux is comprised mostly of sputter products.

Other examples of G : for a perfectly adsorbing process, $G(E,\theta \rightarrow E',\theta')=0$, while for a perfect specular reflection with no energy loss, $G(E,\theta \rightarrow E',\theta')=\delta(-\theta')$. Once a G is calculated for each energetic species in the system, the re-emitted flux can be determined for any given incident flux, by equation 5.1 below. Furthermore, note that a mass balance readily gives v, the deposition (or etch) rate.

$$\Phi_i^{out}\left(\mathbf{r},E',\theta'\right)=\int_0^{\infty}\int_{-\pi/2}^{\pi/2} G_i\left(E,\theta \rightarrow E',\theta'\right)\Phi_i^{in}\left(\mathbf{r},E,\theta\right)dE\,d\theta \qquad (5.1)$$

5.4 Iterative Calculation of Film Growth Rate

Once the atomic scale behavior of the system has been characterized, the next step is to examine transport of material to the feature and withing the feature itself. In this work, gas phase collisions were neglected and transport modeled as a line-of-sight process. The neglect of collisions in the gas phase corresponds to low pressure conditions where the mean free path of the particle is much larger than the dimensions of the feature. Thus, very few collisions would take place within the domain of the feature scale simulation. Within the context of the feature scale model, material arriving from the reactor chamber is treated as arriving from a 'source' just above the substrate surface. The flux from the source of each species, and its angular and energy distrubution are specified as inputs to the model. After leaving the source the impinging particles can experience three possible surface reactions, namely adsorption, sputtering and reflection. From an atomistic point of view, if a highly energized particle or ion hits the surface at a point ***r***, it will loose part of its kinetic energy and will, depending on its impact angle and the surface orientation, either be adsorbed, reflected or will sputter away other surface atoms.

Let us assume the particle is reflected. Then it can subsequently hit another part of the surface, ***r*** and get either adsorbed or again reflected, etc. In such a scenario, the point ***r'*** acts as an additional source. For a typical trench profile these two contributions are illustrated in Fig. 5.5

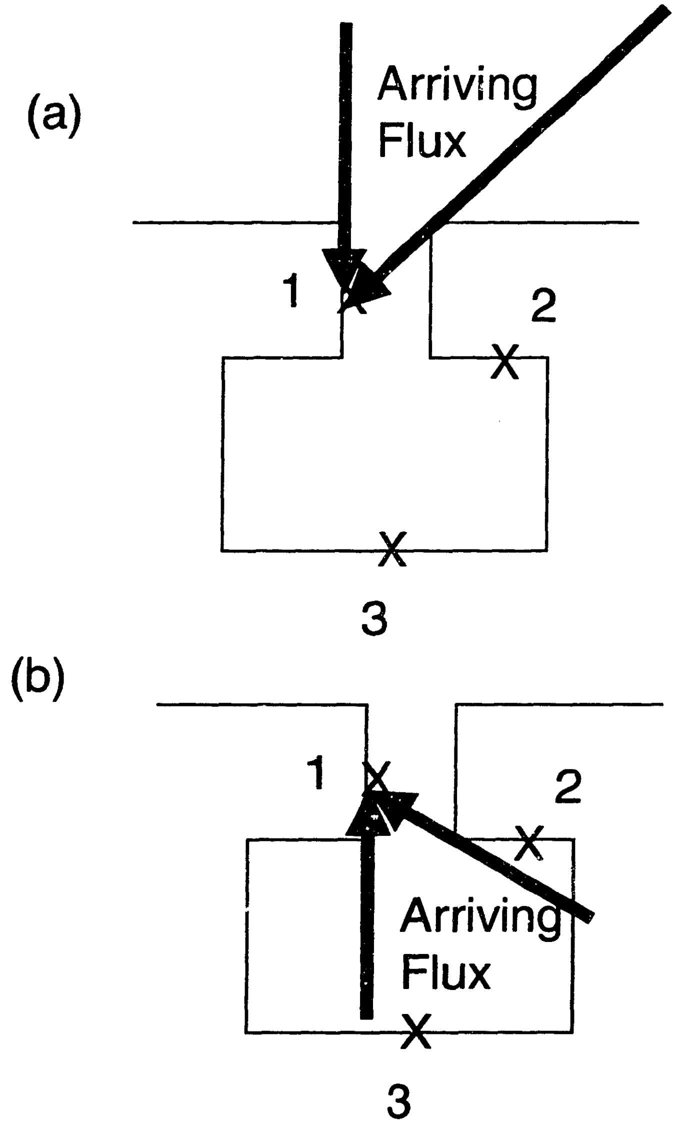


Figure 5. 5 Overhang profile used for this case study. Points (1), (2), and (3) mark reference positions at the sidewall, undercut, and bottom of the feature. Panel (a) depicts the direct flux for point (1). Panel (b) depicts the re-emitted flux that can reach (1).

Panel (a) depicts for a point (1) at the sidewall the range of incident impact angles for the direct flux. The arrows show the directions of the source particles, the dark grey area marks the angular range of the source particles. In panel (b), the origin of the re-emitted flux arriving at (1) is shown. Now the flux is arriving from other points along the surface, and the gray area marks those points that can be 'seen' by (1). Both contributions have to be considered for an accurate description of the front velocity at (1).

We first focus on the contribution of the direct flux and then turn our attention to the reemitted flux. Thus, in every system with microscale topography, a non-unity sticking coefficient and/or a non-vanishing probability for re-sputtering implies that the total arriving flux stems from two possible reaction pathways. The first one is the direct flux, i.e. the flux of atoms that leave the source and arrive at the surface before contacting another part of the feature. The second source of material is the flux due to reemissions, which stems from reflection and sputtering events. We term the contribution of the source as direct flux and the contributions due to reemissions as reemitted flux.

During IPVD growth, in addition to metal atoms arriving from the source, there will also be a significant flux of rare-gas ions, like Ar^+. The source, emits each species, i with the angular and energy dependent vector flux density, $\Phi_i^S(E \ ,\alpha_s)$, where E is the energy of the released particle, α_s is the off-normal angle with respect to the normal of the source plane and i is an index that describes the type of emitted article (i.e.metal 1, metal 2, ..., rare-gas 1, rare-gas2 , ...). There is no restriction with respect to the angular and energy dependence of the source function and it will in general depend on the deposition conditions and on the details of the chamber. $\Phi_i^S(E \ ,\alpha_s)$ can be either described by a simulation or input directly by the user. Furthermore it is assumed that the source is uniform on the length scale of the simulated. Knowing the fluxes from the source, the feature geometry, and *G*, the incident flux of each species any point ***r***, on the feature surface can be written as: Eq. 5.2

$$\Phi_i^{in}(\mathbf{r},E,\theta)=\Phi_i^{S}(E,g[\theta])+\Phi_i^{out}(\mathbf{r}',E,g[\theta]) \tag{5.2}$$

The first term on the right represents the direct flux, the second represents the flux re-emitted from other points ***r'*** of the feature surface. The ray joining emitting points ***r'*** with receiving point ***r***, makes angle theta with the receiving surface normal, and angle $g[\theta]$with the normal of the emitting surface. An obvious difficulty with the equation above is that while the source flux is known a priori, the re-emitted is not. The implicit nature of the equation suggests an iterative approach, and the front velocity is thus obtained with the following scheme:

1. Perform MD simulations to describe surface interactions for each important energetic species.
2. Calculate the flux arriving directly from the source for all points along the surface. Use the molecular dynamics results to determine the angular distribution and amount of reemitted flux, (reflections and sputtering).
3. For each point at the surface consider the reemitted fluxes from the other surface points, in addition calculate the distribution of the flux which is again reemitted.
4. Repeat 2 as long as the reemitted flux is less than a small fraction of the initial flux, such that further iterations will not change the front velocity.

$$\Phi_i^{in,(0)}(\mathbf{r},E,\theta)=\Phi_i^{S}(E,g[\theta]) \tag{5.3}$$

$$\Phi_i^{in,n+1}(\mathbf{r},E,\theta)=\int_0^{\infty}\int_{-\pi/2}^{\pi/2}G_i(E,\theta\rightarrow E',\theta')\Phi_i^{in,(n)}(\mathbf{r}',E',\theta')dE'd\theta' \tag{5.4}$$

During implementation of the equations above we discretize the integrals, and store the source and reemitted fluxes in matrices. This easily allows an angular resolution of 1° and a resolution in energy space of about 1 eV, such that we are constrained only by the available MD or experimental data. We have so far omitted a discussion of surface diffusion, since we want to focus on the effects of the deposition and redeposition. Surface diffusion would yield an opposing effect to the buildup of cusps due to preferential sputtering at certain angles and hence would make the interpretation of the predicted thin film structures more difficult. In addition recent scanning electron microscope pictures show that for low temperature deposition the films have sharp cusps and corners, [16] implying that under these conditions no long range surface diffusion is present, or at least too slow to compete with the dominant processes of deposition and re-sputtering. It should be emphasized that the inclusion of curvature driven surface diffusion - the standard approach in continuum film growth models - according to the ideas developed by Mullins,[17] is a straight forward task within the level-set formalism since curvature information is naturally obtained.

5.5 Example of Al Thin Film Growth

In the following section we demonstrate the capabilities of our new modeling approach and predict thin film topographies under various deposition conditions. We begin by choosing a source representative of typical IPVD conditions. Using the atomistic data obtained from MD calculations we are able to predict Al thin film growth without using adjustable parameters for the deposition rules.

5.5.1 Distribution Function for the Source

In this subsection we discuss an approximation for the energy and angular dependence of an IPVD source based on calculations by Kratzer et al.[18] For conventional sputtering the energy and angular distribution is well described by a Thompson distribution as experimentally confirmed.[19, 20] For ionized physical vapor deposition this picture completely changes. Under IPVD conditions the metal atoms knocked out of the source by argon ions experience ionization as they pass through a high-density plasma. The electric field at the biased substrate collimates and accelerates the ions. The normal velocity of the ions (metal and rare-gas) can be approximated by a constant distribution that starts at the so called self bias and extends to the applied bias voltage[18] The transverse component of the velocity is well described by a Maxwellian distribution with a temperature typically equal to 0.1-0.3 eV. Recent experimental studies on the deposition of Ti revealed a transverse temperatures in the range between 0.13 and 0.18 eV. A large fraction, 80% of the metal atoms is ionized and is described within the model outlined above, the remaining 20%, the neutrals are assumed to follow a Thompson distribution. For the total distribution we superpose the latter contributions. Fig. 5.6 shows the energy and angular dependence of the model source function for 80% ionization, a self bias voltage of 20 eV and a transverse temperature of 0.2 eV.

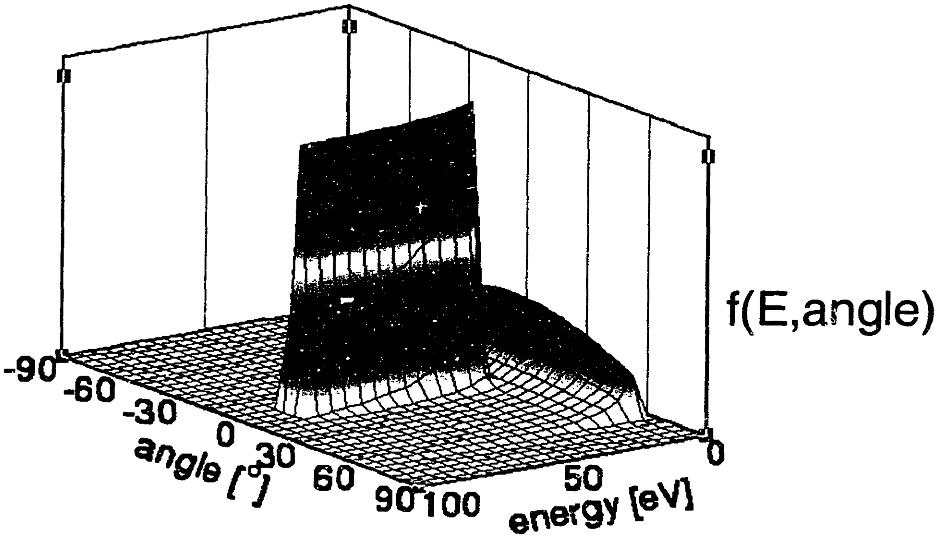


Figure 5. 6 Distribution o f Al particles emitted from an IPVD source as a function of energy and off-normal angle. A self bias voltage of 20 V , an ionization of 80 % and a transverse temperature of 0.2 eV is assumed. The distribution function is shown for a voltage of 80 V

A bias voltage of 80 V is applied, and there are two characteristic regimes: the low energy regime (15 eV) which displays a broad angular distribution due to the neutral atoms and is peaked at low energies (3 eV), but has a long tail towards higher energies. The second one is the high energy regime which is due to the ionized atoms and has in contrast a much narrower angular width. For the larger bias voltage the range of energies extends to higher values and due to the normalization the corresponding peak is less pronounced than in the case of the lower bias voltage. As the mass of Ar and Al is very similar their ion distributions are not very different. Since only ionized argon with a high kinetic energy can etch away Al atoms of the film front or suffer a reflection, thermal Ar atoms will not influence the film growth and are not included in our model.

In the following section we illustrate the importance of using angular and energy resolved surface reaction probabilities in thin film growth modeling. Next we discuss the major differences in the film topographies for conventional and ionized sputter deposition and the last section describes the effect of Ar ion bombardment on the growing film front.

5.5.2 Effect of reemission and spectral resolved reaction probabilities

To gain a better understanding of the importance of the different fluxes(direct and reemitted) we discuss their contribution for a structure representative of those encountered in semiconductor manufacturing. Fig.5.5 shows a side-view of this structure and (1) and (3) mark points at the sidewall and at the bottom respectively. In the following we restrict the discussion for simplicity to the deposition of Al (without the rare-gas Ar) which is emitted according to the distributions described in the previous section. We assume a self-bias voltage of 20 V, an ionization fraction of 80%, an applied bias voltage of 80 V and a transverse temperature of 0.2 eV.

Due to geometrical limitations and the narrow angular distribution of the ionized flux of the source all particles from the source that reach the bottom of the feature, i.e. point(3) will impinge near normal at the surface. According to our molecular dynamics based reaction rates these particles will either be adsorbed or they will sputter other Al atoms away from the surface. There will be hardly any reflection events. Consequently the major contribution to the reemitted flux will be due to etching events. Panel (a) of Fig.5.7 shows as a function of energy and off-normal angle the distribution of flux that is reemitted.

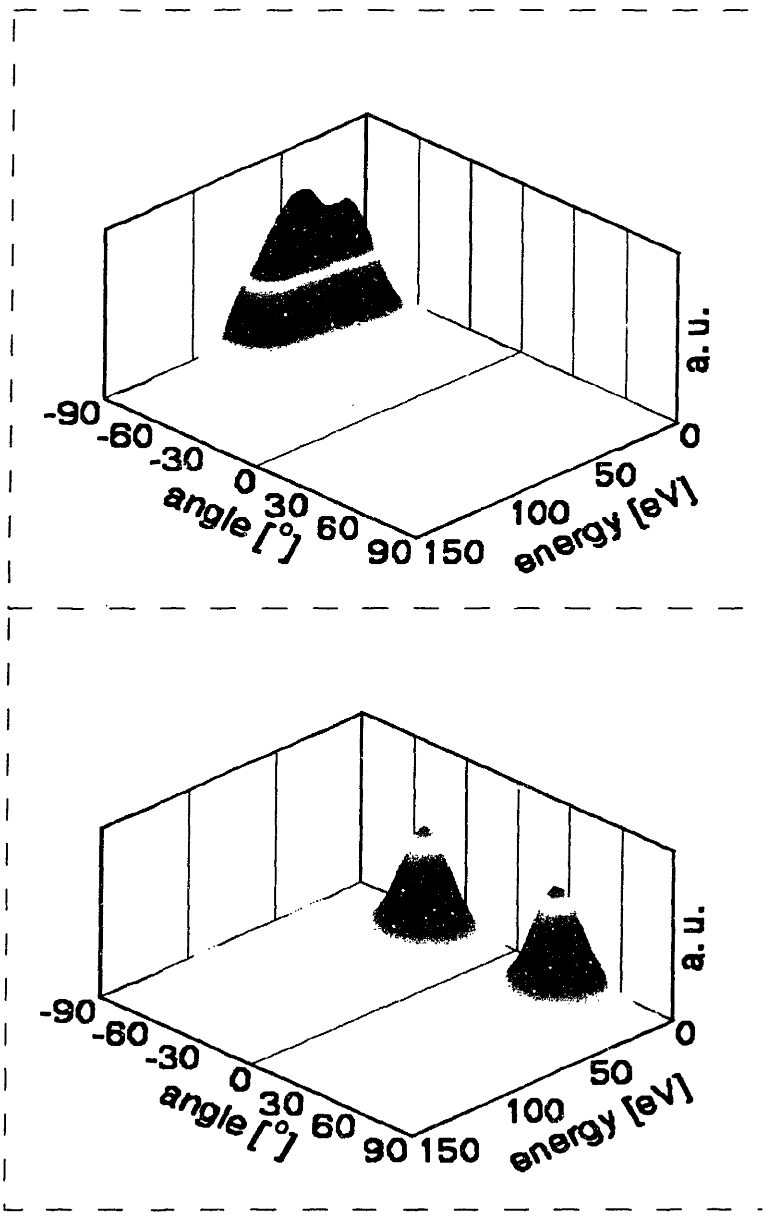


Figure 5. 7 re-emitted flux as a function of angle and energy for the bottom (a) and the sidewall (b) of an overhang structure using molecular dynamics reaction rates. The depicted fluxes correspond to points (1) and (3) respectively in Figure 5-5.

There are two peaks centered around 35°, the preferred direction for reemitted particles leaving the surface.[9] Although the source emits a wide range of energies, most of the particles leaving the surface at (3) have a rather low energy. Due to the wider angular distribution of the non-ionized fraction of the Al atoms these atoms are preferentially adsorbed at the top or side-walls of the structure such that they do not significantly contribute to the reemitted flux at (3).

The reemitted flux at the sidewall of a feature [point (1) in Fig.5.5] is quite different. The ionized contribution to the incoming flux hits the sidewall at near glazing angles. Thus reflections will be the dominant process at this point. In addition a small fraction of the non-ionized particles can participate in etching events. Panel (b) shows the reemitted flux for point (1) at the sidewall. As can be seen the dominant contribution is peaked at rather large off-normal angles and extending to rather high energies. Thus, the flux of metal shows both significant quantitative and qualitataive (in terms incident angles and energetics) variation over the surface of the feature.

The next section will demonstrate how these differences in flux influence the final conformality of the deposited film. In this case, notice that The reemissions from the sidewall will lead to an enhanced filling of the bottom of the feature and the reemission from the bottom will bring material to portions of the trench that can not be reached directly by the source, e.g. point (2) in Fig.5.5

5.5.3 Comparison of PVD and IPVD growth

We first take a closer look at the differences of the thin film properties under conventional magnetron sputtering conditions and ionized physical vapor deposition. For PVD deposition - as discussed earlier the distribution of the arriving particles is described by a Thompson distribution. Due to the geometry of a deposition chamber the angular distribution does not extend to angles up to 90°, but will be limited by typically 60° as suggested by Hamaguchi.[16] Since for the PVD case, most of the atoms have a rather low energy we do not expect a significant contribution to the speed function due to reflection and sputtering events. Only the atoms in the high energy tail of the distribution carry enough kinetic energy not to be adsorbed upon their initial impact. For simplicity, we again restrict our predictions to the deposition of Al neglecting the effect of Ar here. Fig. 5.8 illustrates deposited film profiles for a trench.

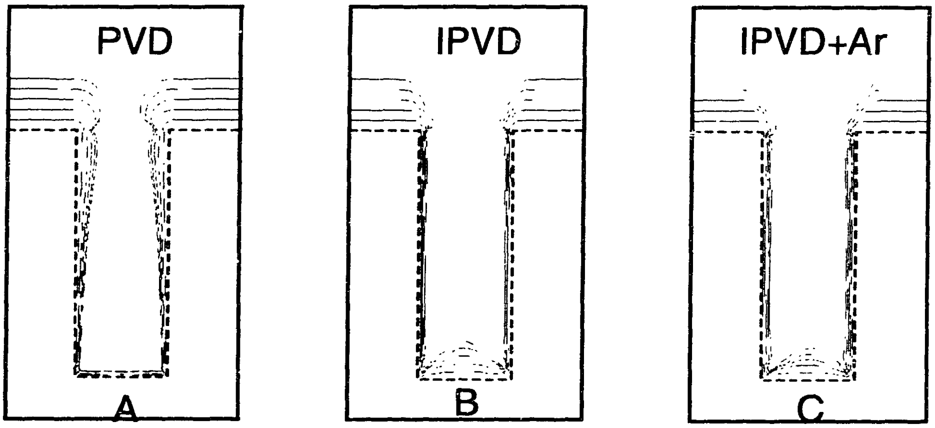


Figure 5. 8 Simulated topographies for an aspect ratio 3 trench, width 0.2m. In panel (a) an Al film under PVD conditions, panel (b) shows results for ionized PVD with 80% ionization fraction. Panel (c) represents the same conditions as (b) with a 1:1 flux ratio of Ar+

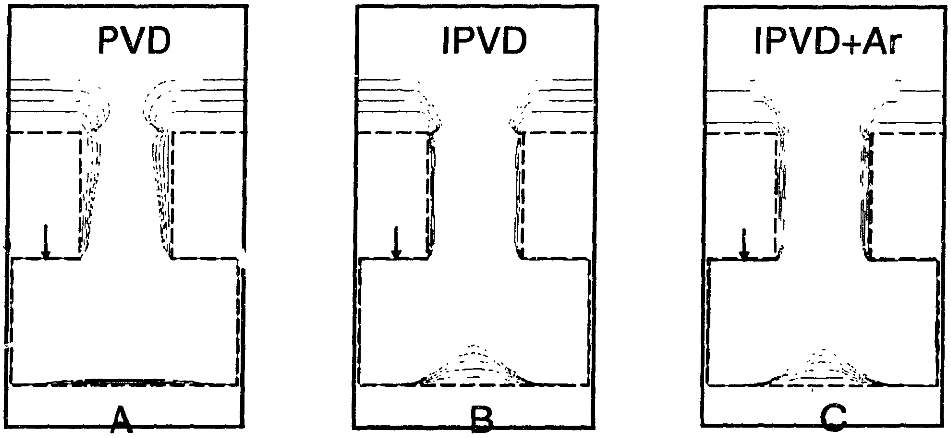


Figure 5. 9Simulated thin film topographies for an overhang structure. Deposition conditions in panels (a), (b), and (c) are identical to those in figure 5.8. The arrows mark positions on the growth front shadowed from the source. Film thickness is ~0 nm, ~2nm and 5nm in panels (a), (b), and (c) respectively.

Panel (a) in Fig. 5.8 and 5.9 shows the resulting thin film topographies deposited on an overhang and a trench with an aspect ratio of 3:1. The solid contour lines show, for constant time intervals, the evolving film front, the dashed line depicts the initial geometry before the metal deposition.

Due to the relatively wide distribution in angular space only few particles can reach the bottom of the trench structure. Most particles are adsorbed when they hit the trench sidewall thus leading to a buildup of material at the opening. This buildup of material further shadows the the bottom region of the trench such that in the course of the deposition less material can make its way to the bottom. Another characteristic is that there is no deposition right below the accumulated material. This is due to the fact that the source does not emit at large off-normal angles and no material can reach this portion of the film. This leads to a cusp in the film right below the opening of the trench. For high aspect ratios the buildup of material at the opening of the trench will finally cause a closure at the top of the structure and a void in the metal film will persist. The hole in the metal film enhances electromigration in the finished device. This adversely affects the reliability of such a structure.

Next we will discuss the effect of the ionized atoms in IPVD deposition. For an IPVD source with a emission distribution as depicted in Fig.5.6 (b) we show the film front at different stages of growth in Fig.5.8 and 5.9 (b). Here Al thin film growth is considered neglecting the effects of Ar. As in panel (a) of Fig.5.8 and 5.9, the solid contour lines show for constant time intervals the evolving film front, the dashed line depicts the initial geometry before the metal deposition. To make a qualitative comparison between the different deposition conditions the front velocity was scaled such for the PVD and IPVD case [panels (a) and (b) respectively] the growth rates on the flat surface were equal. Comparing panels (a) and (b) there are two major differences between the different deposition conditions: first the increased bottom and sidewall coverage for the IPVD source and second the appearance of bevel film front at the opening of the trench structure. Due to the collimation of the impinging atoms there is nearly as much material deposited at the bottom of the trench as on the flat surface. An improvement in sidewall coverage is obtained as material is re-sputtered from the large deposit at feature bottom to the sidewalls.

The pile-up of material at the bottom is beveled towards the edges due to the maximum in the etch rate at angles around 45° [see Fig.5.7]. This maximum in the etch rates also leads to the development of the corner at the opening of the trench structure. Since in the PVD case the energy of most of the impinging metal atoms is too low to etch away surface atoms during impact there is no effect of beveling and the film at the opening of the trench has a smooth shape. The above discussion clearly shows the advantages of the IPVD deposition compared to the PVD deposition. In the following we will address the effect of the Ar atoms.

5.5.4 IPVD Deposition under different Ar fluxes

In this sub-section we investigate the effect of the Ar ions during metal thin film growth. As discussed in Sec.5.51 the Ar and Al atoms have approximately the same distribution at the source.

The ratio of Ar to Al flux depends strongly on the deposition conditions and is unknown for most systems. In the following we will address the effect of the additional

Ar on the film . Ar influences the growing film by sputtering deposited Al atoms from the surface. With this in mind, we expect that the Ar atoms will further reduce the buildup of overhangs at the opening of the structure. In addition, energetic Ar will also sputter material at the bottom of the feature, reducing the 'pile up' at the bottom of the trench. The material sputtered from the bottom of the trench will reach the sidewall and increase the film thickness there. In order to investigate the impact of the Ar ions on the resulting film structures we performed simulations for for an Al/Ar ratio of 2:1 and 1:1.

Using the same deposition conditions as for the Al IPVD deposition described previously and again scaling the front velocity to the same value at the flat portions of the surface we show the resulting thin film topographies in panel (c) of Fig.5.8 and 5.9. Panel (c) corresponds to a Al/Ar ratio of 1:1. It is evident that the additional Ar bombardment according to the discussion above improves the sidewall coverage, reduces the pile up of material at the bottom of the trench and reduces the buildup of overhangs at the opening of the trench. Thus the overall film is more conformal compared to the IPVD deposition without Ar and hence comes closer to the properties the micro-electronic industry desires. We propose that the adjustment of the Ar flux is a promising way to achieve highly conformal metal thin films, if acceptable deposition rates may be maintained.

5.6 Conclusions

We have presented a general approach to implement angular and energy dependent surface reaction rates within the level-set formalism. We begin by presenting molecular dynamics data for hyperthermal Al and Ar atoms interacting with Al surfaces and discuss the energy and angular dependence of the three major surface reactions, namely adsorption, reflection and sputtering. We summarize the surface interactions in a function

G which calculates the flux returning from the surface for a given incident flux. Using the MD data in this fashion, we construct a gereral growth rate model examine Al thin film growth under various deposition conditions. We describe in detail an iterative method to calculate the speed function (growth/etching) rate. Finally, the information is passed to a level set description of the evolving film. The resulting simulator thus allows atomic scale MD information to be directly incorporated in a micron scale descrption of the evolving thin film. This approach is much faster than an atomistic Monte Carlo model, particularly when rare gasses are included in the simulation. Our calculations capture the major differences between IPVD and PVD deposition and display the strengths of deposition with high-energy ions, namely a more conformal sidewall coverage and a better bottom coverage. Profiles with and without Ar ion fluxes have been compared; the additional sputtering of Ar further reduces the buildup of material at the opening of the trench structure and enhances the conformality in overall film thickness. We propose that the Ar flux can be a useful tuning parameter in micro-electronics processing.

1. Cerio, F., *et al.*, *J. Vac. Sci. Technol. A*, 1999. **16**: p. 1863.
2. Dabrowski, J., *et al.*, *Advances in Solid State Physics*, 1998. **38**: p. 595.
3. Hanson, D.E., A.F. Voter, and J.D. Kress, *J. Appl. Phys.*, 1997. **82**: p. 3552.
4. Kress, J.D., *et al.*, *submitted to J. Vac. Sci. Tech.*, .
5. Coronell, D.G., *et al.*, *Appl. Phys. Lett.*, 1998: p. 3860.
6. Frenkel, D. and B. Smit, *Understanding Molecular Simulations: From Algorithims to Applications*. 1996, Boston: Academic Press.
7. Allen, M.P. and D.J. Tildesley, *Computer Simulation of Liquids*. 1996, Oxford: Oxford University Press.
8. Gilmer, G.H., *et al.*, *Materials Science and Engineering B*, 1996. **37**: p. 1-7.
9. Hansen, U. and A. Kersch, *Phys. Rev. B*, 1999. **in press**.
10. Daw, M.S. and M.I. Baskes, *Phys. Rev. Lett.*, 1983. **50**: p. 1285.
11. Daw, M.S. and M.I. Baskes, *Phys. Rev. B*, 1984. **29**: p. 6443.
12. Abrahamson, A.A., *Phys. Rev.*, 1969. **178**: p. 178.
13. Hansen, U., P. Vogl, and V. Fiorentini, *Phys. Rev. B*, 1999. **59**: p. 7856.
14. Hansen, U., P. Vogl, and V. Fiorentini, *Phys. Rev. B*, 1999: p. 5055.
15. Stumpf, R. and M. Scheffler, *Phys. Rev. B.*, 1996: p. 4958.
16. Hamaguchi, S. and S.M. Rossnagel, *J. Vac. Sci. Technol. B*, 1995. **13**: p. 183.
17. Mullins, W.W., *Metall. Mater. Trans. A*, 1995. **26**: p. 1917.
18. Kratzer, M., *et al.*, *submitted to J. Appl. Phys*, 1999.
19. Westwood, W.D., in *Microelectronic Materials and Processes*, R.A. Levy, Editor. 1989, Kluwer: Dordrecht. p. 133.

20. Dullini, E., *nucl. inst. and Meth. B*, 1984. **2**: p. 610.

Chapter 6
Conclusions and Recommendations for Future Work

6.1 Conclusions

In this work a set of techniques for modeling physical processes occurring on widely separated length scales. A general method was found to link atomistic and continuum transport models. The linking method is compatible with either a probabilistic (Monte Carlo) or deterministic (ballistic integral) description of transport at the microscale. A superposition strategy is used to summarize microscale information and simulate the effects of microscale heterogeneity. Monte Carlo and level set profile simulators were used in case studies for chemical vapor deposition of tungsten and aluminum. Plasma etching of polycrystalline silicon was also explored.

Techniques for linking atomic scale and feature scale processes were also presented. Molecular dynamics simulations were used to characterize interactions of energetic particles with the growing film. The resulting method is much faster than the commonly used atom-by-atom approach, while retaining much of the accuracy of atomistic simulation.

6.2 Doing More with Multiscale Models

6.2.1 How much detail ?

In attempting to create multiscale models with resonable computational requirements, several challenges appeared time and again. The first, and most important is the need for the modeler to do a little thinking and determine the appropriate level of detail in the solution.

Beyond this, it is important to realize that in many cases, the entire system needn't be simulated at the smallest level of detail. For example, consider the etching and CVD examples presented in this work. An atomistic perspective was needed only in the vicinity of microscale topography, while the reactor at large was well described with a continuum model. More generally, a set of perspectives can be set up parallel to the discussion of length scales in the first chapter. That is, divisions between scales are made on the basis of physical assumptions that happen to coincide with divisions on a ruler. Much of the multiscale simulation work currently underway in materials science is focused on linking atomistic and continuum models. Models for grain and crystal growth, annealing, and crack propagation frequently require atomistic-continuum linking in some fashion. Although these processes might occur on very different length scales than the gas phase transport problems considered here, many of the key ideas are the same, as microscale and macroscale events interact. In general, the 'coarser' perspective is best, and in many cases most accurate.

For example – if a distribution is known to be Gaussian, there is no reason to simulate it particle by particle or even keep track of the whole distribution, for that matter, the mean and the variance of the distribution will tell the whole story, period. Any point by point simulation of our Gaussian would be subject to sampling error, so interestingly enough, the 'big picture' perspective is also the most accurate as long as the underlying assumptions are true

A hierarchy of perspectives arises where individual events are summarized as a distribution, while distributions might be summarized with a few parameters like mean and varainceThus, whenever possible, it is advantageous to replace probabilistic, atomistic simulations with deterministic simulations of their distributions. This is exactly why the ballistic integral methods in chapter 3 and 4 were more useful than the atomistic MC technique of chapter 2. The big idea, then, is to work from the bottom up, use detailed approaches where needed. This is the only way to maintain a consistent description of the key phenomena at the microsclee. The detailed description is continued until a less detailed one is acceptable. In general, if the domains interact they must be matched in an iterative fashion.

6.2.2 Efficient tools

A second guiding theme in multiscale modeling is not re-doing computations unless absolutely necessary. This concept is simple enough: the hard part is deciding what is absolutely necessary. A good example of this is seen in chapter 2, where the molecular trajectories following the first surface interaction are pre-computed. This was possible because of the Markov property. An even better example is seen in chapter 5, where all the particle-surface interactions are computed by molecular dynamics beforehand and fit to distribution functions for use in deterministic simulation. This method of summarizing atomistic events allows simulations on length scales that would not be attainable using an atom –by atom approach.

Some mathematical methods that already exist appear almost ‘tailor made’ for multiscale analysis. Reduced models could have been used to great advantage in much of this work. In chapters 2,3, and 4 , most of the simulation time was spent matching models iteratively. Each iteration involved solving microscale and reactor scale models repeatedly while making small changes to model boundary conditions. Creating reduced models at the start of simulation could have yielded a large time savings. This issue will become even more important if fully 3-dimensional simulations are used.

A Green's function approach could also be useful. If a Green's function were written for each of the microscale models in chapters 2,3,and 4, and the boundary condition treated as a 'forcing function', I suspect a very general framework for rapid multiscale calculations could be developed.

6.3 Suggestions for Future Work

6.3.1 Continuing efforts in plasma etch

All things considered, the multiscale model of plasma etching presented here while instructive from a methodological point of view has only scratched the surface in terms of achieving quantitative, accurate simulation. There is a lot of additional work possible on this subject, as one of the most technologically important multiscale problems anywhere is almost certainly micro loading during plasma etching. If the linked simulation strategy presented here could be combined with a state of the art equipment model and a more complete description of plasma-surface chemistry, some very important questions could be answered, such as: what is the maximum etching rate possible with acceptable uniformity ? Can pressure, and thus etch rate be increased without incurring loading effects ? If byproducts build up in the reactor, what will the final etch profiles look like. ? Currently, these questions and (many more like them) are addressed by trial and error, a processes that could be speeded up considerably if a linked multiscale simulator was used in tandem. The same, (or very similar) issues exist in plasma enhanced CVD processes, another field of intense practical interest.

6.3.2 Higher pressure CVD processes

As deposition pressures are increased, gas phase diffusion coefficients decrease, while reaction rates often rise. As a result, loading effects during atmospheric pressure CVD are often severe and present an important problem, particularly in the manufacture of compound semiconductor devices. We have already demonstrated that at sufficiently low pressures, (below about 1 torr) a collisionless or view factor model of gas phase transport can also be used to map feature scale details to a 'control surface' compatible with a finite element model. Monte Carlo results suggest that the height of the control surface must be less than 1/3 a mean free path for the collisionless picture to be reasonably accurate. Fortunately, this height allows enough room to relax horizontal gradients in flux to a characteristic length of about 7 mean free paths, resulting in a Knudsen number, Kn of about 0.14, near the rarefied limit for a continuum formulation. In this manner, we summarize reaction and transport on all scales requiring an atomistic approach in a fashion compatible with a continuum formulation.

For low pressures, 7 mean free paths may be a significant distance- 0.15 mm at 1 torr. Thus, a continuum CVD or etch model over a typical wafer (30 cm) is likely employ no more than 150 – 200 nodes without violating kinetic theory.

Higher pressures present a more difficult challenge. At 100 torr placing nodes 7 mean free paths apart would lead to 15,000 to 20,000 nodes over the substrate surface, a situation that is in practice impossible to solve, in spite of being completely correct. This problem is thus distinct from this linking work in that we are addressing a numerical problem rather than the physical one of our previous efforts.

THESIS PROCESSING SLIP

FIXED FIELD: ill. ______ name ______

index ______ biblio ______

► **COPIES:** [✓ Archives] Aero Dewey Eng Hum

Lindgren Music Rotch [✓ Science]

TITLE VARIES: ►☐ ______

NAME VARIES: ►☐ ______

IMPRINT: (COPYRIGHT) ______

► **COLLATION:** 123l

► **ADD: DEGREE:** ______ ► **DEPT.:** ______

SUPERVISORS: ______

NOTES:

cat'r: date:

page: ► T20

► **DEPT:** Chem. Eng.

► **YEAR:** 2000 ► **DEGREE:** Ph.D.

► **NAME:** [illegible]