View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by ScholarBank@NUS

MICROLITHOGRAPHY: CONTROL OF TEMPERATURE AND RESIST THICKNESS

LEE LAY LAY

(B. ENG. (HONS.), NUS)

NATIONAL UNIVERSITY OF SINGAPORE

2003

MICROLITHOGRAPHY: CONTROL OF TEMPERATURE AND RESIST THICKNESS

LEE LAY LAY

(B. ENG. (HONS.), NUS)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2003

Summary

Lithography is one of the key technology drivers for semiconductor manufacturing. As the feature size shrinks, maintaining adequate process latitude for the lithographic processes becomes challenging. In this thesis, in-situ process monitoring and model-based control techniques are used to control the bakeplate temperature and resist thickness. These are two important process variables that can affect the final critical dimension.

A predictive controller is designed to perform a pre-determined heating sequence prior to the arrival of the photomask to eliminate the load disturbance induced by the placement of the cold photomask onto the bakeplate. An order of magnitude improvement in the temperature error is achieved. Using an array of in-situ thickness sensors and advanced control algorithms, a real-time thickness control strategy is implemented to control the resist thickness during softbake. By manipulating the temperature distribution of the bakeplate in real-time, the resist thickness non-uniformity caused by prior coating process is reduced. An average of 10 times improvement in the resist thickness uniformity is achieved across individual wafers and from wafer-to-wafer.

ACKNOWLEDGEMENTS

I would like to express my appreciation to all those who have guided me during my postgraduate study at the National University of Singapore, National Science and Technology Board and Stanford University. Firstly, I wish to express my utmost gratitude to my supervisors, A/P Ho Weng Khuen and A/P Loh Ai Poh, for their wisdom, patience and unfailing guidance throughout the course of my research project. I have indeed benefited tremendously from the many discussions I had with them, without whose help this project and thesis would have been impossible. I would also like to thank Dr. Tan Woei Wan for reading my thesis and giving me many helpful comments.

Secondly, I would like to thank Professor Thomas Kailath for the opportunity to conduct my research at Stanford University and his invaluable advice. I would like to express special thanks to Dr. Charles Schaper for his guidance, help and many useful suggestions, particularly on the thickness control strategy, while I was at Stanford University. I am also grateful to Arthur Tay, Kalid El-Awady, Kenneth and Young Peng for their friendship and invaluable technical assistance to me.

Finally, I would also like to thank my family for their love, encouragement, understanding and support.

Contents

Summary i

Acknowledgement ii

Contents iii

List of Tables vi

List of Figures vii

1 Introduction **1**

1.1 Challenges and Trends in the Semiconductor Industry 1

1.2 Overview of Semiconductor Manufacturing Processes 2

1.3 Overview of Process Control Methods 4

1.3.1 Process Monitoring 5

1.3.2 Statistical Process Control 6

1.3.3 Real-time Feedback Control 7

1.3.4 Run-to-run Control 9

1.4 Temperature Effects in Lithographic Processes 11

1.5 Scope of Thesis 15

1.5.1 Temperature Control for Photomask Fabrication 15

1.5.2 Real-time Thickness Control 17

1.6 Thesis Organization 19

2 Constraint Feedforward Control for Photomask Thermal Processing **20**

2.1 Introduction 20

2.2 Experimental Setup for Photomask Thermal Processing 23

2.3 Optimal Feedforward/Feedback Control Strategy 25

2.4 Implementation of the Constrained Feedforward Controller 31

2.5 Conclusion 335

3 Real-time Predictive Control of Resist Film Thickness Uniformity 37

3.1 Introduction 37

3.2 Experimental Setup 41

3.3 Resist Thickness Estimation 46

3.4 Generalized Predictive Control 54

3.4.1 Identification 55

3.4.2 Control Algorithm 59

3.4.3 Choice of Design Parameters 64

3.4.3.1 Reference Trajectory 64

3.4.3.2 Temperature Limits 65

3.4.3.3 Prediction Horizon and Control Weighting 66

3.5 Experimental Results 69

3.5.1 Conventional Softbake 71

3.5.2 Real-time Thickness Control with GPC algorithm 73

3.6 Summary 76

4 Implementation of Real-time Thickness Control Using Sliding Mode Control 77

4.1 Motivations 77

4.2 Control Structure 79

4.2.1 Identification 84

4.2.2 Sliding Mode Control Algorithm 87

4.2.3 Implementation 92

4.3 Comparison with the GPC Algorithm 94

4.4 Experimental Results 96

4.5 Conclusion 99

5 Conclusion 102

5.1 Review of Objectives and Summary of Results 102

5.2 Scope for Future Developments 104

References 106

Appendix A: Overview of Generalized Predictive Control algorithm 115

Appendix B: Simulation results of Sliding Mode Controller 123

Appendix C: Author’s Publications 131

List of Tables

Table 1.1:	Temperature sensitivity of lithographic thermal processing steps.	14
Table 3.1:	Parameters of a_1 and b_0 for all the three thickness models.	58
Table 3.2:	Parameters of a_1 and b_0 for all the three thermal models.	58
Table 3.3:	Static gain, time constant and dead time of all the process models.	59
Table 4.1:	Summary of the parameters.	97

List of Figures

Figure 1.1:	Major fabrication steps in MOS process flow.	3
Figure 1.2:	Lithography process.	12
Figure 1.3:	Variations of CD with resist thickness.	17
Figure 2.1:	Load disturbance to the bakeplate temperature due to placement of the photomask.	22
Figure 2.2:	Photograph of the bakeplate.	24
Figure 2.3:	Top view and cross section of the bakeplate.	25
Figure 2.4:	Feedforward/feedback control strategy for photomask thermal processing.	26
Figure 2.5:	Result of the identification experiment to obtain the disturbance model.	32
Figure 2.6:	Result of the identification experiment to obtain the plant model.	33
Figure 2.7:	Optimal feedforward control signal.	34
Figure 2.8:	Comparison between runs with and without feedforward control.	36
Figure 3.1:	An example of a swing curve.	38
Figure 3.2:	Schematics of the experimental setup used to control resist thickness in real-time.	42
Figure 3.3:	Cross-section of the experimental setup.	43
Figure 3.4:	Photograph of the experimental setup.	45
Figure 3.5:	(a) Thin film optical model, and (b) variation of the reflectance signal with wavelengths for a particular resist thickness.	47

Figure 3.6: Comparison between the measured and calculated reflectance signal. 51

Figure 3.7: Comparison of the three thickness estimation algorithms. 53

Figure 3.8: Identification experiments. Plots of (a) resist thickness, (b) temperature, (c) change in resist thickness where $y = y''_m - y'_m$, (d) change in temperature where $T = T''_m - T'_m$, and (e) change in heater input, u, with respect to time. 57

Figure 3.9: Simulation results for different values of lambda, λ. 67

Figure 3.10: Simulation results for two sites. 70

Figure 3.11: Conventional softbake with bakeplate maintained uniformly at 90 °C: (a) resist thickness, (b) temperature, (c) heater input, and (d) resist thickness non-uniformity profile of the three sites monitored. Sites 1-3 are represented by the solid, dashed-dotted, and dashed lines, respectively. 72

Figure 3.12: GPC Control with temperature constraints: (a) resist thickness, (b) temperature, (c) heater input, and (d) thickness non-uniformity profile when three sites on the wafer are monitored. Sites 1-3 are represented by the solid, dashed-dotted, and dashed lines, respectively The reference thickness trajectory is given by the dotted line. 75

Figure 3.13: Summary of the experimental runs. 76

Figure 4.1: Control structure of the thickness control strategy using sliding mode control. 80

Figure 4.2: Cascaded control loop structure. 82

Figure 4.3: Identification experiments. Plots of (a) resist thickness, (b) temperature, (c) decrease in resist thickness where $y_1 = y''_m - y'_m$, (d) increase in bakeplate temperature where $T = T''_m - T'_m$, (e) change in heater input, u, with respect to time. The solid lines in (c) and (d) show experimental values and the dashed lines show the calculated responses of the identified models. 86

Figure 4.4:	Sliding mode control for Run #1. Plots of (a) resist thickness, (b) temperature, (c) control signal, and (d) heater input with respect to time when resist thickness at Site A and B are monitored. Site A and B are represented by the solid and dashed lines respectively. The reference thickness trajectory is given by the dotted line in (a).	100
Figure 4.5:	Sliding mode control for Run#2. Plots of (a) resist thickness, (b) temperature, (c) control signal, and (d) heater input with respect to time when resist thickness at Site A and B are monitored. Site A and B are represented by the solid and dashed lines respectively. The reference thickness trajectory is given by the dotted line in (a).	101

Chapter 1

Introduction

1.1 Challenges and Trends in the Semiconductor Industry

The phenomenal growth of the semiconductor industry has been fueled by the increase in productivity year after year, of which a large part of the productivity gains has been the result of lithography improvements in terms of smaller feature sizes, tighter overlays and high-density chips. There is a strong economic incentive for the semiconductor industry to continue to shrink the minimum feature sizes exponentially. Not only can more transistors be packed in a VLSI chip; a decrease in the minimum feature size also results in a significant increase in the switching speed. This in turn translates into a higher selling price for the faster device. In semiconductor manufacturing, the ability to control the critical dimension (CD) and its uniformity is important. Usually, CD control involves controlling the dimension of the smallest feature size in a device such as the gate linewidth. Traditionally, gate CD control is most critical in lithography as any variation in the gate linewidth has a significant impact on the device speed and performance.

According to the International Technology Roadmap for Semiconductor (ITRS) in 1999, gate CD control of 10 nanometer (nm) is required at 100 nm technology node by year 2005 [1]. The ITRS roadmap presents the industry-wide consensus on the R&D efforts needed to meet the challenges of semiconductor manufacturing at a specific minimum

linewidth (technology node). By year 2014, it is estimated that gate CD control of 4 nm is required at the 35 nm technology node. In addition to tighter process specifications, the industry is also moving towards 300-mm wafer high volume production. This further escalates the demands on the manufacturing control for all the VLSI fabrication processes, as the control requirements have to be spread over a larger wafer area. As the feature size approaches sub-100 nm, maintaining adequate and affordable lithographic process latitude that is necessary for post-etch CD control becomes an increasingly challenging and difficult task. Advances in process control methodologies are necessary to achieve such a tight CD control [1].

This Chapter is organized as follows. Section 1.2 gives an overview of some of the important semiconductor processes involved in the VLSI chip fabrication. The trend of the process control methodologies in semiconductor manufacturing is given in Section 1.3. Section 1.4 discusses the effects of temperature on the lithographic processes. The scope of the thesis is given in Section 1.5 and Section 1.6 gives the thesis organization.

1.2 Overview of Semiconductor Manufacturing Processes

Semiconductor manufacturing is a complex and highly integrated industry involving several intermediate processes. Figure 1.1 shows the process flow for fabricating a single MOS transistor. For an advanced integrated circuit fabrication, there may be as many as 350 or more processing steps involved [2]. Some of the critical processes directly influencing the

CD are lithography, etching and deposition processes [2, 3]. A brief description of these processes is given below.

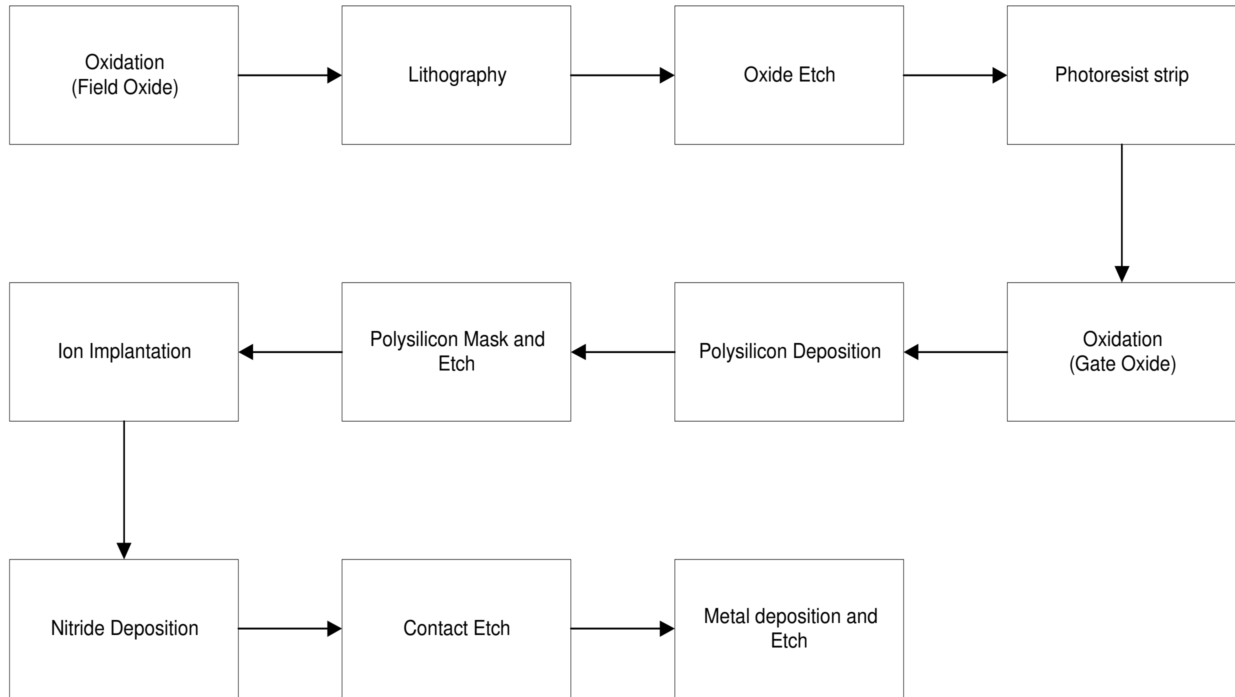


Figure 1.1: Major fabrication steps in MOS process flow.

Lithography is the process that transfers the patterns of the desired circuit from the photomask onto a photosensitive resist film that is coated on top of the wafer substrate. Performed immediately after lithography is the etching process. Etching selectively removes unwanted material from the wafer in areas that are not protected by the resist patterns to form permanent features on the wafer substrate. Some of the commonly used etching techniques are plasma etching and Reactive Ion Etching (RIE). Plasma etching relies on the chemical reaction of the feed gas to remove unwanted material while RIE uses both chemical reaction

of the feed gas and physical reaction due to ion bombardment to remove unwanted material. Deposition is another important process that deposits a thin layer of material on the wafer. Chemical Vapor Deposition (CVD) process is one of the primary methods used to deposit thin and highly uniform films of nitrides, metals and oxide films on the wafer. In CVD process, the reactant gases are introduced into the chamber. The gases undergo chemical reaction with the heated wafer surface to form a thin film of material. Different CVD processes are currently used, including Atmospheric Pressure CVD, Low Pressure CVD and Plasma Enhanced CVD.

Process drifts in any of these processes will affect the final CD. In the next section, an overview of the process control strategies to deal with process drifts in semiconductor manufacturing will be discussed.

1.3 Overview of Process Control Methods

The objective of any process control methods is to reduce variation of the process variables so as to ensure that the wafer-state parameters such as film thickness, CD, etc. is kept within tight process specifications. In recent years, there has been a surge in research interest on process monitoring and control of key semiconductor manufacturing processes due to the need to fabricate VLSI chips with higher feature densities [1]. With total allowed wafer CD variation in the range of only a few nanometers, every nanometer of allowed variation must be carefully assigned in the error budget of each intermediate processing step. Allowing excessive variation in one particular process would impose an unachievable target

for other processes. Higher manufacturing costs may result in the form of scrap, reduced yield and rework.

The major elements of any process control systems are process monitoring and advanced process control techniques. In this section, the process monitoring and control techniques in semiconductor manufacturing are reviewed.

1.3.1 Process Monitoring

The purpose of process monitoring is to provide critical information concerning the process variables that may be used for process diagnostic and feedback control. Traditionally, process monitoring was done in an off-line fashion, based on measurements that were available after several batches of wafers have been processed. In recent years, there has been an increase in the application of in-situ sensors in semiconductor manufacturing processes due to advancement in the in-situ sensor technology. Some examples of the in-situ sensors that have been developed for monitoring the semiconductor manufacturing processes are: ellipsometer [4] and reflection interferometers [5, 6] that measure film thickness, scatterometer that measures CD and particles [7, 8] and photospectrometer that measures the photoactive compounds in the resist film [9].

State estimation algorithms such as Kalman filter and extended Kalman filter have also been implemented to estimate critical internal process variables from the available measurements [5, 10, 11]. Vincent *et al.* [5] uses extended Kalman filter to estimate the etch rate and Caroll *et al.* [11] estimates the development rate from the reflectance measurements

obtained by the reflection interferometers. Palmer *et al.* [10] uses Kalman filter to obtain both the resist thickness and photoactive compound concentration indirectly from the photospectrometer. To provide real-time approximation and prediction of the critical variables, sophisticated models such as neural networks have also been built off-line and implemented on-line for model-based monitoring of the semiconductor manufacturing processes [3, 12-14].

Another important component of an effective process control system is the process control technique itself. In the next section, some of the process control approaches in semiconductor manufacturing are reviewed.

1.3.2 Statistical Process Control

Process drifts are inevitable during fabrication of the integrated circuits. To keep the process within a tight specification, adjustment for unacceptable drift in the process variable is important. One of the most widely used process control methodologies in semiconductor manufacturing is Statistical Process Control [15]. For Statistical Process Control (SPC), critical process observables are monitored using ex-situ metrology to identify abnormal level of variations in materials, equipments, parameters or procedures. Any abnormal variations will be used to detect an "out-of-control" state. When an "out-of-control" situation is detected in the process, sources of the process drift are identified and the process is re-centered via engineering intervention or hardware cleanup to return it to the "in-control"

state. Implementation of SPC is relatively simple and has been successfully implemented for stepper overlay control, plasma enhanced CVD and plasma etching processes [16-18].

However, this approach can be costly and inefficient. This is because SPC detects the process drifts by post examination of processed wafers and adjusts the process for the next batch of wafers when an "out-of-control" situation arises. A large number of wafers can be potentially misprocessed before an "out-of-specification" condition is recognized and corrected. The complexity of the SPC methodology also increases with shrinking error tolerances as the number of inputs that have to be monitored and controlled increases [16]. Furthermore, in situation where process disturbance has a significant effect on the process variation, an open loop approach such as SPC may not be able to compensate for the process disturbance. Hence in recent years, there has been a shift in the control methodology from SPC to more advanced process control techniques, which is the subject of this thesis.

1.3.3 Real-time Feedback Control

Production loss can be significantly reduced by diagnosing the onset of the process drifts on-line and taking a more timely corrective action. One approach is to correlate the wafer-state parameters (e.g. film thickness, CD, etc.) to control variables (e.g. temperature, gas flows, exposure dose) for real-time feedback control. In real-time feedback control, critical processing parameters are measured or inferred in real-time using in-situ sensors. These in-situ measurements are compared against a reference during process control. Any deviation from the reference is compensated and reduced by the use of feedback control.

Since more rapid feedback action takes place when an “out-of-specification” condition occurs, real-time feedback control reduces process variations, wafer scrap and production cost.

Until recently, real-time feedback control technique is focused mainly on the equipments and is confined to simple single loop controllers [3, 19]. Increasingly, real-time feedback control is now being used to improve the fabrication processes [20-24, 27]. By implementing real-time feedback control, Rashap *et al.* [20] and Hankinson *et al.* [21] achieve a better control of the RIE process. Using in-situ process signals and an adaptive nonlinear controller to compute the etch time in real-time, Rietman *et al.* [22] achieves good wafer-to-wafer control of the plasma etch process. A model that predicts linewidth broadening caused by post-exposure delay (time delay between exposure and post exposure bake steps) and an automatic control system have also been employed to reduce CD variation induced by post-exposure delay [23]. Through real-time monitoring of the stepper overlay performance and predictive tuning of the stepper, Ku *et al.* reduces the downtime associated with manual tuning of the stepper [24]. In process control, a higher process capability index is desirable as it means that the percentage of process variables that fall outside the specification limits is lower. Real-time feedback control improves the process capability [25, 26]. Process capability of more than 50 % has been achieved by implementing feedback controller for CD and overlay [27].

Another common form of real-time process control is endpoint detection [28-31]. Endpoint detection uses the real-time process measurements to determine the termination of the process. The in-situ sensors continuously monitor the process parameters so that the process is terminated when the in-situ measurements reach a certain setpoint or decision

criteria (known as endpoint). Endpoint detection has been implemented for the develop, softbake, CVD and RIE processes with improved controllability. Morton *et al.* uses the in-situ ultrasonic sensor to monitor the resist thickness and its properties during the softbake and develop processes [28, 29]. These in-situ measurements are used to detect the endpoint of the softbake and develop processes. Baker *et al.* [30] uses an in-situ surface micro-machined sensor to monitor the film thickness during the RIE process and terminates the etch process when the target thickness is reached. Epitaxial film thickness measurements obtained using Emission Fourier Transform infrared spectroscopy has also been used to determine the endpoint of the CVD process [31]. Process monitoring during CVD also reduces the cost of other related processes such as pre-deposition wafer cleaning and post-deposition material characterization.

1.3.4 Run-to-run Control

Besides real-time feedback process control, run-to-run process control is another control methodology that has received a lot of attention recently. In run-to-run control, feedforward and feedback information between consecutive processes are used in combination with a process model to keep the wafer-state parameters close to their target values. The process model relates the equipment settings (e.g. deposition time) to the wafer-state qualities of interests (e.g. film thickness) and is updated to track the process drifts. Measurement data taken immediately after the processing step is fed back to adjust the

process recipe in the equipment for the next run. The measurements are also sent to the next equipment to adjust its recipe before any deviations from the target become worse.

Run-to-run controller monitors the process on a run-to-run basis and makes small adjustment to the equipment to keep the wafer-state parameters on target. This control approach is most applicable to processes where run-to-run performance is plagued by machine wear or equipment drifts. Some of the common implementation of run-to-run controller are based on exponentially weighted moving average schemes, adaptive algorithms, modified internal control, robust control via worst case framework artificial intelligence control, predictor-corrector control self-tuning controller and linear model predictive control [3, 32]. Run-to-run controllers can also be evaluated by considering the model uncertainty from a probabilistic approach [33].

Run-to-run control has been implemented successfully for processes such as etching [34], photolithography sequence [35], RIE [21], CVD [36] and metal sputtering deposition [37] with improved yield and throughput. Run-to-run control is often used together with real-time control system to control critical process variables. The run-to-run controller determines the equipment settings based on the wafer-state characteristics of the previous run. The real-time controller in the equipment adjusts the control variables to achieve the suggested equipment settings. For example, the real-time controller has been integrated with a run-to-run controller to achieve good etch control for the RIE process [21]. The run-to-run controller determines the plasma variable setpoint based on the wafer characteristics of the previous run. The real-time controller maintains the suggested plasma variables by manipulating the process inputs during the run. The inclusion of a real-time controller keeps

the process in the target range for small disturbances and reduces the amount of work needed for the run-to-run controller.

Given that lithography is the key technology driver in semiconductor industry, this thesis investigates the application of advanced process control methodologies to meet some of the challenges of advanced lithography; focusing particularly on temperature control. This is because lithography involves several bake processes where temperature control is important. In this thesis, in-situ process monitoring and model-based control techniques are used to control the bake process so as to achieve a temperature profile that is repeatable from run-to-run. This is extended to control the resist thickness in real-time during the bake process. A uniform resist thickness distribution across individual wafer and from wafer-to-wafer (run-to-run) is achieved by manipulating the temperature distribution of the bakeplate based on the in-situ resist thickness measurements. In the next section, the effects of temperature on the lithographic processes and the importance of temperature control are discussed.

1.4 Temperature Effects in the Lithographic Processes

Lithography is a manufacturing process that transfers two-dimensional microscopic patterns of the desired circuit from a photomask onto a photosensitive resist film that is coated on top of the wafer substrate. It is also a significant economic factor, representing over 35 % of the chip manufacturing cost. For an advanced integrated circuit, there can be as

many as 20 masking levels involving lithography [48]. Figure 1.2 shows the typical processing steps involved in lithography for a single masking level.

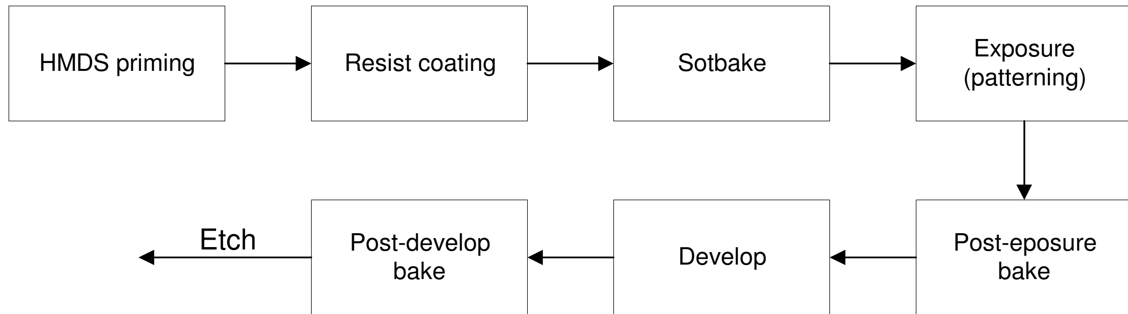


Figure 1.2: Lithography process.

The lithography sequence begins with priming the wafer substrate with hexamethyldisilazane (HMDS) to promote adhesion of the photosensitive resist to the substrate. After HMDS priming, a thin film of resist is coated onto the substrate using a spin coating technique. Liquid resist solution is dispensed onto the wafer in a spinner and spun at a very high speed. The spinning of the wafer causes the centrifugal force, which together with gravity, forces the resist to flow to the edge of the wafer. Surface tension induces the resist flow while viscous force resists the flow. During the spin coating process, solvent in the liquid resist evaporates, resulting in a reduction of the resist film thickness. This is followed by a softbake process, where the heat of the bakeplate removes the residual solvent in the resist film. Using an exposure tool, the resist-coated subtrate is exposed to ultra-violent (UV) radiation to project the desired patterns from the photomask onto the resist film.

A subsequent post-exposure bake reduces the standing waves in the resist film. In the case of the chemically amplified resist, it also activates critical chemical reaction. The resist is next exposed to the developer solution which selectively dissolves the exposed resist for positive-tone resist, and vice-versa for the negative-tone resist. Finally, a post-develop bake is performed to enhance the etch resistance of the resist patterns. After the lithographic process, the wafer undergoes other fabrication processes such as etching, deposition, oxidation, etc. Through a series of etching or deposition processes, the resist patterns formed during lithography are transformed into permanent features in the device.

In lithography, any drifts and variations in the process variables such as exposure doses, temperature, resist thickness, developer concentration, etc. will affect the final CD [2, 15, 48]. Among these process parameters, temperature is one of the most significant process parameters that can affect the CD. The effect of temperature on CD has been studied extensively. For every degree variation in the wafer temperature uniformity during the bake process, the CD can vary by as much as 20 nm [15]. A 9 % variation in CD per 1 °C variation in temperature has been reported for a Deep ultraviolet (DUV) resist [35]. As the width of the feature size continues to shrink, temperature uniformity specification becomes more stringent. Table 1.1 shows the temperature requirements for different thermal processing steps in lithography [36]. For some critical bake processes such as post-exposure bake (PEB), temperature uniformity as stringent as ± 0.1 °C is required. Therefore, temperature control is extremely critical to achieve good CD uniformity.

To achieve the tight temperature specifications required by advanced lithography, one approach is to improve the design of the thermal processing equipment. In Ramanan *et al.* [37], an exhaustive heat transfer analysis of the bake equipment is performed to gain insight

into the heat transport to the wafer during the baking process. This is then used to aid in the design of the bakeplate so that good wafer thermal uniformity during baking can be achieved. Controlled cooling after the bake process is also important. A chillplate has been introduced to improve the repeatability of the bake process [38]. Proximity cooling during photomask fabrication is used to improve the resist CD uniformity in Kushida *et al.* [39]. To enable more uniform resist bake for the photomasks and wafers, independently controlled, multiple-zones bake systems that integrate the baking and cooling process have been designed to achieve better temperature control during the entire bake cycle [40, 41]. Integration of the baking and cooling processes eliminates the use of a robotic arm to transfer wafers between the bakeplate and chillplate. This enables better temperature control throughout the bake cycle [40]. An iterative, self-optimizing algorithm is also developed to determine optimized bake recipes in Dress *et al.* [41].

Table 1.1: Temperature sensitivity of the thermal processing steps.

Thermal Step	**Purpose**	**Temperature Range**	**Precision Required**
HMDS bake	Promote Adhesion	$70-150^{o}C$	$\pm 5^{o}C$
ARC bake	Cure ARC	$90-180^{o}C$	$\pm 1-2^{o}C$
Softbake	Drive off solvent, densify resist, stabilize thickness	$90-140^{o}C$	$\pm 1^{o}C$
Post – exposure bake (PEB)	*i*-line resist: smooth standing waves	$90-180^{o}C$	$\pm 0.5-1^{o}C$
PEB	DUV resist: deblock exposed resist	$90-150^{o}C$	$\pm 0.12-0.5^{o}C$
Post-develop bake	Improve etch stability	$120-180^{o}C$	$\pm 1^{o}C$

1.5 Scope of the Thesis

In this thesis, the application of advanced control algorithm to meet the challenges of some aspects of advanced lithography is investigated. This thesis addresses two areas: 1) Temperature control during photomask fabrication and 2) Real-time thickness control during the bake process.

1.5.1 Temperature Control For Photomask Fabrication

Lithography involves a complex transfer of features from the photomask onto the wafer by the exposure tool. Every feature on the photomask has to be imaged onto the wafer within a predetermined tolerance. With rapid scaling of the feature size, lithographic imaging is being pushed into a regime of non-linear amplification of photomask errors. For example, a 10 nm CD error on the photomask may translate into a 14 nm resist errors on the wafer, after taking into account the reduction factor of the exposure tool. As the feature size shrinks, photomask CD error takes an increasing larger portion of the total wafer CD error budget. The photomask has been identified as the largest contributor to CD error for 0.25 μm DUV lithography [42]. More accurate mask CD control and tightening mask specifications are therefore required. According to the ITRS roadmap, mask CD control of 10 nm and 4 nm are required by year 2002 and 2014 respectively [1].

As in the thermal processing of the wafer, the ability to control temperature within a tight tolerance throughout the bake cycle is important during photomask fabrication. Baking

is performed by placing the photomask on the bakeplate at a specified temperature for a given time. When a photomask, originally at room temperature, is placed onto the bakeplate, the bakeplate experiences a sudden drop in temperature (load disturbance) due to heat transfer from the bakeplate to the cold photomask. In response to this load disturbance, the heating elements of the bakeplate increase the heater powers sharply. If the bakeplate controller is not well designed, there will be an overshoot in the bakeplate temperature before it can be returned to the desired temperature setpoint. This temperature overshoot can cause statistically measurable shifts in CD. Crisalle [43] reports a statistical increase in wafer CD ranging from 0.011 to 0.035 μm due to a temperature overshoot of 3 °C or more. The mass of the photomask is much larger than the wafer. The typical dimension of a wafer is 300 mm diameter and 1 mm thick while the dimension of a leading edge photomask is 6 inch by 6 inch and 0.25 inches thick. Due to the larger mass of the photomask, the amount of heat removed from the bakeplate by the cold photomask will be much greater, resulting in a larger load disturbance. Hence the ability to reject the load disturbance effectively during photomask fabrication is even more critical.

The author's contribution in this research work is as follows. A feedforward/feedback control strategy that is based on a linear programming method is implemented on a bakeplate to perform a pre-determined heating sequence prior to the arrival of the photomask. This minimizes the load disturbance induced by the placement of the cold photomask on the bakeplate. Compared to if only a feedback controller is implemented on the bakeplate, there is a substantial improvement in the integrated-square temperature error using this feedforward control strategy.

1.5.2 Real-time Thickness Control

Resist thickness uniformity is another significant lithographic process parameter that can directly affect the CD distribution across the wafer. The CD varies as a function of the resist thickness [15], as given in Figure 1.3. Hence, the resist thickness has to be well controlled to achieve good CD uniformity.

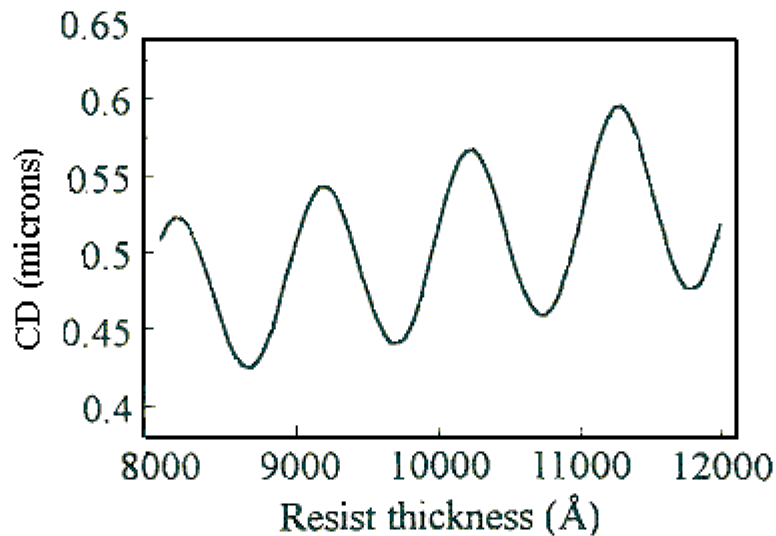


Figure 1.3: Variations of CD with resist thickness

In this thesis, the application of real-time thickness control during softbake to improve the resist thickness uniformity is investigated. Prior to the softbake process, the wafer is spin coated with resist. Typically, a non-uniform resist film is formed on top of the wafer at the end of the coating process. In the conventional approach where the bakeplate temperature is maintained at a constant temperature during softbake, the resist film thickness

non-uniformity remains at the end of the softbake process. In this work, a new technique is implemented to control the resist thickness in real-time during the softbake process. This reduces the resist thickness non-uniformity at the end of the bake process. Given the strong correlation of CD uniformity with resist thickness uniformity, reducing resist thickness non-uniformity will improve CD control.

The contributions on this work are as follow:

1. It has been demonstrated that if the prior coating process results in a non-uniform resist film, conventional softbake approach of maintaining uniform temperature distribution across the bakeplate does not reduce the resist film non-uniformity.
2. A new setup is made to implement a real-time thickness control strategy. Using advanced control algorithms and in-situ resist thickness measurements, resist thickness non-uniformity of less than 1 nm is achieved at the end of the softbake process. In addition to improvement in resist thickness uniformity across individual wafer, the softbake process is also made more repeatable. There is an improvement in resist thickness uniformity from wafer-to-wafer as the same resist thickness trajectory is defined for all wafers.

1.6 Thesis Organization

This thesis consists of 5 chapters and is organized in the following ways. The first chapter covers the introduction, which discusses some of the challenges faced by the semiconductor industry and how the application of control methodologies can meet some of these technical challenges. Chapter 2 describes an optimal feedforward control algorithm that improves the temperature uniformity of the bakeplate by rejecting the load disturbance caused by the placement of cold photomask onto the bakeplate. Chapters 3 and 4 discuss the implementation of a real-time thickness control strategy that compensates for the resist thickness non-uniformity caused by the prior coating process. Two control algorithms have been investigated. The first control algorithm is the Generalized Predictive Control algorithm, whereby constraints on the bake temperature can be readily implemented and is covered in Chapter 3. A simpler and faster implementation using Sliding Mode Control algorithm is investigated in Chapter 4. Chapter 5 summarizes the research work conducted and gives recommendations for future work.

Chapter 2

Constraint Feedforward Control for Photomask Thermal Processing

2.1 Introduction

With shrinking feature sizes, lithographic imaging tool is being pushed to its resolution limits where pattern transfer from the photomask to the wafer is non-linear. Hence, any CD error on the photomask will be non-linearly amplified onto the wafer. As a result, mask CD error contributes to an increasingly larger portion of the wafer CD error budget and tighter CD control has to be imposed during photomask fabrication [42]. To meet the tightening mask specifications, temperature control during thermal processing of the photomask is critical. There have been some researches on improving the mask CD uniformity through temperature control. For example, Kushida *et al.* [39] improved the resist pattern CD uniformity by using proximity cooling during photomask fabrication. Independently controlled, multiple zones bake systems have been designed to enable more precise and uniform resist bake for the photomasks and wafers [40, 41]. The baking and cooling processes are integrated to eliminate the use of a robotic arm to transfer wafers between the bakeplate and chillpate; thereby enabling better temperature control throughout the entire bake cycle. In this Chapter, a simple approach to improve the repeatability of the bake process is investigated.

Thermal processing of the photomask is performed by placing the cold photomask (usually at room temperature) on the heated bakeplate for a given period of time. The bakeplate is maintained at a constant temperature by a feedback controller that adjusts the heater power in response to the reading of the temperature sensor embedded near the surface of the bakeplate. The placement of the cold photomask on the bakeplate creates a load disturbance to the bakeplate. Consequently, the bakeplate temperature drops abruptly as heat is removed from the bakeplate by the photomask. If the bakeplate is under feedback control only, the feedback controller of the bakeplate will respond to the sudden temperature drop by increasing the heater power sharply to reject this load disturbance. Usually, this results in an overshoot of the bakeplate temperature that can affect the CD, before the feedback controller can return the bakeplate temperature to its setpoint [15, 43].

The photomask has a large thermal mass. Typical dimension of a leading edge photomask is a square cross section of 6 inch by 6 inch and a thickness of 0.25 inch, as compared to the wafer dimension of 300 mm diameter and 1 mm thick. Due to its larger thermal mass, the loading effect due to the placement of the cold photomask is more severe with a larger temperature drop. Therefore, a longer time is required for the bakeplate to recover to its original temperature setpoint in photomask thermal processing as compared to wafer processing. To ensure repeatable processing temperature trajectory from photomask-to-photomask, the bake system has to be at steady state before processing subsequent photomask. Figure 2.1 shows the drop in the bakeplate temperature when a photomask is placed on the bakeplate without any feedback control. The heater power is kept at a constant value. With a substantial drop in bake temperature of more than 10 °C, it will take a long time for the bakeplate temperature to be restored to its setpoint. In the case where the

feedback controller is not well-tuned, a large temperature overshoot might also occur. Usually, there is an error budget associated with processing of the photomask [44]. As the photomask goes through many processing steps, errors introduced in each step leads to errors in the final CD. For a specified error tolerance, large errors in other processing steps can be compensated by reducing the temperature errors (deviation of the bakeplate temperature from its setpoint) introduced in the baking steps. One approach is to minimize the temperature overshoot during the bake process.

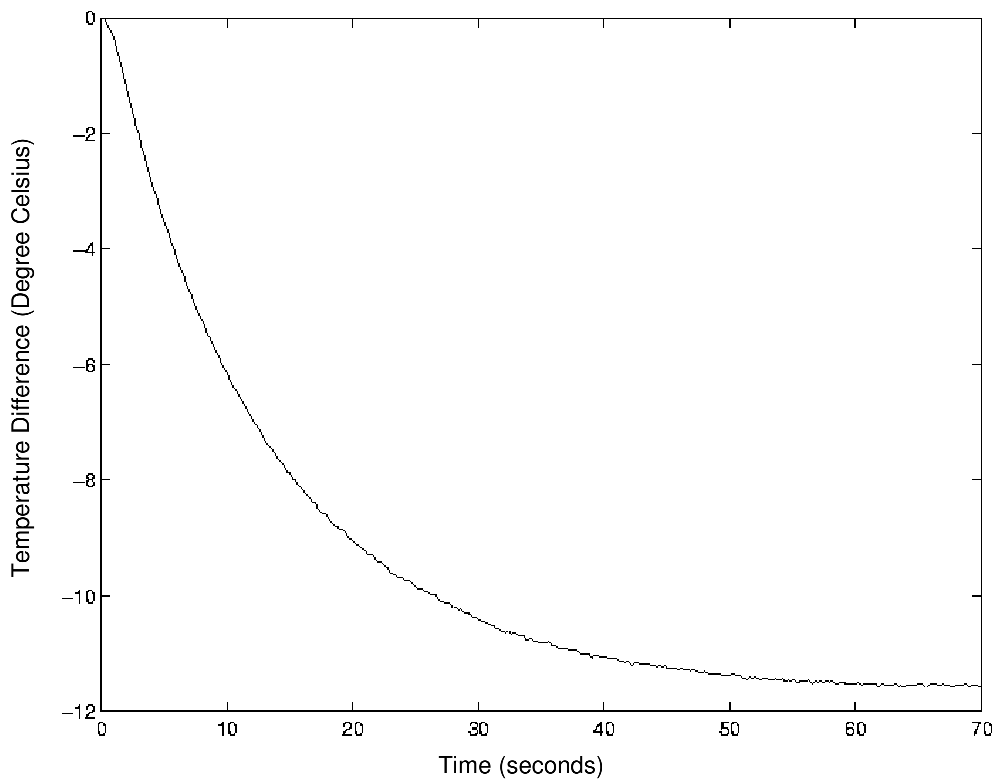


Figure 2.1: Load disturbance to the bakeplate due to placement of the photomask.

The load disturbance arises due to the transfer of heat from the bakeplate to the photomask. It is largely dependent on the thermal mass of the photomask. Therefore, prior

knowledge of the load disturbance is available. For load disturbance with prior knowledge, the solution is to eliminate the load disturbance effectively before it takes effect on the bakeplate, which is the subject of this Chapter. In this Chapter, an optimal predictive controller is designed to perform a pre-determined heating sequence prior to the placement of the photomask on the bakeplate. This is part of a feedback/feedforward strategy to eliminate the load disturbance. Experimental result shows a substantial improvement in the temperature performance of the bakeplate using the feedforward/feedback control strategy, as compared to the use of only a feedback controller. In addition, this model-based control strategy is simple and can be implemented easily on any existing commercial bakeplates with only a slight modification. Section 2.2 describes the experimental setup while Section 2.3 and 2.4 present the formulation and implementation of the feedforward/feedback control respectively. The conclusion of this research work is given in Section 2.5

2.2 Experimental Setup for Photomask Thermal Processing

An optimal feedforward/feedback control strategy is implemented for the baking of the quartz photomask on a multi-zones bakeplate. Figure 2.2 shows the photograph of the bakeplate used for baking a 6 inch by 6 inch photomask. The top view and cross section of the bakeplate is shown in Figure 2.3. The bakeplate consists of a two dimensional 7×7 array of independently controlled resistive heating elements with embedded resistance temperature detectors (RTDs) [40]. Each heating element consists of a thin aluminum square plate with the dimension of 1 inch by 1 inch and 1/16 inch thick. The aluminium plate is

supported by a shaft which is used to house a resistive (AC driven) catridge heater that delivers the heating power to the bakeplate. The RTDs are located within 0.050 inches from the bakeplate surface, providing in-situ temperature measurements during the run. The heating elements are disjoint with small air gaps of approximately 50 mils wide. This reduces the amount of thermal coupling between the heating elements, thereby enhancing the controllability of the bake system. The input to the heater (control signal) has a range of 0-10 V. This input voltage to the heater is varied to adjust the heater power, which in turn creates the change in the bakeplate temperature necessary to reject the load disturbance. The number of zones of the bakeplate can be easily configured, depending on the application. In this application, the bakeplate is configured as a single zone bakeplate.

Figure 2.2: Photograph of the bakeplate.

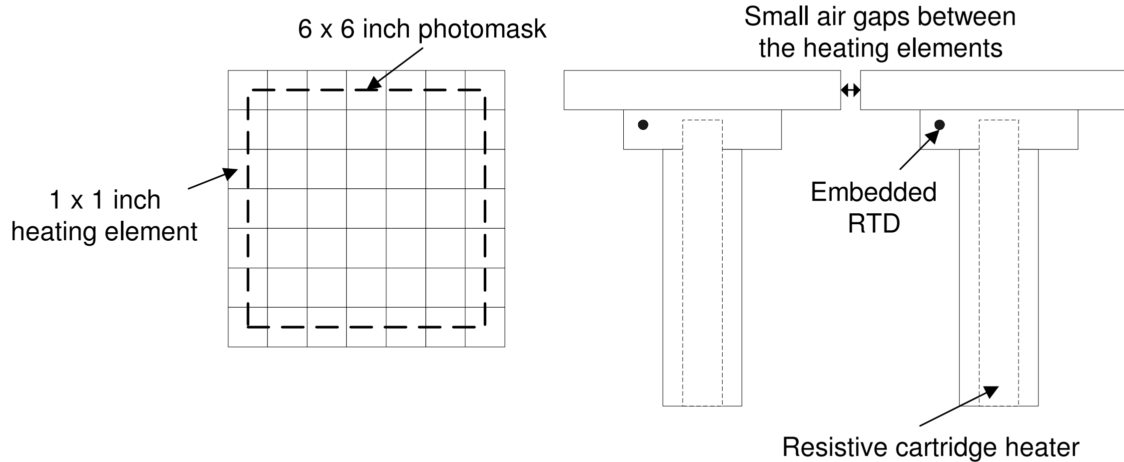


Figure 2.3: Top view and cross-section of the bakeplate.

2.3 Optimal Feedforward/Feedback Control Strategy

In this section, the control strategy used to compensate for the load disturbance induced by the placement of the cold photomask on the bakeplate is developed. The approach is to design a controller that, as best as possible, eliminates the loading disturbance on the bakeplate. Figure 2.4 shows the proposed control system where $G_c(s)$ is the transfer function of the existing feedback controller for the bakeplate and $G_p(s)$ is the transfer function of the bakeplate relating the bakeplate temperature, T, and the heater input, u. The heater input, u, comprises two components: feedback control signal, u_{fb}, and feedforward control signal, u_{ff}. The load disturbance on the bakeplate, which translates into a sudden drop in bakeplate temperature, is denoted as d. As can be seen in Figure 2.1, the load

disturbance can be modeled as a step response as heat is removed by the large thermal mass photomask. That is, the load disturbance, d, is the output of the disturbance model, $G_d(s)$, with a step input such that $u_d = 1\text{ V}$. To reflect the control problem accurately, the load disturbance is incorporated at the output of the plant rather than at the input of the plant. Since the objective of this control strategy is to reject the load disturbance, only the feedforward controller is designed and discussed. No modification is made to the existing feedback controller of the bakeplate, $G_c(s)$, which is a PID controller.

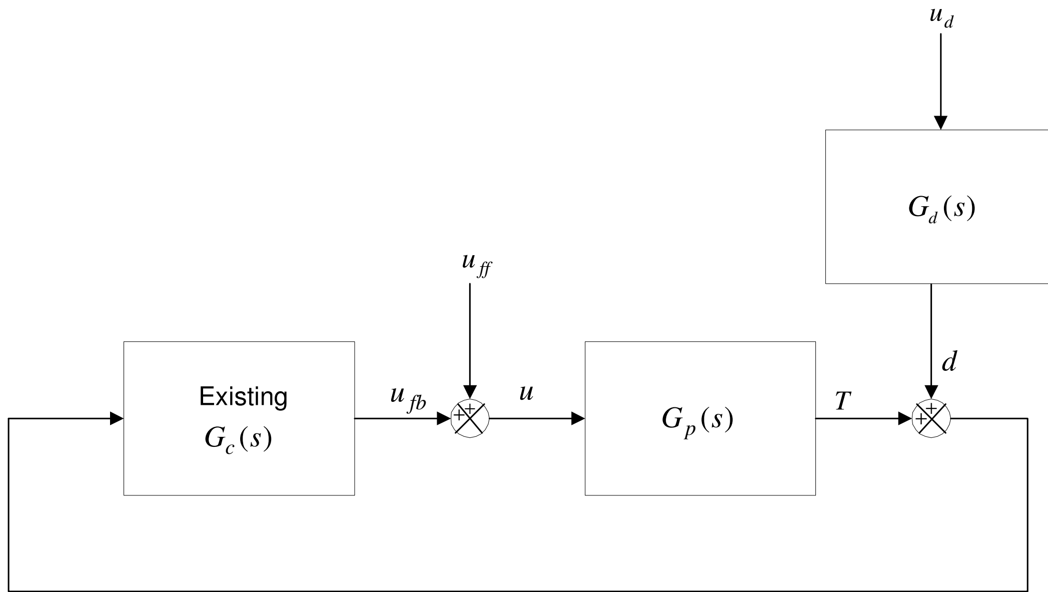


Figure 2.4: Feedforward/feedback control strategy for photomask thermal processing.

From Figure 2.4, it is noted that the loading effect on the bakeplate temperature can be eliminated if the change in bakeplate temperature from its setpoint, T, is equal to the negative of the disturbance d, i.e. $T(s) = -d(s)$. This can be accomplished without any

feedback control, i.e. $u_{fb} = 0$, by adjusting the heater input (control signal) according to the relation

$$u(s) = u_{ff}(s) = -G_p^{-1}(s)d(s) \tag{2.1}$$

This results in a non-causal feedforward control move where control moves are made before the actual placement of the photomask. If the plant model, $G_p(s)$, is known prior to the run, the load disturbance can be eliminated. In practice, the control signal, u, is subjected to saturation of the heater, $u \in [0, U^{\max}]$. If the required control signal is outside the achievable bounds, it is not possible to generate the desired temperature change, T, that is needed to eliminate the disturbance, d. A simple implementation strategy would be to calculate the perfect control moves using Equation (2.1), and then truncate at the upper bound of the achievable heater input voltage, $U^{\max}$. However, this approach will not result in an optimal solution.

In this section, an optimal solution to the control problem is considered. The problem is discretized in a sampled data format where the sampling index is denoted as k. The goal is to find a pre-determined heating sequence, $u_{ff}(k)$, that minimizes the maximum absolute temperature error between the bakeplate temperature and load disturbance, $e(k)$, for $k \in \{0,1,\cdots N\}$ over a finite interval, N. Minimization of the maximum temperature error is chosen as that will have the greatest impact on the recovery time. This is a min-max control problem that can be expressed as

$$\min_{u(k)\in[0,U^{\max}],\, k\in\{0,1,\cdots,N\}} \max |e(k)| \tag{2.2}$$

where $e(k)=T(k)+d(k)$. The solution automatically constrains the heater input, $u(k)$, to its pre-specified lower and upper limits of 0 and $U_{\max}$ respectively. This optimization problem is solved computationally by the use of a model. The transfer function relating change in bakeplate temperature, T, to the heater input, u, is represented as an auto-regressive model [45] such that

$$A(q^{-1})T(k)=B(q^{-1})u(k) \tag{2.3}$$

where q^{-1} is the backward shift operator such that ($q^{-1}T(k)=T(k-1)$), and the $A(q^{-1})$ and $B(q^{-1})$ polynomials are given as

$$A(q^{-1})=1+a_1q^{-1}+\cdots+a_nq^{-n}$$
$$B(q^{-1})=\left(b_0+b_1q^{-1}+\cdots+b_nq^{-n}\right)q^{-n_d}$$

where q^{-n_d} denotes any possible time delay and the order of the polynomials is given by n. The coefficients a_i and b_i are obtained from the identification experiment in Section 2.4. The discrete-time representation given in Equation (2.3) can be expressed in a convolution model at sample time instant, k, such that

$$T(k) = \sum_{j=0}^{k} c_j q^{-j} u(k - n_d)$$

where the coefficient, c_j, is given by

$$c_j = b_j - \sum_{i=1}^{n} a_i c_{j-i}$$

Over a finite interval, N, the input and output signals can be represented as finite-dimensional vectors. Expressed in a Vector-Matrix form, the relationship between the input vector, **U,** and output vector, $\mathbf{T}$, over the interval, N, is given as

$$\mathbf{T} = \Psi\, \mathbf{U}$$

where

$$\mathbf{T} = \begin{bmatrix} T(0) \\ T(1) \\ \vdots \\ T(N) \end{bmatrix}_{(N+1)\times 1}, \quad \mathbf{U} = \begin{bmatrix} u(0-n_d) \\ u(1-n_d) \\ \vdots \\ u(N-n_d) \end{bmatrix}_{(N+1)\times 1}, \quad \Psi = \begin{bmatrix} c_0 & 0 & \cdots & \cdots & 0 \\ c_1 & c_0 & 0 & \cdots & 0 \\ \vdots & & & & \\ c_N & c_{N\text{-}1} & \cdots & c_1 & c_0 \end{bmatrix}_{(N+1)\times(N+1)}$$

The optimization problem in Equation (2.2) can be formulated as a Linear Programming (LP) problem [46] and solved using a standard LP solver as follows.

Letting $z = \max_{k \in \{0,1,\cdots,N\}} |\, e(k)\,|$, the min-max problem given in Equation (2.2) is converted into

$$\min_{\mathbf{U}} \left[\mathbf{0}_{1\times(N+1)}\ , 1\right] \begin{bmatrix} \mathbf{U} \\ z \end{bmatrix}$$

Subject to

$$
\begin{aligned}
e(k) &\le z && \forall k \in \{0,1,\cdots,N\} \\
-e(k) &\le z && \forall k \in \{0,1,\cdots,N\} \\
\mathbf{U} &\le \mathbf{1}_{(N+1)\times 1} U^{\max} && \\
\mathbf{U} &\ge \mathbf{0}_{(N+1)\times 1} &&
\end{aligned}
$$

where $\mathbf{1}_{(N+1)\times 1}$ is a column vector with all its $N+1$ entries equal to 1. Also, $\mathbf{0}_{(N+1)\times 1}$ is a column vector with all its $N+1$ entries equal to 0 while $\mathbf{0}_{1\times(N+1)}$ is a row vector with all its $N+1$ entries equal to 0. The first two constraints can be converted into vector form given as

$$
\begin{aligned}
\mathbf{T}+\mathbf{D} &\le \mathbf{1}_{(N+1)\times 1} z && \text{or} && \mathbf{T}-\mathbf{1}_{(N+1)\times 1} z \le -\mathbf{D} \\
-(\mathbf{T}+\mathbf{D}) &\le \mathbf{1}_{(N+1)\times 1} z && \text{or} && -\mathbf{T}-\mathbf{1}_{(N+1)\times 1} z \le \mathbf{D}
\end{aligned}
$$

where $\mathbf{T} = \Psi \mathbf{U}$ and $\mathbf{D}$ is a column vector with $N+1$ entries such that $\mathbf{D} = [d(0), d(1), \cdots, d(N)]^T$. The superscript, T, denotes the transpose of a vector. The optimal control sequence, $\mathbf{U}$, is obtained by solving the following LP problem:

Minimize

$$
\min_{\mathbf{U}} \left[\mathbf{0}_{1\times(N+1)}\ , 1\right] \begin{bmatrix} \mathbf{U} \\ z \end{bmatrix}
$$

Subject to

$$
\begin{bmatrix} \Psi & -\mathbf{1}_{(N+1)\times 1} \\ -\Psi & -\mathbf{1}_{(N+1)\times 1} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ z \end{bmatrix} \le \begin{bmatrix} -\mathbf{D} \\ \mathbf{D} \end{bmatrix}
$$

$$\mathbf{U} \le \mathbf{1}_{(N+1)\times 1} U^{\max}$$
$$\mathbf{U} \ge \mathbf{0}_{(N+1)\times 1}$$

For any vectors $\mathbf{v}$ and $\mathbf{w}$, $\mathbf{v} \le \mathbf{w}$ means every entry of $\mathbf{v}$ is less than or equal to the corresponding entry of $\mathbf{w}$. For example, the constraint, $\mathbf{U} \le \mathbf{1}_{(N+1)\times 1} U^{\max}$, means that every entries in the vector, $\mathbf{U}$, is less than or equal to $U^{\max}$. That is, $u(k) \le U^{\max} \;\; \forall k \in \{0,1,\cdots,N\}$.

2.4 Implementation of Constraint Feedforward Controller

Before the constraint feedforward controller can be implemented, information about the plant and disturbance models, $G_p(s)$ and $G_d(s)$, must be available. When the photomask is placed on the bakeplate, it induces a load disturbance to the bakeplate and causes the bakeplate temperature to drop. The load disturbance is determined as follows. First, the input to the heater (control signal) is fixed by putting the controller in manual mode. A 6 inch by 6 inch photomask at around room temperature (~24 °C) is then placed on the bakeplate. The sampling interval is 0.2 seconds. The resultant load disturbance is modeled as the output of a transfer function with a step input. By fitting the experimental data in the least squares sense, the disturbance model, $G_d(q^{-1})$, is identified as

$$G_d(q^{-1}) = \frac{-0.0814}{1 - 1.5312q^{-1} + 0.5382q^{-2}}$$

This is obtained by selecting the smallest order of the $A(q^{-1})$ and $B(q^{-1})$ polynomials that gives a reasonably good fit between the experimental and fitted data. The system identification toolbox in MATLAB is used to aid in the selection of the suitable model. Figure 2.5 shows the drop in bakeplate temperature due to the placement of the photomask. The dashed line shows the calculated response using the identified disturbance model, $G_d(q^{-1})$.

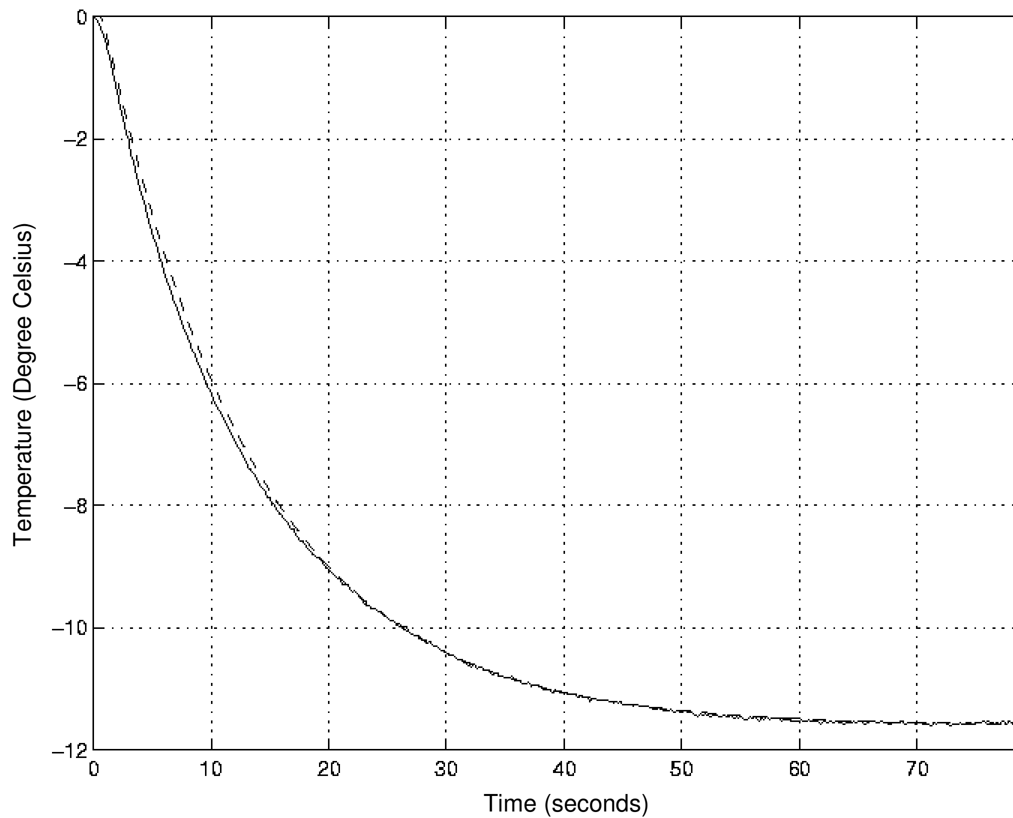


Figure 2.5: Result of the identification experiment to obtain the disturbance model.

The model of the bakeplate, $G_p(q^{-1})$, relating the change in bakeplate temperature, T, to the change in heater input, u, is similarly identified. A pseudo random binary control

sequence [47] is injected into the bakeplate as shown in Figure 2.6 (b). Using least squares estimation, the process model, $G_p(q^{-1})$, is identified as

$$G_p(q^{-1}) = \frac{0.0037}{1 - 2.11q^{-1} + 0.78q^{-2} + 0.83q^{-3} - 0.50q^{-4}}$$

The solid line in Figure 2.6 (a) shows the resulting change in bakeplate temperature while the dashed line shows the calculated response using the identified model.

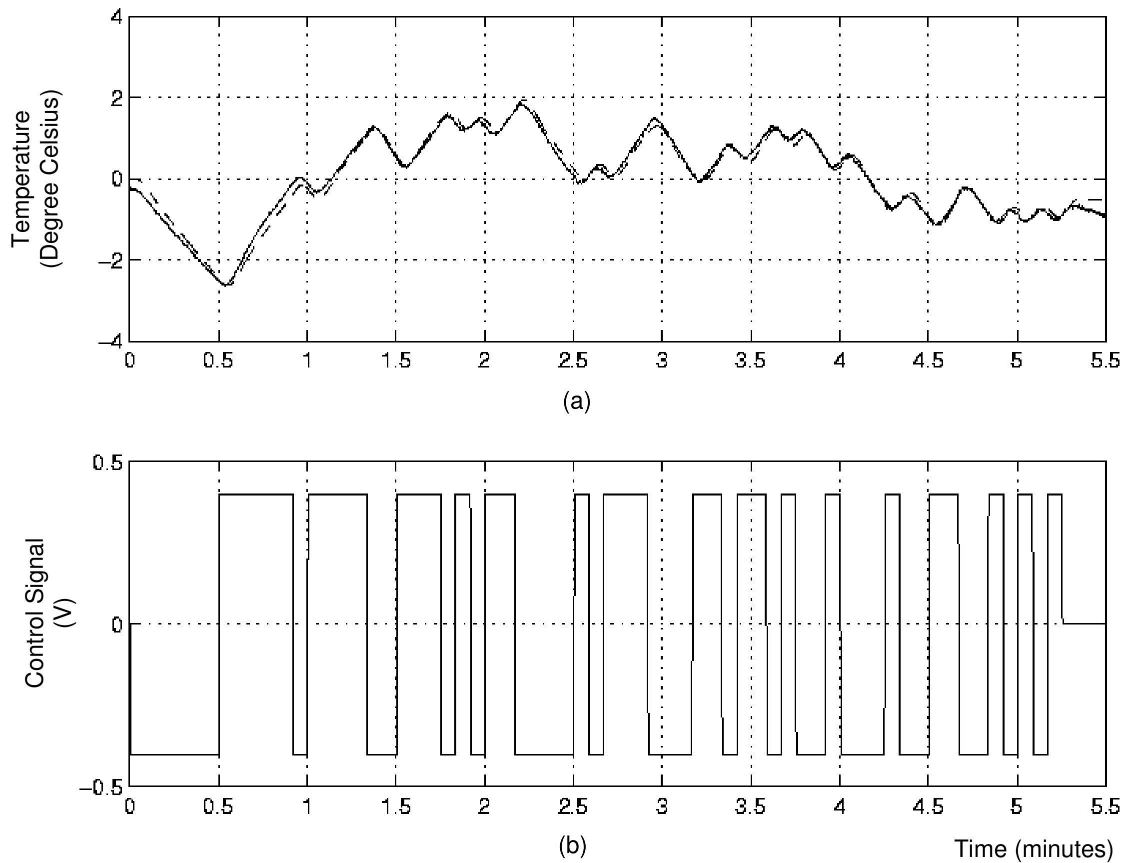


Figure 2.6: Result of the identification experiment to obtain the plant model.

Having obtained the model for the bakeplate and load disturbance, the optimal feedforward control signal, u_{ff}, is computed. At steady state, the feedback control signal, u_{fb}, that is required to maintain temperature at 90°C is 0.5 V. The control signal, $u = u_{fb} + u_{ff}$, is subjected to a heater input voltage constraint where $0 \le u \le 10$. This places a constraint of $-0.5 \le u_{ff}(k) \le 9.5 \ \forall k \in \{0,1,\cdots,N\}$ on the feedforward control signal. Figure 2.7 shows the computed feedforward control signal. Note the control signal is brought to its maximum level and then to its minimum level in a bang-bang control type fashion. This type of response and prediction would be difficult to determine using a trial and error hand tuning method without the use of a model.

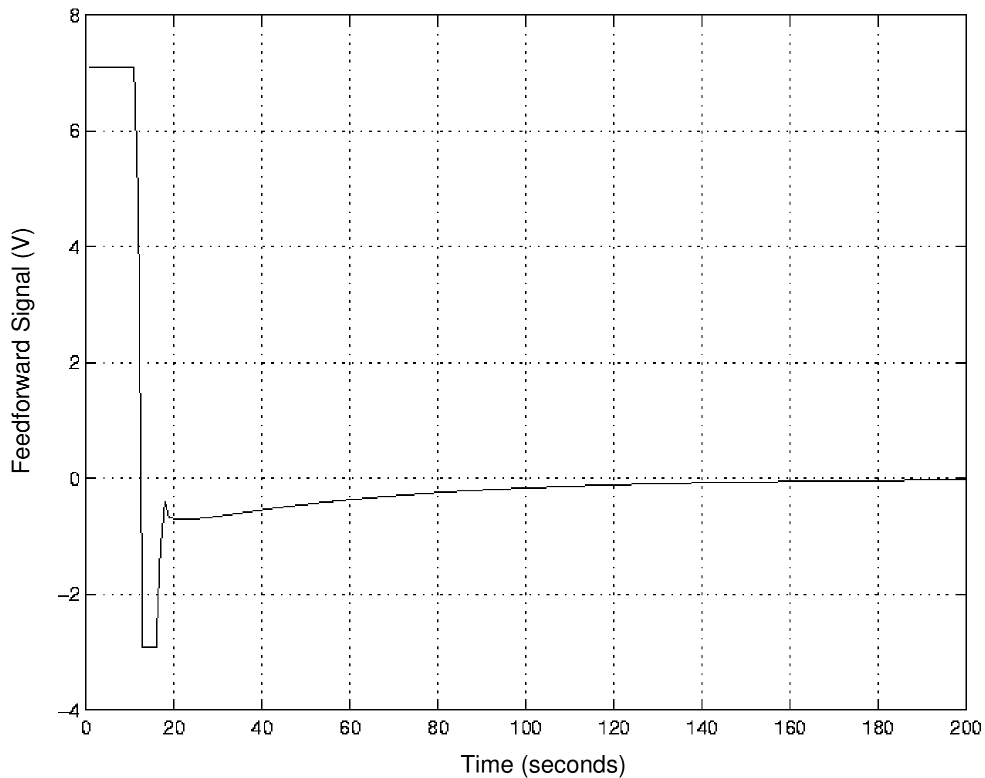


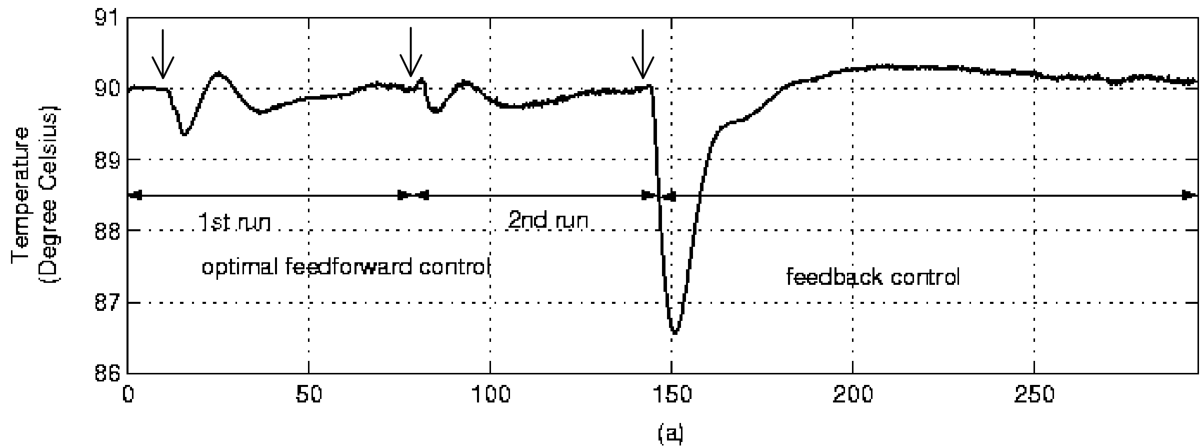
Figure 2.7: Optimal feedforward control signal

Figure 2.8 gives the comparison for the case with and without feedforward control when a 6 inch by 6 inch photomask is placed on the bakeplate. The arrow in Figure 2.8 indicates the time when the photomask is placed on the bakeplate. The first two runs are for optimal feedforward control while the last run is for feedback control only. Feedforward control signal is applied prior to the arrival of the photomask for the first two runs. With only feedback control, the drop in bakeplate temperature is ≈3.5 °C. In contrast, the drop in temperature is ≈0.5 °C with feedforward control. There is a substantial improvement in the integrated-square temperature error (ISE) between the optimal predictive controller and the feedback controller. The ISE for the first two cases of optimal feedforward control and a third case of feedback control are 3.25, 1.63 and 99.66 respectively. Using the optimal feedforward controller, there is about 30 times improvement in the ISE.

2.5 Conclusion

An optimal feedforward control scheme has been designed and demonstrated to minimize the loading effect induced by the placement of a cold photomask on the bakeplate. The elimination of the loading effect is important to current and future generation of DUV resist, which is extremely sensitive to temperature variations. Of more importance is the improvement in the thermal budget. For a specified error budget, larger thermal errors in other processing steps can accommodated due to reduction in temperature errors in the baking steps.

The feedforward control strategy is based on a linear programming method of minimizing the worst-case deviation from the temperature setpoint during the loading effect. This results in a predictive controller that performs a pre-determined heating sequence prior to the arrival of the photomask in order to eliminate load disturbance induced by placement of the cold photomask on the bakeplate. A significant improvement in the integrated-square temperature error between the predictive controller and the feedback controller is achieved.



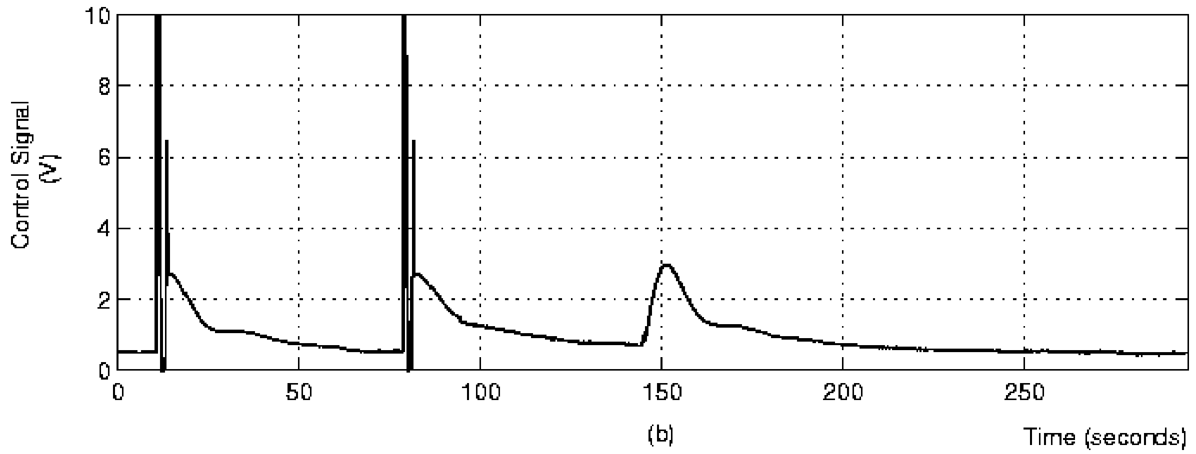


Figure 2.8: Comparison between runs with and without feedforward control.

Chapter 3

Real-time Predictive Control of Resist Film Thickness Uniformity

3.1 Introduction

To form the resist patterns, the wafer substrate is spin-coated with a thin film of resist, followed by a softbake process to remove excess solvent in the resist film. The desired patterns are then patterned onto the resist film by exposing the substrate with deep UV radiation. During this exposure step, some of the incident light propagates through the resist film and reflects at the substrate-resist interface. The phase difference between the incident and reflected light creates an interference effect within the resist film and the total amount of light absorbed by the resist film "swings" up and down as a function of the resist film thickness. An example of such variation of the light absorption as a function of resist thickness, typically known as a swing curve, is shown in Figure 3.1. As a result of this variation of light absorption in the resist film, the final CD also varies with resist thickness [48]. The CD can vary as much as 4 nm for every 1 nm change in resist thickness [49]. The resist thickness has to be well controlled to remain at the extrema of the swing curve where the sensitivity of CD to resist thickness variation is minimized [15]. With shrinking feature size, light of shorter exposure wavelength is used to pattern the resist-coated wafer. As the amplitude and periodicity of the swing curve increases with decreasing exposure wavelength,

control of the resist thickness and its uniformity becomes increasingly more important [49]. Furthermore, the industry is moving towards the use of 300-mm substrate for economic reasons; placing a stringent demand on all lithographic processes as the control requirement is spread over a larger area. To meet the demanding requirement of advanced lithography, better control of the resist film and its properties will be required.

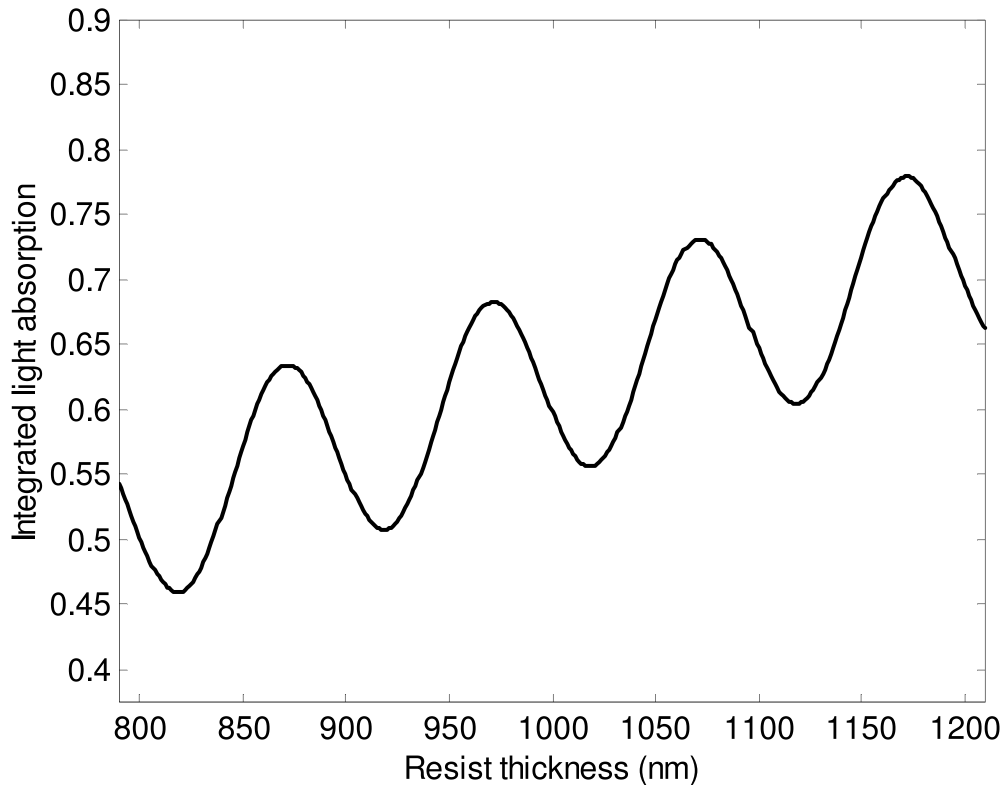


Figure 3.1: An example of a swing curve

The spin coating process affects the uniformity of the resist film coating across the wafer. Low spin speed can result in resist thickness profile, which increases towards the edge of the substrate. High spin speed can result in the opposite effect. For spin coating over topographical features, the resist uniformity is even more significant [50]. Already for the

200 mm substrate, resist thickness uniformity specification is met by having tight controls over important parameters such as relative humidity, chuck temperature, spin speed, exhaust, etc during spin coating [3, 15, 48]. With the use of a larger substrate, the specifications for these parameters become more demanding. The complexity of the spin coating process also increases as the transition from laminar to turbulent flow now occurs at a lower spin speed; limiting the range of useful thickness for any fixed viscosity resist. This transition to turbulent flow during spin coating is largely responsible for the increase in resist thickness non-uniformity at the edge of the wafer [51, 52]. However, it is sometimes necessary to spin the resist at higher speed to obtain the optimum resist thickness, as indicted by the extrema of the swing curve. To achieve better control of the resist thickness, new coating technology and techniques have been investigated [52-55]. Gurer *et al.* [52] decoupled the convective and evaporation mass transfer mechanism during the spin coating process to improve the control of the resist film thickness and its uniformity. A new coating technique using the rotary cup is investigated to provide a fringe-free coating for the photomask by Kushida *et al.* and Tokimitsu *et al.* [53, 54]. Resist coating by spray technology on high topography wafer structure is also investigated by Suriadi *et al.* [55].

Besides the spin coating process, another important process that affects the resist thickness and its uniformity is the softbake process. It is performed after the spin coating process to remove excess solvent from the resist film, reduce standing waves and relax the resist polymer chain into an ordered matrix. As in all bake processes, temperature control [48, 56] during softbake is important. Conventionally, the resist is baked at a fixed temperature with temperature control of ± 1 °C for consistent lithographic performance [48]. In general, the resist thickness formed after the spin coating process will not be uniform. If a

non-uniform resist film is formed during the resist coating process, experiments have shown that maintaining a uniform temperature profile across the bakeplate will not reduce the resist film non-uniformity. In this Chapter, a new technique to monitor and control the resist thickness in real-time during the softbake process is investigated. There have been some researches to control the resist thickness during the bake process through in-situ monitoring of the resist thickness and properties. To study the bake mechanism, Paniez *et al.* [4] uses in-situ ellipsomtery while Fadda *et al.* [57] uses contact angle measurements to monitor the resist thickness in real-time during the bake process. Morton *et al.* [28, 29] uses in-situ ultrasonic sensors to monitor the change in the resist properties to determine if the resist has been sufficiently cured. Metz *et al.* [6] uses in-situ multi-wavelength reflection interferometers to measure the resist thickness versus bake time to determine the optimum bake time. In these approaches, the in-situ measurements are used to determine the completion of the bake process when certain conditions are satisfied, such as when the resist is sufficiently cured or if a specified resist thickness has been achieved.

To implement real-time thickness control during softbake, an array of in-situ thickness sensors is positioned above a multi-zones bakeplate to monitor the resist thickness at various sites on the wafer. The bake temperature is also constrained during softbake to prevent decomposition of the photoactive compound [29, 48]. With these in-situ resist thickness measurements, the temperature profile of the bakeplate is controlled in real-time by manipulating the heater power distribution using the Generalized Predictive Control (GPC) algorithm. This is an advanced control algorithm whereby constraints on the bakeplate temperature can be readily implemented. Often, it results in a non-uniform temperature distribution. This creates different solvent removal rates and densifications at various sites

on the wafer to give a uniform resist thickness distribution across the wafer. For this research, not only are the in-situ resist thickness measurements used to detect the endpoint of the softbake process, they are also used together with the GPC algorithm to improve the resist thickness uniformity by manipulating the bakeplate temperature distribution in real-time. Resist thickness non-uniformity of less than 1 nm at a specified target thickness has been achieved, with an average of 10 times improvement in resist thickness uniformity at the end of the bake process. With the stringent demand of advanced lithography, this ability to squeeze out the last few nanometers of the process is important. This will also help to relax the tight specification imposed during the coating process. Besides silicon wafers, the control strategy can also be applied to photomask and flat panel display manufacturing.

The rest of the chapter is organized as follows. Section 3.2 describes the experimental setup used for real-time thickness control during softbake. The thickness estimation algorithm for monitoring the resist thickness uniformity is given in Section 3.3. The GPC controller is presented in Section 3.4 and the experimental results are discussed in Section 3.5. Section 3.6 summarizes the research work on the real-time thickness control strategy using the GPC algorithm.

3.2 Experimental Setup

Typically, a non-uniform resist film is formed at the end of the spin coating process. Real-time thickness control is implemented during softbake to compensate for the resist

thickness non-uniformity formed during the prior coating process. Figure 3.2 shows the schematics of the experimental setup used to control the resist thickness in real-time. The cross section of the experimental setup is given in Figure 3.3.

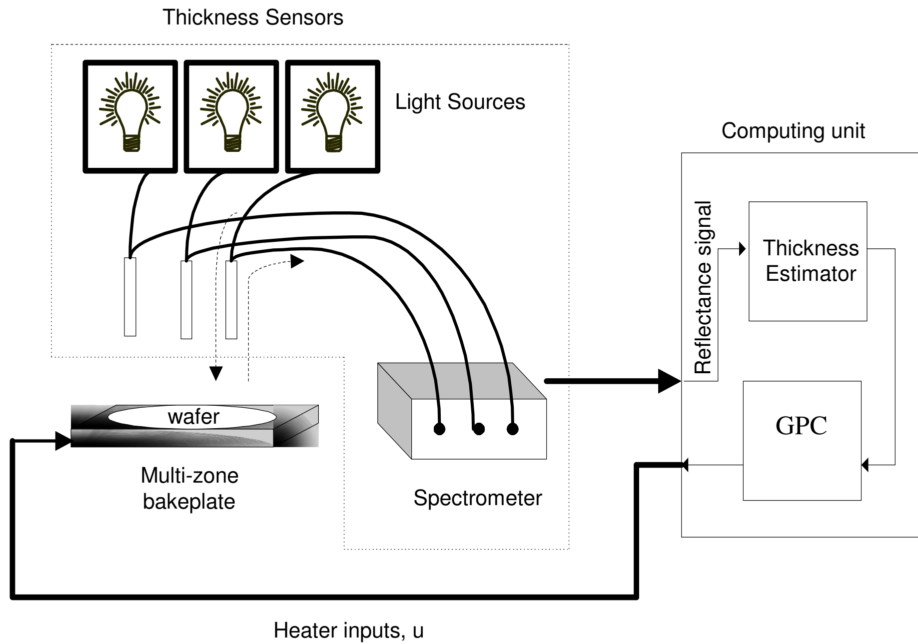


Figure 3.2: Schematics of the experimental setup used to control the resist thickness in real-time.

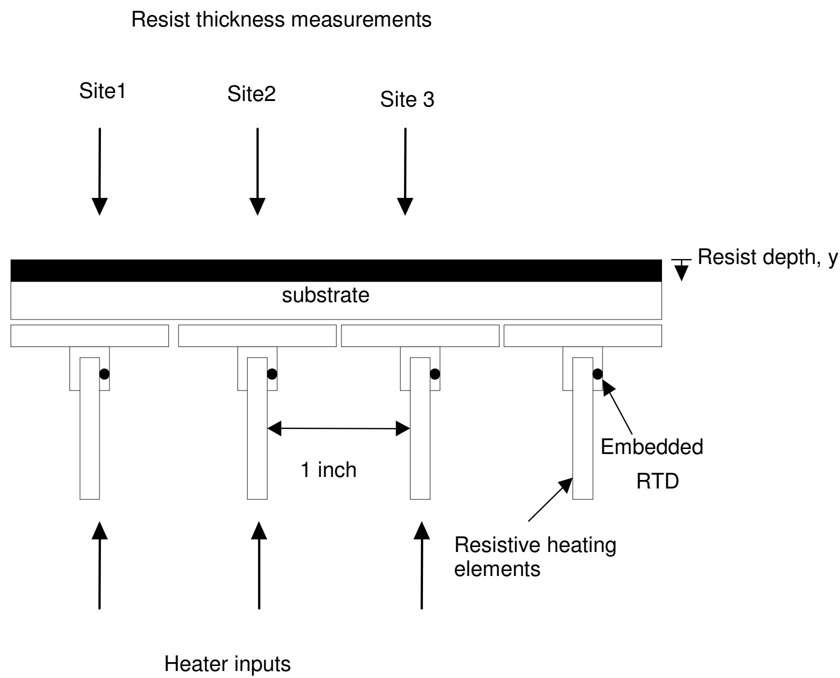


Figure 3.3: Cross-section of the experimental setup.

The experimental setup consists of three main parts: a multi-zones bakeplate, an array of thickness sensors and a computing unit. A brief description of the bakeplate, thickness sensors and computing unit is given in the following.

Multi-zones Bakeplate

The multi-zones bakeplate used in this real-time thickness control strategy is the same as that in Chapter 2. The small thermal mass and fast dynamics of the bakeplate makes it suitable for real-time application. In this application, instead of configuring the bakeplate as a single zone, it is configured into multiple independently controlled heating zones. This

gives us the flexibility to control resist thickness through temperature manipulation at different locations on the bakeplate.

Thickness Sensor

During softbake, the resist-coated wafer is placed on the multi-zones bakeplate and the resist thickness at different sites on the wafer are monitored by an array of in-situ thickness sensors mounted directly above the bakeplate. Three sites, each one inch apart, are monitored and controlled to demonstrate the control strategy (see Figure 3.3).

The thickness sensor has a similar setup as the multi-wavelength development rate monitor in [58]. It comprises a broadband light source (LS-1), a spectrometer with the capability of monitoring the reflected light intensity at three sites simultaneously (SQ2000) and a bifurcated fiber optics reflection probe (R200) from OceanOptics®. The reflection probe consisting of a bundle of 7 optical fibers (6 illumination fibers around 1 read fiber) is positioned above the wafer to monitor the resist thickness in real-time. During softbake, light from the broadband light source is focused on the resist through the illumination fibers of the bifurcated reflection probe and the reflected light is guided back to the spectrometer through the read fiber of the probe.

Computing Unit

Reflectance measurements are acquired through the A/D converter and converted to resist thickness measurements using a thickness estimation algorithm in a Labview

environment. The thickness estimation algorithm is discussed in greater details in Section 3.3. With the availability of the resist thickness measurements, the GPC algorithm computes the heater power distribution that is required to minimize the resist thickness non-uniformity between the the monitored sites. Currently, the setup is for a 4-inch wafer (radius: 2 inches; 3 points monitored). This can be easily scaled to a 12-inch wafer (radius: 6 inches; 7 points monitored). The number of thickness sensors and hence the amount of computation required for the 12-inch wafer is roughly doubled. Figure 3.4 shows the photograph of the experimental setup with a 4-inch wafer placed on top on the multi-zones bakeplate.

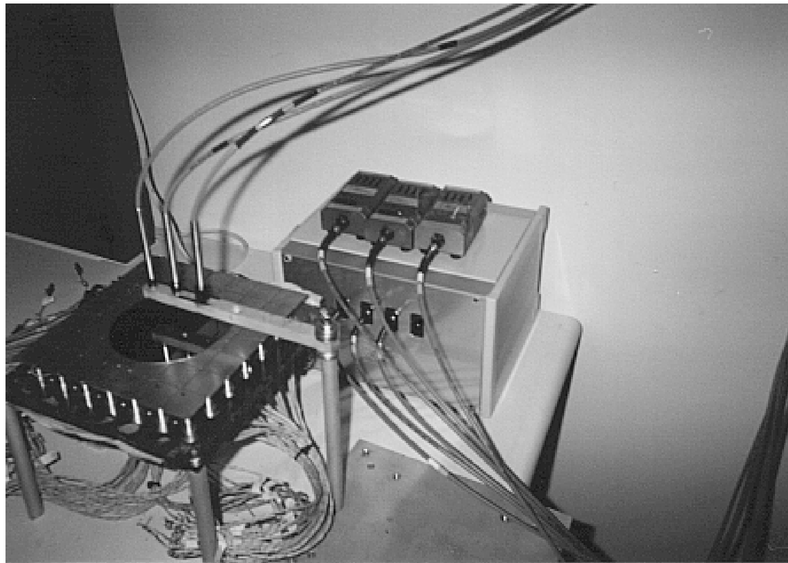
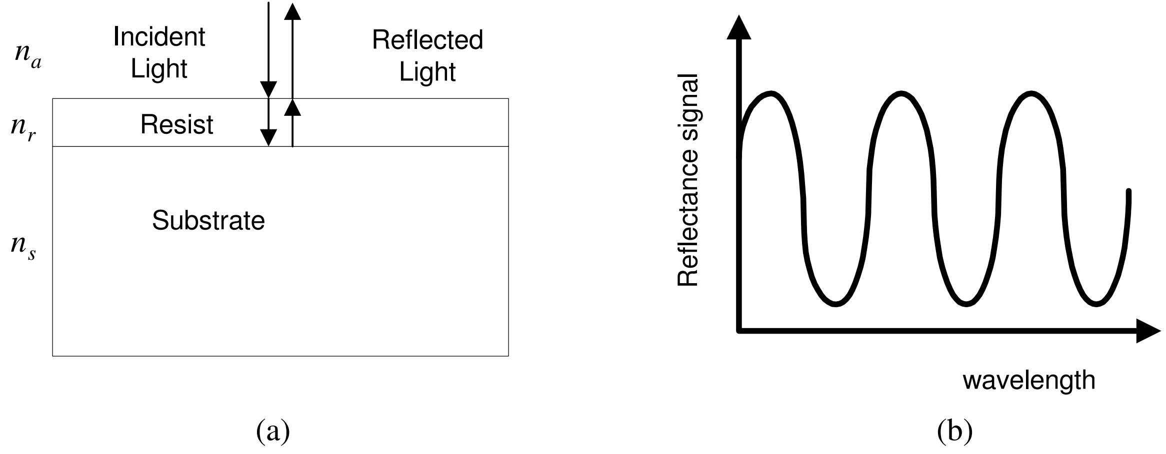


Figure 3.4: Photograph of the experimental setup.

3.3 Resist Thickness Estimation

In Figure 3.2, three thickness sensors enclosed by the dotted lines are used to provide in-situ resist thickness measurements during the bake process. Each thickness sensor comprises a broadband light source and a bifurcated fiber optics reflection probe and is connected to a spectrometer [59].

A thin film optical model is used to estimate the resist thickness from the reflectance signal, as shown in Figure 3.5 (a). The model assumes normally incident light and homogenous thin resist film. During wafer processing, light from the broadband light source is focused normally onto the resist-coated wafer through the illumination end of the bifurcated fiber optics reflection probe while the reflected light is guided back to the spectrometer through the read end of the reflection probe. Some of the incident light reflects at the top resist-ambient interface while part of the incident light propagates through the resist film and reflects at the substrate-resist interface. The additional optical path traveled creates a phase difference between the incident and reflected light. Constructive or destructive interference, which depends on the wavelength of the incident light and resist thickness, occurs in the resist film. Hence the reflectance signal, $h_m(\lambda, y_m)$, observed at the spectrometer also varies as a function of the resist thickness, y_m, and wavelength of the light source, λ. In this thesis, the subscript, m, denotes the measurement data. Figure 3.5(b) shows the typical variation of the reflectance signal with wavelengths for a particular resist thickness.



(a) (b)

Figure 3.5: (a) Thin film optical model, and (b) variation of the reflectance measurement with resist thickness for a particular resist thickness.

At normal incidence, the variation of the reflectance signal with wavelengths, $h(\lambda, y_m)$, can be modeled as [5, 60]

$$h(\lambda, y_m) = \left(p \times \left(\frac{r_2 e^{-i\delta} + r_1 e^{i\delta}}{r_1 r_2 e^{-i\delta} + e^{i\delta}} \right) \left(\frac{r_2 e^{-i\delta} + r_1 e^{i\delta}}{r_1 r_2 e^{-i\delta} + e^{i\delta}} \right)^* \right) \tag{3.1}$$

where

$$r_0 = \frac{n_a - n_s}{n_a + n_s}, r_1 = \frac{n_a - n_r}{n_a + n_r}, \quad r_2 = \frac{n_r - n_s}{n_r + n_s} \text{ and } \delta = \frac{2\pi n_r y_m}{\lambda}$$

and $(\ )^*$ denotes complex conjugation, y_m, is the resist thickness and λ is the wavelength of the incident light. To account for any difference in magnitude between measured and calculated reflectance measurements, a scaling factor, p, is included in Equation (3.1) so that the calculated reflectance measurements is scaled to have the same peak-to-peak amplitude as that of the experimental reflectance measurements at each time instant. Also, n_a, n_r and n_s are the refractive index of air, resist and silicon substrate respectively. The refractive index varies a function of the wavelength, λ. The variation of the refractive index with wavelengths is given by the Cauchy equation [61]:

$$n_i(\lambda) = A_i + \frac{B_i}{\lambda^2} + \frac{C_i}{\lambda^4} \qquad \text{for } i = r \text{ and } s$$

where A_i, B_i and C_i are the Cauchy parameters such that A_r = 1.6116, B_r = 2.33×10^3 nm^{-2} and C_r = 2.99×10^9 nm^{-4} for the Shipley 3612 resist. For silicon wafer, the Cauchy parameters are $A_s = 3.43$, $B_s = 1.33\times10^5$ nm^{-2} and $C_s = 1.9\times10^{10}$ nm^{-4}. The refractive index for air, n_a, is assumed to be constant and equal to 1 for all wavelengths.

The resist thickness, y_m, can be obtained by analyzing the reflectance signal, $h(\lambda, y_m)$, in Equation (3.1) over the range of wavelengths. One approach is to estimate the resist thickness from the reflectance signal using Least Squares Estimation. In this approach, Equation (3.1) is approximated by taking the Taylor series expansion such that

$$h(\lambda, y_m) \approx h(\lambda, y_{m0}) + \left.\frac{\partial h}{\partial y}\right|_{\lambda, y_{m0}} \Delta y_m$$

where y_{m0} is the initial thickness estimate and $\left.\frac{\partial h}{\partial y}\right|_{\lambda, y_{m0}}$ is the derivative computed at λ and y_{m0}. Higher order terms of the Taylor series expansion are assumed to be negligible for very small Δy_m. The derivative, $\left.\frac{\partial h}{\partial y}\right|_{\lambda, y_{m0}}$, is evaluated using forward difference and the estimated resist thickness, y_m^*, is given as

$$y_m^* = y_{m0} + \Delta y_m$$

The change in resist thickness, Δy_m, can be estimated using the least squares estimation method given by

$$\Delta y_m = \left(\frac{\partial \mathbf{h}}{\partial \mathbf{y}}^T \frac{\partial \mathbf{h}}{\partial \mathbf{y}} \right)^{-1} \frac{\partial \mathbf{h}}{\partial \mathbf{y}}^T (\mathbf{h} - \mathbf{h_o})$$

where

$$\frac{\partial \mathbf{h}}{\partial \mathbf{y}} = \begin{bmatrix} \left.\frac{\partial h}{\partial y}\right|_{\lambda_1, y_{m0}} \\ \left.\frac{\partial h}{\partial y}\right|_{\lambda_2, y_{m0}} \\ \vdots \\ \left.\frac{\partial h}{\partial y}\right|_{\lambda_M, y_{m0}} \end{bmatrix} \qquad \mathbf{h} = \begin{bmatrix} h_m(\lambda_1, y_m) \\ h_m(\lambda_2, y_m) \\ \vdots \\ h_m(\lambda_M, y_m) \end{bmatrix} \qquad \mathbf{h_0} = \begin{bmatrix} h(\lambda_1, y_{m0}) \\ h(\lambda_2, y_{m0}) \\ \vdots \\ h(\lambda_M, y_{m0}) \end{bmatrix}$$

For this application, a broadband light source of wavelengths between 480 nm and 850 nm is used to illuminate the wafer. At each sampling instant, 1100 reflectance measurements are obtained at wavelengths between 480 nm and 850 nm such that $M = 1100$, $\lambda_1 = 480$ nm and $\lambda_M = 850$ nm. At the start of the bake process, the initial thickness estimate, y_{m0}, is obtained from the prior coating process. Subsequently, the initial estimate, y_{m0}, is updated with the previous value of y_m^* at every sampling instant. During the bake process, the resist thickness decreases with the bake time due to evaporation of the solvent in the resist film. Therefore, the observed reflectance signal, $h_m(\lambda, y_m)$, also changes with the bake time. Figure 3.6 shows the measured reflectance signal, $h_m(\lambda, y_m)$, plotted against the calculated reflectance signal, $h(\lambda, y_m)$, over the range of wavelengths. The solid line gives the experimental reflectance signal while the dashed line gives the calculated reflectance signal. Using this approach, an estimated resist thickness of 1712.5 nm is obtained from the reflectance signal.

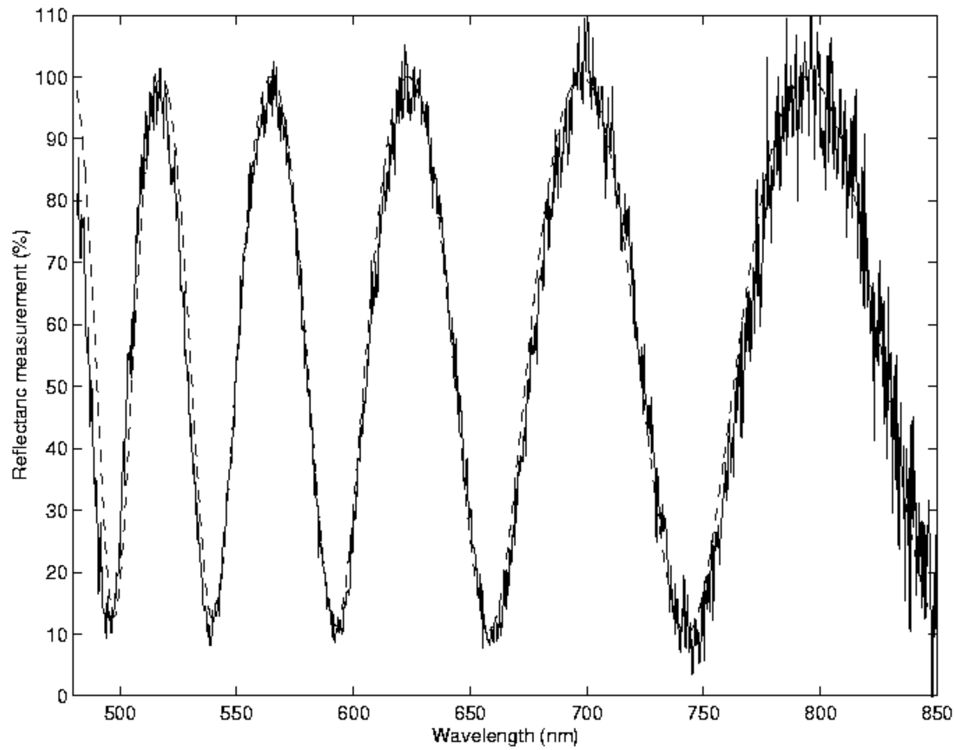


Figure 3.6: Comparison between the measured and calculated reflectance signal.

Besides the Least Squares Estimation approach discussed above, the measured reflectance signal, $h_m(\lambda, y_m)$, can be processed in several other ways to obtain the resist thickness, y_m. In this section, two other data analysis techniques are investigated. They are Nonlinear Least Squares Fit and Fringe Counting. A comparison is made between these three data analysis techniques. Each of these methods is a tradeoff between complexity, robustness and accuracy. In the Nonlinear Least Squares Fit approach, the resist thickness, y_m^*, is obtained by solving the nonlinear least squares optimization problem given as:

$$y_m^* = \arg\left(\min_{450 \le \lambda \le 850} (h_m(\lambda, y_m) - h(\lambda, y_m))^2 \right) \tag{3.2}$$

where $h_m(\lambda, y_m)$ is the measured reflectance signal and $h(\lambda, y_m)$ is given by Equation (3.1). The resist thickness, y_m^*, is taken to be that for which the minimum is achieved. Numerical algorithm such as Levenberg-Marquardt can be used to solve the nonlinear optimization problem. Among the three approaches, this approach is computationally most intensive. The reflectance measurements are converted to resist thickness measurements at a rate of 1 s in the Nonlinear Least Squares Fit approach. In the Least Squares Estimation approach, Equation (3.1) is linearized before solving the optimization problem in Equation (3.2). This reduces the computation time and resist thickness can be obtained every 0.1 s. Both Least Squares Estimation and Nonlinear Least Squares Fit methods require a good initial guess of resist thickness for the algorithms to work well. This is usually not a problem as a reasonably good initial estimate of the resist thickness is obtained from the prior coating process. Another approach that has been investigated is the Fringe Counting technique. As shown in Figure 3.6, the reflectance signal varies with wavelengths in an approximately periodic fashion. The distance between the peaks and valleys in the reflectance signal is a known function of the resist thickness. By noting the distance between the peaks and valleys in the reflectance signal over the wavelength range, the resist thickness at each time instant can be determined. Since only parts of the reflectance measurements are used, it is the fastest algorithm with a sampling rate of 0.02 s. However, it is also most sensitive to measurement noise.

To examine if smaller computational effort of the Fringe Counting and Least Squares Estimation methods comes at the expense of measurement accuracy, the relative accuracy of the resist thickness measurements using these two methods are compared with resist thickness measurements using the Nonlinear Least Squares Fit method. Since no

approximation is made to Equation (3.1) in the Nonlinear Least Squares Fit approach, the resist thickness measurement obtained is most accurate. Denoting the resist thickness measurement using the Nonlinear Least Squares Fit, Least Squares Estimation and Fringe Counting methods as y_{nl}, y_{ls} and y_{fr} respectively, the thickness error is assumed to be $y_{ls} - y_{nl}$ for the Least Squares Estimation approach and $y_{fr} - y_{nl}$ for the Fringe Counting approach. Figure 3.7 shows the thickness errors during a 4 minutes softbake at 90°C. Despite the approximation, the resist thickness estimated using the Least Squares Estimation method is almost the same as that using Nonlinear Least Squares Fit. There is only a negligible average thickness difference of 0.06 nm, as compared to 8.7 nm using the Fringe Counting approach. Therefore, Least Squares Estimation is most suited for real-time application in terms of accuracy and computation time and is used as the algorithm to provide in-situ resist thickness measurements during the bake process.

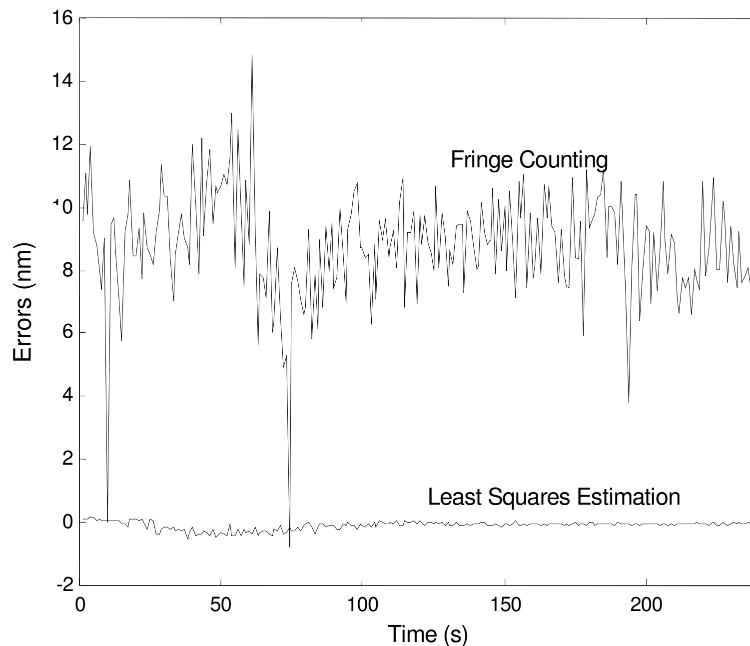


Figure 3.7: Comparison between the three thickness estimation algorithms

3.4 Generalized Predictive Control

When the resist thickness measurements, y_m, are available, the Generalized Predictive Control (GPC) algorithm is used to compute the heater power distribution of the bakeplate that is required to achieve a better resist film thickness uniformity across the wafer. The reference thickness trajectory and the estimated resist thickness are the inputs to the control algorithm while heater power distribution needed to minimize the resist thickness non-uniformity is the output. As baking the resist at excessively high temperature may lead to decomposition of the photoactive compound of the resist and result in degradation of the feature size, the bake temperature has to be constrained to an upper temperature bound during the softbake process [48]. This can be readily implemented in the GPC algorithm.

GPC is a model-based predictive control algorithm that uses the process model to predict the plant output at future instants. In this application, two process models are required for the prediction purposes. One of them is the thickness model relating the change in the resist thickness to the change in the heater input. It is used to predict the resist thickness over a prediction horizon. The other process model is the thermal model that relates the change in the bakeplate temperature to the change in heater input. This model is used to predict the bakeplate temperature to ensure that the bake temperature is constrained during the run. The predictive nature of the GPC algorithm makes it a suitable algorithm for implementing the real-time thickness control strategy. This is important, as solvent removal from the resist film is irreversible.

Before the GPC algorithm can be implemented, the process models have to be identified. Section 3.4.1 discusses the identification of the two process models. The control algorithm is covered in Section 3.4.2 and the simulations are performed for different design parameters in Section 3.4.3.

3.4.1 Identification

The thickness and thermal process models have to be identified before implementing the GPC algorithm. For most industrial processes, it is sufficient to identify a first order model from the step response for the design of the GPC [62]. This is found to be adequate for our application as demonstrated by the experimental results. The process is also fairly linear such that a linear controller such as GPC can be used to control the resist thickness effectively. Resist thickness measurements at three sites were monitored and for each site, two experiments were conducted to identify the process models. The sampling interval was chosen to be 1 second. From this section onwards, to avoid confusion between the change in measurements and the absolute measurements, a subscript m will be used to denote the actual measurements from the in-situ sensors. For example, y_m denotes the resist thickness measurement obtained from the in-situ thickness sensor and T_m is the temperature measurement obtained from the embedded RTD in the bakeplate.

Figure 3.8 shows the identification results for Site 1. In the first experiment, the resist thickness is monitored while the resist is baked at nominal processing temperature of 90 °C. The resist thickness and temperature are denoted as y'_m and T'_m in Figures 3.8(a) and (b) respectively. For the second experiment, a step change in heater input, u, is made 25

seconds after the placement of the wafer on the bakeplate. The change in resist thickness is monitored at the same location on a second wafer. The resist thickness and temperature measurements are denoted as y''_m and T''_m in Figures 3.8(a) and (b) respectively. The change in resist thickness due to the step change in heater input is denoted as y such that $y = y''_m - y'_m$.

The change in bakeplate temperature due to step change in heater input is given as T where $T = T''_m - T'_m$. Figures 3.8 (c) and (d) shows the change in resist thickness and bakeplate temperature due to the step change in heater input respectively. Figure 3.8(e) shows the step change in heater input, u. The solid lines in (c) and (d) show the experimental values and the dashed lines show the calculated thickness and temperature from the identified first order plant models respectively.

Using least squares estimation, the first order plus dead time thickness models are identified for all the three sites such that

$$(1 + a_1 q^{-1}) y(k) = b_0 q^{-d} u(k-1) + \frac{e(k)}{(1 - q^{-1})}$$

where $d = 15$ for all the three sites. The values of a_1 and b_0 parameters for all the three sites are summarized in Table 3.1.

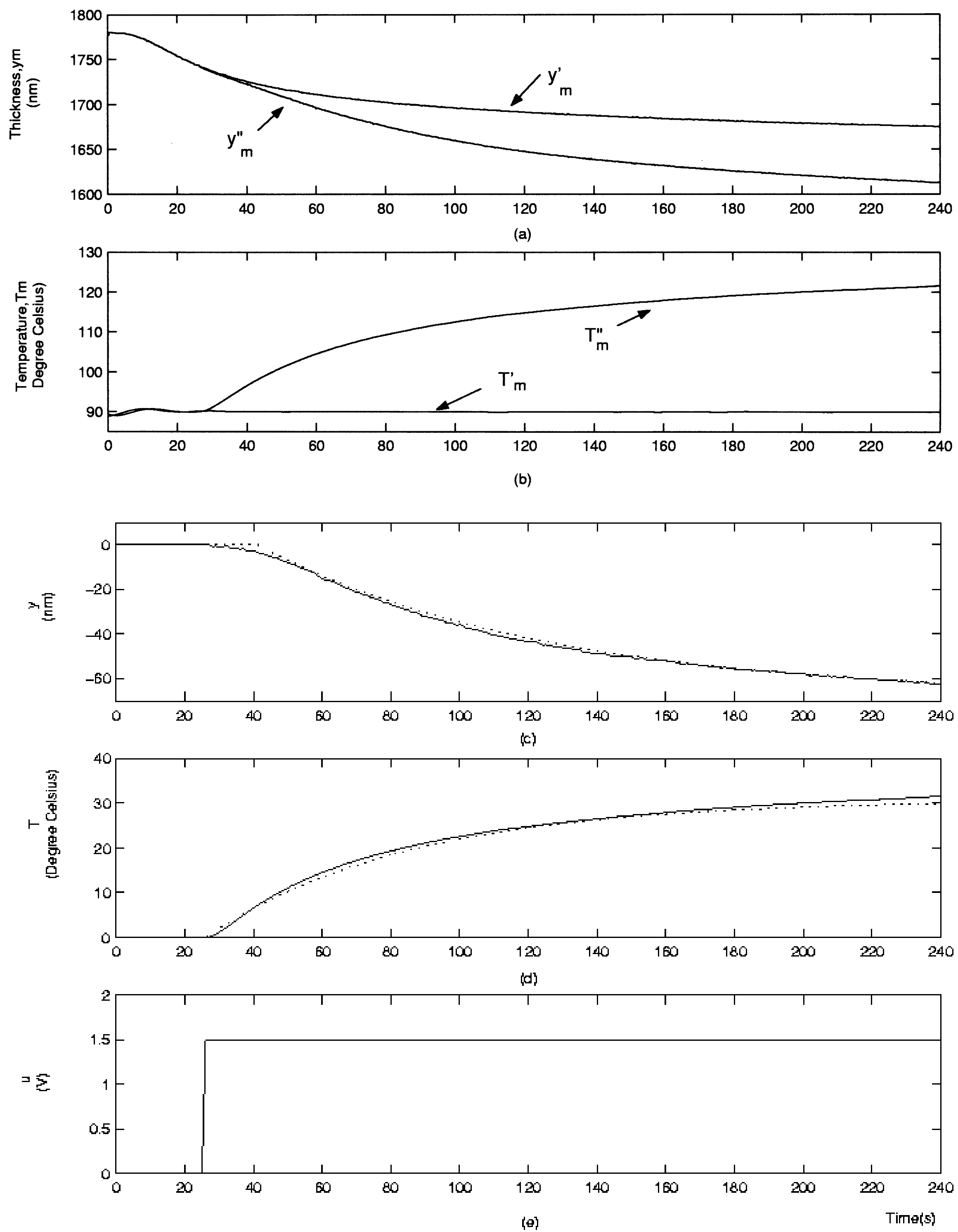


Figure 3.8: Identification experiments. Plots of (a) resist thickness, (b) temperature, (c) change in resist thickness where $y = y''_m - y'_m$, (d) change in temperature where $T = T''_m - T_m$, and (e) change in heater input, u, with respect to time.

Table 3.1: Parameters of a_1 and b_0 for all the three thickness models.

Position on the wafer	a_1	b_0
Site 1	-0.9877	-0.5263
Site 2	-0.9880	-0.5441
Site 3	-0.9882	-0.5252

Likewise, a first order thermal models relating the bakeplate temperatures and heater input voltage are identified at all the three sites such that

$$(1 + a_1 q^{-1})T(k) = b_0 q^{-d} u(k-1) + \frac{e(k)}{1 - q^{-1}}$$

where $d = 1$ for all the three sites. Table 3.2 summarizes the values of a_1 and b_0 parameters for the three thermal models. The values of a_1 and b_0 in Table 3.1 and 3.2 will be used to predict the resist thickness and bakeplate temperature.

Table 3.2: Parameters of a_1 and b_0 for all the three thermal models.

Position on the wafer	a_1	b_0
Site 1	-0.9832	0.3442
Site 2	-0.9820	0.3580
Site 3	-0.9813	0.3408

The discrete process models are converted into continuous models. The static gains and time constants of these continuous time process models are summarized in Table 3.3.

Table 3.3: Static gains, time constants and dead times of the process models.

Positions on the wafer	Thickness models			Thermal models			Sensitivity (nm/°C)
	Static gain (nm/V)	Time constants (s)	Dead Time (s)	Static gain (°C/V)	Time constants (s)	Dead time (s)	
Site 1	-42.7	82.1	15	20.5	59.0	1.0	-2.08
Site 2	-45.3	82.8	15	19.8	55.1	1.0	-2.28
Site 3	-44.5	84.2	15	18.2	53.0	1.0	-2.45

The last column in the Table 3.3 is obtained by dividing the values in 2nd column with the values in 5th column. It shows the sensitivity of the resist thickness to bakeplate temperature, which ranges from -2.08 to -2.45 nm/°C. The resist thickness sensitivity is used to determine the upper temperature bound, $T_{\max}$. In the next section, the control algorithm is discussed. The choice of the reference and the temperature bounds on the bakeplate will be discussed in greater details in Section 3.4.3.

3.4.2 Control Algorithm

GPC has been well discussed in the literature [62-64]. An overview of the GPC algorithm is also given in Appendix A. In this section, only the equations necessary for the implementation are stated.

GPC is a model-based predictive control algorithm that computes an optimal control sequence, $[\Delta u(k), \Delta u(k+1), \cdots, \Delta u(k+N-1)]$ by minimizing a quadratic objective function defined over a prediction horizon, N. The aim is for the predicted future resist thickness,

$\hat{y}_m(k+j\,|\,k)$, to follow the future reference thickness trajectory, $r(k+j)$, for $d+1 \le j \le N+d$ where d denotes the dead time of the thickness model and N is the prediction horizon. To impose the temperature constraints during the softbake process, the optimal control sequence is computed by minimizing the quadratic objective function, J, given as

$$J = \sum_{j=d+1}^{N+d} [\hat{y}_m(k+j\,|\,k) - r(k+j)]^2 + \sum_{j=1}^{N} \lambda[\Delta u(k+j-1)]^2 \tag{3.3}$$

Subject to the temperature constraint

$$T_{\min} \le \hat{T}(k+j\,|\,k) \le T_{\max} \qquad \text{for } d+1 \le j \le N+d$$

where Δu and λ are the change in control signal and control weighting respectively. The control signal in this application is the heater input voltage that is required to minimize the error between the predicted thickness, $\hat{y}_m(k+j\,|\,k)$, and the reference, $r(k+j)$. The objective function in Equation (3.3) also includes a term that penalizes the control effort, Δu, by the choice of the control weighting, λ. By choosing a small value for λ, more aggressive control moves will be made for the resist thickness to track the reference trajectory. For our implementation, the design parameter of λ = 1 is chosen. Given a relatively short softbake time of 1-5 minutes, a prediction horizon, N = 10, and control sampling of 1 s are chosen. A longer prediction horizon improves the accuracy of the computed control sequence but at the expense of higher computational time. For the

temperature constraint equation, $\hat{T}(k+j \mid k)$ is the optimum j-step ahead prediction of the change in temperature of the bakeplate from the nominal bake temperature of 90 °C based on temperature measurements up to sampling instant, k. The lower and upper bound on the change in temperature from the nominal temperature of 90 °C are given as $T_{\min}$ and $T_{\max}$ respectively.

From Appendix A, the prediction of the resist thickness can be expressed as

$$\hat{\mathbf{y}} = \mathbf{G}\mathbf{u} + \mathbf{f}$$

where

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}(k+d+1 \mid k) \\ \hat{y}(k+d+2 \mid k) \\ \vdots \\ \hat{y}(k+d+N \mid k) \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \\ \vdots \\ \Delta u(k+N-1) \end{bmatrix}, \mathbf{G} = \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ g_{N-1} & & \cdots & g_0 \end{bmatrix} \quad \text{and} \quad \mathbf{f} = \begin{bmatrix} f(k+d+1) \\ f(k+d+2) \\ \vdots \\ f(k+d+N) \end{bmatrix}$$

The matrix, $\mathbf{G}$, is step response coefficient matrix and $\mathbf{f}$ is the free response of the thickness model. Similarly, the prediction of bakeplate temperature, $\hat{\mathbf{T}}$, can be expressed as

$$\hat{\mathbf{T}} = \mathbf{G}'\mathbf{u} + \mathbf{f}'$$

where

$$\hat{\mathbf{T}} = \begin{bmatrix} \hat{T}(k+d+1 \mid k) \\ \hat{T}(k+d+2 \mid k) \\ \vdots \\ \hat{T}(k+d+N \mid k) \end{bmatrix}, \mathbf{G}' = \begin{bmatrix} g'_0 & 0 & \cdots & 0 \\ g'_1 & g'_0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ g'_{N-1} & & \cdots & g'_0 \end{bmatrix} \quad \text{and} \quad \mathbf{f}' = \begin{bmatrix} f'(k+d+1) \\ f'(k+d+2) \\ \vdots \\ f'(k+d+N) \end{bmatrix}$$

The matrix, $\mathbf{G}'$, is step response coefficient matrix and $\mathbf{f}'$ is the free response vector for the thermal model. Computation of $\mathbf{G}'$ and $\mathbf{f}'$ are similar to that of $\mathbf{G}$ and $\mathbf{f}$, except that the values of a_1, b_0 and d are obtained from the identified first order thermal models.

With the prediction models, the control sequence is computed through optimization of the objective function given in Equation (3.3). This requires the prediction of the resist thickness, $\hat{y}_m(k+j)$. If y_m is the resist thickness measurement obtained from the in-situ thickness sensor, the change in resist thickness, y, is obtained by subtracting the measured resist thickness from the open loop thickness, y'_m, i.e. $y = y_m - y'_m$. The open loop thickness trajectory, y'_m, is the in-situ resist thickness measurements obtained when the bakeplate temperature is maintained at 90 °C. In this case, there is no implementation of the real-time thickness control strategy. An example of the open thickness trajectory is given as y'_m in Figure 3.8(a). Hence, the prediction of the resist thickness, $\hat{y}_m(k+j)$, in Equation (3.3) can be written as

$$\hat{y}_m(k+j \mid k) = \hat{y}(k+j \mid k) + y'_m(k+j)$$

and the objective function in Equation (3.3) can be expressed as

$$J = \sum_{j=d+1}^{N+d} [\hat{y}(k+j \mid k) - y_d(k+j)]^2 + \sum_{j=1}^{N} \lambda [\Delta u(k+j-1)]^2 \tag{3.4}$$

where $y_d(k+j) = r(k+j) - y'_m(k+j)$. This is subjected to the temperature constraint,

$$T_{\min} \le \hat{T}(k+j \mid k) \le T_{\max} \qquad \text{for } d+1 \le j \le N+d$$

Rewriting Equation (3.4), the optimal control sequence is obtained by minimizing the objective function

$$J = (\mathbf{Gu} + \mathbf{f} - \mathbf{r})^T (\mathbf{Gu} + \mathbf{f} - \mathbf{r}) + \lambda \mathbf{u}^T \mathbf{u} \tag{3.5}$$

Subject to the temperature constraints

$$\mathbf{1_N} T_{\min} \le \mathbf{G'u} + \mathbf{f'} \le \mathbf{1_N} T_{\max}$$

where the vector $\mathbf{1_N}$ is a $N \times 1$ column vector with all its elements equal to 1. The inclusion of the temperature constraints in the GPC algorithm ensures that the bakeplate temperature is constrained to an upper temperature bound of $(90 + T_{\max})$ °C and a lower temperature bound of $(90 + T_{\min})$ °C during real-time thickness control. The minimization of Equation (3.5) can be solved using a commercial QP solver by formulating it as a standard Quadratic Programming (QP) problem, that is minimizing the objective function:

$$J = \frac{1}{2} \mathbf{u}^T \mathbf{H} \mathbf{u} + \mathbf{b}^T \mathbf{u} + \mathbf{f}_0$$

Subject to the constraint

$$\mathbf{Ru} < \mathbf{c}$$

where

$$\mathbf{H} = 2(\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I}) \qquad \mathbf{R} = \begin{bmatrix} \mathbf{G}' \\ -\mathbf{G}' \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} \mathbf{1_N} T_{\max} - \mathbf{f}' \\ -\mathbf{1_N} T_{\min} + \mathbf{f}' \end{bmatrix}$$

$$\mathbf{b}^T = 2(\mathbf{f} - \mathbf{r})^T \mathbf{G} \qquad \mathbf{u} = \left[\Delta u(k) \;\; \Delta u(k+1) \;\; \cdots \;\; \Delta u(k+N-1)\right]^T$$

$$\mathbf{f}_0 = (\mathbf{f} - \mathbf{r})^T (\mathbf{f} - \mathbf{r}) \qquad \mathbf{r} = \left[y_d(k+d+1) \;\; y_d(k+d+2) \;\; \cdots \;\; y_d(k+d+N)\right]^T$$

3.4.3 Choice of Design Parameters

For the implementation of the GPC algorithm, it is necessary to specify some of the important design parameters. In this section, the selection of these design parameters is discussed.

3.4.3.1 Reference Trajectory

During softbake, resist thickness is monitored and the heater power distribution is computed every 1 second. Although resist thickness can be estimated at a sampling rate of 0.1 s, a sampling interval of 1 s is used to allow for additional time needed to compute the heater power distribution using the GPC algorithm. During the first few seconds, the resist thickness increases before it decreases. Similar phenomenon has been observed in Morton [29]. It is believed to be due to resist film softening at the glass transition temperature. After the first few seconds, the resist thickness decreased due to the onset of solvent evaporation.

To avoid this initial peak in the resist thickness, the first control move is introduced at time, $t = 25$ s or at sampling instant, $k = 25$, after the resist thickness has started to decrease.

The choice of reference trajectory is important for the implementation of the GPC algorithm. In conventional softbake process where real-time thickness control strategy is not implemented, the wafer is baked at a constant temperature of 90 °C. The monitored resist thickness measurement during this open loop run is stored as the open loop thickness trajectory, y'_m. An example of the open loop thickness trajectory is shown in Figure 3.8 (a). To obtained the reference trajectory, the open loop thickness trajectory, y'_m, is shifted 10 nm down at the site where the resist thickness is maximum at the start of the first control move, i.e. $r(k) = y'_m(k) - 10$ for $k \geq 25$. Hence at the site where the resist thickness is maximum, the reference, y_d, is a negative step such that $y_d = -10$.

3.4.3.2 Temperature Limits

As discussed in Section 3.4.1, the resist sensitivity to temperature that is given in Table 3.3 can be used to determine the upper temperature limit, $T_{\max}$, in Equation (3.3). Table 3.3 shows that the resist sensitivity ranges from -2.08 to -2.45 nm/°C. To ensure that the resist thickness will track the reference thickness at any point on the wafer, the lowest resist sensitivity of -2.08 nm /°C is used in the selection of $T_{\max}$. When the resist sensitivity is -2.08 nm /°C, the bake temperature has to be increased by 4.8 °C in order for the resist thickness to decrease by 10 nm at steady state. Therefore, the upper temperature constraint must be greater than 4.8 °C, i.e. $T_{\max} \geq 4.8$ to ensure that the resist thickness can track the

reference trajectory. To allow for some inaccuracy in the process models, temperature is constrained to an upper temperature bound of 96 °C ($T_{\max}$ = 6 °C instead of 4.8 °C). The bake temperature is constrained to a lower temperature bound of 86 °C (or $T_{\min} = -4\,^{o}C$). By allowing the bakeplate temperature to cool at a temperature lower than the nominal bakeplate temperature of 90 °C, the resist thickness, y_m, can decrease at a much slower rate than the reference thickness trajectory, r.

3.4.3.3 Prediction Horizon and Control Weighting

The prediction horizon, N, and control weighting, λ, are other important design parameters that have to specified in the GPC algorithm. The choice of the control weighting, λ, determines the aggressiveness of the control effort. Smaller value of λ will result in more aggressive control actions while a larger value of λ means less aggressive control actions. Some simulations are performed to determine the choice of λ. In simulation, the first control move is made at sampling instant, $k = 25$ or at time, $t = 25$ s for a sampling period of 1 s. The reference, y_d, is chosen to be a negative step with magnitude of 10 nm, i.e. $y_d(k) = -10$ for $k \geq 25$. The prediction horizon, N is chosen so that N =10. As discussed in the previous section, temperature constraints of $T_{\min} = -4$ and $T_{\max} = 6$ are chosen.

Figure 3.9 shows the simulation results for $N = 10$ and three different values of λ. In this simulation, y is the change in resist thickness and u is the control signal that is

computed using the GPC algorithm to minimize the error between the simulated change in resist thickness and reference thickness. The change in bakeplate temperature due to the change in heater input, u, is denoted as T. The simulated changes in resist thickness, heater input voltage and temperature when $\lambda = 0.1, 1 \text{ and } 10$ are given by the dashed-dot, solid and dashed lines in Figures 3.9 (a)-(c) respectively.

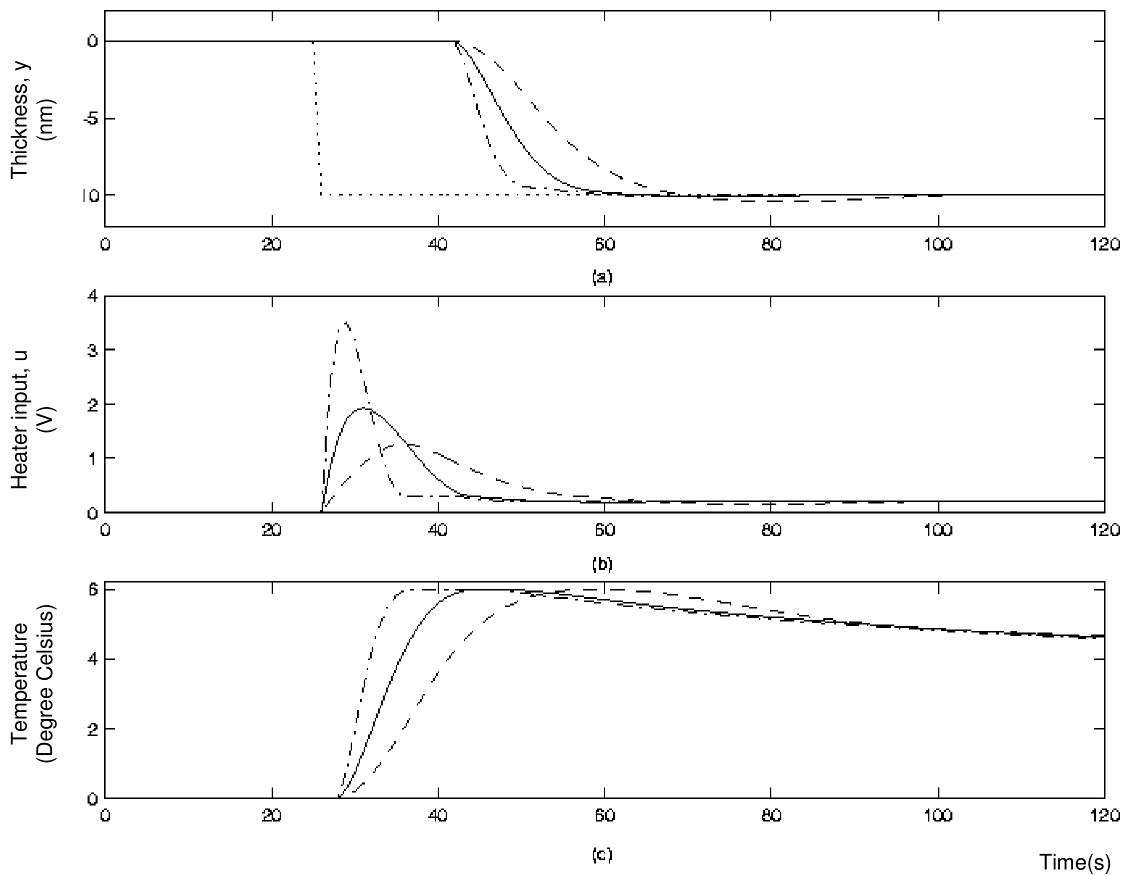


Figure 3.9: Simulation results for different values of lambda, λ

For $\lambda = 1$, the heater input in Figure 3.9 (b) increases rapidly at time, $t = 25$ s, to a maximum heater voltage of 1.9 V. With this set of parameters, the resist thickness tracks the

reference after about 65 seconds, with hardly any overshoot. It is also noted in Figure 3.9 (c) that the bakeplate temperature, T, is saturated at the upper temperature bound for only 4 s. When $\lambda = 0.1$, the performance of the tracking of the reference trajectory is faster than that when $\lambda = 1$. Due to more aggressive control actions, the resist thickness tracks the reference at 62 seconds. However, the maximum heater input also higher at 3.6 V and the bakeplate temperature, T, reaches the upper temperature bound earlier and remains at the upper temperature bound for a longer time. For $\lambda = 10$, the control actions are not aggressive enough, resulting in the resist thickness to fall below the reference. This is illustrated in Figure 3.9 (a). Comparing the simulation results for three sets of parameter, $\lambda = 1$ and $N = 10$ is preferred. This is because it has better tracking performance than $\lambda = 10$. At the same time, the bakeplate temperature when $\lambda = 1$ does not saturate at the upper temperature bound for as long as when $\lambda = 0.1$.

Figure 3.10 shows the simulation results for two sites, Site 1 and 2. The solid and dashed lines in Figures 3.10(a)-(f) give the simulated values for Site 1 and 2 respectively. As discussed in Section 3.4.3.1, the reference is obtained by shifting the open loop thickness trajectory downward by 10 nm. In this simulation, the resist thickness at Site 1 is taken to be maximum while the resist thickness at Site 2 is minimum at the start of the first control move. If the initial resist thickness non-uniformity (difference between the maximum and minimum resist thickness among the monitored sites) is 7.4 nm, the reference, y_d, for Site 1 and 2 are $y_d = -10$ and $y_d = -2.6$ respectively. Figures 3.10 (a), (b) and (c) show the change in resist thickness, heater input and temperature respectively. The absolute resist thickness, heater input and bakeplate temperatures are given in Figures 3.10(d), (e) and (f) respectively. The steady-state heater input voltage needed to maintain the bakeplate

temperature at 90 °c is assumed to be 1 V such that the heater input, u_m, is given as $u_m = u + 1$. Note that although the upper temperature bound of $T_{\max} = 6$ °C, is specified for both sites, only the bakeplate temperature at Site 1 reaches the upper temperature bound, as shown in Figures 3.10 (c) and (f).

3.5 Experimental Results

Eleven experimental runs were conducted. In all the experiments, commercial *i*-line resist, Shipley 3612 was spin coated at 2000 rpm onto a 4-inch wafer prior to the softbake process. Resist thickness measurements at three sites on the wafer were monitored and the target thickness of 1665 nm was chosen. Once the target thickness is reached, the wafer was removed from the bakeplate and chilled. Three conventional softbake runs were conducted where the bakeplate was maintained at 90 °C. Eight runs were conducted where heater power distribution of the bakeplate was manipulated using the GPC algorithm to reduce the resist thickness non-uniformity. Temperature constraints of $T_{\max} = 6$ °C and $T_{\min} = -4$ °C were imposed. Section 3.5.1 shows experimental result for a conventional softbake run while Section 3.5.2 gives the result with GPC control and summarizes the results for all eleven experimental runs.

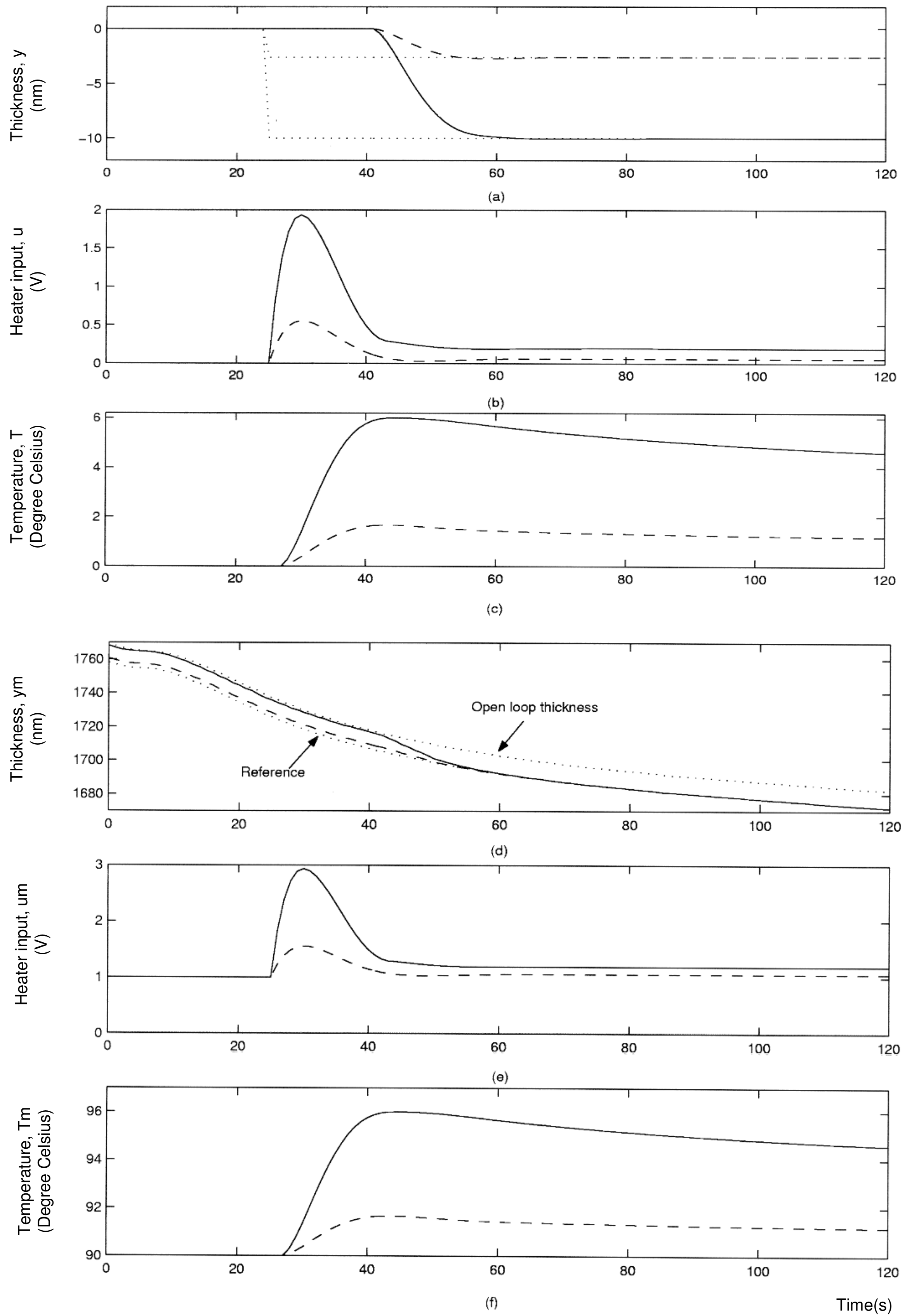


Figure 3.10: Simulation results for two sites

3.5.1 Conventional Softbake

For conventional softbake, temperatures at three sites were maintained at 90 °C. Throughout the softbake process, the resist thickness non-uniformity was measured. Figure 3.11 shows the experimental result of one of the conventional softbake runs.

Figure 3.11 (a) shows the change in resist thickness during the bake process. During the first few seconds, the resist thickness increased before it decreased due to the onset of solvent evaporation. Figures 3.11(b) and 3.11(c) are the bakeplate temperature and heater inputs for all the monitored sites respectively. The resist thickness non-uniformity is shown in Figure 3.11(d). The initial resist thickness non-uniformity was 4.8 nm and final thickness non-uniformity at the chosen target thickness was 4.9 nm. Therefore, resist thickness non-uniformity remains at the end of the conventional softbake process if a non-uniform resist film is formed during the prior coating process.

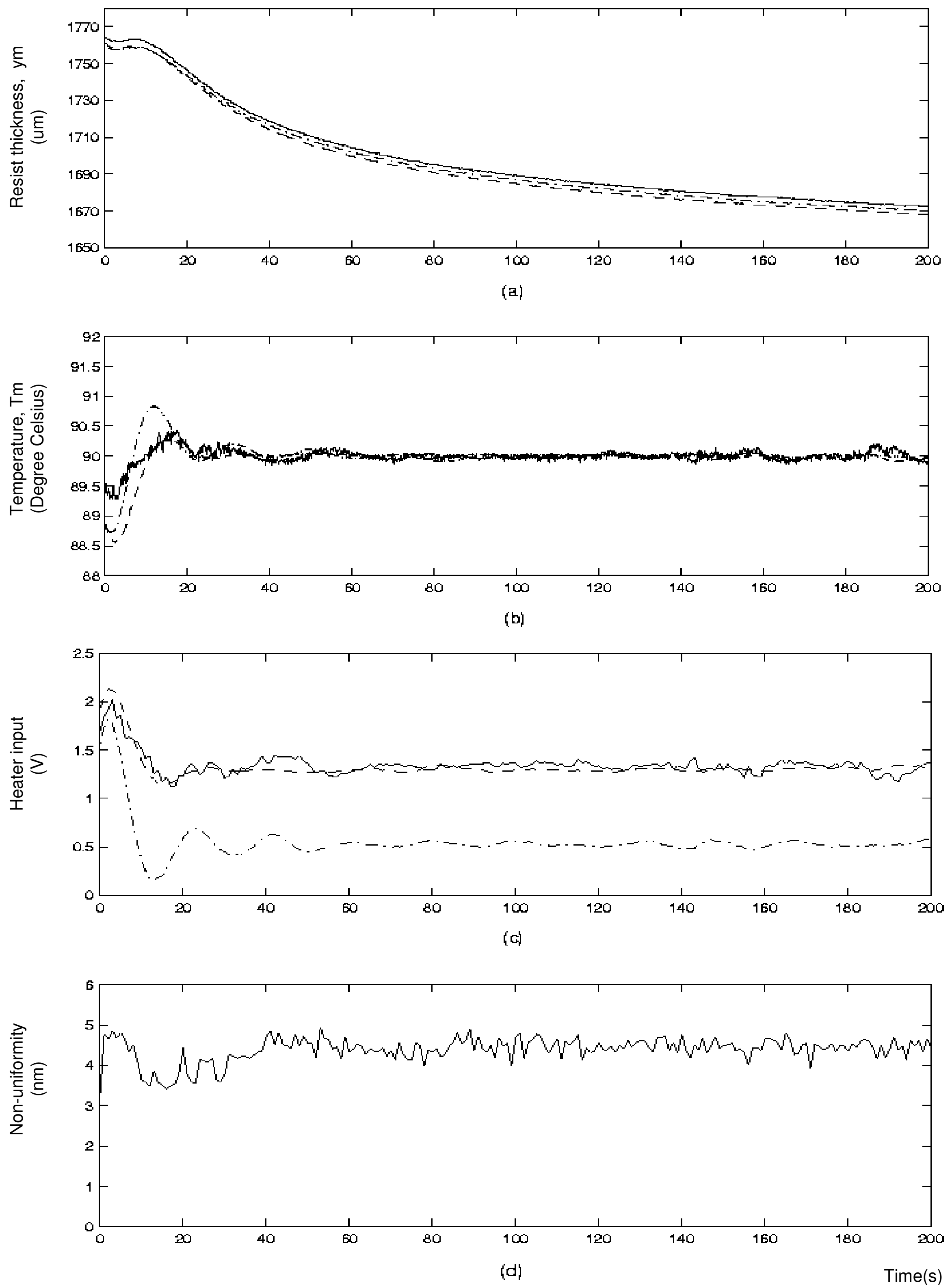


Figure 3.11: Conventional softbake with bakeplate maintained uniformly at 90 °C: (a) thickness, (b) temperature, (c) heater input, and (d) resist thickness non-uniformity profile of the three sites monitored. Sites 1-3 are represented by the solid, dashed-dotted, and dashed lines, respectively.

3.5.2 Real-time Thickness Control with GPC Algorithm.

In this section, GPC with temperature constraints is investigated. Eight experimental runs were conducted using the GPC algorithm. The result of one of the runs using the constrained GPC algorithm is shown in Figure 3.12. Resist thickness at three sites were monitored and made to follow a predefined reference trajectory at the sampling rate of 1 second. The monitored resist thickness and reference trajectory is shown in Figure 3.12(a). Figures 3.12 (b) and (c) show the corresponding bakeplate temperature and heater input respectively. The resist thickness non-uniformity between the three sites was shown in Figure 3.12 (d). As shown in Figure 3.12 (d), the resist thickness at three sites began with an initial thickness non-uniformity of 7.4 nm and converged to the specified target thickness after 155 seconds. Final thickness non-uniformity was 0.9 nm. While the wafer was supposed to be removed from the bakeplate after the resist thickness reached the target thickness, the experiment was allowed to continue in this example for illustration purposes. From Figure 3.12 (d), it can be shown that resist thickness non-uniformity of less than 1 nm was achieved after 100 s.

Figure 3.13 summarizes the results of all the eleven experimental runs. Runs #1−3 are the conventional softbake runs where the bakeplate temperature is maintained at 90 °C. Runs #4−11 are the experiments where heater power distribution of the bakeplate is manipulated using the GPC algorithm to reduce resist thickness non-uniformity. The diamond markers in Figure 3.13 show the initial thickness non-uniformity and the square markers show the thickness non-uniformity at the end of the softbake process. The improvement of the thickness uniformity can be determined as

$$\text{Improvement in thickness uniformity} = \frac{\text{Initial thickness non - uniformity}}{\text{Final thickness non - uniformity}}$$

As shown in Figure 3.13, Runs #1−3 give no significant change in the resist thickness non-uniformity at the end of the softbake process. Runs #4−11 shows an average 9.8 times improvements in resist thickness uniformity.

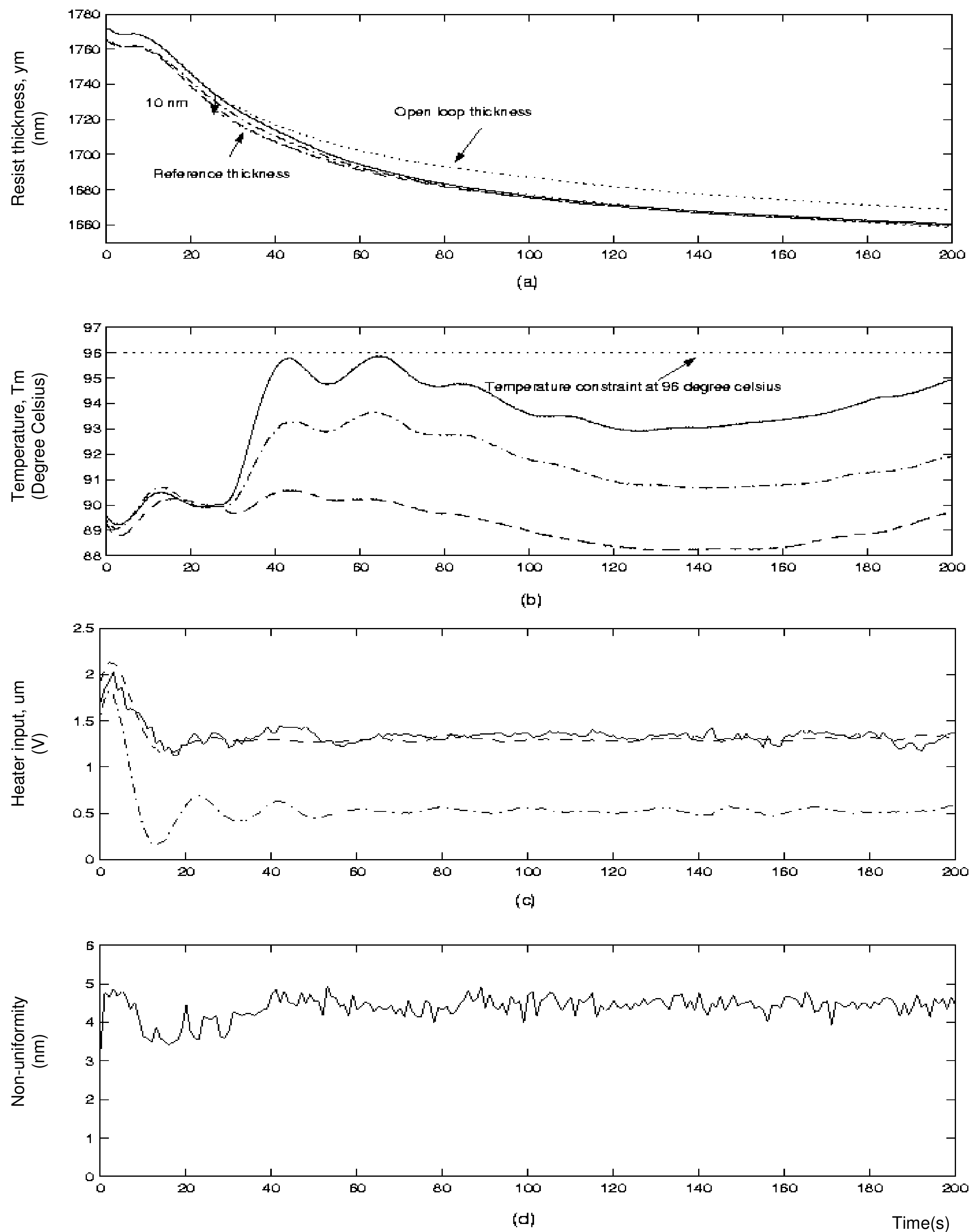


Figure 3.12: GPC Control with temperature constraints: (a) thickness, (b) temperature, (c) heater input, and (d) thickness non-uniformity profile when three sites on the wafer are monitored. Sites 1-3 are represented by the solid, dashed-dotted, and dashed lines respectively. The reference thickness trajectory is given by the dotted line.

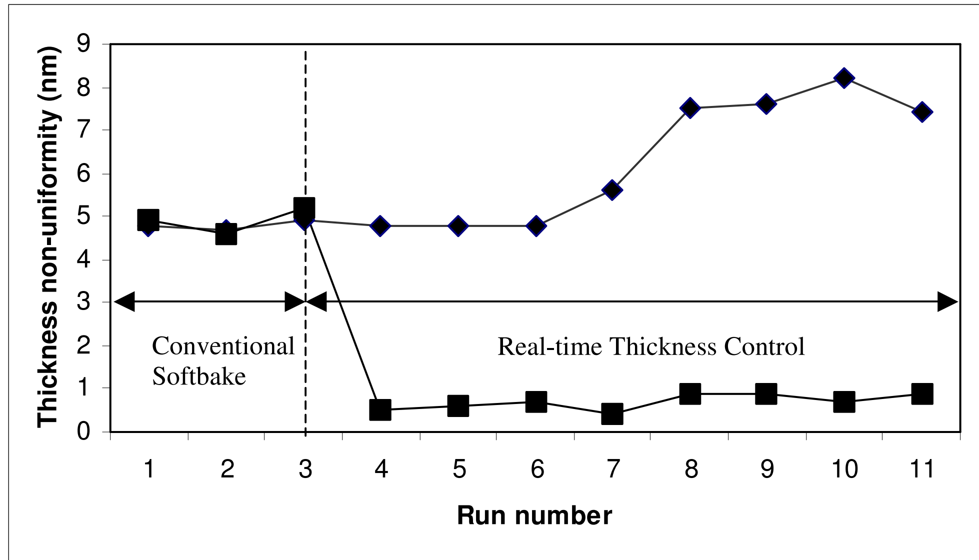


Figure 3.13: Summary of the experimental runs.

3.6 Summary

Real-time thickness control has been implemented using an array of in-situ thickness sensors and predictive control strategy. It has been demonstrated that by maintaining a non-uniform temperature profile through manipulation of the heater power distribution, an average of 10 times improvement in resist thickness uniformity is obtained from wafer-to-wafer (run-to-run) and across individual wafer. Repeatable reduction in resist thickness non-uniformity to less than 1nm at a given target thickness have been obtained. The general control strategy may also be extended to similar applications where stringent film thickness uniformity is required.

Chapter 4

Implementation of Real-time Thickness Control Using Sliding Mode Control

4.1 Motivations

In Chapter 3, it has been demonstrated that the implementation of real-time thickness control strategy during softbake process reduces the resist thickness non-uniformity across the wafer to less than 1 nm at the end of the bake process. The thickness control strategy is implemented using a multi-zones bakeplate, in-situ thickness sensors and the Generalized Predictive Control (GPC) algorithm. During the bake process, the GPC algorithm uses the in-situ resist thickness measurements to compute the heater power distribution of the bakeplate that is required to minimize the resist thickness non-uniformity. To prevent decomposition of the photoactive compound in the resist film, constraints are imposed on the bakeplate temperature during softbake. This results in a constrained GPC control problem which can be formulated as a Quadratic Programming (QP) problem where the optimization is solved using a commercial QP solver. This approach has some disadvantages when the number of controlled sites increases. For instance, if resist thickness at 49 sites have to be controlled during the bake process, tremendous time and effort will be required to identify 49 different process models. The computational time to solve the constrained GPC problem may be excessive. Furthermore, a commercial QP solver is required for the computation of

the heater power distribution. A simpler and faster implementation of the thickness control strategy is investigated in this Chapter. The experimental setup is the same as that in Chapter 3, except that a sliding mode control algorithm is used to compute the heater power distribution.

Sliding mode control is a simple approach to robust control that achieves a certain performance level in the presence of model imprecision. In the formulation of any control problem, there will always exist some discrepancies between the actual plant and the identified mathematical model that is used to represent the plant. These model inaccuracies may arise from unmodelled dynamics from the sensors or actuators (which usually have a much faster dynamics and hence neglected during the modeling process) or as a result of approximating the complex plant behavior with a simplified mathematical model [65, 67]. In this application, a nominal process model is identified and used for all the controlled sites. Any discrepancies between the process models at different sites are treated as modeling inaccuracies. To deal with the modeling inaccuracies, the sliding mode controller is characterized by a suite of feedback control laws and a decision rule. The decision rule, termed as the switching function, uses some measure of current system behavior to determine which particular feedback controller is to be used at that time instant. An ideal sliding mode controller switches between different control laws at an infinitely fast speed. Using this approach, the closed loop response becomes totally insensitive to a particular class of uncertainty.

In recent years, there have been some researches on the application of the intelligent control techniques such as fuzzy control and neural network in semiconductor manufacturing [12-14]. These techniques rely on the learning, in a prescribed manner, the input-output

behavior of the plant to be controlled and are usually applied in areas where the processes are ill-defined, complex, non-linear, time varying and stochastic. In cases where model-based approaches yield high performance control, the value gained from applying such techniques may not be significant [65]. As demonstrated in Chapter 3, model-based predictive control algorithm such as GPC is capable of achieving good thickness control to less than 1 nm. For this application, sliding mode control is chosen for its ability to achieve a good performance in the presence of arbitrary model inaccuracies and minimum implementation complexity [65].

This Chapter is organized as follows. Section 4.2 discusses the design and implementation of the sliding mode control in greater details. The equations necessary for implementing the sliding mode controller are also derived in Section 4.2. A comparison is made between the thickness control strategy using the sliding mode control design and the GPC algorithm in Section 4.3. The experimental results of the thickness control strategy using the sliding mode control algorithm are discussed in Section 4.4 and a conclusion is given in Section 4.5.

4.2 Control Structure

In this Section, the control structure of the thickness control strategy using a sliding mode controller is discussed. Figure 4.1 shows the control structure.

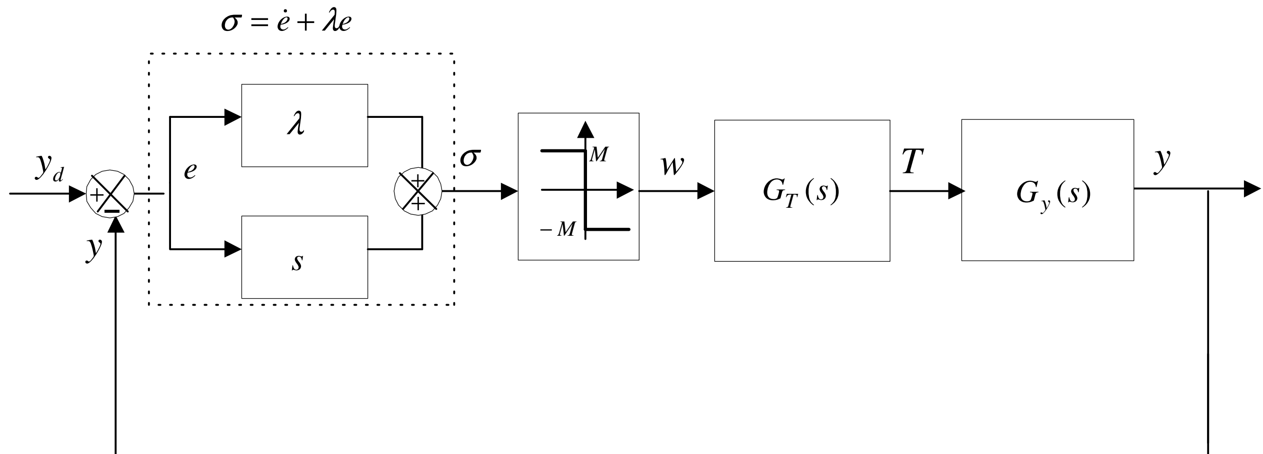


Figure 4.1: Control structure of the thickness control strategy using sliding mode control.

As given in Chapter 3, y is the change in resist thickness from the open loop thickness trajectory and y_d is a negative step change from the open loop thickness trajectory such that $y_d = -Y_d$ where Y_d is the magnitude of the step change. The open loop thickness trajectory is obtained when the resist is baked at a constant temperature of 90 °C. In Figure 4.1, e denotes the resist thickness error given as $e = y_d - y$. The switching function, σ, is a function of the thickness error such that $\sigma = \dot{e} + \lambda e$ where $\dot{e}$ is the time derivative of the thickness error and λ is a design parameter. Based on the switching function, the sliding mode controller switches between two control values such that the control signal, w, takes the values of only M or $-M$. The output of the sliding mode controller, w, is then used to determine the temperature setpoint of the bakeplate that is required to reduce the thickness error.

The process model relating the change in resist thickness, y, to the change in bakeplate temperature, T, is given as $G_y(s)$ while the transfer function relating the change in bakeplate temperature, T, and temperature setpoint, w, is given as $G_T(s)$. If $G_T(s)$ is a first order model with a unity steady state gain, the bakeplate temperature, T, will always be less than or equal to its setpoint, w. Since the output of a sliding mode controller is bounded between the upper and lower bounds of $\pm M$, the bakeplate temperature is also bounded to $\pm M$ °C from the nominal bake temperature of 90 °C. The temperature constraints required in this application can thus be implemented.

Typically, a feedback controller such as a PID controller is used to make the bakeplate temperature follow the temperature setpoint. This results in a cascaded control loop structure, given in Figure 4.2. The switching function, σ, given in Figure 4.1, involves differentiating the thickness error, e. In practice, a pure derivative is not implemented because it will result in a large amplification of measurement noise. The pure differentiator, s, is thus replaced by the term, $\frac{s}{s\tau_f + 1}$, to filter out the measurement noise in Figure 4.2

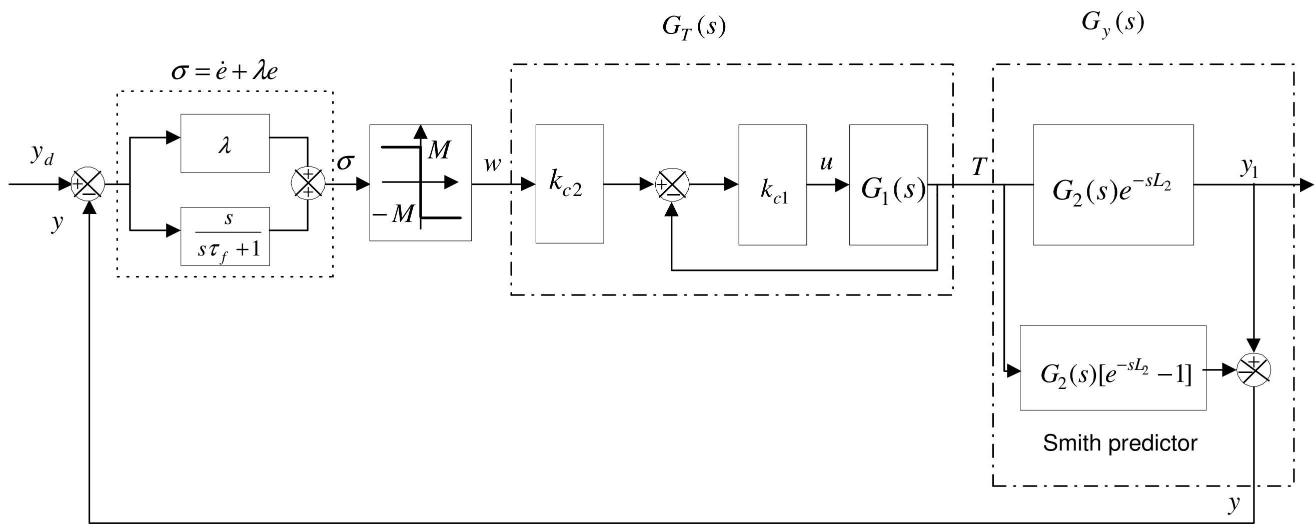


Figure 4.2: Cascaded control loop structure.

From Chapter 3, it is noted that there is a delay before the resist thickness responds to any changes in the bakeplate temperature. Therefore, a first order plus dead-time thickness model is assumed such that

$$\frac{Y_1(s)}{T(s)} = \frac{k_2}{s\tau_2 + 1} e^{-sL_2} = G_2(s)e^{-sL_2}$$

where k_2 is the static gain, τ_2 is the time constant and L_2 is the dead time of the thickness model. The change in resist thickness which occurs only L_2 seconds after the change in bakeplate temperature is denoted as y_1. Note that the static gain, k_2, is negative as an increase in bakeplate temperature will result in a decrease in the resist thickness.

To provide the time-delay compensation, a smith predictor [66] is included such that

$$G_y(s) = \frac{Y(s)}{T(s)} = \frac{k_2}{s\tau_2 + 1} \tag{4.1}$$

where y is the output of the smith predictor, which gives the predicted change in resist thickness without the time delay. For a control sampling of 0.1s, $y(k) = y_1(k + 10L_2)$ where k is the sampling index. In designing the sliding mode controller, complete cancellation of the dead time by the smith predictor is assumed.

The thermal model of the bakeplate relating the change in bakeplate temperature, T, and the heater input, u, is denoted as $G_1(s)$ in Figure 4.2. A first order process model is assumed such that

$$G_1(s) = \frac{T(s)}{U(s)} = \frac{k_1}{s\tau_1 + 1}$$

where k_1 is the static gain and τ_1 is the time constant. A proportional controller, k_{c1}, is used to control the temperature of the bakeplate. Another proportional gain k_{c2} is included so the transfer function, $G_T(s)$, has a unity steady state gain given by

$$G_T(s) = \frac{T(s)}{W(s)} = \frac{1}{s\tau_3 + 1} \tag{4.2}$$

where τ_3 is the time constant of the process model, $G_T(s)$.

Before implementing the sliding mode control design, the process models, $G_1(s)$ and $G_2(s)e^{-sL_2}$, have to be identified first. Letting $\hat{k}_1$ and $\hat{\tau}_1$ be the least squares estimates of

k_1 and τ_1 obtained in the identification experiment for the process model of $G_1(s)$, the time constant, τ_3, and the proportional gain, k_{c2}, are given as $\tau_3 = \dfrac{\hat{\tau}_1}{1+k_{c1}\hat{k}_1}$ and $k_{c2} = \dfrac{1+k_{c1}\hat{k}_1}{k_{c1}\hat{k}_1}$ respectively. The identification experiment will be discussed in greater details in the next section.

4.2.1 Identification

Similar to the identification experiment in Chapter 3, two separate experiments were conducted to identify the process models, $G_1(s)$ and $G_2(s)$, in Figure 4.2. In the first experiment, the resist was baked at the nominal processing temperature of 90 °C (T'_m, Figure 4.3b) and the resist thickness (y'_m, Figure 4.3a) was monitored at a sampling interval of 0.1 s. In the second experiment, a 0.35 V step increase in heater input, u, was made 15 s into the bake process. Measurements of the resist thickness, y''_m, and bakeplate temperature, T''_m, were taken as shown in Figures 4.3(a) and (b) respectively. The resulting changes in the resist thickness and temperature measurements due to the change in the heater input, u, are denoted as y_1 and T respectively such that $y_1 = y''_m - y'_m$ and $T = T''_m - T'_m$. Figures 4.3 (c) and (d) show the decrease in resist thickness, y_1, and the increase in bakeplate temperature, T, due to the step change in heater input, u respectively. The change in heater input, u, is shown in Figure 4.3(e).

Two discrete-time first order models were identified and converted to the continuous time thermal model, $G_1(s)$, and thickness model, $G_2(s)e^{-sL_2}$. The estimated parameters of the thermal model, k_1 and τ_1, are given as $\hat{k}_1 = 20 \ ^oC/V$ and $\hat{\tau}_1 = 33.3$ s respectively.

For the thickness model relating T and y_1, the estimated parameters, k_2, τ_2 and L_2 are $\hat{k}_2 = -2.0 \ nm/^oC$, $\hat{\tau}_2 = 30.3$ s and $\hat{L}_2 = 5$ s respectively. Figures 4.3(c) and (d) show the comparison of the change in resist thickness and temperature calculated using the identified models with the experimental values. Once the process models have been identified, the sliding mode control algorithm can be implemented. In the next section, the necessary equations needed for implementation are discussed.

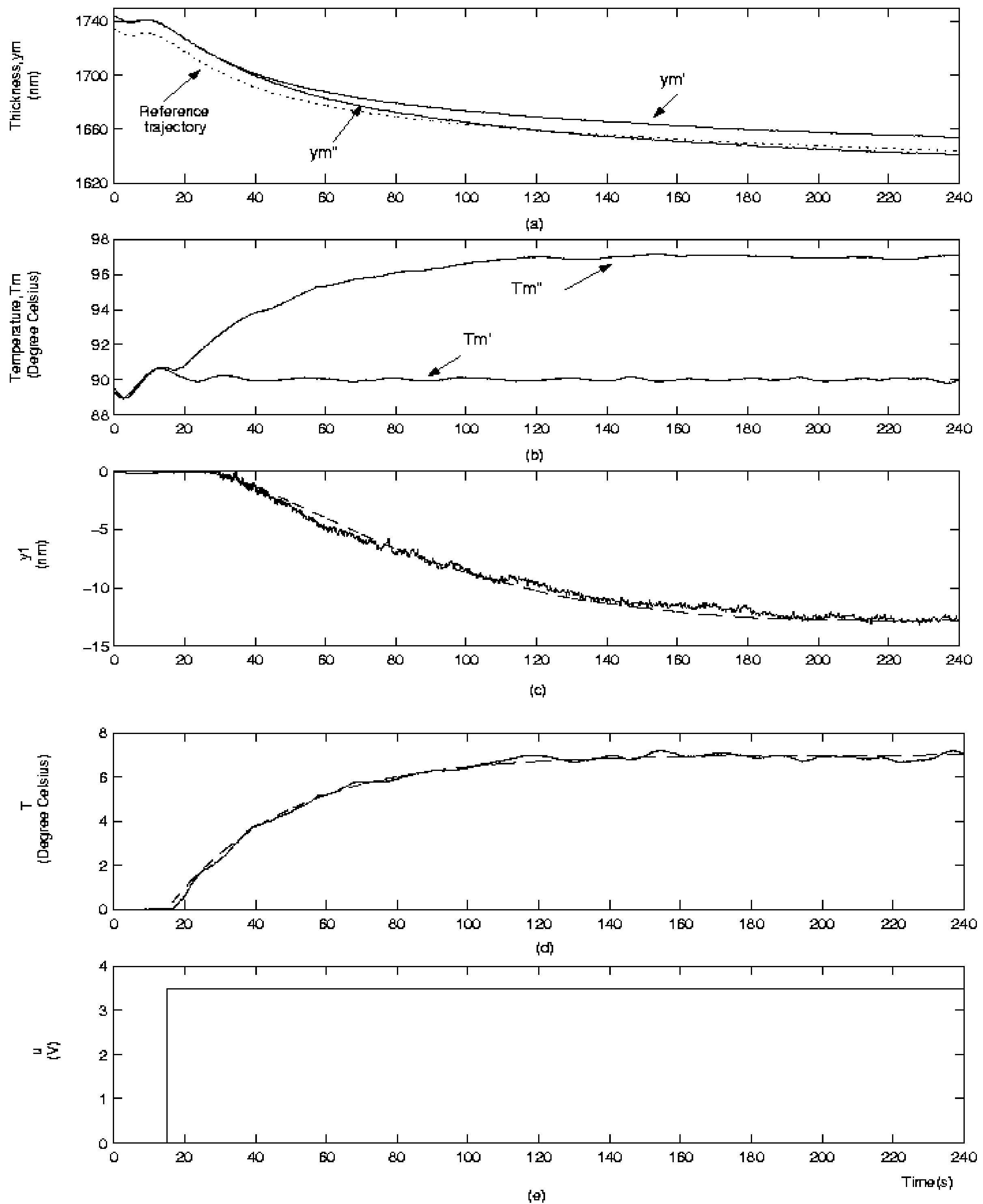


Figure 4.3: Identification experiments. Plots of (a) resist thickness, (b) temperature, (c) decrease in resist thickness where $y_1 = y''_m - y'_m$, (d) increase in bakeplate temperature where $T = T''_m - T'_m$, and (e) change in heater input, u, with respect to time. The solid lines in (c) and (d) show experimental values and the dashed lines show the calculated responses of the identified models.

4.2.2 Sliding Mode Control Algorithm

In this section, the temperature limit, M, used in the sliding mode control algorithm is derived. From Equation (4.1) and (4.2), the process model relating the change in resist thickness to the temperature setpoint is given as

$$\frac{Y(s)}{W(s)} = G_y(s)G_T(s) = \left(\frac{k_2}{s\tau_2 + 1}\right)\left(\frac{1}{s\tau_3 + 1}\right)$$

In time domain, this process model can be expressed as

$$\ddot{y}(t) = \frac{k_2}{\tau_2\tau_3} w(t) - \left(\frac{1}{\tau_2} + \frac{1}{\tau_3}\right)\dot{y}(t) - \left(\frac{1}{\tau_2\tau_3}\right)y(t) \tag{4.3}$$

The switching function or sliding variable, σ, is defined as

$$\sigma(t) = \dot{e}(t) + \lambda e(t) \tag{4.4}$$

where $e(t) = y_d(t) - y(t)$ is the thickness error. The reference is denoted as $y_d(t)$ and λ is a positive constant and a design parameter. The associated sliding mode controller is chosen as

$$w(t) = -M\text{sign}\{\sigma(t)\} \tag{4.5}$$

where the signum function, $\text{sign}\,\{\sigma(t)\}$, is defined as

$$\text{sign}\,\{\sigma(t)\} = \begin{cases} 1 & \text{for } \sigma(t) > 0 \\ -1 & \text{for } \sigma(t) < 0 \end{cases}$$

The purpose of the control action is to ensure that the error trajectory, $e(t)$, is driven towards the sliding surface, $\sigma(t) = 0$, within a finite time, t_σ and remain there for $t \geq t_\sigma$ in order to generate a sliding motion. The reaching time, t_σ, is time when the sliding surface ($\sigma(t) = 0$) is reached for the first time. It gives a conservative estimate of the maximum time necessary to reach the sliding surface [67].

The stability of the closed loop system and the tracking of the reference, $y_d(t)$, can be analyzed by examining the Lyapunov function

$$V(t) = -\frac{\tau_2\tau_3}{2k_2}\sigma^2(t) \qquad \text{for } k_2 < 0 \tag{4.6}$$

Differentiating Equation (4.4) and substituting Equation (4.3) gives

$$\dot{\sigma}(t) = \ddot{y}_d(t) + \lambda\dot{y}_d(t) + \left(\frac{1}{\tau_2} + \frac{1}{\tau_3} - \lambda\right)\dot{y}(t) + \frac{1}{\tau_2\tau_3}y(t) - \frac{k_2}{\tau_2\tau_3}w(t) \tag{4.7}$$

Differentiating Equation (4.6) and substituting Equations (4.5) and (4.7) gives

$$\dot{V}(t) = -\frac{\tau_2\tau_3}{k_2}\sigma(t)\left\{\ddot{y}_d(t) + \lambda\dot{y}_d(t) + \left(\frac{1}{\tau_2} + \frac{1}{\tau_3} - \lambda\right)\dot{y}(t) + \frac{1}{\tau_2\tau_3}y(t)\right\} - M\sigma(t)sign\{\sigma(t)\}$$

Since $\sigma(t)sign\{\sigma(t)\} = |\sigma(t)|$ and $k_2 = -|k_2|$, $\dot{V}(t)$ can be rewritten as

$$\dot{V}(t) = \frac{\tau_2\tau_3}{|k_2|}\sigma(t)\left\{\ddot{y}_d(t) + \lambda\dot{y}_d(t) + \left(\frac{1}{\tau_2} + \frac{1}{\tau_3} - \lambda\right)\dot{y}(t) + \frac{1}{\tau_2\tau_3}y(t)\right\} - M|\sigma(t)| \tag{4.8}$$

To enforce sliding mode for $t > 0$, M should be selected such that $\dot{V} < 0$. From Equation (4.8), M was chosen such that

$$M > \left|\frac{\tau_2\tau_3}{k_2}\left(\ddot{y}_d(t) + \lambda\dot{y}_d(t) + \left(\frac{1}{\tau_2} + \frac{1}{\tau_3} - \lambda\right)\dot{y}(t) + \frac{1}{\tau_2\tau_3}y(t)\right)\right| \tag{4.9}$$

When condition (4.9) is satisfied, sliding mode is enforced and the error trajectory will reach the sliding surface $(\sigma(t) = 0)$ within a finite time. After sliding mode occurs in the manifold $\sigma(t) = 0$, the tracking error $e(t)$ decays exponentially with a time constant equal to $1/\lambda$ [67]. This means that once the system is on the sliding surface, the system response only depends on the design parameter, λ.

The temperature limit, M, that is used for implementation can be derived as follows. In this application, the desired thickness trajectory is chosen to be a negative step of magnitude Y_d such that

$$y_d(t) = -Y_d,\ \dot{y}_d(t) = 0 \text{ and } \ddot{y}_d(t) = 0 \qquad \text{for } t > 0 \tag{4.10}$$

At $t = 0^+$, $y(t) = 0$ and $\dot{y}(t) = 0$ such that $e(0^+) = -Y_d$ and $\dot{e}(0^+) = 0$. From Equation (4.4), $\sigma(0^+) = \lambda e(0^+) = -\lambda Y_d$. Also, from Equation (4.5), $w(0^+) = M$ since $\sigma(0^+) < 0$. For sliding mode control, the control law is continuous until the system first reaches the sliding surface at $t = t_\sigma$, i.e. $w(t) = M$ for $0 < t \le t_\sigma$. After sliding mode has started at $t = t_\sigma$, the control law becomes discontinuous as the ideal sliding mode controller switches infinitely fast between M and $-M$ to remain on the sliding surface. Substituting $w(t) = M$ for $0 < t \le t_\sigma$, with initial condition of $y(0^+) = 0$ and $\dot{y}(0^+) = 0$ in Equation (4.3),

$$y(t) = k_2 M \left\{ 1 + \frac{1}{\tau_2 - \tau_3} \left(\tau_3 \exp\left(-\frac{t}{\tau_3} \right) - \tau_2 \exp\left(-\frac{t}{\tau_2} \right) \right) \right\} \tag{4.11}$$

and

$$\dot{y}(t) = \frac{k_2 M}{\tau_2 - \tau_3} \left(\exp\left(-\frac{t}{\tau_2} \right) - \exp\left(-\frac{t}{\tau_3} \right) \right) \tag{4.12}$$

Substitute Equations (4.10) – (4.12) into Equation (4.7) and integrating $\dot{\sigma}(t)$ from $t = 0^+$ to t_σ gives

$$M = \frac{\sigma(0^+)}{k_2} \frac{(\tau_3 - \tau_2)}{\left\{(\lambda\tau_2 - 1)\exp\left(-\frac{t_\sigma}{\tau_2}\right) + (1 - \lambda\tau_3)\exp\left(-\frac{t_\sigma}{\tau_3}\right) + \lambda(\tau_3 - \tau_2)\right\}} \quad (4.13)$$

where $\sigma(0^+)$ is a function of the initial thickness error given as $\sigma(0^+) = -\lambda Y_d$.

It can be shown that the temperature limit, M, calculated using Equation (4.13) also satisfy the sliding mode condition given in (4.9). From Equation (4.9), the following condition must be true in order to enforce sliding mode.

$$M - M_0(t) > 0 \qquad \forall \mathrm{t} > 0$$

where $$M_0(t) = \left| \frac{\tau_2\tau_3}{k_2}\left(\ddot{y}_d(t) + \lambda\dot{y}_d(t) + \left(\frac{1}{\tau_2} + \frac{1}{\tau_3} - \lambda\right)\dot{y}(t) + \frac{1}{\tau_2\tau_3}y(t)\right)\right| \quad (4.14)$$

Substituting Equations (4.10) to (4.12) into M_0 given in Equation (4.14) for $0 < t \le t_\sigma$,

$$M_0(t) = \left| 1 + \frac{1}{\tau_2 - \tau_3}\left((\lambda\tau_2\tau_3 - \tau_2)\exp\left(-\frac{t}{\tau_3}\right) + (\tau_3 - \lambda\tau_2\tau_3)\exp\left(-\frac{t}{\tau_2}\right)\right)\right| M$$

Let $\tau_2 = \beta\tau_3$ where $\beta > 1$ and substitute it into above equation

$$M_0(t) = \left| 1 + \frac{1}{\beta - 1}\gamma \right| M \qquad \text{where } \gamma = \left((\lambda\beta\tau_3 - \beta)\exp\left(-\frac{t}{\tau_3}\right) - (\lambda\beta\tau_3 - 1)\exp\left(-\frac{t}{\beta\tau_3}\right)\right)$$

Since $(\beta\lambda\tau_3 - 1) > (\beta\lambda\tau_3 - \beta)$ and $\exp\left(-\frac{t}{\beta\tau_3}\right) > \exp\left(-\frac{t}{\tau_3}\right)$, $\gamma < 0$ if $(\beta\lambda\tau_3 - 1) > 0$ or $\lambda > \frac{1}{\beta\tau_3}$. Also, $M_0(t) > 0$. This gives

$$M - M_0(t) = -\frac{1}{\beta - 1}\gamma > 0$$

Hence, as long as $\lambda > \frac{1}{\tau_2}$ where $\tau_2 = \beta\tau_3$, the sliding mode can be enforced. The sliding mode controller takes into account the uncertainty in the plant parameters to achieve "perfect" tracking of the desired thickness trajectory for all plants within the uncertainty region. To prevent decomposition of the resist, the bake temperature is constrained to $90 \pm M$ °C. For implementation, the temperature limit, M, is calculated using Equation (4.13) by specifying the reaching time, t_σ, for the set of plant parameters that requires the largest temperature limit, M, to reach the sliding surface within a finite time. In the next section, the choice of design parameters is discussed in greater details.

4.2.3 Implementation

To filter out the measurement noise, the pure differentiator, s, is replaced by the term, $\frac{s}{s\tau_f + 1}$. To ensure that the inclusion of the filter will not have a significant effect on the closed loop response, the time constant, τ_f, should be sufficiently small so that the filter has a much faster dynamics than the plant. For implementation, τ_f =0.25 s is chosen through

trial and error. Another important design parameter is the selection of the proportional gain for the bakeplate. The proportional gain, $k_{c1} = 0.26$, is chosen so that the time constant between w and T is fast (≈one fifth of the time constant of $G_1(s)$). Choosing a proportional gain $k_{c1} = 0.26$ gives $k_{c2} = 1.2$ and $\tau_3 = 5.4$ s.

The actual processes at different sites are likely to differ from one another and from the process models identified in Section 4.2.1. For simplicity, the parameters of the thermal model of the bakeplate, $G_1(s)$, are assumed to be constant and equal to $\hat{k}_1$ and $\hat{\tau}_1$. As a nominal thickness model, $G_2(s)$, is used for all the controlled sites, a 10% uncertainty in k_2 and τ_2 is assumed such that $0.9\hat{k}_2 \le k_2 \le 1.1\hat{k}_2$ and $0.9\hat{\tau}_2 \le \tau_2 \le 1.1\hat{\tau}_2$. This determines the choice of another important design parameter, λ. In Section 4.2.2, it is shown that $\lambda > \frac{1}{\tau_2}$ is required to ensure the enforcement of the sliding mode. Therefore, to ensure sliding mode is enforced for all plants within a 10% model uncertainties, λ is chosen such that $\lambda > \frac{1}{0.9\hat{\tau}_2}$ or $\lambda > 0.037$. The percentage of model uncertainty depends on the particular equipment setup. The sliding mode control technique will still work for a different model uncertainty, as long as it is adequate for its application. Also, $\lambda = 0.25$ is chosen so that the system will slide along the sliding surface towards the reference with a time constant of $1/\lambda = 4$ s.

Once these design parameters have been chosen, simulations can be performed to examine the behaviour of the system for different k_2 and τ_2 parameters, the effects of digital implementation of a continuous sliding mode control and incomplete cancellation of the time delay by the smith predictor. These will be discussed in the Appendix B. In the next section,

a comparison is made between the implementation of the real-time thickness control strategy using the sliding mode control and GPC algorithm.

4.3 Comparison with GPC Algorithm

For both sliding mode control and GPC algorithms, the choice of the upper and lower temperature limits is important during implementation of the real-time thickness control strategy. The temperature bounds are imposed during implementation of the thickness control strategy to prevent decomposition of the photoactive compound in the resist. In the GPC algorithm, the upper temperature bound is determined by the lowest resist sensitivity. For this application, the resist sensitivity ranges from -2.08 nm/°C to -2.45 nm/°C. For this set of resist sensitivities, the upper temperature bound is fixed at 96 °C ($T_{\max} = 6$) for every runs when the reference is negative step of 10 nm. For the sliding mode control algorithm, the temperature limit is determined by the control requirement to ensure convergence to the sliding surface within a finite time, t_σ. The temperature bound is not only dependent on the reference, y_d , it is also dependent on t_σ. If a bake process is short, t_σ has to be smaller to ensure that resist thickness at all sites convergence to the reference trajectory before the completion of the bake process. Hence, for the same reference, the temperature limit will be different for different choice of t_σ. For example, if the reference is a negative step of 10 nm, the temperature limit, M , that is required for convergence when $t_\sigma = 35s$, $t_\sigma = 60s$ and $t_\sigma = 90s$ is $M = 8.8$ °C, 6.7 °C and 6.0 °C using Equation (4.13) respectively. If $t_\sigma = 90s$

is chosen for the sliding mode implementation, the bake temperature will be constrained to between 84 °C and 96 °C. If the first control move is made 15 s after the placement of the wafer on the bakeplate, the resist thickness will first reach the sliding surface at time, $t \leq 105s$. The reaching time, t_σ, is a conservative estimate of the maximum time necessary to reach the sliding surface [67]. The actual time the resist thickness first reaches the sliding surface and start tracking the reference trajectory might be less than 90 s after making the first control move. In comparison, the resist thickness will start to track the reference after 65 s when the upper temperature limit in the GPC algorithm is chosen to be 96 °C (see Figure 3.9 in Section 3.4.3.3). To ensure that the time that the resist thickness tracks the reference using the sliding mode implementation is comparable to that using the GPC algorithm, a shorter reaching time such as $t_\sigma = 60s$ may be chosen. In that case, the upper temperature bound has to be set at a slightly higher value of 96.7 °C.

One of the motivations for using the sliding mode control algorithm is to provide a simpler and faster implementation of the thickness control strategy to make it suitable for real-time application. Using functions from the MATLAB toolbox (tic.m, toc.m) on a 450 MHz, Pentium III processor, a comparison of the computation time using the Generalized Predictive Control (GPC) and sliding mode control algorithm can be made. While the QP problem takes 12.1 ms to compute the control signal for one site, the time to find a solution for sliding mode control is only 0.16 ms. We observe an 75 times improvement in computation time. In the experiment, two sites on the wafer were monitored and controlled at a sampling interval of 0.1 s. In practice, more sites have to be controlled to achieve good performance. For example, it was found that 49 independently temperature controlled sites are desirable for temperature control [40]. In this case, thickness at 49 sites can be

controlled. The computation time required for the GPC and sliding mode controller will then be about 0.6 s (49×12.1 ms) and 0.008 s (49×0.16 ms) respectively. It is clear that the GPC algorithm cannot meet the 0.1 s sampling required.

The 49-zones bakeplate [40] was built for a photomask mask of 6×6 inch. If we scale it up to a 12-inch wafer, then about 154 sites have to be controlled. In this case, the computational intervals for the GPC and sliding mode control are about 1.9 s and 0.025 s respectively. The 1.9 s is the time required just for the numerical calculation using the GPC algorithm. It does not include the time required for the thickness estimation nor the software interface, which would then result in sampling interval greater than 1.9 s. Thus, the sliding mode approach is better in terms of computation time and is more suitable for real-time application. In terms of tracking performance, the experimental result shows the tracking performance using sliding mode control algorithm and GPC algorithm are comparable. At the end of the bake process, resist thickness non-uniformity of less than 1 nm was also achieved.

4.4 Experimental Results

To demonstrate the control strategy, the resist thickness at two sites were monitored and made to track a reference trajectory using in-situ thickness sensors and the sliding mode control algorithm. The first control move was made 15 s after the placement of the wafer. The parameters used in the experiment have been discussed earlier and are now summarized in Table 4.1.

Table 4.1: Summary of the parameters.

$\hat{k}_1 (^oC/V)$	20
$\hat{k}_2 (nm/^oC)$	-2.0
$k_{c1} (V/^oC)$	0.26
k_{c2}	1.2
$\hat{\tau}_1 (s)$	33.3
$\hat{\tau}_2 (s)$	30.3
$\tau_3 (s)$	5.4
$L_2 (s)$	5.0
$t_\sigma (s)$	35
$\lambda (s^{-1})$	0.25

Figure 4.4 shows the first experimental run for 60 s. At the end of the 60 s bake, the wafer was removed from the bakeplate. The initial thickness error, $e(0^+)$, at Site A and B on the wafer are 2.5 nm and 5.4 nm respectively. Using the parameters in Table 4.1 and Equation (4.13) gives $M = 2.2$ and $M = 4.7$ for Site A and Site B respectively. Figure 4.4(b) shows the bakeplate temperature for both sites and Figure 4.4(c) shows the control signal of the sliding mode controller, w, which determines the temperature bounds for the bakeplate. By having the sliding mode controller in the outer control loop, the bake temperature for each site is constrained to an upper bound of $90 + M$ °C (92.2 °C for Site A and 94.7 °C for Site B). Notice in Figure 4.4(c), the control signal, w, switches from the

upper temperature bound to the lower temperature bound at $t = 42.5$ s or 27.5 s after the first control move. This is less than the specified worst case t_σ that is chosen to be 35 s after the first control move. Figure 4.4(d) shows the heater input, u. At the end of the softbake process, the resist thickness at the two sites converged to the target thickness of 1680 nm at $t = 60$ s.

For a non-repeatable coating process where the resist thickness and its non-uniformity varies from wafer-to-wafer, a real-time thickness control strategy helps to improve the repeatability of the process in terms of resist thickness requirements. Figure 4.5 shows the results of another experimental run. Site A and B has an initial error of 7.9 nm and 3.3 nm respectively. The temperature limit, M, is calculated such that $M = 6.9$ for Site A and $M = 2.9$ for Site B. Again, Site A and B first reaches the sliding surface at $t = 43.9$ s and $t = 45.7$ s respectively; i.e. 28.9 s (Site A) and 30.7 s (Site B) after the first control move. Again, this is less than the specified worst case t_σ of 35 s. In this run, the resist thickness and its non-uniformity prior to the softbake process is different from the previous run. At the end of the softbake process, the resist thickness at the two monitored sites converged to the same target thickness of 1680 nm and the resist thickness difference between the two sites was reduced. In both experimental runs, chattering phenomenon exist as the sliding mode controller switches between the two control limits of $\pm M$. However, as long as resist thickness non-uniformity between different sites is less than 1 nm, chattering phenomenon is still acceptable for this application. Modification of the control law to eliminate chattering phenomenon may reduce the tracking precision of sliding mode control.

4.5 Conclusion

Real-time thickness control has been implemented using an array of in-situ thickness sensors and a sliding mode control algorithm. The sliding mode control algorithm is used to compute the heater power distribution needed to reduce the resist thickness non-uniformity. In addition to improvement in the resist thickness uniformity across individual wafers, the softbake process is also made more repeatable as the same thickness trajectory is defined for all wafers; resulting in an improvement in the resist thickness uniformity from wafer-to-wafer. There is about a 75 times improvement in the computation time using this approach. This general control strategy is simple and may also be extended to similar applications demanding stringent film thickness uniformity.

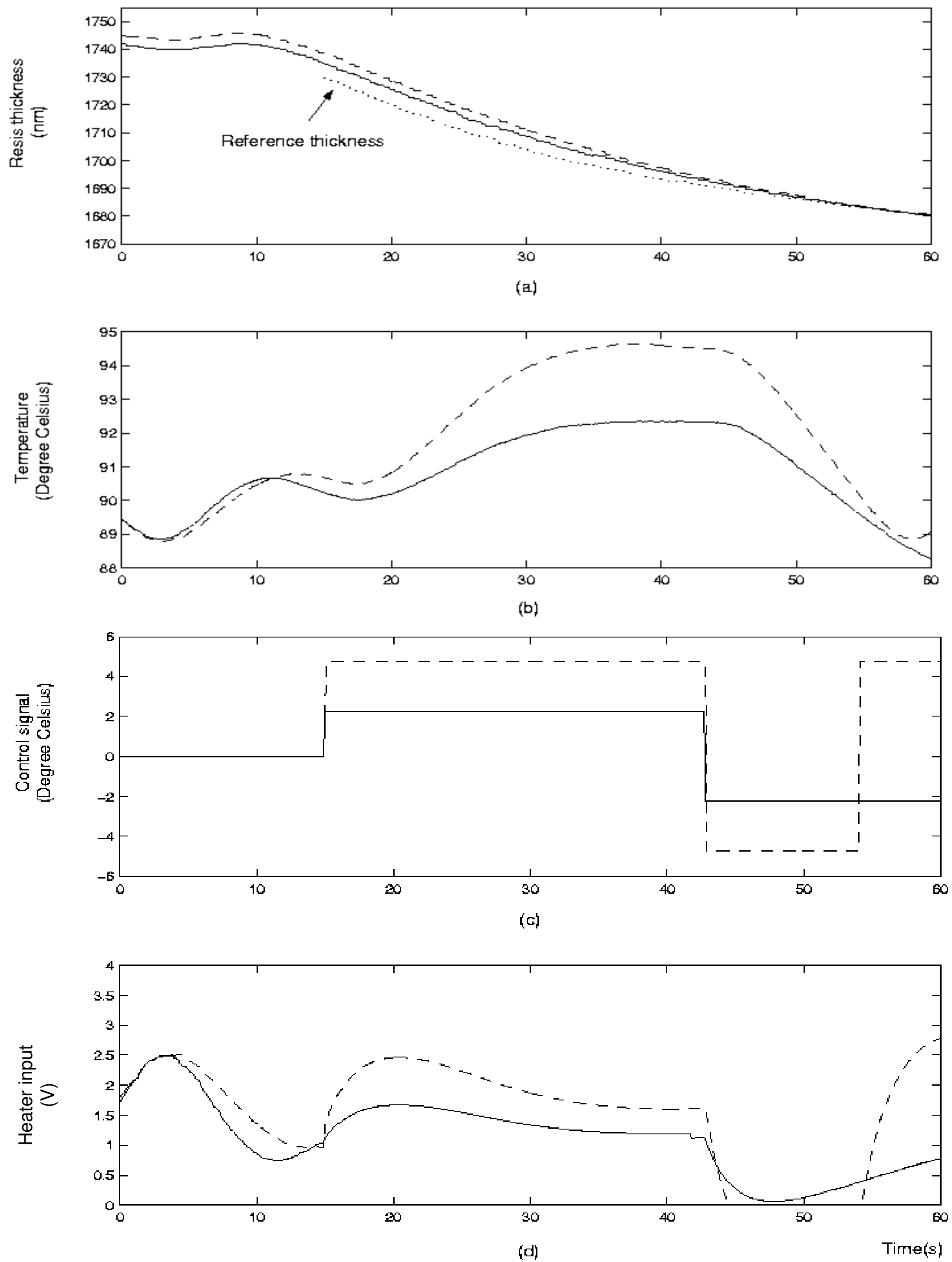


Figure 4.4. Sliding mode control for Run #1. Plots of (a) resist thickness, (b) temperature, (c) control signal, and (d) heater input with respect to time when resist thickness at Site A and B are monitored. Site A and B are represented by the solid and dashed lines respectively. The reference thickness trajectory is given by the dotted line in (a).

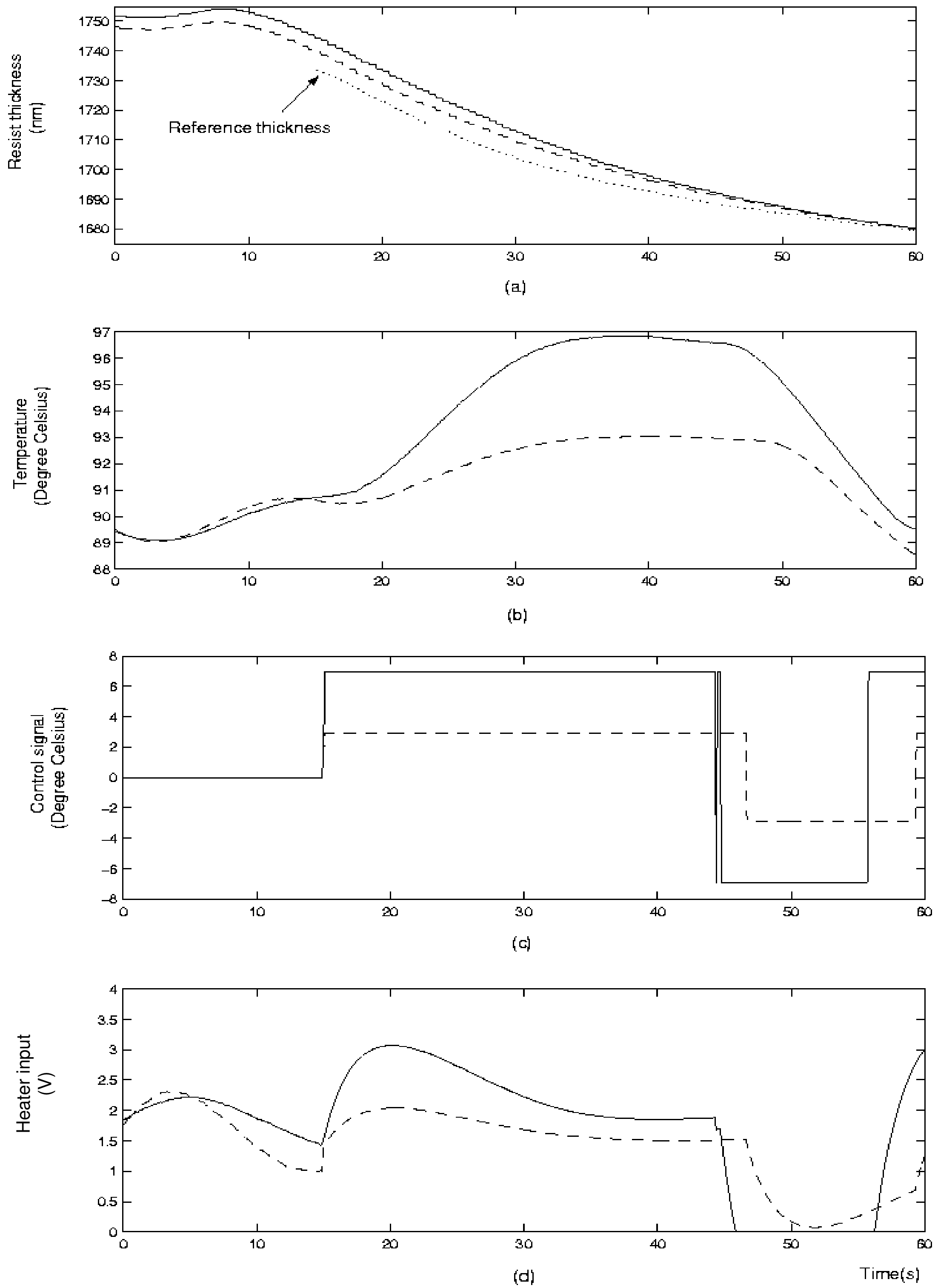


Figure 4.5: Sliding mode control for Run#2. Plots of (a) resist thickness, (b) temperature, (c) control signal, and (d) heater input with respect to time when resist thickness at Site A and B are monitored. Site A and B are represented by the solid and dashed lines respectively. The reference thickness trajectory is given by the dotted line in (a).

Chapter 5

Conclusion

5.1 Review of Objectives and Summary of Results

The trend in the semiconductor industry is towards to the use of more advanced process control methods to meet the tightening process specifications with the continual shrinking of the feature sizes. This thesis examines the application of advanced control algorithm to meet the challenges of some aspects of advanced lithography, particularly on the bake process. Advanced control is applied to two areas: photomask thermal processing and real-time thickness control. In this Section, the results are summarized. The scope for future developments is looked into in the next section.

In Chapter 2, an optimal predictive controller is designed to effectively eliminate the loading effects induced by the placement of cold photomask on the bakeplate and improve the bake process repeatability. Due to large thermal mass of the photomask, the placement of the cold photomask on the bakeplate result in a great drop in the bakeplate temperature (load disturbance). The predictive controller is part of a feedforward-feedback control scheme where the feedforward control is based on a linear programming method of minimizing the worst-case deviation from the temperature setpoint during loading effect. By predicting the impact of load disturbance and increasing the heater power prior to the placement of the cold photomask onto the bakeplate, the feedforward controller can reject the load disturbance induced by placement of the cold substrate. With the effective rejection of the load

disturbance by the feedforward controller, the feedback controller is able to maintain the bakeplate temperature at the setpoint during the bake process. This approach seeks to improve the repeatability of the temperature response from wafer-to-wafer and from plate-to-plate (matching). A significant improvement in the bakeplate temperature error is achieved using the optimal predictive controller as compared to the conventional approach of using only the feedback controller for temperature control of the bake process.

In Chapter 3 and 4, a real-time thickness control strategy is implemented for the softbake process to achieve good resist thickness uniformity across the wafer and from wafer-to-wafer. The approach is to use an array of in-situ thickness sensors positioned above a multi-zones bakeplate to monitor the resist thickness. With these in-situ thickness measurements, the temperature profile of the bakeplate is controlled in real-time by manipulating the heater power distribution using suitable control algorithm. Often, it results in a non-uniform temperature distribution and creates different solvent removal rates and densifications at various locations on the wafer to give a uniform thickness distribution across the wafer. Resist thickness non-uniformity of less than 1 nm has been achieved. In Chapter 3, the Generalized Predictive Control (GPC) algorithm is used to compute the heater power distribution needed to minimize the resist thickness non-uniformity. To obtain a simpler and faster implementation of the real-time control strategy, sliding mode control algorithm is used for the implementation of the thickness control strategy in Chapter 4. Like the implementation using the GPC algorithm, resist thickness uniformity of less than 1 nm is also achieved. In addition to that, there is about a 75 times improvement in the computation time using this approach over the GPC algorithm. Besides improvements in the resist thickness uniformity across individual wafers, the softbake process is also made more

repeatable as the same thickness trajectory is defined for all wafers; resulting in an improvement in the resist thickness uniformity from wafer-to-wafer.

Both feedforward control and real-time resist thickness control are model-based control technique. An accurate model must be obtained before they can work well. This will be the main problem for implementing these two techniques in the semiconductor industry

5.2 Scope for Future Developments

The focus of this thesis is on the bake process involved in lithography. The same experimental setup and control strategy can also be extended to the development process. This is because the development process is also a strong function of temperature. In order to implement this for the development process, it is important to be able to monitor the development process on-line. While commercial development rate monitor is available, a more advanced data analysis is required to estimate the resist thickness from the reflectance signals during development. This is because reflectance signal used for resist thickness measurement may be distorted due to the topography of wafers, absorbing resist residue in developer, developer layer and many other factors [68]. Given that the development process is the last step in formation the resist patterns, the ability to control the development process through real-time manipulation of development temperature can reduce CD variations.

In the real-time thickness control application, the control algorithm manipulates the bakeplate temperature so that a non-uniform temperature distribution will result in a uniform

resist thickness distribution. This idea can be extend to control the CD since CD is a function of resist thickness. Similarly, the resist thickness can be manipulated in real-time so that a non-uniform resist film can give rise to a more uniform CD distribution. This can help to compensate for any CD variation caused by variations of other process variables.

Also, the research work done on the photomask thermal processing can also be extended to wafer processing. For thermal processing of the wafer, the wafer will rest on the pins for proximity baking. There is a small air gap between the wafer and the bakeplate surface. If there is no warpage, this air gap will be uniform. The load disturbance caused by placement of the wafer on the bakeplate will be the same from run-to–run. However, if the wafer warps for a particular run, the air gap between the wafer and the bakeplate will not be uniform. Hence, the load disturbance due to the placement of this wafer will be different from other runs. The shape of the load disturbance can be used to provide information of the wafer warpage and corrective actions can be taken to prevent warpage for subsequent runs. This is important as wafer warpage is expected to be more critical with the use of large wafer size.

References

[1]. "International Technology Roadmap for Semiconductors: lithography", Semiconductor Industry Association, 1999.

[2] M. Quirk and J. Serda, Semiconductor Manufacturing Technology, Prentice Hall, 2001.

[3]. S. Limanond, J. Si and K. Tsakalis, "Monitoring and Control of Semiconductor Manufacturing Processes", IEEE Control Systems, December 1998.

[4] P. J. Paniez, A. Vareille, P. Ballet, B. Mortini, "Study of Bake Mechanisms by Real-time In-Situ Ellipsometry", Proc. SPIE, vol. 3333, pp. 289-300, 1998.

[5] T. L. Vincent, P. P. Khargonekar, F. L. Terry, Jr., "An Extended Kalman Filtering-Based Method of Processing Reflectometry Data for Fast In-Situ Etch Rate Measurements", IEEE Trans. Semiconduct. Manuf., vol. 10, no. 1, pp. 42-51, 1997.

[6] T. E. Metz, R. N. Savage and H. O. Simmons, "In-Situ Thickness Measurements for Real-Time Monitoring and Control of Advanced Photoresist Track Coating Systems", Proc. SPIE, vol. 1594, pp. 146-152, 1991.

[7] M. L. Miller, "Use of Scatterometric Measurements for Control of Lithography", Ph.D. dissertation, University of California, Santa Barbara.

[8] C. J. Raymond, "Angle-resolved Scatterometry for Semiconductor Manufacturing", Micorlithography World, pp. 18-23, Winter 2000.

[9] G. Barna, L. Loewenstein, R. Robbins, S. O'Brien, A. Lane, D. White, M. Hanratty, J. Hosch, G. Shinn, K. Taylor and K. Branker, "MMST Manufacturing Technology –

Hardware, Sensors, and Processes", IEEE Trans. Semicond. Manuf., vol. 7, no. 2, pp. 115-125, 1994.

[10] E. Palmer, W. Ren, C. J. Spanos, and K. Poolla, "Control of Photoresist Properties: A Kalman Filter Based Approach", IEEE Trans. Semiconduct. Manuf., vol. 9, no. 2, pp. 208-214, 1996.

[11] T. S. Caroll and W. F. Ramirez, "Development of Positive Optical Photoresist: Adaptive Control", Chemical Engineering Science, vol. 48, no. 12, pp. 2239-2250, 1993 (June).

[12] B. Kim, G. S. May, " An Optimal Neural Network Process Model for Plasma Etching", IEEE Trans. Semicond. Manuf., vol. 7, no. 1, pp. 12-21, 1994.

[13] E. A. Rietman and E. R. Lorry, "Use of Neural Networks in Modeling Semiconductor Manufacturing Processes: An Example for Plasma Etch Modeling", IEEE Trans. Semicond. Manufact., vol. 6, no. 4, pp. 343-347, 1993.

[14] J. Si and Y-L. Tseng, "Control Relevant RIE Modeling by Neural Networks from Real-Time Production State Sensor Measurements", Proc. Amer. Contr. Conf., pp. 1583-1587, Albuquerque, June 1997.

[15] H. J. Levison, Lithography Process Control, SPIE Optical Press, 1999.

[16] G. E. Flores, W.W. Flack and S. Avlakeotes, "Statistical Process Control for Monitoring Stepper Overlay", Microlithography World, pp. 14-19, Spring, 1995.

[17] P. K. Mozumder and G. G. Barna, "Statistical Feedback Control of Plasma Etch Process", IEEE Trans. Semicond. Manuf., vol. 7, no.1, pp. 7-11, 1994.

[18] P. K. Mozumder, S. Saxena and D. J. Collins, "A Monitor Wafer Based Controller for Semiconductor Processes", IEEE Trans. Semicond. Manuf., vol. 7, no. 3, pp. 400-411, 1994.

[19] T. F. Edgar, S. W. Butler, W. J. Campbell, C. Pfeiffer, C. Bode, S. B. Hwang, K. S. Balakrishnan, J. Hahn, "Automatic Control in Microelectronics Manufacturing: Practices, Challenges, and Possibilities", Automatica (36), pp. 1567-1603, 2000.

[20] B. A. Rashap, M. E. Elta, H. Etemad, J. P. Fournier, J. S. Freudenberg, M. D. Giles, J. W. Grizzle, P. T. Kabamba, P. P. Kargonekar, S. Lafortune, J. R. Moyne, D. Teneketzis and F. L. Terry, "Control of Semiconductor Manufacturing Equipment: Real-Time Feedback Control of a Reactive Ion Etcher", IEEE Trans. Semicond. Manuf., vol. 8, no. 3, pp. 286-297, 1995.

[21] M. Hankinson, T. Vincent, K. B. Irani and P. P. Khargenekar, " Integrated Real-Time and Run-to-Run Control of Etch Depth in Reactive Ion Etching", IEEE Trans. Manuf., vol. 10, no. 1, pp. 121-130, 1997.

[22] E. A. Rietman and S. H. Patel, "A Production Demonstration of Wafer-to-Wafer Plasma Gate Etch Control by Adaptive Real-Time Computation of the Over-Etch Time from in-situ Process Signals", IEEE Trans. Semicond. Manuf., vol. 8, no. 3, pp. 302-308, 1995.

[23] C-Y. Ku, T. F. Lei, J. M. Shieh, T. B. Chiou, Y. C. Chen, "Real-Time Control to Prevent CD Variation Induced by Post Exposure Delay", Proc. SPIE, vol. 4182, pp. 40-47, 2000.

[24] R. M. Newcomb, J. Pellgrini and B. Singh, "A New Paradigm for Real-time Stepper Performance Monitoring," Microlithography World, pp. 5-9, Autumn, 1996.

[25] K. El-Awady, C. Schaper and T. Kailath, “Improvements in C_{pk} using Using Real-Time Feedback Control”, IEEE Trans. Semicond. Manuf., vol. 9, no. 1, pp. 87-94, 1996.

[26] O. D. Patterson and P. P. Khargonekar, “Methodology for Real-time Feedback Variable Selection for Manufacturing Process Control: Theoretical and Simulation Results”, Proc. SPIE, vol. 3507, pp. 30-41, 1998.

[27] J. Sturtevant, “ Manufacturing Implementation of a Feedback Controller for CD and Overlay”, Microlithography World, pp. 22-26, Summer, 1999.

[28] S. L. Morton, F. L. Degertekin, and B. T. Khuri-Yakub, “Ultrasonic Monitoring of Photoresist Processing”, Proc. SPIE, vol. 3677, pp. 340-347, 1999.

[29] S. L. Morton, “Ultrasonic Sensor for Photoresist Process Monitoring”, Ph.D dissertation, Stanford University, 1999.

[30] M. D. Baker, F. R. Williams and G. S. May, “A Novel In Situ Monitoring Technique for Reactive Ion Etching Using a Surface Micromachined Sensor”, IEEE Trans. Semicond. Manufact., vol. 11, no. 2, pp. 254-264, 1998.

[31] Z. H. Zhou and R. Reif, “Epi-Film Thickness Measurements Using Emission Fourier Transform Infrared Spectroscopy-Part II: Real-Time in Situ Process Monitoring and Control”, IEEE Trans. Semicond. Manuf., vol. 8, no. 3, pp. 340-345, 1995.

[32] T. F. Edgar, W. J. Campbell, C. Bode, “Model-based Control in Microelectronics Manufacturing,” Proc. 38th Conf. Decision and Control, pp. 4185-4191, 1999.

[33] E. S. Hamby, P. T. Kabamba, P. P. Khargonekar, “A Probabilistic Approach to Run-to-Run Control”, IEEE Trans. Semicond. Manuf., vol. 11, no. 4, pp. 654-668, 1998.

[34] S. W. Butler and J. Stefani, “Supervisory Run-to-Run Control of Polysilicon Gate Etch Using In-situ Ellipsometry”, IEEE Trans. Semicond. Manuf., vol. 7, no. 2, pp. 193-201, 1994.

[35] S. Leang. S. Ma, J. Thompson, B. J. Bombay and C. J. Spanos, “A Control System for Photolithographic Sequences”, IEEE Trans. Semicond. Manuf., vol. 9, no. 2, pp. 191-207, 1996.

[36] J. Stefani, S. Poarch, S. Saxena and P.K. Mozumder, “Advanced Process Control of a CVD Tungsten Reactor”, IEEE Trans. Semicond. Manuf., vol. 9, no.3, pp. 366-383, 1996.

[37] T.H. Smith, D.S. Boning, J. Stefani and S. W. Butler, “Run by Run Advanced Process Control of Metal Sputter Deposition”, IEEE Trans. Semicond. Manuf., vol. 11, no. 2, pp. 276-284, 1998.

[35] J. Sturtevant, S. Holmes, T. VanKessel, P. Hobbs, J. Shaw, and R. Jackson, “Post-Exposure Bake as a Process-Control Parameter for Chemically-Amplified Photoresists”, Proc. SPIE, vol. 1926, pp. 106-114, 1993.

[36] J. Parker and W. Renken, “Temperature Metrology for CD Control in DUV Lithography”, Semicond. Intl, vol. 20, no. 10, pp. 111-116, 1997.

[37] N. Ramanan, F. F. Liang and J. B. Sims, “Conjugate heat-transfer analysis of 300-mm bake station,” Proc. SPIE, vol. 3678, pp. 1296-1306, 1999.

[38] G. MacBeth, “Prebaking Positive Photoresists”, Proc. Kodak Microelectronics Seminar, pp. 87-92, 1982.

[39] Y. Kushida, Y. Usui and H. Shirai, “Improvement of ZEP Process for Advanced Mask Fabrication”, Proc. SPIE, vol. 4066, pp. 261-268, 2000.

[40] K. El-Awady, “Spatially Programmable Thermal Processing Module for Semiconductors”, Ph.D. dissertation, Stanford University, 2000.

[41] P. Dress, T. Gairing, W. Saule, U. Dietze and J. Szekeresch, “Improved Baking of Photomasks by a Dynamically Zone-Controlled Process Approach”, Proc. SPIE, vol. 4409, pp. 356-363, 2001.

[42] J. P. Kuijten, F. Duray and T. D. Kinderen, “Analysis of Reticle Contributions to CD Uniformity for 0.25 μm DUV Lithography”, Proc. SPIE-Int. Soc. Eng., vol. 3334, pp. 620-628, 1998.

[43] O. D. Crisalle, “Improvements in Photolithography Performance by Controlled Baking”, Proc. SPIE, vol. 921, pp. 317-325, 1998.

[44] A. E. Braun, “Track Systems Meet Throughput and Productivity Challenges”, Semicond. Intl., vol. 21, no. 2, pp. 63-68,1998.

[45] G.C. Goodwin and S. S. Sin, Adaptive, Filtering, Prediction, and Control, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[46] T. F. Edgar, D. M. Himmelblau and L.S. Lasdon, Optimization of Chemical Processes, McGraw-Hill International, 2001.

[47] I. D. Landau, System Identification and Control Design, Prentice Hall, 1990.

[48] J. R. Sheats, B. W. Smith, Microlithography: Science and Technology, Marcel Dekker Inc., New York, 1998.

[49] T. A. Brunner, “Optimization of Optical Properties of Resist Processes”, Proc. SPIE, vol. 1466, pp. 297-308, 1991.

[50] L. M. Peurrung and D. B. Graves, "Spin Coating over Topography", IEEE Trans. Semicond. Manuf., vol. 6, no. 1, pp. 72-76, 1993.

[51] A.B. Charles, J. G. Maltabes, S. R. Hornig, T. S. Schedel, D. Ganz, S. schmidt, L. Grant, G. Hraschan, K. E. Mautz, R. Otto, "Current State of 300 mm Lithography in Pilot Line Environment", Proc. SPIE, Vol. 3882, pp. 140, 1999.

[52] E. Gurer, T. Zhong, J. Lewellen, E. Lee, "A Novel Spin Coating Technology for 248 nm/193 nm DUV Lithography and Low-k spin on Dielectrics of 200/300 mm wafers", Proc. SPIE, vol. 3999, pp. 805-817, 2000.

[53] Y. Kushida, Y. Usui, T, Kobayashi, and K. Shigematsu, "Advanced Resist Coating Technology for Mask Manufacturing Process", Proc. SPIE, vol. 3679, pp. 1001-1008, 1999.

[54] T. Tokimitsu, "Consideration About Possibility of Spin Cup Method", Proc. SPIE, vol. 3412, pp. 196-200, 1998.

[55] A. Suriadi and T. Luxbacher, "Photolithography on Micromachined 3-D Surfaces Using Spray Coating Technology of Photoresist", Proc. SPIE, vol. 4404, pp. 245-253, 2001.

[56] W. K. Ho, A. Tay and C. D. Schaper, "Optimal Predictive Control with Constraints for the Processing of the Semiconductor Wafers on Large Thermal-Mass Heating Plates", IEEE Trans. Semicond. Manuf., vol. 13, no. 1, pp. 88-96, 2000.

[57] E. Fadda, C. Clarisse, P. J. Paniez, "Study of Bake Mechanisms in Novolak Based Photoresist Films: Invetigation by Contact Angle Measurements", Proc. SPIE, vol. 2714, pp. 460-468, 1996.

[58] C.L. Henderson, S.A.Sheer, P.C. Tsiartas, B.M. Rathsack, J.P. Sagan, R.R. Dammel, A.Erdmann, C.G. Willson, "Modeling Parameter Extraction for DNQ-Novolac Thick Film Resist," Proc. SPIE, vol. 3333, pp. 256-267, 1998.

[59] Ocean Optics Inc., Product Catalog, 2002.

[60] G. R. Fowles, Introduction to Modern Optics, Dover, New York, 1975.

[61] M.Born, E. Wolf, Principles of Optics, Pergamon Press, Oxford, 1980.

[62] E. F. Camacho, C. Bordons, Model Predictive Control, Springer-Verlag, London, 1999.

[63] D. E. Seborg, T. F. Edgar, D. A. Mellichamp, Process Dynamics and Control, John Wiley & Sons, New York, 1989.

[64] K. W. Lim, W. K. Ho, T. H. Lee, K. V. Ling, W. Xu, "Generalized Predictive Controller with Pole Restriction", IEEE Proc. Parts D, Control Theory and Applications, vol. 145, no. 2, pp. 219-225, 1998.

[65] Vadim Utkin, Jürgen Guldner and Jingxin Shi, Sliding Mode Control in Electromechanical Systems, Taylor & Francis, 1999.

[66] K. J. Åstrom, B. Wittenmark, Computer-Controlled Systems: Theory and Design, Prentice Hall, 1997.

[66] Jean-Jacques E. Slotine, Weiping Li, Applied Nonlinear Control, Prentice Hall, 1991.

[67] Christopher Edwards and Sarah K. Spurgeon, Sliding Mode Control: Theory and Applications, Taylor & Francis, 1998.

[68] Michael P.C. Watts, S. Williams, P. McCarthy, M. Tsai, " A Model for End Point Detect During the Development of Photoresists", Proc. SPIE, vol. 1087, pp. 332-343, 1989.

Appendix A

Overview of Generalized Predictive Control (GPC) algorithm

GPC is a model-based predictive control algorithm that computes an optimal control sequence, $[\Delta u(k), \Delta u(k+1), \cdots, \Delta u(k+N-1)]$ by minimizing a quadratic objective function defined over a prediction horizon, N. A major element of the GPC algorithm is the use of the process model to calculate the predicted plant output at future instants. In this application, two process models are identified. One of them is the thickness model relating the change in the resist thickness to the change in the heater input. It is used to predict the resist thickness over a prediction horizon. The prediction of the resist thickness sequence, $\hat{y}_m(k+j|k)$ for $j = d+1, 2, \cdots N+d$, is separated into two parts: free response and forced response, as shown in Figure A.1. The free response corresponds to the prediction of the resist thickness due to past control sequence, $\Delta u(k-j)$ for $j > 0$, while the forced response corresponds to the prediction of the resist thickness based on future optimal control moves, $\Delta u(k+j)$ for $0 \leq j < N$.

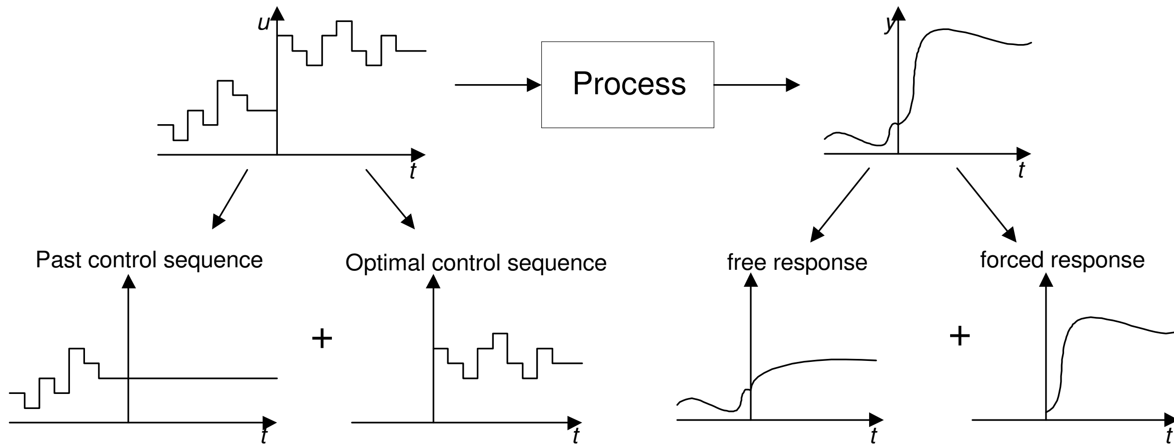


Figure A.1: Free and forced responses

The other process model is the thermal model that relates the change in the bakeplate temperature to the change in heater input. This model is used to predict the bakeplate temperature to ensure that the bake temperature is constrained during the run. With the resist thickness and temperature measurements predicted over the prediction horizon, the optimal control sequence, $\Delta u(k), \Delta u(k+1), \cdots, \Delta u(k+N)$, is computed by the GPC algorithm by minimizing the quadratic objective function, J

$$J = \sum_{j=d+1}^{N+d} [\hat{y}(k+j \mid k) - y_d(k+j)]^2 + \sum_{j=1}^{N} \lambda [\Delta u(k+j-1)]^2 \qquad \text{(A.1)}$$

Subject to the temperature constraint

$$T_{\min} \leq \hat{T}(k+j \mid k) \leq T_{\max} \qquad \text{for } d+1 \leq j \leq N+d$$

where Δu and λ are the change in control signal and control weighting respectively. The control signal in this application is the heater input voltage that is required to minimize the error between the predicted thickness, $\hat{y}_m(k+j\,|\,k)$, and the reference, $r(k+j)$. Also, $\hat{T}(k+j\,|\,k)$ is the optimum j-step ahead prediction of the change in temperature of the bakeplate from the nominal bake temperature of 90 °C based on temperature measurements up to sampling instant, k. The lower and upper bound on the change in temperature from the nominal temperature of 90 °C are given as $T_{\min}$ and $T_{\max}$ respectively.

In the presence of the temperature constraints, the solution has to be obtained using more computationally taxing numerical algorithms. Although a sequence of control moves are computed, only the first control signal of the sequence, $\Delta u(k)$, is used at each sampling instant. A new control sequence is recomputed when a new resist thickness measurement is available.

In the GPC algorithm, the process model is always required for the prediction of the plant output. An Integrated Controller Auto-Regressive Moving Average (CARIMA) model is assumed such that

$$A(q^{-1})y(k) = B(q^{-1})q^{-d}u(k-1) + C(q^{-1})\frac{e(k)}{1-q^{-1}} \tag{A.2}$$

where $e(k)$ is the zero mean white noise, d is the dead time and q^{-1} is the backward shift operator. The control signal and model output are given as $u(k)$ and $y(k)$ respectively. In this application, the control signal is the heater input voltage while the model output may be

either the resist thickness or the bakeplate temperature. The polynomials in the backward shift operator, $A(q^{-1})$ and $B(q^{-1})$ are given as:

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_{na} q^{-na}$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \cdots + b_{nb} q^{-nb}$$

where n_a and n_b are the order of the polynomials. For simplicity, $C(q^{-1})$=1 is chosen.

Consider the Diophantine equation:

$$1 = E_j(q^{-1})\tilde{A}(q^{-1}) + q^{-j} F_j(q^{-1}) \qquad \text{with } \tilde{A}(q^{-1}) = \Delta A(q^{-1}) \text{ and } \Delta = 1\text{-}q^{-1}$$

where E_j and F_j are polynomials with degrees of $j-1$ and n_a respectively. The polynomials, $E_j(q^{-1})$ and $F_j(q^{-1})$, can be obtained recursively by dividing 1 by $\tilde{A}(q^{-1})$ until the remainder of the division can be factorized as $q^{-j} F_j(q^{-1})$.

Multiplying Equation (A.2) by $\Delta E_j(q^{-1}) q^j$ gives

$$\tilde{A}(q^{-1}) E_j(q^{-1}) y(k+j) = E_j(q^{-1}) B(q^{-1}) \Delta u(k+j-d-1) + E_j(q^{-1}) e(k+j) \qquad \text{(A.3)}$$

Using the Diophantine equation, Equation (A.3) can be written as

$$(1 - q^{-j} F_j(q^{-1})) y(k+j) = E_j(q^{-1}) B(q^{-1}) \Delta u(k+j-d-1) + E_j(q^{-1}) e(k+j)$$

This can be rewritten as

$$y(k+j) = F_j(q^{-1})y(k) + E_j(q^{-1})B(q^{-1})\Delta u(k+j-d-1) + E_j(q^{-1})e(k+j) \qquad \text{(A.4)}$$

As the degree of the polynomial, $E_j(q^{-1}) = j-1$, the noise terms in Equation (A.4) are all in the future. The best prediction of $y(k+j)$ is therefore given as:

$$\hat{y}(k+j \mid k) = G_j(q^{-1})\Delta u(k+j-d-1) + F_j(q^{-1})y(k) \quad \text{where } G_j(q^{-1}) = E_j(q^{-1})B(q^{-1})$$

which can also be written as

$$\hat{y}(k+d+j \mid k) = G_j(q^{-1})\Delta u(k+j-1) + F_j(q^{-1})y(k+d) \qquad \text{(A.5)}$$

where $G_j(q^{-1}) = E_j(q^{-1})B(q^{-1})$.

For a first order plant, $A(q^{-1}) = 1 + a_1 q^{-1}$ and $B(q^{-1}) = b_0$ in Equation (A.2). In this case, $\tilde{A}(q^{-1}) = 1 - (1-a_1)q^{-1} - a_1 q^{-2}$ and the polynomials, $E_1(q^{-1})$ and $F_1(q^{-1})$, are obtained at the first step of the division such that

$$E_1(q^{-1}) = 1$$
$$F_1(q^{-1}) = (1-a_1) + a_1 q^{-1}$$

Continuing the division of 1 by $\tilde{A}(q^{-1})$ until the remainder is now factorized as $q^{-2}F_2(q^{-1})$ gives

$$E_2(q^{-1}) = 1 + (1 - a_1)q^{-1}$$
$$F_2(q^{-1}) = (1 - a_1 + a_1{}^2) + (a_1 - a_1^2)q^{-1}$$

Similarly,

$$E_3(q^{-1}) = 1 + (1 - a_1)q^{-1} + (1 - a_1 + a_1^2)q^{-2}$$
$$F_3(q^{-1}) = (1 - a_1 + a_1{}^2 - a_1^3) + (a_1 - a_1{}^2 + a_1^3)q^{-1}$$

Hence for a first order plant, the polynomials, $E_j(q^{-1})$ and $F_j(q^{-1})$ can be computed as

$$E_j(q^{-1}) = E_{j-1}(q^{-1}) + \left(\sum_{i=0}^{j-1}(-a_1)^i\right)q^{-j+1} \qquad \text{where } E_1(q^{-1}) = 1$$
$$F_j(q^{-1}) = \sum_{i=0}^{j}(-a_1)^i - \left(\sum_{i=1}^{j}(-a_1)^i\right)q^{-1} \tag{A.6}$$

Also, $G_j(q^{-1}) = b_0 E_j(q^{-1})$ is a polynomial with degree $j-1$ such that

$$G_j(q^{-1}) = g_0 + g_1 q^{-1} + g_2 q^{-2} + \cdots + g_{j-1}q^{-j+1} \qquad \text{where } g_j = b_0 \sum_{i=0}^{j}(-a_1)^i \tag{A.7}$$

Substituting Equations (A.6) and (A.7) into Equation (A.5), the prediction of the resist thickness, $\hat{y}(k+j)$ for $d+1 \le j \le N+d$, can be expressed as

$$\hat{y}(k+d+1\,|\,k) = g_0\Delta u(k) + \sum_{i=0}^{1}(-a_1)^i\,\hat{y}\big(k+d\big) - (-a_1)\,\hat{y}(k+d-1)$$

$$\hat{y}(k+d+2\,|\,k) = g_o\Delta u(k+1) + g_1\Delta u(k) + \sum_{i=0}^{2}(-a_1)^i\,\hat{y}(k+d) - \sum_{i=1}^{2}(-a_1)^i\,\hat{y}(k+d-1) \qquad \text{(A.8)}$$

$$\vdots$$

$$\hat{y}(k+d+N\,|\,k) = g_o\Delta u(k+N-1) + \cdots + g_{N-1}\Delta u(k) + \sum_{i=0}^{N}(-a_1)^i\,\hat{y}(k+d) - \sum_{i=1}^{N}(-a_1)^i\,\hat{y}(k+d-1)$$

In vector form,

$$\hat{\mathbf{y}} = \mathbf{G}\mathbf{u} + \mathbf{f} \qquad \text{(A.9)}$$

where

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}(k+d+1\,|\,k) \\ \hat{y}(k+d+2\,|\,k) \\ \vdots \\ \hat{y}(k+d+N\,|\,k) \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \\ \vdots \\ \Delta u(k+N-1) \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ g_{N-1} & & \cdots & g_0 \end{bmatrix} \quad \text{and} \quad \mathbf{f} = \begin{bmatrix} f\big(k+d+1\big) \\ f\big(k+d+2\big) \\ \vdots \\ f\big(k+d+N\big) \end{bmatrix}$$

The matrix, $\mathbf{G}$, is made up of N columns of the plant's step response coefficient, g_i for $i = 0, 1, \cdots, N-1$. The free response vector, $\mathbf{f}$, is the part of the response that does not depend on the future control actions. From Equation (A.9), the free response, $f(k+d+j)$, for $1 \le j \le N$ is given as

$$f(k+d+j) = \sum_{i=0}^{j}(-a_1)^i\,\hat{y}(k+d) - \sum_{i=1}^{j}(-a_1)^i\,\hat{y}(k+d-1) \qquad \text{(A.10)}$$

It can be shown in the following that the free response in Equation (A.10) is not dependent on the future control. For a first order plant,

$$\hat{y}(k+d) = -a_1 \hat{y}(k+d-1) + b_o u(k-1)$$

$$\hat{y}(k+d-1) = -a_1 \hat{y}(k+d-2) + b_0 u(k-2)$$

Subtracting the above two equations and re-arranging it gives

$$\hat{y}(k+d) = (1-a_1)\hat{y}(k+d-1) + a_1 \hat{y}(k+d-2) + b\Delta u(k-1)$$

Similarly,

$$\hat{y}(k+d-1) = (1-a_1)\hat{y}(k+d-2) + a_1 \hat{y}(k+d-3) + b\Delta u(k-2)$$

As shown above, $\hat{y}(k+d)$ is dependant only on past control action, $\Delta u(k-1)$, and $\hat{y}(k+d-1)$ depends on $\Delta u(k-2)$. Therefore, the free response $f(k+d+j)$ given in Equation (A.10) does not depend on future control moves.

Similarly, the prediction of bakeplate temperature, $\hat{\mathbf{T}}$, can be expressed as

$$\hat{\mathbf{T}} = \mathbf{G}'\mathbf{u} + \mathbf{f}' \tag{A.11}$$

where

$$\hat{\mathbf{T}} = \begin{bmatrix} \hat{T}(k+d+1 \mid k) \\ \hat{T}(k+d+2 \mid k) \\ \vdots \\ \hat{T}(k+d+N \mid k) \end{bmatrix}, \ \mathbf{G}' = \begin{bmatrix} g'_0 & 0 & \cdots & 0 \\ g'_1 & g'_0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ g'_{N-1} & & \cdots & g'_0 \end{bmatrix} \ \text{and} \ \mathbf{f}' = \begin{bmatrix} f'(k+d+1) \\ f'(k+d+2) \\ \vdots \\ f'(k+d+N) \end{bmatrix}$$

Appendix B

Simulation Results of Sliding Mode Controller

Simulations are performed to examine the system behaviour for different plant parameters. The chattering phenomenon are also discussed here.

B.1 Different plant parameters

Figure B.1 shows the simulation result of the system response for a nominal thermal model, $G_1(s)$, where $k_1 = \hat{k}_1$ and $\tau_1 = \hat{\tau}_1$, and three different thickness models, $G_2(s)$, such that the plant parameters are within a 10 % uncertainty range, i.e. $0.9\hat{k}_2 \le k_2 \le 1.1\hat{k}_2$ and $0.9\hat{\tau}_2 \le \tau_2 \le 1.1\hat{\tau}_2$. The reference is chosen to be a negative step such that $y_d(t) = -2.5$ nm $\forall t > 0$. The first control move is made at $t = 15$ s. Among all the thickness models within the 10% uncertainty region, the thickness model with the parameters, $k_2 = 0.9\hat{k}_2$ and $\tau_2 = 1.1\hat{\tau}_2$, has the smallest static gain and the slowest dynamics. Therefore, to ensure that all the plants will reach the sliding mode surface within a finite time, t_σ, this set of parameter is chosen in the calculation of the temperature limits, M, in Equation (4.13). By specifying $t_\sigma = 35$ s after the first control move and considering the parameters of

the thickness model to be $k_2 = 0.9\hat{k}_2$ and $\tau_2 = 1.1\hat{\tau}_2$, $M = 2.2$ is obtained from Equation (4.13).

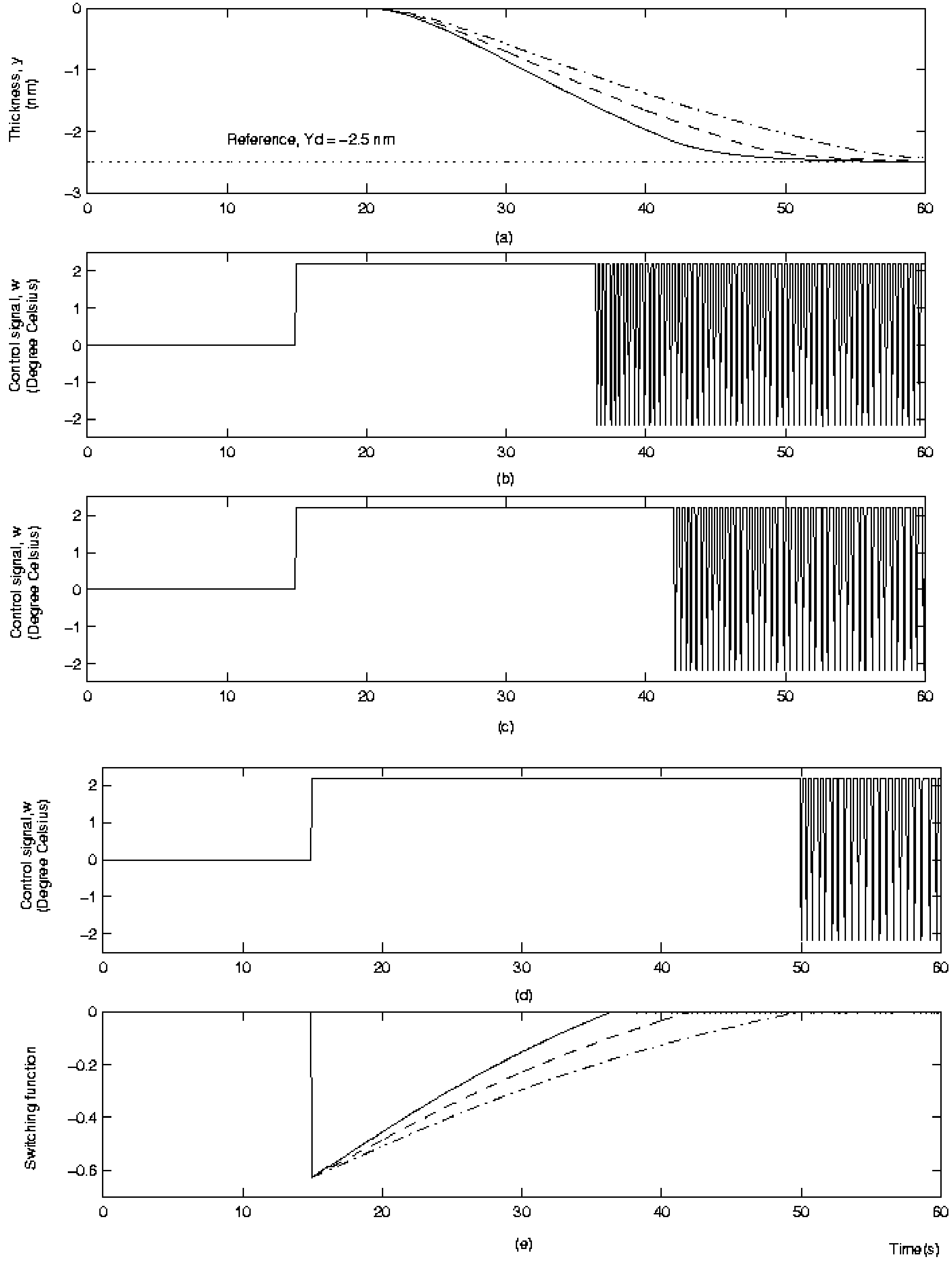


Figure B.1: Simulation results for different k_2 and τ_2 parameters.

Figure B.1 (a) shows the decrease in the resist thickness, y in response to the control signal, w. The solid lines in Figures B.1 (a) and (e) show the simulated change in resist thickness and switching function for the plant parameters, $k_2 = 1.1\hat{k}_2$ and $\tau_2 = 0.9\hat{\tau}_2$ respectively. Figures B.1 (b) is the corresponding control signal. The dashed lines in Figures B.1 (a) and (e) show the simulated change in resist thickness and switching function for the plant parameter, $k_2 = \hat{k}_2$ and $\tau_2 = \hat{\tau}_2$ respectively. The corresponding control signal is shown in Figure B.1 (c). For the plant parameters, $k_2 = 0.9\hat{k}_2$ and $\tau_2 = 1.1\hat{\tau}_2$, the simulated change in resist thickness and switching function is represented by the dashed-dotted lines in Figures B.1 (a) and (e) respectively. The control signal is shown in Figure B.1 (d).

Note that for the plant parameters, $k_2 = 0.9\hat{k}_2$ and $\tau_2 = 1.1\hat{\tau}_2$, the sliding surface ($\sigma = 0$) is first reached at $t = 35$ s after the first control move or at time, $t = 50$ s. This is expected as the temperature limit, M, is obtained by assuming this set of parameters. After the sliding surface is first reached at $t = 50$ s, the sliding mode controller, w, switches infinitely fast between M and $-M$. For the plant parameters, $k_2 = 1.1\hat{k}_2$ and $\tau_2 = 0.9\hat{\tau}_2$, the resist thickness tracks the reference, $y_d(t)$, in the shortest time as this set of parameter has the largest gain and fastest dynamics. In this case, the sliding surface is first reached at $t = 36$ s. For the nominal thickness model where $k_2 = \hat{k}_2$ and $\tau_2 = \hat{\tau}_2$, the sliding mode is first reached at $t = 42$ s. Therefore, it can be seen that by specifying t_σ using $k_2 = 0.9\hat{k}_2$ and $\tau_2 = 1.1\hat{\tau}_2$ in the calculation of M, we can ensure that all the plants within

the uncertainty region will reach the sliding surface within the specified time, $t = 50s$ or $t_{\sigma} = 35s$.

B.2 Digital Implementation

In most sliding mode control application, chattering is a very common phenomenon. Chattering describes the undesirable system oscillations caused by system imperfections or digital implementation of the continuous sliding mode control. In this section, the impact of digital implementation of a continuous time controller is discussed. To implement a continuous control law digitally, a zero order hold (ZOH) is typically used such that $w(t) = w(k)$ for $kT_s \le t < (k+1)T_s$ where T_s is the sampling interval and $w(k)$ is the computed control signal at sampling instant, k.

Figure B.2 shows the response of the nominal thickness model when the continuous sliding mode control law is implemented digitally at the sampling interval of 1 s. Figure B.2(b) shows that when the continous sliding mode controller is implemented directly with a ZOH at $T_s = 1$s, control signal switches between $\pm M$ with finite frequency. Also, the switching function σ does not remain zero after the sliding mode surface is first reached at $t = 42$s. As a result, the resist thickness, y, will not be able to track the reference, $y_d(t)$ at the end of 60s. As shown in Figure B.2(a), there is a steady-state error of about 0.1 nm as a result of digitial implementation of the continous sliding mode control law. One approach to reduce the chattering phenomenon due to the digital implementation of the continous sliding mode control algorithm is to reduce the sampling interval, T_s. For our application, $T_s = 0.1$s is chosen.

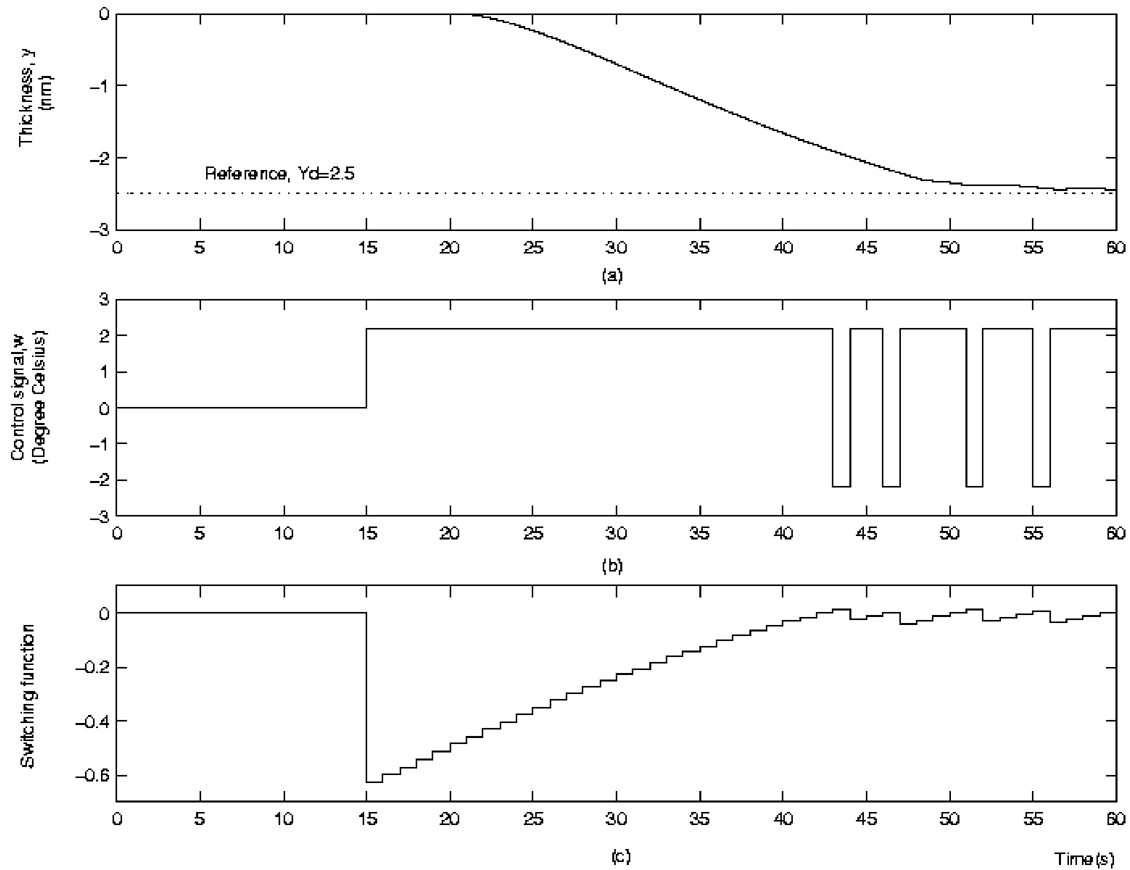


Figure B.2: Chattering due to digital implemetntation with $T_s = 1$ s.

Figure B.3 shows the simulation result for the nominal thickness model with $T_s = 0.1$s. In continous sliding mode control, the controller switches between the two control limits at infinite frequency to keep the system on the sliding surface. However, for digitial implemementation with the introduction of ZOH, this infinite switching is not achievable. Hence, it is no longer possible to achieve the result that an ideal sliding mode controller can achieve. One solution is to increase the sampling frequency so that the resulting signal will once again approach the ideal sliding mode control signal. As shown in Figure B.3, this can be achieved when $T_s = 0.1$ s is chosen.

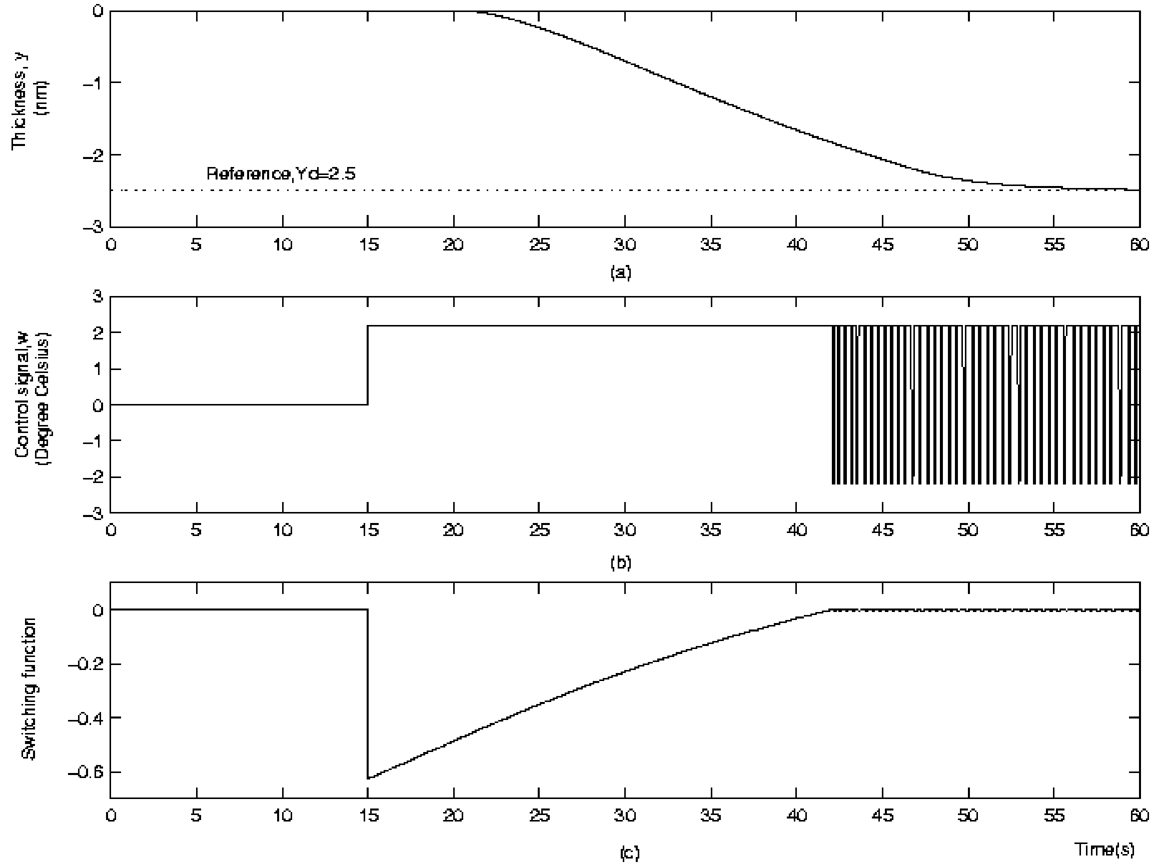


Figure B.3: Simulation of the system response with $T_s = 0.1$ s.

B.3 Unmodelled delay

For a time-delay system, a smith predictor is included to compensate for the dead-time in the thickness model so that the controller can be designed for a system without any time delay. The sliding mode control design assumes a complete cancellation of the delay using smith prediction. In this section, the effect of incomplete cancellation of the delay by the smith predictor is discussed.

Figure B.4 (a) shows the tracking of the resist thickness when there is an unmodeled delay of 1 s and 5 s. The solid line shows the resist thickness when there is a 1 s delay and

the dash line is the resist thickness when there is a 5 s delay. As a result of the unmodelled delay, the system response is oscillatory and the amplitude of the oscillation increase with the unmodelled delay. The control signal, w, for an unmodeled a delay of 1 s and 5 s, is shown in Figures B.4(b) and (c) respectively. As shown in Figures B.4 (b) and (c), the controller is not able to switch between the two control limits of M and $-M$ at an infinite frequency like an ideal sliding mode controller.

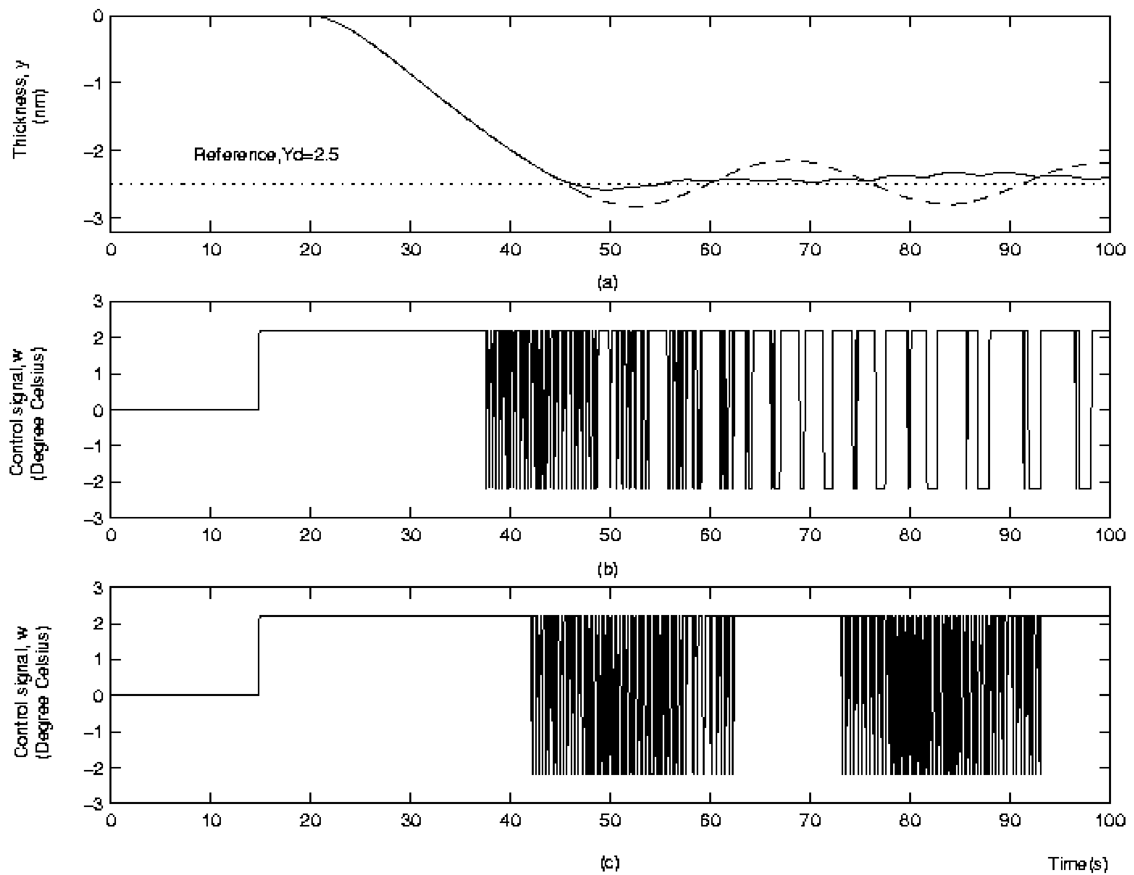


Figure B.4: Chattering due to unmodeled delay for $k_2 = 1.1\hat{k}_2$ and $\tau_2 = 0.9\hat{\tau}_2$.

Another source of chattering is due to unmodelled dynamics. An ideal sliding mode controller tries to achieve "perfect" performance in the presence of arbitrary parameter inaccuracies by switching between different control laws at a very fast speed, resulting in extremely high control activity. The high control activity may excite the neglected dynamics from the sensors or actuators and cause chattering phenomenon. To eliminate chattering, modification of the control laws has to be made to achieve effective trade-off between tracking performance and parametric uncertainty. However, in specific applications, particularly those involved in the control of electric motors and drives, the unmodified control laws has also been used directly. Given that tracking precision is important for our application, the ideal sliding mode control is implemented even though this might give rise to some chattering phenomenon.

Appendix C

Author's Publications

[1] W. K. Ho, A. Tay and L. L. Lee, "Constraint Fedforward Control for Thermal Processing of Quartz Photomasks in Microelectronics Manufacturing", *Journal of Process Control,* vol. 14, pp. 1-9, 2003.

[2] L. L. Lee, C. Schaper and W. K. Ho, "Real-time Control of Photoresist Thickness Uniformity During the Bake Process", *Proc. SPIE*, vol. 4182, pp. 54-64, 2000.

[3] L. L. Lee, C. Schaper, and W. K. Ho, "Real-time Predictive Control of Photoresist Film Thickness Uniformity", *IEEE Trans. Semicond. Manuf.*, vol. 15, no. 1, pp. 51-59, 2002.

[4] W. K. Ho, L. L. Lee and C. Schaper, "On Control of Resist Film Uniformity in the Microlithography Process," *15th IFAC World Congress*, Barcelona, 2002.

[5] W. K. Ho, L. L. Lee, A. Tay and C. Schaper, " Resist Film Uniformity in the Microlithography Process", *IEEE Trans. Semicond. Manuf.*, vol. 15, no. 3, pp. 323-330, 2002.

[6] W. K. Ho, A. Tay, L. L. Lee and C. Schaper, "On Control of Resist Film Uniformity in the Microlithography Process", submitted to *Control Engineering Practice* for publication.

[7] C. D. Schaper, K. El-Awady, T. Kailath, A. Tay, L. L. Lee, W. K. Ho and S. Fuller, "Processing Chemically Amplified Resists on Advanced Photomasks Using a Thermal Array", *Microelectronic Engineering,* vol. 71, pp. 63-68, 2004.