Multiscale Modeling of Thin Film Deposition Processes

by

Gwang-Soo Kim

M.Sc. in Chemical Engineering Practice
Massachusetts Institute of Technology, 1999

B.S., Chemical Engineering
Korea Advanced Institute of Science and Technology, 1997

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2002

© 2002 Massachusetts Institute of Technology
All rights reserved

Author

Kwang-Soo Kim

Department of Chemical Engineering
August 9, 2002

Certified by

Klavs F. Jensen
Lammot du Pont Professor of Chemical Engineering
Thesis Supervisor

Accepted by

Daniel Blankschtein
Professor of Chemical Engineering
Chairman, Committee for Graduate Student
DISCLAIMER OF QUALITY

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available. If you are dissatisfied with this product and find it unusable, please contact Document Services as soon as possible.

Thank you.

Some pages in the original document contain color pictures or graphics that will not scan or reproduce well.
Multiscale Modeling of Thin Film Deposition Processes

By

Gwang-Soo Kim

Submitted to the Department of Chemical Engineering
On July 16, 2002, in partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

Abstract
Ionized physical vapor deposition (IPVD) and electrochemical deposition (ECD) are two major thin film deposition processes in the microelectronics industry. The ion fluxes with high kinetic energies in IPVD process involve complex surface interactions that affect overall topology of the microscale features. Copper ECD process involves complex surface reactions and transport phenomena that ranges over different length scales. In this work, predictive simulation tools for these two processes have been developed by investigating the surface reaction and the transport phenomena in IPVD and ECD processes.

In the IPVD process, molecular dynamics (MD) techniques with embedded-atom potentials are used to study the surface reactions for atoms with high impinging energies (30 – 50 eV). The surface reaction rates are combined with ballistic transport and level set methods. The resulting tool demonstrates the effect of the kinetic energy driven surface diffusion on the feature profile evolution.

For the ECD process of copper, detailed surface kinetic mechanisms are developed based on the competitive adsorption/desorption model in the presence of three representative additives, poly ethylene glycol (PEG) and bis-(sodium sulfopropyl) (SPS) and chloride. The proposed kinetic mechanism is capable of describing the synergistic effect of different additives on the copper deposition. Statistically designed experiments were performed with the rotating disk electrode (RDE) apparatus. A hydrodynamic model was developed for RDE and is used to fit the kinetic parameters that are independent of the transport effect.

A reactor scale model is developed based on the Galerkin finite element method. The model includes momentum transport, transient mass transport, potential distribution and detailed surface kinetic mechanisms. The experimental film thickness uniformity on the blank wafer with commercial electrochemical deposition cell is compared with the simulation result. The reactor scale model is used to investigate the various effects on the film thickness uniformity including terminal effects and mass transport effects. The analysis shows the qualitative difference between two effects and how they can be eliminated. Also, the reactor scale simulation tool is used to model the pulse plating process. Improved performance of the pulse plating over the constant current operation suggests that the relaxation period is the critical parameter that determines the film thickness uniformity.

A computationally efficient feature scale model is developed. Mass transport, potential distribution and detailed surface reactions are included in the model. A
combination of level-set and boundary element methods is used to solve the governing equations for feature scale evolution. The level set method is used to model the temporal change of the growing film front, and the boundary element method is used to simulate liquid phase transport. As a consequence of this coupling approach, an explicit remeshing step, which is often computationally very expensive, is not needed for the liquid phase transport equation. The experimental data obtained from the commercial ECD cell is compared with the simulation results, and they show agreements for features with 0.15 and 0.3 μm widths. The domain decomposition and the parallel iterative solver are used to reduce the computational costs in solving large feature scale problems. The structure of the Jacobian matrix after domain decomposition is close to a block diagonal matrix, which is convenient in parallelization schemes. The parallel version of the feature scale tool is used to include net diffusive flux between asymmetric features.

A multiscale simulation tool is developed to model the ECD cell with patterned wafer. The domain is decomposed into reactor, pattern and feature scales. The decomposition criteria are suggested to maximize the numerical efficiency. The feature scale simulations are used to represent the pattern scale. The pattern scale is linked with the reactor scale in a computationally efficient manner. The consistent mass transport model in all the length scales is linked with the moving boundary algorithm based on the level set method to include the effect of the film growth. A developed simulation tool is used to model the effect of the terminal effect on the feature profile evolution on the different locations of the wafer. The analysis shows that the feature at the edge of the wafer grows faster than one at the center, as the current distribution on the blank wafer suggests.

The techniques and models developed in this thesis can be used to investigate systematically the underlying transport and the reactions in the IPVD and the copper ECD system and to help in the design and the optimization of these processes.

Thesis Supervisor : Klavs Jensen
Title: Lammot Du Pont Professor of Chemical Engineering
Acknowledgements

Indeed, time flies like an arrow. I remember when I first came to MIT like yesterday, but almost five years have passed. There were so many fun and hard moments that have made me mature personally and professionally. All of this was possible because there was always somebody to help me. Though I can never thank them enough, I would like to express my gratitude to them in this narrow space.

First, I would like to thank my thesis advisor, Dr. Klavs Jensen. He introduced me to challenging and exciting projects that I have enjoyed last five years. He taught me how to become a good researcher. He emphasized necessary professional virtues, such as devotion to do work in the right way, emphasis on professional knowledge, being an independent researcher and a good collaborator. He patiently guided me through the past five years when it was indeed the toughest time in my life.

I have to thank many people who helped me to conduct my projects. Most of all, I would like to thank Maria Nemirovskaya. She was always willing to discuss technical issues and most of the time she knew how to do things right. Her helpful discussions on personal matters surely helped me go through my toughest times. I would like to thank many people at Motorola, who helped me in my Summer internship in 2000. It was folks in Motorola who encouraged me to conduct the copper electrochemical deposition project and shared their great resources with me. Especially, I would like to thank Tushar Merchant. He was a great mentor, not only on technical matters, but also career issues. I also would like to thank John D’Urso, Len Boruki, and Bill Johnson at Motorola for their support. I would like to thank people in Novellus Systems, Inc., including Larry Gochberg, Jack McInery, and Jon Reid for their help in the summer internship, 2001. I also would like to thank my SRC mentors, Karson Knutson from Intel and Jiong-Ping Lu from TI, for their help in doing my project. There are other people who directly collaborated with me at MIT. They are Uwe Hanson, Vijay Gupta, and Carlo Cavallotti. I enjoyed working with them, and their research experience was helpful. Of course, I had a wonderful time with folks in the KFJ group, especially ones who shared 66-513, Cyril Delattre, Hang Lu, Thomas Gervia, and recently joined Nuria De Mas. By now, they must be immune to Kimchee smell. I also would like to thank our great secretary, Joan Chisholm. With her help, my work was at least twice as easy.

I know I couldn’t have gone through the last five years at MIT without my friends. First of all, I would like to thank Jongmin Lee. He has been a wonderful confidant, and together we have shared times of celebration and grief. I would like to thank Jiaein Choi. As undergraduate students in KAIST, we shared great moments by working, studying, exercising and eating together. Now, he has come to the US to pursue his PhD. Though in Atlanta, GA, the distance did not prevent us from continuing our friendship. There are too many others to list individually, including classmates, practice schoolmates, and other personal friends. Since they’ll feel jealous if I drop their names, let me list the first letter of their family name in an alphabetical order without repeating the same characters. A, B, C, F, H, I, J, K, L, M, N, O, P, R, S, T, and W.
I feel deeply in debt to my old professors in KAIST. They not only provided me professional knowledge, but also the passion toward challenging research, which made me eventually obtain PhD. Even after my graduation, they kept in touch with me and gave me valuable advices. First, I would like to thank professor S.-I. Woo. It was he who taught me what was research and how to become a sincere researcher. I also would like to thank professor S.-B. Park for encouraging me to try new things and reach out all the time. I would like to thank professor S.-W. Park for his advice on different issues in life. He suggested me many great books and especially one he recommended in the class of 1996, “Seven habits of highly effective people” influenced my life throughout my stay in MIT. I would like to thank professors Chang, Ihm, Kim, Lee, Park, and Yang for their support and helpful advice.

More than anybody, it is my family who I thank most for their unconditional love and support. My parents have silently supported me and always wanted the best for me throughout my life. It was my parents who taught me not to want anything without hard working. My two older brothers have also supported me, filling in for their absent youngest brother. My new twin nieces make me particularly happy. Thank you to my aunts and uncles for their emotional support.

Finally, I would like to thank SRC for their generous financial support during my stay in MIT and providing me a number of opportunities to collaborate with people in its host company through of Summer internship.
지난 5년간의 결실입니다.
어머님과 아버님께 현사합니다.
# Table of Contents

1. Introduction and Motivation ................................................................. 18  
   1.1 Introduction .................................................................................. 18  
      1.1.1 Seed layer deposition process: ionized physical vapor deposition .... 20  
      1.1.2 Copper electrochemical deposition ........................................... 21  
   1.2 Motivation for Modeling ................................................................. 22  
   1.3 Issues in Modeling ......................................................................... 24  
   1.4 Thesis Objectives .......................................................................... 25  
   1.5 References .................................................................................... 26  
2. Feature Scale Simulation of Ionized Physical Vapor Deposition Process ......................................................................... 27  
   2.1 Introduction .................................................................................... 27  
   2.2 Methodology .................................................................................. 28  
      2.2.1 IPVD feature scale simulation .................................................... 28  
      2.2.2 Molecular dynamics simulations .............................................. 29  
      2.2.3 Put them together ..................................................................... 31  
   2.3 Results and Discussions ................................................................ 32  
      2.3.1 Feature scale simulation for Cu IPVD ........................................ 32  
      2.3.2 Analysis of kinetic energy driven long range diffusion ............... 33  
      2.3.3 Effect of long range diffusion on feature scale simulation .......... 38  
   2.4 Conclusion ..................................................................................... 38  
   2.5 References ..................................................................................... 39  
3. A Systematic Study of Surface Kinetic Mechanism for Copper Electrochemical Deposition in the Presence of Additives ..................................................................................... 41
4.3.6 Time integration .......................................................... 83
4.4 Experimental Procedure .................................................. 83
4.5 Results and Discussions .................................................. 85
  4.5.1 Comparison of simulation result with experimental data .......... 85
  4.5.2 Parameter study: factors affecting the uniformity of thin film deposition ..... 89
  4.5.3 A case study: pulse plating simulation ................................ 97
4.6 Conclusions ..................................................................... 100
4.7 Appendix: Kinetic Parameters for PPG/SPS/Cl⁻ ........................................ 103
4.8 Reference ......................................................................... 103

5. Feature Scale Simulation of Copper Electrochemical Deposition .......... 106
  5.1 Introduction ..................................................................... 106
  5.2 Governing Equation .......................................................... 107
    5.1.1 Mass transport .......................................................... 107
    5.1.2 Potential distribution .................................................. 109
    5.1.3 Boundary condition .................................................. 110
    5.1.4 Effects of geometry change ......................................... 113
  5.2 Description of Numerical Method .......................................... 113
    5.2.1 Transport equations: boundary element method .................. 114
    5.1.2 Geometry update: level set method ................................ 117
    5.2.3 Parallelization: domain decomposition with artificial boundaries .... 119
  5.3 Experiments ..................................................................... 122
  5.4 Results and Discussion .................................................. 123
    5.4.1 Comparison between experimental data and simulation results .... 123
5.4.2 Parallel computing strategies .................................................. 126

5.5 Conclusion .................................................................................. 130

5.6 References .................................................................................. 132

6. Multiscale Simulation of Copper Electrochemical Deposition Processes ........................................ 134

6.1 Introduction ................................................................................ 134

6.2 Methodology .............................................................................. 137

6.2.1 Problem description: domain decomposition ....................... 137

6.2.2 Decomposition into reactor and pattern scales .................... 138

6.2.3 Decomposition of pattern scale and assembling flux ............ 141

6.2.4 Linking reactor and pattern scales ........................................ 143

6.2.5 Putting all together ............................................................... 147

6.3 Results and Discussions ............................................................. 147

6.3.1 Simulation condition ............................................................. 148

6.3.2 Effect of transient mass transport on the multiscale simulation ........................................... 150

6.3.3 Effect of wafer scale uniformity on feature growth .............. 152

6.3.4 Effect of feature density ......................................................... 156

6.4 Conclusion ................................................................................. 158

6.5 Appendix .................................................................................... 160

6.6 References ................................................................................ 161

7. Conclusion and Recommendations for Future Work ................................................................. 163

7.1 Conclusion ................................................................................ 163

7.2 Recommendations for Future Work .......................................... 166
List of Figures

Figure 1-1. Copper interconnect structures. (a) Three dimensional view (b) Cross section ................................................................. 18

Figure 1-2. Processes to deposit interconnect ................................................................. 19

Figure 1-3. Ionized physical vapor deposition process ...................................................... 20

Figure 1-4. Comparison of the flux between PVD and IPVD processes .......................... 21

Figure 1-5. Copper electrochemical deposition process ................................................... 22

Figure 2-1. Comparison of string algorithm and level set method. (a) String algorithm and geometry merging at constant growth rate. (b) For the level set method and geometry merging at constant growth rate .................................................. 28

Figure 2-2. Surface reactions represented by MD cells ....................................................... 29

Figure 2-3. Flow diagram of IPVD feature scale simulation .............................................. 31

Figure 2-4. Qualitative comparison between feature scale simulation and experimental data .................................................................................................................. 32

Figure 2-5. MD cell for analysis of long range diffusion. For an adsorbing atom, the diffusion length is analyzed. For a desorbing atom, the desorbing angle and energy are also monitored ................................................................. 34

Figure 2-6. Different surfaces to check the effect of surface roughness ............................ 34

Figure 2-7. Diffusion length for an adsorbing atom at different impinging energies (left: 35 eV, right: 50eV) and angles on (100), (111) and (112) surfaces ........................................... 35

Figure 2-8. Diffusion length for a desorbing atom at different impinging energies (left: 35 eV, right: 50eV) and angles on (100), (111) and (112) surfaces ........................................... 36
Figure 2-9. Distribution of diffusion length, desorbing energy and angle for a desorbing atom with different impinging energy and angle, on different surfaces. 37

Figure 2-10. The inclusion of long range diffusion in the feature scale simulation. 38

Figure 2-11. Cu IPVD features in 0.05 mm trench with aspect ratio 3:1. Long range diffusion removes the groove under the overhang. 39

Figure 3-1. Schematic of rotating disk electrode. 42

Figure 3-3. Computational domain of hydrodynamics simulation for the rotating disk electrode. 46

Figure 3-4. Conformation of (a) monomer, (b) dimer and (c) trimer. 53

Figure 3-5. Configuration of SPS. 54

Figure 3-6. Phenomena on copper surface. 55

Figure 3-7. Current-Voltage response with PEG concentration change. Cl⁻ 30 ppm, SPS 0 ppm. 57

Figure 3-8. Current-Voltage response with SPS concentration change. Cl⁻ 30 ppm, PEG 200 ppm. 58

Figure 3-9. Current-Voltage response with SPS concentration change. Cl⁻ 30 ppm, PEG 0 ppm. 59

Figure 3-10. Steric hindrance of SPS. 60

Figure 3-11. Current-Voltage response with Cl⁻ concentration change. SPS 40 ppm, PEG 200 ppm. 60

Figure 3-12. Current-Voltage response with PEG concentration change. Cl⁻ 30 ppm, SPS 0 ppm. 65
Figure 3-13. Current-Voltage response with SPS concentration change. Cl’ 30 ppm,

PEG 200 ppm........................................................................................................................................ 66

Figure 3-14. Current-Voltage response with SPS concentration change. Cl’ 30 ppm,

PEG 0 ppm........................................................................................................................................ 67

Figure 3-15. Current-Voltage response with Cl’ concentration change. SPS 40 ppm,

PEG 200 ppm........................................................................................................................................ 68

Figure 4-1. Computational domain and boundary condition of reactor scale simulation 85

Figure 4-2. Computational mesh for the reactor scale simulation................................. 86

Figure 4-3. Sheet resistivity of wafer at different times. Real time corresponding to the

legend is in Table 4-2......................................................................................................................... 87

Figure 4-4. Comparison between experimental and predicted thickness distributions at

different times. Y-axis is in log scale. ................................................................................................. 88

Figure 4-5. Relative nonuniformity of experimental and predicted results.................. 89

Figure 4-6. Current-Voltage response of electroplating cell with conditions given in

Table 4-3.............................................................................................................................................. 90

Figure 4-7. Influence of terminal effect on the variation in secondary current distribution

for two different seed layer thickness. (a) 50 nm, (b) 1 μm.............................................................. 91

Figure 5-1. Computational domain and boundary condition of feature scale simulation

............................................................................................................................................................... 111

Figure 5-2. Computational domain of boundary element method. Ω is internal domain

and Γ is the boundary points................................................................................................................ 114
Figure 5-3. Boundary condition change. (a) Corner point over specified (b) Closely spaced points avoid overspecification (c) Decomposition of flux avoids overspecification.

Figure 5-4. Combination of level set and boundary element methods for feature scale simulation.

Figure 5-5. Adjusting mesh scheme in finite element method. When opening is pinches off, the numerical quality of mesh is poor.

Figure 5-6. (a) Original geometry and (b) decomposed domain with artificial boundary.

Figure 5-7. Initial geometry of feature scale simulation indicated by dots.

Figure 5-8. Comparison between experimental and simulated feature profiles.

Figure 5-9. Computational domain for long strip with height of 0.205 μm and width of 1.8 μm.

Figure 5-10. Wall clock time of one Newton step with different number of domains (solid line) and parallelization efficiency with different number of domains (dashed line).

Figure 5-11. Number of non-zeroes in the Jacobian matrix (solid line) and number of unknowns for different number of domains (dashed line).

Figure 5-12. Non zero pattern of Jacobian matrix with three artificial boundaries.

Figure 5-13. The feature scale simulation results with four trenches in four domains. Domain boundaries lie at the half way point, which is symmetric plane.

Figure 5-14. Comparisons of feature simulation with interaction case and without interaction case.
Figure 6-1. Different sub-domains of the copper electrodeposition process, (a) reactor scale, (b) pattern scale and (c) feature scale. ................................................................. 136

Figure 6-2. The fluxes on top of feature scale as a function of domain size δ. .......... 140

Figure 6-3. Decomposition of patterns into different segments to resolve variation in the reactor scale. ........................................................................................................ 141

Figure 6-4. Unknowns along the artificial boundaries, Γ_R and Γ_F. .......................... 142

Figure 6-5. Difference in the reactor and feature nodes ........................................ 146

Figure 6-6. Schematic of overall numerical method ................................................ 147

Figure 6-7. Computational mesh for multiscale simulation .................................. 149

Figure 6-8. Feature geometry and three parameters, W, H, W_1 that characterize the geometry ........................................................................................................... 149

Figure 6-9. Feature geometry used for transient mass transport example. .............. 150

Figure 6-10. Concentration profile of Cu^{2+} and feature profiles at (a) beginning of simulation (b) after 2000 seconds elapsed. Concentration profiles within the boundary layer is magnified for better visibility. .............................................. 151

Figure 6-11. Surface concentration of Cu^{2+} at different times at the location x in Figure 6-10 ................................................................. 152

Figure 6-12. Pattern placement for terminal effect, a=1 cm, b=1 cm...................... 152

Figure 6-13. Different feature profiles at the center (left) and at the edge (right) due to the uneven boundary layer thickness ....................................................... 153

Figure 6-14. Pattern placement for mass transport effect, a=2 cm, b=8 cm ............. 154

Figure 6-15. Different feature profiles at the center (left) and at the edge (right) due to the uneven boundary layer thickness ....................................................... 155
Figure 6-16. Pattern placement for mass transport effect, a=5cm, b=5cm.................. 156

Figure 6-17. Different feature profiles at the center (left) and at the edge (right) due to the different boundary pattern density................................................................. 157

Figure 6-18. Geometry for example................................................................. 160
List of Tables

Table 3-1. Experimental conditions for kinetic experiments ........................................ 43
Table 3-2. Variables and equations for rotating disk electrode simulation ...................... 51
Table 3-3. Radius of monomer, dimer and trimer .......................................................... 52
Table 3-4. Radius of PEG .................................................................................................. 53
Table 3-5. Diffusivity of species ....................................................................................... 55
Table 3-6. Proposed kinetic mechanism .......................................................................... 63
Table 3-7. Converged kinetic parameters ........................................................................ 64
Table 4-1. Experimental condition ................................................................................... 84
Table 4-2. Measurement time .......................................................................................... 84
Table 4-3. Simulation condition for parameter study ....................................................... 89
Table 5-1. Operating condition for feature scale evolution experiment ......................... 122
Table 5-2. Feature evolution measurement point ............................................................. 122
Table 6-1. Operating condition for the simulation ........................................................... 148
Table 6-2. Geometry of features used in feature density effect simulation .................... 156
Chapter

1. Introduction and Motivation

1.1 Introduction

Solid state thin films are used in several different applications\(^1\) – storage devices, such as hard disk and CD-ROM; optical application, such as light emitting diode and solid state lasers; micro- and nanoscale devices, such as MEMS; and integrated circuits, such as central processing units and random access memories. The integrated circuit is our main interest.

![Copper interconnect structures](image)

(a) Three dimensional view  (b) Cross section

Figure 1-1. Copper interconnect structures. (a) Three dimensional view (b) Cross section

Modern integrated circuits (IC’s) consist of numerous logic devices. In order for IC’s to be fully functional, all the logic devices have to work and be connected through metal wire, i.e. a so-called interconnect. The structure of a six-level interconnect from
IBM is shown in Figure 1-1. As the number of devices increases and the dimension of devices shrinks, the perfect interconnect deposition becomes more and more difficult.

![Interconnect diagram]

Figure 1-2. Processes to deposit interconnect.

The deposition process for interconnect is summarized in Figure 1-2. On the silicon notch, the diffusion barrier is deposited, which prevents metal atom from diffusing to silicon. Metal diffusion into silicon leads to the breakdown of insulator. Then, a thin metal film is deposited on top of the wafer. High-energy bombardment of metal is used to deposit the seed layer by a process called ionized physical vapor deposition. The function of the seed layer is to provide even nucleation sites so that subsequent metal can deposit evenly through trenches or vias. Seed layer material and subsequently deposited materials are typically the same. Then, the features are filled by much faster deposition processes. Electrochemical deposition process is usually used for copper. Excess
deposition of copper is removed by the next step called chemical mechanical polishing (CMP).

Our main interest is the deposition of the main interconnect materials—seed layer deposition and feature filling steps. Each process will be briefly described in the next sections.

1.1.1 Seed layer deposition process: ionized physical vapor deposition

A seed layer is required to ensure even distribution of metal nuclei so that the subsequent film can be deposited uniformly without voids. Typically, physical vapor deposition (PVD) and ionized physical vapor deposition (IPVD) are used to deposit the seed layer.

![Diagram of ionized physical vapor deposition process](image)

**Figure 1-3.** Ionized physical vapor deposition process
PVD is characterized by its isotropic flux on top of a feature. The isotropic flux is not suitable for small features with a high aspect ratio due to the buildup of metals at the opening of the features. This difficulty can be overcome by introducing a highly directed flux. In the IPVD process shown in Figure 1-3, the atoms are ionized in the plasma, and are brought to the surface orthogonally due to an electric field parallel to the wafer surface. In this way the highly directed flux can penetrate deeper than conventional PVD, as illustrated in Figure 1-4. The interaction between the impinging flux and already deposited film is the determining factor that controls the shape of the feature evolution.

![Diagram](image)

Figure 1-4. Comparison of the flux between PVD and IPVD processes

1.1.2 Copper electrochemical deposition process

Previously, aluminum was the metal of choice for interconnects. However, with the success of inexpensive copper deposition with electrochemical deposition (ECD) processes, more conductive copper became a dominant interconnect material. The higher conductivity of copper creates less heat and consumes less power.

The ECD reactor shown in Figure 1-5 has the wafer at the top, which can be rotated. The inlet is at the bottom of the reactor. Depending on the vendor, the inlet flow
can come through a diffuser or a jet stream. The anode, which is copper metal, is located at the bottom of the reactor. The thin seed layer covering the surface of the wafer act as the cathode.

The advantage of the ECD processes lies in its supreme feature-filling property called superfilling\(^3\); with addition of additives, the copper film grows faster at the bottom of the feature than outside, which ensures feature-fill without voids. The additives also control the grain structure of copper, producing a brighter film. It is not completely understood how these additives affect the deposition process in not completely understood.

![Diagram of Copper Electrochemical Deposition Process](image)

**Figure 1-5. Copper electrochemical deposition process**

1.2 Motivation for Modeling

The fabrication of an interconnect involves many processing steps. As demonstrated from Figure 1-2, even one layer of thin film involves five steps. Therefore, the six level interconnect shown in Figure 1-1 involves at least 30 steps. The large number of processing steps would result in small overall yield. For instance, if there are 30 processes involved and each of them has a yield of 90%, the overall yield is merely 4.24%. The yield is improved to 21.46 % if individual yield becomes 95%. This
example shows that each process has to be highly optimized in order to increase overall yield. For this purpose, it is necessary to identify what determines the yield of the device.

The first factor influencing the yield of the process is the existence of voids in a feature. The voids can be created from any of the deposition steps in Figure 1-2. It could be caused by the geometric constraint of a small width and high aspect ratio trench. In such geometry, the film tends to pinch off at the opening. Or, the void in the feature may be caused by uneven seed layer deposition that precedes the electrochemical deposition step, which creates nonuniform nucleation at the metallization step. If there is high nucleation density at the opening, it will pinch off at the top. If there is no metallization at the bottom, there is no electric contact at the top and the bottom of the circuit level, which is especially detrimental to the multilevel structure.

Another important parameter determining yield of the process is the growth rate uniformity in the wafer scale. If there is any nonuniformity between the center and the edge of the wafer, there can be differences in feature filling. Such nonuniform deposition leads to the failure of devices, particularly those located at the edge of the wafer.

In the microelectronics industry, the design and optimization of the process is typically performed with trial-and-error type experiments. These experiments are often time-consuming and costly. In addition, such an approach may be a quick solution to a particular problem, but does not provide any understanding of the process. Therefore, any systematic issue, such as fundamental design faults of the reactor cannot be captured.
Our objective is to develop a predictive simulation tool, so that it can provide a guideline for design and optimization of the IPVD and the ECD processes, thereby reducing time and resources.

1.3 Issues in Modeling

In the IPVD process, the surface reactions of high kinetic energy ion fluxes need to be understood. The surface reaction rates then need to be combined with the feature scale simulation tool based on fundamental transport models and the moving boundary algorithm.

In ECD processes, the predictive simulation tool requires detailed surface reaction mechanisms and rigorous description of the transport phenomena. The modeling of the transport phenomena in the ECD process is complicated by different length scales involved in the system, e.g. reactor scale \( \sim 0.1 \text{m} \) and feature scale \( \sim 0.1 \cdot 10^{-6} \text{m} \), since the individual features and the reactor chamber need to be modeled simultaneously. The most straightforward way would be to model all the system simultaneously. A similar system (high-pressure chemical vapor deposition) has been modeled by Nemirovskaya, et al.\(^4\). They pointed out that the two-dimensional problem requires \( O \left( 10^7 \right) \) unknowns. Even with a parallel computer, that would take a significant amount of time. Furthermore, the features are continuously growing, requiring remeshing at every time step, which is a computationally demanding procedure. Alternatively, the multiscale simulation approach could be used. In this approach, the domain is decomposed into sub-domains based on its length scale and different models used for different sub-domains. Then each is linked together to ensure consistency among them. This approach does not
require excessive meshing near the feature-reactor interface. Also, if the sub-domains requiring remeshing are grouped together, the whole computational domain does not go through the remeshing steps. Furthermore, the most suitable governing equations and numerical methods can be assigned to each sub-domain model. In conclusion, the multiscale simulation approach can be used to model the transport phenomena in the ECD process. For successful multiscale simulation, the simulation tool for each length scale needs to be developed. Then, an efficient linking scheme needs to couple different sub-domain models consistently.

1.4 Thesis Objectives

The first objective of this thesis is to develop a predictive feature scale simulation tool for the IPVD process by combining detailed surface reaction rates with the rigorous transport and the moving boundary routines. The surface reaction rates of high kinetic energy ions will be calculated with the molecular dynamics simulations.

The second objective of this thesis is to develop a predictive simulation tool for the copper ECD process. Since the detailed surface reactions are not available in the literature, a systematic approach will be taken to develop the kinetic mechanism and obtain kinetic parameters. Then the multiscale simulation approach will be used to model the transport phenomena in the system. Various components of the multiscale simulation including the reactor scale simulation tool and the feature scale simulation tool will be developed. Then, all the information will be coupled together to link different length scale models. The resulting multiscale simulation tool will be linked with the moving boundary algorithm to predict the feature profile evolution in the different locations on
the wafer. First, a feature scale simulation tool will be developed for the IPVD process. The surface reaction rates will be obtained from molecular dynamics simulations and the surface reaction rates will be combined with rigorous transport and moving boundary routines.

1.5 References


2. Feature Scale Simulation of Ionized Physical Vapor Deposition Process

2.1 Introduction

Ionized physical vapor deposition (IPVD) is widely used in the deposition of metals for ultra large-scale integration (ULSI)\textsuperscript{1-3}, since it improves the filling characteristics of deep trenches compared to conventional sputtering (PVD). A large fraction of metal atoms is ionized by high-density plasma in IPVD and an additionally applied bias collimates the ionized metal atoms. The ratio of the collimated flux (ions) to the diffuse flux ( neutrals) is critical to the filling properties of trenches. Adjusting the bias voltage controls the kinetic energy of the impinging metal ions, and the plasma density regulates the portion of ionized versus neutral metal atoms.

The technological importance of the IPVD process has motivated increasing interest in atomistic scale models of the deposition process\textsuperscript{4-6}. For practical applications, the atomistic information needs to be coupled with the \( \mu \)m scale feature scale simulation.

In this chapter, a feature scale simulation tool is developed for IPVD process. The feature scale model is based on the ballistic transport, the level set method and the surface reaction rates. Using the developed feature scale simulation tool, an example Cu IPVD feature profile is produced and compared with experimental data. Kinetic energy driven long-range diffusion effects are studied with molecular dynamics simulation techniques and their effects on the feature profile are delineated.
2.2 Methodology

2.2.1 IPVD feature scale simulation

Transport phenomena of metal ions and neutral atoms in the feature scale can be considered as purely collisionless at low pressures (15-30 mTorr³). At this condition, the mean free path of the particles is of the order of several mm, whereas the dimensions of

(a) String algorithm

(b) Level set method

Figure 2-1. Comparison of string algorithm and level set method. (a) String algorithm and geometry merging at constant growth rate. (b) For the level set method and geometry merging at constant growth rate.
the feature are of the order of 0.1 μm. These particles can be reflected, etch the surface or adsorb on the surface upon collision with the surface. The transport equation should be coupled with the geometry of the feature, and the surface reaction rates\textsuperscript{7,8}.

Adsorption, incorporation and sputtering control the local geometry of the film. Once the net growth rate (i.e. incorporation less sputtering) is obtained, it can be used to update the geometry. String algorithm and level set method have been used in the past to advance profiles. A comparison between the string algorithm and level set method is shown in Figure 2-1. String algorithm has been used as a moving boundary algorithm to update the feature evolution in the past\textsuperscript{9,10}. However, it has been difficult for modeling the geometry merge, since an additional algorithm is needed to remove overlapping regions (See Figure 2-1(a)). Conversely, a second algorithm is not needed when the level set method is used, since geometry merging falls out naturally from the formulation of a hyper surface evolving with a deposition speed\textsuperscript{11-15}. In this work, the level set method is used to update the geometry.

2.2.2 Molecular dynamics simulations

Understanding the interaction between an impinging atom and the previously deposited film is required to predict the feature profile evolution in the IPVD process. Adsorption,
reflection and sputtering (see Figure 2-2) have been studied using molecular dynamics (MD) as a function of impinging angle and energy\textsuperscript{14}.

In this work, the embedded atom method with published parameters is used in the simulations. Details of the underlying potential energy function and the setup of the MD runs can be found in references\textsuperscript{15-17}.

The following model system is constructed to simulate the thin film deposition process. Slabs of metal atoms, 10 layers deep, with periodic boundary conditions in the lateral directions, were used to describe the Al surfaces. The top layer of atoms was arranged to represent (100), (111) and (112) surfaces. The (100) and the (111) surfaces are atomically flat surfaces, whereas the (112) surface is a highly stepped surface consisting of terraces three atomic rows wide. All atoms were allowed to move freely except those of the three bottom layers, which were fixed during the course of the simulation. The time integration of the governing equations (Newton’s 2\textsuperscript{nd} law) was done by Runge-Kutta of second order and the total energy was monitored to discern whether it was conserved\textsuperscript{18,19}.

Each MD run was started with an atom impinging on the top of the slab system. For each incident angle and energy, a large number of trajectories (typically 200) were analyzed to get statistically meaningful results. The position and the velocity of the impinging atom and of all slab atoms were monitored during each run. The force exerted on the migrating atom was used to discern whether or not the atom escapes from the interaction field of the surface. The kinetic energy of the atom provided a measure of the initial energy transferred to the substrate. This approach made it possible to distinguish between adsorbing and desorbing atoms. For adsorbing atoms, only the diffusion length
was recorded. For desorbing atoms, desorption kinetic energy and angle were analyzed in addition to the diffusion length. This information was recorded in terms of the average and the standard deviation, and subsequently used as an input for the level set feature scale simulator. The database for the level set feature scale simulator already contained the necessary adsorption, reflection and sputtering probabilities. The current work made it possible to also include the long-range diffusion phenomenon. For each combination of impinging angle and energy, the probabilities of these different events are calculated and stored as a lookup table.

2.2.3 Putting all together

The flow diagram of the feature scale simulation is given in Figure 2-3. The flux to the point on the surface is determined by the contribution from the source and from the other points in the feature. Depending on the distribution of the energy and the angle, each

![Flow diagram of IPVD feature scale simulation](image)

Figure 2-3. Flow diagram of IPVD feature scale simulation
probability of the adsorption, the reflection and the sputtering is obtained from the MD lookup table described in the previous section\textsuperscript{14}. The net growth rate at the point is obtained by the difference between the adsorption and the sputtering probabilities. The reflected and the sputtered material can reach other points in the feature. Calculating the re-emitted flux is repeated until its contribution becomes negligible. Then, the feature geometry is updated using the net growth rate.

2.3 Results and Discussions

2.3.1 Feature scale simulation for Cu IPVD

In order to test the feature scale simulation tool, published surface reaction rate data is used in the simulation for Cu\textsuperscript{5}. The molecular dynamics data in the paper are sticking

![Figure 2-4. Qualitative comparison between feature scale simulation and experimental data.](image)
coefficients, sputtering yields, and thermal equilibrium coefficients. The source flux on top of the feature is the combination of the isotropic flux and the nonisotropic flux. The ratio between these two fluxes is dependent upon equipment design. In this work it is arbitrarily set to 80% for lack of any information. Using this value, the feature scale simulation is performed with the initial feature width of 0.15 μm and an aspect ratio of 3.

The feature scale profile result is shown in Figure 2-4. The feature scale result is compared with published experimental data\textsuperscript{20}. Qualitatively, the buildup of large overhang at the opening of the feature and the buildup of the triangular structure at the bottom of the trench is well reproduced in the simulation. The main discrepancy is formation of a groove under the overhang observed in the simulation result. This groove is formed because the overhang prevents the mass transport from the source, the so-called shadowing effect. Long-range diffusion of Cu atoms could be a possible means of mass transport to the groove. If the surface diffusion length is significant compared to the feature size, the diffusion can have a significant effect on the feature profile. To explain this issue, MD is used to analyze the diffusion length at different kinetic energies and impinging angles. In the following study, an Al system is studied instead of Cu, since the EAM potential for Cu is not yet publicly available.

2.3.2 Analysis of kinetic energy driven long range diffusion

Recent experiments have revealed that at grazing angles of incidence atoms exhibit so-called long-range diffusion\textsuperscript{21,22}. Zhou and Wadley\textsuperscript{23} used molecular dynamics to investigate the angular and energy dependence of long-range diffusion lengths for incoming atoms with a kinetic energy of less than 20 eV. They also explored the effect
of surface roughness by means of a stepped surface. Under common operating conditions of an IPVD system, there is a significant portion of atoms with energies higher than 20 eV\textsuperscript{24}. Consequently, we explore higher energies of the incoming atoms and evaluate the effect of long-range diffusion on the evolution of surface topology using molecular dynamics technique. For high incident kinetic energies we find that the diffusion lengths significantly increase, and simultaneously the sticking coefficients decrease. Thus, atoms can travel large distances before desorbing.

This long-range diffusion effect can have a significant effect on the feature

![Adsorbing Atoms](image)

![Desorbing Atoms](image)

Figure 2-5. MD cell for analysis of long range diffusion. For an adsorbing atom, the diffusion length is analyzed. For a desorbing atom, the desorbing angle and energy are also monitored.

![Surfaces](image)

Figure 2-6. Different surfaces to check the effect of surface roughness.
scale evolution, especially for smaller features. The effect of this long-range diffusion on the feature scale evolution has not been previously explained.

Using MD simulations, long-range diffusion effects were monitored at specified kinetic energies and angles (cf. Figure 2-5). The top layer of atoms was arranged to represent (100), (111) and (112) surfaces shown in Figure 2-6. The (100) and the (111) surfaces are atomically flat surfaces, whereas the (112) surface is a highly stepped surface consisting of terraces three atomic rows wide.

The diffusion length for an adsorbing atom was simulated for two different impinging energies. The diffusion length is analyzed for the different angles and summarized in Figure 2-7. On average, the diffusion length increases with the impinging kinetic energy. The diffusion length is longer when the impinging angle is higher, and the trajectory starts becoming parallel to the copper surface. The diffusion length is rather similar for (100) and (111) surfaces. At 35 eV, there is essentially no difference between (100) and (111) surfaces. Diffusion lengths on those two surfaces are consistently longer than (112) surface.

![Graphs showing diffusion length vs. incident angle](image1)

(a) Impinging energy = 35 eV

(b) Impinging energy = 50 eV

Figure 2-7. Diffusion length for an adsorbing atom at different impinging energies (left: 35 eV, right: 50eV) and angles on (100), (111) and (112) surfaces.
It is noteworthy that the diffusion length can reach up to 500 Å or 0.05 μm. Considering the feature size of 0.1 μm, the diffusion length could potentially have a significant effect on the feature scale. However, the adsorption probability is small at high kinetic energies and high angles. Consequently, long-range diffusion of high energy adsorbing atoms is not going to have a significant effect.

For desorbing atoms, the angle and the energy of desorption need to be computed in addition to the diffusion length. Figure 2-8 shows the diffusion length for different angles and energies for desorbing atoms. The diffusion length for desorbing atoms is generally shorter than for adsorbing atoms. The maximum distance is around 20 to 25 nm, a significant portion of the dimension of a contemporary feature (~200nm). Also, this value is significantly larger than the diffusion length (~0.5nm) driven by thermal energy. Figure 2-9 shows the distribution of diffusion length, desorbing energy and desorbing angle as a function of angle of incidence. At the incident angle of 60 degrees, the diffusion length is mono-dispersed near 0, which indicates that the atoms

![Figure 2-8. Diffusion length for a desorbing atom at different impinging energies (left: 35 eV, right: 50 eV) and angles on (100), (111) and (112) surfaces.](image-url)
are reflected without traveling a long distance on the surface. As the angle increases, the frequency of larger diffusion length increases and ultimately the distribution of length becomes very wide at incident angles near 90 degrees. The desorbing kinetic energy follows a Gaussian distribution at the smallest incident angle. However, as the impinging angle increases, the averages of desorbing energy average shift to higher values and the distribution becomes skewed. This behavior can be explained by the atoms having a large portion of momentum component parallel to the surface at higher angles. The angle of desorption does not appear to display any significant trends with incident angle.

![Graphs showing distribution of diffusion length, desorbing energy, and angle for different incident angles.](image)

Figure 2-9. Distribution of diffusion length, desorbing energy and angle for a desorbing atom with different impinging energy and angle, on different surfaces.
2.3.3 Effect of long range diffusion on feature scale simulation

Figure 2-10 illustrates the impact of long-range diffusion behavior on transport in a trench. The diffusion allows the atom flux to penetrate deep into the trench. Two simulations were performed to demonstrate the effect of long-range diffusion on the evolution of feature profiles (Figure 2-11). The feature profile without the long-range diffusion effect has a groove under the overhang as in Figure 2-4, while long-range diffusion eliminates this groove.

![Diagram showing long range diffusion](image)

Figure 2-10. The inclusion of long range diffusion in the feature scale simulation.

2.4 Conclusion

A feature scale simulation tool for ionized physical vapor deposition was described. Molecular dynamics simulations were used to calculate the surface kinetic mechanism. The feature scale simulation tool containing the level set method and the ballistic transport mechanism was combined with surface reaction rates. Long-range diffusion rates were calculated by MD simulations for different surfaces, angles of incidence, and kinetic energies. The results demonstrate that long-range diffusion effects could have a
significant effect on sub-micron feature scale evolution. The effect appears to smooth out effects of shadowing.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2-11.png}
\caption{Cu IPVD features in 0.05 mm trench with aspect ratio 3:1. Long range diffusion removes the groove under the overhang.}
\end{figure}

2.5 References


Chapter

3. A Systematic Study of Surface Kinetic Mechanism for Copper Electrochemical Deposition in the Presence of Additives

3.1 Introduction

In the simulation of copper electrochemical deposition, surface reaction rates must be assigned as a boundary condition to specify the consumption rate of cupric ion under different operating conditions. Since it is known that the consumption rate for cupric ion changes with different additive concentrations, it is necessary to include the effect of the additive species on the surface kinetic mechanism. Constructing a rigorous kinetic mechanism is a daunting task due to numerous possible ways that additives and cupric ions could interact on the surface. Furthermore, even such a detailed mechanism will have a limited usage, since different proprietary additives are used by plating companies and the same mechanism will not be generally applicable to different sets of additive chemistry. Therefore, we first choose a representative system and then focus on developing a systematic methodology for realizing surface kinetic mechanisms for copper electrochemical deposition.

In this work, we selected three representative additives, polyethylene glycol (PEG), bis-(sodium sulfropropyl) disulfide (SPS), and chloride ions. We aim at developing surface kinetic mechanism that is consistent with a collection of experimental data covering a wide range of operating conditions. Also, we try to keep the reaction as simple as possible, unless there are trends that cannot be explained by the proposed kinetic mechanism. Finally, we use the experimental data and the hydrodynamic model
for the experimental apparatus to extract the kinetic parameters that are not dependent upon transport phenomena.

3.2 Experiments

3.2.1 Kinetic Experiments

The rotating disk electrode (RDE)\textsuperscript{1-5} (see Figure 3-1) is chosen as the experimental apparatus to obtain current-voltage response under different operating conditions. A DC voltage is applied between the working and the counter electrodes, and the current is measured. The measured current represents the surface consumption rate of cupric ions, which is governed by surface reaction rates and transport. The working electrode is rotating at high speeds (~O(1000RPM)) to create a thin boundary layer. The enhanced transport created by the disk rotation allows the system to operate in surface reaction limited regime, which is needed to measure intrinsic surface reaction rates\textsuperscript{1}. However,
when the applied voltage is too high or the rotation speed is too low, the current approaches the limiting value and transport of cupric ions limits the observed rate. Thus, in order to analyze surface reaction rates under wide range of operating conditions, the transport effect must be decoupled from the observed current densities by using a rigorous transport analysis. This approach will be addressed in Section 3.3.

RDE equipment from Radiometer Analytical, using an EDI101 rotating disk electrode with the B70A600 cell was used. The working electrode was a 5mm diameter platinum disk. \( \text{H}_2\text{SO}_4 \) and \( \text{CuSO}_4 \) concentrations were fixed and three concentration levels were examined for three additives, SPS, PEG and Cl\(^-\). A full factorial experiment was designed for the concentration combination. The experimental conditions are summarized in Table 3-1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEG Concentration</td>
<td>0, 100, 200 ppm</td>
</tr>
<tr>
<td>SPS Concentration</td>
<td>0, 40, 80 ppm</td>
</tr>
<tr>
<td>Cl(^-) Concentration</td>
<td>30, 60, 90 ppm</td>
</tr>
<tr>
<td>Current Density</td>
<td>0, 1, 10, 20, 40, 60, 80 mA/cm(^2)</td>
</tr>
<tr>
<td>Rotation Speed</td>
<td>1500 RPM</td>
</tr>
<tr>
<td>( \text{CuSO}_4 ) Concentration</td>
<td>0.34 M</td>
</tr>
<tr>
<td>( \text{H}_2\text{SO}_4 ) Concentration</td>
<td>1.65 M</td>
</tr>
</tbody>
</table>

Before the current-voltage response was measured, a ~10 nm layer of Cu was deposited on the Pt electrode, so that surface reactions occurred on the Cu surface. The voltage corresponding to each applied current density was recorded over a period of time. A typical voltage response graph is shown in Figure 3-2. Transient effects in the system prevent the voltage from immediately converging to its steady state value. Thus, a stop criterion was selected in the data acquisition software to locate the end point of
measurements by identifying when the voltage change in two consecutive time steps was smaller than the selected tolerance. Then, the voltage-time curve was visually inspected to check whether the stop tolerance was strict enough to catch any slowly varying trends. In this way, the steady state voltage was recorded for each specified current. Finally, the voltage corresponding to 20 mA/cm² was measured again at the end of each experiment to ensure the reproducibility. In total, 30 sets of experiments were performed and each set of experiments took around 15 minutes.

3.3 Modeling of Rotating Disk Electrode

The current density value from the rotating disk electrode reflects the consumption rate of cupric ions. The consumption rate is dictated by inherent surface reaction rates as functions of surface concentrations and potential. These local variables can only be

![Graph of Voltage vs Time](image)

Figure 3-2. Typical voltage responses from rotating disk electrode experiments: 30 ppm Cl⁻, 40 ppm SPS, 200 ppm PEG, 20 mA/cm².
found by solving transport equations. In the next subsections, the transport equations and required parameters are developed.

3.3.1 Rotating disk electrode simulation

Modeling of the rotating disk electrode reproduces the current densities for specified operating conditions. The results can be used to decouple transport effect and fit kinetic parameters from experimental data. The following analysis is based on John Newman's original analysis\textsuperscript{3}. The assumptions are:

- The electrode is much smaller than the rotating disk electrode cell, and thus the system can be handled as a rotating disk within a semi-infinite medium.
- The majority of potential drop occurs outside of the diffusion boundary layer. Therefore, the potential distribution is only calculated outside of the diffusion boundary layer. The deviation caused by neglecting potential drop within the diffusion layer is accounted for by including a concentration overpotential.

By introducing these assumptions, the potential and the concentration are decoupled in the governing equation, but still coupled in the boundary condition. The computational domain is shown in Figure 3-3. The axisymmetry of the system implies that only radial and axial variations of concentration and potential need to be computed.
Figure 3-3. Computational domain of hydrodynamics simulation for the rotating disk electrode

Under the first assumption, the velocity field near the rotating disk can be given by the classical von Karman rotating disk solution:

\[ v_r = ary\Omega\sqrt{\Omega/\nu} \]  \hspace{1cm} (3.1)

\[ v_y = -ay^2\Omega\sqrt{\Omega/\nu} , \]  \hspace{1cm} (3.2)

where \( a = 0.51023 \). The convection induced by disk rotation is the dominant mass transport mechanism outside the diffusion boundary layer. The migration is negligible compared to the convection despite the existence of potential gradient. Within the diffusion boundary layer, the diffusion plays the dominant role. Only axial component of diffusion is included, since the thickness of the diffusion boundary layer is negligible compared to the disk radius. The second assumption implies that the migration within
the diffusion layer can be neglected. Consequently, the governing equation for mass transport is reduced to the convective diffusion equation of the form:

\[ \nu_r \frac{\partial c}{\partial r} + \nu_y \frac{\partial c}{\partial y} = D \frac{\partial^2 c}{\partial y^2} \]  

(3.3)

Inserting Equations (3.1) and (3.2) into Equation (3.3) turns Equation (3.3) into the following equation:

\[ ay\sqrt{\frac{\Omega}{\nu}} \left[ r \frac{\partial c}{\partial r} - y \frac{\partial c}{\partial y} \right] = D \frac{\partial^2 c}{\partial y^2} \]  

(3.4)

At the center axis of the boundary, the concentration field is symmetric, therefore

\[ \frac{\partial c}{\partial r} = 0, r = 0 \]  

(3.5)

Far away from the disk, the concentration equals the bulk concentration.

\[ c = c^\infty, \text{ as } y \rightarrow \infty \]  

(3.6)

On the surface of the disk \((y=0)\), the consumption of different species is specified as a sum of the elementary surface reactions in which species \(i\) participates:

\[ -D \frac{\partial c}{\partial y} \bigg|_{y=0} = \sum_{j=1}^{\text{num}} n_j r_j \]  

(3.7)

The surface reaction rate \(r_j\) is dependent upon the potential values on the surface. The potential distribution outside of the diffusion layer satisfies Laplace’s equation:

\[ \nabla^2 \Phi = 0 \]  

(3.8)

Far away from the disk, the potential is the same as value at the anode, which can be arbitrarily defined as zero. The potential on the cathode surface of electrode is defined as \(V\), which is the applied voltage. However, this value is not the same as the liquid phase potential \(\Phi\) due to additional potential drops that occur in the diffusion boundary layer.
and in the double layer, defined as concentration and surface overpotentials, respectively.

The continuity of potential at the solid-liquid interface can be written as follows:

\[ V = \Phi_0 + \eta_c + \eta_s \quad (3.9) \]

The concentration overpotential is computed as follows\(^4\):

\[ \eta_c = -\frac{RT}{ZF} \ln \left( \frac{c_\infty}{c_0} \right) - I \left( 1 - \frac{c_0}{c_\infty} \right) \quad (3.10) \]

The surface overpotential is the driving force for the surface reaction, and can be expressed by the Butler-Volmer formula for electrochemical reactions\(^1\). This value is coupled with the current values on the surface. The current density can be expressed in several equivalent manners.

\[ i = \frac{nFD}{1-t} \frac{\partial c}{\partial y} \bigg|_{y=0} = -k_w \frac{\partial \Phi}{\partial y} \bigg|_{y=0} = \frac{nF}{1-t} \sum_{j=1}^{j=\text{max}} v_j r_j \quad (3.11) \]

Combining all the equations, the observable average current density is computed by the integral.

\[ i_{\text{avg}} = \frac{1}{r_0} \int_{r=r_0}^{r=r_6} f_n(V, \Omega, R; k_i) \quad (3.12) \]

Equation (3.12) gives the relationship between the observable current density and the kinetic parameters \(k_i\) for a given operating condition. Using this relationship, one can find the optimum set of \(k_i\)'s that make the predicted \(i_{\text{avg}}\) close to the experimental value. Since the function in Equation (3.12) is nonlinear, it has to be solved numerically, as described in the next section.
3.3.2 Numerical simulation of rotating disk electrode

In order to predict average current density, \( i_{avg} \) at different operating conditions, Equations (3.4) to (3.12) need to be solved numerically in the specified geometry and with the specified operating conditions. The numerical method is based on Newman’s original method\(^3\).

Since the governing equation Equation (3.4), and the boundary conditions in Equations (3.5) and (3.6) are linear, the infinite series solution approach is applicable. Using the boundary layer thickness as a characteristic length in the \( y \) direction, the coordinate \( y \) can be normalized as follows:

\[
\xi = y (aN / 3D)^{1/3} \sqrt{\Omega / \nu} \tag{3.13}
\]

The concentration can then be expressed as a series solution:

\[
c = c^* \left[ 1 + \sum_{m=0}^{\infty} A_m (r / r_0)^{2m} \theta_m (\xi) \right] \tag{3.14}
\]

\( \theta \) satisfies the differential equation

\[
\theta'' + 3 \xi^2 \theta' - 6m \xi \theta_m = 0 \tag{3.15}
\]

with boundary conditions

\[
\theta_m = 1 \text{ at } \xi = 0 \tag{3.16}
\]

\[
\theta_m = 0 \text{ at } \xi = \infty. \tag{3.17}
\]

The only nonlinear part is the boundary condition on the copper surface, Equation (3.7). This boundary condition is used to find the coefficients, \( A_m \).

\[
D \frac{\partial c}{\partial y} \bigg|_{y=0} = D c^* \left( \frac{aN}{3D} \right)^{1/3} \sqrt{\frac{\Omega}{\nu}} \sum_{m=0}^{\infty} A_m \left( \frac{r}{r_0} \right)^{2m} \theta_m (0) = \sum_{j=1}^{j=N_{cpp}} V_j r_j \tag{3.18}
\]
A similar approach is used to find the potential distribution, but instead of using cylindrical coordinate, elliptic coordinates are used for convenience. The elliptic coordinates are related to the cylindrical coordinate \( r \) and \( y \) as follows.

\[
y = n_0 \xi \eta \quad \text{and} \quad r = r_0 \sqrt{(1 + \xi^2)(1 - \eta^2)}
\]

(3.19)

Using this coordinate, Equation (3.8) can be expressed as the following series solution:

\[
\Phi = \frac{RT}{ZF} \sum_{n=0}^{\infty} B_n P_{2n}(\eta) M_{2n}(\xi)
\]

(3.20)

\( P_n \) is a Legendre polynomial and \( M_n \) is a Legendre function of imaginary argument satisfying the differential equation

\[
\frac{d}{d\xi} \left[ (1 + \xi^2) \frac{dM_{2n}}{d\xi} \right] = 2n(2n+1)M_{2n}
\]

(3.21)

with the boundary condition

\[
M_{2n} = 1 \text{ at } \xi = 0 \text{ and } M_{2n} = 0 \text{ at } \xi = \infty.
\]

(3.22)

Using Equations (3.11) and (3.20), the current can be expressed as follows:

\[
i = -k_m \frac{\partial \Phi}{\partial y} \bigg|_{y=0} = \frac{k_m}{r_0 n_0} \frac{RT}{ZF} \sum_{n=0}^{\infty} B_n P_{2n}(\eta) M_{2n}(0)
\]

(3.23)

Since the consumption of cupric ions and current has a linear relationship

\[
i = \frac{nFD}{1-t} \frac{\partial c}{\partial y} \bigg|_{y=0},
\]

(3.24)

combining (3.18), (3.23) and (3.24) produces the relation between \( A_m \) and \( B_n \):

\[
B_n = \frac{\pi}{4} N \sum_{n=0}^{\infty} Q_{n,m} A_m
\]

(3.25)

The definition of \( N \) and \( Q_{n,m} \) in Equation (3.25) can be found in Newman’s original reference\(^3\). Although (3.14) and (3.20) are infinite series, sufficient numerical accuracy
can be obtained by truncating the infinite series to contain only \( n \leq 10 \) and \( m \leq 10 \). The set of equations that need to be determined is summarized in Table 3-2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B_{m, n=0, 10} )</td>
<td>(3.25)</td>
</tr>
<tr>
<td>( A_{m, m=0, 10} )</td>
<td>(3.18)</td>
</tr>
<tr>
<td>( i_s, k=0, 10 )</td>
<td>(3.23)</td>
</tr>
<tr>
<td>( c_0, k=0, 10 )</td>
<td>(3.14)</td>
</tr>
<tr>
<td>( \Phi_{0s, k=0, 10} )</td>
<td>(3.20)</td>
</tr>
<tr>
<td>( \eta_s, k=0, 10 )</td>
<td>(3.9)</td>
</tr>
<tr>
<td>( \eta_{co, k=0, 10} )</td>
<td>(3.10)</td>
</tr>
<tr>
<td>( i_{avg} )</td>
<td>(3.12)</td>
</tr>
</tbody>
</table>

The nonlinear system is solved by the Newton-Raphson method.

3.3.3 Diffusion coefficient

Accurate values of diffusion coefficients are crucial parameters for a rigorous mass transport simulation. Here, \( Cu^{2+} \), PEG, SPS, and Cl\(^-\) will be considered. The diffusivity of \( Cu^{2+}, 4.6 \) and Cl\(^-\) are well documented, but the diffusivities of PEG and SPS require special attention.

The diffusivity of PEG varies with its molecular weight. There is no reported value for PEG with molecular weight of 4600 \( g \cdot mol^{-1} \). Two methods can be used to obtain these values. The first method is to use a correlation in the literature reported for PEG\(^7\). Using this correlation, the diffusivity is as follows:

\[
D_{PEG} = 1.25 \cdot 10^{-4} \cdot MW^{0.55} = 1.18 \cdot 10^{-6} cm^2 / sec
\]  

(3-26)

The second method is based on the Stokes-Einstein correlation\(^8\) and assumes a spherical solute with the radius as a parameter. An approximation to the spherical radius can be obtained by computing the PEG equilibrium conformation and the radius of the sphere
enclosing this conformation. Quantum chemistry calculations including the effect of solvent and ions would be a rigorous approach. For simplicity, an approximate method is used to estimate the volume of PEG. First, we assume that the shape of the PEG molecule is a perfect sphere and there are no voids between stacks of monomers. In that case, the total volume of PEG is the volume of a repeat unit times the number of repeat units.

The volume of a repeat unit can be estimated as the volume difference between the dimer and the monomer. The monomer volume is not used directly because the additional hydrogen at the end of the molecule could inflate the PEG radius. The trimer and monomer volumes are used as well as a "sanity check". The volume of the monomer, the dimer and the trimer are calculated using the nominal radii of those molecules, which is defined as the radius where the electron density is 0.001 electron/bohr.\(^3\). Restricted Hartree-Fock theory and the 3-21G basis function set is used to optimize geometry using Gaussian 98\(^{\text{\textsuperscript{\textregistered}}}\). The obtained equilibrium shapes of the monomer, the dimer and the trimer are given in Figure 3-4.

<table>
<thead>
<tr>
<th></th>
<th>Radius (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monomer</td>
<td>2.89</td>
</tr>
<tr>
<td>Dimer</td>
<td>3.42</td>
</tr>
<tr>
<td>Trimer</td>
<td>3.71</td>
</tr>
</tbody>
</table>

Based on these calculations, the volume of the repeat unit, defined as the volume difference between dimer and monomer, is 46.34 Å\(^3\). If we use the volume difference between trimer and dimer, the volume of the repeat unit is 66.45 Å\(^3\). The radius of polymer is calculated as follows.
(Unit volume of repeating unit)*(number of repeating unit) = \(\frac{4}{3}\pi R^3\) \hspace{1cm} (3.27)

The radius of PEG is summarized in Table 3-4.

<table>
<thead>
<tr>
<th>Unit volume</th>
<th>Radius of PEG (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimer-Monomer</td>
<td>10.6</td>
</tr>
<tr>
<td>Trimer-Dimer</td>
<td>13.76</td>
</tr>
</tbody>
</table>

Kelly and West estimated the radius of PEG with MW of 3350 using their microbalance experiment\(^{11}\). The radius they calculated is 8.5 Å. Since our molecular weight is larger (4600), it makes sense that our radius is larger than their calculation. If we back-

![Diagram of molecular structures (a) monomer, (b) dimer, (c) trimer](image)

**Figure 3-4. Conformation of (a) monomer, (b) dimer and (c) trimer**

calculate the radius corresponding to 3350 MW PEG, the radius would be 9.43 Å and 10.64 Å depending on which unit volume to use.
Using the radius of 9.43 Å and the viscosity of pure water, 1 cp = 10⁻³ Pa sec, the diffusivity of PEG is calculated using Stokes-Einstein equation, 

\[ D = \frac{k_B T}{6\pi \eta R} = 2.06 \cdot 10^{-6} \text{ cm}^² / \text{sec} \]  

(3.28)

This value is approximately twice as large as that calculated from the correlation, primarily because in reality the volume of PEG is increased by the void volume within the chain.

The diffusion coefficient of SPS is not available in the literature and its manufacturer, Raschig in Germany, would not provide this information. A similar approach to that used for PEG is therefore used to calculate the diffusivity of SPS. The equilibrium shape of SPS is calculated using Gaussian 98™. Its equilibrium shape is given in Figure 3-5. The calculated radius is 5.22 Å. It is noteworthy to mention that

Figure 3-5. Configuration of SPS
even though there is large difference in the molecular weight, the calculated radii of SPS and PEG are not much different from each other. Using the computed SPS radius in the Stokes-Einstein equation gives the diffusivity:

\[
D = \frac{k_BT}{6\pi\eta R} = 4.18 \cdot 10^{-6} \text{ cm}^2/\text{sec}
\]  

(3-29)

The diffusion coefficients are summarized in Table 3-5.

<table>
<thead>
<tr>
<th></th>
<th>Diffusivity (10^{-6} \text{cm}^2/\text{sec})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cl^{-}</td>
<td>20.3</td>
</tr>
<tr>
<td>SPS</td>
<td>4.18</td>
</tr>
<tr>
<td>PEG</td>
<td>2.06</td>
</tr>
<tr>
<td>Cu^{2+}</td>
<td>6.48*</td>
</tr>
</tbody>
</table>

*: The value in Table 3-5 is for 1.78M H_2SO_4 and 0.11M CuSO_4.

3.4 Results and Discussion

3.4.1 Surface kinetic mechanism

Competitive adsorption/desorption models well established for other surface reaction systems, such as chemical vapor deposition. Figure 3-6 illustrates the physical processes

![Surface diffusion](image1)

![Adsorption](image2)

![Desorption](image3)

![Incorporation](image4)

Figure 3-6. Phenomena on copper surface
taking place on the copper surface. The solute adsorbs on the copper surface. The adsorbate can then diffuse on the surface or desorb back to the liquid phase. The rate processes can be formulated as follows:

*Adsorption:*

\[ A + \theta_v \rightarrow A_s \]  \hspace{1cm} (3.30)

*Desorption:*

\[ A_s \rightarrow A + \theta_v \] \hspace{1cm} (3.31)

The additive species can be incorporated into the copper film if the growth rate of copper is higher than the surface diffusion or the desorption rate. This incorporation phenomenon is supported by the existence of impurities in the copper film. Roha et al. analyzed incorporation reaction rates as a function of current density values and the size of the additive molecule\(^\text{12}\). In this study, his general formula will be used to describe the incorporation reaction rate.

*Incorporation:*

\[ A_s \rightarrow A_{\text{incorporated}} + \theta_v \] \hspace{1cm} (3.32)

Equations (3.30), (3.31) and (3.32) form the basis for the surface reaction expressions for all the components. It has yet to be verified whether these formulae can explain the experimental observation qualitatively. Once it is verified that the proposed mechanism can explain the experimental trends qualitatively, then a rigorous parameter fitting procedure can be used to obtain the parameters for quantitative agreement with data.

Typical qualitative characteristics of competitive adsorption and desorption arise from the components independently competing for available sites. That is, the
addition of additives would produce smaller current density at the same voltage, since the adsorption of additives results in fewer available sites for cupric ion adsorption. Thus, examining the experimental current-voltage response could indicate whether the competitive adsorption-desorption reaction scheme is appropriate.

Figure 3-7 shows a typical current-voltage response graph. Changing only the PEG concentration shifts the graph in the cathodic direction, indicating a higher voltage will be required to obtain the same current. This observation is consistent with other published results\textsuperscript{11,13,14} and fits with the competitive adsorption-desorption scheme.

![Figure 3-7. Current-Voltage response with PEG concentration change. Cl\textsuperscript{-} 30 ppm, SPS 0 ppm](image)

Figure 3-8 shows the current-voltage responses when the SPS concentration is varied. Opposite to the observation in the PEG case, the current-voltage response curves move to the anodic direction as SPS concentration increases. The current density
increases at the specified potential as SPS concentration increases, which is contradictory to the basic competitive adsorption/desorption argument. Figure 3-9 shows the current-voltage response in slightly different conditions. The conditions in the Figure 3-9 are identical to those for Figure 3-8 except there is no PEG added. In this case, the addition of SPS does not affect the current-voltage response. Again, this observation is not consistent with a simple competitive adsorption/desorption model. Comparing Figures 3-8 and 3-9 reveals that I-V characteristic with 80 ppm of SPS in Figure 3-8 overlaps with Figure 3-9 with no PEG present.

Figure 3-8. Current-Voltage response with SPS concentration change. Cl 30 ppm, PEG 200 ppm
Figure 3-9. Current-Voltage response with SPS concentration change. Cl 30 ppm, PEG 0 ppm

These results show that addition of SPS with PEG increases the current density, whereas SPS alone has no significant effect on current-voltage response.

A fundamental understanding of these phenomena would require a detailed study of adsorption/desorption phenomena of additives at the molecular scale. However, that level of detailed information is not available. Therefore, we propose a simple model consistent with the experimental observations.

The experimental data can be interpreted as SPS preventing PEG from adsorbing by steric hindrance. Since SPS is believed to adsorb more strongly than PEG, even a small concentration of SPS on the surface can effectively prevent PEG from adsorbing. However, SPS does not hinder cupric ion adsorption due to the relatively small size of cupric ions. This process is illustrated in Figure 3-10.
This steric hindrance of SPS can be formulated by introducing new vacant sites near the adsorbed SPS and specifying this new site is only available to the cupric ions. The formation of a new vacant site near SPS is expressed as:

$$SPS_z + \theta_v \leftrightarrow SPS_z - \theta_v$$

(3.33)

A representative current-voltage response is shown in Figure 3-11 for the case of
Cl. The qualitative trend is similar to Figure 3-8, so that Equation 3-33 can be used to explain the Cl trends analogous to the SPS adsorption process.

At this point, the model includes competitive adsorption/desorption models with the additional reaction for a new surface site. Details of the surface reaction mechanism are summarized below. For PEG, the adsorption, the desorption and the incorporation reactions need to be considered. Because of the large size of PEG, it can occupy more than 1 site in contrast to other smaller species. The actual number of sites is dependents upon the size of PEG (i.e., the molecular weight of PEG), and it is left as a parameter to be fit to data.

\[
P_{\text{EG}} + n \cdot \theta_v \rightarrow P_{\text{EG}} s \\
P_{\text{EG}} s \rightarrow P_{\text{EG}} + n \cdot \theta_v \\
P_{\text{EG}} s \rightarrow P_{\text{EG incorporated}} + n \cdot \theta_v
\]

For SPS, adsorption, desorption, incorporation and creation of the preferential spots for cupric ion are the reactions to be considered.

\[
S_{\text{PS}} + \theta_v \rightarrow S_{\text{PS}} s \\
S_{\text{PS}} s \rightarrow S_{\text{PS}} + \theta_v \\
S_{\text{PS}} s \rightarrow S_{\text{PS incorporated}} + \theta_v \\
S_{\text{PS}} s + \theta_v \leftrightarrow S_{\text{PS}} s - \theta_v
\]

Similar reaction processes have to be included for Cl.
\[ \text{Cl}^- + \theta_v \rightarrow \text{Cl}^- \text{s} \]  
\[ (3.41) \]

\[ \text{Cl}^- \text{s} \rightarrow \text{Cl}^- + \theta_v \]  
\[ (3.42) \]

\[ \text{Cl}^- \text{s} \rightarrow \text{Cl}^- \text{incorporated} + \theta_v \]  
\[ (3.43) \]

\[ \text{Cl}^- \text{s} + \theta_v \leftrightarrow \text{Cl}^- \text{s} - \theta_v \]  
\[ (3.44) \]

For cupric ions, a two-step reductions mechanism is included.

\[ \text{Cu}^{(II)} + \theta_v + e^- \leftrightarrow \text{Cu}^{(I)} \text{s} \]  
\[ (3.45) \]

\[ \text{Cu}^{(I)} \text{s} + e^- \leftrightarrow \text{Cu}^{\text{incorporated}} + \theta_v \]  
\[ (3.46) \]

The preferential deposition sites created by SPS and Cl\(^-\) are described.

\[ \text{Cu}^{(II)} + \text{SPS}_v - \theta_v + e^- \leftrightarrow \text{SPS}_v - \text{Cu}^{(I)} \text{s} \]  
\[ (3.47) \]

\[ \text{SPS}_v - \text{Cu}^{(I)} \text{s} + e^- \leftrightarrow \text{Cu}^{\text{incorporated}} + \text{SPS}_v - \theta_v \]  
\[ (3.48) \]

\[ \text{Cu}^{(II)} + \text{Cl}^- \text{s} - \theta_v + e^- \leftrightarrow \text{Cl}^- \text{s} - \text{Cu}^{(I)} \text{s} \]  
\[ (3.49) \]

\[ \text{Cl}^- \text{s} - \text{Cu}^{(I)} \text{s} + e^- \leftrightarrow \text{Cu}^{\text{incorporated}} + \text{Cl}^- \text{s} - \theta_v \]  
\[ (3.50) \]

In the above equations, SPS\(_r\)-Cu(I)\(_r\) and Cl\(_r\)-Cu(I)\(_r\) are cuprous ions that adsorb on SPS\(_r\)-\(\theta\) and Cl\(_r\)-\(\theta\), respectively.

The reaction rates are dependent upon the applied potential, specifically the surface overpotential. The reaction rate constant can be written as follows:
\[ k = k_0 \exp \left( -\frac{nF\beta\eta}{RT} \right) \]  

(3.51)

For simplicity, the potential dependency is considered only for cupric ion reduction. The reaction rates are summarized in Table 3-6 with parameters to be fitted identified in the right hand column.

<table>
<thead>
<tr>
<th>No</th>
<th>Reactant</th>
<th>Product</th>
<th>Reaction Types</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cu(II) + e⁻ + θ_v e⁻ ↔ Cu(I)_{ads}</td>
<td>Reductive</td>
<td>k₁, k₂, β₁</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Cu(I)_{ads} e⁻ ↔ Cu + θ_v</td>
<td>Reduction</td>
<td>k₂, k₂, β₂</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Cu(II) + SPS_{ads} e⁻ ↔ SPS_{ads}-Cu(I)_{ads}</td>
<td>Reductive</td>
<td>k₁, k₁, β₁</td>
<td></td>
</tr>
<tr>
<td></td>
<td>θ_v + e⁻ ↔ SPS_{ads}</td>
<td>Adsorption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SPS_{ads}-Cu(I)<em>{ads} e⁻ ↔ Cu + SPS</em>{ads} - θ_v</td>
<td>Reduction</td>
<td>k₂, k₂, β₂</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Cu(II) + Cl_{ads} - θ_v e⁻ ↔ Cl_{ads}-Cu(I)_{ads}</td>
<td>Reductive</td>
<td>k₁, k₁, β₁</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cl_{ads}-Cu(I)<em>{ads} e⁻ ↔ Cu + Cl</em>{ads} - θ_v</td>
<td>Adsorption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Cl_{ads}-Cu(I)<em>{ads} + e⁻ ↔ Cu + SPS</em>{ads} - θ_v</td>
<td>Reduction</td>
<td>k₂, k₂, β₂</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>SPS + θ_v ↔ SPS_{ads}</td>
<td>Adsorption</td>
<td>k_{sps1}, k_{sps-1}</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>SPS_{ads} + θ_v ↔ SPS_{ads} - θ_v</td>
<td>Adsorption</td>
<td>k_{sps2}</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>SPS_{ads} ↔ Incorporated + θ_v</td>
<td>Incorporation</td>
<td>k_{sps}</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Cl + θ_v ↔ Cl_{ads}</td>
<td>Adsorption</td>
<td>k_{c1}, k_{c1-1}</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Cl_{ads} + θ_v ↔ Cl_{ads} - θ_v</td>
<td>Adsorption</td>
<td>k_{cl2}</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Cl_{ads} ↔ Incorporated + θ_v</td>
<td>Incorporation</td>
<td>k_{cl}^{-i}</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>PEG + n·θ_v ↔ PEG_{ads}</td>
<td>Adsorption</td>
<td>k_{peg1}, k_{peg-1}</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>PEG_{ads} ↔ Incorporated + n·θ_v</td>
<td>Incorporation</td>
<td>k_{peg}^{-i}</td>
<td></td>
</tr>
</tbody>
</table>
3.4.2 Kinetic parameter fitting

Using the reactor scale model developed for the RDE apparatus, one can formulate the current density as a function of unknown kinetic parameters as in Equation (3.12). The error to be minimized is defined as the square sum of the difference between predicted and experimental current densities under the same operating conditions. Simplex algorithms were initially used to minimize the error and the Powell algorithm was used to improve the convergence speed\textsuperscript{15}. The kinetic parameters resulting from the parameter fitting are summarized in Table 3-7.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>$0.205 \cdot 10^{-7}$</td>
<td>$\text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{-1}$</td>
<td>$0.501 \cdot 10^{-5}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_2$</td>
<td>$0.220 \cdot 10^{-12}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{-2}$</td>
<td>$0.454 \cdot 10^{-5}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>$0.511$</td>
<td></td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>$0.796$</td>
<td></td>
</tr>
<tr>
<td>$k_{sps1}$</td>
<td>$0.719 \cdot 10^{-3}$</td>
<td>$\text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{sps\cdot 1}$</td>
<td>$0.205 \cdot 10^{-7}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{sps2}$</td>
<td>$0.288 \cdot 10^{-5}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{spsi}$</td>
<td>$0.838 \cdot 10^{-9}$</td>
<td>$\text{mol} \cdot \text{cm}^{-1} \cdot \text{A}^{-1}$</td>
</tr>
<tr>
<td>$k_{cl1}$</td>
<td>$0.785 \cdot 10^{-5}$</td>
<td>$\text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{cl\cdot 1}$</td>
<td>$0.101 \cdot 10^{-7}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{cl2}$</td>
<td>$0.103 \cdot 10^{-1}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{cl \cdot 1}$</td>
<td>$0.339 \cdot 10^{-7}$</td>
<td>$\text{mol} \cdot \text{cm}^{-1} \cdot \text{A}^{-1}$</td>
</tr>
<tr>
<td>$k_{peg1}$</td>
<td>$1.42 \cdot 10^{6}$</td>
<td>$\text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{peg\cdot 1}$</td>
<td>$0.170 \cdot 10^{-10}$</td>
<td>$\text{mol} \cdot \text{cm}^{-3} \cdot \text{sec}^{-1}$</td>
</tr>
<tr>
<td>$k_{peg \cdot 1}$</td>
<td>$0.679 \cdot 10^{-15}$</td>
<td>$\text{mol} \cdot \text{cm}^{-1} \cdot \text{A}^{-1}$</td>
</tr>
<tr>
<td>$n_{peg}$</td>
<td>8.38</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-12 shows that the model with the fitted parameters reproduces the effect of PEG concentration (Figure 3-7). The predictions shown in Figures 3-13 and 3-14 mirror the experimental observations in Figure 3-8 and 3-9 respectively. Figure 3-15 shows the effect of the Cl$^-$ concentration on the current densities. In general, the fitted results are
consistent with experimental data; especially, the trends for PEG and SPS are matched well. Additional data will be needed to reduce the small discrepancy between experimental results and fitted values, in particular the lack of sensitivity to Cl' concentration (Figure 3-15).

Figure 3-12. Current-Voltage response with PEG concentration change. Cl' 30 ppm, SPS 0 ppm
Figure 3-13. Current-Voltage response with SPS concentration change. Cl⁻ 30 ppm, PEG 200 ppm
Figure 3-14. Current-Voltage response with SPS concentration change. Cl 30 ppm, PEG 0 ppm
Figure 3-15. Current-Voltage response with Cl⁻ concentration change. SPS 40 ppm, PEG 200 ppm
3.5 Conclusion

A surface reaction mechanism for copper deposition in the presence of additives has been proposed based on a competitive adsorption and desorption model. The kinetic mechanism includes the steric hindrance of SPS preventing PEG adsorption, leading to an increase in current density. Statistically designed experiments were performed. A tertiary current distribution model is used to model the current density in the rotating disk electrode system. The kinetic parameters for PEG, SPS and Cl\(^-\) were extracted using nonlinear parameter fitting with Simplex and Powell algorithms.

There are three advantages of this kinetic mechanism formulation. First, the reaction rate k’s can expressed with Arrhenius form A·exp(-E/RT). By performing experiments at different temperatures, the activation energy E and the pre-exponential factor A can be obtained. This could be useful in thermal processes, such as electroless plating. Second, the nature of the competitive adsorption/desorption model readily enables addition of species in the reaction mechanism. This feature would be useful in describing multi-component systems, such as an alloy deposition processes. Finally, the proposed kinetic mechanism is applicable to both reactor and feature scale simulation.

**List of Symbols**

- \(c\) concentration, mol/cm\(^3\)
- \(k\) electric conductivity, S/cm
- \(n\) number of vacant sites taken by PEG adsorption
- \(r\) radial coordinate
- \(v\) velocity
y  axial coordinate

$c^\infty$  bulk concentration, mol/cm$^3$

$k_t$  surface reaction rates

$r_0$  radius of rotating disk

$r_i$  homogeneous reaction rate, mol/cm$^3$ sec

$u_i$  mobility of species i

$z_i$  charge in species i

$D$  diffusivity

$F$  Faraday constants, 96484.6 C/mol

$R$  gas constants, 8.314J/cm$^3$ K

$Z$  number of charges carried by cupric ions

$Cu(I)_{ads}$  adsorbed Cu(I) concentration

$Cu(II)$  surface concentration of Cu(II)

$SPS_{ads}$  adsorbed SPS concentration

$SPS_{ads}-\theta_v$  deposition sites near $SPS_{ads}$

$SPS_{ads}-Cu(I)_{ads}$  Cu(I) concentration adsorbed near $SPS_{ads}$

$Cl_{ads}$  adsorbed Cl concentration

$Cl_{ads}-\theta_v$  deposition sites near $Cl_{ads}$

$Cl_{ads}-Cu(I)_{ads}$  Cu(I) concentration adsorbed near $Cl_{ads}$

$PEG_{ads}$  adsorbed PEG concentration

$N_i$  molar flux of species i, mol/cm$^2$ sec
Greek Letters

\( \eta \)  
transformed elliptic coordinate

\( \mu \)  
viscosity, cp

\( \nu \)  
kinematic viscosity

\( \xi \)  
transformed radial coordinate

\( \zeta \)  
transformed axial coordinate

\( \nu_{ij} \)  
stoichiometric coefficients

\( \eta_c \)  
concentration overpotential, Volts

\( \eta_s \)  
surface overpotential, Volts

\( \theta_v \)  
vacant sites

\( \Omega \)  
rotation speed of disk

\( \Gamma_i \)  
Boundary

\( \Phi_S \)  
potential in seed layer, Volts

\( \Phi_L \)  
potential in liquid phase, Volts

3.6 References


6. J. D'Urso, Personal Communication


A. Frisch and M. J. Frisch, Gaussian 98, User's Reference 2nd Ed. (Gaussian Inc., PA, 1999).


Chapter

4. Reactor Scale Simulation of Copper Electrochemical Deposition Process

4.1 Introduction

Reactor scale simulation is a critical component in the multiscale simulation of the copper electrochemical deposition process by predicting potential and concentration distributions for different operating condition. Since potential and concentration fields near the wafer affect the filling of features, predicting feature fill at different operating condition requires accurate reactor scale modeling. Moreover, the reactor model must be based on the fundamental governing equations and account for all factors influencing potential and concentration distributions in order to be predictive.

The operation of an electrochemical deposition cell is affected by several factors. A large potential drop within the seed layer at the beginning of the deposition process, the so-called terminal effect, produces nonuniform film thickness\(^{1,3}\). Inlet concentration of additives, flow effects, and geometry of the reactor fall into the general category of mass transport effects. Nonuniform mass transport of cupric ions and additive species leads to growth rate variations. Predicting reactor operation at different additive concentrations requires including detailed chemistry mechanisms such as the one developed in Chapter 3. Also, slow mass transport in the liquid phase\(^{4,6}\) implies that transient effects have to be considered. A reactor scale simulation tool that combines all these effects has not been developed previously.
Development of a reactor scale simulation tool for a blank wafer (i.e. unpatterned wafer) is the first step. Since pattern effects need not to be included, the whole system can be described in one computational domain and a single set of governing equations. Also, the predicted film thickness variation across the wafer can be compared to experimental data as a validation of the reactor scale simulation tool.

The content of this chapter is as follows. The reactor scale model is formulated in terms of the governing equations, which are solved by the finite element method. The developed simulation tool is used to investigate various factors that affect film thickness uniformity. Finally, pulse plating is simulated to provide insight into the way pulse plating helps improve film thickness uniformity.

4.2 Governing Equations and Boundary Conditions

4.2.1 Governing equations

Simulation of electroplating cells requires complete description of concentration and potential distribution\(^7\,^8\). Thermal effect is neglected in this work, since the process is operated isothermally. The velocity field is needed to predict mass transport by convection. It is given by the Navier-Stokes equation.

\[
\rho \left( \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = -\nabla p + \mu \nabla^2 \mathbf{v} \quad (4.1)
\]

\[
\nabla \cdot \mathbf{v} = 0 \quad (4.2)
\]

In the simulation of the chemical vapor deposition (CVD) where reactants are in low concentrations in an inlet gas, the dilute approximation, is used to simplify the solution strategy. With this assumption, the Navier-Stokes equations may be decoupled from the chemical species conservation equations\(^9\). However, in the simulation of the
electrochemical deposition process, the dilute approximation is not applicable, since the concentration of a major solute, the sulfuric acid, is typically order of 2 M in comparison with the solvent concentration, 55 M. This problem can be circumvented by considering water and sulfuric acid as one pseudo-component. This assumption is reasonable, since the sulfuric acid is added to increase the conductivity of the electroplating bath without directly participating any of the plating reactions, and the concentration therefore remaining nearly constant. All the other components including cupric sulfate and additives have considerably lower concentrations than the solvent. So, the dilute approximation may be applied in the pseudo-component treatment.

For a given electrochemical deposition cell geometry and inlet flow condition, the Navier-Stokes equations are solved and then the velocity profile is substituted to the species transport equation to describe the convective mass transport. The transport properties of each component should be those in the water-sulfuric acid mixture, not in the pure water. The transport properties in the water-sulfuric acid mixture are given in Chapter 3 and the references therein.

The concentration distribution of each species is described by the following equation:

$$\frac{\partial c_k}{\partial t} = \nabla \cdot N_k + R_k$$  \hspace{1cm} (4.3)

$R_k$ represents the homogeneous reactions of species $k$. $N_k$ is the flux of species $k$ defined by:

$$N_k = -z_k u_k F c_k \nabla \Phi_L - D_k \nabla c_k + \nu c_k$$  \hspace{1cm} (4.4)

In Equation (4.4), $\nu$ is the velocity vector obtained from the solution of Equations (4.1) and (4.2). The potential in the liquid phase, $\Phi_L$ is found from the conservation of charge.
Current in the liquid phase is formulated with the gradients of concentration and potential as follows.

\[ i = -F^2 \nabla \Phi_L \sum_k z_k^2 u_k c_k + F \sum_k z_k D_k \nabla c_k \equiv -k_L \nabla \Phi_L - F \sum_k z_k D_k \nabla c_k \]  \hspace{1cm} (4.5)

By applying Equation (4.5), the conservation of charge takes the following form::

\[ \nabla \cdot (k_L \nabla \Phi_L) + F \sum_k z_k \nabla \cdot (D_k \nabla c_k) = 0 \]  \hspace{1cm} (4.6)

In the bulk liquid phase where the concentration gradients of the charged species are negligible, the diffusive contribution to the current is negligible and the potential distribution reduces to the Laplace equation\(^8\). In the copper film, the current is expressed only with the potential gradient:

\[ \nabla \cdot (k_g \nabla \Phi_g) = 0 \]  \hspace{1cm} (4.7)

4.2.2 Boundary conditions

Boundary conditions for the velocity are as follows. On the reactor walls, the velocity is specified as zero. On the wafer, the velocity in the angular direction is specified based on the wafer rotation speed and the distance from the center of the wafer. At the inlet of the electrochemical deposition cell, either average flow velocity calculated from total volumetric flow rate or a fully developed flow profile is specified.

Boundary conditions for the species transport are as follows. At the inlet of the reactor, a constant concentration of each species is specified. On the inert reactor wall, no flux boundary condition is specified. On the copper surface, the consumption of each species in surface reactions is balanced by diffusion to the surface:

\[ Flux_k = -D_k \nabla c_k \cdot n = \sum_j v_{kj} r_j \]  \hspace{1cm} (4.8)
Here $k$ is the species number, $j$ is the reaction number, $v_{kj}$ is the stoichiometric coefficient and $r_j$ is the surface reaction rate in mol/cm$^2$/sec. The surface reactions used in this work are given in Chapter 3. The set of the surface reactions includes the reduction of the cupric ion, and the adsorption, the desorption and the inclusion reactions of additives.

The boundary condition for the electrical potential is as follows. The reactor inlet is considered as an anode and a constant potential is specified as a boundary condition. If the inlet is different from the anode, a different anode can be specified. On the inert wall, a no flux boundary condition is specified. The boundary condition on the copper surface is a matching condition between the current value from the potential distribution and that from the mass transport. Additional concentration and surface overpotentials are specified on the copper surface. The concentration overpotential is the potential variation due to the concentration variation near the electrode. In the excess supporting electrolyte, the concentration overpotential can be described by the following equation:\(^8\):

$$\eta_C = \frac{RT}{nF} \ln \frac{C_M}{C_M^*}$$  \hspace{1cm} (4.9)

The surface overpotential accounts for the potential drop due to the slow surface reaction. The surface overpotential, $\eta_S$, is given by a balance between the potential in liquid phase, $\Phi_L$, the concentration overpotential, $\eta_C$, and the potential in the solid phase $\Phi_S$.

$$\Phi_S + \eta_S + \eta_C = \Phi_L$$  \hspace{1cm} (4.10)

Due to the potential drop across the wafer, $\Phi_S$ varies for each location. At the edge of the wafer, either potential or current is specified as a boundary condition to the potential of
the copper film. These potential and current can be either constant or function in time depending upon the operation mode.

4.3 Numerical Methods

The governing equations are solved by the finite element method\textsuperscript{10-15} implemented in Fortran 77. Details of the numerical methods are described in the following sections.

4.3.1 Mesh generation

A commercial mesh generation software, ICEM CFD\textsuperscript{16} is used to generate the FEM mesh. The mesh generation step is composed of two steps: reactor geometry drawing and mesh generation. The reactor geometry is defined with computer aided design (CAD) software. Different materials for each volute element such as wall, fluid and seed layer can be assigned. Different properties and governing equations for different materials or phases are assigned in the main program. Different boundaries, such as inlet, outlet, anode, cathode, and wafer surface can be indicated. Boundary conditions corresponding to each physical boundary are assigned in the simulation code. The mesh generation is performed with a module, called Hexa\textsuperscript{TM} in ICEM CFD. Hexa creates a mesh file composed of list of elements and nodal coordinates based on the reactor geometry. The mesh file is read from the main program.

4.3.2 Finite element discretization

Within each element, the variables are approximated with the following equations:
\[
\tilde{v}(\tilde{x}) = \sum_{i=1}^{N_q} \tilde{v}_i \phi^i(\tilde{\xi}) , \\
c^i(x) = \sum_{i=1}^{N_q} c^i \phi^i(\tilde{\xi}) , \\
\Phi(\tilde{x}) = \sum_{i=1}^{N_q} \Phi^i(\tilde{\xi}) ,
\]

(4.11)  
(4.12)  
(4.13)

where \( \tilde{x} \) are global position vector and \( \tilde{\xi} \) are local position vector whose component is normalized between 0 and 1. The interpolation function \( \phi^i(\tilde{\xi}) \) is biquadratic Lagrangian polynomial function in two-dimensional cases. \( N_q \) is the node number for biquadratic elements, which is equal to 9 in two dimensional cases\textsuperscript{13,14}. For pressure, there is only first order derivative in the governing equation shown in Equation (4.1), and thus a linear interpolation function is used.

\[
p(\tilde{x}) = \sum_{i=1}^{N_t} p^i \psi^i(\tilde{\xi})
\]

(4.14)

\( N_t \) is the number of node in the linear element, which is 4 in the two dimensional case. The interpolation \( \psi^i(\tilde{\xi}) \) is bilinear Lagrangian polynomial function. Isoparametric mapping is used for the interpolation of the coordinates. The coordinate vector \( \tilde{x} \) is approximated with the same interpolation function to other variables.

\[
\tilde{x} = \sum_{i=1}^{N_q} \tilde{x}_i \phi^i(\tilde{\xi})
\]

(4.15)
4.3.3 Weak form of the governing equation

The weighted-integral form of the governing equation is used to generate the algebraic equation and the weight functions are selected to be the same as the approximation function, or so called Galerkin models\textsuperscript{10-15}.

The weak form of each governing equation is:

Continuity:
\[
\int_{\Omega} (\nabla \cdot v) \psi^i dV = 0
\] (4.16)

Momentum balance:
\[
\int_{\Omega} \left( \frac{\partial v}{\partial t} + v \cdot \nabla v \right) \phi^i dV = -\int_{\Omega} \nabla p \phi^i dV - \int_{\Omega} \mu \nabla \phi \cdot \nabla v dV + \int_{\partial \Omega} \mu \phi \nabla v dS
\] (4.17)

Species conservation:
\[
\int_{\Omega} \frac{\partial c_k^i}{\partial t} \phi^i dV = \int_{\partial \Omega} N_k \phi^i dS - \int_{\Omega} \nabla \phi \cdot \nabla v dV + \int_{\Omega} R_k \phi^i dV
\] (4.18)

Potential distribution in the liquid phase:
\[
0 = \int_{\partial \Omega} \left( -k_L \nabla \Phi_L - F \sum_k z_k D_k \nabla c_k \right) \phi^i dS
\]
\[-\int_{\Omega} \left( -k_L \nabla \Phi_L - F \sum_k z_k D_k \nabla c_k \right) \cdot \nabla \phi^i dV
\] (4.19)

Potential distribution in the solid phase:
\[
0 = \int_{\partial \Omega} (k_s \nabla \Phi_s \phi^i dS - \int_{\Omega} (k_s \nabla \Phi_s) \cdot \nabla \phi^i dV
\] (4.20)

Equations (4.16) to (4.20) and the approximated variables, Equations (4.11) to (4.15) are combined to generate a system of differential algebraic equation, which is solved by the methods described in Sections 4.3.5 and 4.3.6.
4.3.4 Terminal effect

The potential distribution equation in the copper film is described by Equation (4.7). The thickness of copper film is order of microns, and the dimension of electrochemical deposition cell is order of 10 cm. Resolving the two phases in a single computational domain would require very high refinement of the mesh near the wafer and within the copper film. Furthermore, the film thickness changes as the film grows. Remeshing could in principle be done at each time step, but this would be slow and computationally expensive. Deligianni et al. developed a method that efficiently addresses both issues\(^1\).

In cylindrical coordinates, Equation (4.7) takes the form:

\[
\frac{1}{r} \frac{\partial}{\partial r} \left( k_s r \frac{\partial \Phi_s}{\partial r} \right) + \frac{\partial}{\partial z} \left( k_s \frac{\partial \Phi_s}{\partial z} \right) = 0 \tag{4.21}
\]

\[
0 \leq z \leq \delta(r, t) \tag{4.22}
\]

Here \(\delta(r, t)\) is the thickness of copper film on the wafer and varies with radial position and time. Its evolution over time is expressed as follows.

\[
\frac{\partial \delta(r, t)}{\partial t} = \frac{i(r) MW_{Cu}}{nF \rho_{Cu}} \tag{4.23}
\]

The axial coordinate \(z\) can be rescaled by the following transformation.

\[
\zeta = \zeta_c \cdot \left( \frac{z}{\delta(r, t)} \right), \tag{4.24}
\]

\[
0 \leq \zeta \leq \zeta_c, \tag{4.25}
\]
where $\zeta$ is an arbitrary constant. With this transformation, the new axial coordinate $\zeta$ is bounded. Note that $\zeta$ can be arbitrary defined as a large enough constant. Applying this coordinate transform, the derivative with respect to $z$ takes the form:

$$
\frac{\partial \Phi_z}{\partial z} = \frac{\partial \Phi_z}{\partial \zeta} \frac{\partial \zeta}{\partial z} = \frac{\partial \Phi_z}{\partial \zeta} \frac{\zeta_c}{\delta(r,t)}
$$

(4.26)

$$
\frac{\partial}{\partial z} \left( k_s \frac{\partial \Phi_s}{\partial z} \right) = \frac{\partial}{\partial \zeta} \left( k_s \frac{\partial \Phi_s}{\partial z} \right) \frac{\partial \zeta}{\partial z} = k_s \left( \frac{\zeta_c}{\delta(r,t)} \right)^2 \frac{\partial^2 \Phi_s}{\partial \zeta^2}
$$

(4.27)

By defining the nonisotropic conductivity

$$
k_{s,r} \equiv k_s, \text{ and}$$

(4.28)

$$
k_{s,\zeta} \equiv k_s \left( \frac{\zeta_c}{\delta(r,t)} \right)^2,
$$

(4.29)

the original governing equation in seed layer is converted to :

$$
k_{s,r} \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \Phi_z}{\partial r} \right) + k_{s,\zeta} (r,t) \frac{\partial^2 \Phi_z}{\partial \zeta^2} = 0
$$

(4.30)

The axial conductivity $k_{s,\zeta}$ is updated at each time step with new local copper film thickness. With this new formulation, the copper film can be resolved in the same mesh with the liquid phase and the complexity arising from the change in copper film geometry is avoided.

4.3.5 Solution of nonlinear algebraic equation:

The steady state version of the weak form of the equations is highly nonlinear and needs to be solved iteratively. The Newton-Raphson method is used to solve the systems of the
nonlinear equations arising from the FEM approach\textsuperscript{17}. At iteration \(n+1\), the Jacobian (\(J\)) and residual (\(R\)) are computed with the solution from previous iteration, \(n\). The solution vector \(u^{n+1}\) is updated with the vector \(\delta\) that is the solution of the following equation.

\[
J(u^n) \cdot \delta = -R(u^n) \tag{4.31}
\]

\[
u^{n+1} = u^n + \delta \tag{4.32}
\]

This procedure repeated until the norm of the correction vector \(|\delta|\) becomes smaller than the specified error tolerance. If residual and Jacobian are computed correctly, the size of norm decreases quadratically at each iteration in a neighborhood of the solution. The linear system Equation (4.31) is solved with a frontal solver developed for finite element computation\textsuperscript{18,19}.

4.3.6 Time integration

The system of nonlinear equations (4.16-20) is a typical example of differential-algebraic equation (DAE)\textsuperscript{20,21}. The potential distribution is time invariant, whereas species transport equations are dependent upon time. The DAE solver, DSL48S, is used to solve this system\textsuperscript{22,23}. A consistent initial condition that satisfies the original DAE needs to be provided. In this work, the steady state solution at zero current is used as an initial condition.

4.4 Experimental Procedure

Experiments were performed at Novellus Systems, Inc., San Jose, CA with Novellus SABRE\textsuperscript{TM} electroplating tool with a 50 nm of copper seed layer deposited with physical
vapor deposition (PVD) and a constant current of 6 A was applied until a desired film thickness was obtained. The experimental conditions are summarized in Table 4-1.

<table>
<thead>
<tr>
<th><strong>Parameter</strong></th>
<th><strong>Value</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation Speed</td>
<td>100 RPM</td>
</tr>
<tr>
<td>Flow Rate from center jet</td>
<td>8 slm</td>
</tr>
<tr>
<td>H₂SO₄ Concentration</td>
<td>1.78 M</td>
</tr>
<tr>
<td>CuSO₄ Concentration</td>
<td>0.11 M</td>
</tr>
<tr>
<td>HCl Concentration</td>
<td>50 PPM</td>
</tr>
<tr>
<td>PPG 1200 Concentration</td>
<td>300 PPM</td>
</tr>
<tr>
<td>SPS Concentration</td>
<td>6 PPM</td>
</tr>
<tr>
<td>Applied current</td>
<td>6 A</td>
</tr>
<tr>
<td>Seed layer thickness</td>
<td>50 nm</td>
</tr>
<tr>
<td>Wafer size</td>
<td>191.5 mm diameter</td>
</tr>
</tbody>
</table>

*: Poly ethylene glycol

The thickness of copper film was analyzed at six different times as shown in Table 4-2. The experiments were performed with a single wafer. Copper was deposited on a single wafer for a specified period of time, the sheet resistance was measured and then the wafer was plated again until the next analysis point. This approach saves resources compared to running 6 different cases with different durations in parallel. However, the sequential method affects the profile evolution, since interrupting the operation replenishes consumed additives and cupric ion and therefore decreases the mass transport effect.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Thickness (µm)</th>
<th>Coulomb</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1µ</td>
<td>0.095</td>
<td>74.06</td>
<td>12.34</td>
</tr>
<tr>
<td>0.2µ</td>
<td>0.19</td>
<td>148.11</td>
<td>24.69</td>
</tr>
<tr>
<td>0.5µ</td>
<td>0.47</td>
<td>366.38</td>
<td>61.06</td>
</tr>
<tr>
<td>1.0µ</td>
<td>0.99</td>
<td>771.74</td>
<td>128.62</td>
</tr>
<tr>
<td>2.0µ</td>
<td>1.98</td>
<td>1543.49</td>
<td>257.25</td>
</tr>
</tbody>
</table>

The measurement of the thickness was performed with a 121-point sheet resistance measurement. Measurement points were distributed evenly through the azimuthal
coordinate at each radial position to mitigate any azimuthal variation. Sheet resistance measurements are simpler to perform than typical profilometry, and are unaffected by wafer bending\textsuperscript{24}. The sheet resistance was converted into film thickness by the following equation:

\[
\text{thickness} = 2.2 \cdot 10^{-6} \Omega \cdot \text{cm}/\text{sheet resistance},
\]

where \(2.2 \cdot 10^{-6} \Omega \cdot \text{cm}\) is the resistivity of bulk copper.

4.5 Results and Discussions

4.5.1 Comparison of simulation result with experimental data

Schematics of the electrochemical deposition cell and the computational domain are shown in Figure 4-1. Only the cross section is used in the computational domain, since the electrochemical deposition cell was considered to be axisymmetric. The FEM computational mesh is shown in Figure 4-2. Since PPG was used instead of PEG, as in the original kinetic experiments (see Chapter 3), a new set of kinetic parameters were fitted to experimental data by using the same procedure as described in Chapter 3. The

![Figure 4-1. Computational domain and boundary condition of reactor scale simulation](image)

85
new kinetic parameters are given in Table in Appendix. The turbulent flow induced by
the high rotation speed of the wafer (100 RPM) is expected to create a well-mixed
condition and thus, uniform species concentrations. Thus, a secondary current
distribution model equivalent to a CSTR approximation except for the spatial variation in
the potential is used to model the system. In Sections 4.5.2 and 4.5.3, convective mass
transport effects are considered in the simulation at lower wafer rotation speed, where the
flow field is in the laminar flow regime.

![Diagram of Copper film, Cathode, Outlet, and Inlet, anode]

Figure 4-2. Computational mesh for the reactor scale simulation

Figure 4-3 shows the experimental sheet resistance at different positions on
the wafer at different times. The data in Figure 4-3 are average over sheet resistance data
for different azimuthal angles and fixed radial position. A comparison between
experimental and predicted thickness distributions is shown in Figure 4-4. Predicted and
measured results agree at all time. Figure 4-5 shows the relative uniformity vs. thickness.
The relative uniformity is defined as:
\[ \text{average} = \frac{\int_0^R \delta r \, dr}{2\pi \int_0^R r \, dr} \equiv \bar{\delta} \]  
\[ \text{standard deviation} = \sqrt{\frac{\int_0^R (\delta - \bar{\delta})^2 r \, dr}{2\pi \int_0^R r \, dr}} \]  
\[ \% \text{ Nonuniformity} = \frac{\text{(standard deviation)}}{(\text{average})} \times 100\% \]

The simulation underpredicts the peak values and slightly overpredicts at the higher thickness, but good agreement is achieved overall.

It is encouraging that with the secondary current distribution model, the simulation results matches the experimental film thickness evolution well. This can be

Figure 4-3. Sheet resistivity of wafer at different times. Real time corresponding to the legend is in Table 4-2.
explained by the CSTR, i.e. well mixed assumption, being realized by the turbulence. The measurement method also influenced the result. Each measurement interrupted the electroplating cell providing enough time for the concentration field to replenish. Nevertheless, the good agreement between simulation and experiment indicates that the potential distribution and the surface kinetic mechanisms properly represent the system. However, when the mass transport effects become important, a more rigorous mass transport mechanism including convection need to be included.

It is also noteworthy to emphasize the general applicability of the surface kinetic mechanism. Although our original kinetic mechanism was developed for the PEG/SPS/CT system, the present comparison shows that it can be extended to PPG/SPS/CT.

![Graph](image)

**Figure 4-4.** Comparison between experimental and predicted thickness distributions at different times. Y-axis is in log scale.
4.5.2 Parameter study: factors affecting the uniformity of thin film deposition

The developed reaction level model can be used to investigate various causes of film thickness variations in the electrochemical deposition cell. The conditions chosen for the present case study are summarized in Table 4-3.

![Graph showing % nonuniformity vs. film thickness](image)

Figure 4-5. Relative nonuniformity of experimental and predicted results.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation Speed</td>
<td>0 RPM</td>
</tr>
<tr>
<td>Flow Rate from center jet</td>
<td>0.01 slm</td>
</tr>
<tr>
<td>H₂SO₄ Concentration</td>
<td>1.78 M</td>
</tr>
<tr>
<td>CuSO₄ Concentration</td>
<td>0.11 M</td>
</tr>
<tr>
<td>Cl Concentration</td>
<td>50 PPM</td>
</tr>
<tr>
<td>PPG 1200 Concentration</td>
<td>30 PPM</td>
</tr>
<tr>
<td>SPS Concentration</td>
<td>6 PPM</td>
</tr>
<tr>
<td>Applied voltage</td>
<td>0.01 V</td>
</tr>
<tr>
<td>Seed layer thickness</td>
<td>1 µm</td>
</tr>
<tr>
<td>Wafer size</td>
<td>191.5 mm diameter</td>
</tr>
</tbody>
</table>
A very small flow rate is intentionally chosen to reduce numerical difficulty, since with realistic flow rates (~ 10 slm), the computational mesh needs to be fine enough to resolve a thin mass transport boundary of order 100 μm. The low flow rate implies that the limiting current of the system is expected to be small. This is indeed the case; at the applied voltage of 0.1 V, the current became limited at 0.16 A (see Figure 4-6).

![Graph](image)

Figure 4-6. Current-Voltage response of electroplating cell with conditions given in Table 4-3.

*Nonuniformity attributed to seed layer thickness: terminal effect*

Film thickness variations caused by the terminal effect has been demonstrated in Section 4.5.1. The terminal effect is the primary cause of film thickness variation in the absence of mass transport. Figure 4-7 compares the uniformity of the current density at two different seed layer thickness. In the seed layer of 50 nm, there is a large variation in current density, and nonuniform potential distribution follows. With significantly higher thickness, 1 μm, the current density is nearly flat and the potential distribution is uniform.
Figure 4-7. Influence of terminal effect on the variation in secondary current distribution for two different seed layer thickness. (a) 50 nm, (b) 1 μm
Figure 4-8. Comparison between secondary and tertiary current distribution in current density. The tertiary current distribution shows larger variation contributed to the irregular boundary layer thickness created by the outlet.
Figure 4-9. Effect of applied voltage on the uniformity of the current density. (a) 0.01 V (b) 0.3V. As system reaches to the mass transfer controlled regime, the nonuniformity due to uneven mass transport boundary layer thickness magnifies.
Figure 4-10. Nonuniformity due to additive. With strong additive effect or strong suppression reduces the nonuniformity.
Figure 4.11. Effect of the reactor configuration on the uniformity of the current density at two different applied voltage. (a) 0.01 V (b) 0.3 V. At higher voltage, the nonuniformity increases.
Variations caused by the mass transport boundary layer

In this case, the thick seed layer of 1 μm is chosen so that there is no terminal effect. The same case as in Figure 4-7 and the corresponding case with mass transport effects, or tertiary current distribution are compared in Figure 4-8. Although the secondary current distribution is nearly flat, the tertiary current distribution shows variation especially at the edge of the wafer. The explanation is in the concentration contour of the cupric ion. At the edge of the wafer, the boundary layer thickness is smaller due to the higher flow velocity created near the exit.

Film thickness variations created by the applied voltage

With a higher applied voltage of 0.3 V relative to the case in Figure 4-8, there is a significant increase in current density at the edge of the wafer (see Figure 4-9). The results demonstrate that in a purely mass transport regime, such as the limiting current density, variations in mass transport boundary layer thickness have a dramatic effect on film growth rate uniformity.

Film thickness variations induced by additives

Adding a additive to the system influence the current distribution. In this case study, the effect of PPG is simulated. PPG blocks the active site and reduces the current density at the applied voltage. A variation of PPG concentration will therefore change the current density. In this simulation, the adsorption reaction rate $k_{\text{ppg}}$ in Appendix (Section 4.7) is changed instead of changing the concentration of PPG. Current density values for corresponding cases are shown in Figure 4-10. The effect is similar to increasing voltage
(Figure 4-9). With lower PPG concentration, the system is shifted to the limiting current regime, since the current density is increased with lower PPG concentration.

The effect of the PPG on the current density is expected from the qualitative trends of the PPG. The effect of the other additives can be guessed in the same way. For instance, if the bulk concentration of the SPS changes, the current density will change in a opposite trends to the PPG (see Chapter 3).

*Influence of reactor geometry*

The radius of the inlet of the cell is reduced by approximately 10 times. With a lower applied voltage, the current density is almost flat (Figure 4-11), which supports the observation that the system is less affected by the mass transport if the system is away from the limiting current regime. However, if the applied voltage is 0.3 V, a substantial film thickness variation is observed. The characteristic shape of the current density can be rationalized in terms of the contour plot of cupric ion concentration with the streamlines of the flow field. Because of the location of the inlet, the boundary layer thickness is very small at the center of the wafer. At the edge of the wafer, the boundary layer thickness is also thin because of the presence of the outlet.

4.5.3 *A case study: pulse plating simulation*

The developed reaction simulation allows explaining technically important operating modes, specifically the pulse plating mode. In pulse plating, currents are applied intermittently. Pulse plating can alleviate variation in film thickness, since the non-depositing phase provides time for the concentration field to recover. Modeling of the
pulse plating requires rigorous handling of transient mass transport effects and detailed surface reactions—the electroplating essentials emphasized in thesis.

The same simulation conditions for Figure 4-11 are used to maximize the effect of the mass transport. The pulsed current is applied as a boundary condition at the edge of the copper film. Two different pulse regimes are considered (Figure 4-12).

In the first case, the current changes as a step function between 0 and $i_{\text{max}}$ for the same duration repeatedly for ten periods. The period of the depositing mode was estimated from a separate simulation in order to estimate the time when the cupric ions start to deplete locally, which roughly corresponds to the characteristic diffusion time scale. In the second case, the constant current is applied for half the duration of the first

Case 1:

![Diagram of Case 1: Pulsed Current with $\tau=14000$ sec for ten periods and Case 2: Constant Current $i_{\text{max}}=0.078$A]

Figure 4-12. Input current pulse for comparison. Case 1: pulsed current with $\tau=14000$ sec for ten periods, Case 2: constant current. $i_{\text{max}} = 0.078$A.
case in order to deposit the same amount of materials.

The concentration profile of case 1 does not have a steady state since the pulsed current force the system to oscillate. Figure 4-13 shows the local variation of cupric ions at two different positions corresponding to case 1. The slow mass transport implies that the local concentration does not respond immediately to the current change. Instead, serrate forms of cupric ion concentration time variations were observed. The concentration decreases when the current is on, and the concentration increases when the current is off. These current-off periods provide time for the system to smooth out gradients. As can be seen in Figure 4-13, point 1 has thinner diffusion layer, so that the concentration recovers nearly to the bulk concentration. However, at position 2, the concentration recovers only up to 80 % of the bulk concentration. The surface

![Diagram showing oscillation of cupric ion concentration at different location](image)

Figure 4-13. Oscillation of cupric ion concentration at different location. Solid line: point 1, dotted line: point 2
concentration would be recovered to a higher value, if the relaxation time was increased. In this way, the system could be optimized to produce a more uniform film thickness.

Figure 4-14 shows the standard deviation in film thickness as a function of average film thickness. As expected, variations in film thickness at constant current are higher than the pulse plating, in this specific case, twice as high. It is worthwhile to compare variations caused by mass transport and the terminal effect. The variations created by the terminal effect are eventually eliminated as the seed layer thickens. On the contrary, the nonuniformity created by the mass transport persists even after a long periods of time.

![Graph of standard deviation vs. thickness](image)

*Figure 4-14. Nonuniformity at different condition.*

4.6 Conclusions

Reactor scale simulations for copper electrochemical deposition have been developed based on FEM solutions of the governing equations. For the multiscale simulation, the
reactor scale simulations provide boundary conditions for the feature scale. The factors affecting variation in film thickness including terminal effect, mass transport boundary layer thickness, additives, reactor geometry and applied voltage were delineated and simulated with potential and mass transport analysis using the simulations. These effects are commingled in actual systems so that a quantitative simulation tool is necessary to identify the main causes of film thickness variations and optimize deposition conditions. The terminal effect attributed to the high initial resistance of the seed layer was analyzed and it disappeared as the seed layer grew. However, nonuniformity caused by mass transport effects persists at long time. Finally, the simulation tool was used to predict the performance of pulse plating. The simulations showed that pulse plating provided time for the system to relax the concentration profiles. The developed reactor scale simulations allow investigation of the behavior of electroplating cells. Moreover, it serves a critical component in the multiscale simulation of copper electrochemical deposition system with patterned wafer.

List of Symbols

\( v \) \hspace{1cm} \text{velocity} \\
\( p \) \hspace{1cm} \text{pressure} \\
\( z \) \hspace{1cm} \text{axial coordinate} \\
\( r_j \) \hspace{1cm} \text{surface reaction rates, mol/cm}^2/\text{sec} \\
\( F \) \hspace{1cm} \text{Faraday constants, 96484.6 C/mol} \\
\( c_k \) \hspace{1cm} \text{concentration of species k, mol/cm}^3 \\
\( k \) \hspace{1cm} \text{electric conductivity, S/cm}
\( k_L \) electric conductivity in electrochemical deposition bath, \( \Omega^{-1} \text{sec}^{-1} \)

\( k_S \) electric conductivity in solid copper, \( \Omega^{-1} \text{sec}^{-1} \)

\( k_{S,r} \) transformed electric conductivity in radial direction

\( k_{S,\zeta} \) electric conductivity in axial direction

\( u_k \) mobility of species \( k \)

\( z_k \) charge in species \( k \)

\( C_M \) metal ion concentration, \( \text{M} \)

\( C_{M}^{\omega} \) bulk concentration of metal ion, \( \text{M} \)

\( D_k \) diffusivity of species \( k \)

\( N_k \) molar flux of species \( k \), \( \text{mol/cm}^2 \text{sec} \)

\( R_k \) homogeneous reaction rate, \( \text{mol/cm}^3 \text{sec} \)

**Greek Letters**

\( \delta \) seed layer thickness at different time and location on the wafer

\( \rho \) density

\( \mu \) viscosity, cp

\( \zeta \) transform axial coordinate in seed layer

\( \nu_{ij} \) stoichiometric coefficients

\( \eta_c \) concentration overpotential, Volts

\( \eta_s \) surface overpotential, Volts

\( \phi^i \) biquadratic interpolation for node \( i \)
\( \psi^i \) bilinear interpolation function for node i

\( \Gamma_i \) boundary

\( \Phi_S \) potential in seed layer, Volts

\( \Phi_L \) potential in liquid phase, Volts

4.7 Appendix: Kinetic Parameters for PPG/SPS/Cl⁻

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k_1 )</td>
<td>0.641\cdot10^{-7}</td>
<td>sec⁻¹</td>
</tr>
<tr>
<td>( k_{-1} )</td>
<td>0.604\cdot10^{-5}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( k_2 )</td>
<td>0.126\cdot10^{-12}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( k_{-2} )</td>
<td>0.618\cdot10^{-5}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.511</td>
<td></td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.796</td>
<td></td>
</tr>
<tr>
<td>( k_{sp1} )</td>
<td>0.399\cdot10^{-5}</td>
<td>sec⁻¹</td>
</tr>
<tr>
<td>( k_{sp-1} )</td>
<td>0.211\cdot10^{-7}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( k_{sp2} )</td>
<td>0.185\cdot10^{-6}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( k_{sp1} )</td>
<td>0.163\cdot10^{-6}</td>
<td>mol·cm⁻³·A⁻¹</td>
</tr>
<tr>
<td>( k_{ci1} )</td>
<td>0.952\cdot10^{-5}</td>
<td>sec⁻¹</td>
</tr>
<tr>
<td>( k_{ci1} )</td>
<td>0.968\cdot10^{-8}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( k_{ci2} )</td>
<td>0.932\cdot10^{-2}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( k_{ci} )</td>
<td>0.345\cdot10^{-7}</td>
<td>mol·cm⁻³·A⁻¹</td>
</tr>
<tr>
<td>( k_{pp1} )</td>
<td>0.1</td>
<td>sec⁻¹</td>
</tr>
<tr>
<td>( k_{pp-1} )</td>
<td>0.202\cdot10^{-8}</td>
<td>mol·cm⁻³·sec⁻¹</td>
</tr>
<tr>
<td>( k_{pp1} )</td>
<td>0.375\cdot10^{-22}</td>
<td>mol·cm⁻³·A⁻¹</td>
</tr>
<tr>
<td>( n_{ppg} )</td>
<td>1.38</td>
<td></td>
</tr>
</tbody>
</table>

4.8 Reference


Chapter

5. Feature Scale Simulation of Copper Electrochemical Deposition

5.1 Introduction

In the previous chapter, reactor scale modeling of copper electrochemical deposition was discussed. The reactor scale model provides detailed information needed to predict deposition on a blank wafer. Predictive simulations of deposition on a patterned wafer would, however, require a multiscale simulation tool. Feature scale modeling is addressed in this chapter as a step toward developing a multiscale simulation capability.

Important physical phenomena on the feature scale (∼ 0.1 μm) include mass transport, potential distribution, and surface reactions. The continual geometry change caused by deposition of copper in the narrow space within trenches and vias is another important consideration. The underlying fundamental balance equations need to be formulated to have predictive simulation of feature scale deposition. Moreover, since the feature scale model ultimately will be a component in the multiscale model, the numerical method should be suitable for multiscale scheme simulations. Multiscale simulations involve iteration between the different simulation modules, so all building block modules have to be robust and fast.

In this chapter, the governing equations for feature scale modeling are developed and the underlying assumptions addressed. Boundary conditions and their relation to multiscale simulation are characterized. Numerical methods are selected on the basis of computational efficiency and robustness.
5.2 Governing Equation

5.1.1 Mass transport

Species transport in the liquid phase is carried out by convection, diffusion and migration. The general species conservation equation is as follows\(^1\):

\[
\frac{\partial C_i}{\partial t} = \nabla \cdot (-z_i u_i F C_i \nabla \Phi - D_i \nabla C_i + \nu C_i) + R_i \tag{5.1}
\]

Although this equation could in principle be solved numerically, it is worthwhile to eliminate insignificant terms. The importance of each term is determined by comparing the relative magnitudes of each term.

The convective mass transport in the feature scale is created by wafer rotation and film growth. Convection created by the wafer rotation is insignificant since the relative velocity of fluid with respect to the wafer surface is negligible near the surface and within trenches. Convection created by film growth can be assessed by estimating the Peclet number

\[
P_e = \frac{L_c V_c}{D} = \frac{L_c^2}{t_c D}, \tag{5.2}
\]

where the nominal velocity created by the feature geometry change is represented as the ratio of the feature dimensions over the total time to fill the feature, or \(L_c/t_c\). Substituting representative values, \(L_c = 0.1 \mu m\), \(t_c = 100 \text{ sec}\), and \(D = 1 \cdot 10^{-6} \text{ cm}^2/\text{sec}\),

\[
P_e = \frac{(0.1 \cdot 10^6 \mu m)^2}{100 \text{sec} \cdot 10^{-40} \text{ m}^2/\text{sec}} = 10^{-2} \tag{5.3}
\]

This value is fairly small, indicating that convection effects from filling of feature can be neglected.
The relative magnitude of migration over diffusion can be approximated as follows:

\[
\frac{\text{migration diffusion}}{\text{diffusion}} = \frac{z_i u_i FC_i \nabla \Phi}{D_i \nabla C_i}
\]  

(5.4)

In order to estimate this ratio, the value of each parameter needs to be estimated. The potential gradient may be approximated by the current density value and the conductivity of electroplating bath as follows:

\[
\nabla \Phi = \frac{i}{k} 
\]  

(5.5)

Under typical operating condition, \(i = 10 \text{mA/cm}^2\), \(k = 60 \text{\Omega}^{-1} \text{cm}^{-1}\),

\[
\nabla \Phi = \frac{10 \text{mA/cm}^2}{60 \text{\Omega}^{-1} \text{cm}^{-1}} = 1.67 \times 10^{-4} \text{V/cm}
\]  

(5.6)

Assuming that the majority of the concentration gradient lies within feature scale, the concentration gradient can be estimated as follows:

\[
\nabla C_i = \frac{C_i^{\text{bulk}}}{L_c}
\]  

(5.7)

Finally, the mobility of species \(i\) can be estimated using the Nernst-Einstein equation\(^1\),

\[
D_i = RT u_i
\]  

(5.8)

Inserting Equations (5-6), (5-7), and (5-8) into Equation (5-4), the ratio of the migration and the diffusion becomes

\[
\frac{\text{migration diffusion}}{\text{diffusion}} = \frac{z_i F \nabla \Phi}{RT / L_c} \left| \frac{96485 \text{C/mol} \cdot 1.67 \cdot 10^{-2} \text{V/cm}}{8.314 \text{J/mol} \cdot \text{K} \cdot 300 \text{K} / 0.1 \cdot 10^{-4} \text{cm}} \right| = 6.46 \times 10^{-6}
\]  

(5.9)

This value indicates that the migration is negligible compared to the diffusion.
The above analysis assumes, however, that the species is consumed on the electrode surface, and there is major concentration gradient within feature scale computational domain. On the other hand, if the species is not consumed or created in case of the supporting electrolyte, the concentration gradient is small and the migration is the dominant mass transport mechanism.

The ratio between the diffusion and the time derivative terms can be expressed as follows:

\[
\frac{\text{time derivative}}{\text{diffusion}} = \frac{\partial C_i}{\partial t} \left[ \frac{C_i^{\text{Bulk}}}{D_i} \right] = \frac{L_c^2}{D_i t_c} \quad (5.10)
\]

This is exactly the same formula as Equation (5.2), and thus the time derivative term can be neglected in estimating concentration distribution. In other words, the pseudo-steady state approximation applies.

In summary, the species transport equation (5.1) in the feature scale without homogeneous reaction with constant diffusivity satisfies the Laplace equation as follows:

\[
0 = \nabla \cdot (\nabla C_i) \quad (5.11)
\]

This equation needs to be solved for cupric ion and additive species. For the supporting electrolyte, the concentration is assumed to be homogeneous within the computational domain, since the species is neither consumed nor generated, and its bulk concentration is so high that any variation of concentration due to migration can be neglected.

5.1.2 Potential distribution

The conservation of charge satisfies the following equation¹:

\[
\nabla \cdot k \nabla \Phi + F \sum_i z_i \nabla \cdot (D_i \nabla C_i) = 0 \quad (5.12)
\]
When the electroplating bath includes excess amount of supporting electrolyte, the electric conductivity of the electroplating bath is constant and migration of the supporting electrolyte dominates the charge transfer. In this case, the potential obeys the Laplace equation\textsuperscript{1,2}.

\[ 0 = \nabla \cdot (\nabla \Phi) \]  
(5.13)

It is important to note that both the species transport equation (5.11) and the potential distribution equation (5.13) satisfies the Laplace equation, which indicates that the same numerical method could be used.

\subsection*{5.1.3 Boundary condition}

The computational domain and the boundary conditions of the feature scale are shown in Figure 5-1. At the interface between feature and reactor scale \( \Gamma_i \), typically bulk concentrations and the average current density are applied as boundary conditions\textsuperscript{2}. However, the concentration at the boundary in general differs from the bulk concentration unless mass transport from reactor scale is very fast compared to the consumption rate on the copper surface. Also, the current density values are different at different locations on the wafer due to variations in the potential and concentration distribution on the wafer.

One possible way of avoiding these uncertainties in the boundary condition is placing top boundary \( \Gamma_i \) on the plane where there is no concentration or potential variation. Since variations in potential and concentration distributions mostly occur within the diffusion boundary layer, changes in potential and concentration could be minimized by choosing a domain size \( \delta \) larger than the diffusion layer thickness. The
increased domain size would, however, defy all assumptions in Sections 5.1.1 and 5.1.2, since as the size increases, the relative importance of each term increases. The larger domain size further gives rise to convective mass transport between features, and thus the whole feature field on the wafer will have to be modeled simultaneously with a drastic increase in computational effort. Therefore, the domain size $\delta$ should be kept well below the mass transport boundary layer thickness, so that numerical methods remain tractable and the convective coupling between features is still insignificant. In the latter case, the unknown interface values must be determined by matching the consumption at the feature scale and the transport from the reactor scale through a multiscale simulation. For now, it will be assumed that interfacial concentration and current values are known. Chapter 6 will combine reactor and feature scale in self-consistent multiscale simulations.

\[ -k \nabla \Phi \cdot \vec{n} = i_{\text{a,\kappa}} \]

\[ C_M = C_M^\delta \]

\[ C_A = C_A^\delta \]

\[ \Gamma_1 \]

\[ \nabla \Phi \cdot \vec{n} = 0 \]

\[ \nabla C \cdot \vec{n} = 0 \]

\[ \delta \text{(domain size)} \]

\[ \Gamma_2 \]

\[ \Gamma_3 \]

\[ -k \nabla \Phi \cdot \vec{n} = i \]

\[ D_i \nabla \dot{C}_i \cdot \vec{n} = r_i \]

Figure 5-1. Computational domain and boundary condition of feature scale simulation
Symmetric boundary condition or no flux boundary condition is applied for both species and potential at the boundary between neighboring features \( \Gamma_2 \), typically. This approach is valid only if neighboring features have identical shape. Otherwise, the net flux between features should be included. This issue will be discussed in detail later in Section 5.4.

On the copper surface \( \Gamma_3 \), the boundary conditions are more complex. The boundary condition for species transport is given as follows.

\[
-D \frac{\partial c}{\partial y} \bigg|_{y=0} = \sum_{j=1}^{j-\text{total}} V_j r_j ,
\]

where \( i \) is the species number and \( j \) is the reaction number. The surface reactions used in this work are given in Table 3-6. This set of reactions includes the reduction of cupric ion, and the adsorption, the desorption and the inclusion reaction of additives.

The potential distribution on the copper surface is the related to the surface reaction rates as follows:

\[
i = \frac{nFD}{1-t} \frac{\partial c}{\partial y} \bigg|_{y=0} = -k_w \frac{\partial \Phi}{\partial y} \bigg|_{y=0} = \frac{nF}{1-t} \sum_{j=1}^{j-\text{total}} V_j r_j
\]

The transfer number \( t \) is neglected for high concentration of supporting electrolyte.

In the excess supporting electrolyte, the concentration overpotential can be approximated by the following equation:

\[
\eta_c = -\frac{RT}{ZF} \left[ \ln \left( \frac{c_w}{c_0} \right) \right]
\]

The surface overpotential is derived from a balance between the potential in liquid phase, the concentration overpotential, and the potential in the solid phase:
\[
\Phi_s + \eta_s + \eta_c = \Phi_t
\]

(5.17)

When the seed layer thickness is small, there is a significant potential drop within the seed layer, and therefore the seed layer potential \( \Phi_s \) in Equation (5.17) varies spatially (cf. Chapter 4). This variation can only be captured by linking the feature scale with the reactor scale simulation.

5.1.4 Effects of geometry change

With plating of copper, the geometry of the feature changes over time. The coordinate of the evolving interface surface \( \tilde{x} \) can be tracked by using the following equation:

\[
\frac{d\tilde{x}}{dt} = \frac{\text{Flux}_{Cu}}{\rho_{Cu}} = \frac{i}{nF\rho_{Cu}} \cdot \tilde{n},
\]

(5.18)

where \( \tilde{n} \) is the normal vector from the surface. Under usual operating conditions, it takes \( \sim 100 \) seconds to fill a 0.1 \( \mu \)m trench. This time is significantly larger than the diffusion time scale in the feature, \( O(10^{-4} \text{ sec}) \). Consequently, species transport may be considered as a steady state.

The feature shape evolution could go through drastic morphology change, such as when the opening of a feature pinches off. Therefore, the simulation method must be able to handle transient changes in feature topology efficiently and robustly.

5.2 Description of Numerical Method

The governing equations that describe the feature scale reaction and transport processes are

- the species transport equation (5.11) and its boundary conditions (5.14),
• the potential distribution equation (5.13) and its boundary conditions, (5.15), (5.16) and (5.17), and
• the geometry change (5.18).

For simplicity, simulations of field variables and geometry changes are solved in a sequential manner. In the following section, numerical methods are described for obtaining concentration and potential distributions as well as geometry update. Particular emphasis is placed on the robustness and the efficiency of each numerical method.

5.2.1 Transport equations: boundary element method

Recall that concentration and potential distributions satisfy the Laplace equation. The boundary element method (BEM) is an accurate and efficient tool for solving the Laplace equation.

For illustration purpose, let’s consider the formulation of the BEM with a single field variable $T$ that satisfies the Laplace equation in the computational domain shown in Figure 5-2.

![Figure 5-2. Computational domain of boundary element method. $\Omega$ is internal domain and $\Gamma$ is the boundary points.](image)
\[ \nabla^2 T = 0 \quad (5.19) \]

We introduce the weighting function \( G_i \), which satisfies the following equation.

\[ \nabla^2 G_i + \delta(x-x_i) \cdot \delta(y-y_i) = 0, \quad (5.20) \]

where \( \delta \) is the Dirac delta function, \((x, y)\) are the observation points and \((x_i, y_i)\) are node points. The solution of Equation (5.20) is known as an analytic function \(^4\). Starting from the weak form of the Laplace equation (5.19) with \( G_i \) as a weighting function, the derivative operator originally applied to the field variable \( T \) is transferred to the weighting function \( G_i \) with additional surface integration term by applying Green’s theorem twice

\[ \int_{\Omega} G_i \nabla^2 T d\Omega = \int_{\Omega} T \nabla^2 G_i d\Omega + \int_{\Gamma} \left[ G_i \frac{\partial T}{\partial n} - T \frac{\partial G_i}{\partial n} \right] d\Gamma = 0, \quad (5.21) \]

where \( i \) is the node number. On each boundary, either \( \partial T / \partial n \) or \( T \) is specified from the boundary condition. Alternatively, a nonlinear relationship between \( \partial T / \partial n \) and \( T \) can be applied. By selecting the node point \((x_i, y_i)\) along the boundary \( \Gamma \), systems of nonlinear equations, which are with function of only boundary values of \( \partial T / \partial n \) or \( T \), are obtained.

If the node point \((x_i, y_i)\) is chosen within the computational domain \( \Omega \), the internal values can be expressed as a function of the boundary values. Thus, finding only boundary values \( \partial T / \partial n \) or \( T \) is our interest.

The numerical solution of Equation (5.21) is constructed as follows. The variable \( T \) is discretized by using interpolation functions

\[ T(\eta) = \sum_{i=m}^{N_element} \sum_{k=1}^{Norder} T^{i,k} \psi_k(\eta) \quad (5.22) \]
\[ \frac{\partial T}{\partial n}(\eta) = \sum_{i=1}^{\text{Nelement}} \sum_{k=1}^{\text{Norder}} \frac{\partial T^{i,k}}{\partial n} \Psi_k(\eta) \]  

(5.23)

The quadratic interpolation functions take the form:

\[ \Psi_1(\eta) = 2\eta^2 - 3\eta + 1, \]  

(5.24)

\[ \Psi_2(\eta) = 4\eta - 4\eta^2, \text{ and} \]  

(5.25)

\[ \Psi_3(\eta) = 2\eta^2 - \eta, \]  

(5.26)

where \(0 \leq \eta \leq 1\). Substituting Equations (5.22) and (5.23) into Equation (5.21) produces the following result:

\[ -d_i T_i + \int_G \left[ \sum_{i=1}^{\text{Nelement}} \sum_{k=1}^{\text{Norder}} \frac{\partial T^{i,k}}{\partial n} \Psi_k(\eta) \right] \frac{\partial G_i}{\partial n} \ d\Gamma = 0 \]  

(5.27)

Introducing the Gaussian quadrature as an approximation to the line integrals gives:

\[ -d_i T_i + \sum_{g=1}^{N_g} G_i \left[ \sum_{i=1}^{\text{Nelement}} \sum_{k=1}^{\text{Norder}} \frac{\partial T^{i,k}}{\partial n} \Psi_k(\eta_g) \right] w_g - \sum_{g=1}^{N_g} \left[ \sum_{i=1}^{\text{Nelement}} \sum_{k=1}^{\text{Norder}} T^{i,k} \Psi_k(\eta_g) \right] \frac{\partial G_i}{\partial n} w_g = 0 \]  

(5.28)

When boundary conditions are applied, Equation (5.28) constitutes a complete set of equations for determining values of \( \partial T / \partial n \) and \( T \) at each node.

Special attention needs to be paid to situations where boundary conditions change form (See Figure 5.3). For example, when the boundary condition changes from Dirichlet to Neumann, the system becomes over-specified at the interface (corner), since both \( \partial T / \partial n \) and \( T \) are specified (see Figure 5.3 (a)). In order to avoid this over-specification, Brebbia described one resolution shown in Figure 5-3 (b). In this method, two closely spaced nodes are placed near the corner and different boundary conditions are assigned to the two nodes\(^5\). The error decreases as the distance between the two nodes is reduced. A more elegant approach described by Ramachandran is shown in
Figure 5-3 (c). The flux at the corner is decomposed into two parts, $\frac{\partial T}{\partial n}$ and $\frac{\partial T}{\partial n}$, so that $\frac{\partial T}{\partial n}$ and $T$ are assigned, but $\frac{\partial T}{\partial n}$ is not. In this way, only one unknown is available at each node.

The BEM described above applies to both concentration and potential distributions. The coupled systems of nonlinear equations are solved with the Newton-Rhapson method. The solution of the BEM equations produces the current density that can be converted into the growth rate of the copper film.

![Diagram](image)

Figure 5-3. Boundary condition change. (a) Corner point over specified (b) Closely spaced points avoid overspecification (c) Decomposition of flux avoids overspecification

5.1.2 Geometry update: level set method

The growth velocity calculated from Equation (5.18) is used to update the coordinate of growing film with the level set method. The level set method is faster than conventional string methods, and allows identical implementation for deposition or
etching. Moreover, the topological merging is naturally obtained without additional algorithmic logic.

The procedure for coupling of the BEM and level set methods is summarized in Figure 5-4. The original feature profile initializes the level set function by first calculating the distance between an arbitrary point \((x, y)\) and the nearest point to the feature profile and second assigning the sign to the distance based on the location of the arbitrary point \((x, y)\) (e.g. if \((x, y)\) is in liquid phase, the positive sign to the distance; otherwise, the negative sign is assigned to the distance). It also creates a mesh for BEM by adding boundary points. The solution of the BEM produces the growth rate of copper film, which is used to update the level set values. Then, the new zero level set, the feature

![Diagram](image)

**Figure 5-4.** Combination of level set and boundary element methods for feature scale simulation
profile in the next time step, is extracted with contouring algorithm. This procedure is repeated until the required feature profile is obtained.

Other than the aforementioned natural coupling between the BEM and the level set method, another advantage of the suggested numerical method manifests itself if the finite element method (FEM) is considered as a substitute for the BEM. With FEM, the constantly growing feature wall requires the adjustment of mesh nodal coordinates with either adaptive meshing\(^9\) or remeshing schemes\(^{10,11}\). Adaptive meshing schemes are computationally inexpensive, but they could create an ill-conditioned mesh, if the opening of the feature nearly pinches off as illustrated in Figure 5-5. A remeshing scheme could resolve this problem with a very refined mesh near the pinchoff point. However, the remeshing increases the number of variables significantly. Furthermore, the remeshing step needs to be repeated until the desired numerical accuracy is obtained for a given geometry. In conclusion, avoiding an internal mesh by application of BEM eliminates problems associated with remeshing.

![Figure 5-5. Adjusting mesh scheme in finite element method. When opening is pinches off, the numerical quality of mesh is poor.](image)

5.2.3 Parallelization: domain decomposition with artificial boundaries

Despite the advantages of BEM-level set combination described above, large memory and long convergence times are needed to handle large scale features, primarily because
of the dense data structure of the BEM. The level set method is hardly a numerical bottleneck - at least in two-dimensional problems. One way of reducing memory requirements and convergence time is by using a parallel iterative solver with a domain decomposition scheme similar to that used to simulate three-dimensional reactor flows\textsuperscript{12}. As in the reactor case, domain decomposition and preconditioning need to be addressed.

![Diagram](image)

Figure 5-6. (a) Original geometry and (b) decomposed domain with artificial boundary

The domain decomposition scheme is illustrated in Figure 5-6. Domain decomposition introduces artificial boundaries within the computational domain and allows the application of BEM in each separate domain.

\[
\int_{\Omega_1} \nabla^2 G_r \psi d\Omega + \int_{\Gamma_1} \left[ G_r \frac{\partial T}{\partial n} - T \frac{\partial G_r}{\partial n} \right] d\Gamma = 0
\]  

(5.29)

\[
\int_{\Omega_2} \nabla^2 G_r \psi d\Omega + \int_{\Gamma_2} \left[ G_r \frac{\partial T}{\partial n} - T \frac{\partial G_r}{\partial n} \right] d\Gamma = 0
\]  

(5.30)
From Figure 5-6, one variable $T_i$ in domain $\Omega_1$ originally depends on the entire variables on the boundary $\Gamma$. After domain decomposition is applied, however, it depends only upon variables $\Gamma_1 + \Gamma_+$, which reduces the number of off-diagonal terms in Jacobian matrix. The artificial boundary introduces the additional variables on $\Gamma_+$ and $\Gamma_-$, which must be matched between the two adjacent domains. Consistency can be obtained by the following procedure$^{13-15}$:

- assume boundary condition at the artificial boundary
- solve each domain independently with the boundary condition
- update boundary condition with calculated flux either with direct substitution or with relaxation factor

Nevertheless, such a successive iteration procedure does not guarantee convergence, and it is likely to be slow if it converges at all. A much faster in convergence would be possible if the artificial boundary conditions are handled as explicit unknowns and the exact Jacobian and residual including all neighboring domains are used simultaneously. In doing so, the originally dense Jacobian matrix is turned into a structured block diagonal matrix by each region coupled with common boundary condition$^{5,16}$.

Iterative solvers have been used extensively in the boundary element method$^{17-19}$. Since the BEM formulation does not have any inherent singularities in the diagonal, though it is not completely diagonally dominant system, Krylov subspace methods (e.g. generalized minimum residual (GMRES) and conjugate gradient (CG)) with simple preconditioners (e.g. Jacobi) work well$^{20}$. In this work, residual and Jacobian matrix with near block structure are formulated in parallel. The GMRES method with a simple preconditioner (Jacobi) is then used to solve the resulting linear system. A
general-purpose parallel iterative subroutine from Argon National lab (PETSc)\textsuperscript{21-23} is used to perform the actual computations.

5.3 Experiments

The operating conditions for the performed feature scale experiments (at Novellus Inc., San Jose, CA) are summarized in Table 5-1. The operating conditions are similar to those for the reactor scale experiments in Chapter 4. The main differences are the thicker seed layer and the applied current that is changed as a step function rather than a constant value.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation Speed (RPM)</td>
<td>100 RPM</td>
</tr>
<tr>
<td>Flow Rate from center jet</td>
<td>8 slm</td>
</tr>
<tr>
<td>H\textsubscript{2}SO\textsubscript{4} Concentration</td>
<td>1.78 M</td>
</tr>
<tr>
<td>CuSO\textsubscript{4} Concentration</td>
<td>0.11 M</td>
</tr>
<tr>
<td>Cl\textsuperscript{-} Concentration</td>
<td>50 PPM</td>
</tr>
<tr>
<td>PPG 1200 Concentration</td>
<td>10 PPM</td>
</tr>
<tr>
<td>SPS Concentration</td>
<td>6 PPM</td>
</tr>
<tr>
<td>Applied current</td>
<td>1 A /11 sec + 3 A</td>
</tr>
<tr>
<td>Seed layer thickness</td>
<td>150 nm</td>
</tr>
<tr>
<td>Wafer size</td>
<td>191.5 mm diameter</td>
</tr>
</tbody>
</table>

The secondary electron microscopy (SEM) pictures of the feature profiles were taken at the times specified in Table 5-2. The number of Coulombs was converted into the equivalent time by using the total current and the size of the wafer.

<table>
<thead>
<tr>
<th>Thickness ((\mu)m)</th>
<th>Coulomb</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.08</td>
<td>62.36</td>
<td>28.12</td>
</tr>
<tr>
<td>0.18</td>
<td>140.32</td>
<td>54.11</td>
</tr>
<tr>
<td>0.24</td>
<td>187.09</td>
<td>69.70</td>
</tr>
<tr>
<td>0.48</td>
<td>374.18</td>
<td>132.06</td>
</tr>
</tbody>
</table>
5.4 Results and Discussion

5.4.1 Comparison between experimental data and simulation results

The experimental data were obtained using the operating condition in Section 5.3. SEM pictures of three different feature widths, 0.15 μm, 0.3 μm, and 0.6 μm after depositing 0.08 μm of copper are shown in Figure 5-6. The feature with 0.15 μm width is almost filled, since the total growth thickness exceeds the width of features. The feature with 0.3 μm width is almost filled especially in the bottom part where the width of the tapered region is thinner. The feature with 0.60 μm width shows higher film thickness at the bottom of the trench.

Using this initial feature profile of each feature, a feature scale simulation was performed and the results compared to the experimental data in Figure 5-8. The computational domain size was set to 4 μm. At the top of the computational boundary, the bulk concentration and the average current density were applied as boundary conditions. In these operating conditions, the simulation results show almost conformal growth. For 0.15 and 0.3 μm trenches, the predictions are close to experimental results, but the 0.60 μm trench predictions deviates from experimental data. The pronounced growth at the bottom of the trench is not obtained in the simulation. Identifying the cause of this deviation will require more study of atomistic growth phenomena beyond the scope of this thesis.
Figure 5-7. Initial geometry of feature scale simulation indicated by dots.
Figure 5-8. Comparison between experimental and simulated feature profiles
5.4.2 Parallel computing strategies

To improve the convergence speed of the feature scale simulation, a parallel version of the feature scale simulator with artificial boundary was developed. The computational domain with one artificial boundary is illustrated in Figure 5-9. The thickness of the domain is 0.205 μm and the width of domain is 1.8 μm. The physical parameters and the operating conditions are the same with Table 3-2 and 3-3. The simulation was performed

![Figure 5-9](image_url)

Figure 5-9. Computational domain for long strip with height of 0.205 μm and width of 1.8 μm

![Figure 5-10](image_url)

Figure 5-10. Wall clock time of one Newton step with different number of domains (solid line) and parallelization efficiency with different number of domains (dashed line).
with SGI Origin 2000 machine with 195 MHz R10000 and 100 MBs ethernet connection. The number of processors for the parallel computing was set to the number of domains. GMRES was used with a restart parameter of 1000, and the absolute error tolerance and relative error tolerance was fixed to the machine precision.

Figure 5-10 shows the wall clock time and the parallel efficiency. The parallel efficiency is defined as:

$$\text{parallel efficiency} = \frac{\text{wall clock time}}{\text{wall clock time for 1 domain/number of domains}}$$  \hspace{1cm} (5.31)

The wall clock time is measured for the first Newton iteration. The wall clock time decreases as the number of domains increases. The parallel efficiency stays near 1, indicating that the inter-processor communication is small.

![Graph showing number of non-zeroes in Jacobian matrix and number of variables for different number of domains](image)

Figure 5-11. Number of non-zeroes in the Jacobian matrix (solid line) and number of unknowns for different number of domains (dashed line).
Figure 5-11 shows the number of total unknowns and the number of non-zeroes in the Jacobian matrix. As the number of processors, or equivalently number of domain increases, the number of unknown increases linearly, since new variables are introduced along the artificial boundary as the number of domain increases. However, the increase is not significant compared to the original number of unknowns as long as the width is much larger than thickness of the computational domain. As the number of domains increases, the number of non-zeroes in the Jacobian matrix decreases. This behavior can be understood from the nonzero-pattern of Jacobian matrix in Figure 5-12. The matrix structures are nearly block diagonal, and the off-diagonal term indicates the additional variables along the interface. Only if the number of these interface variables is

![Figure 5-12. Non zero pattern of Jacobian matrix with three artificial boundaries.](image-url)
kept small, the size of Jacobian decreases in proportion to the number of domain.

The parallel version of the feature scale tool can also be used to simulate large-scale feature scale phenomena. Figure 5-13 shows simulation of four identical trenches modeled with four artificial domains. Since the initial trenches have identical shape, the evolved feature profiles are also identical. A group of asymmetric trenches can be modeled in the same way. Figure 5-14 shows simulation of three trenches with different widths. In this case, the domain is decomposed into four sub-domains and the rigorous result is shown in the figure on the right. The figure on the left is obtained by neglecting the diffusive flux between artificial domains. The film at the right domain is thicker, since the flux cannot penetrate to the left domains. As illustrated above, the parallel version of the feature scale code does more than speeding up the convergence.

As pointed out in Section 5.1.3, the symmetry boundary condition along interface $\Gamma_2$ cannot be used unless the neighboring features have identical shape as in Figure 5-13. In general, as shown in Figure 5-14, the diffusive mass flux at the interface must be calculated by simulating neighboring features simultaneously.

Figure 5-13. The feature scale simulation results with four trenches in four domains. Domain boundaries lie at the half way point, which is symmetric plane.
5.5 Conclusion

A feature scale model has been developed based upon governing equations composed of a general surface chemistry model, species transport, potential distribution, and geometry change. The efficiency and the robustness of the numerical methods were emphasized in the selection of numerical methods for modeling this system. The combination of level set and boundary element methods formed complementary tools, and had the advantage of not requiring remeshing. Parallelization with domain decomposition was applied to the feature scale simulation tool to further improve its efficiency. This domain decomposition technique could also be used to account for the diffusive mass flux between asymmetric neighboring features. The uncertainty that still remains at the top boundary condition and at the seed potential need to be eliminated by linking feature scale code with reactor scale code. The derived feature scale model could provide an efficient building block for multiscale simulation tool as a predictive tool.
### List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>time</td>
</tr>
<tr>
<td>$k$</td>
<td>conductivity of electroplating bath, S/cm</td>
</tr>
<tr>
<td>$n$</td>
<td>number of charge transferred in the reaction</td>
</tr>
<tr>
<td>$t_c$</td>
<td>characteristic time to fill a feature</td>
</tr>
<tr>
<td>$u_i$</td>
<td>mobility of species i</td>
</tr>
<tr>
<td>$F$</td>
<td>Faraday constant, 96485 C/mol</td>
</tr>
<tr>
<td>$Pe$</td>
<td>Peclet number</td>
</tr>
<tr>
<td>$R$</td>
<td>gas constant, 8.31441 J/mol/K</td>
</tr>
<tr>
<td>$T$</td>
<td>arbitrary field variable</td>
</tr>
<tr>
<td>$C_i$</td>
<td>molar concentration of species i</td>
</tr>
<tr>
<td>$C_M$</td>
<td>metal ion concentration, M</td>
</tr>
<tr>
<td>$C_{M}^{m}$</td>
<td>bulk concentration of metal ion, M</td>
</tr>
<tr>
<td>$D_i$</td>
<td>diffusivity of species i, cm$^2$/sec</td>
</tr>
<tr>
<td>$G_i$</td>
<td>fundamental solution to Laplace equation</td>
</tr>
<tr>
<td>$L_c$</td>
<td>characteristic length of feature scale</td>
</tr>
<tr>
<td>$R_i$</td>
<td>homogeneous reaction rate of species i, mol/cm$^3$/sec</td>
</tr>
<tr>
<td>$V_c$</td>
<td>characteristic velocity</td>
</tr>
<tr>
<td>$Z_i$</td>
<td>number of charges of species i</td>
</tr>
</tbody>
</table>

### Greek Letters

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>Dirac delta function</td>
</tr>
<tr>
<td>$\rho$</td>
<td>density</td>
</tr>
</tbody>
</table>
\( \eta \)  
non-dimensional coordinate along the boundary

\( \eta_s \)  
surface overpotential

\( \eta_c \)  
concentration overpotential

\( \Omega \)  
computational domain

\( \Gamma \)  
boundary of computational domain

\( \Phi_L \)  
potential in the liquid phase

\( \Phi_S \)  
potential in the seed layer

\( \Psi \)  
quadratic interpolation function

5.6 References


6. Multiscale Simulation of Copper Electrochemical Deposition Processes

6.1 Introduction

Copper electrochemical deposition processes are governed by physical phenomena that occur over a wide range of length and time scales. Since these phenomena are inherently coupled, a predictive simulation tool needs to be capable of describing the multi-length and -time scale physics. Multiscale simulation strategies have been used to model low pressure chemical vapor deposition (LPCVD)\textsuperscript{1-4} and high pressure chemical vapor deposition processes (HPCVD)\textsuperscript{5,6}. For LPCVD processes, the physical processes in the feature and the reactor scales require different models. The multiscale simulation enables different physical descriptions of model domain and a consistent linking between different models. For HPCVD processes, the physical process in the feature and the reactor scales can be described with the same set of continuum models. Thus, the multiscale simulation is solely motivated by numerical considerations, specifically avoiding excessive meshing near patterned regions. Irrespective of such motivational difference in the two processes, a common multiscale methodology is used.

The essence of the multiscale simulation is as follows. The system is decomposed into a number of sub-domains based on its length scale. Each sub-domain is modeled by the most suitable governing equations and numerical methods. Then, each length scale is coupled by sharing common flux and concentration across domain boundaries to preserve
consistency. Consequently, successful multiscale simulation requires models for individual scales, as well as consistent and numerically efficient linking schemes.

In the case of the copper electrochemical deposition processes, most modeling studies for patterned wafers have been focused on either individual feature scale simulations with a constant concentration boundary condition along the interface between the reactor and the feature scales\textsuperscript{7-11}, or a so-called pattern scale simulation where the feature effect is averaged out and represented as an increased surface area due to patterns\textsuperscript{12}. The individual feature scale simulation cannot provide information on the effect of the operating condition on the feature profile evolution, not to mention differences in feature profiles across the wafer. The pattern scale simulation does not provide any information on the behavior of individual features and the effect of the film growth. A multiscale simulation that includes the reactor and the feature scale has been reported\textsuperscript{13,14}, but no detailed description of rigorous multiscale simulation have been reported to date.

The required components of multiscale simulation of copper electrochemical deposition processes are a surface kinetics, a reactor scale, a feature scale models and a consistent linking scheme. These required components have been developed in earlier chapters of this thesis. The detailed surface chemistry model was developed in Chapter 3. The reactor scale simulation tool was developed in Chapter 4, with detailed mass transport and potential distribution equations\textsuperscript{15} using Galerkin finite element methods\textsuperscript{16-21}. For the feature scale simulation developed in Chapter 5, the governing equation could be simplified to the Laplace equation because of the small feature scale. The boundary element method was used to solve the Laplace equation\textsuperscript{22} and the level set method was
employed to follow the topology change$^{23-25}$. The combination of the boundary element method and the level set method reduced the computational cost by eliminating the remeshing step. The only missing component is the linking scheme.

In this chapter, a rigorous multiscale simulation tool for the copper electrochemical deposition processes will be developed. First, quantitative criteria for decomposing the overall computational domain into different sub-domains will be given. Resulting multiscale strategy will be combined with a computational approach predicting feature evolution at different locations on the wafer. Finally, the multiscale simulation

Figure 6-1. Different sub-domains of the copper electrodeposition process, (a) reactor scale, (b) pattern scale and (c) feature scale.
tool will be used to analyze various factors influencing feature evolution at different locations on the wafer.

6.2 Methodology

6.2.1 Problem description: domain decomposition

The computational domain for the copper electrochemical deposition processes can be decomposed into following three length scales\textsuperscript{26}.

- Reactor scale (work piece scale)
- Pattern scale
- Feature scale.

Figure 6-1 shows a schematic of these scales in the electrochemical deposition processes. The simulation of the copper electrochemical deposition processes needs to include all the information of different sub-domains. However, the sub-domains involve widely different length scales. For instance, the reactor scale is order of 10 cm (0.1m), the pattern scale is of order of 1 mm (0.1-10\textsuperscript{2}m) and the feature scale is order of 0.1 \textmu m (0.1-10\textsuperscript{4}m). These widely differing length scales make it very difficult to model all the sub-domains in a single computational domain, even though the same set of governing equations describes the transport phenomena of all the scales. Furthermore, the film growth within the feature scale would require remeshing of the whole computational domain at every time step. Since the geometry change is localized, it is numerically advantageous to handle the feature scale separately through a multiscale simulation.

As a first step of applying the multiscale simulation strategy to the copper electrochemical deposition processes, quantitative criteria for domain decomposition
needs to be described. Then, a consistent linking scheme of different sub-domains needs to be developed. Finally, the completed multiscale simulation tool needs to be combined with level set based simulation of feature evolution.

6.2.2 Decomposition into reactor and pattern scales

The pattern scale shown in Figure 6-1 includes feature regions of the wafer. As a result, the reactor scale does not involve any moving boundary. As long as this criterion is fulfilled, the boundary between reactor and pattern scales can be placed on top of the patterns with arbitrarily determined height. However, if the numerical efficiency is to be maximized, a few more constraints need to be added to the determination of the domain height. The height of the pattern scale determines the relative importance of convective and transient mass transport to pure diffusion in the pattern scale. In order to simplify the governing equation, it needs to be investigated how to minimize the convective and the transient mass transport effects in the pattern scale. Also, the resolution of the mesh along the interface between the reactor and the pattern scale is determined by the domain height, since the height determines the variation of the flux on top of the pattern\(^3,5\). Therefore, the height of the pattern scale should be established in a quantitative manner considering all these effects. Three factors to determine the height of the pattern scale are given in the following paragraphs.

First, it is advantageous to place the boundary within the diffusion layer.

\[
\delta < \delta_{\text{diffusion boundary layer}}
\]  

(6.1)

where \(\delta\) is the domain height of the pattern scale. The mass transport boundary layer thickness can be estimated either from the reactor scale simulation or existing
correlations\textsuperscript{27,28}. If the domain size is larger than the mass transport boundary layer, the convection plays a role within the computational domain of the pattern scale, which increases coupling between individual features in radial directions. This feature coupling would be a significant problem, since individual features can then not be modeled independently. Additionally, convection must be included in the governing equation of the feature scale transport, and the computational advantages of the Laplace equation are lost.

Second, the domain size should be small enough so that the diffusion within the feature scale can be much faster than the film growth. The characteristic times for the diffusion and the film growth are as follows:

\begin{equation}
    t_{\text{diffusion}} = \frac{\delta^2}{D} \tag{6.2}
\end{equation}

\begin{equation}
    t_{\text{film growth}} = \frac{\delta}{\text{(growth rate)}} \tag{6.3}
\end{equation}

If the diffusion is faster than the film growth, or

\begin{equation}
    t_{\text{diffusion}} \ll t_{\text{film growth}}, \text{equivalently } \delta \ll \frac{D}{\text{(growth rate)}}, \tag{6.4}
\end{equation}

the mass transport in the feature scale can be considered as pseudo-steady state process.

Third, the height of the pattern scale need to be large enough so that the flux at the top of the domain does not have any fluctuations as illustrated in Figure 6-2. If there is a high variation of the flux, the reactor scale mesh will have to be fine enough to resolve it. This in turn would require the reactor scale mesh to resolve sub-micron variations on feature scale. However, if the domain height is sufficiently large, the flux at the top boundary varies smoothly on the pattern scale due to the lateral diffusion, as illustrated in
Figure 6-2 and can then be represented by the reactor scale mesh. Based on the analytic solution of mass transport in an infinite strip (See Section 6.5 Appendix), the height of the domain required for reducing the variation of flux is as follows

$$\delta \geq \frac{W}{\pi},$$

(6.5)

where $W$ is the half width of the feature domain.

The height of the feature scale is determined by Equations (6.1), (6.3) and (6.5). Combining all of these equation results in the following relations.

$$\frac{W}{\pi} \leq \delta \ll \min\left(\frac{D}{(\text{growth rate})}, \delta_{\text{diffusion boundary layer}}\right)$$

(6.6)

The optimum size is determined by the geometry of the feature scale and the boundary layer thickness, the growth rate. The growth rate and the diffusion boundary layer

![Flux diagram](image)

Flux at the top of the boundary as $\delta$ increases

![Domain size $\delta$](image)

Domain size $\delta$

Figure 6-2. The fluxes on top of feature scale as a function of domain size $\delta$. 
thickness changes for different operating conditions. Thus, it should be clearly checked whether this relationship is satisfied.

### 6.2.3 Decomposition of pattern scale and assembling flux

The pattern scale is composed of many features. Resolving all features in one computational domain is computationally expensive. If the concentration of the species at the top of the domain is constant and the constituent features have identical shapes, the whole pattern can be represented by a single feature scale simulation. However, the concentration is likely to vary along the top boundary of the pattern scale. In this work, the pattern scale is decomposed into different segments and then each segment is

---

**Figure 6-3.** Decomposition of patterns into different segments to resolve variation in the reactor scale.
represented by one feature scale simulation. In this way, the variation of the concentration at the top of the domain can be resolved and the pattern scale is described by small number of feature scale simulations. The decomposition of the pattern into different segments is illustrated in Figure 6-3. Since the variation of the concentration at the boundary is not known a priori, the number of the segments needs to be iteratively increased until there is no more change in the flux response.

Generating the flux map on the pattern scale from the combined effects of the individual feature scale simulations is the reverse process of the decomposition. The flux from each feature scale is duplicated to represent each segment. Then, the fluxes from each segment are patched together to form a flux map of the pattern scale.

Simply duplicating the flux from the individual feature to represent the segments is applicable only when the neighboring features are identical, in which case symmetry conditions can be imposed on the side boundaries. If the features are not identical, which is likely to happen at the edge of wafer, the net diffusive flux between the features needs to be included. The methods introduced in Chapter 5 as well as those developed by Nemirovskaya et al.\textsuperscript{5} can be used to account for the net diffusive flux between features.

![Diagram](Image.png)

Figure 6-4. Unknowns along the artificial boundaries, $\Gamma_R$ and $\Gamma_F$. 
In the first method, the neighboring features need to be modeled simultaneously. In the second, the neighboring features can be modeled independently. In either case, the number of features is so large that it is not practical to attempt to consider all the features. Physical insight must be used to judge the appropriate level of treatment in each case.

6.2.4 Linking reactor and pattern scales

Once the boundary between the reactor and the pattern scales is determined from Section 6.2.2 and the pattern scale is represented by a number of feature scale models, the next step is to specify its proper boundary conditions. Typically, flux is specified for reactor scale simulation and concentration for feature scale simulation. However, these values at the artificial boundary are not known a priori. Thus, a proper iteration scheme needs to be devised to find concentration and flux that satisfy both the reactor and the feature scale models.

The domain decomposition into the reactor and the pattern scales, and the newly introduced variables along the artificial boundaries are illustrated in Figure 6-4. Along the new boundaries, \( \Gamma_R \) and \( \Gamma_F \), the unknowns are as follows:

\[
\begin{align*}
c^R, \frac{\partial c^R}{\partial n} & \text{ along } \Gamma_R \\
c^F, \frac{\partial c^F}{\partial n} & \text{ along } \Gamma_F
\end{align*}
\]  

(6.7) (6.8)

These four unknowns need to be determined. Consistency between reactor and feature scale requires the following relationship:

\[
c^R = c^F
\]  

(6.9)

\[
\frac{\partial c^R}{\partial n} = \frac{\partial c^F}{\partial n}
\]  

(6.10)
The models for the reactor and the feature scales provide two additional relationships. With the reactor scale model, the interfacial concentration can be calculated with the given interfacial flux,

\[ c^R = f\left(\frac{\partial c^R}{\partial n}\right), \tag{6.11} \]

where \( f \) implies the reactor scale model that determines boundary concentration \( c^R \). With a feature scale model, the interfacial flux can be obtained with the given interfacial concentration,

\[ \frac{\partial c^F}{\partial n} = g(c^F), \tag{6.12} \]

where \( g \) represents the feature scale model that determines boundary flux \( \frac{\partial c^F}{\partial n} \). Thus, all four constraints, Equations (6.9-12), are given to determine four unknowns in Equation (6.7) and (6.8). The simplest iteration method to determine all the unknowns is as follows. First, Equations (6.9) and (6.12) are combined to reduce one variable, resulting in the following equation:

\[ \frac{\partial c^F}{\partial n} = g(c^R) \tag{6.13} \]

Combining Equations (6.10) and (6.11) reduces another variable and results in the following equation,

\[ c^R = f\left(\frac{\partial c^F}{\partial n}\right). \tag{6.14} \]

Then, assuming the flux from the pattern scales is zero, Equation (6.14) is solved. The resulting interfacial concentration \( c^R \) is used to solve Equation (6.13). The flux herein, in turn, is employed to calculate the interfacial concentration using Equation (6.14). This iteration is repeated until Equations (6.9) and (6.10) are satisfied. This linking method is the simplest one that assures the consistency between reactor and pattern scales. This method is typically useful if the functional form of the reactor and the pattern scales
cannot be accessed, e.g. for commercial codes. However, the numerical method converges linearly, if it converges at all, and diverges if the system is stiff.

Nemirovskaya et al. illustrated a more sophisticated method that drastically improves the convergence\(^5\). The methodology is described as follows. The boundary condition along the interface between the reactor and the pattern scales is

\[
-D \nabla c \cdot \vec{n}\bigg|_{\text{interface}} = \text{flux}\bigg|_{\text{interface}},
\]

where \text{flux} is the flux at the interface between reactor and pattern scales. These values are calculated from the pattern scale using the method described in Section 6.2.3. The weak form of the governing equation for the species transport is

\[
\int_{\Omega} \frac{\partial c_k}{\partial t} \phi' dV = \int_{\partial \Omega} \text{flux} \cdot \phi' dS - \int_{\Omega} N_k \cdot \nabla \phi' dV + \int_{\Omega} R_k \phi' dV. \tag{6.16}
\]

This equation is based on Equation (4.18), where \text{flux} is substituted for \(N_k\) in the boundary integration. In order to obtain optimum convergence, the Jacobian of Equation (6.16) needs to be provided. It is given as

\[
\int_{\Omega} \frac{\partial^2 c_k}{\partial t \partial c_{j,l}} \phi' dV = -\int_{\partial \Omega} \frac{\partial \text{flux}}{\partial c_{j,l}} \phi' dS - \int_{\Omega} \frac{\partial N_k}{\partial c_{j,l}} \cdot \nabla \phi' dV + \int_{\Omega} \frac{\partial R_k}{\partial c_{j,l}} \phi' dV, \tag{6.17}
\]

where \(c_{j,l}\) is the concentration of species \(j\) in node \(l\). In Equation (6.17), the sensitivity of \text{flux} with respect to the interfacial concentration is the key factor determining the convergence of the system. Numerical perturbation is used to calculate the sensitivity of \text{flux} with respect to \(c_{j,l}\). In principle, the perturbation must be performed with respect to both indices \(j\) and \(l\), which would make the number of perturbations equal to the number of the variables per node multiplied by the number of the interfacial nodes. In this work, however, the concentration of \(j\) for all nodes is perturbed simultaneously. Thus, the total number of perturbations is equal to the number of variables in the interfacial nodes. The
formulated problem is solved by the Newton-Raphson method. Even though the simplified Jacobian does not produce exact quadratic convergence, the problem converges super-linearly even for stiff systems. Since this linking scheme illustrated above requires a large number of feature scale simulations, numerical efficiency is particularly crucial. This is not the case for the reactor scale simulations, especially in 2D, which converge very fast. Therefore, it is particularly important to have an efficient feature scale simulations. The developed boundary element –level set strategy (Chapter 5) provides such a method.

The reactor and the pattern scales involve vastly different mesh refinement. These differences along the reactor and pattern interface are illustrated in Figure 6-5. The proper interpolation rule needs to be defined in order to exchange information between two scales. The information passed from the reactor to the feature scales includes the interfacial concentration, the liquid phase potential, and the seed layer potential. Since the reactor scale nodes are much sparser than the pattern scale, linearly interpolated values, based on the location of features underneath the reactor scale, are used. When passing information from the pattern scale to the reactor scale, the flux values at the equivalent location in the pattern scale is transferred. However, this approach does not guarantee overall mass balance between two domains. To insure mass conservation, the flux underneath reactor scale nodes is integrated and the average values

![Diagram](image_url)

Figure 6-5. Difference in the reactor and feature nodes.
are assigned at the reactor scale node.

6.2.5 Putting all together

So far the multiscale simulation approach emphasizes the consistent mass transport between different sub-domains. In order to produce practical information, the multiscale simulation must be combined with predictions of the feature profile evolution based on level set methods\textsuperscript{23-25}. The overall numerical scheme is illustrated in Figure 6-6. In this approach, the consistent mass transport computations need to be satisfied at every time step and the geometry of the feature updated. The procedure is repeated until the required time elapses.

6.3 Results and Discussions

The feature profile evolution is strongly influenced by operating conditions and shows

![Figure 6-6. Schematic of overall numerical method](image-url)
different filling properties at different locations on the wafer. Identifying main causes of the nonuniform film growth is crucial in optimizing the process. Previous modeling attempts for individual feature scale simulations\textsuperscript{7-11} or averaged feature pattern effect embedded in the reactor scale\textsuperscript{12} do not provide sufficient information; the first approach is not capable of studying the effect of operating conditions on the feature scale evolution and the latter cannot predict the behavior of individual features and the effect of film growth on the feature scale. In the following sections, a number of examples will be given to illustrate different factors that influence the feature profile evolution at different locations on the wafer. The result will also show how the multiscale simulation tool developed in this chapter overcomes the limitations of the previous modeling approaches.

6.3.1 Simulation condition

The computational mesh for the reactor scale is illustrated in Figure 6-7. Different features can be placed on top of the wafer. The feature pattern shown in Figure 6-8 is characterized by three parameters, W, H, and \( W_1 \). Typical simulation conditions are summarized in Table 6-1. The simulation results given in the following example used this simulation conditions unless mentioned otherwise.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation Speed</td>
<td>0 RPM</td>
</tr>
<tr>
<td>Flow Rate from center jet</td>
<td>0.01 slm</td>
</tr>
<tr>
<td>( \text{H}_2\text{SO}_4 ) Concentration</td>
<td>1.78 M</td>
</tr>
<tr>
<td>( \text{CuSO}_4 ) Concentration</td>
<td>0.11 M</td>
</tr>
<tr>
<td>Cl Concentration</td>
<td>50 PPM</td>
</tr>
<tr>
<td>PPG 1200 Concentration</td>
<td>30 PPM</td>
</tr>
<tr>
<td>SPS Concentration</td>
<td>6 PPM</td>
</tr>
<tr>
<td>Applied voltage</td>
<td>0.01 V</td>
</tr>
<tr>
<td>Seed layer thickness</td>
<td>50 nm</td>
</tr>
<tr>
<td>Wafer size</td>
<td>191.5 mm diameter</td>
</tr>
</tbody>
</table>
Figure 6-7. Computational mesh for multiscale simulation

Figure 6-8. Feature geometry and three parameters, $W$, $H$, $W_1$ that characterize the geometry
6.3.2 Effect of transient mass transport on the multiscale simulation

The liquid phase transport is characterized by its slow mass transport. For the operation with blank wafers, the transient mass transport effect is important either at the beginning of the operation or when there are external perturbations, such as pulse plating described in Chapter 4. For the operation with patterned wafers, however, the film growth in the feature changes the effective surface area over the time, which acts as a perturbation to the system. Therefore, a fully transient simulation needs to be performed to model the electrochemical deposition process with the patterned wafer, even in the absence of external perturbations.

![Diagram](image)

**Figure 6-9.** Feature geometry used for transient mass transport example.

For the multiscale simulation with fully transient mass transport effect, the computational mesh shown in Figure 6-7 is used. A feature shown in Figure 6-9 are chosen and placed on top of the wafer. The initial surface area of this feature is 2.56 times larger surface area than the blank wafer. The steady state solution with the initial feature geometry is used as an initial condition for the transient problem.
Figure 6-10 shows the concentration profile of cupric ions near the wafer and the feature profiles at the corresponding times. The concentration of cupric ions at the position \( \times \) is shown in Figure 6-11. Concentration of Cu\(^{2+}\) increases as the feature grows, since the effective surface area is reduced as the feature grows, which in turn reduces the consumption of the cupric ions.

Figure 6-10. Concentration profile of Cu\(^{2+}\) and feature profiles at (a) beginning of simulation (b) after 2000 seconds elapsed. Concentration profiles within the boundary layer is magnified for better visibility.
6.3.3 Effect of reactor scale variations on feature growth

In Chapter 4, the terminal effect and the mass transport were illustrated as two major causes of the uniformity in the reactor scale. In this section, it will be illustrated how
these effects also affect the feature profile evolution.

*Influence of terminal effect on feature evolution*

Both current and potential distributions for a blank wafer with 50 nm seed layer are shown in Figure 4-7(a). The large potential drop within the seed layer leads to a current density difference between the edge and the center of the wafer. Patterns are placed in two different locations on the wafer shown in Figure 6-12 in order to maximize the difference. Identical features are placed in these locations, and the features are characterized with following parameters.

\[
W = 0.15 \, \mu m
\]

\[
H = 0.195 \, \mu m
\]

\[
W_1 = 0.075 \, \mu m
\]

The initial surface area of the feature is 2.3 times larger than that of the blank wafer. The feature profiles that are grown for 200 seconds at the different locations are

![Diagram](image.png)

Figure 6-13. Different feature profiles at the center (left) and at the edge (right) due to the uneven boundary layer thickness
shown in Figure 6-13. The feature at the center of the wafer grows slower than the feature at the edge, which is consistent with the blank wafer results.

Since the nonuniform growth rate in the feature is primarily caused by the terminal effect in this example, one could attempt to mitigate the nonuniformity by removing the terminal effect by using a thicker seed layer. As shown in Chapter 4, however, the seed layer thickness needs to be at least 1 \( \mu \)m in order to remove the terminal effect. Considering the size of the feature (currently 0.13 \( \mu \)m) is much smaller than 1 \( \mu \)m, this is not a valid approach. Thus, other methods need to be investigated with the multiscale simulation tool.

*Effect of mass transport boundary layer*

Mass transport boundary layer thickness uniformity is determined by the reactor geometry and the flow condition. Figure 4-11 in Chapter 4 showed the distribution of the growth rate across the wafer in the presence of uneven mass transport boundary layer thickness. We expect that the uneven mass transport boundary layer thickness will affect

![Diagram](image)

Figure 6-14. Pattern placement for mass transport effect, \( a=2 \text{cm}, b=8 \text{cm} \)
the feature profile evolution at different locations on the wafer. In order to illustrate the point, the wafer is divided into two parts and the same features to the previous example are placed in each region as shown in Figure 6-14.

The transient multiscale simulation was performed to observe the feature growth. The feature profiles at different times at two different locations are shown in Figure 6-15. The feature corresponding to the smaller boundary layer thickness grows much faster than one under the thicker boundary layer thickness.

The cause of the uneven feature profile evolution is nonuniform boundary layer thickness, primarily initiated by the small inlet. Thus, redesigning the reactor is the best way of eliminating nonuniform growth. However, in case when the reactor geometry cannot be changed, the modification of the operating condition can be used as a remedy. As shown in Figure 4-11(a), the growth rate can be reduced to obtain uniform growth rate. As an alternative, pulsed current in Section 4.5.3 can be used to eliminate the mass transport effect without changing growth rate significantly

![Diagram](image)

Figure 6-15. Different feature profiles at the center (left) and at the edge (right) due to the uneven boundary layer thickness
6.3.4 Effect of feature density

Example in Section 6.3.2 shows that the uniformity in the reactor scale affects the feature profile evolution. The other factor that influences the feature profile is the density of the feature. In order to demonstrate the point, the simulation condition needs to be chosen so that there are no variations in the reactor scale. For this purpose, the seed layer thickness of 1 μm was used and the concentration profile of cupric ions and additives in the reactor scale is assumed to be uniform. Then, two different features are chosen, each of which has different effective surface area as summarized in the Table 6-2. Feature type 1 is placed on the location a in Figure 6-16 and feature type 2 on the location b.

![Diagram showing features a and b with dimensions](image)

Figure 6-16. Pattern placement for mass transport effect, a=5cm, b=5cm

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>H</th>
<th>W₁</th>
<th>Effective area density</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.45 μm</td>
<td>0.15 μm</td>
<td>0.075 μm</td>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
<td>0.15 μm</td>
<td>0.3 μm</td>
<td>0.075 μm</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 6-2. Geometry of features used in feature density effect simulation
The feature profile evolution with the applied voltage of 0.3 V is shown in Figure 6-17. The pattern 1 at the location a in Figure 6-16 grows much faster than the pattern 2 at the location b. This example shows that the nonuniform growth can be initiated solely by the difference in the pattern density in the absence of the reactor scale variation. Optimization of such a process could be much harder than the case with uniform pattern density, since removing the wafer scale nonuniformity does not guarantee uniform film growth. More sophisticated dynamic optimization will need to be used with the multiscale simulation tool developed in this chapter.

![Figure 6-17. Different feature profiles at the center (left) and at the edge (right) due to the different boundary pattern density](image-url)
6.4 Conclusion

In this chapter, a rigorous multiscale simulation tool for copper electrochemical deposition was presented. Since the transport phenomena of different length scales could be described with the same set of governing equations, multiscale simulation of the copper electrochemical deposition process was motivated by the numerical efficiency. Thus, domain decomposition and linking of sub-domains were performed to maximize numerical efficiency. The resulting multiscale simulation tool was capable of simulating consistent mass transport within its computational domain including reactor, pattern and feature scales. In order to provide the feature profile evolution on the different locations on the wafer, the multiscale simulation tool was coupled with the feature profile evolution strategy based on the level set method.

Simulation tool was used to see the effect of the transient mass transport. The simulation tool was used also to investigate the growth rate uniformity in the feature scale. First, two major effects influencing the reactor scale uniformity, the terminal effect and the mass transport effect was investigated to see how large impact they have on the feature evolution. Second, in the absence of the reactor scale uniformity, how the feature density affects the feature profile evolution at different applied voltage. This example showed that the optimizing the reactor scale does not guarentee the uniform film growth of the patterns, - the only way is to use the multiscale simulation approach.

List of Symbols

\[ t \] time
\( f \)  
nonlinear relationship for reactor scale

\( \Phi \)  
flux between reactor and pattern scales

\( g \)  
nonlinear relationship for feature scale

\( c_k \)  
concentration of species \( k \)

\( c_{j,l} \)  
concentration of species \( j \) at node \( l \)

\( c^R \)  
interfacial concentration at the reactor scale

\( c^F \)  
interfacial concentration at the feature scale

\( t_{\text{film growth}} \)  
characteristic time to fill a feature

\( t_{\text{diffusion}} \)  
characteristic diffusion time scale

\( D \)  
diffusivity, \( \text{cm}^2/\text{sec} \)

\( W \)  
half width of the feature scale geometry

\( L_c \)  
characteristic length of feature scale

\( N_k \)  
flux of species \( k \), \( \text{mol/cm}^2/\text{sec} \)

\( R_k \)  
homogeneous reaction rate of species \( k \), \( \text{mol/cm}^3/\text{sec} \)

**Greek Letters**

\( \delta \)  
domain size of the feature scale

\( \Phi \)  
Lagrangian biquadratic interpolation function

\( \Omega \)  
computational domain

\( \Gamma \)  
boundary of computational domain
6.5 Appendix

In order to model the concentration change, the following geometry is considered in Figure 6-18. The solution to the model system is,

\[ C = C_0 \]
\[ y \to \infty \]
\[ \nabla C \cdot \hat{n} = 0 \]
\[ \nabla^2 C = 0 \]
\[ \nabla C \cdot \hat{n} = 0 \]
\[ x = 0 \]
\[ x = W \]
\[ y = 0 \]
\[ C = C_0 \left( 1 - \frac{x}{W} \right) \]

Figure 6-18. Geometry for example

\[ C(x, y) = C_0 + C_0 \sum_{n=1}^{\infty} B_n e^{-\frac{ny}{W}} \cos\left( \frac{n\pi x}{W} \right) \]

\[ B_n = 2 \cdot \frac{1 - (-1)^n}{n^2 \pi^2} \]

In order for the flux in y component does not have any variation in x direction, the exponent of the leading term should vanish. Thus,

\[ e^{-\frac{ny}{W}} \to 0, \text{ or } \frac{ny}{W} > 1 \]

Therefore, the minimum distance of the y coordinate is

\[ y > \frac{W}{\pi} \]
6.6 References


Chapter

7. Conclusion and Recommendations for Future Work

7.1 Conclusion

The goal of this thesis was developing predictive simulation tools for ionized physical vapor deposition (IPVD) and the electrochemical deposition (ECD) processes. In order to achieve the goal, we have developed the models based on fundamental transport phenomena and obtained surface kinetic parameters that represent wide operating conditions.

Transport phenomena in the feature scale of IPVD process was described by the ballistic transport, since the mean free path of the process is much larger than the size of the feature. The geometry change of the feature was modeled with the level set method. The important parameter to the IPVD feature scale simulation tool was the surface reaction rates. In particular, the interaction of the high kinetic energy ion fluxes with the already deposited film played a key role, since they could not only adsorb or reflect, but also etch the metal film and diffuse long distance. The analysis of the surface reactions at different incident energies and angles was performed with molecular dynamics (MD) simulations with embedded atom potential.

The copper electrochemical deposition process involves a number of proprietary additives to obtain desired film quality. In this study, three representative additives, poly ethylene glycol (PEG), bis-(sodium sulfoprophyll) (SPS), and chloride were used. The detailed kinetic mechanism was developed based on the competitive adsorption/desorption model. This kinetic mechanism was shown to capture the
synergistic effect of different additives on the copper deposition. Careful RDE experiments were performed to obtain the experimental data used to extract the kinetic parameters of the model using non-linear regression techniques.

The governing equations of the reactor scale model included momentum transport, transient mass transport, potential distribution, and detailed surface kinetic mechanism. The coupled nonlinear equation was discretized with Galerkin finite element methods and the time integration was performed with a differential algebraic equation (DAE) solver. In order to validate the model, the simulation result was compared with the experimental film thickness uniformity on the blank wafer and shown to be in agreement. The validated simulation tool was used to investigate the various effects on the film thickness, e.g., terminal effect and mass transport effects. The nonuniformity caused by the terminal effect vanished as the film grows, but the nonuniformity created by the mass transport effects persisted over time. As one way of removing the nonuniformity by the mass transport effect, the pulse plating was modeled and its result showed 50% improvements in nonuniformity compared to the constant current case.

In the feature scale simulation, the same set of governing equations to the reactor scale still could be applied, but the small computational size simplified the governing equations into the Laplace equation. Based on the simplified governing equation, a computationally efficient feature scale model was developed using a combination of level-set and boundary element methods. The level set method was used to model the temporal change of the growing film front and the boundary element method was used to simulate liquid phase transport. The use of these two methods led to two main synergistic advantages, the first being that of a straightforward interface in which
the zero level set, which is part of the level set solution describing the growing surface front, could be directly used as a computational mesh for the boundary element method. In turn, the solution for the boundary element method, in the form of mass flux or local current density, could be put into the level set method as a front advancing velocity. As a consequence of this coupling approach, an often very computationally expensive explicit remeshing step was not needed for the liquid phase transport equation. The simulation result was compared with the experimental data obtained from the commercial ECD cell and shown to be in agreement for features with 0.15 and 0.3 \(\mu\)m widths. In order to improve the computational efficiency, the domain decomposition and the parallel iterative solver were used in solving large feature scale problems. The domain decomposition made the structure of the Jacobian matrix into a block diagonal matrix. The parallel version of the feature scale tool was used to include net diffusive flux between asymmetric features.

Based on the surface kinetics, the reactor scale and the feature scale models, a multiscale simulation tool was developed to model the ECD cell with a patterned wafer. The domain was decomposed into reactor, pattern and feature scales based on its length scale. The feature scale simulations were used to represent the pattern scale. The pattern scale was linked with the reactor scale in a computationally efficient manner. This multiscale simulation scheme made a consistent mass transport model in all the length scales. This model was combined with the moving boundary algorithm based on the level set method to predict the feature profile evolution. A developed simulation tool was used to model the terminal effect and the mass transport effect on the feature profile evolution on the different locations of the wafer for uniform pattern density. In the
presence of the terminal effect, the feature at the edge of the wafer grew faster than one at the center as the current distribution on the blank wafer suggested. In the presence of uneven mass transport boundary layer, the feature grew faster under thinner boundary layer. Finally, the effect of uneven pattern density on the feature profile evolution was studied. The feature with small effective surface area grew faster than the one with higher effective surface area. Optimizing uneven growth rate caused by the uneven pattern density would need to require sophisticated dynamic optimization.

7.2 Recommendations for Future Work

The reactor scale model in this work can be improved in many aspects. First, the Navier-Stokes equation in the reactor scale model applies only for the laminar flow. In the commercial ECD reactor, the flow regime can fall in the turbulent regime with the marginal rotation speed of wafer (~ 100 RPM) because of the high density of the water and the larger wafer size (300 mm). Thus, a different turbulent model can be applied to the reactor scale model if the model needs to be used in wider operating condition. Second, the reactor geometry was assumed to be axisymmetric in this thesis. However, in many commercial reactors, the ECD cell chamber may not be perfectly axisymmetric, e.g. due to the location of the outlets. Thus, a three-dimensional model needs to be developed in order to model more complex ECD cells.

The effect of different additives in this study was performed by means of a semi-empirical approach. In order to fully understand the role of the additives, their chemical interaction with the cupric ion on the copper-water interface should be understood. For example, an ab-initio study can be performed on the interaction of SPS
molecule with the cupric ions on the copper surface to demonstrate how the presence of SPS can affect the deposition of the cupric ions. In this work, SPS was assumed to be a steric hindrance, but more direct validation will be more convincing.

Detailed understanding of the interaction of additives and cupric ions on the copper surface can eventually allow the detailed surface morphology evolution study with kinetic Monte Carlo (KMC) simulation. This simulation can provide insight into the effect of different additives on the nucleation, and thus the microstructure, of the copper deposit. The nucleation behavior can give insight into the grain growth in the copper deposit. This grain growth affects the electromigration properties of the copper deposit, which determines the lifetime of the semiconductor devices. Finally, these findings can be linked to the current multiscale simulation framework, thereby allowing the prediction of the device property as a function of the operating condition.