

Multi-scale Modeling of Chemical Vapor Deposition: From Feature to Reactor Scale

by

Jonathan Jilesen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Mechanical Engineering

Waterloo, Ontario, Canada, 2009

©Jonathan Jilesen 2009

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Multi-scale modeling of chemical vapor deposition (CVD) is a very broad topic because a large number of physical processes affect the quality and speed of film deposition. These processes have different length scales associated with them creating the need for a multi-scale model. The three main scales of importance to the modeling of CVD are the reactor scale, the feature scale, and the atomic scale. The reactor scale ranges from meters to millimeters and is called the reactor scale because it corresponds with the scale of the reactor geometry. The micrometer scale is labeled as the feature scale in this study because this is the scale related to the feature geometries. However, this is also the scale at which grain boundaries and surface quality can be discussed. The final scale of importance to the CVD process is the atomic scale.

The focus of this study is on the reactor and feature scales with special focus on the coupling between these two scales. Currently there are two main methods of coupling between the reactor and feature scales. The first method is mainly applied when a modified line of sight feature scale model is used, with coupling occurring through a mass balance performed at the wafer surface. The second method is only applicable to Monte Carlo based feature scale models. Coupling in this second method is accomplished through a mass balance performed at a plane offset from the surface.

During this study a means of using an offset plane to couple a continuum based reactor/meso scale model to a modified line of sight feature scale model was developed. This new model is then applied to several test cases and compared with the surface coupling method. In order to facilitate coupling at an offset plane a new feature scale model called the Ballistic Transport with Local Sticking Factors (BTLFSF) was developed. The BTLFSF model uses a source plane instead of a hemispherical source to calculate the initial deposition flux arriving from the source volume. The advantage of using a source plane is that it can be made to be the same plane as the coupling plane. The presence of only one interface between the feature and reactor/meso scales simplifies coupling. Modifications were also made to the surface coupling method to allow it to model non-uniform patterned features.

Comparison of the two coupling methods showed that they produced similar results with a maximum of 4.6% percent difference in their effective growth rate maps. However, the shapes of individual effective reactivity functions produced by the offset coupling method are more realistic, without the

step functions present in the effective reactivity functions of the surface coupling method. Also the cell size of the continuum based component of the multi-scale model was shown to be limited when the surface coupling method was used.

Thanks to the work done in this study researchers using a modified line of sight feature scale model now have a choice of using either a surface or an offset coupling method to link their reactor/meso and feature scales. Furthermore, the comparative study of these two methods in this thesis highlights the differences between the two methods allowing their selection to be an informed decision.

Acknowledgements

The author wishes to thank his advisor Dr. Fue-Sang Lien for his guidance and support, the Natural Sciences and Engineering Research Council of Canada for financial support, Arthur F. Church for his donations to the Arthur F. Church Graduate Scholarship, and SHARCNET for the use of their computer network.

Dedication

Dedicated to my loving wife, Kaholi Jilesen for her patience and support during this time of having little.

Table of Contents

List of Figures	ix
List of Tables	xii
Nomenclature	xiii
Chapter 1 Introduction.....	1
1.1 Background	2
1.2 Outline	3
Chapter 2 Reactor Scale	5
2.1 Theory	6
Navier Stokes equation.....	7
Energy Equation	7
Species balance equations	7
2.2 Numerical Method.....	10
2.2.1 Finite Volume Method	10
2.2.2 Quick Scheme.....	12
2.2.3 Upwind differencing.....	14
2.3 Boundary Conditions.....	14
2.4 Geometry and Grid.....	16
2.5 Validation	18
Chapter 3 Feature Scale.....	22
3.1 Theory	23
3.1.1 Surface Tracking	30
3.1.2 Automatically varying time step.....	31
3.2 Validation	34
3.3 Local Sticking Coefficients	37
3.4 Application to non-first order chemistry	38
3.5 Source Plane Height Sensitivity	40
3.6 Closure.....	42
Chapter 4 Coupling	43
4.1 Theory	44
4.1.1 Downscaling of reactor scale to feature scale.....	44
4.1.2 Homogenization of feature scale results to reactor scale.....	47

4.1.3 Limitations of coupling methods.....	50
4.1.4 Use of Meso Scale.....	63
4.1.5 Solution Algorithm.....	64
4.2 Results.....	66
4.2.1 Silicon Deposition.....	68
4.2.2 Tungsten Deposition.....	75
4.2.3 Transient Results.....	78
4.3 Discussion.....	82
4.4 Closure.....	85
Chapter 5 Conclusions and Contributions.....	87
5.1 Conclusions and Contributions.....	87
5.2 Recommendations for future work.....	91
Appendices.....	93
Appendix A Calculation of transport properties.....	94
Appendix B Chemical reaction kinetics.....	101
Bibliography.....	103

List of Figures

Figure 2-1 Labeling of cells for QUICK scheme.	12
Figure 2-2 Geometry and grid for reactor scale. [(A) Inlet, (B) Outlet, (C) Reacting Surface, (D) Non-reacting wall ($T=T^0$), (E) Non-reacting wall ($dT/dx=0$), (F) Non-reacting wall].....	17
Figure 2-3 Contour plots for the silane mole fraction. Left: without thermal diffusion. Right: with thermal diffusion. (□) Digitized results from Kleijn <i>et al.</i> [□]	19
Figure 2-4 Comparison of growth rates for hydrogen reduction of silane with hydrogen as the carrier gas. No thermal diffusion: current code (—), Kleijn 1989 [□] (□). Thermal diffusion: current code (- - -), Kleijn 1989 [□] (Δ).	20
Figure 2-5 Comparison of growth rates for hydrogen reduction of silane with a hydrogen-nitrogen mixture as the carrier gas. Current code course mesh (—), Current code fine mesh (· · · ·), Kleijn 1989 (□).	21
Figure 3-1 Schematic showing the definition of angles θ_a , θ_b , w_1 and w_2	29
Figure 3-2 Schematics showing the definition of angles used in calculation of 'transmission probability': (A) for BTLFSF, Equation (3-12) (B) for BTRM, Equation (3-13).	30
Figure 3-3 Variation of N (—) and τ (····) as feature closes: $AR = 1.0$, $\epsilon_0 = 0.01$	33
Figure 3-4 Film profiles for a first order deposition reaction in a stepped trench for a sticking of 0.1 using: (A) a fixed time step (B) a variable time step. Every time step is plotted for the variable time step case while every fourth time step is plotted for the fixed time step case.	34
Figure 3-5 Instantaneous step coverage as a function of feature closure for a first order deposition reaction in a $1\mu\text{m}$ by $1\mu\text{m}$ trench. (- - -) Cale and Raupp [□] , (—) BTLFSF.	36
Figure 3-6 Comparison of BTLFSF (— - —) results with the HVS (• • • •) and MC (x x x) results of Yun and Rhee [□] . The solid line indicates the initial profile (—).	37
Figure 3-7 Variation in step coverage with feature aspect ratio for Order 0.5 (∇), Order 1.0 (□), Order 2.0 (O), and Eley-Rideal (◇) kinetics.	40
Figure 3-8 Change in surface flux prediction of BTLFSF with source plane height.	41
Figure 4-1 Estimate of error in surface flux resulting from the assumption of a Maxwell distribution.	47
Figure 4-2 Schematic showing definition of angle θ	49

Figure 4-3 Dependence of ξ distribution on source plane height for aspect ratio 10 feature with a width of $0.1\mu\text{m}$, where $h = \lambda$ (Δ), $h = 2/3 \lambda$ (\circ), and $h = 1/3 \lambda$ (\square).....	53
Figure 4-4 Variation of ξ function with aspect ratio (AR) when the source plane height, $h = 1/3 \lambda$ and the feature width is $0.1 \mu\text{m}$. [(\square) AR = 10, (\circ) AR = 2, (Δ) AR = 1, (\diamond) AR = 0.5.]	54
Figure 4-5 Variation of ξ function with feature width for feature with AR of 10 and source plane height, $h = 1/3 \lambda$. [width = $0.1\mu\text{m}$ (\square), $0.25 \mu\text{m}$ (Δ), $0.5 \mu\text{m}$ (\circ)].....	56
Figure 4-6 ξ function resulting from linear superposition of five features with a source plane height of $h = 1/3 \lambda$ and the features widths are $0.1 \mu\text{m}$. [Single feature (\square), five features with pitch of $2\mu\text{m}$ (Δ), five features with pitch of $6\mu\text{m}$ (\circ)]	57
Figure 4-7 Dependence of ξ distribution on sticking factor for a aspect ratio 10 feature with a source plane height, $h = 1/3 \lambda$ and a feature width of $0.1 \mu\text{m}$, where $\epsilon_0 = 0.5$ (\circ), $\epsilon_0 = 0.1$ (\diamond), $\epsilon_0 = 0.01$ (Δ), and $\epsilon_0 = 0.001$ (\square).....	58
Figure 4-8 ξ profiles for surface and offset methods and their representation on the reactor scale, $\bar{\xi}$ for $0.1 \mu\text{m}$ feature with AR 10 and source plane height of $1/3\lambda$. [surface coupled ξ profile (\square), $\bar{\xi}$ profile (solid line); offset coupled ξ profile (Δ), $\bar{\xi}$ profile (dashed line)].....	62
Figure 4-9 Effect of solving Navier-Stokes equation at meso scale on meso scale surface velocity profiles. (- - -) Velocity field interpolated from reactor scale. (—) Velocity field found using meso scale Navier-Stokes equation.....	64
Figure 4-10: Algorithm for combining multiple scale simulation for transient simulations.....	67
Figure 4-11 Growth rate results with features uniformly distributed over reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) reactor scale with coupling at source plane, (\circ) reactor scale coupled directly to feature scale, (- - -) meso scale with coupling at surface, and (—) meso scale with coupling at source plane.	70
Figure 4-12 Variation of silane mass fraction on surface without features (- - -) and with features (—) for test case 1. [(Δ) Coupling at surface, (\square) Coupling at source plane]. Vertical lines indicate reactor scale cell width.	71
Figure 4-13 Growth rate results for uniformly clustered features within reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) Reactor scale with coupling at source plane, (- - -) Meso scale with coupling at surface, and (—) Meso scale with coupling at source plane.	72

Figure 4-14 Feature scale geometry for test case with non-uniformly distributed features.	74
Figure 4-15 Growth rate results for non-uniformly distributed features within reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) Reactor scale with coupling at source plane, (- - -) Meso scale with coupling at surface, and (—) Meso scale with coupling at source plane.....	74
Figure 4-16 Variation of silane mass fraction on surface without features (- - -) and with features (—) for test case 3. (Δ) Coupling at surface, (\square) Coupling at source plane.....	75
Figure 4-17 Tungsten growth rate results with features uniformly distributed over reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) Reactor scale with coupling at source plane, (- - -) Meso scale with coupling at surface, and (—) Meso scale with coupling at source plane.....	77
Figure 4-18 Variation of tungsten mass fraction on surface with coupling at surface (- - -) and with coupling at source plane (—).	78
Figure 4-19 Profile of deposited silicon displayed at 50 s intervals for features distributed as shown in Figure 4-14.	79
Figure 4-20 The reduction of the meso scale effective growth rate profiles displayed at 50 s intervals. [Offset coupling method (- - -), Surface coupling method (—)]	80
Figure 4-21 Silane mass fraction profile close to substrate surface displayed at 50 s intervals. [T = 0 s (thick solid line), T = 250 s (thick solid line with Δ), T = 1000 s (dashed line)]	81
Figure 4-22 Variation of growth rate with mass fraction over the substrate surface for silane gas chemistry (Left) and tungsten hexafluoride (Right).	83
Figure 4-23 Silane mass fraction profiles for substrate uniformly covered in features (—) compared with that of a featureless substrate (- - -).	85

List of Tables

Table 2-1 Coefficients resulting from QUICK scheme.....	13
Table 2-2 Coefficients resulting from upwind differencing.....	14
Table 4-1 Variation of ξ with aspect ratio with sticking. ($\epsilon_0 = 0.0009$)	60
Table 4-2 Variation of ξ with feature width. ($\epsilon_0 = 0.0009$)	60
Table 4-3 Variation of ξ with sticking factor for 0.1 μm wide feature	61
Table A-1 Fitting constants for transport properties [□]	95
Table A-2 Fitting constants for binary diffusion coefficient [□]	96
Table A-3 Fitting constants for binary thermo diffusion ratios [□]	96
Table A-4 Molecular mass and Lennard-Jones force parameters for tungsten deposition test case [□] ...	96

Nomenclature

Roman Letters

A	area
\tilde{A}	surface arc length per unit depth
AR	aspect ratio
a	coefficient
a_{kj}	thermal diffusion factor
B	desired time step ratio
C	convected mass flux per unit area
C_0	stream velocity
c_p	specific heat at constant pressure
D	diffusion conductance
D'	effective diffusion coefficient
D_{kj}	thermal diffusion factor
d	mean diameter of gas molecules
ds	discrete segment of surface
F	flux across source plane
f	mole fraction
G	species generation term
g	gravitational acceleration
K	proportionality constant
$K_{i,f}^g$	forward gas phase reaction rate constant
$K_{i,b}^g$	reverse gas phase reaction rate constant
Kn	Kudsen number
j	diffusion flux
j^C	flux driven by concentration gradient
j^T	flux driven by temperature gradient
L	characteristic dimension, width of reactor scale cell
M	number of discrete surface sections
m	molecular mass
N	number of re-emission stages to reach convergence

N_A	Avogadro's number
n	number of particles per unit volume
\mathbf{n}	normal vector
P	pressure
$ P_{ij} $	distance between points
$p(v)$	probability density function
q	transmission probability
R	universal gas constant
R^g	gas phase net molar reaction rate
R^s	net molar surface reaction rate
r	radius
S	source term
T	temperature
t	time
\mathbf{v}	velocity vector
w', v', u'	thermal velocity components
x	spatial co-ordinate
Y	mass fraction, spatial co-ordinate
Z_{coll}	collision frequency

Greek Symbols

α	flag indicating direction of flow across a face
β	inverse of most probable thermal speed
χ	visibility indicator
ε	sticking factor
ε_0	base sticking factor
ϕ	variable
Γ	interface diffusion coefficient
η	flux of particles into surface
φ	angle
λ	mean free path, thermal conductivity
μ	viscosity
ν_{ki}	net stoichiometric coefficient
ν	forward stoichiometric coefficient

ν'	reverse stoichiometric coefficient
θ	variable, angle
ρ	density
σ	collision diameter
τ	adjustment to growth rate time increment
ξ	effective reactivity
$\xi \cdot dx$	area under the ξ function within reactor scale cell

Abbreviations

ALD	atomic layer deposition
BTLSF	ballistic transport with local sticking factors
BTRM	ballistic transport and reaction model
CVD	chemical vapor deposition
HVS	hemispherical vapor source
ICs	integrated circuits
KD	Knudsen diffusion
KDM	Knudsen diffusion model
KTRM	kinetic transport and reaction model
LPCVD	low pressure chemical vapor deposition
MC	Monte Carlo
MD	molecular dynamics
MOCVD	metal organic chemical vapor deposition
PECVD	plasma enhanced chemical vapor deposition
QUICK	quadratic upwind differencing scheme
RSF	reactive sticking factors
SIMPLE	semi-implicit method for pressure linked equations
STREAM	simulation of turbulent Reynolds-averaged equations for all mach numbers

Chapter 1

Introduction

Chemical vapor deposition (CVD) is a process by which thin films are deposited on a surface. CVD is used in the production of such products as solar cells, semi conductors, MEMs and nanostructures like carbon nanotubes. Other processes which deposit similar films such as sputtering, sublimation or evaporation do not provide as high a quality of film^[1]. The major benefits of CVD are that it can produce excellent uniformity and the fact that it is capable of conformal deposition where the surface topography is coated while maintaining its features.

Simulation of CVD is important because CVD processes are often not scalable making lab results difficult to reproduce at industrial scales. Through simulation insight into the physics behind the process can be gained and these insights can be used to better understand difficulties with scaling and other unexplained experimental observations. Simulation is also beneficial during optimization because it can relatively quickly predict how changes to the input gas or reactor geometry will affect gas transport and growth rates.

One of the challenges of simulating CVD is the multi-scale nature of the problem. For example one of the major applications of CVD is the manufacturing of integrated circuits (ICs) for the microelectronics industry. While the wafers on which the film is deposited may be hundreds of centimeters across, the dies for the ICs will likely be close to one centimeter in diameter and the microscopic features within those dies (which make up the components of the ICs) can be less than a micron wide. Furthermore, if the crystal structure is of interest modeling of the movement of atoms on the surface may even become important. The importance of physical processes varies over this large range of scales making it difficult for a single model to account for all of the significant processes. Instead different models are used at the scales where they work best and coupled to form a heterogeneous multi-scale model. The communication between these scales is an important issue and the major focus of this thesis is the coupling between the reactor (wafer) scale and the feature scale.

The issue of coupling between (continuum) reactor and (non-continuum) feature scale models has been addressed in the past with the most successful approaches being those of Gobbert *et al.*^[2] and Rodgers and Jensen^[3]. However, these two models are very different in their approach and no direct

comparison has been conducted. In this thesis modified versions of these two approaches are rigorously compared. The modifications to the models made in this thesis allow the method of Rodgers and Jensen to be applied to a ballistic transport feature scale model. Having both models using the same feature scale model makes it easier to compare the coupling methods. The adaptation of the Rodgers and Jensen coupling method also required that a new ballistic feature scale model be developed, which used an offset plane as opposed to the hemispherical source used in previous feature scale models. Adaptations were also made to Gobbert *et al.* coupling method to increase the flexibility of their model.

1.1 Background

In Section 1.0 it was mentioned that one of the major applications of CVD is the manufacturing of ICs for the microelectronics industry. During the manufacturing process of ICs, CVD is used to deposit a series of thin films of material on the surface of a wafer (substrate). By etching away unwanted material in between deposition processes or using selective deposition, the topography of the depositions are controlled and the components which make the integrated circuits are built up.

There are many different materials which can be deposited by CVD such as tungsten, copper, nickel, silicon and silicon oxide. There are also many different CVD processes such as low pressure (LPCVD), atomic layer deposition (ALD), plasma enhanced (PECVD), and metal organic (MOCVD). These processes take different lengths of time to complete and produce different films. The process of particular interest for this study is LPCVD. This is because the pressures associated with LPCVD result in large mean free paths and create the need for using particle based methods to model the deposition within the features.

While the large mean free paths of LPCVD make it appealing for demonstration purposes, the methods applied in this paper also apply to other CVD processes. The higher pressures of the other process, however, will mean that the size of the features will have to be smaller before particle based methods can be applied due to the smaller mean free path.

There are many different reactors which are used with the LPCVD process such as barrel, stagnation, and impinging jet. These reactor designs have different advantages such as increased production rate or increased control over the deposition process. The reactor design used in this study was an

impinging jet reactor because this type of reactor can provide very uniform concentration distributions over the surface of the wafer, making it easier to observe changes in the concentrations due to the presence of features.

When simulating CVD one of the major goals is accurate prediction of the growth of the deposition within features. The thickness of deposition within these features is important to the performance of the final product. One of the applications for LPCVD of Tungsten is the filling contact holes^[4]. These contact holes connect different layers within the IC and can be submicron in diameter and have large aspect ratios. In the case of contact filling it is very important that voids do not form in the contact holes because they will adversely affect the contacts resistance and may even cause failure of the component. The accuracy of the feature filling predictions can be increased through multi-scale modeling with coupled reactor and feature scales^[2].

The prediction of film morphology is another application of multi-scale modeling with coupled reactor and micro (feature) scales^[5]. However, this application requires very different models from those used in this thesis and is beyond the scope of this thesis.

1.2 Outline

Chapter 2 deals with the reactor scale. It reviews the history of modeling CVD at the reactor scale and the theory related to the reactor scale. The numerical methods used in the current study for the reactor scale are then discussed, as well as the reactor scale geometry and boundary conditions used throughout this study. The results of the reactor scale are then verified against results from other numerical studies.

Chapter 3 focuses on the feature scale modeling of CVD. Past developments in the modeling of the feature scale are discussed as well as what is considered to be state of the art. The theory behind a new modified line of sight model developed for this study is discussed along with features such as automatically varying time stepping and local sticking factors. The feature scale code is verified against existing models and the sensitivity of the results to source plane height is discussed.

Chapter 4 pertains specifically to the coupling between reactor and feature scales. Relevant literature is reviewed and the theory behind the two main coupling methods is discussed along with the

modifications made in this study. The limitations on the validity of both methods with regards to continuity are discussed and factors which affect this limit are investigated. The results of the two different methods are compared for three different feature patterns and two different chemistries. Transient results are also presented with concentration and growth rate profiles changing as the deposition grows.

Finally, Chapter 5 summarizes the results and highlights the contributions of this study. Recommendations for areas of further study on coupling between reactor and features scales are also made.

Chapter 2

Reactor Scale

Chemical vapor deposition reactors come in many different shapes and sizes with some reactors only producing a single wafer per run while others produce batches of wafers. The different reactor designs have distinguishing features such as increased process control or high throughput and operate at different temperatures and pressures. The geometry of these reactors allows microchip manufacturers to control the deposition characteristics. This is done through careful selection of the inlet flow conditions and the use of channels and baffles to direct the flow of fluid to the surface of the wafers. Simulation of flow at this scale is important because it allows the effects of different geometry choices to be tested quickly and cost effectively, with the alternative being trial and error.

The first attempts to model CVD at this scale were performed using 2D analytical models^[6,7]. These analytical models made assumptions such as plug flow and uniform temperature in order to simplify the problem. Later numerical methods were introduced allowing more physics to be included such as multi-component diffusion, multi-reaction schemes, and varying gas properties^[8,9]. Jensen 1987^[10] provides a detailed review of the early history of CVD modeling.

State of the art reactor scale models are 3D and have mixed convection, variable fluid properties detailed gas and surface phase chemical models, thermal diffusion, multi-component diffusion, radiation modeling and turbulence modeling^[11]. The inclusion of additional physical processes makes the numerical model more realistic. However, depending on the importance of the process to a specific flow the accuracy of the predictions may not increase. For example thermal diffusion is very important in cold wall reactors while not significant in hot wall reactors. Furthermore the computational cost of performing simulations increases with each additional component included.

Multi-scale modeling of CVD is very computational costly in nature, with the computational cost increasing as the scale decreases. The increased cost with decreased scale is due to the fact that at smaller scales there is more information to be processed. This is why systems which require less complex reactor scale models, such as impinging jet LPCVD, are selected for demonstrating multi-scale modeling of CVD.

In this study impinging jet LPCVD of silicon was selected. This chapter covers the theory, numerical methods and validation of the reactor scale portion of the multi-scale code. The selection of impinging jet LPCVD allows for simplifying assumptions to be made which allow computational power to be saved for use at the feature scale.

2.1 Theory

The dimensions of the reactor scale geometry are such that for pressures above 100 Pa the Knudsen number, Kn will be less than $0.01^{[11]}$, which indicates that the gas can be treated as a continuum. The local Knudsen number is calculated from

$$Kn = \frac{\lambda}{L}, \quad (2-1)$$

where λ is the mean free path and L is the characteristic dimension or length scale. The mean free path is calculated as

$$\lambda = \frac{RT}{\sqrt{2}\pi d^2 N_A P}, \quad (2-2)$$

where R is the gas constant, T is the temperature, d is the mean diameter of the gas molecules, N_A is Avogadro's number and P is the pressure. The characteristic dimension is often chosen to be some dimension such as the channel width in order to create an overall Knudsen number for the entire flow. However, this can be misleading in complex flows where there may be a large range of characteristic length scales for different processes. It is more precise to use a characteristic length scale calculated from the macroscopic gradients at the location of interest^[12],

$$L|_x = \frac{\theta}{\left. \frac{d\theta}{dx} \right|_x}, \quad (2-3)$$

where θ is the variable to be resolved using a continuum based model and x is the location of interest. When the gas of the reactor is treated as a continuum then continuum based equations such as the Navier Stokes equation can be used to describe the fluid flow.

Navier Stokes equation

The continuity and Navier-Stokes equations used in this study assume stationary flow and can be written as

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}), \quad (2-4)$$

and

$$\frac{\partial (\rho \mathbf{v})}{\partial t} = -\nabla \cdot (\rho \mathbf{v}^2) + \nabla \cdot (\mu \nabla \mathbf{v}) - \nabla P + \rho \mathbf{g}, \quad (2-5)$$

where \mathbf{v} is the mass-averaged velocity, P is pressure, ρ is density, μ is viscosity, t is time, and \mathbf{g} is gravitational acceleration.

The density is calculated using

$$\rho = \frac{P_0 m}{RT}, \quad (2-6)$$

where P_0 is the inlet pressure, m is molecular mass of the fluid, R is the gas constant and T is temperature.

Energy Equation

Assuming that the heat of reaction viscous dissipation and Dufour effects are negligible the energy equation can be written as

$$c_p \frac{\partial (\rho T)}{\partial t} = -c_p \nabla \cdot (\rho \mathbf{v} T) + \nabla \cdot (\lambda \nabla T), \quad (2-7)$$

where c_p is the specific heat at constant pressure and λ is the thermal conductivity. Equation (2-7) also assumes that the emission and adsorption of radiation by the gases is negligible; this is a good assumption due to the low gas pressures of LPCVD. The assumptions applied to Equation (2-7) were the same assumptions used by Kleijn et al. ^[19].

Species balance equations

The species equation which accounts for diffusion flux and chemical reactions can be written as

$$\frac{\partial(\rho Y_k)}{\partial t} = -\nabla \cdot (\rho \mathbf{v} Y_k) - \nabla \cdot \mathbf{j}_k + G_k, \quad (2-8)$$

where Y_k represents the mass fraction of species k in the gas phase, \mathbf{j}_k is the diffusion flux and G_k is the generation term in the gas phase^[11]. The diffusion flux can be split into the flux driven by concentration gradients, \mathbf{j}_k^c and the flux driven by temperature, \mathbf{j}_k^T such that

$$\mathbf{j}_k = \mathbf{j}_k^c + \mathbf{j}_k^T, \quad (2-9)$$

where

$$\mathbf{j}_k^c = -\rho D'_k \nabla Y_k - \rho Y_k D'_k \nabla \ln(Y_k) + m Y_k D'_k \sum_{\substack{j=1 \\ j \neq k}}^N \frac{\mathbf{j}_j^c}{m_j D_{kj}}, \quad (2-10)$$

and

$$\mathbf{j}_k^T = -D_k^T \nabla \ln(T), \quad (2-11)$$

where m is the average mole mass of the mixture, T is the temperature, D_{kj} is the binary diffusion coefficient, D'_k is an effective diffusion coefficient and D_k^T is the thermal diffusion coefficient. The diffusive mass flux defined in Equation (2-10) requires that the diffusive fluxes be known. These fluxes can be calculated by solving a series of Stefan-Maxwell equations iteratively. An alternative to this approach is to use Wilke's^[13] approximation,

$$\mathbf{j}_k^c = -\rho D_k^{eff} \nabla Y_k, \quad (2-12)$$

where D_k^{eff} is a different effective diffusion coefficient than D'_k with

$$D_k^{eff} = (1 - f_k) \left(\sum_{\substack{j=1 \\ j \neq k}}^N \frac{f_j}{D_{kj}} \right)^{-1}, \quad (2-13)$$

where f is the species mole fraction. This approximation is most valid for dilute species ($f_k, Y_k \ll 1.0$). The thermal diffusion coefficients are given by

$$D_k^T = \sum_{\substack{j=1 \\ j \neq k}}^N \rho Y_k Y_j D_{kj} a_{kj}, \quad (2-14)$$

where a_{kj} is the thermal diffusion factor. The thermal diffusion factor is dependant on the gas concentrations and temperature and is either found using fitting relations or predicted from kinetic theory. The details of how these coefficients are calculated is shown in Appendix A.

The generation term G_k is dependent on the summation of the reactions producing species k . Some of these reactions are reversible and will consume species k instead of producing it. The resulting equation is

$$G_k = m_k \sum_{i=1}^N v_{ki} R_i^g, \quad (2-15)$$

where m is the molecular mass, v is the net stoichiometric coefficient, R is the net molar reaction rate and N is the total number of chemical reactions that produce species k .

The net gas phase reaction rate for reaction i , R_i^g , is given by Kleijn^[11]

$$R_i^g = k_{i,f}^g \prod_{\text{reactants}} \left(\frac{P f_j}{RT} \right)^{v_{ji}'} - k_{i,b}^g \prod_{\text{products}} \left(\frac{P f_j}{RT} \right)^{v_{ji}''}, \quad (2-16)$$

where $k_{i,f}^g$ is the forward reaction rate constant of reaction i , $k_{i,b}^g$ is the reverse reaction rate constant, f_j is the mole fraction of species j , R is the gas constant, v_{ji}^{\prime} is the forward stoichiometric coefficient for species j in reaction i and $v_{ji}^{\prime\prime}$ is the reverse stoichiometric coefficient for species j in reaction i .

The reaction rate constants, activation energies and thermo-chemical data needed to describe both the surface and gases phase reactions can be obtained from either experimental data or classical

theoretical models^[1]. If an atomic scale model is linked with the reactor scale model this data can also be obtained through the use of density function theory and *ab initio computational* chemistry calculations^[14].

While gas phase reactions will be important for future work, in this study the gas-phase reactions are assumed to be negligible. This is generally a good assumption for LPCVD^[11] and the impinging jet geometry selected further reduces the impact of gas-phase reactions. Furthermore, with the focus of this study being the coupling between reactor and feature scales the additional computational cost associated with including more complex chemistry was not justified.

2.2 Numerical Method

The transport equations for the reactor scale were solved using STREAM^[15], a finite volume based in-house code. In order to solve the transport of the scalar properties the local velocity field must first be solved. STREAM accomplishes this by applying the SIMPLE^[16] algorithm to the momentum and continuity equations. STREAM uses a collocated grid. Checker board oscillations are avoided by using the collocated approach of Rhie and Chow^[17].

2.2.1 Finite Volume Method

The steady state form of Equations (2-5), (2-7) and (2-8) can be written in general as

$$\nabla \cdot (\rho \mathbf{v} \phi) = \nabla \cdot (\Gamma_\phi \nabla \phi) + S_\phi, \quad (2-17)$$

where ϕ is either \mathbf{v} , T , or Y . Integrating Equation (2-17) over a control volume using Gauss Theorem gives

$$\int_A \mathbf{n} \cdot (\rho \mathbf{v} \phi) dA = \int_A \mathbf{n} \cdot (\Gamma_\phi \nabla \phi) dA + \int_{CV} S_\phi dV, \quad (2-18)$$

Applying Equation (2-18) to the control volume of node P in Figure 2-1 gives

$$\left[(\rho u \phi - \Gamma_\phi \nabla \phi)_e - (\rho u \phi - \Gamma_\phi \nabla \phi)_w \right] \Delta y + \left[(\rho v \phi - \Gamma_\phi \nabla \phi)_n - (\rho v \phi - \Gamma_\phi \nabla \phi)_s \right] \Delta x = \bar{S}_\phi \Delta V \quad (2-19)$$

where \bar{S} is the average of source S over the control volume. The velocities u , and v in Equation (2-19) are the advecting velocities. In the SIMPLE algorithm these velocities are assumed to be known either from the results of a previous iteration or an initial guess. In order to discretise Equation (2-19) the values for u , v , and Γ along with the gradient of ϕ must be known at the cell faces. Central differencing, a linear approximation method is used to find these values. Using the east face as an example the results of central differencing are

$$(\rho u)_e = \frac{(\rho u)_P + (\rho u)_E}{2}, \quad (\Gamma)_e = \frac{(\Gamma)_P + (\Gamma)_E}{2}, \quad (2-20)$$

and

$$\left(\frac{d\phi}{dx} \right)_e = \frac{\phi_E - \phi_P}{\Delta x_{PE}}. \quad (2-21)$$

The discrete form of Equation (2-19) is shortened by defining variables C and D to represent the convective mass flux per unit area and the diffusion conductance at the cell faces respectively. Where at the east face

$$C_e = (\rho u)_e \Delta y \quad (2-22)$$

and

$$D_e = \frac{\Gamma_e}{\Delta x_{PE}} \Delta y. \quad (2-23)$$

C and D at the other faces are defined in the same way. Applying these substitutions to Equation (2-19) gives

$$\begin{aligned} C_e \phi_e - D_e (\phi_E - \phi_P) - C_w \phi_w + D_w (\phi_P - \phi_W) \\ + C_n \phi_n - D_n (\phi_N - \phi_P) - C_s \phi_s + D_s (\phi_P - \phi_S) = \bar{S}_\phi \Delta V \end{aligned} \quad (2-24)$$

The Equation (2-24) is not completely discretised, the face values for the advected general property ϕ are still unknown. The scheme used in this study to find most of these values was the quadratic upwind differencing scheme, QUICK.

2.2.2 Quick Scheme

On a uniform mesh the QUICK scheme gives a third order accurate scheme in terms of truncation error^[18]. This is because it uses three nodes to interpolate for cell face values. Figure 2-1 shows the grid points surrounding a node, which are used to calculate its face values. The QUICK scheme also includes upwinding so that the nodes used in the quadratic interpolation of the face value change depending on the direction of convective mass flow across the face. For example the east face value for ϕ is either

$$\phi_e = \frac{6}{8}\phi_P + \frac{3}{8}\phi_E - \frac{1}{8}\phi_W, \tag{2-25}$$

when C_e is greater than zero or

$$\phi_e = \frac{6}{8}\phi_E + \frac{3}{8}\phi_P - \frac{1}{8}\phi_{EE}, \tag{2-26}$$

when C_e is less than zero. The values for the west, north and south values are similarly defined and dependant on their respective fluxes.

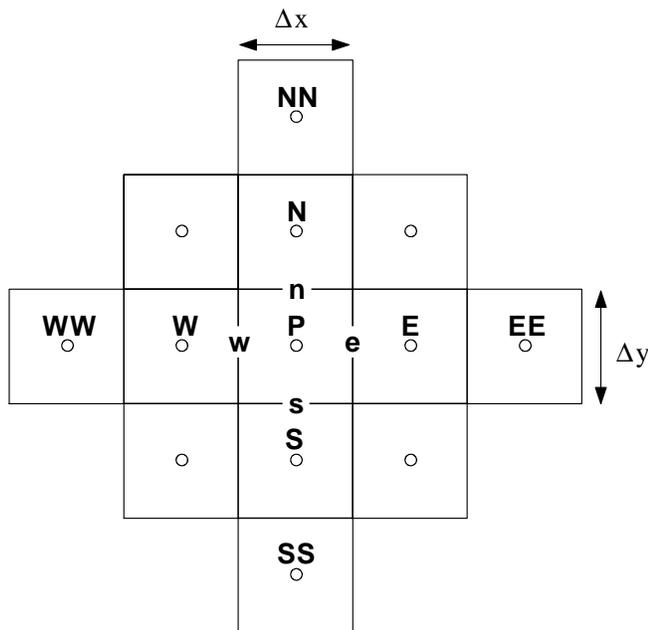


Figure 2-1 Labeling of cells for QUICK scheme.

Due to the convective mass flow dependence of the face values resulting in two possible equations for every face flux Equations (2-25) and (2-26) can not be substituted directly into Equation (2-24).

Instead an intermediate step introduces a flag variable α at each face which is either 0 or 1 depending on the sign of C . This means that at the east face $\alpha_e = 1$ for $C_e > 0$ and $\alpha_e = 0$ for $C_e < 0$.

After the introduction of the variable α the face value equations resulting from the QUICK scheme can be substituted into Equation (2-24). The resulting equation written in the standard form for discretised equations is

$$a_P \phi_P = a_W \phi_W + a_E \phi_E + a_{WW} \phi_{WW} + a_{EE} \phi_{EE} + a_S \phi_S + a_N \phi_N + a_{SS} \phi_{SS} + a_{NN} \phi_{NN}, \quad (2-27)$$

where the values of the coefficients for the right hand side are defined in Table 2-1.

Table 2-1 Coefficients resulting from QUICK scheme.

a_W	$D_w + \frac{6}{8} \alpha_w C_w + \frac{1}{8} \alpha_e C_e + \frac{3}{8} (1 - \alpha_w) C_w$
a_E	$D_e - \frac{3}{8} \alpha_e C_e - \frac{6}{8} (1 - \alpha_e) C_e - \frac{1}{8} (1 - \alpha_w) C_w$
a_{WW}	$-\frac{1}{8} \alpha_w C_w$
a_{EE}	$\frac{1}{8} (1 - \alpha_e) C_e$
a_S	$D_s + \frac{6}{8} \alpha_s C_s + \frac{1}{8} \alpha_n C_n + \frac{3}{8} (1 - \alpha_s) C_s$
a_N	$D_n - \frac{3}{8} \alpha_n C_n - \frac{6}{8} (1 - \alpha_n) C_n - \frac{1}{8} (1 - \alpha_s) C_s$
a_{SS}	$-\frac{1}{8} \alpha_s C_s$

a_{NN}	$\frac{1}{8}(1 - \alpha_n)C_n$
----------	--------------------------------

The value of a_p in Equation (2-27) is calculated as

$$a_p = a_w + a_E + a_{WW} + a_{EE} + a_S + a_N + a_{SS} + a_{NN} + (C_e - C_w) + (C_n - C_s). \quad (2-28)$$

2.2.3 Upwind differencing

For the species equation upwind differencing was used instead of the QUICK scheme. This was due to some instability problems associated with the QUICK scheme leading to non-physical negative values for species concentrations. The coefficients for Equation (2-27) which result from the use of the upwind differencing scheme instead of QUICK are shown in Table 2-2. The upwind differencing scheme is only a first order scheme so the coefficients of the extended nodes EE, WW, SS, and NN are zero.

Table 2-2 Coefficients resulting from upwind differencing.

a_w	$D_w + \max(C_w, 0.0)$
a_E	$D_e + \max(0.0, -C_e)$
a_S	$D_s + \max(C_s, 0.0)$
a_N	$D_n + \max(0.0, -C_n)$

2.3 Boundary Conditions

The inflow boundary condition is prescribed by a flow rate, Q in sccm (standard cubic centimetres per minutes), where the standard temperature is $T^0 = 273.15\text{K}$ and the standard pressure is $P^0 = 1 \text{ atm}$.

The inlet pressure, P_{in} and temperature, T_{in} are also prescribed. Based on this information the velocity normal to the inlet is calculated as

$$v_{in} = \frac{10^{-6}}{60} \frac{P^0}{T^0} \frac{T_{in}}{P_{in}} \frac{1}{A_{in}} Q, \quad (2-29)$$

where A_{in} is the inlet area. The velocity component parallel to the inlet is set to zero. The species concentrations Y_k at the inlet are fixed and the species diffusion set to zero. This allows the mass flow of species into the reactor to be controlled.

The species flux at non-reacting walls is also set to zero. This is accomplished by setting

$$\mathbf{n} \cdot (\mathbf{j}_k^C + \mathbf{j}_k^T) = 0, \quad (2-30)$$

where \mathbf{n} is the wall normal. The no-slip condition is also applied to these walls so that $\mathbf{v} = 0$. The walls are also treated as isothermal with their temperatures prescribed.

In the test cases presented in this study only the substrate is a reacting surface. This is due to the substrate being the only surface hot enough for deposition to occur. The temperature of the substrate is fixed at $T = T_s$. Assuming an isothermal reacting surface is usually a good assumption because the reactors are designed to facilitate this.

The velocity component tangential to the substrate surface assumes the no-slip condition similar to the non-reacting walls. However, the loss of mass from the gas phase due to the surface reaction causes the normal velocity to be non-zero. The normal component of the wall velocity is described by the equation

$$\mathbf{n} \cdot \mathbf{v} = \frac{1}{\rho} \sum_{l=1}^L m_l \sum_{i=1}^N \nu_{li} R_i^S, \quad (2-31)$$

where L is the total number of species, N is the number of surface reactions, ν_{li} is the stoichiometric coefficient and R_i^S is the surface reaction rate. The species concentrations are also affected by the surface reaction. The net total mass flux of species k normal to the wafer is described by

$$\mathbf{n} \cdot (\rho Y_k \mathbf{v} + \mathbf{j}_k^C + \mathbf{j}_k^T) = m_k \sum_{i=1}^N \nu_{ki} R_i^S. \quad (2-32)$$

This equation can be rearranged to give

$$\frac{dY_k}{dn} = \frac{1}{\rho D_k^{eff}} \left(-m_k \sum_{i=1}^N \nu_{ki} R_i^S + \mathbf{n} \cdot (\mathbf{j}_k^T + \rho Y_k \mathbf{v}) \right), \quad (2-33)$$

which provides a boundary condition on the wafer surface for the species equation.

Surface chemistry models are more complex than gas phase reactions because they have to model the adsorption, migration, accommodation, and desorption of molecules along with the chemical reaction.

Kleijn *et al.*^[1] listed three approaches used to model surface reactions. These are:

- a small number of lumped surface processes, with fitted lumped reaction constants
- reactive sticking factors
- a large number of elementary, reversible surface processes

The surface reactions in this study are described with a combination of lumped surface processes and reactive sticking factors. Lumped surface processes are used at the reactor scale while reactive sticking factors are used at the feature scale. Appendix B shows the details of the chemical reaction kinetics used in this study.

The outlet flow was assumed to be fully developed which results in the gradients normal to the exit for velocity, temperature and species concentration being zero.

$$\frac{\partial \rho \mathbf{v}}{\partial n} = 0, \quad \frac{\partial T}{\partial n} = 0, \quad \frac{\partial f}{\partial n} = 0. \quad (2-34)$$

These were the boundary conditions used by Kleijn *et al.*^[1]. In order to accelerate convergence the exit velocity was multiplied by the ratio of inlet to outlet mass fluxes. Similarly the species mole fractions are multiplied by the ratio of the number of particles entering the reactor to those leaving the reactor. This ratio also took into consideration the surface reaction. In this way both global mass and species conservation are ensured.

2.4 Geometry and Grid

The reactor scale model consisted of an impinging jet type reactor. While the make up of the inlet gas and the temperatures of the reactor varied for different test cases, the geometry remained constant.

The inlet sheath was 0.05 m in diameter, and was 0.115 m long so that it came within 0.025 m of the susceptor. The reactor scale chamber was 0.42 m in diameter, while the cylinder on which the wafer was positioned was 0.32 m in diameter. The wafer itself was 0.125 m in diameter and the distance

between the top of the reactor chamber and the wafer was 0.14 m. These geometries were selected to match those used by Kleijn *et al.*^[19]. However, there was a major difference in the inlet used. Kleijn *et al.*^[19] had their in flow coming from the outer radius and being channeled into the inlet sheath. The inclusion of this geometry would have greatly increased the complexity of the grid generation and the dimensions were not provided. For these reasons the inlet condition was simplified so that the inlet to the solution domain was moved to the top of the inlet sheath.

The grid was also designed to be similar to that of by Kleijn *et al.*^[19]. The axisymmetric grid used throughout this study consists of 50 cells in the z direction and 70 cells along the radius. The grid was skewed, as shown in Figure 2-2, to attain higher resolution in regions of interest, such as near the wafer surface.

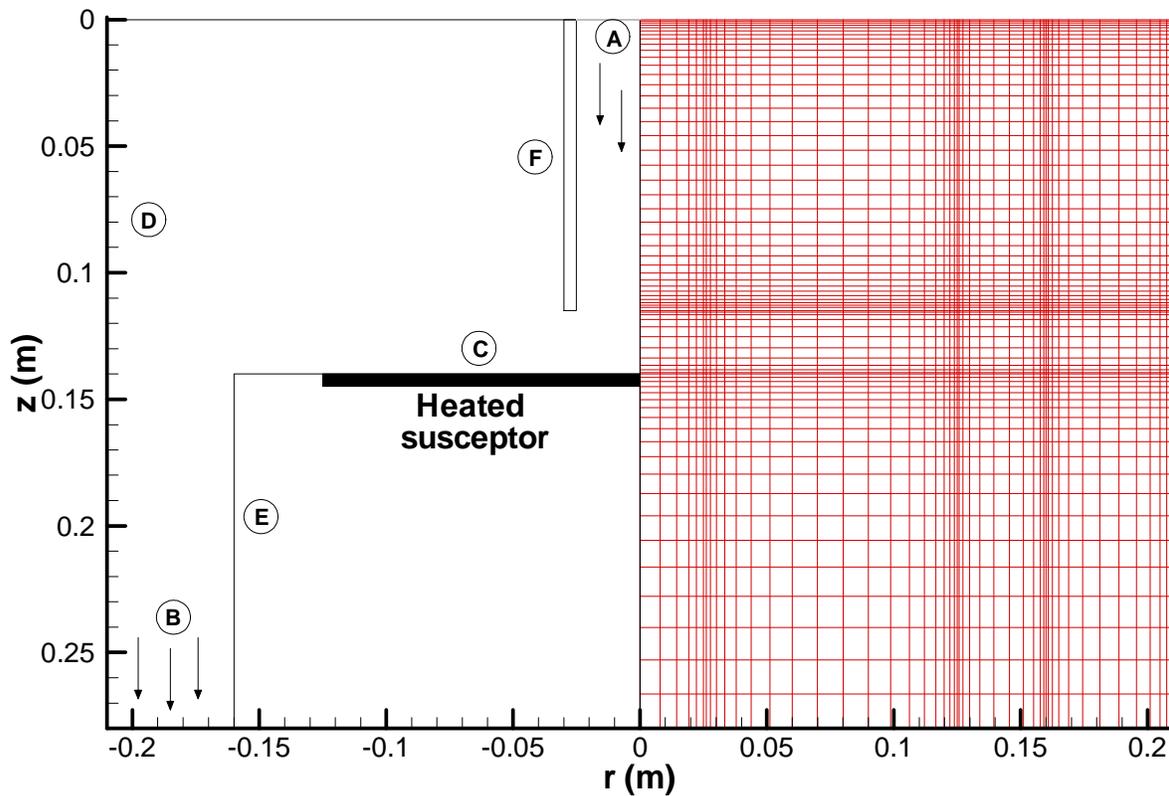


Figure 2-2 Geometry and grid for reactor scale. [(A) Inlet, (B) Outlet, (C) Reacting Surface, (D) Non-reacting wall ($T=T^0$), (E) Non-reacting wall ($dT/dx=0$), (F) Non-reacting wall]

2.5 Validation

Prior to using STREAM^[15] to model chemical vapor deposition major modifications had to be made to the code. These modifications included adding a subroutine to handle the species balance equations and another to calculate the temperature dependant transport properties such as C_p , D_k^{eff} , and D_k^T . In order to ensure that these modifications to the code were working properly two test cases from Kleijn *et al.*^[19] were reproduced using STREAM. The first validation test case was for LPCVD of silicon from a hydrogen silane gas mixture. The inlet flow rate was 0.2 slm, the wafer temperature was 1000 K, the pressure was 1 torr, and the inlet gas and wall temperatures were 300 K. The composition of the inlet gas was 90 mole percent (m/o) hydrogen and 10 m/o silane. This test case was chosen because contours for the temperature, and silane concentration were provided along with the surface growth rate profiles. Results with and without thermal diffusion were also available.

Figure 2-3 shows contour plots of the silane concentrations predicted by the reactor scale model with and without thermal diffusion taken into account. The results from the study of Kleijn *et al.*^[19] are also plotted on this figure for comparison purposes. Through examination of Figure 2-3 it can be seen that there is good agreement between the concentration results of the current study and those of Kleijn *et al.*^[19] independent of whether thermal diffusion is included. Both the shapes of the contours and value of the contours agree well especially in the most important region near the wafer surface.

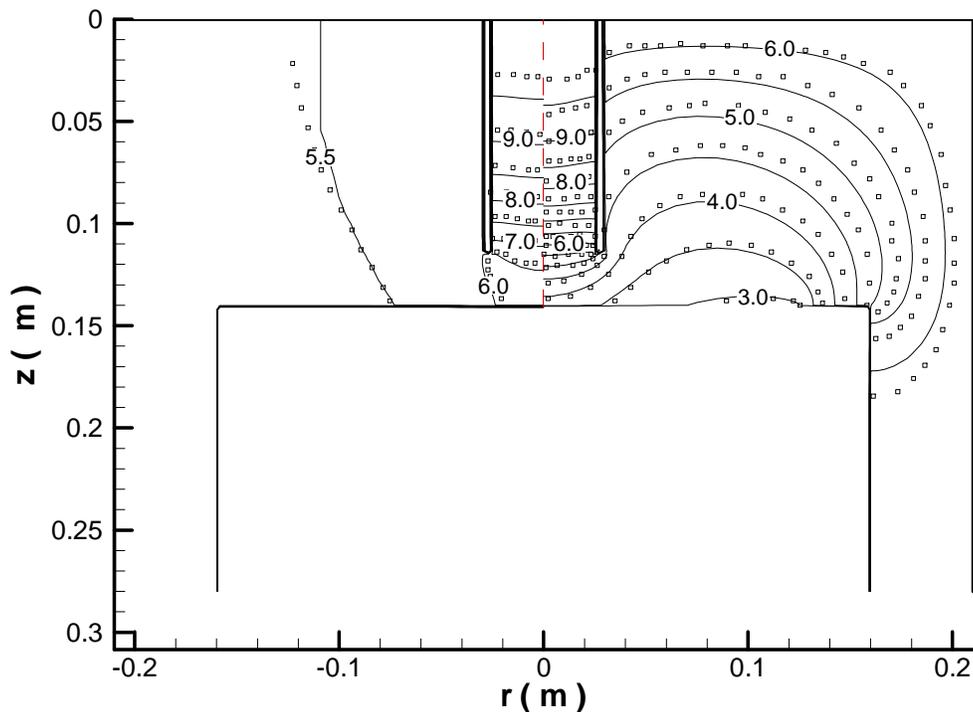


Figure 2-3 Contour plots for the silane mole fraction. Left: without thermal diffusion. Right: with thermal diffusion. (□) Digitized results from Kleijn *et al.* ^[19].

In addition to comparing the concentration contour plots the growth rate profiles produced by STREAM were also compared with those reported by Kleijn *et al.* ^[19]. Figure 2-4 shows the results of both studies with and without thermal diffusion. The results of the two codes were in good agreement with the difference between the results produced by the current study and those of Kleijn *et al.* ^[19] was less than 2%. The fact that the large impact of the thermal diffusion on the results was accurately taken into account also confirmed that the portions of the code accounting for thermal diffusion and variable fluid properties were working properly.

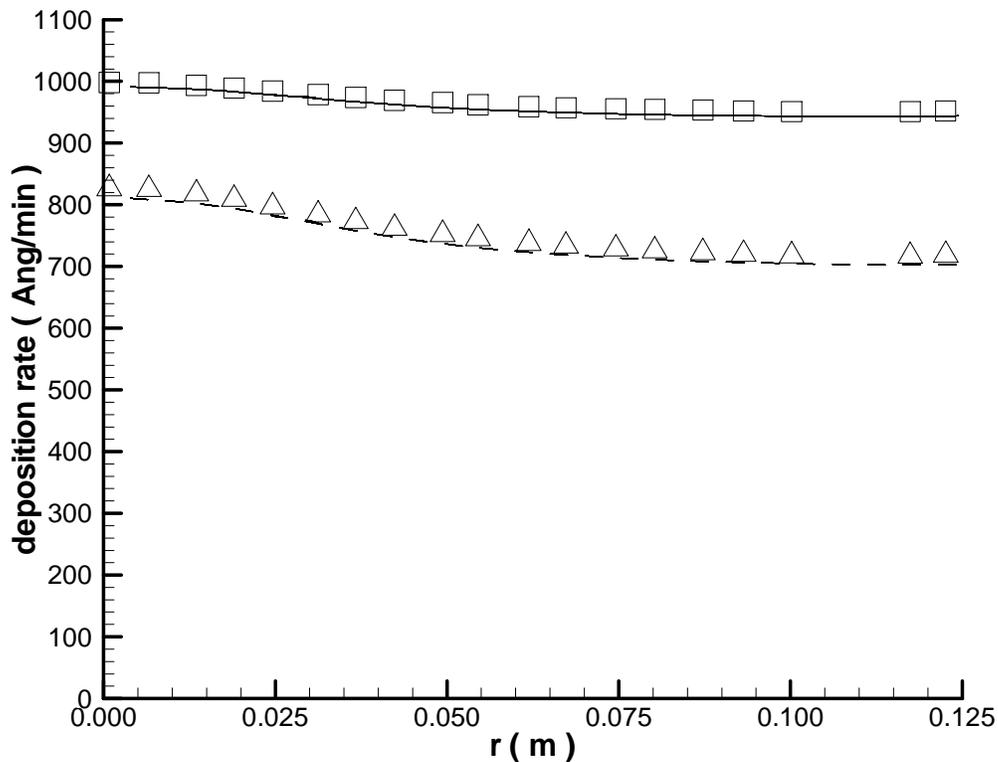


Figure 2-4 Comparison of growth rates for hydrogen reduction of silane with hydrogen as the carrier gas. No thermal diffusion: current code (—), Kleijn 1989^[19] (□). Thermal diffusion: current code (---), Kleijn 1989^[19] (Δ).

In anticipation of the desire to study more complex chemistries than the two species model covered in the first validation test case a second test case was performed where Nitrogen was introduced as third species. The introduction of nitrogen into the carrier gas meant that effects of multi-component diffusion had to be taken into account.

The inlet flow rate for this second test case was 0.2 slm, the wafer temperature was 1000 K, the pressure was 1 torr, and the inlet gas temperature and wall temperature were 300 K. The composition of the inlet gas was 45 mole percent (m/o) hydrogen, 45 mole percent (m/o) nitrogen and 10 m/o silane.

Despite being inert, the introduction of nitrogen had an impact on the surface growth rate profile near the edge of the wafer dropping the growth rate an additional 50 Ang/min. Figure 2-5 shows a comparison of the growth rate profile results of STREAM and the digitized results of Kleijn *et al.*^[19]. There was a difference of approximately 1.5% in the results. This level of agreement was seen to be acceptable for the purposes of this study.

In addition to using this test case to compare with the results of previous studies it was also used to test whether the grid was adequately refined. This was done by increasing the grid resolution by a factor of 2 and comparing the results. The results generated by the finer grid are plot in Figure 2-5 along with the results of the coarser grid. Comparison of the results generated by the two different grids showed that increasing the grid density by a factor of 2 in each direction resulted in a difference of less than 1% in the growth results.

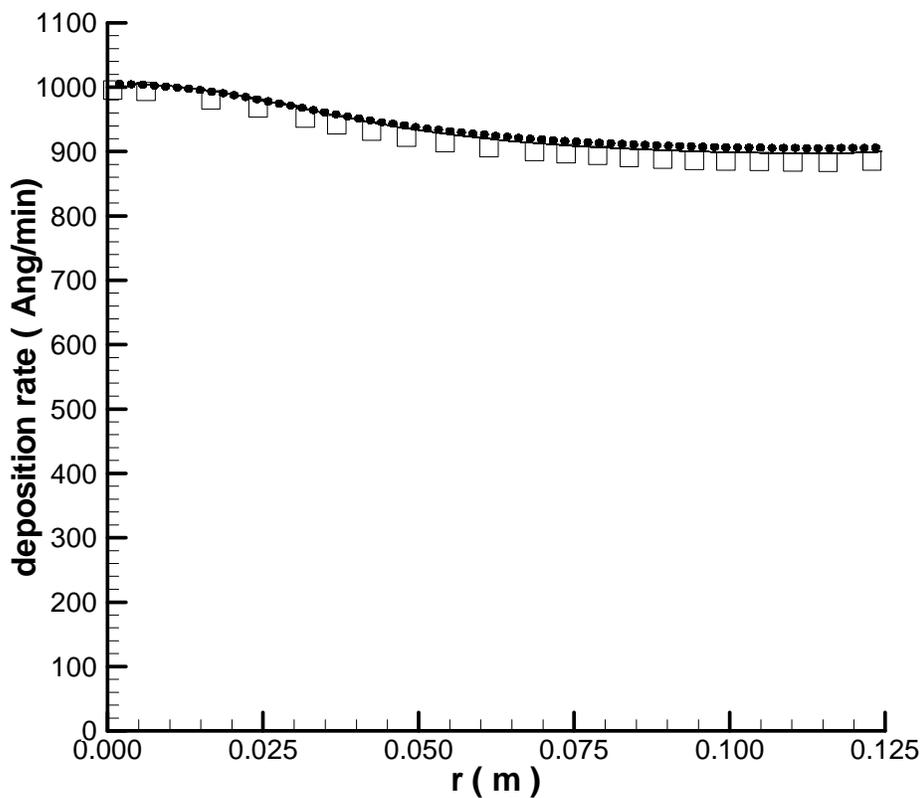


Figure 2-5 Comparison of growth rates for hydrogen reduction of silane with a hydrogen-nitrogen mixture as the carrier gas. Current code course mesh (—), Current code fine mesh (· · ·), Kleijn 1989 (□).

Chapter 3

Feature Scale

Feature scale modeling of CVD is an important component of the multi-scale modeling of CVD. This is because it is at the feature scale that conformality of the deposition can be observed. Conformality is the ability of the film to exactly reproduce the surface topography on which it is deposited. Having good conformality is important because poor conformality results in the formation of voids which can cause failure of the final product.

The first attempts at modeling the feature scale were simple line-of-sight models^[20-24]. These simple line of sight models assumed a sticking factor of 1.0 with no re-emission. These initial attempts were followed by the development of the Monte Carlo (MC)^[25-32], Knudsen diffusion (KD)^[33-41] and modified line-of-sight models^[42-49].

MC methods arrive at deposition predictions by tracking the trajectories of a statistically representative sample of molecules. These methods provide the best physical understanding of the process and are capable of handling gas phase collisions and complex gas phase chemistry. The drawback of MC methods is their relatively high computational cost especially for depositions with low sticking factors^[46].

KD methods treat the flow within the features as a continuum with the dominate transport being Knudsen diffusion. KD methods are one-dimensional assuming a constant feature cross-section and are unable to account for particles bouncing back from the bottom of the feature^[19]. For high aspect ratio features the impact of the bottom of the feature becomes less significant and there is less variation in cross-section. The major benefit of KD methods is low computational cost, this is why Lankhorst *et al.*^[50] recently used KD methods in their multi-scale simulation of atomic layer deposition (ALD).

Modified line of sight models such as the ballistic transport and reaction (BTRM) and the simplified hemispherical vapor source (HVS) models enabling the modeling of non-unity sticking factors. This is accomplished through successive re-emission and re-deposition steps. With each successive re-

emission fewer and fewer particles are re-deposited allowing a converged solution to be obtained. These improved line of sight models are significantly less computationally costly than MC methods and are able to handle more complex geometries than KD methods^[46]. These improved line-of-sight models are, however, limited to the ballistic transport flow regime due to their inability to model particle-particle collisions.

Recently a new model called the Kinetic Transport and Reaction Model (KTRM)^[51-57] has been developed. This model uses a Galerkin method to simulate the transient linear Boltzmann equation. Simulation of the transient linear Boltzmann equation accounts for particle-particle collision and allows KTRM to predict the kinetic density. Since KTRM allows particle-particle collisions it is valid over a larger range of Knudsen numbers than the improved line-of-sight models, ($Kn \leq 1.0$). The ability to predict the kinetic density allows KTRM to model transient processes such as ALD. The KTRM method is limited to flows with a dominant gas phase species^[57]. The large computational cost of KTRM has also lead the developers of the method to recommend that BTRM be used for large Knudsen number flows^[57].

State of the art feature scale models are able to simulate complex three dimensional geometries^[51,54,58-60] and processes such as ALD^[61-63] where transients in concentration are more important. However, the feature scale model developed in the current study is a 2-D ballistic transport model called the BTLFSF (Ballistic Transport with Local Sticking Factors) model^[64]. The reason for choosing a 2-D ballistic transport based model was to keep the computational cost of the feature scale down. This is because in order to achieve multi-scale solutions within a reasonable time having an efficient feature scale model is very important. This is highlighted by Lankhorst *et al.* in their recent work^[50] where they used a 1-dimensional Knudsen Diffusion Model (KDM) for multi-scale modeling of CVD and ALD.

3.1 Theory

The modified line of sight model used to model the feature scale in this study was the BTLFSF^[64]. The BTLFSF model was developed for this study based on concepts taken from three different models. The three models are the HVS model developed by Yun and Rhee^[65], BTRM from Cale *et al.*^[42,43,66] and the MC method used by Rodgers and Jensen^[3]. The goal of developing this new model was to develop a simple but accurate ballistic transport based feature scale model where the initial deposition

arrives from a source plane. The importance of the deposition arriving from a source plane will be discussed in the section on coupling. The idea of using a source plane to calculate the initial deposition was taken from Rodgers and Jensen^[3], while the processes for formulating the initial deposition equation was based on the work of Cale *et al.*^[42,43,66].

While the solution algorithm of the BTLSF model was based on the HVS^[65] model, there are three major differences between the BTLSF model and the HVS model. These are the treatment of the initial surface deposition, the use of local sticking factors, and the way in which the 'differential transmission probability' is calculated. The initial deposition flux at point x in the HVS model was calculated using the equation^[49],

$$\eta^0(x) = 2K\sqrt{(\cos w_1 - \cos w_2)^2 + (\sin w_2 - \sin w_1)^2}, \quad (3-1)$$

where K is a proportionality constant while w_1 and w_2 are angles describing the shadowing of the deposition sight as shown in Figure 3-1. This equation is based on particles arriving at the surface from a hemispherical source. While BTRM also uses a hemispherical source volume its equation for initial deposition does not have the same form as that of the HVS model. The BTLSF model is more similar to the BTRM model in its calculation of initial deposition than the HVS model because it uses the same concepts from physical chemistry^[67] as the BTRM model. For this reason the assumptions included in BTRM:

- 'The frequency of collisions between gas phase species is small relative to the frequency of collisions between gas phase species and surfaces; i.e. intra-feature transport is by free molecular flow'^[60].
- 'Deposition occurs by heterogeneous reactions between gas phase species and the evolving film surface'^[60].
- 'The film grows slowly relative to the redistribution of fluxes to the feature surfaces caused by film evolution'^[60].
- 'The open end of the feature is exposed to a gas from which species travel to the surface arriving with well defined flux distributions'^[60].
- 'Reactive sticking factors do not depend on the incident angle or collision history of the impinging molecules'^[60].

are also applicable to the BTLSF model.

In the derivation of the equation for the initial deposition in BTRM spherical co-ordinates are used. This allows the BTRM model to be applied to three dimensional problems. However, in order to keep the computational cost down, only 2-D geometries are considered in this study. One way this reduces

the cost is it allows cylindrical co-ordinates to be used instead of spherical co-ordinates, which greatly simplifies the description of the shadowing affects of the feature geometry. The perspective of the BTLSF model is also different from that of BTRM, with BTLSF having the initial deposition arriving from a source plane offset from the wafer surface similar to the source plane used in the study done by Rodgers and Jensen^[3]. The use of a source plane instead of a hemisphere source allows for better linking between scales in multi-scale simulations with the feature scale source plane being made up of cell faces from the reactor or meso scales. In this way the source plane acts as the interface between the feature scale and the reactor or meso scale. Gradients in the flux crossing the source plane from the larger scale can also be reproduced using this method.

Since the geometry considered in this study is two dimensional the source plane is discretized as line sources. The frequency with which particles leaving a line source at \mathbf{x}' collide with the area ds can then be written as

$$Z_{coll}(\mathbf{x}' \rightarrow \mathbf{x}) = \frac{n(\mathbf{x}')}{2\pi} \left[\int_0^\infty p(v) v dv \int_{\theta_a}^{\theta_b} \cos \theta d\theta \right], \quad (3-2)$$

where $n(\mathbf{x}')$ is the number of particles per unit volume or number density at position \mathbf{x}' , $p(v)$ is the probability density function (PDF) for the particle speed, θ_a and θ_b are angles defined in Figure 3-1 and v is the particle speed. Note that all angles in this study are defined as counter clockwise positive.

The PDF for a Maxwellian point source is

$$p(v) = \left(\frac{m}{2RT\pi} \right)^{3/2} 4\pi v^2 e^{-\frac{mv^2}{2RT}}, \quad (3-3)$$

while that of a Maxwellian line source is

$$p(v) = \frac{m}{2RT} v e^{-\frac{mv^2}{2RT}}, \quad (3-4)$$

where R is the universal gas constant, T is the temperature, and m is the molecular weight mass. The difference in their probability distributions causes the mean speed of a line source to vary from that of the point source, so that the mean speed for particles leaving a Maxwellian line source is

$$\int_0^{\infty} p(v)v dv = \int_0^{\infty} \frac{m}{RT} v^2 e^{\frac{-mv^2}{2RT}} dv = \frac{\pi}{4} \left(\frac{8RT}{\pi m} \right)^{\frac{1}{2}}. \quad (3-5)$$

Substitution into Equation (3-2) gives

$$Z_{coll}(x' \rightarrow x) = \frac{n(x')}{4} \left[\left(\frac{8RT}{\pi m} \right)^{\frac{1}{2}} \frac{(\sin \theta_b - \sin \theta_a)}{2} \right]. \quad (3-6)$$

If we assume that the number density, $n(x')$ of particles along the source plane is constant, then n is no longer a function of x' and the first few terms of Equation (3-6) can be grouped as the flux of species k particles across the source plane

$$F_k = \frac{n_k}{4} \left(\frac{8RT}{\pi m} \right)^{\frac{1}{2}}, \quad (3-7)$$

where the subscript k has been added to indicate the flux of an individual species, k .

Equation (3-6) is only the contribution from a single discrete line source on the source plane. In order to determine the total flux of species k particles arriving at point x from the source plane Equation (3-6) must be integrated across the source plane. However, the entire source plane is not necessarily visible as seen in Figure 3-1. The surface area, ds shown in Figure 3-1 is only visible to the portion of the source plane highlighted by a thicker line segment. This area is bounded by the angles w_1 and w_2 defined in Figure 3-1. Taking shadowing into account the initial surface flux at point x of species k is

$$\eta_k^0(x) = \frac{F_k}{2ds} \int_L^R (\sin \theta_b - \sin \theta_a) dx', \quad (3-8)$$

where

$$L = x - \Delta s_x / 2 + \tan(w_1)(y' - (y - \Delta s_y / 2)),$$

and

$$R = x + \Delta s_x / 2 + \tan(w_2)(y' - (y + \Delta s_y / 2)),$$

where Δs_x and Δs_y are the x and y components of ds . This equation can be solved analytically to give

$$\begin{aligned} \eta_k^0(x) = \frac{F_k}{2ds} & \left[- \left((L-x+\Delta s_x/2)^2 + (y'-y+\Delta s_y/2)^2 \right)^{1/2} \right. \\ & + \left((L-x-\Delta s_x/2)^2 + (y'-y-\Delta s_y/2)^2 \right)^{1/2} \\ & + \left((R-x+\Delta s_x/2)^2 + (y'-y+\Delta s_y/2)^2 \right)^{1/2} \\ & \left. - \left((R-x-\Delta s_x/2)^2 + (y'-y-\Delta s_y/2)^2 \right)^{1/2} \right] \end{aligned} \quad (3-9)$$

Not all of the particles which collide with an area on the surface stick to the surface, a fraction of the particles are re-emitted from the surface. ϵ_k , the sticking factor is defined as the fraction of species k particles which stick when they collided with an area of the surface; the rest of the particles are re-emitted. Some of the particles which are re-emitted from one surface may collide with a second surface. These particles add to the total number of particles colliding with the second surface. Assuming diffuse re-emission, which means that the particles colliding with the surface reach thermal equilibrium with the surface before being re-emitted, the total flux of particles into a section, i of the surface becomes

$$\eta_{k,i}^{Total} = \eta_{k,i}^0 + \sum_{j=1, j \neq i}^M \left[(1 - \epsilon_{k,j}) \eta_{k,j}^{Total} \chi_{j,i} q(j,i) \right], \quad (3-10)$$

where $\chi_{j,i}$ is one if site j is visible from site i and zero otherwise. The subscript j on the sticking factor, ϵ in Equation (3-10) indicates that the sticking factors are calculated locally for each discrete section of the feature surface, M sections in total. Having the factors calculated locally means that the sticking factors vary over the feature surface depending on the flux and temperature distribution. This makes the BTLFSF model different from the HVS model because the HVS model like many models uses a single sticking factor for each species over the entire feature scale.

In this study Equation (3-10) was solved using direct quadrature. This meant that the re-deposition portion of Equation (3-10) was broken up into a series of emission and re-deposition stages such that Equation (3-10) became

$$\eta_{k,i}^{Total} = \eta_{k,i}^0 + \sum_{l=1}^N \left[\sum_{j=1, j \neq i}^M [(1 - \varepsilon_{k,j}) \eta_{k,j}^l \chi_{j,i} q(j,i)] \right], \quad (3-11)$$

where $\eta_{k,j}^l$ was the flux of species k arriving from site j from the l^{th} re-emission stage and N was the total number of stages required to reach convergence.

Another difference between the HVS model and the BTLFSF model is the definition of the 'transmission probability', $q(j,i)$. For the BTLFSF model $q(j,i)$ is defined as

$$q(j,i) = \frac{1}{2} \int_{\theta_{i,j}^L}^{\theta_{i,j}^R} \cos \theta d\theta, \quad (3-12)$$

where $\theta_{i,j}^R$ and $\theta_{i,j}^L$ are shown in Figure 3-2a. This 'transmission probability' is calculated based on the visibility of the emitting surface, j from the receiving surface, i while that of the HVS model is calculated based on the visibility of the receiving surface, i from the emitting surface, j . These two transmission probabilities become identical as the size of the elements ΔS_j and ΔS_i approach zero.

The BTRM model uses a different definition of the 'transmission probability' than both the BTLFSF and HVS models. The BTRM model defines $q(j,i)$ in two dimensions as

$$q(j,i) = \int_{\Delta s_j} \frac{\cos \theta_{i,j} \cos \varphi_{j,i}}{2|P_{i,j}|} ds \approx \frac{\cos \theta_{i,j} \cos \varphi_{j,i}}{2|P_{i,j}|} \Delta s_j, \quad (3-13)$$

where $|P_{i,j}|$ is distance between points i , and j . $\theta_{i,j}$ and $\varphi_{j,i}$ are shown in Figure 3-2b. The difference between the BTRM 'transmission probability' and that of the BTLFSF model is that the BTRM model approximates the plane angle, $d\theta$ as

$$d\theta = \frac{\cos \varphi_{j,i}}{|P_{i,j}|} ds, \quad (3-14)$$

Additionally $|P_{i,j}|$, $\theta_{i,j}$, and $\varphi_{j,i}$ are kept constant across the area ΔS_j . Usually these are good approximations. However, if the distance between the segments $|P_{i,j}|$ becomes smaller than $\Delta S_j/2$

there is a potential for the calculated 'transmission probability' to be greater than 1. Over estimation of the 'transmission probability' in this way can lead to numerical instability when solving Equation (3-10). The BTLFSF definition of 'transmission probability' not only ensures that the individual transmission probabilities are less than one but also that the summation of transmission probabilities over j is less than or equal to 1.

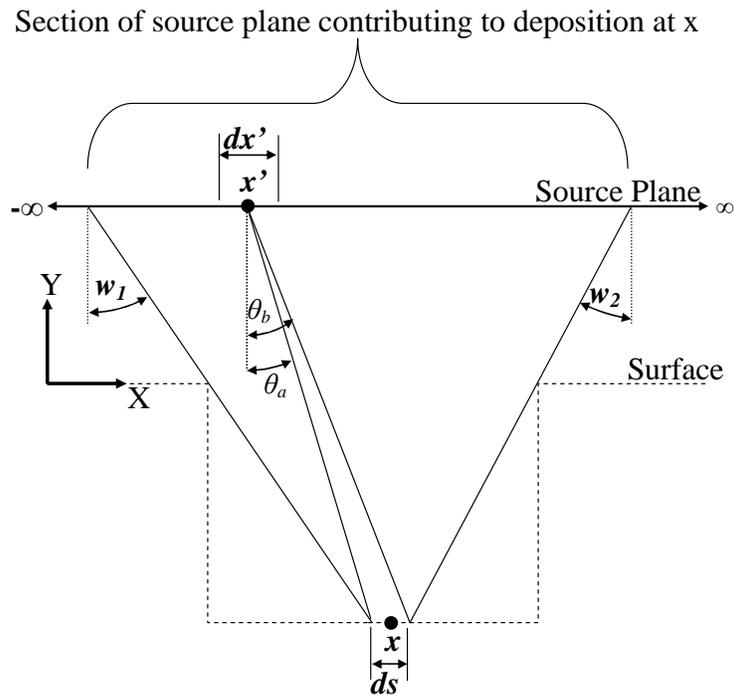


Figure 3-1 Schematic showing the definition of angles θ_a , θ_b , w_1 and w_2 .

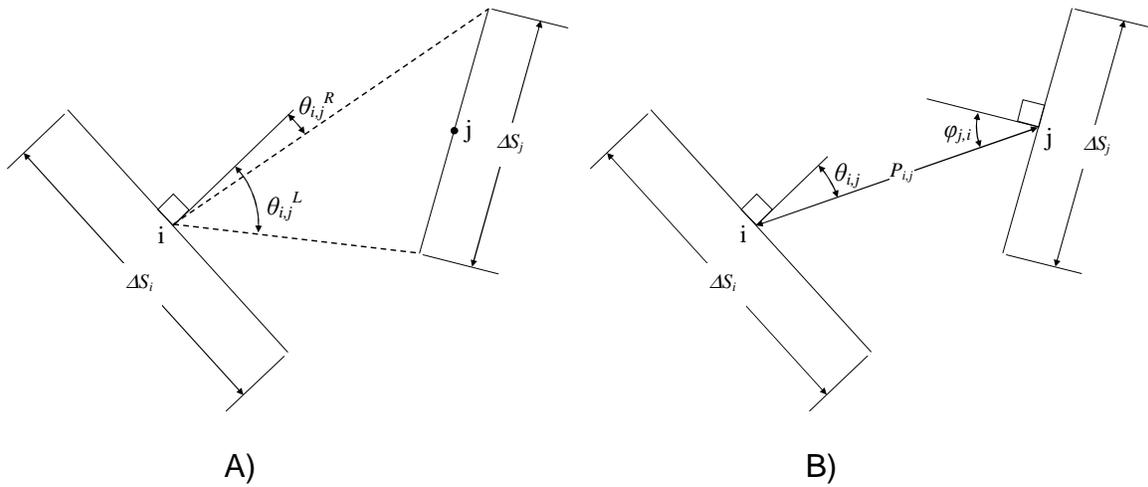


Figure 3-2 Schematics showing the definition of angles used in calculation of 'transmission probability': (A) for BTLSF, Equation (3-12) (B) for BTRM, Equation (3-13).

3.1.1 Surface Tracking

Surface tracking is a major component of feature scale modeling. The surface tracking method used in this study is a string algorithm similar to that of Neureuther *et al.* [68]. String methods can be quite accurate [69] and were seen as the easiest option to implement. The major difficulty with string methods is 'looping' and this is avoided by updating the node spacing every time the surface is advanced and careful control of the time step used. However, if this model were to be extended to the atomic scale a more accurate method would have to be introduced as sting methods are unable to capture nano-structural details [32].

An example of a more complex model which could be used if the work in this study were to be extended to the atomic scale is the level set method of Sethian [70,71]. While much more complex than string methods Level set methods have been shown to be capability of modeling the evolution of faceted crystals and capturing micro-structural details such as grain boundaries [72]. A level set method was first applied to the modeling of deposition by Adalsteinsson and Sethian [73]. These methods are popular because they offer a highly accurate means of tracking complex interface motion free from the inherent looping problem associated with traditional string algorithms [72]. Level set methods are also more easily extended to 3 dimensions than string methods.

3.1.2 Automatically varying time step

As previously mentioned BTLFSF uses direct quadrature to solve for the total surface fluxes. This means that the deposition process is broken down into successive deposition, emission, and re-deposition stages. The total number of stages required to reach the convergence criteria depends on the convergence criteria. The feature scale convergence criteria used in this study is that the maximum value of η^N , the flux contributions from the final deposition stage for each segment had to be less than one ten-thousandth of the maximum η^0 , initial surface fluxes,

$$MAX(\eta^N) < \frac{MAX(\eta^0)}{10000}. \quad (3-15)$$

The number of stages, N required to reach the convergence criteria also varies depending on the width of the feature neck. Initially when the neck of the feature is large the segments on the walls of the feature are far apart. When segments are far apart the transmission probability between them is small. This means that N will also be small. However, when the neck of the feature is almost closed, the wall segments on the walls of either side of the neck are very close together as well as being close to parallel. This means that the transmission probability between these segments will be high. If the sticking factor is low this high transmission probability will result in particles being passed between the segments many times before they escape or stick, resulting in a large N . Figure 3-3 shows how N increases with non-dimensional time due to the closure of an aspect ratio 1.0 feature with a base sticking factor, $\varepsilon_0 = 0.01$. The non-dimensional time shown on the x -axis is the time normalized by the time it takes for the feature to close. Examining Figure 3-3 it can be seen that at $t = 1.0$, N decreases sharply because the feature closes and solution of the surface deposition becomes easier. By making the time step dependent on N it can be automatically decreased as the feature closes. This allows for accurate capturing of the feature's "time of closure" with less time steps. The use of automatically decreasing time increments has been mentioned in the past^[21], however, the details of its implementation have not been discussed. Most likely the time increments were selected based on estimations of the growth rate and width of the feature throat.

Having the time step dependent on N is especially useful for complex geometries where there may be more than one time when the time step size should be reduced to increase accuracy. For example, when the double stepped feature shown in Figure 3-4 is filled using a sticking factor of 0.1 two voids

will form as seen in Figure 3-4a. A small time step must be used each time one of the voids closes in order to accurately capture the locations at which the closures take place. When a fixed time step is used this means a small time step must be used for the entire simulation, which is very computationally costly. A variable time step can be calculated based on,

$$\Delta t = \tau \Delta t_{t=0} = \left[1 - (1 - B) \left(\frac{\frac{N_{t=0}}{N_t} - 1}{\frac{N_{t=0}}{N_{t=1.0}} - 1} \right) \right] \Delta t_{t=0}, \quad (3-16)$$

where B is the desired time step ratio,

$$B = \frac{\Delta t_{t=0}}{\Delta t_{t=1.0}}, \quad (3-17)$$

and $\Delta t_{t=0}$ is the time step at $t=0$. An example of the resulting variation of τ with time when $B=0.1$ can be seen in Figure 3-3. Defining the time step in this way allows for a larger initial time step to be used as seen in Figure 3-4b. This saves a lot of computational time by greatly reducing the number of time steps required.

Since $N_{t=0}$ and $N_{t=1.0}$ are unknown before the simulation is performed they are approximated using Equations 4 and 5,

$$N_{t=0} = 39.28AR\epsilon_0^{-1.978\epsilon_0} \quad (3-18)$$

$$N_{t=1.0} = 12.789AR\epsilon_0^{-0.8143} \quad (3-19)$$

where AR is the aspect ratio of the feature based on the width of feature entrance and the maximum depth of the feature and ϵ_0 is the base sticking factor. These equations were arrived at through a series of parametric studies, which found that $N_{t=0}$ and $N_{t=1.0}$ were mainly dependent on the aspect ratio and base sticking factor.

Varying the time stepping scheme based on N works best for small base sticking factors, $\epsilon_0 \leq 0.1$. This is because $N_{t=1.0}$ is large for small sticking factors resulting in a large smooth variation between $N_{t=0}$ and $N_{t=1.0}$. Simulations with large sticking factors require every few re-deposition steps, even when the feature is about to close. This results in very little variation in N with t , which reduces the effectiveness of using N to control the time step.

The method described here is for use with a single feature. In multi-scale modeling there may be multiple features with different geometries. The time it takes to fill these different features will vary and the number of iterations to reach convergence for each feature will also vary. The same time step must also be used for all the features since it also corresponds to the reactor scale time step. One possible solution would be to use the minimum time step resulting from the different N values of the features. This would ensure the best predications for all the feature closures. However, the computational cost of the additional accuracy would be fairly high. Determining the best balance accuracy and computational cost for time step selection in multi-scale simulations requires further research.

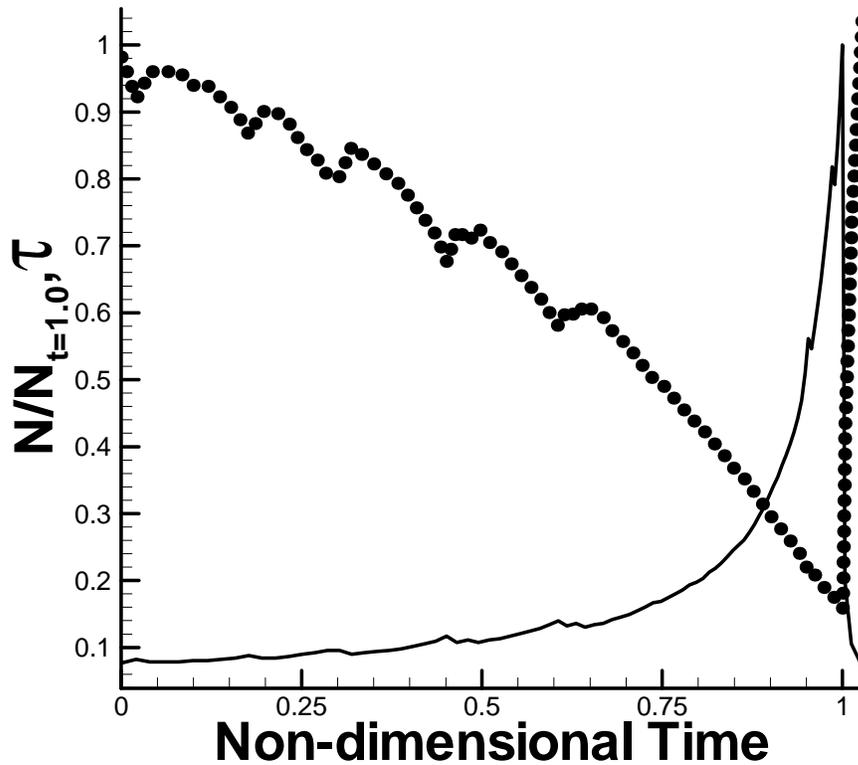


Figure 3-3 Variation of N (—) and τ (····) as feature closes: $AR = 1.0$, $\epsilon_0 = 0.01$.

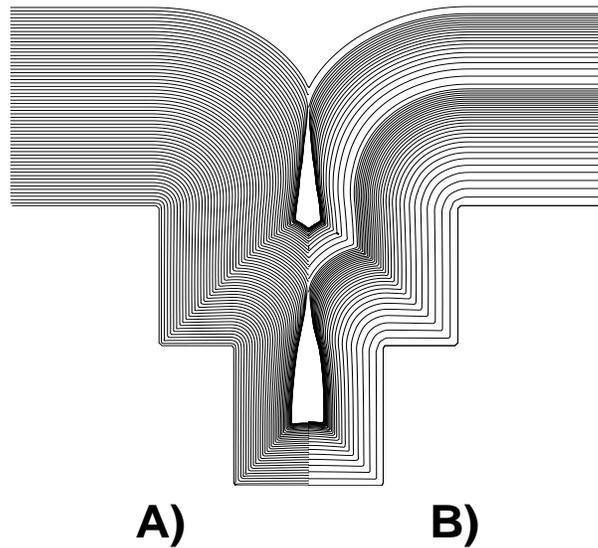


Figure 3-4 Film profiles for a first order deposition reaction in a stepped trench for a sticking of 0.1 using: (A) a fixed time step (B) a variable time step. Every time step is plotted for the variable time step case while every fourth time step is plotted for the fixed time step case.

3.2 Validation

The results of Cale and Raupp^[42] showing the dependence of instantaneous step coverage on the fraction of feature closure were selected for verification of the BTLFSF model. Cale and Raupp defined instantaneous step coverage as a ratio of film deposition rates^[42]. The numerator deposition rate was located on the feature side wall where the wall and base intersect while the denominator was located at a flat surface exterior to the feature.

Cale and Raupp's results were chosen because they cover a range of sticking factors and reflect the change in the flux distribution as the feature closes. The test case consisted of simulating three first order deposition reactions within a $1\mu\text{m}$ by $1\mu\text{m}$ trench. These three simulations had different base sticking factors, which were 0.01, 0.1, and 1.0. The magnitude of the sticking factors has a large impact on the conformality of the deposition, with smaller sticking factors producing more uniform depositions. Large sticking factors resulted in lower instantaneous step coverage with more material being deposited on the surface and entrance of the feature than inside the feature. As the feature closed this caused necking, further reducing the deposition rate within the feature and the instantaneous step coverage. Examining Figure 3-5, which shows the results generated using both

models, it can be seen that the BTLFSF model produced very similar results to those produced by Cale and Raupp^[42] using their pre-BTRM model.

Comparison of the BTLFSF model with the results of Cale and Raupp^[42] only benchmarks the BTLFSF model's results against those of pre-BTRM. Figure 3-6 contrasts the results of the BTLFSF model with a test case performed by Yun and Rhee^[65] using the HVS model and MC methods. This test case was for tungsten LPCVD deposited within a 3.6 aspect ratio trench using a global sticking factor of 0.29. Examining Figure 3-6 it can be seen that the BTLFSF profile more closely resembles the MC results than the HVS model. The main discrepancy between the BTLFSF model results and those generated using MC methods occurred at the entrance to the feature. The MC results predicted that the feature would close faster resulting in a smaller opening than was produced by BTLFSF. The width of the opening predicted by the BTLFSF, however, matched that of HVS.

The node spacing used in these test cases was 0.0001 μm which resulted in 816 initial surface nodes for the HVS comparison test case. The number of surface nodes varied throughout the simulation as the surface area of the feature changed. This node spacing was selected after a grid convergence study, showing that it produced a grid independent solution. The initial time steps used in the simulations were selected so that it would take approximately 200 time steps for the features to close.

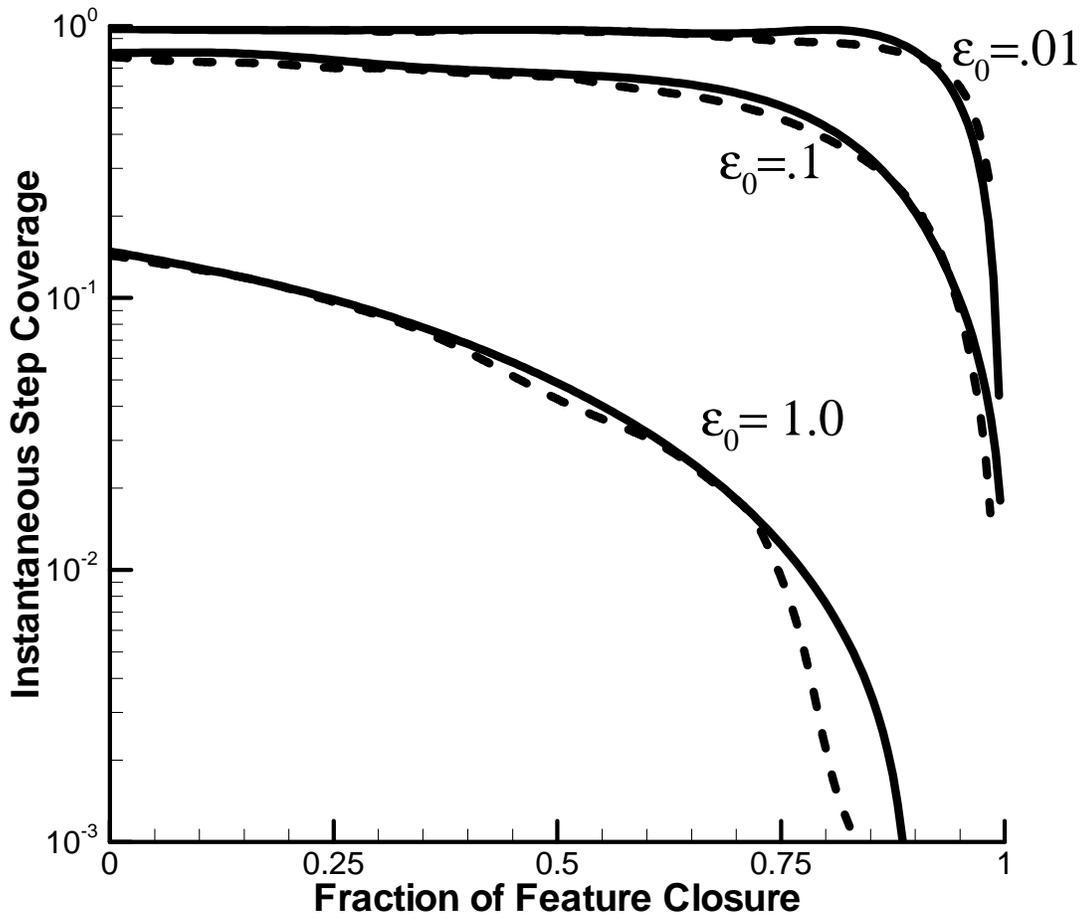


Figure 3-5 Instantaneous step coverage as a function of feature closure for a first order deposition reaction in a $1\mu\text{m}$ by $1\mu\text{m}$ trench. (- - -) Cale and Raupp^[42], (—) BTLFSF.

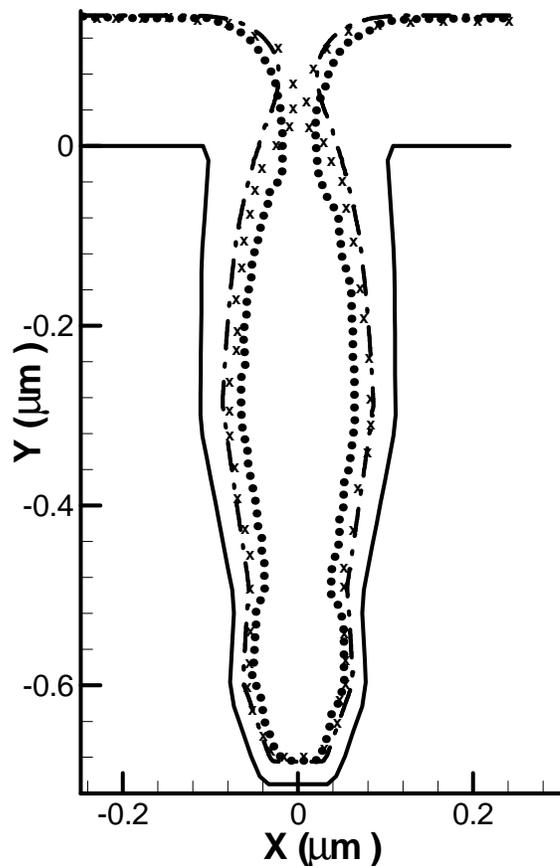


Figure 3-6 Comparison of BTLFSF (— — —) results with the HVS (•••••) and MC (x x x) results of Yun and Rhee^[65]. The solid line indicates the initial profile (——).

3.3 Local Sticking Coefficients

Whenever possible chemical reaction mechanisms should be used to model the reactions on the surface, however, there are many processes for which no chemical reaction mechanisms are available. An alternative to using chemical reaction mechanisms is to use a 'sticking factor'. Reactive sticking factors (RSF) describe the fraction of an impinging flux which will be deposited on the surface. They are calculated based on overall rate expressions or simple chemical models which take into consideration parameters such as the flux of molecules colliding with the surface, the temperature of these molecules and the temperature of the surface. While the flux over the entire surface within a

feature is not consistent, many codes including the HVS model use a single RSF calculated based on a single representative flux value, where the representative flux is often the flux entering the feature scale model. Using a single RSF approximates the reaction as being a first order reaction where the sticking factor is independent of the incident flux. When there is very little variation in the surface flux within the solution domain this approximation will provide adequate results. However, while this may initially be the case the surface flux distribution will change as the feature fills and the aspect ratio increases. Depending on discrepancy between the actual flux dependence of the chemistry and the first order approximation large errors may result. The error may also be compounded by the fact that the surface flux distribution is also dependent on the magnitude of the sticking factors. The use of local sticking factors accounts for changes in the surface flux conditions which reduces the error in the prediction of the surface topography. The use of locally calculated sticking factors also increases the validity of the approach^[60].

In previous studies the use of variable local sticking factors has been shown to affect the prediction of deposition topography^[42] and the minimum deposition rate for non-first order reactions^[43]. However, the use of local sticking factors increases the computational cost due to the iterative process by which a converged solution for the reaction dependant surface fluxes is reached. This additional cost increases with aspect ratio and decreases with increased sticking factors.

BTLSF is designed to use local sticking factors as seen in Equation (3-10). However, during the multi-scale simulation global sticking factors are used because the focus of the multi-scale simulations is the linking and not the accuracy of the deposition prediction at the feature scale.

3.4 Application to non-first order chemistry

In order to demonstrate the ability of the BTLFSF model to apply non-first order kinetics to different aspect ratio features a series of test cases were conducted. These test cases included applying second and half order reaction kinetics to rectangular features with aspect ratios varying from 1 to 100. The second and half order kinetics were applied by defining the local sticking factor for surface segment i in a similar way to Cale and Raupp^[42] with

$$\varepsilon_i = \min \left(\varepsilon_0 \left(\frac{\eta_i}{\eta_0} \right)^m, 1.0 \right), \quad (3-20)$$

where η_i is flux for segment i and m is the reaction order. η_0 and ε_0 are reference values for the fluxes and sticking factors based on the flux for a flat surface segment outside of the feature.

Second and half order kinetics were chosen because they do not contain chemistry specific constants such as those present in more realistic kinetics such as Langmuir-Hinshelwood or Eley-Rideal. However, a third series of test cases were also run using Eley-Rideal kinetics based on the reaction rate equation presented by Chaara *et al*^[74],

$$R_{WF_6,i}^S = 0.002 \exp \left(\frac{-8300}{T} \right) \frac{P_{H_2,i}^{1/2} P_{WF_6,i}}{1 + 1000 P_{WF_6,i}}, \quad (3-21)$$

where $P_{H_2,i}$ and $P_{WF_6,i}$ are the local partial pressures and T is the surface temperature. Using this reaction rate equation the local sticking factors were calculated by dividing the local reaction rates by the appropriate local surface flux. This third series of test was conduct in order to demonstrate that the BTLFSF model can also handle more complex kinetic models.

Figure 3-7 shows the results of these test cases. The base sticking factor for the first and second order reactions in the present study was 0.01. While the surface temperature and source plane partial pressures for Eley-Rideal test case series were the same as those by Chaara *et al*.^[74] in their example 1: $T = 925\text{K}$, $P_{H_2} = 0.9091\text{Torr}$, $P_{WF_6} = 0.0909\text{Torr}$.

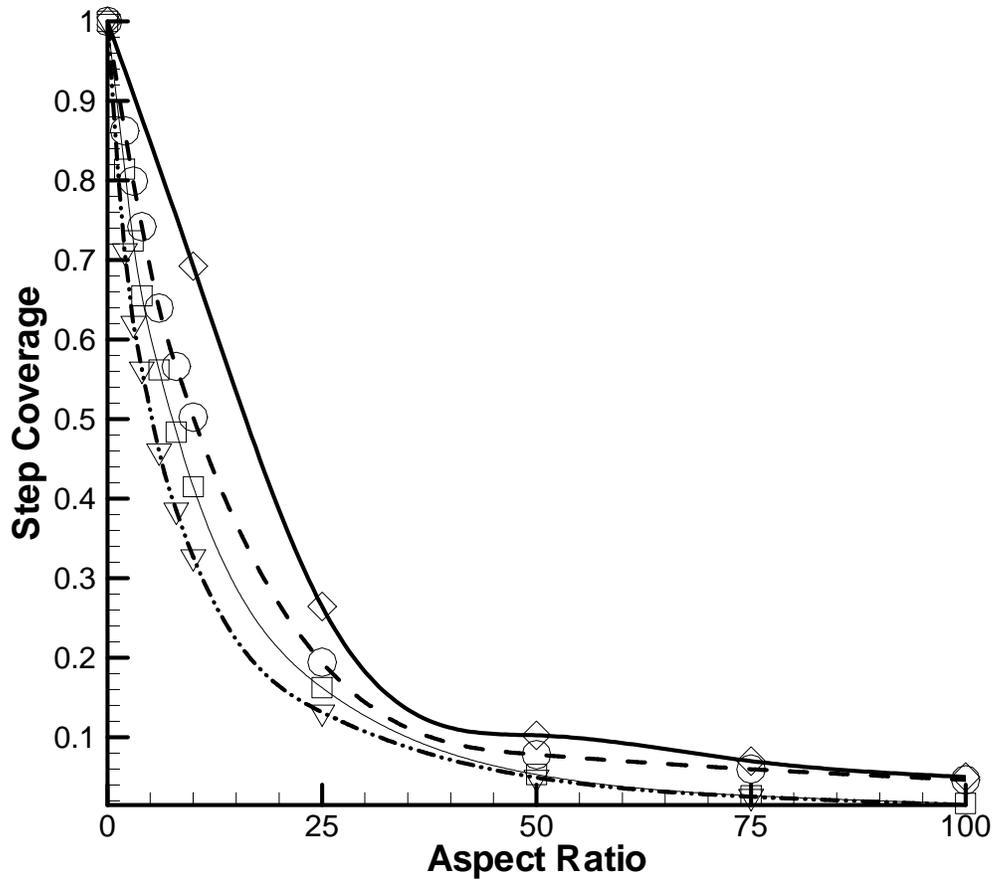


Figure 3-7 Variation in step coverage with feature aspect ratio for Order 0.5 (▽), Order 1.0 (□), Order 2.0 (○), and Eley-Ridel (◇) kinetics.

3.5 Source Plane Height Sensitivity

One of the benefits of the BTLFSF model is that coupling between the feature scale and reactor scale can be made easier for coupling at an offset plane. This is because the source plane of the feature scale can be the offset coupling plane. One concern with this method is that the height of the source plane above the surface may affect surface deposition prediction. Since there are no gas-phase collisions in the BTLFSF model the source plane height can only affect the initial surface flux. In order to determine the sensitivity of the initial flux into the source plane height a series of trials were

conducted with the source plane at different heights. The initial flux into the surface is independent of the sticking factor and is directly proportional to the flux across the source plane. The source plane height, h used in many of the calculations in this study is $1/3$ of the mean free path. By defining the error in the flux as

$$\text{flux error} = \frac{\eta^0 - \eta_{h=1/3\lambda}^0}{\eta_{h=1/3\lambda}^0} \cdot 100\% , \quad (3-22)$$

the flux error can be made independent of the flux across the source plane. This means that the error in the calculation of initial flux is only dependent on the geometry of the problem and the numerical method. By examining Figure 3-8 which shows the results of these trials it can be seen that the source plane height has very little impact on the surface deposition.

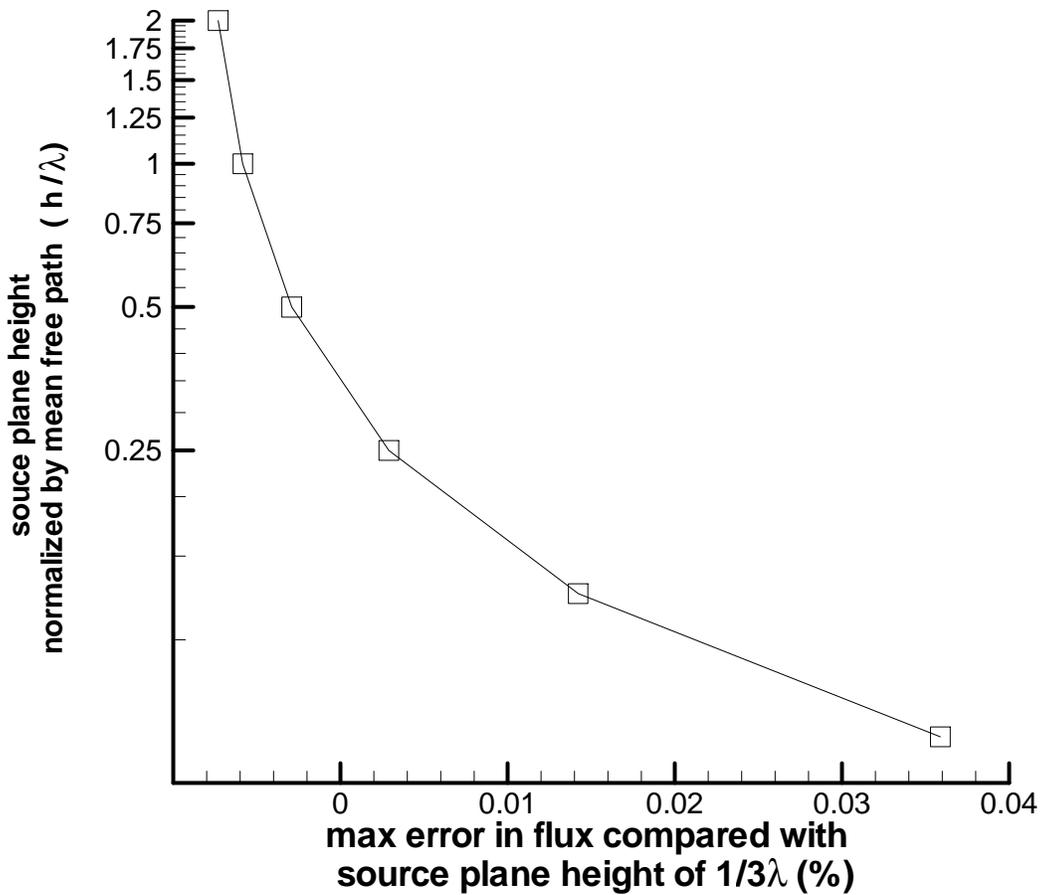


Figure 3-8 Change in surface flux prediction of BTLSE with source plane height.

The independence of the surface deposition on the source plane height is also due to the assumption that the number density, $n(\mathbf{x}')$ of particles along the source plane is constant. If $n(\mathbf{x}')$ varied over the source plane then the initial deposition would become dependent on the source plane height in the same way that the flux returning across the source plane is to be shown in the chapter on coupling.

3.6 Closure

In this chapter the development of a new feature scale model, BTLFSF is discussed. The key feature of this new model is the fact that while it is a ballistic transport based model it uses a source plane similar to that used in MC methods to calculate its initial deposition, instead of the hemispherical source used by most ballistic transport methods.

Another novel feature of the BTLFSF model is the ability to incorporate automatically varying time stepping. This is done by coupling the time step increment to the number of emission-deposition cycles required to meet the criteria for deposition convergence. The reason this works is because as the distance between two elements of the discretized surface decreases the probability of re-deposition increases. This means that the time step increment can be decreased as the entrance to features close to more accurately capture the geometry at closure.

In order to increase confidence in the results of the BTLFSF^[64] model is validated against other numerical models. Through the validation the BTLFSF model is shown to produce results comparable to HVS, BTRM and MC simulations. The BTLFSF model is also demonstrated to be able to handle high aspect ratio features, and non-first-order chemistry, such as the Eley-Rideal model for tungsten deposition.

Finally the deposition is shown to be independent of source plane height as long as the number density is constant over source plane. This will become important later when coupling between the reactor and feature scale models is discussed.

Chapter 4

Coupling

When modeling the flow of gases within a CVD reactor, the length scales associated with the reactor scale allow for continuum based equations such as the Navier-Stokes equations to be used. The same is often not true of the feature scale, where the length scales of features can be so small that the local Knudsen numbers indicate the requirement of particle based methods for modeling the fluid flow. While fundamental particle based methods could be developed to model both the reactor and feature scales, this approach would be extremely computationally costly^[75]. A solution to this difficulty is to use a heterogeneous multi-scale model which combines a continuum based reactor scale model and a particle based feature scale model.

The coupling between the continuum based reactor and particle based feature scale models is important in multi-scale modeling of CVD because it determines how and what information is passed between these two scales. A simple way to couple the two scales is to solve the scales sequentially^[4,76]. First the continuum transport model is solved without taking into account the effects of surface topography, and then the feature scale model is solved based on the results of the reactor scale model. However, this method of coupling is inaccurate because surface topography can greatly increase the surface area exposed to the gas, which may increase the deposition reaction rate^[77]. This means that in order to have an accurate multi-scale model there must be bi-directional coupling. When the models for the different scales are bi-directionally coupled information is passed to and from both models. The reactor scale model supplies the feature scale model with information on the gases entering the feature scale domain while the feature scale model supplies information on the surface reaction rate to the reactor scale model.

A recent literature survey on multi-scale modeling of CVD by Kleijn *et al.*^[1] cites two different approaches currently being used to bi-directionally link macroscopic reactor and microscopic feature scales. The first of these approaches, proposed by Gobbert *et al.*^[2], couples the feature and reactor scales through the net flux into the surface. The second approach mentioned by Kleijn *et al.*^[1] was proposed by Rodgers and Jensen^[3] and uses effective reactivity maps to link the reactor and feature scales. These effective reactivity maps are generated based on the net flux across a source plane offset

from the surface. Rodgers and Jensen^[3] claimed that their approach was more flexible than that of Gobbert *et al.*^[2]. However, no direct comparison of the two different approaches has been demonstrated.

This work presents a multi-scale CVD model where coupling can be performed either at the reaction surface as in Gobbert *et al.*^[2] or at an offset plane as in Rodgers and Jensen^[3]. In order to accomplish this, the ballistic transport based features scale model BTLSF^[64] was modified so that it could generate effective reactivity maps similar to those produced by Rodgers and Jensen^[3] using Monte Carlo (MC) methods. An alternative method of generating effective reactivity maps based on coupling at the surface is also proposed in this study. The benefit of this would be an increase in the flexibility of the surface based coupling method.

4.1 Theory

4.1.1 Downscaling of reactor scale to feature scale

The first attempts at sequentially coupling the reactor and features scales consisted of passing concentration, temperature, and pressure information from the reactor scale to the feature scale^[4,78]. This is similar to the downscaling coupling used in this study. When MC or line of sight models are used to model the feature scale these parameters are used to define the source flux such that

$$F_k = \frac{n_k}{4} \left(\frac{8RT}{\pi m} \right)^{\frac{1}{2}}, \quad (4-1)$$

where the number density, n of a specific species is calculated from

$$n_k(x) = \frac{P_k}{RT}. \quad (4-2)$$

In the development of the BTLSF model discussed in Section 3.1, Equation (3-8) assumed that the number density and therefore flux was constant across the entire source plane. This means that only one pressure, temperature, and concentration value for each species are required for each feature scale model. The values used in these calculations are the face values of the appropriate reactor scale cell.

The assumption of a constant number density is a good assumption for the test cases presented in this study because the features are located in regions where there is very little gradient in the concentration, pressure or temperature. The BTLFSF model can be modified to allow for variation of the number density over the source plane. However, this most likely would require numerical integration of Equation (3-8) and increase computational cost.

Another assumption inherent in Equation (4-1) is that the flux crossing the source plane has a Maxwell distribution. The assumption of a Maxwell distribution means that the gas near the wall is assumed to be stationary. However, from the reactor scale model, Equation (2-31), we know that the mean velocity normal to the surface is non-zero due to the surface reaction. Since the mean velocity is non-zero it would be more accurate to use the advected Maxwell distribution^[12]. In Cartesian coordinates the resulting flux of molecules across the source plane is

$$F_k = \frac{n_k \beta^3}{\pi^{3/2}} \int_{-\infty}^{\infty} \exp(-\beta^2 w'^2) dw' \int_{-\infty}^{\infty} \exp(-\beta^2 v'^2) dv' \int_{-c_0 \cos \theta}^{\infty} (u' + c_0 \cos \theta) \exp(-\beta^2 u'^2) du' \quad (4-3)$$

where c_0 is the stream velocity, θ is the angle which the stream velocity makes with the source plane normal and β is the inverse of the most probably thermal speed,

$$\beta = \left(\frac{2RT}{m_k} \right)^{-1/2}. \quad (4-4)$$

Integration of Equation (4-3) results in

$$F_k = \frac{n_k}{2\pi^{1/2} \beta} \left[\exp(-\beta^2 c_0^2 \cos^2 \theta) + \pi^{1/2} \beta c_0 \cos \theta \{1 + \operatorname{erf}(\beta c_0 \cos \theta)\} \right]. \quad (4-5)$$

Since there is no horizontal velocity component to the mean velocity $\theta = 0$, this allows Equation (4-5) to be simplified so that

$$F_k = \frac{n_k}{2\pi^{1/2} \beta} \left[\exp(-\beta^2 c_0^2) + \pi^{1/2} \beta c_0 \{1 + \operatorname{erf}(\beta c_0)\} \right].$$

(4-6)

When the flow was assumed to be stationary this equation can be simplified further to give

$$F_k = \frac{n_k}{2\pi^{1/2}\beta} = \frac{n_k}{4} \left(\frac{8RT}{\pi n_k} \right)^{1/2},$$

(4-7)

which is equivalent to Equation (4-1). In order to quantify the error which results from ignoring the impact of the velocity normal to the wall on the flux distribution, the fluxes resulting from Equations (4-6) and (4-7) are compared for a range of temperatures and normal velocities. Figure 4-1 shows the flux error calculated as

$$\text{flux error} = \frac{|F_k^{\text{advected Maxwell}} - F_k^{\text{Maxwell}}|}{F_k^{\text{Maxwell}}} \times 100\%$$

(4-8)

for a species with a molecular weight of 297.84 g/mol. This is the heaviest species used in this study and represents the maximum flux error since the flux error increases with molecular weight. The maximum velocity normal to the reacting surface in this study was found to be approximately 0.02 m/s. This means that the maximum flux error for this study should not be more than approximately 0.015%, which is acceptable.

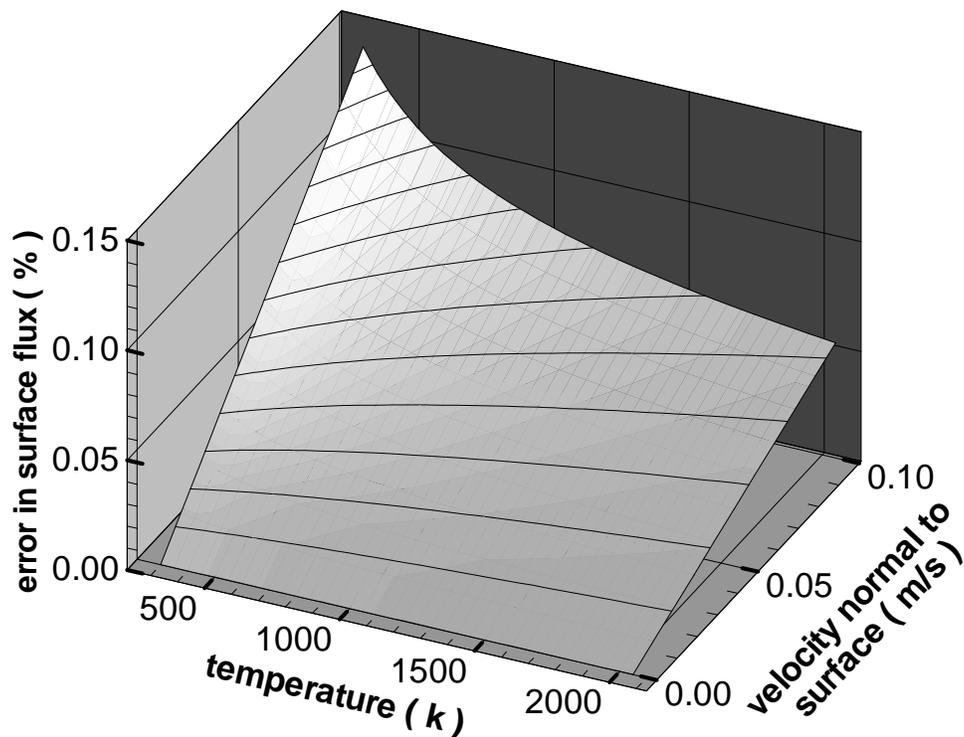


Figure 4-1 Estimate of error in surface flux resulting from the assumption of a Maxwell distribution.

4.1.2 Homogenization of feature scale results to reactor scale.

The major advantage of the bi-directionally coupled methods is the ability to homogenize the feature scale reaction rate results back to the reactor scale. Without homogenization of these results to the reactor scale the impact of the features on the apparent reaction rates will not be related to the reactor scale, reducing the accuracy of the entire model. The theory behind the two different methods of homogenizing the feature scale results to the reactor scale used in this study are discussed in this section.

4.1.2.1 Offset plane method

The concept of using a plane offset from the surface as the interface between the reactor and feature scales was first proposed by Rodgers and Jensen^[3]. In their method they generate an effective reactivity map for the reactor scale based on MC feature scale results. However, the method described is only applicable when a MC feature scale model is used. In order to apply the effective reactivity map based linking to a multi-scale code using a ballistic feature scale, a new method for determining effective reactivity maps had to be developed.

The effective reactivity is the ratio of the net flux across the source plane with a feature present to that which would be expect were no feature present. An effective reactivity map indicates the change in the local effective reactivity over the source plane. This map can then be homogenized to the reactor scale to produce an effective reactivity map at the source plane.

The local effective reactivity can be calculated as

$$\xi(x) = \frac{(F - \eta_{exit}(x))}{\epsilon_0 F}, \quad (4-9)$$

where $\xi(x)$ is the effective reactivity, F is the flux entering the source plane from the source volume, $\eta_{exit}(x)$ is the local flux exiting across the source plane, and ϵ_0 is the fraction of particles which will react on a homogenous surface.

The flux exiting across the source plane can be calculated as

$$\eta_{exit}(x) = \frac{1}{2} \int_0^\pi (1 - \epsilon) \eta \cos(\theta) d\theta, \quad (4-10)$$

where ϵ is the sticking factor for the surface within angle $d\theta$, η is the flux arriving at the surface segment located at angle θ , and angle θ is defined in Figure 4-2. The BTLSF code used in this study for the feature scale was developed to use a source plane instead of a hemispherical source. The use of a feature scale model with a source plane is important to the calculation of effective reactivity maps because it allows the location of the interface between the two models to be consistent. The

height of the source plane above the map may also affect the distribution of particles arriving at the deposition surface as discussed in Section 3.5.

In order to apply an effective reactivity map to the reactor scale the local effective reactivities must be averaged over the reactor scale cell face. Averaging ξ gives

$$\bar{\xi} = \frac{R^S}{R_{flat_plate}^S} = \frac{\int_L (\xi(x) - 1) dx + 1}{L},$$

(4-11)

where L is the width of the reactor scale cell.

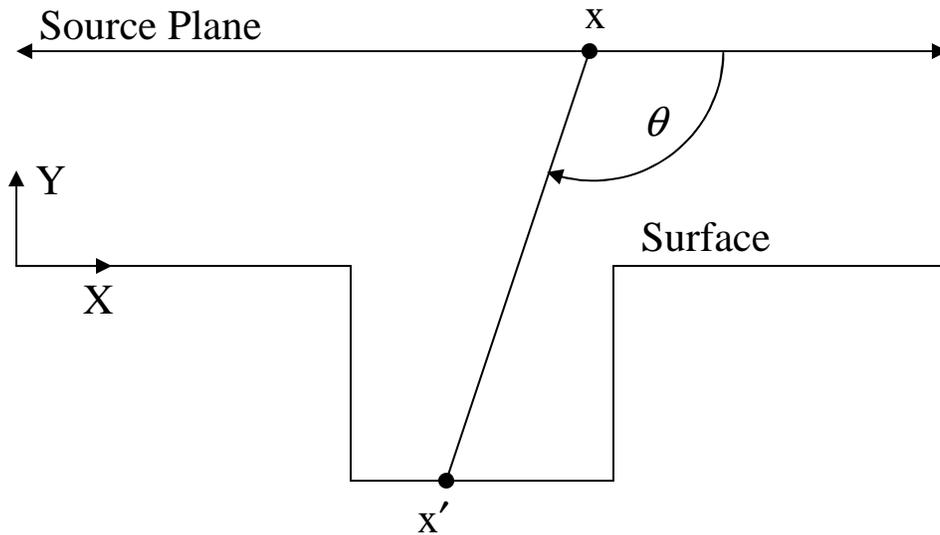


Figure 4-2 Schematic showing definition of angle θ

4.1.2.2 Surface plane method

The first study with full integration between the reactor scale and the feature scale was done by Gobbert *et al.* [79]. In their study they linked a ballistic based feature scale model to a continuum based FEM model through the net flux into the surface. This is accomplished by having linked feature scale simulations at each node on the boundary with the deposition surface. The net fluxes returned by

these feature scale simulations were assumed to represent the net fluxes into the deposition surface at the associated FEM boundary nodes.

Similar results to those of Gobbert *et al.* ^[79] can be attained by assuming that the average reaction rate per unit area, r can be approximated by integrating the reaction rate over the feature scale surface and dividing by the width of the feature,

$$\bar{\xi} = \frac{R^S}{R_{flat_plate}^S} = \frac{\int \varepsilon \eta d\tilde{A}}{A \varepsilon_0 F}, \quad (4-12)$$

where \tilde{A} is the arc length along the feature surface, A is the width of the feature scale domain, ε is the local sticking factor and η is the flux into the surface. Both this method and that of Gobbert *et al.* ^[79], implicitly assume that the entire surface of the associated reactor scale cell is patterned with identical features.

In order to extend these methods to include simulation of non-uniformly patterned features a means of generating effective reactivity maps for surface coupled features is proposed. This method requires that, when different features or regions with no features are present, the effect of the different regions be combined by weighting their contribution to $\bar{\xi}$ by the fraction of L which they cover. Such that

$$\bar{\xi} = \frac{R^S}{R_{flat_plate}^S} = \frac{\sum_{k=1}^{NF} \int \varepsilon \eta d\tilde{A}_k + \left(L - \sum_{k=1}^{NF} A_k \right) \varepsilon_0 F}{L \varepsilon_0 F}, \quad (4-13)$$

where NF is the number of features. Combining $\bar{\xi}$ values for different cells can result in an effective reactivity map for the reactor scale without the use of a source plane.

4.1.3 Limitations of coupling methods

The results of both linking methods are that the effective reaction rate at the reactor scale can be defined as

$$R^S(\mathbf{x}) = \bar{\xi}(\mathbf{x}) \cdot R_{flat_plate}^S. \quad (4-14)$$

This allows the effective growth rate of the reactor scale to account for the extra material deposited on the surface due to the presence of features. However, in order to couple the feature scale with a continuum based reactor scale in this manner the gradients in the reactivity map must be resolvable at the source plane. Whether or not the gradients can be resolved is determined by the associated Knudsen number. Using Birds^[12] definition the local Knudsen number can be calculated as

$$Kn(\mathbf{x}) = \frac{\lambda(\mathbf{x})}{L(\mathbf{x})}, \quad (4-15)$$

where $\lambda(x)$ is the local mean free path and $L(x)$ is the local gradient-length. The reason for using this definition of Knudsen number is that when the Navier-Stokes equation is obtained by the Chapman-Enskog approach, the Knudsen number used in the expansion is based on a length scale determined by flow gradients. Calculation of the local mean free path require that the local temperature, pressure, and concentrations be used in Equation (2-2) such that

$$\lambda(\mathbf{x}) = \frac{RT(\mathbf{x})}{\sqrt{2\pi}d^2(\mathbf{x})N_A P(\mathbf{x})}, \quad (4-16)$$

where $d(x)$ is the average particle diameter. The local length scale determined from the flow gradients is

$$L(\mathbf{x}) = \left. \frac{\phi}{\frac{d\phi}{d\mathbf{x}}} \right|_x, \quad (4-17)$$

where ϕ is a flow property. Usually ϕ is either the density or temperature but Rodgers and Jensen^[3] used $\bar{\xi}$. The choice of $\bar{\xi}$ is a reasonable choice in this case because like with density and temperature the gradients of $\bar{\xi}$ can indicate the variation from collisional equilibrium.

The use of local Knudsen numbers means that the continuum model may be applied if the Knudsen number is less than 0.2^[12]. This means that L must be greater than 5λ . If we make the conservative assumption that $\bar{\xi}$ is 1.0, then the gradient of $\bar{\xi}$ will be resolvable at the reactor scale if

$$\left| \frac{d\bar{\xi}}{dx} \right| < \frac{1}{5\lambda}.$$

(4-18)

4.1.3.1 Offset Coupling Method

The gradient of $\bar{\xi}$ must be kept below this limit to keep the reactor scale results valid. The factors which affect $\bar{\xi}$ for the offset coupling method are the source plane height, the geometry of the features, the spacing of the features, the sticking factor of the gas species, and the reactor scale cell size. By controlling these factors the gradient of $\bar{\xi}$ can be controlled. However, some of these properties are given or restricted limiting their usefulness in controlling the value of $\bar{\xi}$.

The height of the source plane has a great impact on the gradient of ξ and therefore $\bar{\xi}$. Figure 4-3 shows how the shape of the effective reactivity profile changes with the source plane height. As the source plane height decreases the gradient of effective reactivity increases. This means that a resolvable gradient could be achieved by increasing the source plane height above the surface. However, as the source plane height is increased the probability of gases phase collisions increases. Rodgers and Jensen^[3] stated that a surface offset of $1/3\lambda$ was sufficiently close to ensure that gases phase collisions would not be significant. This is very important to ballistic based methods like BTLSEF because they assume collisionless flow. A low probability of gases phase collisions is also important to MC methods because with each gases phase collision the computational cost increases. This is why Rodgers and Jensen^[3] selected a surface offset of $1/3\lambda$ in their study.

Rodgers and Jensen^[3] also noted that, with regards to the gradients in ξ , collisionless flows represented a worst cases scenario and that allowing gas phase collisions would smooth out the effective reactivity map thus making them easier to resolve with continuum based models.

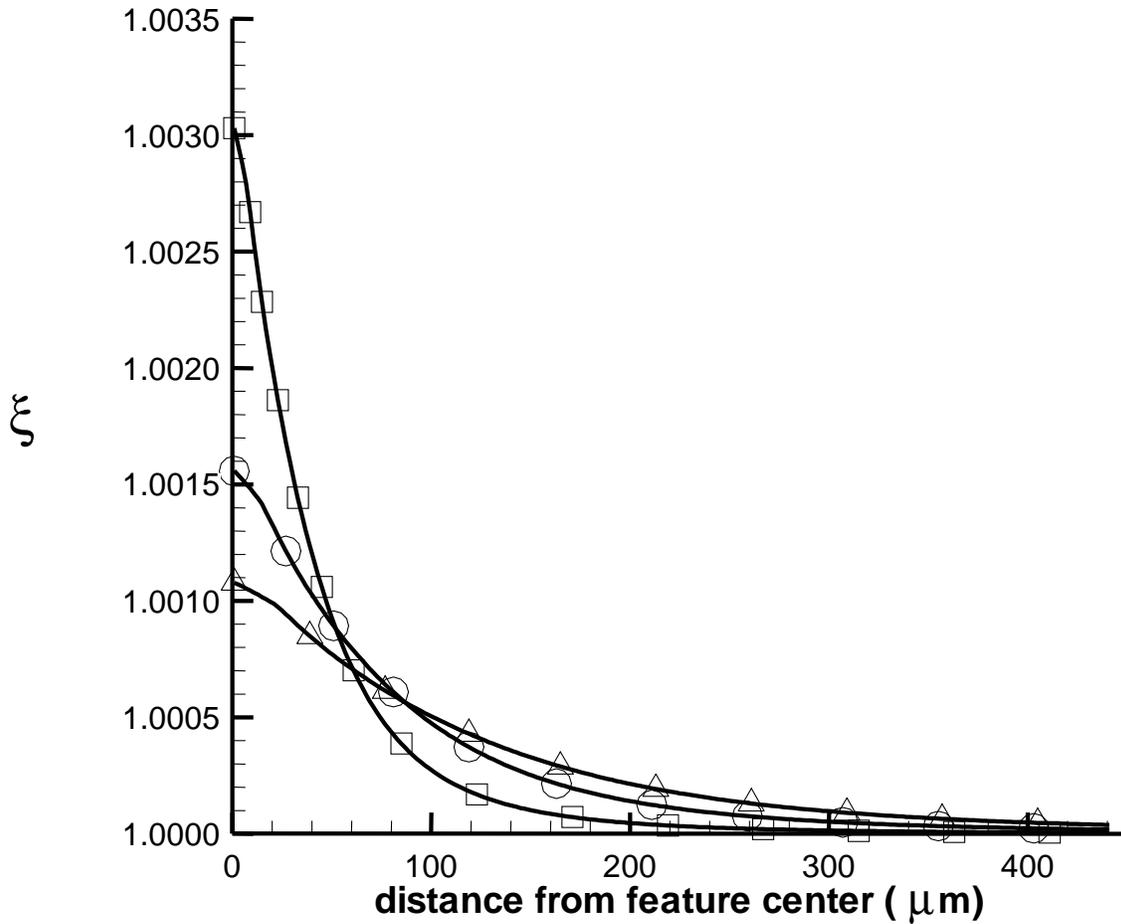


Figure 4-3 Dependence of ξ distribution on source plane height for aspect ratio 10 feature with a width of $0.1\mu\text{m}$, where $h = \lambda$ (Δ), $h = 2/3 \lambda$ (\circ), and $h = 1/3 \lambda$ (\square).

The gradient of ξ is also affected by the shape of the individual features and the patterning of the features. Increasing the aspect ratio of features increases the gradient in the ξ function associated with the feature. This can be seen in Figure 4-4. Keeping the aspect ratio constant and varying the width of the features allows the effect of the feature width on the gradient of ξ to be observed. Figure 4-5 shows the ξ function resulting from three features with different widths and an aspect ratio of 10. Examining Figure 4-5 it can be seen that varying the width of the feature affects the ξ function in the same way as the aspect ratio, with the gradient increasing as the feature width increases. The increase

in the ξ function is directly proportional to the increase in width such that if the width of the feature is increased by a factor of 2 so will the height of the ξ function. This is different from variations in the aspect ratio where the dependence of the ξ function is more complex. Knowledge of the effect of feature geometry on the ξ function does not help to keep the gradient within limits. This is because while the aspect ratios of the geometries in this study are only selected for demonstration purposes, in reality they would be predetermined based on product requirements.

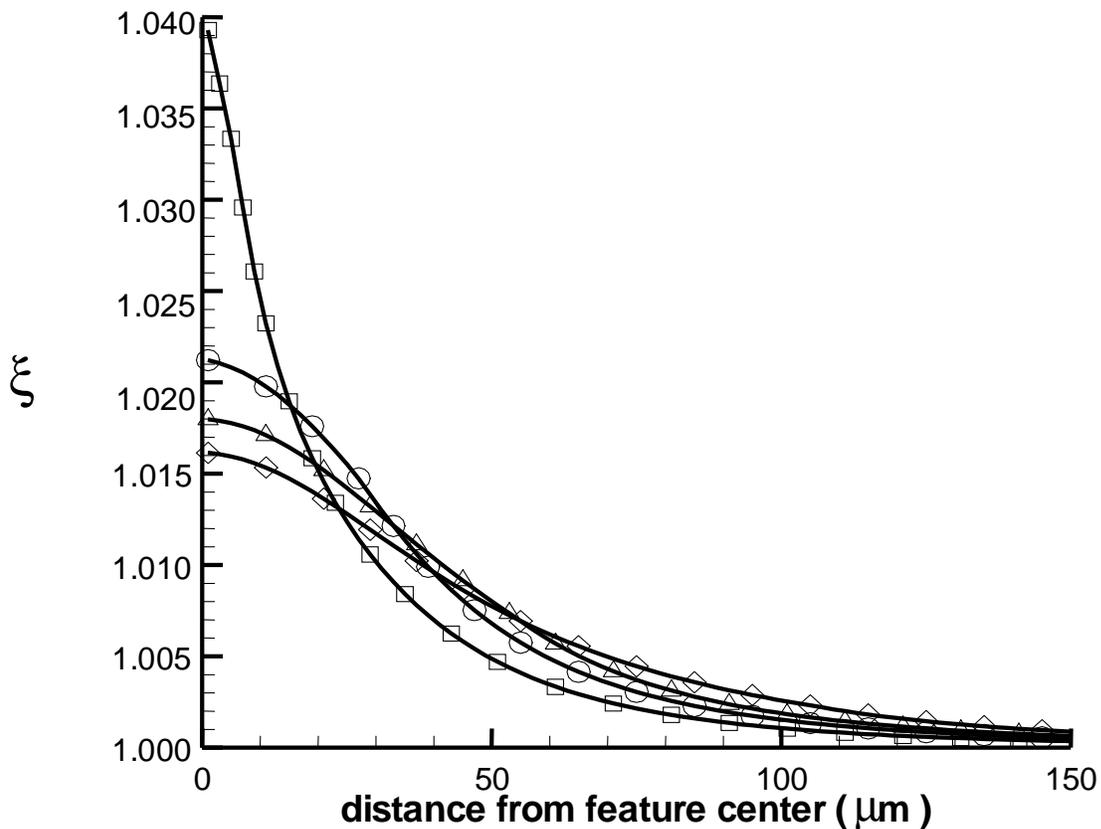


Figure 4-4 Variation of ξ function with aspect ratio (AR) when the source plane height, $h = 1/3 \lambda$ and the feature width is $0.1 \mu\text{m}$. [(\square) AR = 10, (\circ) AR = 2, (Δ) AR = 1, (\diamond) AR = 0.5.]

Since it is the combined ξ which results in the $\bar{\xi}$ function seen at the reactor scale the spacing of features is also important. The area of the source plane affected by the presence of a feature is much larger than the actual feature. This can be seen in Figure 4-3 through Figure 4-7 where the region

affected by the presence of a feature may be as much as 1000 times greater than the width of the feature. This means that unless features are spaced more than 1000 times their width apart, the resultant combined ξ function peak value will be larger than the individual feature value. As with the study of Rodgers and Jensen^[3] the combined ξ functions are determined through superposition. The normal spacing of features is much less than 1000 times the width of the features. In fact the feature spacing or pitch is often of the same order of magnitude as the feature width and many features are clustered together. This means that the slope of the combined ξ function is much larger than that of an individual feature. Figure 4-6 shows the combined ξ function for five features with a pitch of $2\mu\text{m}$. These features all have an aspect ratio of 10 and a width of $0.1\mu\text{m}$. Comparing the ξ function resulting from a single feature and that of the five combined features it can be seen that the combined ξ function is almost exactly the same as multiplying the original ξ function by a factor of 5. The difference is that the peak of the superimposed features is some what more round than that which would results from amplifying the original feature. This rounding results from the fact that the ξ functions of the individual features are not exactly aligned and the rounding increases with the pitch of the features.

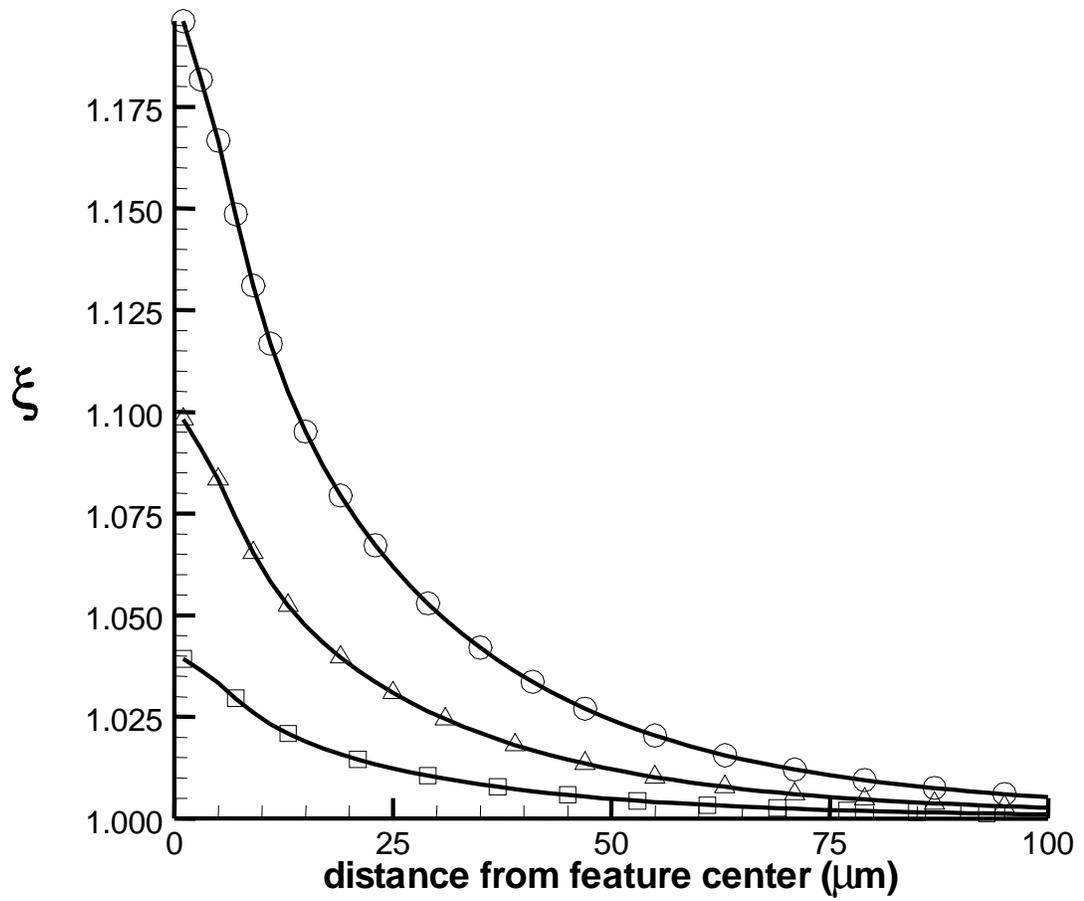


Figure 4-5 Variation of ξ function with feature width for feature with AR of 10 and source plane height, $h = 1/3 \lambda$. [width = $0.1 \mu\text{m}$ (\square), $0.25 \mu\text{m}$ (Δ), $0.5 \mu\text{m}$ (O)]

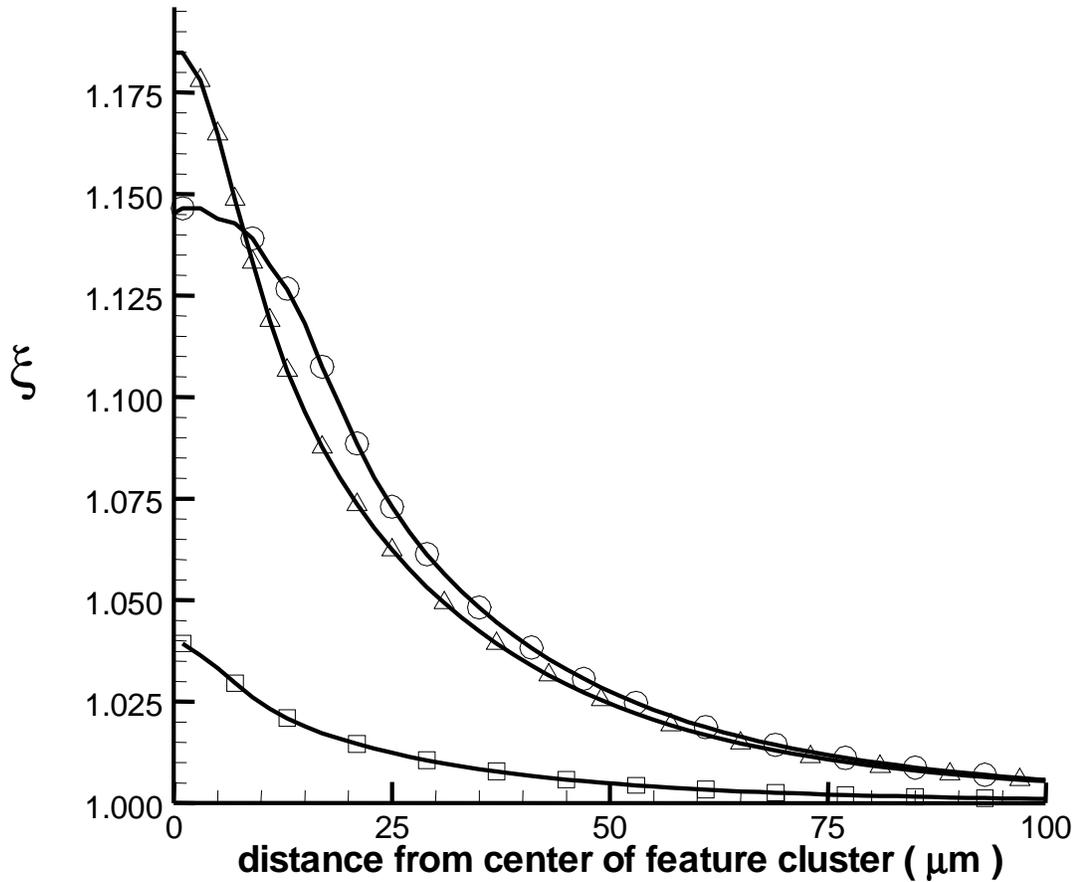


Figure 4-6 ξ function resulting from linear superposition of five features with a source plane height of $h = 1/3 \lambda$ and the features widths are $0.1 \mu\text{m}$. [Single feature (\square), five features with pitch of $2\mu\text{m}$ (Δ), five features with pitch of $6\mu\text{m}$ (O)]

Another factor which affects the gradient of ξ is the surface chemistry. Species with high surface reactivity and therefore large sticking factors have smaller gradients in ξ than those with small sticking factors as seen in Figure 4-7. The chemistry like the geometry is however normally predetermined. This means that the only factor affecting the gradient in $\bar{\xi}$ which can be used to ensure the restrictions on Knudsen number are satisfied is the reactor scale cell size.

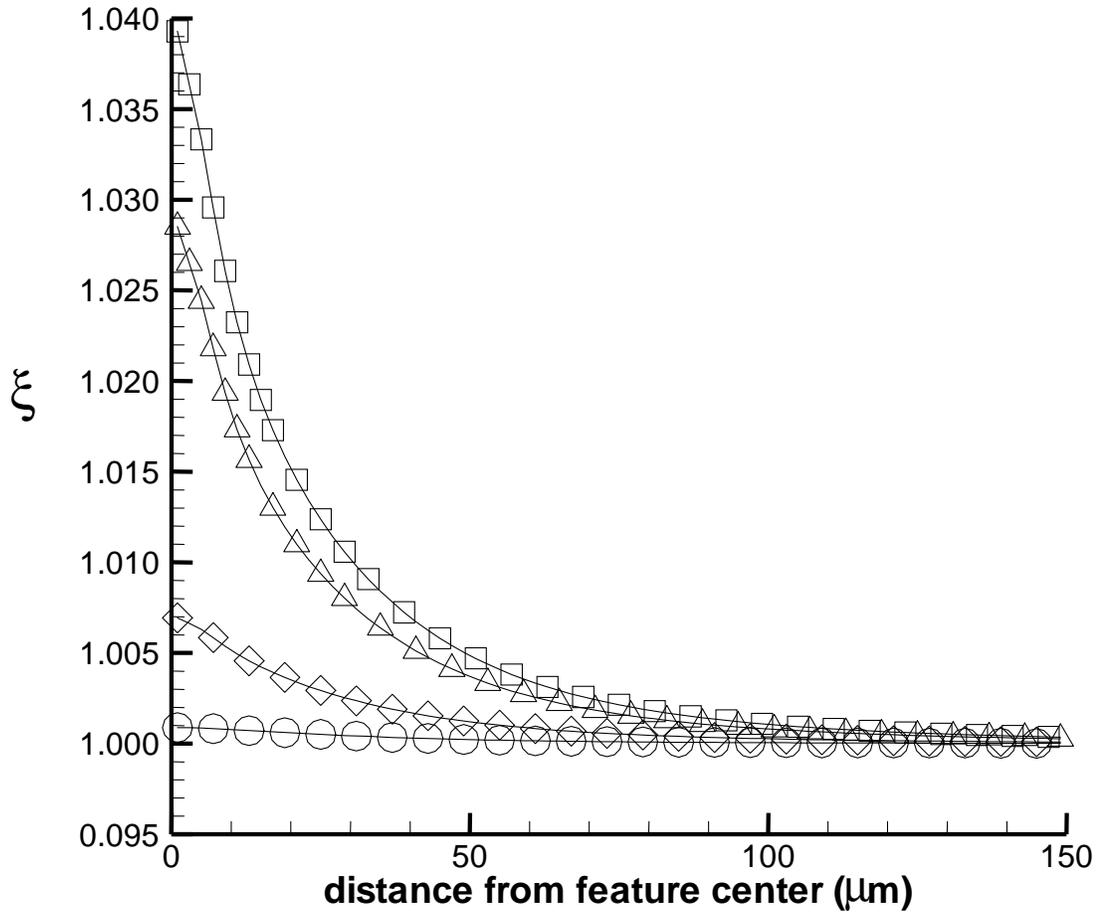


Figure 4-7 Dependence of $\bar{\xi}$ distribution on sticking factor for a aspect ratio 10 feature with a source plane height, $h = 1/3 \lambda$ and a feature width of $0.1 \mu\text{m}$, where $\varepsilon_0 = 0.5$ (\circ), $\varepsilon_0 = 0.1$ (\diamond), $\varepsilon_0 = 0.01$ (Δ), and $\varepsilon_0 = 0.001$ (\square).

The width of the reactor scale cells corresponds the denominator dx in Equation (4-18). Since the area of influence of the feature is limited, the gradient of $\bar{\xi}$ can be reduced by increasing the size of dx to include a larger area. This is due to the fact that the area under the $\bar{\xi}$ curve remains the same while the distance which it is averaged over increases. However, once dx becomes longer than the width of the area of influence of the feature the results may become indistinguishable from those of the surface

coupled method. For example if the width of the reactor scale cells associated with the ξ functions shown in Figure 4-7 was 300 μm , then the entire ξ profile could fit within the reactor cell. However, if the boundary of the reactor scale cell happened to fall within the ξ profile the $\bar{\xi}$ values generated by the two coupling methods could still be different.

4.1.3.2 Surface coupling method

Similar to the offset coupling method the gradient of $\bar{\xi}$ resulting from the surface coupling method is dependent on the geometry of the features, the spacing of the features, the sticking factor of the gas species, and the reactor scale cell size. The way in which these properties affect $\bar{\xi}$ is different than with the offset coupling method because the shape of the ξ function associated with the surface coupling method is always rectangular. This means that there are only two dimensions which can change. Furthermore, the width of the ξ function is always the same as that of the feature scale domain.

The shape of the actual ξ function is only important if the boundary of a reactor scale transects the feature scale. The feature scale is much narrower than the reactor scale cells making this unlikely. This means that the more important aspect of the ξ function is the area under the curve, $\xi \cdot dx$ because this represents the actual impact of the ξ function on the $\bar{\xi}$ function if the reactor cell size is kept constant. Where $\xi \cdot dx$ is defined as

$$\xi \cdot dx \equiv \int_L \xi dx, \tag{4-19}$$

and L is the reactor scale cell width.

When offset coupling is used the aspect ratio of the features has been shown to have a large impact on the shape of the ξ function and the maximum slope. The area under the curves, however, remains fairly constant. The area also remains fairly constant when the surface coupling method is applied as seen in Table 4-1. Since details of the return flux distribution are not present in the surface coupling method only $\xi \cdot dx$ is important for determining scaling length. This means that while the height of the

ξ function is affected by the aspect ratio, aspect ratio has very little impact on the effective slope of the ξ function when surface coupling is used.

Table 4-1 Variation of ξ with aspect ratio with sticking. ($\epsilon_0 = 0.0009$)

AR	ξ	Width of Feature	Width of Feature Scale	$\xi \cdot dx$
10	9.661	0.1	0.2	1.932
2	1.990	0.5	1	1.990
1	0.997	1	2	1.995
0.5	0.499	2	4	1.996

While aspect ratio does not have a large impact on the slope of ξ when surface coupling is applied, the shape of the geometry of the features is still important. Similar to the offset coupling method the area under the curve is directly proportional to the width of the features when the aspect ratio is kept constant. This means that an increase in the size of the feature by a factor of 5 will result in an increase in the effective slope of the ξ function by a factor of 5. This is reflected by the increase of $\xi \cdot dx$ in Table 4-2 with feature width.

Table 4-2 Variation of ξ with feature width. ($\epsilon_0 = 0.0009$)

Width of Feature	ξ	AR	Width of Feature Scale	$\xi \cdot dx$
0.1	9.661	10	0.2	1.932
0.25	9.662	10	0.5	4.831
0.5	9.664	10	1	9.664

In the offset coupling method the reactivity of the depositing gas represented by the sticking factor was shown to have a significant impact on the shape of the ξ function. Examining Table 4-3 it can be seen that there is a large variation of $\xi \cdot dx$ with sticking factor when surface coupling is used. This indicates that the sticking factor has a large impact on the effective slope of the surface coupling ξ function and therefore the scaling length.

Table 4-3 Variation of ξ with sticking factor for 0.1 μm wide feature

Sticking Factor ϵ_0	ξ	AR	Width of Feature Scale	$\xi \cdot dx$
0.5	0.35	10	0.2	0.070
0.1	2.13	10	0.2	0.426
0.01	7.23	10	0.2	1.446
0.001	9.63	10	0.2	1.926

The width of the reactor scale cells has the greatest controllable impact on the gradient of $\bar{\xi}$. As mentioned in Section 4.1.3.1 the width of the reactor scale cells corresponds to the denominator, dx in Equation (4-18). The effect of changing dx for surface coupling is different than with the offset coupling method. The reason for this is the ξ function of the offset coupling method is much wider than that of the surface coupling method. This makes it much more probable that the ξ function of the offset method will be distributed over more than one reactor scale cell. For example, Figure 4-8 shows the ξ profiles for an aspect ratio 10 feature for both coupling methods and the resulting $\bar{\xi}$ functions. The ξ function for the surface coupling method is only as wide as the feature scale domain. Since the ξ function is a step function for the surface coupling method this makes the potential slope of the $\bar{\xi}$ function infinite. In order to keep the Knudsen number below the continuity limit the width of the reactor scales is limited to

$$dx \geq \sqrt{5\lambda(Eta \cdot dx)}, \quad (4-20)$$

where ξdx is the area under the ξ curve. The reactor scale cell widths in Figure 4-8 meet the minimum of this criterion.

Although the ξ functions associated with the surface coupled method are very narrow and there is not likely to be any overlap with between ξ functions the spacing or pitch of features on the surface does affect the limitations on the width of the reactor scale cells. This is because the ξ functions of multiple features will lie within a single reactor scale cell. The ξdx values from the features within the same reactor cell will be summed when calculating the $\bar{\xi}$ function and the $\bar{\xi}$ value for the reactor scale

cell will increase. The increase in $\bar{\xi}$ may increase the size of reactor scale cell required to keep the Knudsen number within the continuum limits.

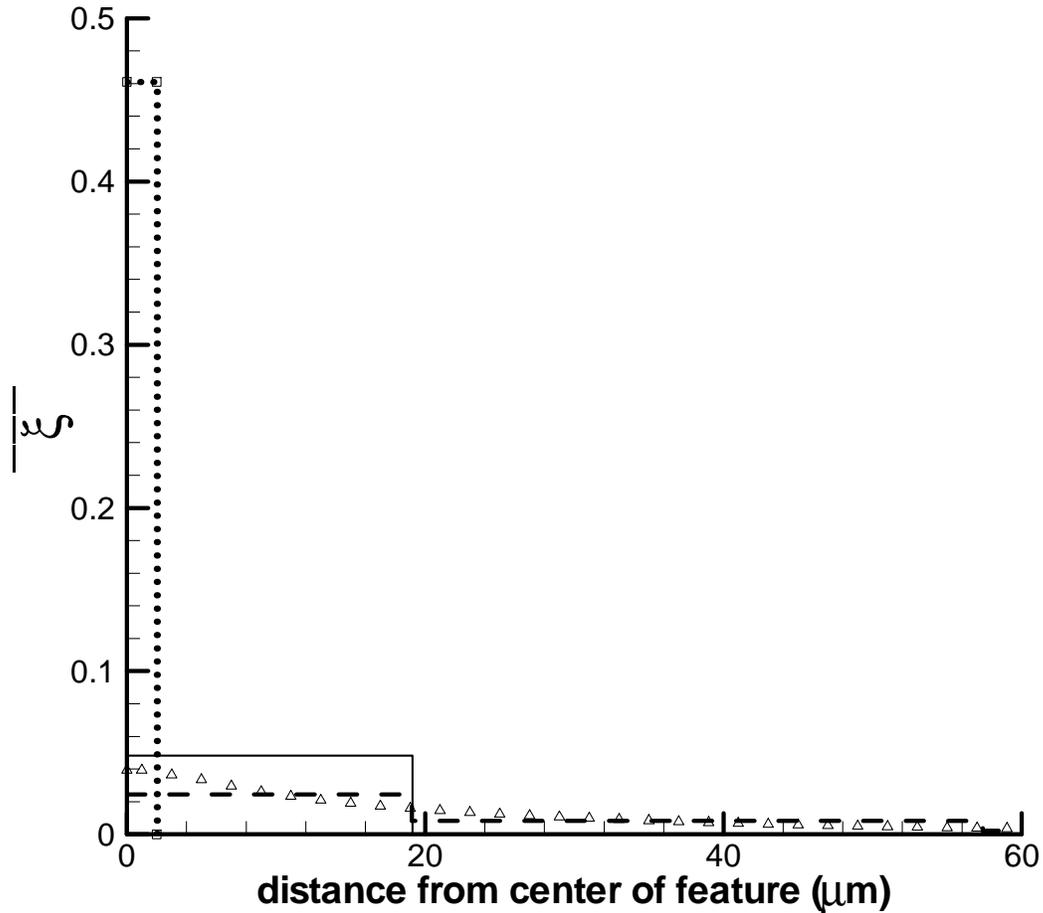


Figure 4-8 ξ profiles for surface and offset methods and their representation on the reactor scale, $\bar{\xi}$ for $0.1 \mu\text{m}$ feature with AR 10 and source plane height of $1/3\lambda$. [surface coupled ξ profile (\square), $\bar{\xi}$ profile (solid line); offset coupled ξ profile (Δ), $\bar{\xi}$ profile (dashed line)]

The ξ function of the offset coupling method is not a step function and the maximum slope is finite. This means that as long as the maximum slope of the offset methods ξ function is less than the maximum allowable slope defined in Equation (4-18) the width of the reactor scale cells is not limited

by the Knudsen number constraint. The cell size would still be limited, however, since it must be much larger than the mean molecular spacing. This is to ensure that the macroscopic properties based on the molecules within the cell are independent of the cell size^[12]. The mean molecular spacing is very small so this limit is much smaller than the width constraint imposed by the Knudsen number on the surface coupled reactor cells.

4.1.4 Use of Meso Scale

In between the reactor scale and the feature scale a meso scale can be introduced. This scale is at the limit of continuum based modeling and provides the best resolution possible with a continuum model. Gobbert *et al.*^[79] used this in between scale to distinguish between clusters of features within a die. When using an unstructured grid the meso scale can be incorporated into the reactor scale model through local mesh refinement^[80]. However, when a structured Cartesian grid is used, as in this study, a large number of cells are required to incorporate the meso scale. A solution to this problem used by Gobbert *et al.*^[79] was to solve the meso scale separately from the reactor scale.

The main difference between the meso scale of Gobbert *et al.*^[79] and that of the current study is that the velocity, pressure and temperature fields are simulated at the meso scale in the current study while they were interpolated in the study of Gobbert *et al.*^[79]. It is important to solve for the continuity and Navier-Stokes equations at the meso scale because the velocity field is significantly impacted by species consumption as shown in Figure 4-9. The variation in velocity field can also impact the species consumption through Equation (2-32).

Figure 4-9 shows the velocity fields, which result when the continuity and Navier-Stokes equations are solved and when values are interpolated from the reactor scale results, for a surface patterned by three regions of features. Comparing these two velocity profiles it can be seen that there is an increase in the velocity normal to the surface near the clustered features due to the presence of the features.

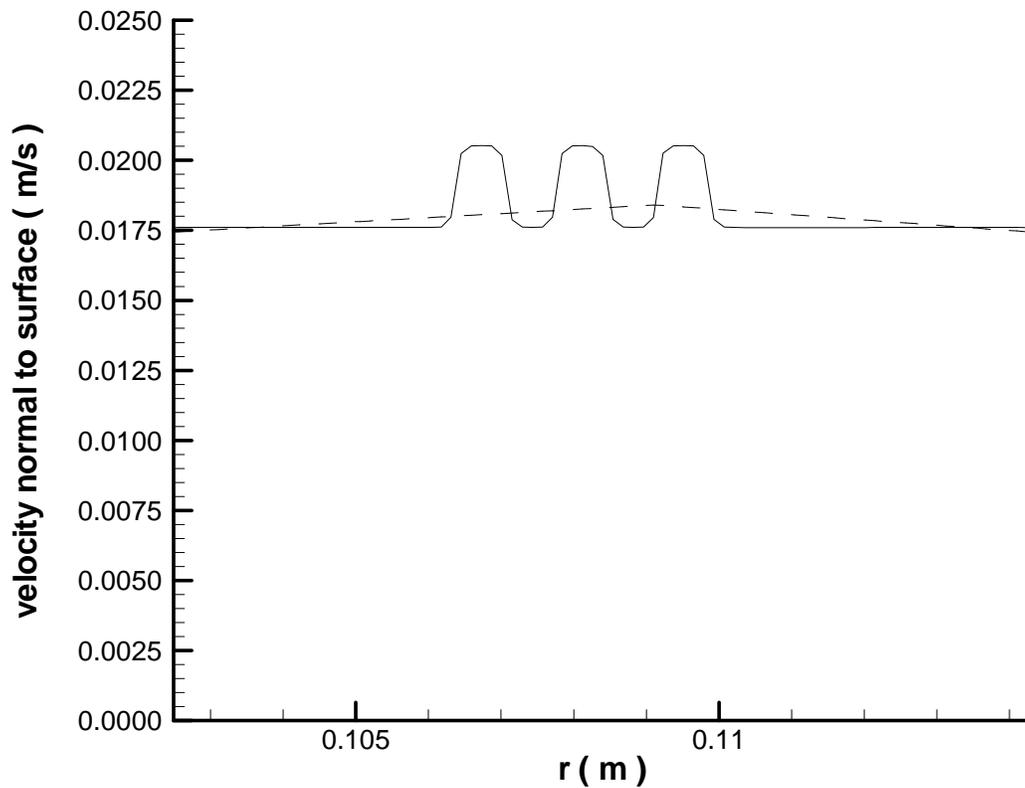


Figure 4-9 Effect of solving Navier-Stokes equation at meso scale on meso scale surface velocity profiles. (- - -) Velocity field interpolated from reactor scale. (—) Velocity field found using meso scale Navier-Stokes equation.

4.1.5 Solution Algorithm

The multi-scale code used in this study is a three-scale deposition model similar to that of Gobbert *et al.*^[79]. The reactor and meso scale models are both simulated using the finite volume code STREAM^[15]. While the feature scale is modeled using BTLFSF^[64], a ballistic transport model which uses a source plane in the calculation of the initial deposition.

The use of a finite volume code instead of a finite element code means that the coupling between the different scales is different from that of Gobbert *et al.*^[79]. The solution algorithm requires that the reactor scale is iterated to obtain an initial guess. This initial guess assumes that no features are

present on the surface since coupling with the feature scale has not yet occurred. The reactor scale results are used to prescribe flux conditions for three of the boundaries of the meso scale and interpolated over the meso scale to give an initial guess to the solution at the meso scale. The final boundary condition for the meso scale cells, for the faces adjacent the surface, are dependent on the associated effective reactivity map. The effective reactivity map for the initial iteration of the meso scale is calculated based on the meso scale data resulting from the interpolated reactor scale values. Using these initial boundary conditions the meso scale is iterated with the feature scale being updated at a predetermined interval. The preset number of iterations between feature scale runs used in this study was 300. The feature scale results could have been updated at each iteration of the meso scale but this would have greatly increased the computational cost. The meso scale solver continues to iterate until either convergence is attained or the maximum number of meso scale iterations is reached. This is followed by iteration at the reactor scale. The reactor scale iterates until it converges with the meso scale results being updated every so often. In this study the meso scale was updated every 300 iterations of the reactor scale. Having the reactor scale and meso scale coupled in this way meant that the meso scale could converge without having a floating outflow boundary.

The criteria for convergence for this study are that the sum of the absolute cell residuals be less than 10×10^{-5} . This constraint was found to be too lax for the species equation so the sum of the absolute cell residuals for concentration was required to be less than 10×10^{-9} . In addition to these requirements, convergence at the meso scale and reactor scale required that the results of the smaller scales not change when the converged results were applied as boundary conditions to the smaller scales.

Figure 4-10 shows a flowchart describing the multi-scale solution algorithm. The use of this algorithm is only possible due to the total integration of all three scales. If separate codes had been used full convergence at each scales would have been require before the results could be passed between the scales as in the solution algorithm used by Gobbert *et al.*^[79].

Pseudo steady-state results were used for the initial conditions in the transient simulations with changing feature topography. Pseudo steady-state results meant that the steady state form of the transport equations were used at the reactor and meso scales. When the feature geometry was advanced however the transient terms were applied to the transport equations with the time step

corresponding to the growth at the feature scale. This is the same approach used for transient simulation of the coupled system in Merchant *et al.* ^[59]. The use of the pseudo steady-state results as an initial condition allows simulation of the initial unstable start-up conditions to be avoided. ‘Since the growth rates are much smaller than residence times for the flow this is a reasonable approximation.’ ^[59]

4.2 Results

The geometry for the reactor scale in all of the multi-scale test cases is based on the geometry from Kleijn *et al.* ^[19]. This is because many of the previous multi-scale studies did not include the details of the reactor scale geometry as they were not considered to be important to the focus of their papers. The other benefit of using the Kleijn geometry is that the boundary conditions and grid implementation have already been validated in the reactor scale chapter.

In the present test cases there is also only a single meso scale per reactor scale. This was consistent with Gobbert *et al.* ^[79]. However, the methods presented here can easily be applied cases with multiple meso scales.

The region of the substrate surface with the most uniform concentration distribution was selected as the location of the meso scale for the test cases presented. The reason for selecting this region was that gradients in the concentration make it more difficult to detect the impact of features at the meso and reactor scales. This is not because the impact of the features is lessened but rather because it is easier to see changes in the concentration gradients when the gradient is close to zero.

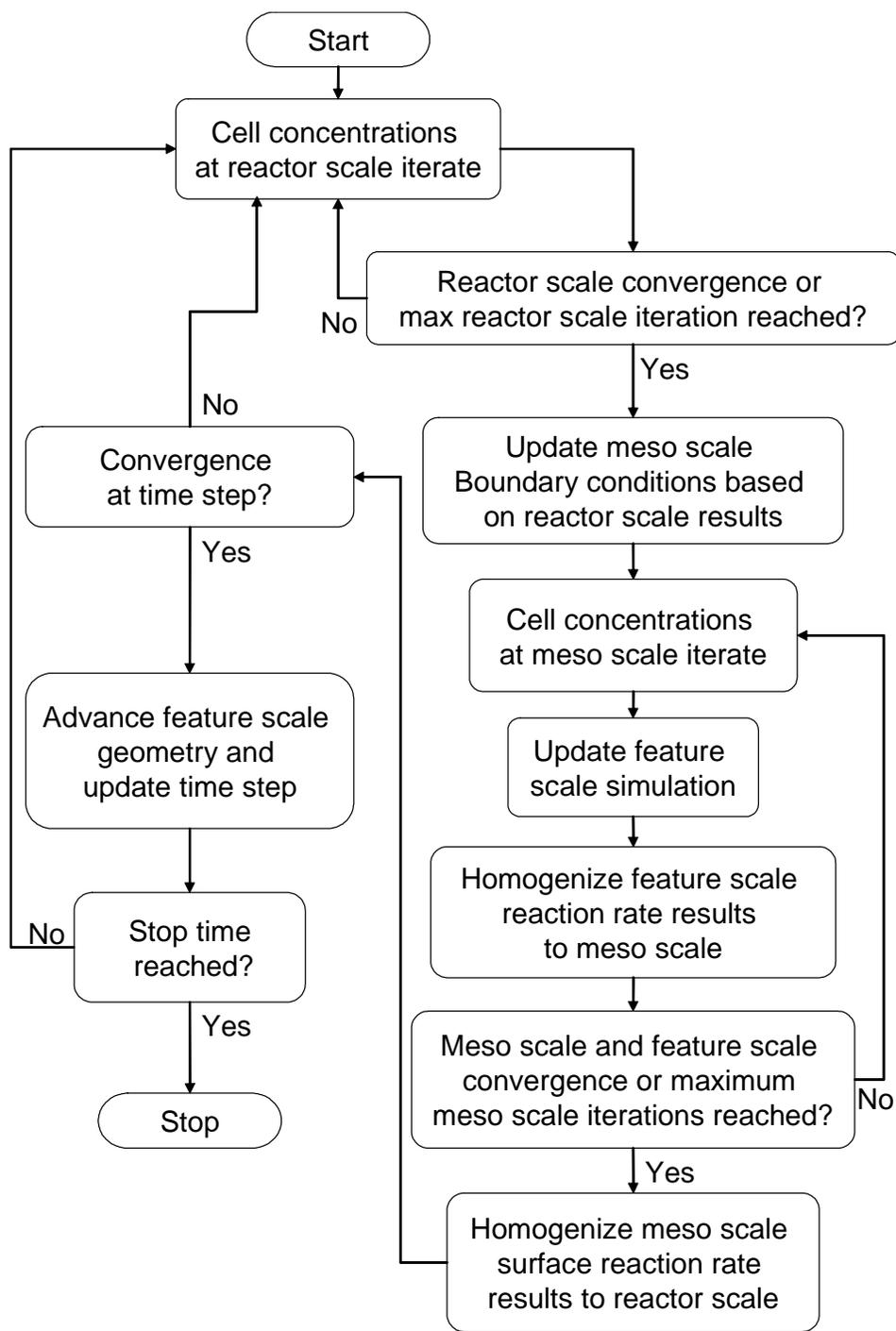


Figure 4-10: Algorithm for combining multiple scale simulation for transient simulations.

The results of two different chemistries are presented in this chapter; this was to see if there would be any impact of the chemistry on the conclusions. Hydrogen reduction of silane gas was selected as the first chemistry. This was because it was the chemistry used for the validation of the reactor scale and all of the transport properties were readily available from the literature^[19]. The second chemistry chosen was tungsten deposition. The reason for selecting this deposition was that it was used by Rodger and Jensen^[3]. Tungsten hexafluoride also has a much higher sticking coefficient than silane gas which greatly impacts the effective reactivity maps. Thermal deposition of silicon dioxide from tetraethyl orthosilicate used by Gobbert et al.^[79] was not selected because of its complexity and the fact that Merchant *et al.*^[59] report that they were unable to detect feature to feature or even cluster to cluster affects with this chemistry.

4.2.1 Silicon Deposition

Using the geometry and results from the reactor scale test case as a starting point, three multi-scale test cases were designed. As with the reactor scale test case the conditions for these test cases were an inlet flow rate of 0.2 slm, a wafer temperature of 1000 K, and a reactor pressure of 1 torr, and an inlet gas temperature and wall temperature of 300 K. The inlet gas was composed of 90 mole percent [m/o] hydrogen and 10 [m/o] silane.

The first two of these test cases share a similar format to those of Gobbert et al.^[79]. The first test case has features uniformly patterned over the surface of a reactor scale cell. This test case is important for determining the impact of including the meso scale on the reactor scale results. The second test case has clusters of uniformly patterned features. Clustering the features allows the benefits of the meso scale to be demonstrated because without the meso scale the clusters would not be resolvable. The final silicon chemistry based test cases shows the differences of the current coupling methods for non-uniformly patterned features. This final test case is beyond the capabilities of the method originally developed by Gobbert et al.^[79].

During each test case the results of the surface and offset plane coupling methods developed in this study are compared.

4.2.1.1 Uniform patterning of features over a reactor scale cell

The larger the region of the surface covered by features, the larger the impact of the features on the species concentrations. For this reason the first test case chosen for this study was one where the features are patterned uniformly over the entire surface of a reactor scale cell. The patterned features were infinite rectangular trenches with a width of 0.1 μm and a depth of 1 μm . The features were spaced 12 μm apart with flat areas separating the features. The reactor scale cell was 6 mm wide and was located so that its left edge was 106 mm from the center of the reactor. Covering an entire reactor scale cell with features also made it possible to show that the presence of a meso scale did not affect the results as in Gobbert *et al.* ^[2].

The two different linking methods were applied to this test case with the growth rate results being shown in Figure 4-11. The reactor and meso scale results for both linking methods are shown in Figure 4-11 so that the impact of the linking methods on the different scales can be seen. For this first test case there is almost no difference at either scale in the results produced using the two different linking methods. However, if the growth rates at the leading and trailing edges of the feature cluster are examined closely (Figure 4-11 right) it can be seen that the effective growth rate of the meso scale cells adjacent to these edges varies slightly depending on the coupling method used. The reason for these variations is the edges of the reaction rate profile are smoothed slightly when the offset plane coupling method is used. This results in a small (~1.4%) increase in the effective deposition rate of the first cell just outside of the blanketed region and a similar decrease in the deposition rate of the cells just inside the blanketed region. This effect, however, diminishes quickly beyond the boundary cells and so does not have a great effect on the reactor scale results as seen in Figure 4-11.

Reaction rate results without a meso scale are also shown in Figure 4-11 and it can be observed that the presence of a meso scale does not affect the reactor scale results. While changes to the growth rate profile are interesting, the effect of these changes on the species concentrations is more important since changes to the species concentrations affect the growth rate and conformality of the deposition. Figure 4-12 shows how the mass concentration of silane next to the surface changes at the meso scale due to the presence of features. The meso scale concentration profiles produced by both surface and offset coupling methods are also shown in Figure 4-12. However, there is no discernable difference in the concentration results produced by the two methods, which are within about 0.33%.

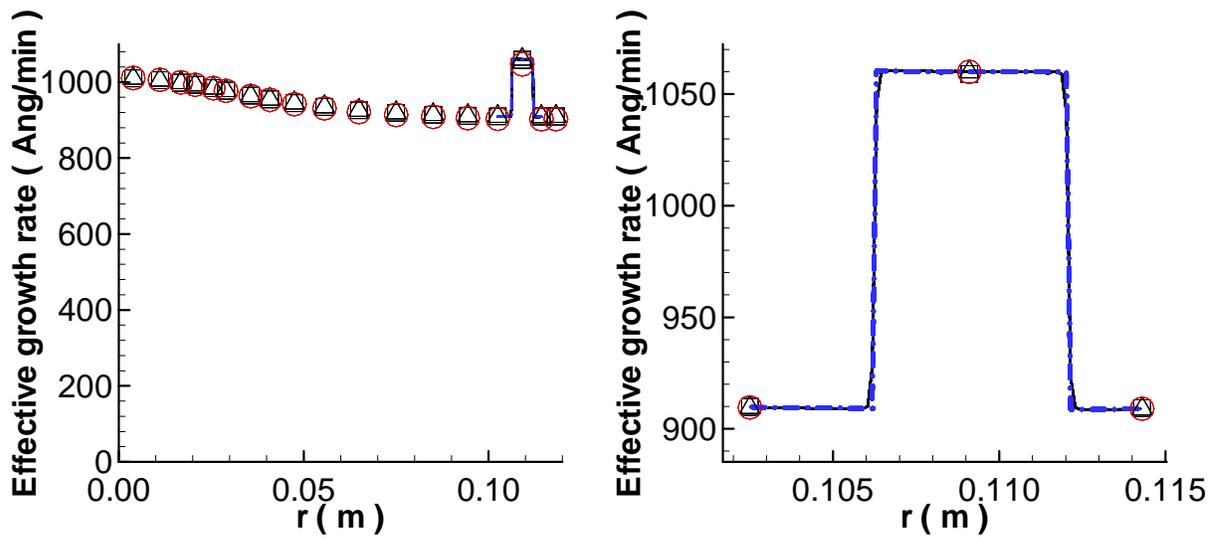


Figure 4-11 Growth rate results with features uniformly distributed over reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) reactor scale with coupling at source plane, (\circ) reactor scale coupled directly to feature scale, (- - -) meso scale with coupling at surface, and (—) meso scale with coupling at source plane.

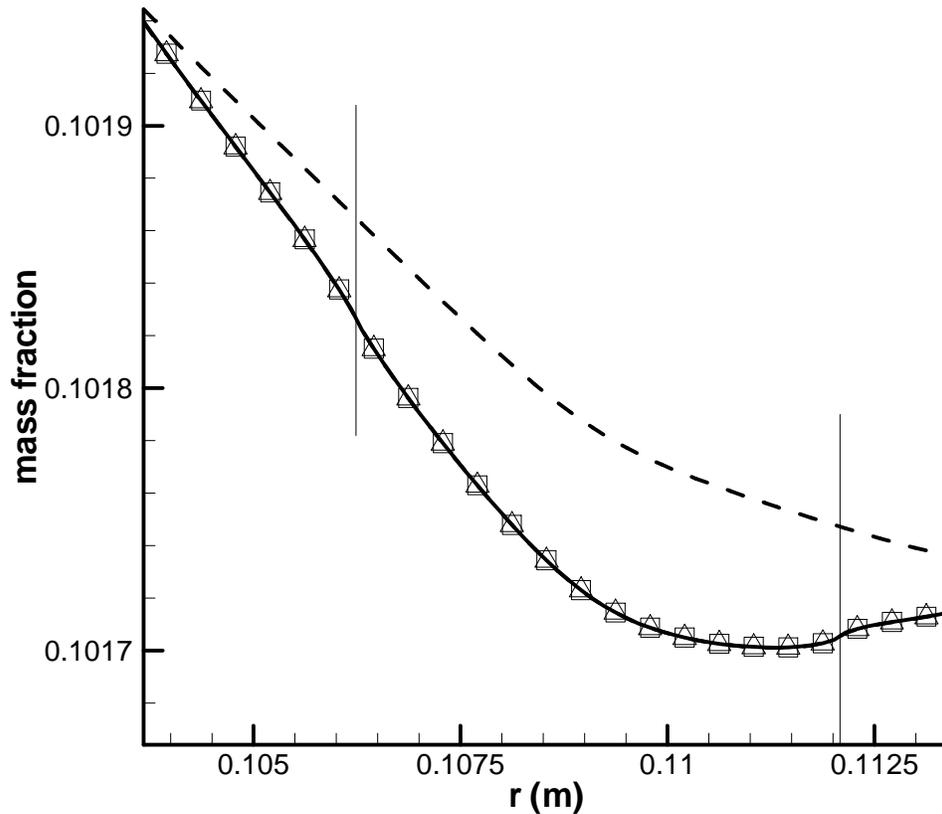


Figure 4-12 Variation of silane mass fraction on surface without features (- - -) and with features (—) for test case 1. [(Δ) Coupling at surface, (□) Coupling at source plane]. Vertical lines indicate reactor scale cell width.

4.2.1.2 Uniformly clustered features

Gobbert *et al.* ^[79] showed how the use of a meso scale allowed them to capture the influence of clusters of features which were too small to be accurately modeled using a two scale model. This second test case uses a similar clustering of features in an attempt to detect differences in the results produced by the two linking methods.

The same spacing and $0.1 \mu\text{m}$ by $1 \mu\text{m}$ features that were used in the first test case were also used in this second test case. Instead of having the entire surface related to the reactor scale cell covered with features, only three regions of the surface were coated with features. These regions or clusters were 0.7 mm long and spaced 0.7 mm apart.

Having clusters of features instead of blanketing the entire cell with features results in there being more leading and trailing edges for the coupling to affect. As a result the difference in the meso scale growth rate profiles became more evident as seen in Figure 4-13 with the transition effective growth rates changing by approximately 2% depending on the coupling method used. The impact on the reactor scale is the same, however, for both methods due to the averaging over the reactor scale cell width.

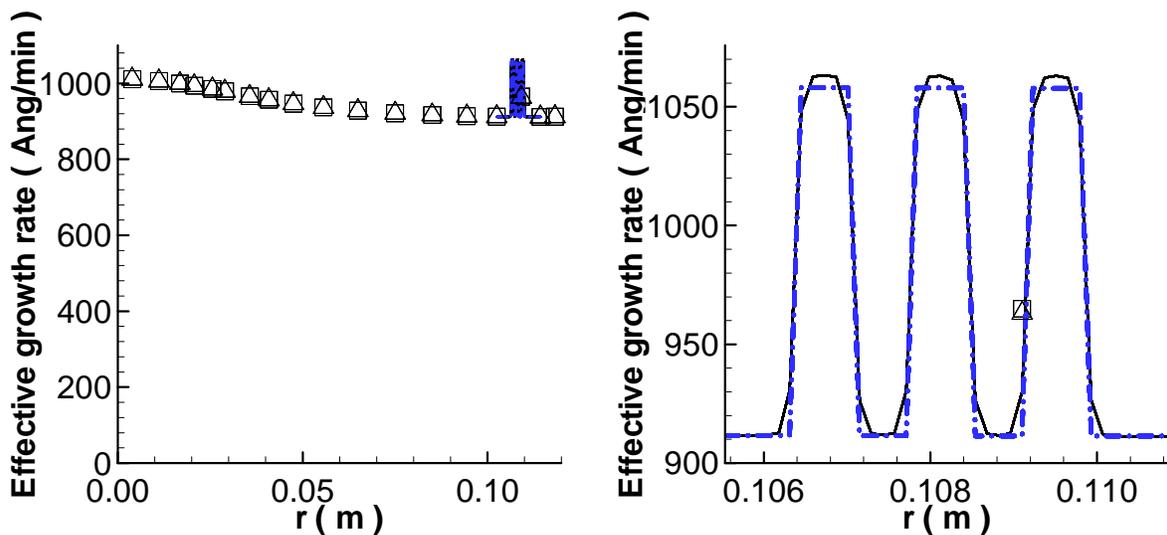


Figure 4-13 Growth rate results for uniformly clustered features within reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) Reactor scale with coupling at source plane, (---) Meso scale with coupling at surface, and (—) Meso scale with coupling at source plane.

In addition to the edge values changing the transition growth rates, the choice of method also affected the peak values of the effective growth rate. The reason for this has to do with how the area of the surface within each meso scale cell covered by features is calculated. For the offset coupling method this calculation is simple because an effective reactivity map is generated for the entire meso scale surface. This effective reactivity map takes into consideration the location of features relative to the meso scale control volumes. Without knowledge of the feature locations it is difficult to know how many features are located in each meso scale cell. During the application of the surface coupling for Figure 4-13 the impact of the features was calculated by generating an effective reactivity per unit area based on the ratio of the surface covered by features. Such that

$$r = \frac{.2\mu m \cdot r_{feature} + 11.8\mu m \cdot r_{flat_plate}}{12\mu m}, \quad (4-21)$$

since the features were spaced $12\mu m$ apart and the feature scale domain was $0.2\mu m$ across. This resulted in there being effectively 11.56 features per meso scale cell. However, there were actually 12 features per meso scale cell. Generating effective reactivity maps when surface coupling is used solves this difficulty since the number of features per meso scale cell can then be calculated.

4.2.1.3 Non-uniformly distributed features

The linking method of Gobbert *et al.*^[79] only dealt with features which were uniformly patterned across meso scale cells. No method was proposed for dealing with geometries which did not result in uniform patterning of features across the entire meso scale cell. The surface coupling method proposed in this study, however, does allow for non-uniform distribution of features even within a single meso scale cell. The approach proposed by Rodgers and Jensen^[3] using an offset source plane also does not have difficulty with non-uniform patterning of features. Their flexibility in feature placement results from superposition of the different ‘effective reactivity’ maps. In order to highlight the difference between coupling at the surface and coupling at an offset plane, a highly non-uniform distribution of features was selected. The distribution of features is shown in Figure 4-14. In addition to being non-uniform the features were located to one side of a meso scale cell, maximizing the spill-over of the ‘effective reactivity’ map into the adjacent meso scale cell. The resulting growth rate profiles are shown in Figure 4-15. While there is a 4.6% difference in the meso scale peak growth rates, the difference becomes approximately 0.17% at the reactor scale. This is due to the features only covering a small region of the surface. The impact of the meso scale concentration profiles is also small as seen in Figure 4-16. The differences in the concentration profiles produced by the two different methods shown in Figure 4-16 is significant when compared to the overall impact of the features of the third test case. However, the difference will not significantly affect surface deposition because it is only 0.001% of the concentration. While a 1×10^{-6} variation in the mass fraction seems small it is of the same order of magnitude as those found by Gobbert *et al.*^[79] in their study of clustered features.

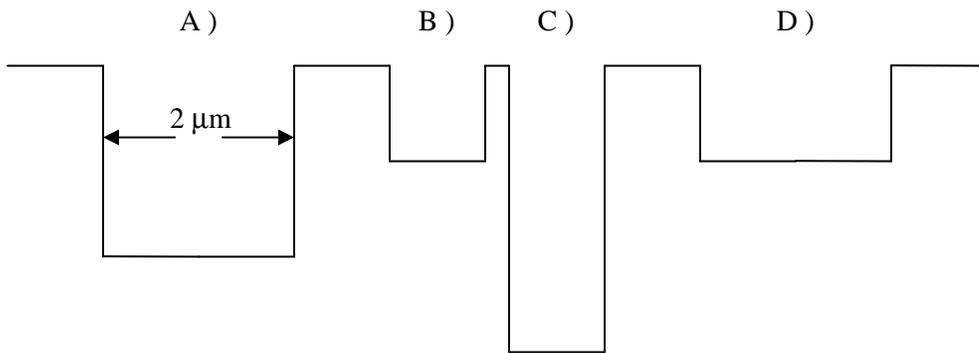


Figure 4-14 Feature scale geometry for test case with non-uniformly distributed features.

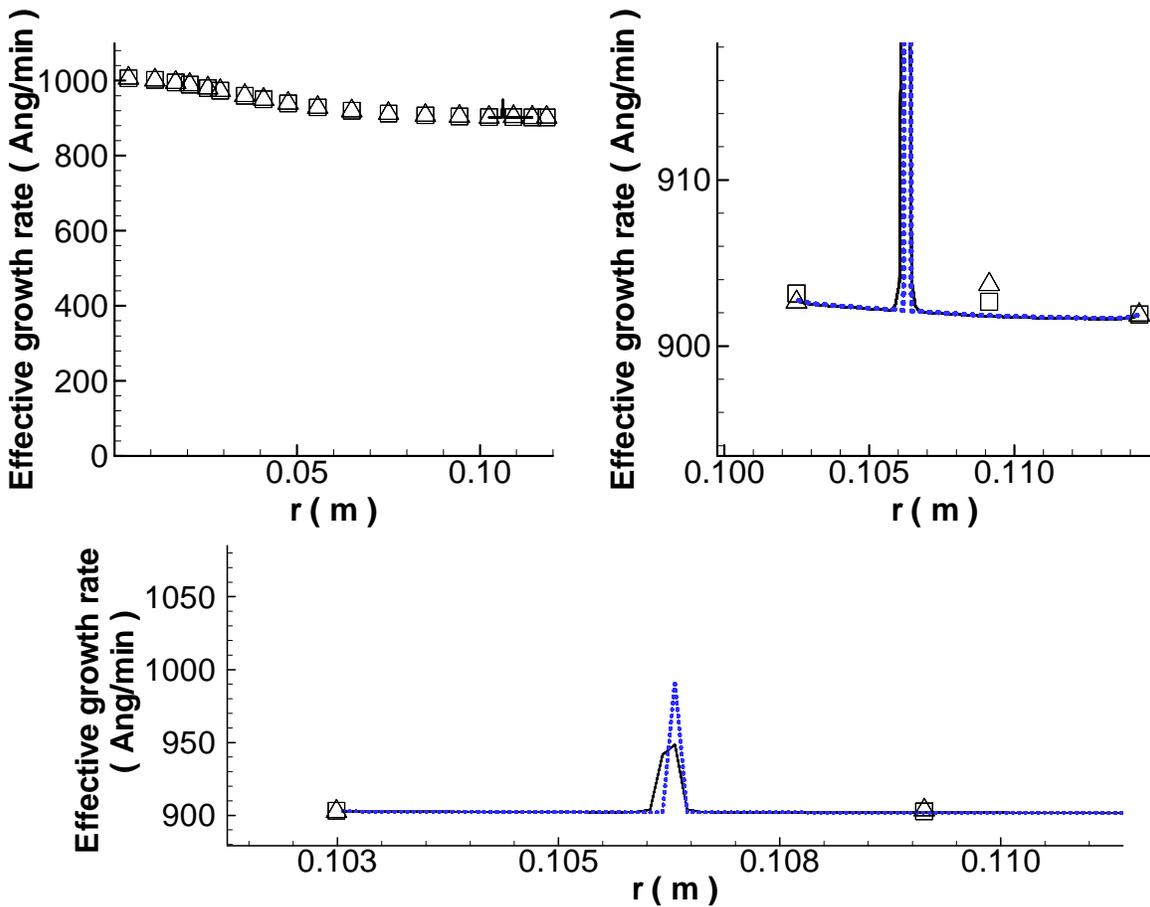


Figure 4-15 Growth rate results for non-uniformly distributed features within reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) Reactor scale with coupling at source plane, (- -) Meso scale with coupling at surface, and (—) Meso scale with coupling at source plane.

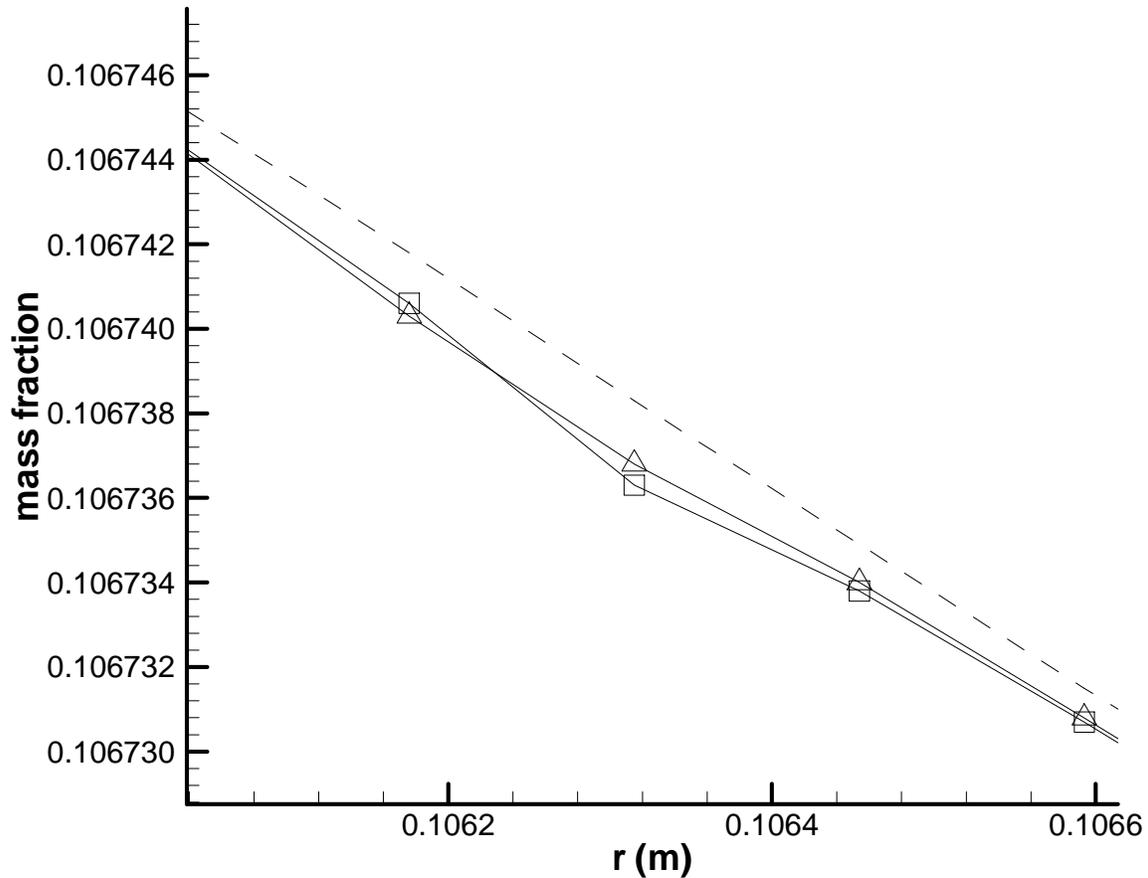


Figure 4-16 Variation of silane mass fraction on surface without features (- - -) and with features (—) for test case 3. (Δ) Coupling at surface, (□) Coupling at source plane.

4.2.2 Tungsten Deposition

The previous test cases were all for silane deposition. In order to test the generality of the observations made for the silicon deposition a test case with hydrogen reduction of tungsten was conducted. The same geometry that was used for the silane reactor was also used for the tungsten simulations. The conditions for this test case were an inlet flow rate of 0.2 slm, a wafer temperature of 750 K, and a reactor pressure of 1 torr, and an inlet gas temperature and wall temperature of 300 K. The inlet gas was composed of 89.55 mole percent [m/o] hydrogen, 9.95 [m/o] argon and 0.5 [m/o] tungsten hexafluoride.

The highly non-uniform deposition profile of Figure 4-17 shows that this is a very poor reactor design for this deposition process. This is most likely due to the tungsten hexafluoride having a much higher sticking factor than the silane gas with approximately 20% of the impinging molecules reacting with the surface. The higher sticking factor makes the tungsten hexafluoride deposition process much different than that of silane which is good for testing the generality of the observations made for the silane chemistry.

4.2.2.1 Uniform patterning of feature over a reactor scale cell

The single test case chosen to repeat with the tungsten hexafluoride chemistry was the case with features uniformly patterned over an entire reactor scale cell. The patterned features were infinite rectangular trenches with a width of 0.1 μm and a depth of 1 μm . The features were spaced 12 μm apart with flat areas separating the features. The reactor scale cell was 6 mm wide and was located so that its left edge was 106 mm from the center of the reactor. This is the same patterning of features applied with silane chemistry in Section 4.2.1.1.

Comparing the offset and surface coupling results for this test case it can be seen that the difference between the methods is small ($\approx 36\%$) as with the silane chemistry. However, the difference is more significant when compared to the effect of the presence of the features, with the difference being 28.7% of the total change in effective growth rate. The resulting difference in the mass fraction of tungsten hexafluoride was 4×10^{-6} [g/g].

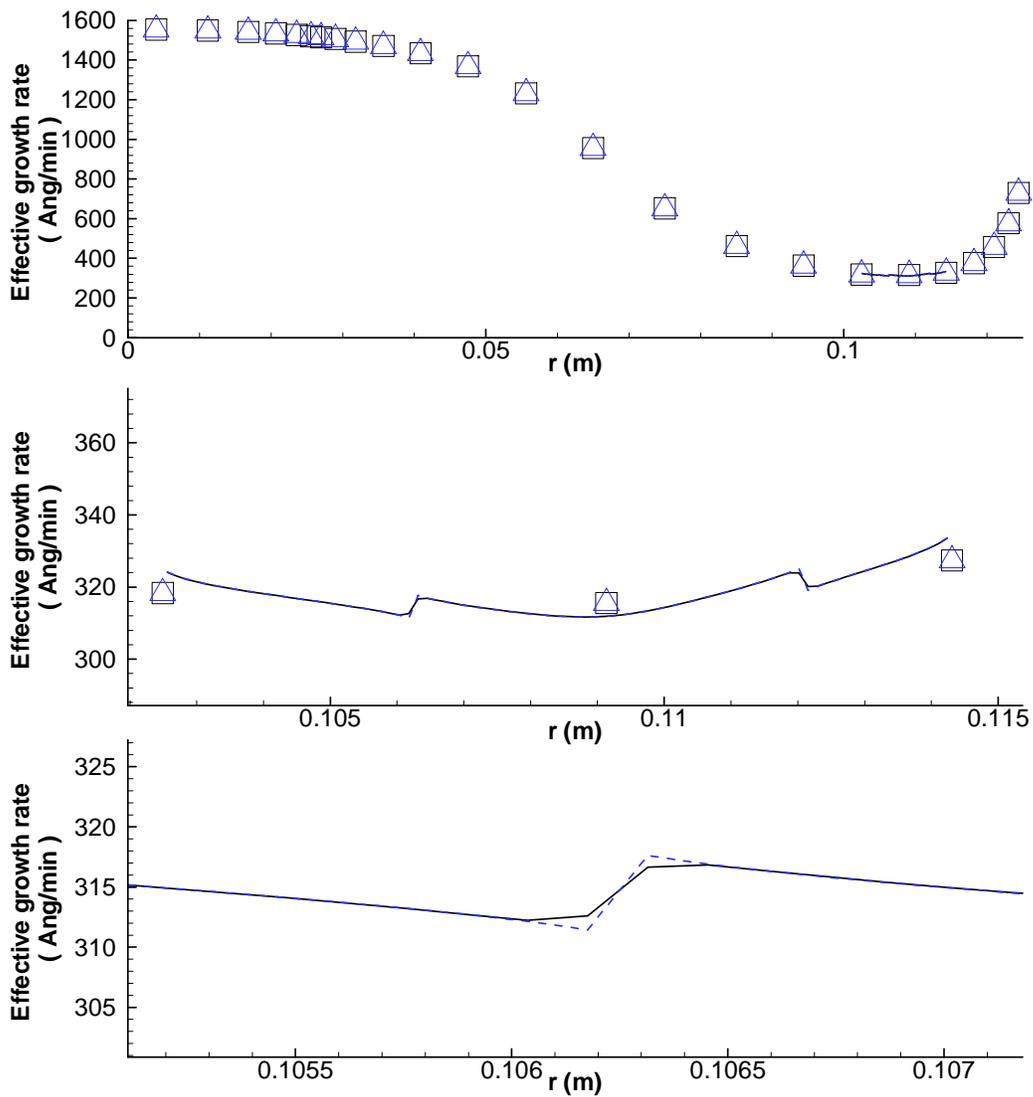


Figure 4-17 Tungsten growth rate results with features uniformly distributed over reactor scale cell. (Δ) Reactor scale with coupling at surface, (\square) Reactor scale with coupling at source plane, (- - -) Meso scale with coupling at surface, and (—) Meso scale with coupling at source plane.

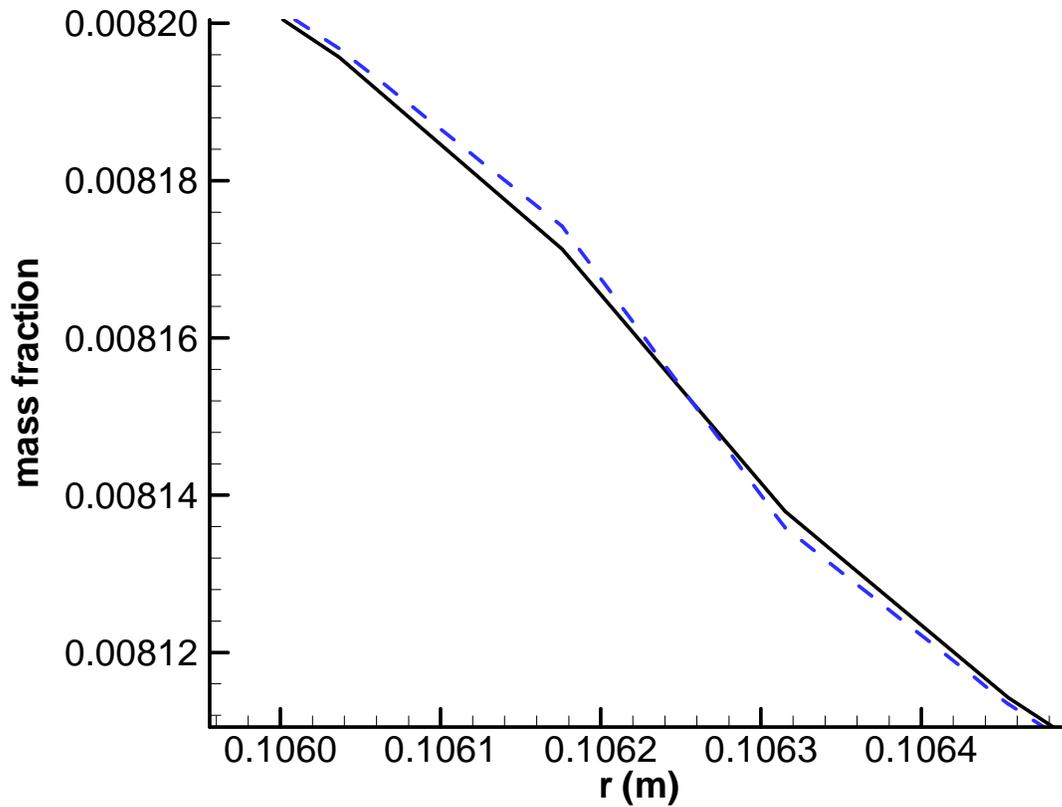


Figure 4-18 Variation of tungsten mass fraction on surface with coupling at surface (- - -) and with coupling at source plane (—).

4.2.3 Transient Results

Deposition within features changes the surface topography. For example Figure 4-19 shows how the geometry of the features from Section 4.2.1.3 change over time due to the deposition of silicon. This test case was chosen as an example because of the low number of features and the variety of features.

The low aspect ratio of the feature in Figure 4-19 A helps to prevent void formation. Figure 4-19's B, C and D results all show the possibility for void formation as deep narrow channels form prior to feature closure. While the features close the surface area available for deposition changes. The moment that the narrow channels close the surface area available for deposition drops sharply. This sudden change in surface availability has a noticeable affect on the effective growth rate as shown in

Figure 4-20. The sudden drops in the effective growth rate at 335 s and 665 s correspond to feature closures. The drop at 335 s is more drastic because both features B and C close at this time.

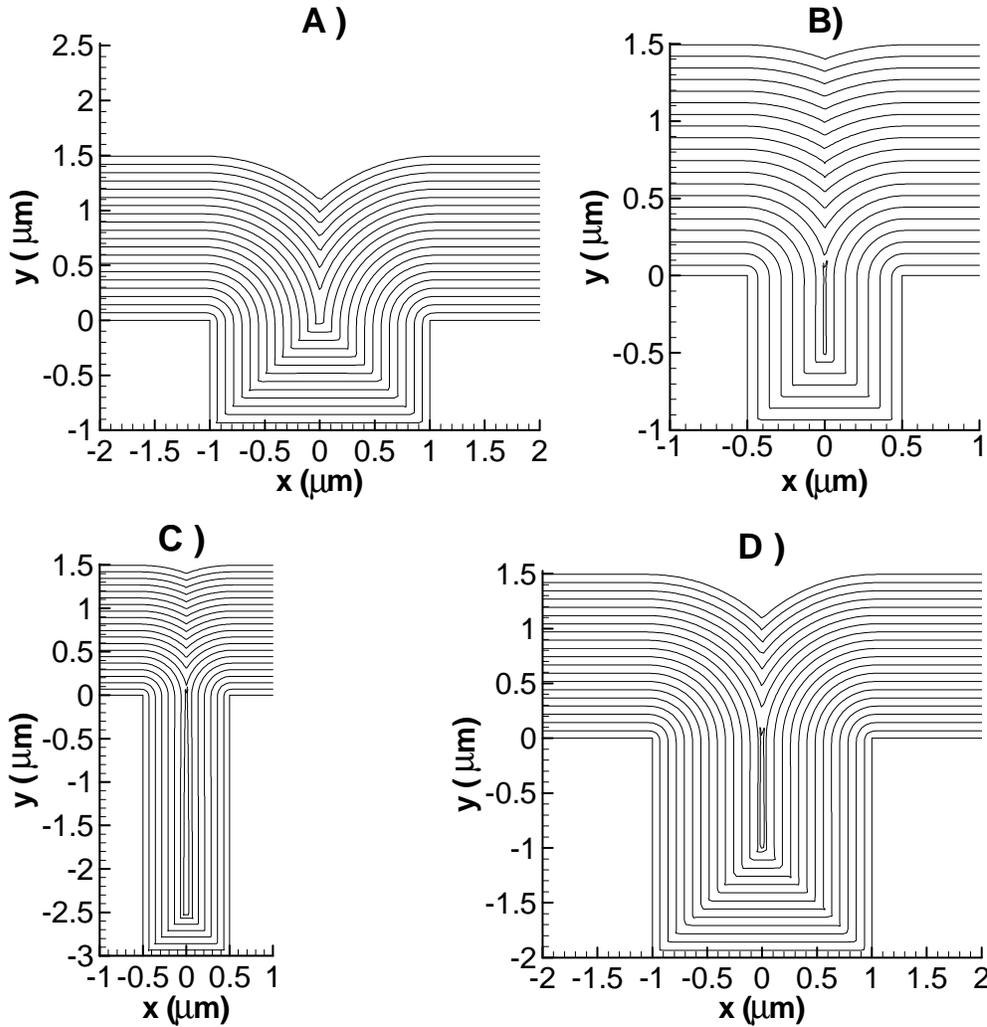


Figure 4-19 Profile of deposited silicon displayed at 50 s intervals for features distributed as shown in Figure 4-14.

The profiles shown in Figure 4-19 are representative of the result from both offset and surface coupling methods. The total difference in the deposition thickness for the two methods at the final time step of 1000 s is very small, approximately 5×10^{-12} m.

The effective growth rate profiles for both coupling methods are shown in Figure 4-20. The surface coupled results show a larger change in growth rate because the effect of the features is concentrated

into a single cell making their impact more noticeable. The effect of the different coupling methods was not strong enough to change the time of closure for the features. This meant that the jumps in the effective growth rate occurred at the same time step for both methods.

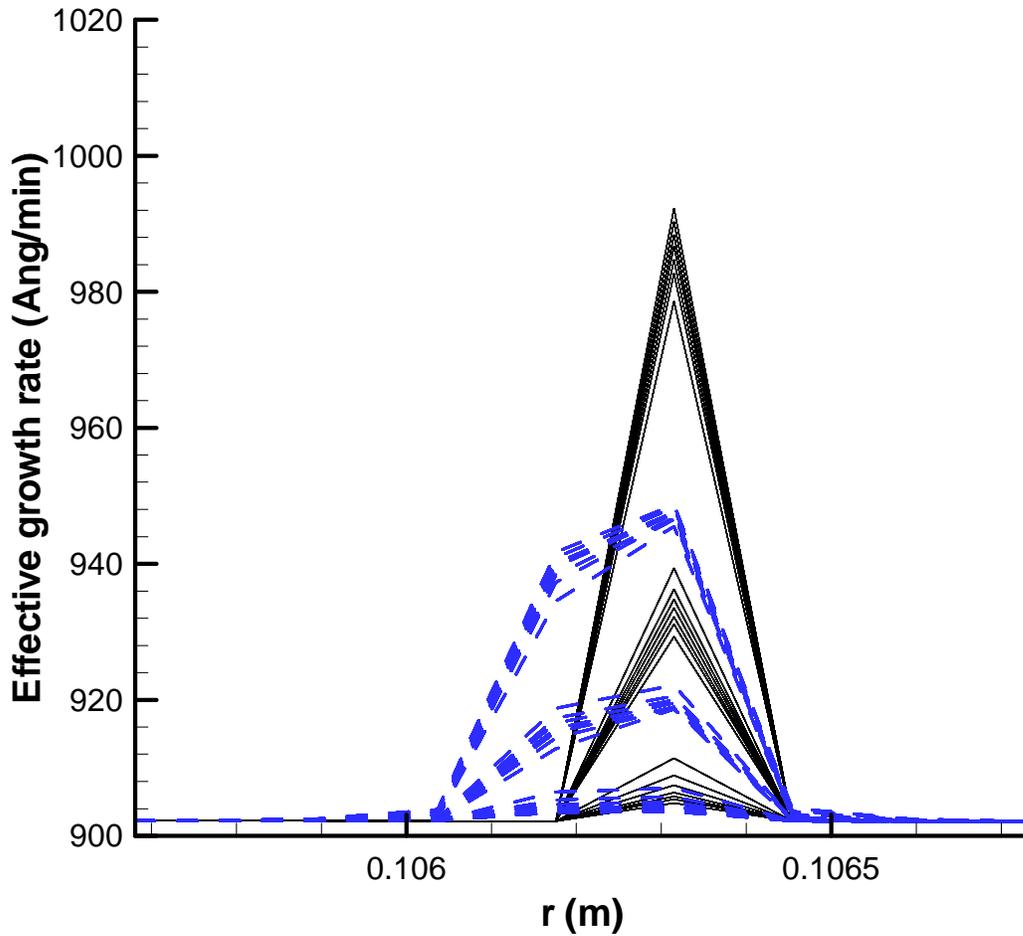


Figure 4-20 The reduction of the meso scale effective growth rate profiles displayed at 50 s intervals. [Offset coupling method (- - -), Surface coupling method (——)]

Changes to the effective growth rate also affect the species concentrations. Figure 4-21 shows the variation of the silane profile near the substrate as the features close when surface coupling is used. The results of the surface coupling method were chosen for this figure due spike in species concentration being more pronounced. Examining Figure 4-21 it can be seen that initially the concentration starts at the pseudo steady-state the values from Figure 4-16. Following the initial time

step there is a sudden drop in concentration similar to the drop reported by Merchant *et al.* [59] in their study of the deposition of silicon dioxide from tetraethoxysilane. After the sudden drop in concentration the concentration continues to drop until it reaches a minimum around 250 s. After reaching this minimum the concentration begins to increase. Coinciding with the closure of feature there are sudden increases in the species concentration. As the species concentration increases, approaching the flat plate concentration levels, the spike associated with the presence of features decreases, becoming smooth after all of the features have closed.

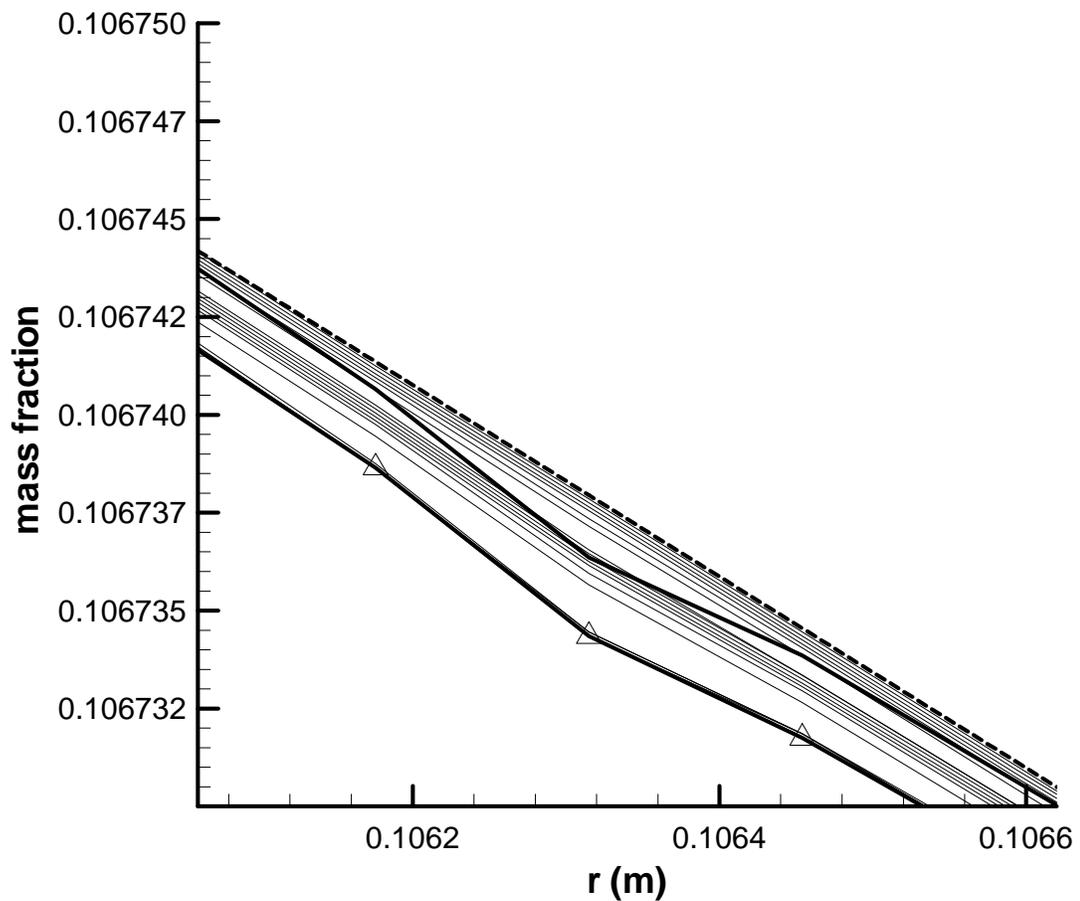


Figure 4-21 Silane mass fraction profile close to substrate surface displayed at 50 s intervals. [T = 0 s (thick solid line), T = 250 s (thick solid line with Δ), T = 1000 s (dashed line)]

4.3 Discussion

In previous sections the presence of features has been shown to have a significant impact on the effective growth rate. However, there was almost no change to the growth of the deposition within the features as a result of these changes in effective growth rate. In order to understand why this is a broader view of the problem must be taken. Changes to the effective growth rate affect the growth of the features through their impact on the species concentrations. Large effective growth rates indicate an increase in the consumption of depositing species. However, the impact of features is fairly local and reactant molecules from the surrounding gas diffuse towards the sink reducing the impact of the increased effective reaction rate on the species concentration. Any reduction in the local concentration of species increases the gradient with the surrounding gases and increases the rate with which species diffuse into the affected reactor or meso scale cell. This balancing effect is present in Equation (2-32) from the reactor scale section.

Looking at this problem from the molecular perspective the distribution of particles leaving across the source plane into the source volume is greatly affected by the presence of features. However, particles leaving across the source plane undergo several collisions before returning (if they return) across the source plane. These collisions smooth out the distribution of particles as noted by Rodgers and Jensen^[3]. A large number of features is required to substantially affect the source volume and the distribution of particles entering the feature domain across the source plane.

Another reason for the limited impact of the effective growth rate on the feature growth rates is the relative insensitivity of the feature growth rate to changes in the reactant mass fraction. Figure 4-22 shows the dependence of growth rate on the mass fraction of silane and tungsten hexafluoride.

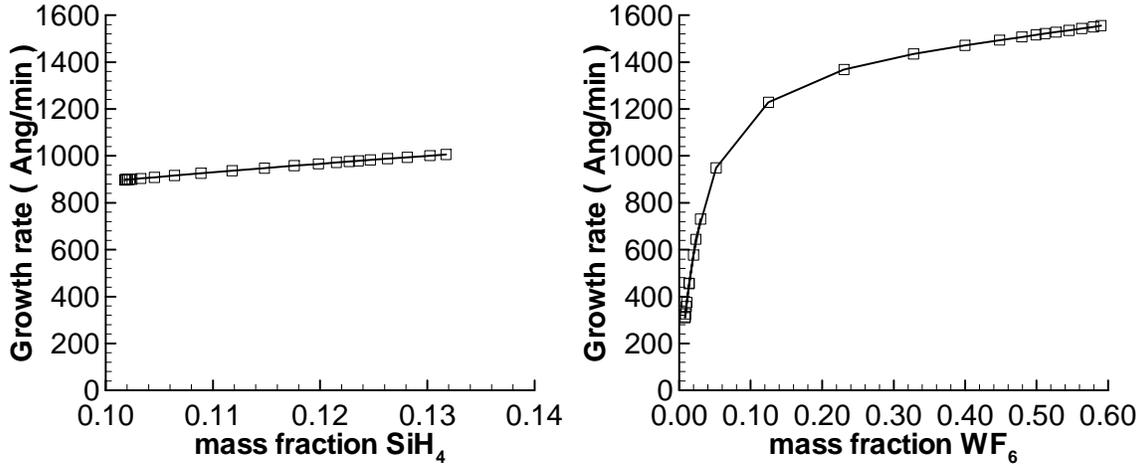


Figure 4-22 Variation of growth rate with mass fraction over the substrate surface for silane gas chemistry (Left) and tungsten hexafluoride (Right).

The dependence of the growth rate for silane and tungsten on their concentrations are different as seen in Figure 4-22. The region where the simulated features are located on the substrate also has lowest depositing gas mass fraction. The dependence of the growth rate on mass fraction at this low mass fraction is approximately 3700 [Ang/min]/[g/g] for silane deposition and 22800 [Ang/min]/[g/g] for tungsten hexafluoride deposition. While these values may seem large they are not large enough for small groups of features to significantly impact the growth rate. This is because the changes in the mass fraction due to the presence of the features are much less than 1 [g/g]. For example the 4×10^{-6} [g/g] change in tungsten hexafluoride mass fraction due to the use of different coupling methods visible in Figure 4-18 only results in a change of 0.09136 [Ang/min] in the actual growth rate. This is fairly insignificant since the overall tungsten growth rate at that location was approximately 309 [Ang/min]. Further the largest variation in silane concentration for the silane test cases is 1×10^{-4} [g/g]. This is in Figure 4-12 where the mass fraction profile of a reactor scale cell covered with features is compared with that of a reactor scale cell with no features. Varying the mass fraction of silane by 1×10^{-4} [g/g] results in a change of the actual growth rate of 0.366 [Ang/min], which is also insignificant since the growth rate at that location is approximately 900 [Ang/min].

If a change of 1% in the growth rate was deemed to be a significant impact this would require the mass fraction of silane to change by 0.00243 [g/g] and tungsten hexafluoride by 0.000136 [g/g]. In the silane reactor this change represents approximately the difference between the silane mass

fraction at the center of the reactor and that at the outer edge of the substrate. This means that if only one way coupling were used, two feature scale simulations would be sufficient to accurately predict the variation in deposition along the substrate surface.

While the results of the test cases in the results section did not show the presence of features having a large enough impact on the concentration to significantly affect the closure time or the deposition topography this is not the case in general. The impact of the features can be increased by either increasing their density or by increasing the area which they cover. Increasing the impact of the features in this way can result in them having a significant enough impact on the concentration to affect closure times and deposition topography. For example Figure 4-23 shows what the concentration profile would have looked like if the entire substrate were coated in features in the same manner as the cell was in Section 4.2.1.1. Comparing this profile with the feature-less profile, it can be seen that the presence of the features results in a 0.003 [g/g] drop in the concentration of silane at the surface. This is a significant change in the concentration and would result in a more than 1% change in the growth rate demonstrating the requirement of two-way coupling.

The conditions under which the presence of features can significantly impact the feature growth rate do not coincide with those which highlight the difference between the two linking methods.

Increasing the density of features greatly increases the gradient of the ζ function, thereby increasing the cell width required to satisfy the Knudsen number constraint. Furthermore, the difference in the concentration profiles of the two different coupling methods become less significant as more of the surface is covered with features. This means that while the offset coupling method can produce greater insight into the effect of features on the concentration profile, the level of accuracy required for predicting changes in the surface topography can be attained using either coupling method.

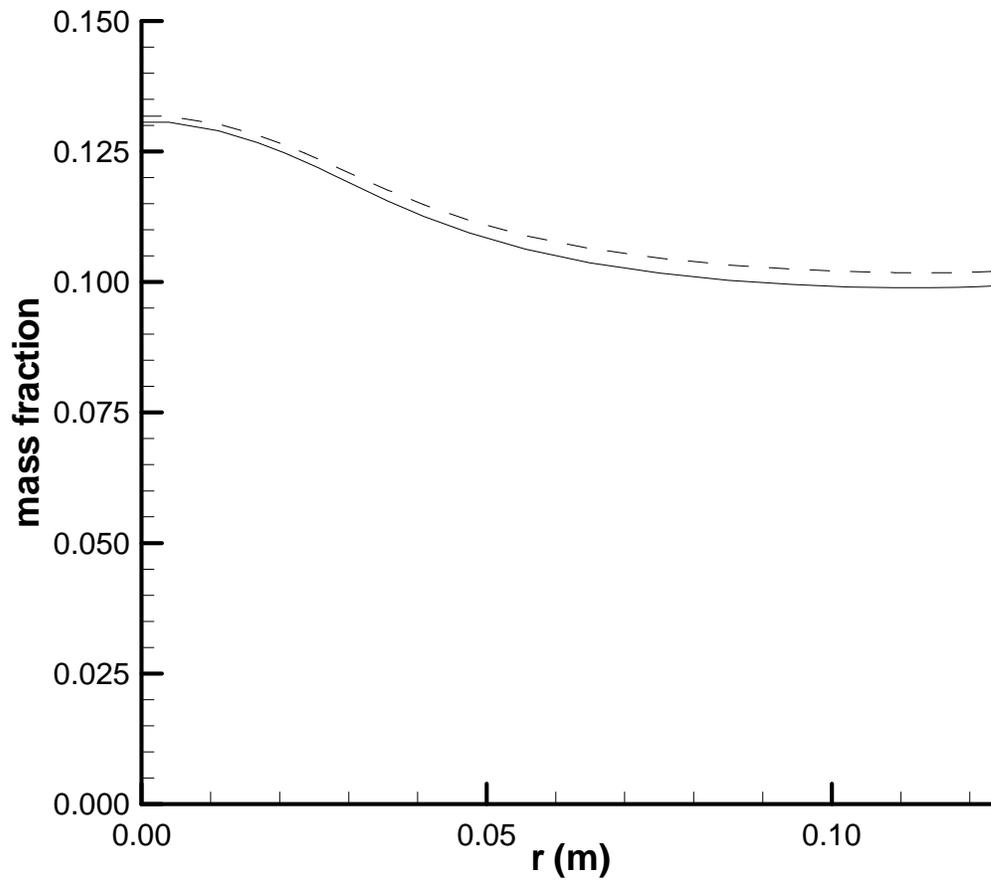


Figure 4-23 Silane mass fraction profiles for substrate uniformly covered in features (—) compared with that of a featureless substrate (---).

4.4 Closure

In this chapter coupling between reactor and feature scale models was investigated. The use of a meso scale to reduce the separation of scales was discussed and it was shown that the inclusion of a meso scale did not affect the results at the reactor scale. Two different methods of coupling reactor/meso and feature scales were introduced. One method coupled the scales at the surface while the other used an offset plane as the interface between the two scales.

Both coupling methods were based on previous methods but were novel in their implementation. The offset coupling method was particularly novel because this type of coupling had never been performed with a ballistic based feature scale model.

Comparing the two models showed that while both methods produced sufficiently accurate results for predicting closure times and topography changes in the features, the offset method produced more realistic concentration profile details. Due to the additional costs associated with the offset coupling method the surface coupling method is more practical for most simulations. However, the offset method also had the additional benefit of allowing for much finer grid resolution than the surface coupling method.

The major difference between the surface coupling method used in this study and that of previous studies is the use of an effective reactivity function. These effective reactivity functions are generated by summing the $\xi \cdot dx$ values of individual features. The use of these maps during coupling allows non-uniformly distributed features to be modeled and ensures that the correct numbers of features are assigned to each reactor/meso scale cell. Incorrect distribution of features can result in loss of accuracy as seen in Section 4.2.1.2.

Chapter 5

Conclusions and Contributions

The chemical vapor deposition process has many different characteristic lengths because of the multitude of phenomenon which impact the quality of the film deposition. These characteristic lengths range from the order of meters at the reactor scale to angstroms at the atomic scale. The large range of length scale makes it impossible for a single model to be used to simulate all of the significant phenomenon. Instead multiple different models are used and coupled together to form a heterogeneous multi-scale model. The focus of this thesis is the coupling between the reactor scale model which accounts for the impact of the reactor geometry on the deposition process and the feature scale model which accounts for the impact of feature geometries.

Up until now there have been two main coupling methods used to pass information between the continuum based reactor scale and the molecular transport based feature scale. These two methods were a method proposed by Gobbert *et al.*^[2] where coupling took place at the deposition surface and a method proposed by Rodgers and Jensen^[3] where coupling took place at an offset plane. The Gobbert *et al.*^[2] method was used in conjunction with ballistic transport based feature scale models, while the coupling method of Rodgers and Jensen^[3] only worked with MC based feature scale models.

Since the offset surface coupling method was not designed to work with ballistic transport based feature scale methods, no direct comparison of these two methods was conducted. Instead the choice of coupling method was determined by the choice of feature scale modeling. The offset coupling method did seem to have some advantages over the surface coupling method because the effective reactivity maps of the offset coupling method allowed for the use of superposition for determining the combined effect of differently shaped features^[3]. The surface coupling method, however, was only applicable when entire reactor/meso scale were covered with uniformly patterned features^[75]. In the following section the results, conclusions and contributions of this thesis are summarized.

5.1 Conclusions and Contributions

In Chapter 3 the development of a new feature scale model, BTLSF is discussed. The main unique feature of BTLSF is that it uses a source plane in the calculation of initial deposition instead of the

hemispherical source used by other model. This helps with the calculation of effective reactivity maps because the interface for flux entering and leaving the feature scale domain is the same. These effective reactivity maps allow for coupling similar to that of Rodgers and Jensen^[3], to be applied to a modified line of sight model.

It is also shown in Chapter 3 that the number of iterations required in order to reach the surface flux convergence criteria with direct quadrature increases as the features close. The ability to use this dependence to automate the selection of time step size is then demonstrated. By applying this method to a double stepped feature, the benefit of having the time step tied to the number of direct quadrature iterations is demonstrated.

During investigation of the application of BTLFSF to non-first-order chemistries the new model demonstrated the ability to model features with aspect ratios as high as 100. BTLFSF was also shown to be able to model more complex chemical models with more realistic kinetics such as the Eley-Rideal kinetics.

The sensitivity of the surface deposition predictions of BTLFSF to source plane height is also investigated. This investigation showed that when the flux crossing the source plane is uniform the source plane height has very little effect on the surface flux distribution. However, if the flux crossing the source plane was non-uniform the source plane height is expected to have more impact as with the return flux distribution discussed in Chapter 4.

Chapter 4 focused on the coupling between reactor and feature scales. Two methods of coupling were considered the first using an offset plane as the coupling interface and the second using the surface of the wafer. Comparing these two methods it was found that they produced similar results. The offset method produced more convincing effective reactivity profiles and provided better insight into physical processes. However, for most applications the simpler surface coupling method would suffice.

The limitations on the coupling methods were also investigated. It was found that in order to couple with the continuum based model the Knudsen number associated with the gradients in the effective reactivity profile would have to be less than one fifth of the inverse of the mean free path. For the

surface coupled method this meant that the change in the effective reactivity from one cell to the next at the meso or reactor scale could not be more than one fifth of the ratio between the cell width and the mean free path. Since the effective reactivity for the surface coupling method is fixed by the geometry this effectively limits the minimum valid cell size used for the continuum base model.

The effect of the Knudsen number restriction is different for the offset coupling method because the shapes of the effective reactivity functions are different. While the shape of the effective reactivity functions for the surface coupled method are like a rectangular function, the shape of the effective reactivity maps for the offset method are more similar to a Dirac delta function. This means that while the slope of the surface coupled effective reactivity function is infinite the slope of the offset coupling method is finite. If the finite slope of the offset coupling method is less than one fifth of the inverse of the mean free path then the Knudsen number restriction is satisfied for all meso and reactor cell widths.

The factors which affected the slope of the effective reactivity map for both methods were investigated. It was found that for the offset coupling method the factors which affect the slope of the effective reactivity map are the source plane height, the sticking factor as well as the aspect ratio, width, and spacing of the features. Increasing both the source plane height and the sticking factor decreases the slope of the effective reactivity. Increasing the spacing between features can also decrease the slope of the effective reactivity. While increasing the aspect ratio or width of the features increases the slope, however, in a realistic application the feature spacing, aspect ratio and width would all be fixed. The source plane height is not fixed and can be used to decrease the slope of the effective reactivity function. The height of the source plane is limited though to around one third of the mean free path because of the need for collisionless flow in the gas phase of the feature scale. This means that if the slope of the effective reactivity is too large to satisfy the Knudsen number limitation then the only factor which can be changed to allow coupling with the continuum based scale is the meso or reactor scale cell width. With larger reactor scale cell widths the difference between the coupling methods is lost during the homogenization process.

Since the slope of the effective reactivity function associated with the surface coupling method is infinite the only factor which can be controlled to ensure that the Knudsen number limit is satisfied is the cell width. However, the minimum cell width which can be used is affected by factors similar to

those which affected the slope of the effective reactivity for the offset coupling method. Examining the effects of sticking factor, aspect ratio, feature width, and feature spacing on the minimum cell width it was found that the minimum cell width decreased with increasing sticking factor and feature spacing. Aspect ratio was found not to impact the minimum cell width and increasing the feature width increased the required cell width.

While normally the feature's spacing and geometry would be predetermined the geometry used in the test cases of Chapter 4 were specifically selected in order that the slope of the effective reactivity could be resolved by the offset coupling method without restricting the meso scale cell width. The geometry was selected in this way to maximize the difference between the two methods. However, selecting this resolvable geometry also meant that the impact of the features on the concentration was greatly reduced.

Testing with silane deposition chemistry showed that the difference between the two methods was most evident when the patterning of features changed. It was also found that the use of a reactivity map was important with both methods. This ensures that the right number of features are included in each meso scale cell. The test case which showed the most difference between the methods was the non-uniform distribution where the features were clustered to one side of the meso scale cell. The difference between these two methods was still too small to significantly affect the profile evolution even after 1000s.

Applying the coupling methods to tungsten deposition chemistry showed similar results to those found with the silane chemistry. However, the higher sticking factor associated with the tungsten chemistry created the possibility that the feature spacing could have been reduced to a more realistic value.

The significance of ignoring the component of the velocity resulting from the deposition of mass on the surface when selecting the distribution for particles entering the feature scale domain was investigated. It was found that for the temperatures and reaction rates involved in this study the use of a Maxwell distribution was acceptable.

5.2 Recommendations for future work

While having collisionless flow at the feature scale greatly reduces the computational cost it also makes it difficult to conceptually link feature and reactor scales. Without collisions there can be no gas phase reactions, thermal diffusion or concentration diffusion. It is also hard to explain the connection between the particle based models and the continuum properties such as viscosity. Without gas phase collisions there is also no way for particles exiting the source plane to return. In a way linking with the reactor or meso scale does allow particles to return across the source plane. However, the scale separation between these two models is so large that too much information is lost to detect such small things as feature to feature effects. One possible solution may be to include an intermediate model between the feature and reactor scales. This model would have to be applicable for Knudsen numbers near 1. Most models with this capability are fairly computationally costly but the geometry of this intermediate scale will be very simple because the complex feature geometries would be taken care of by the feature scale model.

Another probably more important need for further work is the application of the current model to more complex chemistries with gas phase reactions. The reason for this is gas phase reactions can produce radicals which have much higher sticking factors than those of the main species. The small differences in the concentration distributions which result from the choice of surface versus offset coupling may be more important in these cases. This is because while the concentrations of radical species are very small, they can have a large impact on the topography resulting from deposition^[76].

There are many feature scale geometries which require a three dimensional code to accurately predict the way in which their topography will grow^[81]. The feature scale model developed to facilitate coupling with a source plane in this study does not easily allow for three dimensional simulation of feature filling. However, the source plane coupling method discussed in this thesis could be applied to three dimensional geometries. This would require calculation of the return flux across a source plane and would result in an effective reactivity surface rather than profile. The reactor scale would also be required to be three dimensional in nature. This truly three dimensional multi-scale model would be able to simulate very realist geometries including the patterning of dies on the wafer surface.

In order to verify the accuracy of the BTLFSF model comparison with experimental results is required. The two-dimensional limitation of the BTLFSF model would necessitate a very specialized test setup

in order to minimize three dimensional effects at the feature scale. In the experiment the benefits of using an offset plane for the calculation of the initial deposition could be highlighted by having a large concentration gradient parallel to the feature cross section and examining growth rates within the feature. This would force the use of numerical integration of Equation (3-8).

Appendices

Appendix A

Calculation of transport properties

Many transport properties are dependant on temperature and gas composition. In this study these dependencies are accounted for in two different ways. The first method was through the use of curve fits to experimental data. This method was used for all the test cases involving silane gas, with the fitting relationships and constants being taken from Kleijn *et al.*^[19]. The second way in which the temperature and species concentration dependencies were accounted for in this study was through the use of kinetic theory. Kinetic theory was used to calculate the transport properties for the tungsten deposition test case.

The curve fit used to approximate the experimental values for the temperature dependent viscosity, thermal conductivity, and specific heat capacity is

$$\phi = C_0 + C_1 \cdot T + C_2 \cdot T^2 \quad (\text{A-1})$$

where ϕ is either μ , λ , or c_p . Table A-1 shows the values used for fitting constants C_0 , C_1 , and C_2 . The concentration dependence of these properties also had to be taken into account. For the specific heat capacity this was done through a simple weighting based on the mass fractions of the different species so that

$$c_p = \sum_{k=1}^N Y_k c_{pk} \quad (\text{A-2})$$

The calculation of the mixture viscosity and thermal conductivity was more complex. In to be consistent with Kleijn *et al.*^[19] the method for calculating the viscosity and thermal conductivity recommended by Bretsznajder^[82] was used. This meant that

$$\mu = \sum_{k=1}^N \frac{f_k \mu_k}{\sum_{j=1}^N f_j \Phi_{kj}}, \quad (\text{A-3})$$

where

$$\Phi_{kj} = \frac{1}{\sqrt{8}} \left(1 + \frac{m_k}{m_j} \right)^{-1/2} \left(1 + \left(\frac{\mu_k}{\mu_j} \right)^{1/2} \left(\frac{m_k}{m_j} \right)^{1/4} \right)^2$$

(A-4)

and that

$$\lambda = \alpha \sum_{k=1}^N \lambda_k f_k + (1 - \alpha) \left(\sum_{k=1}^N \frac{f_k}{\lambda_k} \right)^{-1}$$

(A-5)

The value used in this study for α was 0.80, however, Kleijn *et al.* ^[19] stated that the value could vary between 0.32 and 0.80 depending on the mole fraction of light gasses in the mixture.

Table A-1 Fitting constants for transport properties^[19].

Property	Gas	C ₀	C ₁	C ₂
μ	SiH ₄	1.47E-06	3.66E-08	-6.81E-12
	N ₂	5.73E-06	4.37E-08	-9.28E-12
	H ₂	3.11E-06	2.06E-08	-3.54E-12
λ	SiH ₄	-2.12E-02	1.45E-04	-1.31E-08
	N ₂	8.54E-03	6.23E-05	-4.34E-09
	H ₂	1.05E-01	3.21E-04	-2.50E-09
C_P	SiH ₄	4.74E+02	3.26E+00	-1.08E-03
	N ₂	9.83E+02	1.58E-01	1.69E-05
	H ₂	1.47E+04	-1.01E+00	1.29E-03

The binary diffusion coefficients for the different gas couples were dependent on both the pressure and the temperature. The equation used to approximate this dependence was

$$D_{kj} = D_{jk} = (C_0 + C_1 \cdot T + C_2 \cdot T^2) / P,$$

(A-6)

with the fitting constants being given in Table A-2.

Table A-2 Fitting constants for binary diffusion coefficient^[19].

Property	Gas Pair	C ₀	C ₁	C ₂
D_{kj}	SiH ₄ , N ₂	-9.64E-01	6.25E-03	8.20E-06
	SiH ₄ , H ₂	-2.90E+00	2.06E-02	2.81E-05
	N ₂ , H ₂	-3.20E+00	2.44E-02	3.37E-05

The thermal diffusion factor, a_{kj} is highly dependent on species concentration as well as temperature.

The fitted curve used to account for this was

$$a_{kj} = -a_{jk} = C_0 f_k f_j (1 + C_1 f_k + C_2 f_k^2 + C_3 f_k^3) (1 + C_4 \exp(C_5 T)) \quad (\text{A-7})$$

and the fitting constants required for this equation are provided in Table A-3.

Table A-3 Fitting constants for binary thermo diffusion ratios^[19].

Gas Pair	C ₀	C ₁	C ₂	C ₃	C ₄	C ₅
N ₂ , SiH ₄	-5.15E-02	1.86E-01	2.94E-02	4.43E-03	-1.69E+00	-4.94E-03
H ₂ , N ₂	-2.71E-01	7.27E-01	-3.57E-01	9.87E-01	-1.61E+00	-9.15E-03
H ₂ , SiH ₄	-2.74E-01	1.01E+00	-1.04E+00	1.90E+00	-1.70E+00	-6.35E-03

Since the curve fitting data for the transport properties of the gas mixture used in the tungsten deposition test case was not readily available the properties had to be calculated through the use of kinetic theory. This process was far more involved than that used in for the silicon deposition test cases and required information about the Lennard-Jones force parameters of the gases. These parameters are given in Table A-4.

Table A-4 Molecular mass and Lennard-Jones force parameters for tungsten deposition test case^[11].

Species	m [kg/mole]	σ [Ang]	ϵ/k [K]
Ar	0.03994	3.542	93.3
H ₂	0.00202	2.827	59.7
HF	0.02001	3.138	330.0
WF ₆	0.29784	5.210	338.0

Using the Lennard-Jones force parameters the viscosity can be calculated as

$$\mu_k = 8.441 \cdot 10^{-5} \frac{(m_k T)^{1/2}}{\sigma_k^2 \Omega_\mu}, \quad (\text{A-8})$$

where σ_k is the collision diameter and Ω_μ is a dimensionless collision integral function. The values for this function along with the other dimensionless collision integral functions, Ω_D , A^* , B^* and C^* are dependant on the reduced temperature, T^* and were interpolated from table values found in the appendix of Hirschfelder *et al.* ^[83]. The reduced temperature is

$$T^* = kT / \varepsilon_k, \quad (\text{A-9})$$

where ε_k is the maximum energy of attraction. The values of ε_k / k are one set of Lennard-Jones force parameters supplied in Table A-4, where k is Boltzmann's constant. The collision diameters are also found in Table A-4.

Once the individual species viscosities were calculated they were combined using Equation (A-3) in the same way as they were for the silicon deposition test cases. However, the method used to combine the individual species thermal conductivities was different from the method used in the silicon test cases with

$$\lambda = \frac{\sum_{k=1}^N f_k \lambda_k}{\sum_{j=1}^N f_j \Psi_{kj}} \quad (\text{A-10})$$

and

$$\Psi_{kj} = \frac{1}{\sqrt{8}} \left(1 + \frac{m_k}{m_j} \right)^{-1/2} \left(1 + \left(\frac{\lambda_k}{\lambda_j} \right)^{1/2} \left(\frac{m_k}{m_j} \right)^{1/4} \right)^2. \quad (\text{A-11})$$

This method is more consistent with the method used for combining the viscosities and removes the uncertainty related to the α term in Equation (A-5). The individual species thermal conductivities were calculated from

$$\lambda_k = \frac{15}{4} \frac{R}{m_k} \mu_k. \quad (\text{A-12})$$

The binary diffusion coefficients of the species were calculated using the equation suggested by Wilke and Lee^[84],

$$D_{kj} = \left[6.77 - 0.0492 \left(\frac{m_k + m_j}{m_k m_j} \right)^{1/2} \right] \cdot 10^{-4} \frac{T^{3/2}}{P \sigma_{kj}^2 \Omega_D} \left(\frac{m_k + m_j}{m_k m_j} \right)^{1/2}. \quad (\text{A-13})$$

The collision diameter, σ_{kj} is found from the values provided in Table A-4 with

$$\sigma_{kj} = \frac{1}{2} (\sigma_k + \sigma_j). \quad (\text{A-14})$$

The calculation of the thermal diffusion factors using kinetic theory is very complex due to the strong dependence on the assumed intermolecular potential. If molecules were approximated as rigid elastic spheres then Ω_μ , A^* , B^* and C^* could have been approximated as being unity. However, this would have likely resulted in an over estimation of the thermal diffusion ratio. Instead the values of Ω_μ , A^* , B^* and C^* were interpolated from tables in the appendix of Hirschfelder *et al.*^[83]. In order to do this the reduced temperature for the gas pairs had to be calculated with

$$T_{kj}^* = kT / \epsilon_{kj} \quad (\text{A-15})$$

with

$$\epsilon_{kj} = (\epsilon_k \epsilon_j)^{1/2}. \quad (\text{A-16})$$

With the interpolated collision integral functions the binary thermal diffusion factors could now be calculated as

$$a_{kj} = \frac{1}{6\Theta_{kj}} \frac{S_{kj}f_k - S_{jk}f_j}{X_\lambda + Y_\lambda} (6C_{kj}^* - 5) \quad (\text{A-17})$$

with

$$S_{kj} = \frac{\Theta_{kj}}{\Theta_k} \left(\frac{m_k + m_j}{2m_j} \right) - \frac{15}{4A_{kj}^*} \left(\frac{m_j - m_k}{2m_k} \right) \quad (\text{A-18})$$

$$\Theta_{kj} = 0.00263 \frac{\left(T \frac{m_k + m_j}{2m_k m_j} \right)^{1/2}}{\sigma_{kj}^2 \Omega_\mu} \quad (\text{A-19})$$

$$\Theta_k = 0.00263 \frac{\left(\frac{T}{m_k} \right)^{1/2}}{\sigma_k^2 \Omega_\mu} \quad (\text{A-20})$$

$$X_\lambda = \frac{f_k^2}{\Theta_k} + \frac{f_k f_j}{\Theta_{kj}} + \frac{f_j^2}{\Theta_j} \quad (\text{A-21})$$

$$Y_\lambda = \frac{f_k^2}{\Theta_k} U_k + \frac{f_k f_j}{\Theta_{kj}} U_{kj} + \frac{f_j^2}{\Theta_j} U_j \quad (\text{A-22})$$

$$U_k = \frac{4}{15} A_{kj}^* - \frac{1}{12} \left(\frac{12}{5} B_{kj}^* + 1 \right) \frac{m_k}{m_j} + \frac{1}{2} \frac{(m_k - m_j)^2}{m_k m_j} \quad (\text{A-23})$$

$$U_{kj} = \frac{1}{15} A_{kj}^* \frac{(m_k + m_j)^2}{m_k m_j} \frac{\Theta_{kj}^2}{\Theta_k \Theta_j} - \left(\frac{1}{5} B_{kj}^* + \frac{1}{12} \right) - \frac{5}{32 A_{kj}^*} \left(\frac{12}{5} B_{kj}^* - 5 \right) \frac{(m_k - m_j)^2}{m_k m_j}$$

(A-24)

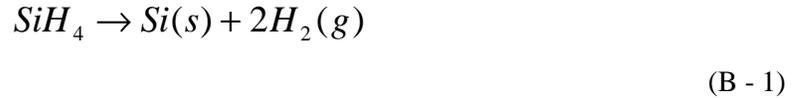
The component thermal diffusion factors were substituted into Equation (2-14) to approximate the multi-component thermo diffusion coefficient. While Equation (2-14) is only an approximation, the additional cost associated with calculating the exact multi-component thermal diffusion coefficient^[11] was not seen to be necessary for this study. The use of Equation (2-14) also kept the method more similar to that used in silicon deposition test cases.

Appendix B

Chemical reaction kinetics

The surface chemical reactions were described in two different methods in this study. The method used at reactor scale was simple one step reaction model, while at the feature scale reactive sticking factors were used.

The simple one step reaction models used at the reactor scale for the silicon deposition cases can be written as



here all the details of the surface reaction are lumped into a single reaction. The equation used to describe the rate at which this reaction occurs was

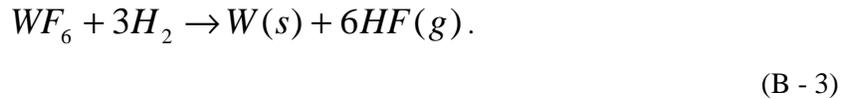
$$R_{SiH_4}^S = \frac{1.6 \cdot 10^{-4} \cdot P_{SiH_4} \exp\left(-\frac{18500}{T}\right)}{1 + K_H P_{H_2}^{1/2} + K_S P_{SiH_4}}, \quad (B - 2)$$

where

$$K_H = 0.19 \text{ Pa}^{-1/2}$$

$$K_S = 0.70 \text{ Pa}^{-1}.$$

This equation was selected in order to be consistent with Kleijn *et al.*^[19]. The surface reaction for the tungsten deposition test case was also described by a single step model,



This was the same reaction model as used by Rodgers and Jensen^[3]. The rate equation for this reaction proposed by Chaara *et al.*^[74] was

$$R_{WF_6}^S = 0.002 \exp\left(\frac{-8300}{T}\right) \frac{P_{H_2}^{1/2} P_{WF_6}}{1 + 1000 P_{WF_6}}.$$

(B - 4)

When global sticking factors were used at the features scale they were calculated based on the single step reaction equations described above and the flux across the feature scale source plane such that

$$\epsilon_k = \frac{R_k^S}{\frac{P_k}{\sqrt{2\pi n_k RT}}}$$

(B - 5)

In order to implement local sticking factors the local fluxes, $\eta_{k,i}$ have to be converted into partial pressures. From kinetic theory

$$P_{k,i} = \eta_{k,i} \sqrt{2\pi n_k RT}$$

(B - 6)

These local partial pressures can now be substituted into Equations (B - 2) or (B - 4) to give local reaction rates. These local reaction rates can then be divided by the local fluxes to give local sticking factors such that

$$\epsilon_{k,i} = \frac{R_{k,i}^S}{\eta_{k,i}}$$

(B - 7)

Bibliography

-
- [¹] C.R. Kleijn, R. Dorsmand, K.J. Kuijlaars, M.Okkerse, H. van Santen, 'Multi-scale modeling of chemical vapor deposition processes for thin film technology', *Journal of Crystal Growth*, **303** (2007) 362-380.
- [²] M.K. Gobbert, C.A. Ringhofer, T.S. Cale, 'Mesoscopic scale modeling of microloading during low pressure chemical vapor deposition', *J. Electrochem. Soc.* **1996**, 143, 2624-2631.
- [³] S. Rodgers, K. Jensen, 'Multiscale modeling of chemical vapor deposition', *J. Appl. Phys.*, **83**, (1998), 524-530.
- [⁴] A. Hasper, J. Holleman, J. Middelhoek, C. R. Kleijn, C. J. Hoogendoorn, 'Modeling and optimization of step coverage of tungsten LPCVD in trenches and contact holes' *J. Electrochem. Soc.*, **138** (1991) 1728-1738.
- [⁵] C. Cavallotti, E. Pantano, a. Veneroni, M. Masi, 'Multiscale simulation of silicon film growth', *Cryst. Res. Technol.*, **40** (2005) 958-963.
- [⁶] W.H. Shepherd, 'Vapor Phase Deposition and Etching of Silicon', *J. Electrochem. Soc.*, **112** (1965) 988-994.
- [⁷] Eversteijn, P. Severin, C. van den Brekel, H. Peek, 'A Stagnant Layer Model for the Epitaxial Growth of Silicon from Silane in a Horizontal Reactor', *J. Electrochem. Soc.*, **117** (1970) 925-931.
- [⁸] K. Jensen, D. Graves, 'Modeling and analysis of low pressure CVD reactors' *J. Electrochem. Soc.*, **130** (1983) 1950-1957.
- [⁹] K. Roenigk, K. Jensen, 'Analysis of multicomponent LPCVD processes', *J. Electrochem. Soc.*, **132** (1985) 448-454.
- [¹⁰] K. Jensen, 'Micro-reaction engineering applications of reaction engineering to processing of electronic and photonic materials', *Chem. Eng. Sci.*, **42** (1987) 923-958.
- [¹¹] C.R. Kleijn, 'Chemical Vapor Deposition Processes' in *Computational Modeling in Semiconductor Processing*, edited by M. Meyyappan, (Artech House, Norwood, MA, 1995), p.97-229.
- [¹²] G. A. Bird, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford University Press Inc., New York **1994**.
- [¹³] C. R. Wilke, 'Diffusional properties of multicomponent gases', *Chem. Eng. Prog.*, **46** (1950), 95-104.
- [¹⁴] M. Grujicj, S.G. Lai, 'Multi-length scale modeling of CVD of diamond Part I A combined reactor-scale/atomic-scale analysis', *Journal of Materials Science*, **35** (2000) 5359-5369.
- [¹⁵] F.S. Lien, M.A. Leschziner, 'A general nonorthogonal collocated finite-volume algorithm for turbulent-flow at all speeds incorporating 2nd-moment turbulence-transport closure .1. Computational implementation', *Comput. Method Appl. M.* **1994**, 114, 123-148.
- [¹⁶] S.V. Patankar, D.B. Spalding, 'A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows', *Int. J. Heat Mass Transfer*, **15** (1972) 1787-1806.
- [¹⁷] C.M. Rhie, W.L. Chow, 'Numerical study of the turbulent flow past an aerofoil with trailing edge separation', *AIAA J.*, **21** (1983) 1525-1532.

-
- [18] H. K. Versteeg, W. Malalasekera, *Computational Fluid Dynamics*, (Prentice Hall, New York, 1995)
- [19] C.R. Kleijn, T.H. van der Meer, C.J. Hoogendoorn, 'A mathematical model for LPCVD in a single wafer reactor', *J. Electrochem. Soc.*, **136** (1989) 3423-3433.
- [20] R.M. Levin, L.E. Evans-Lutterrodt, 'The step coverage of undoped and phosphorus-doped SiO₂ glass-films', *J. Vac. Sci. Technol. B*, **1** (1983) 54-61.
- [21] S.M. Sze, in *VLSI Technology*, (McGraw-Hill, New York, 1983), pp. 111-113.
- [22] K. Skidmore, 'Applying photoresist for optimal coatings', *Semiconductor Int.*, **11** (1988), 57-62.
- [23] J.E.J. Schmitz, R.C. Ellwanger, A.J.M. van Dijk, in *Tungsten and Other Refractory Metals for VLSI III*, edited by V. A. Wells (MRS, Pittsburgh, PA, 1988), p. 55.
- [24] P.E. Riley, V.D. Kulkarni, E.D. Castel, 'Limitation of low-temperature low-pressure chemical vapor deposition of SiO₂ for the insulation of high-density multilevel metal very large scale integrated circuits', *J. Vac. Sci. Technol. B*, **7** (1989), 229-232.
- [25] G.A. Bird, 'Molecular gas dynamics', Oxford University Press, Oxford (1976)
- [26] M. Ikegawa, J. Kobayashi, 'Deposition profile simulation using the direct simulation Monte Carlo method', *J. Electrochem. Soc.*, **136** (1989) 2982-2986.
- [27] M.J. Cooke, G. Harris, 'Monte Carlo simulation of thin-film deposition in a rectangular groove', *J. Vac. Sci. Technol. A*, **7** (1989) 3217-3221.
- [28] A. Yuuki, Y. Matsui, K. Tachibana, 'A study on radical fluxes in silane plasma CVD from trench coverage analysis', *Jpn. J. Appl. Phys.*, **28** (1989) 212-218.
- [29] A. Yuuki, T. Kawahara, Y. Matsui, K. Tachibana, in *Rapid Thermal Annealing/Chemical Vapor Deposition and Integrated Processing*, MRS Symp. Ser. No. 146, edited by D. Hodul, J.C. Gelpey, M.L. Green, T.E. Seidel, (MRS, Pittsburgh, PA, 1989), p.121.
- [30] H. Wulu, K. Sarawat, J. McVittie, 'Simulation of mass transport for deposition in via holes and trenches', *J. Electrochem. Soc.*, **138** (1991) 1831-1840.
- [31] L. Cheng, J. McVittie, K. Sarawat, 'New test structure to identify step coverage mechanisms in chemical vapor deposition of silicon dioxide', *Appl. Phys. Lett.*, **58** (1991) 2147-2149.
- [32] D.G. Coronell and K. F. Jensen, 'Monte Carlo simulations of very low pressure chemical vapor deposition', *J. Computer-Aided Materials Design*, **1** (1993) 3-26.
- [33] C.M. McConica, S. Churchill, in *Tungsten and Other Refractory Metals for VLSI Applications III*, edited by V.A. Wells (MRS, Pittsburgh, PA, 1988), p. 257.
- [34] S. Chatterjee, C.M. McConica, 'Prediction of step coverage during blanket CVD tungsten deposition in cylindrical pores', *J. Electrochem. Soc.*, **137**, (1990) 328-335.
- [35] G.B. Raupp, T.S. Cale, 'Step coverage prediction in low-pressure chemical vapor deposition', *Chem. Mater.*, **1** (1989) 207-214.

-
- [36] T.S. Cale, F.A. Shemansky, G.B. Raupp, in *Tungsten and Other Refractory Metals for VLSI IV*, edited by C. M. McConica and R.A. Blewer (MRS, Pittsburgh, PA, 1989), p. 183.
- [37] G.B. Raupp, T.S. Cale, H.P.W. Hey, 'Step coverage prediction in plasma-enhanced deposition of silicon dioxide from TEOS', *Proceedings of the 6th International VLSI Multilevel Interconnecting Conference*, (IEEE, New York 1989), 488.
- [38] F.A. Shemansky, M.K. Jain, T.S. Cale, and G.B. Raupp in *Rapid Thermal Annealing/Chemical Vapor Deposition and Integrated Processing*. MRS Symp. Ser. No. 146, edited by D. Hodul, J.C. Gelpey, M.L. Green, T.E. Seidel, (MRS, Pittsburgh, PA 1989), p. 173.
- [39] J.E.J. Schmitz, W.L.N. van der Sluys, A.H. Montree, in *Tungsten and Other Advanced Metals for VLSI/ULSI Applications V* (MRS, Pittsburgh, PA, 1990), p. 117.
- [40] T.S. Cale, M.K. Jain, G.B. Raupp, 'Programmed rate processing to increase throughput in LPCVD', *J. Electrochem Soc.*, **137** (1990) 1526-1533.
- [41] T.S. Cale, M.K. Jain, G.B. Raupp, 'Maximizing step coverage during blanket tungsten low pressure chemical vapor deposition', *Thin Solid Films*, **193** (1990) 51-60.
- [42] T.S. Cale, G.B. Raupp, 'Free molecular transport and deposition in cylindrical features', *J. Vac. Sci. Tech B.*, **8** (1990) 649-655.
- [43] T.S. Cale, G.B. Raupp, T.H. Gandy, 'Free molecular transport and deposition in long rectangular trenches', *J. Apply. Phys.*, **68** (1990) 3645-3652.
- [44] T.S. Cale, G.B. Raupp, 'A unified line-of-sight model of deposition in rectangular trenches', *J. Vac. Sci. Tech. B*, **8** (1990) 1242-1248.
- [45] T.S. Cale, T.H. Gandy, G.B. Raupp, 'A fundamental feature scale model for low pressure deposition processes', *J. Vac. Sci. Technol. A*, **9** (1991) 524-529.
- [46] M.M. IslamRaja, M.A. Cappellis, J.P. McVittie, K.C. Sarawat, 'A 3-dimensional model for low-pressure chemical-vapor-deposition step coverage in trenches and circular vias', *J. Appl. Phys.*, **70** (1991) 7137-7140.
- [47] T.S. Cale, G.B. Raupp, T.H. Gandy, 'Ballistic transport-reaction prediction of film conformality in tetraethoxysilane O₂ plasma enhanced deposition of silicon dioxide', *J. Vac. Sci. Technol. A*, **10** (1992) 1128-1134.
- [48] M. Chacara, T. Cale, 'A method to estimate local deposition conditions using film profiles in features', *Thin Solid Films*, **220** (1992) 19-23.
- [49] J.-H. Yun, S.-W. Rhee, 'Deposition with a hemispherical vapor source model', *J. Electrochem. Soc.*, **144** (1997), 1803-1807.
- [50] A. Lankhorst, B. Paarhuis, H. Terhorst, P. Simons, C. Kleijn, 'Transient ALD simulations for a multi-wafer reactor with trench wafers', *Surface & Coatings Technology*, **201** (2007), 8842-8848.

-
- [51] M.K. Gobbert, S.G. Webster, T.S. Cale, 'Transient adsorption and desorption in micrometer scale features', *J. Electrochem. Soc.*, **149** (2002) G461.
- [52] M.K. Gobbert, V. Prasad, T.S. Cale, 'Modeling and simulation of atomic layer deposition at the feature scale', *J. Vac. Sci. Technol. B*, **20** (2002) 1031-1043.
- [53] M.K. Gobbert, V. Prasad, T.S. Cale, 'Predictive modeling of atomic layer deposition on the feature scale', *Thin Solid Films*, **410** (2002) 129-141.
- [54] M.K. Gobbert, T.S. Cale, 'A kinetic transport and reaction model and simulator for rarefied gas flow in the transition regime', *J. Comput. Phys.*, **213** (2006) 591-612.
- [55] M.K. Gobbert, S. G. Webster, T.S. Cale, 'A Galerkin method for the simulation of the transient 2-D/2-D and 3-D/3-D linear Boltzmann equation', *Journal of Scientific Computing*, **30** (2007), 237-273.
- [56] M.K. Gobbert, T.S. Cale, 'Modeling multiscale effects on transients during chemical vapor deposition', *Surface & Coating Technology*, **201** (2007) 8830-8837.
- [57] T.S. Cale, M. O. Bloomfeild, M.K. Gobbert, 'Two deterministic approaches to topography evolution', *Surface & Coating Technology*, **201** (2007), 8873-8877.
- [58] D.G. Coronell, K.F. Jensen, 'Monte Carlo simulations of very low pressure chemical vapor deposition', *J. Computer-Aided Mats. Design*, **1** (1993), 3-26.
- [59] T. Merchant, M. Gobbert, T. Cale, L. Borucki, 'Multiple scale integrated modeling of deposition processes', *Thin Solid Films*, **365** (2000), 368-375.
- [60] T.S. Cale, T.P. Merchant, L.J. Borucki, A.H. Labun, 'Topography simulation for the virtual wafer fab', *Thin Solid Films*, **365** (2000), 152-175.
- [61] J.-Y. Kim, J.-H. Ahn, S.-W. Kang, J-H Kim, 'Step coverage modeling of thin films in atomic layer deposition', *J. Appl. Phys.*, **101** (2007) 073502.
- [62] R.G. Gordon, D. Hausmann, E. Kim, J. Shepard, 'A Kinetic Model for Step Coverage by Atomic Layer Deposition in Narrow Holes or Trenches', *Chemical Vapor Deposition*, **9** (2003), 73-78.
- [63] S.G. Webster, M.K. Gobbert, J.-F. Remacle, T.S. Cale, *Parallel numerical solution of the Boltzmann Equation for atomic layer deposition, Lecture Notes in Computer Science, vol. 2400*, Spriner-Verlag, (2002).
- [64] J. Jilesen, F.-S. Lien, 'Planar source line of sight model with automatically adjusting time increment and local sticking factors', *Chemical Vapor Deposition*, (2009), in press.
- [65] J. Yun, S. Rhee, 'Simplified Simulation of Step Coverage in chemical vapor deposition with a hemispherical vapor source model', *J. Electrochem. Soc.*, **144** (1997), 1803-1807.
- [66] T. S. Cale, V. Mahadev, in *Modeling of film deposition for mircoelectronic applications* (Eds: S. Rossnagel, A. Ullman), (Academic Press, New York, 1996).
- [67] D.A. McQuarie, J.D. Simon, *Physical Chemistry: a Molecular Approach*, (University Science Books, Sausalito, California 1997).

-
- [68] A. R. Neureuther, C.H. Ting, C.-Y. Liu, 'Application of line-edge profile simulation to thin-film deposition processes', *IEEE Trans. Electron. Devices*, **ED-27** (1980), 1449-1455.
- [69] D. Adalsteinsson, J. Sethian, 'A Level set approach to a unified model for etching, deposition, and lithography 1: Algorithms and two-dimensional simulations', *J. Comp. Phys.*, **120** (1995) 128-144.
- [70] S. Osher, J. Sethian, 'Fronts Propagating with Curvature- Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations', *J. Comp. Phys.*, **101** (1988) 12-49.
- [71] J.A. Sethian, *Level Set Methods*, (Cambridge University Press, Cambridge England 1996).
- [72] H. Hwang, T. Govindan, M. Meyyappan, 'Feature Profile Evolution Simulation Using a Level Set Method', *J. Electrochem. Soc.*, **146** (1999), 1889-1894.
- [73] D. Adalsteinsson, J. Sethian, 'A fast Level set method for propagating interfaces', *J. Comp. Phys.*, **118** (1995) 269-277.
- [74] M. B. Chhara, T.S. Cale, 'A method to estimate local deposition conditions using film profiles in features', *Thin Solid Films*, **220** (1992), 19-23.
- [75] K.F. Jensen, S.T. Rodgers, R. Venkataramani, 'Multiscale modeling of thin film growth' *Current Opinion in Solid State & Materials Science*, **3** (1998), 562-569.
- [76] S. Kinoshita, S. Takagi, T. Kai, J. Shiozawa, K. Maki, 'Multiscale analysis of silicon low-pressure chemical vapor deposition', *Japanese Journal of Applied Physics*, **44** (2005), 7855-7862.
- [77] A. Dollet, 'Multiscale modeling of CVD film growth—a review of recent works', *Surface and Coating Technology*, **177-178** (2004), 245-251.
- [78] T.S. Cale, J.-H. Park, T.H. Gandy, G.B. Raupp, M.K. Jain, 'Step coverage predictions using combined reactor scale and feature scale models for blanket tungsten LPCVD', *Chemical Engineering Communications*, **119** (1993) 197-220.
- [79] M. K. Gobbert, T. P. Merchant, L. J. Borucki, T. S. Cale, 'A multiscale simulator for low pressure chemical vapor deposition', *J. ElectroElectrochem. Soc.*, **144** (1997), 3945-3951.
- [80] T.S. Cale, M.O. Bloomfield, D.F. Richards, K.E. Jensen, M.K. Gobbert, 'Integrated multiscale process simulation', *Comp. Mat. Sci.*, **23** (2002), 3-14.
- [81] U.-H. Kwon, W.-J. Lee, 'Three-dimensional deposition topography simulation based on new combinations of flux distribution and surface representation algorithms', *Thin Solid Films*, **445** (2003), 80-89.
- [82] S. Bretznajder (translated by J. Bandrowski), 'Prediction of Transport and Other Physical Properties of Fluids', (Pergamon Press, Oxford 1971).
- [83] J. O. Hirschfelder, C.F. Curtis, R. B. Bird, *Molecular Theory of Gases and Liquids*, (John Wiley & Sons Inc., New York, NY, 1954).
- [84] C.R. Wilke, C.Y. Lee, 'Estimation of diffusion coefficients for gases and vapors', *Industrial and Engineering Chemistry*, **47** (1955), 1253-1257.