ABSTRACT

Title of Document: SENSOR BASED ATOMIC LAYER DEPOSITION FOR RAPID PROCESS LEARNING AND ENHANCED MANUFACURABILITY

Wei Lei, Doctor of Philosophy, 2006

Directed By: Professor Gary W. Rubloff
Department of Materials Science and Engineering and the Institute for Systems Research

In the search for sensor based atomic layer deposition (ALD) process to accelerate process learning and enhance manufacturability, we have explored new reactor designs and applied in-situ process sensing to W and HfO_2 ALD processes. A novel wafer scale ALD reactor, which features fast gas switching, good process sensing compatibility and significant similarity to the real manufacturing environment, is constructed. The reactor has a unique movable reactor cap design that allows two possible operation modes: (1) steady-state flow with alternating gas species; or (2) fill-and-pump-out cycling of each gas, accelerating the pump-out by lifting the cap to employ the large chamber volume as ballast. Downstream quadrupole mass spectrometry (QMS) sampling is applied for in-situ process sensing of tungsten ALD process. The QMS reveals essential surface reaction dynamics through real-time signals associated with byproduct generation as well as precursor introduction and depletion for each ALD half cycle, which are then used for process learning and optimization. More subtle interactions such as imperfect surface saturation and reactant dose interaction are also directly observed by QMS, indicating

that ALD process is more complicated than the suggested layer-by-layer growth. By integrating in real-time the byproduct QMS signals over each exposure and plotting it against process cycle number, the deposition kinetics on the wafer is directly measured. For continuous ALD runs, the total integrated byproduct QMS signal in each ALD run is also linear to ALD film thickness, and therefore can be used for ALD film thickness metrology.

The in-situ process sensing is also applied to HfO_2 ALD process that is carried out in a furnace type ALD reactor. Precursor dose end-point control is applied to precisely control the precursor dose in each half cycle. Multiple process sensors, including quartz crystal microbalance (QCM) and QMS are used to provide real time process information. The sensing results confirm the proposed surface reaction path and once again reveal the complexity of ALD processes.

The impact of this work includes: (1) It explores new ALD reactor designs which enable the implementation of in-situ process sensors for rapid process learning and enhanced manufacturability; (2) It demonstrates in the first time that in-situ QMS can reveal detailed process dynamics and film growth kinetics in wafer-scale ALD process, and thus can be used for ALD film thickness metrology. (3) Based on results from two different processes carried out in two different reactors, it is clear that ALD is a more complicated process than normally believed or advertised, but real-time observation of the operational chemistries in ALD by in-situ sensors provides critical insight to the process and the basis for more effective process control for ALD applications.

SENSOR BASED ATOMIC LAYER DEPOSITION FOR RAPID PROCESS LEARNING AND ENHANCED MANUFACTURABILITY

By

Wei Lei

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Professor Gary W. Rubloff, Chair
Professor Raymond A. Adomaitis
Professor Ichiro Takeuchi
Professor Lourdes G. Salamanca-Riba
Professor John Cumings

Dedication

To my father Yuanzhi Lei, my mother Qifeng Li, my wife Mei Yang and my son Terrence

Acknowledgements

I would like to acknowledge Mr. Laurent Henn-Lecordier for his contribution to the W and HfO_2 ALD work. I am also thankful for his patient training and continuous help throughout my research.

I would like to acknowledge Dr. Mariano Anderle and his group in ITC-irst, Italy for their characterization work and delightful discussions. I am also grateful to Dr. Louis Frees and Dr. Robert Ellefson of Inficon for their technical support.

I am especially grateful to my advisor Dr. Gary W. Rubloff. He gave me the chance to work in an area I really like, provided me with great working environment and guided me throughout my study in his group. I also want to thank my previous and current group members: Yijun Liu, Soon Cho, Jae-ouk Choo, Jung Jin Park, Yuhong Cai, Ramaswamy Screenivasan and Erin Robertson for their help in my work.

This work has been partially supported by the National Institute of Standards and Technology Chemical Science and Technology Laboratory (60NANB2D0094), the National Science Foundation (DMR0231291), and the Autonomous Province of Trento, Italy through MicroCombi project.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Atomic layer deposition

Atomic layer deposition (ALD) is a gas phase deposition technique that was first developed in the 1970s for electroluminescent flat panel displays.[1] Now it has emerged as the prime candidate for depositing critical ultra-thin layers in semiconductor manufacturing, including metallic diffusion barrier layers, high K dielectrics and optoelectronics materials.[2,3] ALD holds similar promise in other technology frontier areas such as micro-electro-mechanical systems (MEMS) and bio-technology.[3-5] The power of ALD comes from the self-limiting nature of ALD deposition reaction: unlike chemical vapor deposition (CVD), reactants are supplied to the reactor in alternating sequences during the ALD process, rather than simultaneously as in CVD. As each reaction between reactant and substrate surface is self-limiting, in principle only one monolayer of film will grow in one process cycle. The self-limiting character of ALD process produces excellent conformality, precise thickness and high uniformity, features which have stimulated intense efforts to master ALD for manufacturing.

In order to maintain its self-limited behavior, it is important to fully separate different reactant exposures in ALD process so no reactant mix could happen and lead to CVD reactions. This is done by introducing a purge step between each reactant exposure to remove un-reacted reactant before another exposure. A typical ALD process flow is illustrated in figure 1.1, where a nondescript metal oxide (MO_2) ALD process is used as an example. In order to initiate ALD reactions, the substrate surface normally is chemically treated to have the appropriate reactive surface groups.[6-8] During the ALD

process, the metal precursor ML_4 is first introduced into the reactor. It reacts with the reactive surface groups (e.g. –OH) and covers the substrate with one monolayer of product -ML_3. Once the surface reaction is saturated, the remaining precursor in gas phase can't further react with the substrate. The remaining gas phase precursor and product HL are then pumped out of the reactor in the following purge step. After that, reactant H_2O is introduced into the reactor and reacts with -ML_3 on the substrate surface. This again is a self-limiting surface reaction and only one monolayer of film (MO_2) is deposited after that. The remaining H_2O molecules and product HL are also removed in the following purge step. The two reactant exposure steps and two purge steps constitute one ALD process cycle. As each reaction between reactant and the substrate surface is self-limiting, in principle only one monolayer of film will grow for every process cycle. The process flow is then repeated until the targeted film thickness is reached. Attention should be paid on the following aspects to avoid any non-ALD film growth: (a) proper reactant exposure, as insufficient exposure may cause non-saturated adsorption and overexposure may cause physisorbed reactant layers; (b) sufficient purging, as incomplete purge causes non-self-limited reactions of both reactants and results in CVD-like growth; and (c) proper deposition temperature, as too low deposition temperature may cause physisorbed reactant molecules and too high temperature may cause excessive reactant decomposition and deleterious desorption.

Given the stringent standards for quality and throughput in semiconductor manufacturing, the small growth rate inherent from the self-limiting reaction mechanism also poses serious challenges for ALD process to be implemented into manufacturing.[9,10] In addition to the small growth rate, a variety of effects-steric hindrance, substrate surface

status, temperature-dependent reactant adsorption/reaction, reactant decomposition etc.- complicate ALD growth and limit its application.[11-13] Ideally there should be a linear relationship between ALD film thickness and ALD cycle number, but in practice the film thickness increase per cycle is affected by all the effects as listed above and may vary in different film growth stages because the substrate characteristics are changed by depositing new materials[14,15], as shown in figure 1.2. In most cases, the substrate surface is not completely homogeneous and some surface groups are more reactive than others. Consequently, ALD film growth will tend to start from those more reactive surface groups and form isolated islands instead of continuous film at the beginning, as shown in figure 1.3 (a). The islands will expand their size as ALD process proceeds, and the island number can also increase as less reactive surface groups participate in the reaction. Eventually these islands will coalesce with each other and form a continuous film, as shown in figure 1.3 (b). And generally the film morphology is not changed after the continuous film is formed.

The period where ALD film growth is in island growth mode is therefore called nucleation region (or incubation region). In the nucleation region, the growth rate normally keeps increasing because more reactant can participate in the reaction as islands grow with more reactive surface groups available. After continuous film is formed, the growth rate becomes constant because the film surface area is constant (and the number of reaction groups is constant). This period therefore is referred as linear growth region. The length of the nucleation region strongly depends on the substrate surface condition, and its corresponding growth rate is normally less than one monolayer per cycle. But even in the linear growth region, because the growth rate can be affected by a number of

factors such as steric hindrance and temperature, in most cases the actual growth rate is different from one monolayer per cycle. The nucleation region does not always have a smaller growth rate than the linear growth region. For some ALD processes, such as $Co(acac)_3/O_2$ on silica[16] and the plasma enhanced $TiCl_4/H_2$ on platinum[17], the substrate can actually serve as catalyst and the growth rate is higher when the substrate is not fully covered by the deposited materials. No matter whether the growth rate is inhibited or enhanced by the substrate, the non-constant growth rate is not desired for ALD processes.

Although the ALD reaction is fairly complicated on the atomic scale, its overall self-limiting character still grants it with much better conformality, uniformity and thickness control ability than other conventional thin film processes like CVD and PVD. To better understand and control ALD processes, one need to understand the detailed mechanism during each of its four steps, which requires extensive work on surface chemistry, reaction kinetics, mass transportation and film characterization.

1.2 In-situ ALD process sensing

In order to use the inherent advantages of ALD processes, the process conditions need to be carefully optimized to avoid non-self-limited film growth. In-situ observation of the ALD process enables both fundamental understanding and real time monitoring of ALD process, and will benefit ALD process significantly. A number of in-situ characterization methods have been applied to study the process chemistry and reaction mechanism in ALD processes. Quartz crystal microbalance (QCM) is probably the most widely used method so far[2,5,17,18]. It is sensitive enough to detect adsorption/reaction during ALD half-cycles by measuring the mass change and associated acoustic resonant frequency of a quartz crystal. When a mass is deposited on the quartz crystal, the

resonant frequency changes as $\Delta f = -C\Delta m$, where Δf is the change in the resonating frequency of the quartz crystal, Δm is the mass change and C is a crystal dependent constant. QCM can be used to monitor film growth rate and thickness in real time and it is very easy to use.

However QCM measurements also have some intrinsic limitations: the resonating frequency of the quartz crystal depends not only on the mass change but also on the viscosity of the gas, stress in the forming film, pressure and temperature. When only thin films are deposited, all but the temperature effect can be neglected. So it is very important for QCM to operate under constant temperature throughout the whole ALD process. The absolute limit for the QCM measurement is the Curie point of quartz (573ºC), where the piezoelectric properties are lost. In addition to the temperature limitation, QCM measures the mass change on the quartz, and film deposition on the quartz may be different from film deposition on the wafer surface because surface status are different, especially in the nucleation region. Despite its easy use in furnace type reactors where the quartz can be kept at the same temperature as the substrate, QCM is difficult to implement in cold wall substrate heating ALD reactors usually found in manufacturing environment. It also can't provide any chemically specific information about the surface reaction. Other diagnostic methods, which employs optical, surface analytical, and electrical measurements, can also be used in-situ in ALD process, including Fourier transform infrared (FTIR) spectroscopy[19], Auger electron spectroscopy (AES)[20,21], reflectance difference spectroscopy (RDS)[22], surface photo-absorption (SPA)[23] and electrical resistivity measurement[24]. However, these techniques have also

been limited to a research context, and they pose serious problems for integration into manufacturing.

Quadrupole mass spectrometry (QMS) is a widely used gas phase chemical sensor. By providing quantitative, chemically specific information, QMS can provide extensive real-time information regarding process chemistry, reactor condition and wafer state metrology. QMS got its name because it uses a quadrupole ion analyzer. Quadrupole is a device that is made up of four rods of circular or, ideally, hyperbolic cross-section which are perfectly parallel to each other.[25] It uses the stability of the trajectories in oscillating electric field to separate ions according to their m/z ratio (where m is the ion mass and z is the ion charge). By using quadrupole ion analyzer, QMS has a compact structure, relative low cost (about 30-40K US$) and reasonably high resolution (the "unit resolution with 15% valley" is less than 1 atomic mass unit). It is readily integrated with manufacturing tools, and it already serves widely for contamination control, fault detection, and advanced process control applications in semiconductor manufacturing. Quantitative use of QMS in chemical processes has led to demonstrations of thickness/deposition rate metrology with precision of order 1% and to associated end-point control in a variety of semiconductor manufacturing processes[26-28], including thermal and plasma-enhanced chemical vapor deposition.

QMS has also been proven of substantial value in tracking ALD process chemistry for a significant variety of ALD processes, such as Al_2O_3, TiN, Ti(Al)N and others[29-36]. One major challenge faced by the QMS in-situ sensing of ALD processes is the QMS response time: as the exposure time is short (in the order of seconds) and gas switching is frequent in ALD processes, the QMS setup needs to be well configured to

minimize its gas residence time. The low reactant conversion rate in ALD processes also requires an optimized QMS setup to achieve high sensitivity. In addition to that, reactant/product fragmentation reactions can be complicated in some ALD processes, especially when metal organic precursors are used, and the interpretation of mass spectra may also be complicated.

1.3 ALD reactor and manufacturability

ALD reactor design plays an important role in the enhancement of ALD manufacturability. One key requirement for ALD reactors is the small reactor size. Because gas switching time, which is shorter in reactors with smaller volume, comprises a significant portion of ALD cycle time, the use of small reactor size can significantly reduce the ALD cycle time and improve the manufacturing throughput.

ALD reactor operating pressure is another important factor. Based on the working pressure and flow conditions, there are basically two types of ALD reactors: molecular flow type reactors[19,37,38] and viscous flow type reactors[18,39,40]. Molecular flow type ALD reactors work under high vacuum environment, and exhaust the reactor to high vacuum to separate different reactant exposures. The un-reacted reactant and reaction products are pumped out during high vacuum periods under molecular flow condition. Molecular flow type ALD reactors usually have complicated structures and a big reactor volume due to the use of load-lock chamber and high vacuum pumping system. In addition to that, pumping under molecular flow condition is slow, so the throughput is usually an issue for molecular flow type reactors. However, the high vacuum condition is preferred for in-situ process sensing, especially for many in-situ process sensors such as AES, the high vacuum environment is required. Viscous flow type ALD reactors use periods of inert gas

purging under viscous flow conditions to separate different precursor exposures. The inert gas carries the reactant/product molecules out of the process chamber in a form of wave-fronts so reactant exposures can be spaced more closely together in time. Therefore gas switching is more efficient under viscous flow conditions with significant carrier gas (purge gas) flow. Although the viscous flow condition means higher pressure (normally more than 1 torr) and limits the application of many in-situ process sensors, most ALD reactors used in both research and manufacturing environment are operating under viscous flow conditions, as the throughput is the primary concern in most cases.

Based on the heating method, ALD reactors generally can be divided into two types: furnace type hot wall reactors or substrate heated cold wall reactors. Most viscous flow ALD reactors used in research environment are hot-wall furnace type reactors. The advantages of viscous flow reactors with furnace heating are the simpler design, lower cost and shorter cycle time. However, furnace type reactors generally are not suitable for investigating ALD process issues faced in manufacturing environment. There are two primary reasons: first, substrate heated reactors are dominant in main-stream semiconductor thin-film manufacturing tools, and second, most furnace type reactors in research environment use small size substrates (e.g. 1 $inch^2$) because the reactor volume and the furnace size will be substantially increased if wafer-scale substrates are used, which not only slows down gas switching but also increases the reactor cost. For substrate heating production scale tools working in viscous flow region, their advantages include large wafer size (200mm to 300 mm in diameter), cluster tool integration, small wall to wafer surface ratio (preferred for QMS sensing) and large throughput. But such

tools are also complicated, expensive and difficult to maintain in university research environment.

1.4 Motivation

ALD has recently become an important player in semiconductor manufacturing. As discussed in section 1.1, although in principle ALD process provides a layer-by-layer growth with good control ability, in practice it is considerably more complicated. The successful application of in-situ sensing will accelerate ALD process learning and benefit process optimization. So far most ALD in-situ sensing work is limited to research context, but both research and wafer-scale manufacturing can benefit if proper sensors and sensing strategy are used. Therefore the primary motivation of this work is the direct observation of ALD surface reaction during dynamic process cycles for rapid process learning, and extending the real-time sensing, metrology, and process control benefits of QMS to wafer scale manufacturing.

The implementation of in-situ process sensing will be facilitated if the reactor is properly designed. Sensor based ALD reactor designs that have good process control ability and fast gas switching can not only apply in-situ sensing for process learning, but also enhance ALD process manufacturability directly. So exploring ALD reactor designs for better in-situ sensor integration and manufacturability enhancement is another motivation of this work.

1.5 Research objectives, challenges and contributions

The ultimate goal of this work is to develop an in-situ ALD process sensing strategy for rapid process learning and enhanced manufacturability. The work has three

components. (1) We have designed, built, and operated a novel wafer scale ALD reactor (called UHV-ALD reactor), which features fast gas switching and good process sensing compatibility; by employing full wafers, a small reaction volume, and load-locked entry, this tool bears significant similarity to the real manufacturing environment. (2) We have applied the UHV-ALD reactor to W ALD process and used QMS to observe process dynamics and film growth kinetics in real-time, and we employ this information for process learning, process optimization and film thickness metrology. (3) We have also constructed a different, more common ALD reactor based on a hot wall furnace (called furnace-ALD reactor), and employed in-situ sensing using QMS and QCM, with the goal of advanced understanding and control of HfO_2 ALD process.

Although significant effort has been devoted to in-situ sensing in ALD processes, their application to manufacturing has not been explicitly addressed. While QMS has been used in CVD manufacturing research and applications, ALD is more challenging because of the short exposure cycles for reactant gases, the smaller signals to be expected for byproduct generation at the wafer surface, and the presence of corresponding parasitic reactions at the walls.

The impact of this work is threefold. (1) It develops new ALD reactor designs which enable the implementation of in-situ process sensors for rapid process learning and enhanced manufacturability. (2) It demonstrates in the first time that in-situ QMS can reveal detailed process dynamics and film growth kinetics in wafer-scale ALD process, and thus can be used for ALD film thickness metrology. (3) Based on results from two different processes carried out in two different reactors, it is clear that ALD is a more complicated process than normally believed or advertised, but real-time observation of

the operational chemistries in ALD by in-situ sensors provides critical insight to the process and the basis for more effective applications of ALD.

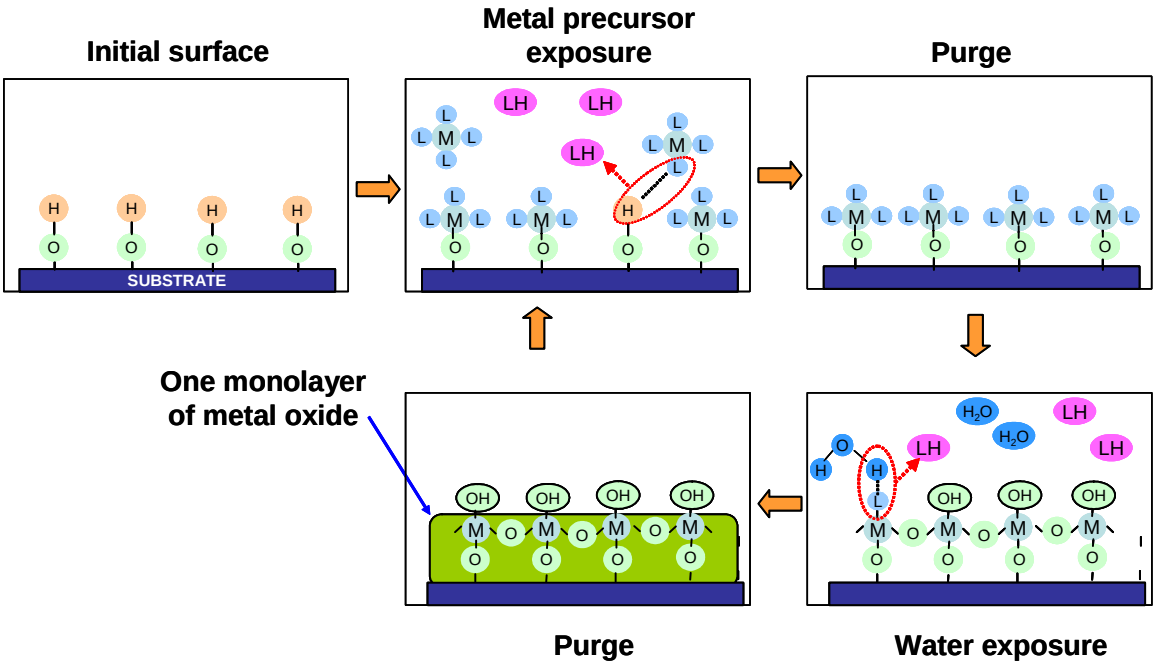


Figure 1.1 ALD process schematic

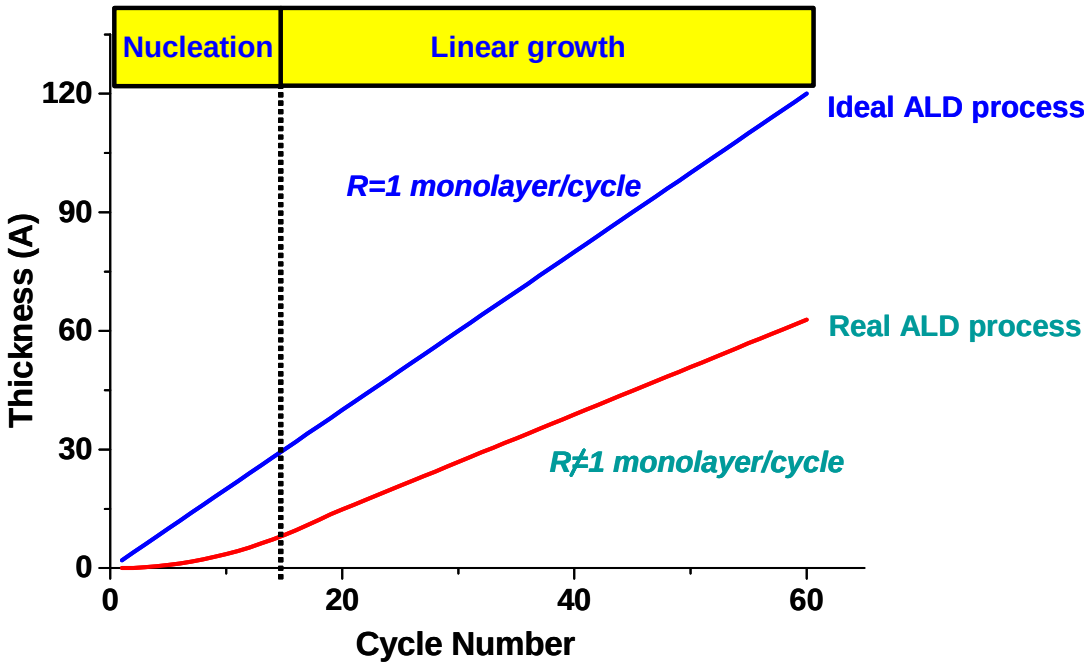


Figure 1.2 Ideal and actual ALD process curves

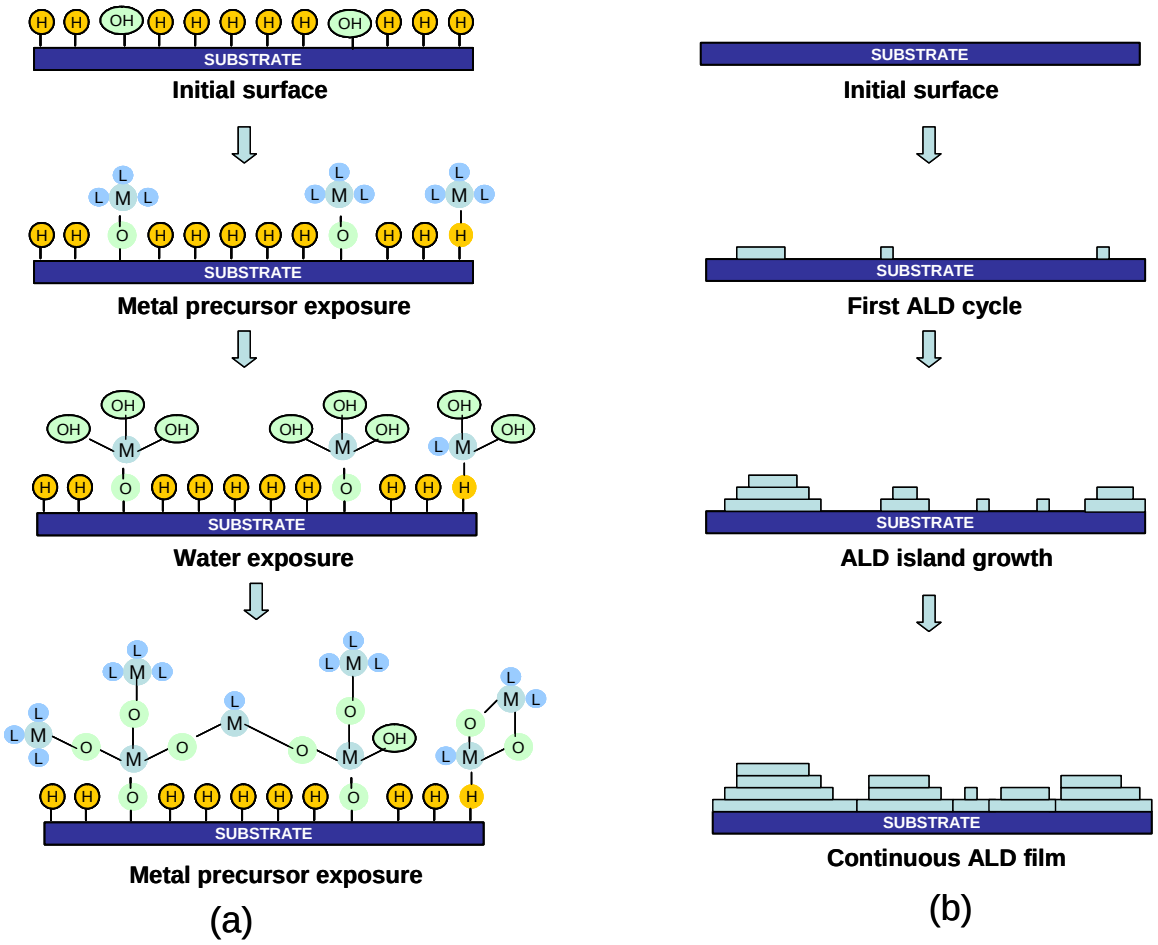


Figure 1.3 Schematic of ALD island growth: (a) as atomic scale; (b) as mesoscopic scale

Chapter 2 Novel wafer scale ALD reactor design

2.1 Overview

In order to develop in-situ process sensing for rapid process learning and enhanced manufacturability, a novel wafer-scale ALD reactor (UHV-ALD reactor) is constructed to simulate manufacturing environment with good process sensing compatibility. Figure 2.1 shows the schematic and the picture of this reactor. It is composed of two standard stainless steel 8 inch 6-way compact flange (CF) crosses, which are used as the load-lock chamber and the UHV chamber respectively. The UHV chamber is supported by a turbopump and a backing mechanical pump, and has a base pressure of 10^{-9} torr. The load-lock serves as a buffer chamber between the UHV chamber and the outside environment. A pneumatic gate valve is used to separate the UHV chamber and the load-lock chamber. Wafers are loaded to the load-lock chamber first before they are transferred to the UHV chamber by a linear motion transfer arm. During wafer unloading, wafers are also transferred to the load-lock chamber first before it is moved out of the reactor system. Therefore, the UHV chamber will not be directly exposed to the outside environment during operation and a clean process environment is guaranteed. The high vacuum condition in the UHV chamber would also make it possible to use various surface analysis techniques such as AES in situ in the future.

This load-locked UHV system provides solutions for the clean process environment requirement, but it also adds restrictions on the UHV chamber size. The use of gate valves and a turbo pump on the UHV chamber makes it difficult to reduce the UHV chamber size; the UHV chamber also should have enough space for the transfer arm to move. By using some non-standard components such as a rectangular gate valve

and a robotic distribution center, the chamber size can be substantially reduced in manufacturing tools. But this approach is too expensive and too complicated for research tools.

A novel concept –embedded "mini-reactor"– is used to solve this problem. The embedded "mini-reactor" provides the small reactor volume required for ALD processes while the surrounding UHV chamber provides clean environment and enough space for the transfer arm to move. Most components used in this reactor are standard and are available from suppliers. All the pneumatically driven components are controlled by Labview™ program. Therefore the cost and the complexity of this system are affordable in university research environment. More details about the "mini-reactor" will be described in a later section.

2.2 Substrate heater

As main-stream thin film deposition tools in industry use substrate heating, a substrate heater is also used in this system. It is a customized substrate heater with a commercially available, enclosed heater element wound in a spiral groove located on a stainless steel holder. The heater schematic is shown in figure 2.2.

The reactor gas inlet and outlet are integrated on the substrate heater to carry gases from outside the reactor through stainless steel tubes to the heater and mini-reactor, and then the reverse for exhaust (Figure 2.3). Reactants and carrier gases enter the substrate heater holder from its bottom, flow across the substrate heater from the gas inlet opening and are pumped out from gas outlet located on the other side of the substrate heater. The gas inlet is opened along the round edge of the substrate heater. The opening

is about 1mm wide and 4.4 inch long. The gas outlet has exactly the same size. This configuration is expected to create a uniform gas flow under viscous flow conditions.

The substrate heater can use up to 4 inch diameter wafers. We uses 4 inch diameter wafers in which case they are supported by the edge of the substrate heater holder and not in direct contact with the heater element (shown in figure 2.2), so the wafer is heated primary by heat radiation. If a smaller size wafer is used, it lays directly on the heater element and the wafer is heated primarily by heat conduction. In the latter case, the heating efficiency is high but the temperature uniformity on the wafer may suffer as some parts of the wafer are in direct contact with the heater element while others are not. Radiation heating reduces the temperature non-uniformity across the wafer but the heating efficiency also decreases.

Thermocouples are attached to the center and the edge of the substrate heater holder. A 4 inch diameter SensArray^{TM} wafer instrumented with thermocouples is also used to measure temperature differences across the wafer and the between wafer and the thermocouple affixed to the heater holder. When the center thermocouple temperature, referred as process temperature, is 325ºC, the average wafer temperature is 350ºC with a temperature non-uniformity of better than 10%.

2.3 "Mini-reactor" design

A movable stainless cap is used to cover the substrate heater and form a well-sealed small volume "mini-reactor" inside the UHV chamber. The gas inlet and outlet are integrated on this specially designed substrate heater, so the "mini-reactor" can work under viscous flow condition, independent of the UHV chamber. The "mini-reactor" uses wafer-scale substrate heater (can hold up to 4 inch diameter wafer); has small reactor

volume (about 0.2 L) and can work under viscous flow condition, so it is similar to the manufacturing environment in many aspects. A detailed schematic of the UHV chamber and the "mini-reactor" is shown in figure 2.3 (a) and its picture is shown in figure 2.3 (b).

This size of the "mini-reactor" is only limited by the substrate heater and can be made very small, thus enables relatively rapid cycling of gases in the "mini-reactor". The wafer-scale small reactor also has a large "wafer surface area to reactor surface area" ratio that benefits in-situ QMS sensing as discussed in Chapter 4. For the novel "mini-reactor", its volume is about 0.2 L and the "wafer surface area to reactor surface area" ratio is about $1/3$.

The stainless steel cap is controlled by a pneumatic motion device located outside the UHV chamber. While not leak-tight in the vacuum technology sense, the actual sealing of the "mini-reactor" is achieved by using the pneumatic motion device to push the cap against the substrate heater. When the pressure of the "mini-reactor" is about 3 torr, the outside UHV chamber has a pressure of about 5×10^{-4} torr. The pressure of the "mini-reactor", referred as process pressure, can be controlled by changing the inlet gas flow rate. The process pressure is about 3 torr when the incoming gas flow is 100 sccm.

When the UHV-ALD reactor system is at idle status, the cap is lifted up so that the ALD "mini-reactor" is pumped by the UHV system. The UHV chamber provides a clean reactor condition to better control surface chemistry in the ALD process. In-situ QMS process sensing can also benefit from the UHV environment and the large "wafer surface area to reactor surface area" ratio. The small ALD mini-reactor has no constraint on the wafer loading/unloading: during wafer loading/unloading, the reactor cap is lifted high above the substrate heater and the wafer transfer arm can move freely.

2.4 Modes of operation

The UHV-ALD reactor can be used like a conventional viscous flow ALD reactor. In this mode, the reactor cap is closed all the time with gas simply entering and leaving the mini-reactor from its own gas delivery and pumping systems. The process schematic is shown in figure 2.4 (a). As the gas residence time is short in the ALD mini-reactor (about 0.16 sec with 100 sccm gas flow), a fast gas switching can be achieved under viscous flow conditions. In practice we have used this mode most frequently, and all the data presented here is got from this mode.

The UHV-ALD reactor can also be used in a new mode of operation as shown in figure 2.4 (b): fill-and-pumpout cycling of each gas. In this mode, the "ALD mini-reactor" exhaust line valve is closed at all time. During the exposure step, the reactor cap is in contact with the substrate heater and the reactant is supplied to the wafer surface through the "ALD mini-reactor" gas inlet. The gas inlet valve then is closed, and after the surface reaction saturates, the reactor cap is lifted up by the outside motion device. There are huge pressure and volume differences between the ALD mini-reactor and the outside UHV chamber: P(mini-reactor) ~3 torr, P(UHV chamber) ~10^{-5} torr, V(mini-reactor) ~ 0.2L, V(UHV chamber) ~25L, therefore all the gas species inside the ALD mini-reactor rapidly diffuse into the UHV chamber. After that, the reactor cap is lowered on the substrate heater for another reactant exposure. The main advantage of this "fill-and-pumpout cycling" operation mode is that the pumpout is accelerated by lifting the reactor cap to employ the large UHV chamber as ballast. But this operation mode doesn't work well for the in-situ QMS sensing. Under this operation mode, since the gas exhaust of ALD mini-reactor is closed and QMS is not directly attached to the mini-reactor (QMS is

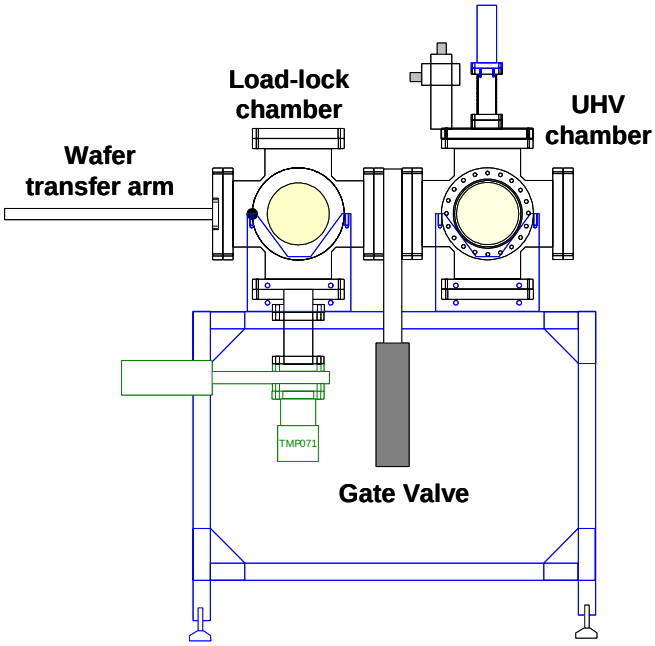
actually downstream to the mini-reactor as the mini-reactor is inside the UHV chamber), the gas phase change in the immediate wafer area is not readily reflected by the gas phase change in the gas sampling line. In addition to that, even though the reactor cap is pushed against the substrate heater, there is still gas leaking from the ALD mini-reactor to the UHV chamber. This can causes variations in the reactor pressure, which also affects the QMS sensing.

2.5 Implementation of in-situ mass spectrometry

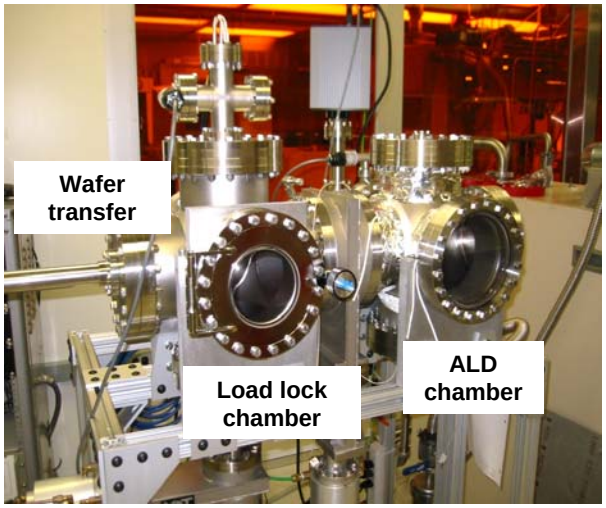
A differentially pumped quadrupole mass spectrometry (QMS) is setup downstream to the ALD reactor to directly observe ALD process details for the conventional viscous flow operation, as shown in figure 2.5. QMS has been readily integrated with manufacturing tools and successfully used for process monitoring, measuring deposition rates and establishing thickness metrology in a variety of semiconductor manufacturing processes.[26-28,41-44] Ideally QMS should be located as close as possible to the wafer surface in order to get the best sensing result. In the UHV-ALD reactor, however, it is not possible to directly attach QMS to the "ALD mini-reactor" because the "ALD mini-reactor" is located inside a UHV chamber. Therefore a downstream setup is used. As shown in figure 2.5, the QMS is located outside the UHV chamber but as close as possible to the "ALD mini-reactor" wherever space allows.

The working pressure of QMS is less than $1x10^{-5}$ torr but the process pressure of our ALD process is about 3 torr, so an orifice is needed to reduce the pressure from 3 torr to $1x10^{-5}$ torr. The orifice size can be calculated based on the QMS tuobopump pumping speed and the pressure reduction requirement. A 35 μm orifice is used in our setup: with this orifice, the QMS working pressure can be maintained at about $6x10^{-6}$ torr when the

process pressure is 3 torr. The orifice location should also be carefully chosen to reduce the existence of dead-volume. When the orifice is located at location B, the gas sampling line becomes a “dead-volume” because the whole gas delivery line is not directly pumped and the gas residence time inside it is significant, and the observed dynamics lags behind what is happening inside the “mini-reactor”. When the orifice is located at location A, the whole sampling line is under molecular flow condition, and the gas phase change in the “mini-reactor” will be readily reflected by the gas phase change inside the gas sampling line, so the “dead-volume” effect disappears. If the orifice is located anywhere between position A and position B, there will still be a “dead volume” in the gas sampling line, so position A is the best place to locate the orifice.



(a)



(b)

Figure 2.1 (a) schematic and (b) picture of the novel ALD reactor system

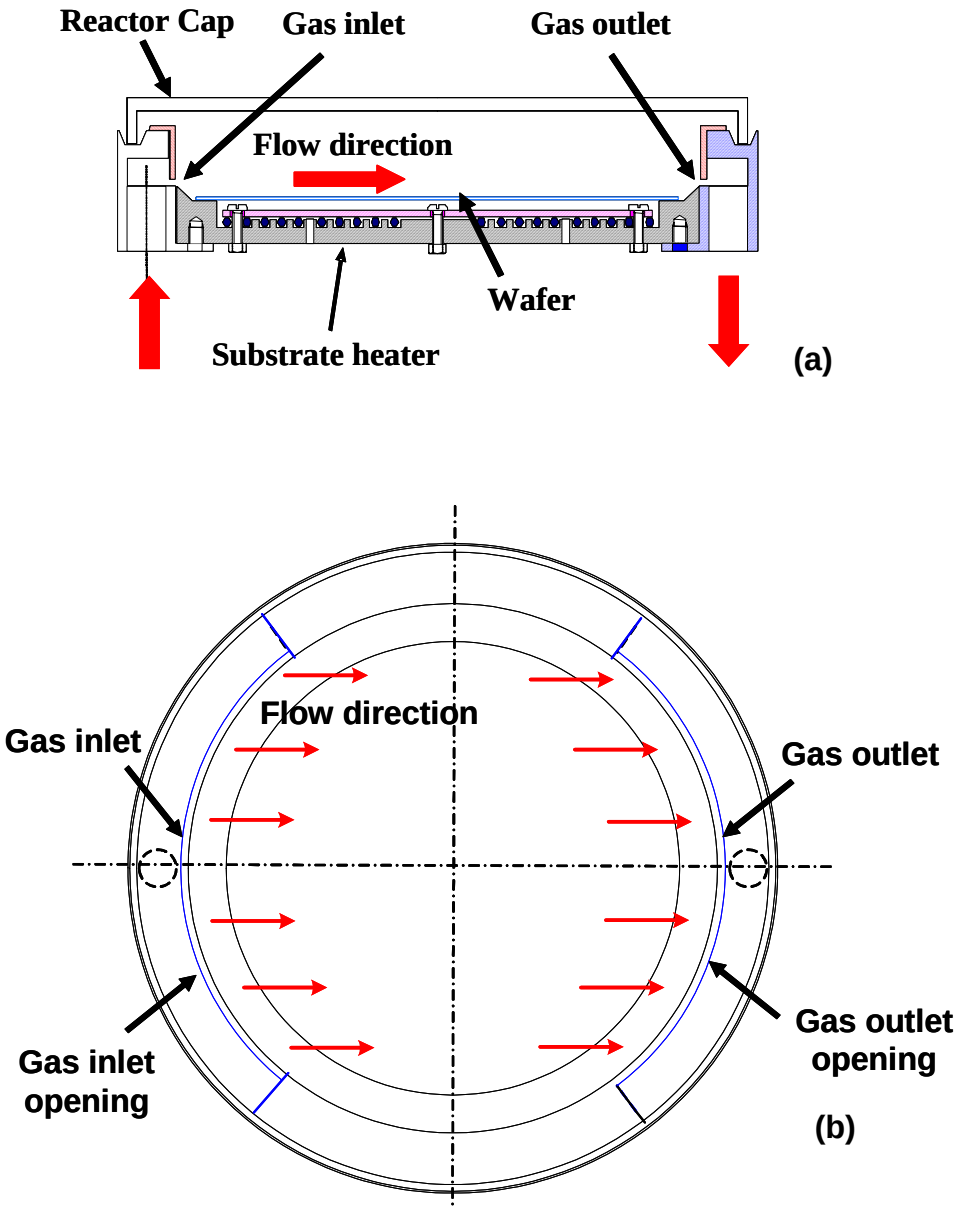
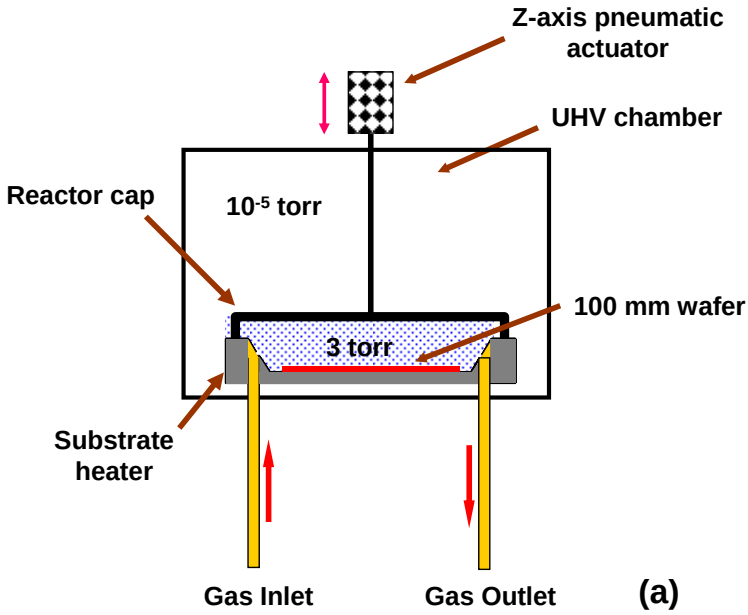


Figure 2.2 Substrate heater schematic: (a) sideview; (b) top-view



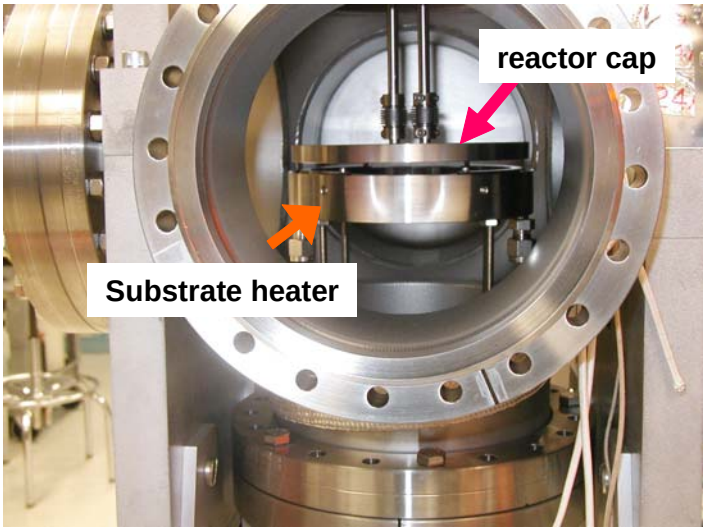


Figure 2.3 UHV chamber and ALD mini-reactor: (a) schematic; (b) picture

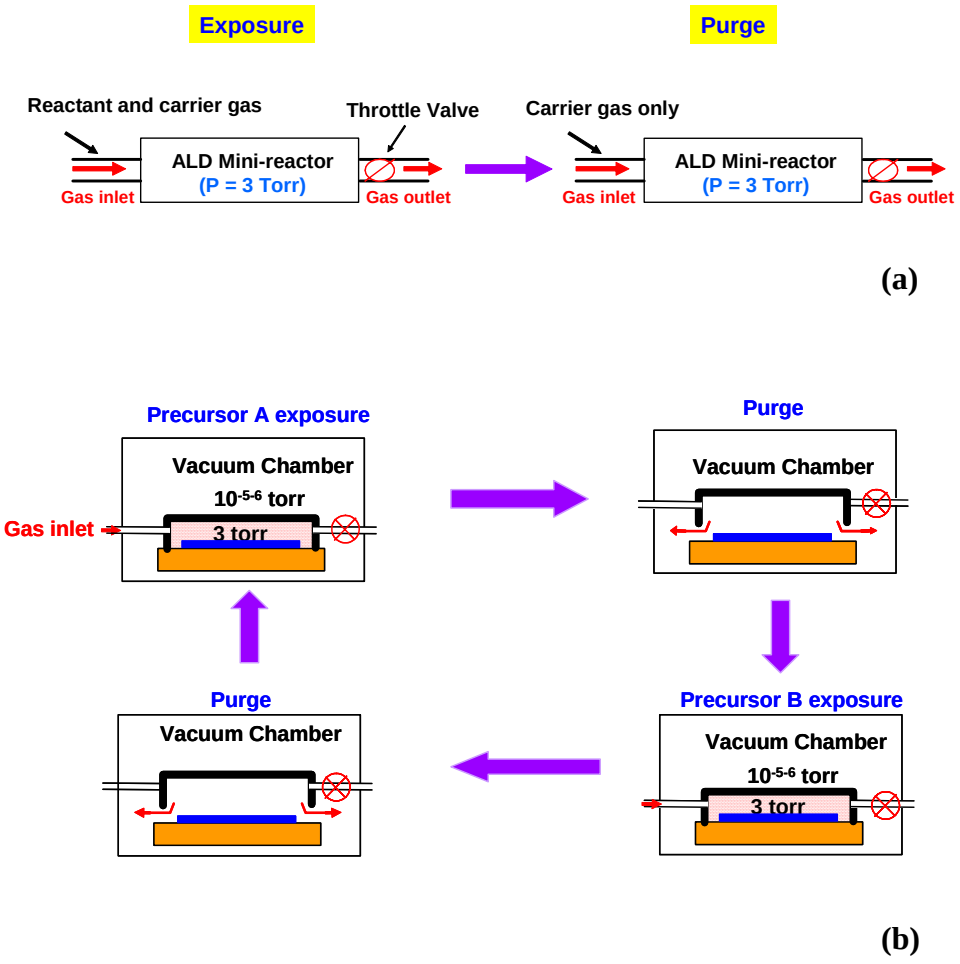


Figure 2.4 ALD process schematics: (a) viscous flow mode; (b) fill-and-pumpout gas cycling mode

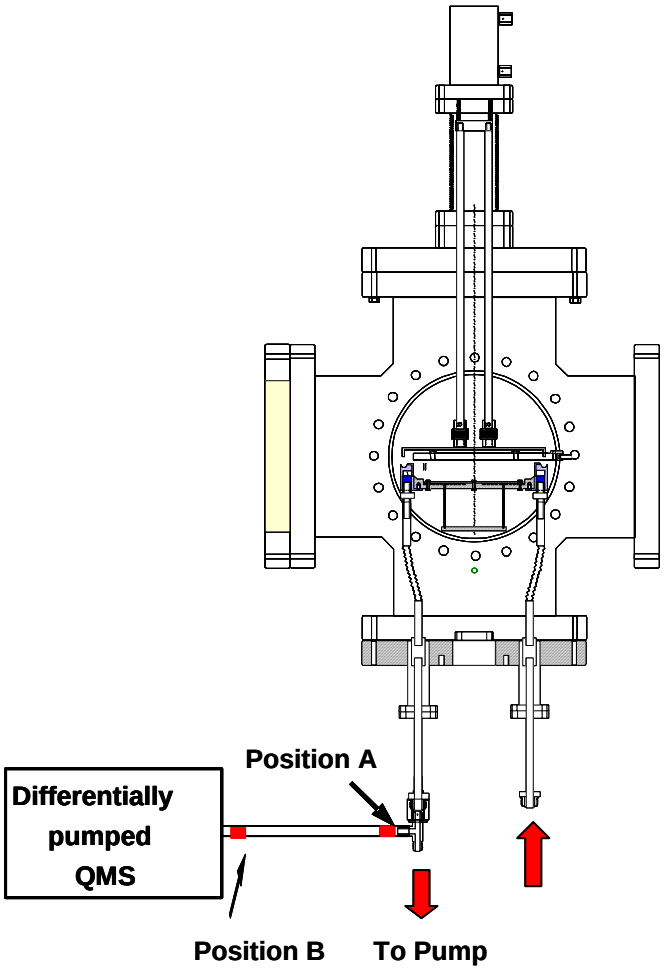


Figure 2.5 In-situ QMS setup

Chapter 3 Real time observation and optimization of W ALD process

The UHV-ALD reactor is constructed to simulate manufacturing environment with good sensing compatibility. It is integrated with a differentially pumped QMS system, and is used to study tungsten (W) ALD process dynamics. The results apply QMS to the UHV-ALD reactor, using reactant fluxes and exposure time compatible with the QMS response times (of order 1 sec). This strategy directly reveals the dynamics of the ALD surface reactions. From this, ALD process recipes are straightforwardly optimized. The results also indicate how process temperature and the relative gas doses for the two ALD reactants influence film growth on the wafer.

3.1 Experimental conditions

All W ALD experiments presented here were performed on 4 inch (100 mm) Si(100) wafers with 500 Å thermally grown SiO_2. The wafers were treated in dilute HF solution to provide a suitable surface condition for nucleating the W ALD reaction. HF solution is known to break Si-O bonds to form Si-H and Si-OH bonds[45], facilitating the nucleation. Wafers were dipped in $HF:H_2O$ = 1:25 solution for 30 seconds and then blown dry in N_2 before introducing them through the load lock. The etch rate of the dilute HF solution on thermally grown SiO_2 is about 100 Å per minute[46], so the as-prepared wafers for ALD studies had a somewhat thinner SiO_2 layer (about 450 Å) with exposed Si-H and Si-OH bonds on the surface.

Reactants used as ALD precursors are WF_6 and dilute SiH_4 (1% SiH_4 with 99% N_2). Research grade Ar gas is used as a carrier gas (99%) during WF_6 exposure and as the purge gas between reactant half-cycles, while the N_2 diluent in the SiH_4 source served as

its carrier. Thus both precursors have 1% concentration in their diluents. Wafers were first exposed to SiH_4 to deposit Si species before the first WF_6 exposure. The process temperature was 325ºC, which corresponds to an actual wafer temperature ≈350ºC. The default dynamic process cycle is given in Table 3.1. Unless otherwise specified, all the experiments reported here were conducted under these default process conditions.

Ex-situ film characterization included both electrical and compositional measurements for the ALD film. A custom 4-point probe resistance probe scanned the wafer surface to obtain a sheet resistance map, thus indicating both thickness and uniformity for the ALD film. For each 4 inch wafer, the map consisted of 608 measurements evenly distributed across the wafer surface.

Film chemical composition was measured by dynamic secondary ion mass spectrometry (SIMS) carried out at IRST, Italy. SIMS measurements were carried out using a Cameca SC Ultra at 1kev of Cs^+ primary beam impact energy and monitoring the secondary MCs^+ molecular ions (where M is the element of interest). This analytical method strongly reduces matrix effects associated to dynamic SIMS and allows quantitative depth profiles to be obtained.[47,48] The primary beam was rastered over an area of 250×250μm^2 and the secondary ions were collected from an inside area of 100×100μm^2. An electron beam was used to compensate charging effects which can occur at the insulating silicon oxide layer.

In order to evaluate the tungsten film thickness all the analyses were stopped at the tungsten/silicon dioxide interface, identified as the 50% decrease of the WCs^+ signal. Then the crater depths were measured by a mechanical profilometer (Tencor P15).

3.2 Results

3.2.1 Observation of ALD process dynamics

For W CVD process using WF_6 and SiH_4 at process temperatures below 400ºC, SiF_4 and H_2 are the main byproducts, following the overall reaction $3\ SiH_4 + 2\ WF_6 \rightarrow 2\ W + 3\ SiF_4 + 6\ H_2$, with very little HF generation[49,50]. One can imagine that ALD behaves in a chemically similar fashion as does CVD, except that the state of the reaction is maintained the same at all surface sites. Therefore, we expected SiF_4 and H_2 as primary byproducts in W ALD from WF_6 and SiH_4. Indeed, previous W ALD studies using WF_6 and Si_2H_6 also revealed H_2 and SiF_4 as reaction byproducts[20,35].

Dynamic mass spectrometry results through a single W ALD cycle are shown in Fig. 3.1, obtained after previous cycles to establish a steady-state pattern. For low wafer temperature (60ºC) as shown in Fig. 3.1 (a), little surface reaction is seen for either the SiH_4 or WF_6 exposures. The SiH_4 reactant is seen clearly, with an accompanying H_2 signal as expected from the molecular fragmentation which occurs as part of the ionization process in the mass spectrometer. During the WF_6 exposure cycle, a modest WF_6 signal is observed when data is plotted with log scale, along with a comparable amount of SiF_4 (although both species are almost invisible on linear scale). Consistent with our previous experience in mass spectrometric sensing of these processes, we attribute the existence of broad WF_6, SiF_4, and HF traces in Fig. 3a to wall reactions in the system, in which WF_6 is adsorbed on walls, reacts with SiH_4 there, and provides a slowly varying background.[26] Since both SiH_4 and WF_6 reactants are introduced at 1% concentrations with their carrier gases, the difference in signals between them is associated not only with different cross-sections for mass spectrometry, but also with

more subtle – but important – wall reactions which occur in such systems, at chamber walls and delivery tubes.

Upon increasing the process temperature to 325ºC, the spectra change profoundly, as seen in Fig. 3.1 (b). Reaction product signals for the half-cycles are now clearly seen, including H_2 product signal from SiH_4 exposure and SiF_4 product signal from WF_6 exposure. These signals are considerably larger than those attributed to low temperature wall reactions in Fig. 3a. Although a previous study showed that SiF_4 was produced when SiH_4 was exposed to a fluorinated tungsten surface,[50] we find only a very small SiF_4 signal peak during SiH_4 exposure in figure 3b, which is negligible compared to the SiF_4 peak that is generated by WF_6 exposure. This suggests that the majority of fluoride atoms are released during the WF_6 exposure, forming SiF_4, rather than during SiH_4 exposure under our process conditions. The signal level of another possible byproduct, HF, is found only at the noise level throughout the ALD cycle, indicating that little or no HF is generated in the surface reaction, consistent with previous studies.[50]

Figure 3.1 (b) also reveals the essential process dynamics of the W ALD process at 325°C. At the beginning of the SiH_4 exposure, the SiH_4 reactant signal is weak (at the noise level) due to depletion, while the H_2 reaction product signal increases rapidly as it is released by surface reaction. The H_2 product signal increases to a plateau, then decreases sharply, consistent with saturation of the surface reaction. This self-limiting feature of the reaction half-cycle is precisely what distinguishes ALD. Analogous behavior is seen for the WF_6 half-cycle exposure, where the SiF_4 signal increases sharply upon WF_6 exposure, then decreases sharply as the surface reaction is completed.

Unfortunately the WF_6 signal is small, as discussed above, and cannot be seen on the linear scale of Fig. 3.1 (b), though it is more apparent on a log scale.

The SiH_4 precursor signal, seen in Fig. 3.1 (b) after completion of the surface reaction, decreases rapidly to noise level when the inert gas purge step begins. The response time of this signal (~2 s) demonstrates the fast gas switching achieved in the ALD reactor, which is reached because of the small volume (0.2 L) of the mini-reactor and the relative large gas flow (100 sccm).

3.2.2 W ALD film deposition

Figure 3.2 (a) demonstrates the linear relationship between film thickness measured by 4 point probe (described earlier) and the cycle number for W ALD process at 325°C. Following a nucleation regime of about 20 ALD process cycles, the deposition thickness is linear in the number of cycles, with more than one monolayer deposited in each cycle; this is consistent with other W ALD work under similar temperature region[51,52].

This W ALD process displays good uniformity (10% thickness variation) across 4 inch wafer, as shown in figure 3.2 (b). Spreading resistance wafer maps obtained with a 4 point probe indicate that the resistivity of W ALD film is 157 μΩ·cm. While this value is considerably larger than that for bulk W resistivity (5.29 μΩ·cm), it is close to the resistivity value reported for W ALD films in previous W ALD work (122 μΩ·cm)[19]. It is conceivable that this difference may be associated with microstructure and/or impurity effects.

3.2.3 ALD Process Cycle Time Optimization

Minimizing reactant exposure and purge time is critical in ALD to achieve sufficiently high throughput for manufacturing. The detailed process dynamics revealed by real-time QMS sensing can be used for ALD process cycle time optimization, since the self-limiting surface reaction is directly observed, as seen in Fig. 3.1(b).

Under the experimental conditions of Fig. 3.1(b), the surface reaction time for SiH_4 exposure is about 10 sec, while the surface reaction time for WF_6 exposure is about 8 seconds. Both of these are significantly less than the exposure times assigned in the process recipe in Fig. 3.1(b). Furthermore, the SiH_4 reactant signal decreases to noise level with about 3 sec during the purge, again smaller than the purge time used; we expect a similar situation for the WF_6 purge cycle. This means that the overall process cycle time can be substantially reduced, for both reactant exposure cycles and both purges, thereby increasing throughput. Reactant exposure times were decreased to 10 seconds for SiH_4 exposure and 8 seconds for WF_8 exposure for the QMS results shown in Fig. 3.3. This causes little change in the QMS signals for the H_2 and SiF_4 product signals (though of course the SiH_4 reactant signal is not visible after SiH_4 reaction saturation because the exposure cycle is then terminated).

3.2.4 ALD depletion

A primary character of ALD is its ability to achieve good conformality and uniformity, so that reactant depletion phenomena which limit traditional chemical vapor deposition processes are avoided. However, this character is achieved only when the self-limiting reaction characteristics of each half-cycle exposure are attained, which in

turn defines a minimum exposure dose of reactants (pressure x time) that must be met or exceeded at each point on the wafer. Our reactor geometry offers a path to observe this situation directly through across-wafer uniformity measurements carried out by post-process sheet resistance mapping.

Fig. 3.4 shows sheet resistance profiles across the 4" wafer along the gas flow direction for a 70 cycle W ALD deposition process at various exposure times, for constant flow rates, temperature, and pressure conditions. From Fig. 3.1 (b) we know that the minimum exposure times for surface saturation at these conditions are 10 sec SiH_4 exposure and 8 sec WF_6 exposure. The QMS results identify these as saturation exposures for the entire wafer (and indeed, surrounding areas at temperatures near to the wafer temperature). Accordingly, profiles for [SiH_4:WF_6] = 15sec:20sec and 10sec:8sec are quite flat (uniform) across the wafer in the gas flow direction. However, reducing these exposure times to [8sec:6sec] and [5sec:4sec] cause substantial thinning of the deposited W film in the downstream region of the wafer. This is because an insufficient number of reactant molecules (exposure dose) have been introduced to the reactor to achieve saturation across the entire surface, so that average deposition per cycle or average thickness after multiple cycles must be less than that expected for full saturation coverage on each half-cycle. Furthermore, since the 3 torr pressure regime is not fully in the molecular flow regime and the gas flows from one side of the wafer to another, the incomplete deposition is reflected as depletion across the wafer; i.e., the reactant in each half-cycle is preferentially consumed by the upstream side of the wafer, leaving inadequate reactant to achieve saturation exposures at the downstream side.

In the discussion of Fig. 3.1 (b) above, we noted that [SiH_4:WF_6] = 10sec:8sec exposures produced saturation coverage during the two half-cycles. This is confirmed by Fig. 3.4, which demonstrates that smaller exposures degrade the deposited film thickness. Clearly the QMS sensor measures integrated reaction product generation, and it appears to do so independent of details of gas transport, such as the flow across the wafer surface.

3.2.5 Precursor dose effects

The above results show that sufficient precursor exposure times for each reactant are needed to achieve surface reaction saturation if the benefits of ALD are to be reached. Accordingly, further shortening of process cycle time should be possible using higher reactant gas flows. This is illustrated in Fig. 3.5 (a), where increasing WF_6 flow rates at constant total pressure causes more rapid saturation of the surface reaction in WF_6 exposure, evidenced by the faster rise and fall of SiF_4 QMS reaction product signatures. Fig. 3.5 (a) also shows the corresponding behavior of H_2 product signals during the SiH_4 exposure. Strikingly, the H_2 product signal decreases with increasing WF_6 flow rate although no process changes are made in SiH_4 exposure. If the WF_6 exposure only produced a saturated coverage condition for that half-cycle, but accomplished this in a shorter time, then there should be no change in the behavior for the SiH_4 half-cycle represented in Fig. 3.5 (a), in contradiction to experiment.

The H_2 product data in Fig. 3.5 (a) demonstrates that more complex behavior occurs. Sheet resistance mapping reveals that with 5 sccm WF_6 flow rate or higher, W film uniformity degrades significantly and the average film thickness also decreases. Fig. 3.5 (b) shows the dependence of the integrated H_2 product signal generated by the SiH_4 exposure half-cycle as a function of the WF_6 flow rate, consistent with the results of Fig.

3.5 (a). In addition, Fig. 3.5 (b) demonstrates that for a 70 cycle W ALD process with all other process parameters constant, increasing WF_6 flow rate leads to decreasing film thickness, again a result in apparent contradiction to the simple picture of ALD. These surprising observations suggest that the ALD process may involve more complex dose-dependencies between the two exposure half-cycles involved in the ALD process.

While a credible chemical pathway to explain this is not certain at this time, one can imagine the higher WF_6 exposures could lead to multilayer coverage on the surface above that expected for single-layer saturation, and that the SiH_4 exposure results in Fig. 3.5 (a) and (b) are a consequence. Multilayer coverage involving e.g. the WF_6 exposure could depend on flow rate or concentration of the WF_6 dose at given temperature, as occurs for physisorbed or chemisorbed layers in steady-state equilibrium between adsorption and desorption in the case of desorption isotherms for simple surface behavior. For example, incident SiH_4 could lead to film formation and H_2 product generation at surface sites where only a single layer was produced by the WF_6 exposure, but that other sites with multilayer WF_6 might block the deposition reaction by taking different reaction paths with higher energy barriers and lead to desorption of other products. This picture is merely a suggestion of the kinds of complexity suggested by the data and by the possibility of imperfect self-limiting adsorption/reaction for an ALD half-cycle.

3.2.6 Temperature effects

Process dynamics revealed by QMS show a distinct temperature-dependence, as illustrated in Fig. 3.6. Higher temperatures cause increasing generation of reaction products as seen for both WF_6 and SiH_4 exposure half-cycles. Figure 3.7 (a) illustrates SIMS tungsten depth profiles for four W ALD samples grown with the same number of

ALD cycles but at different deposition temperatures. The data clearly shows an increase of film thickness with deposition temperature, verifying the in-situ QMS measurements. Fluorine profiles, not shown in the plot, has a considerable low level (< 5%) throughout the ALD film with a pile-up at the tungsten/oxide interface, indicating chemical residuals left on the substrate after the cleaning procedure before the ALD process. Increasing ALD growth rates with temperature have been previously reported in various ALD processes, including W ALD using WF_6 and Si_2H_6[52].

A natural explanation for this behavior is that the surface reactions which transform species and cause product desorption are thermally activated, leading to a notable temperature-dependence. In Fig. 3.7 (b) the ALD growth rates obtained from reaction product signals in QMS are used to derive thermal activation energies for the growth on Arrhenius plots, which are about 2.57 kcal/mol for the SiF_4 signal (WF_6 exposure) and 2.05 kcal/mol for the H_2 signal (SiH_4 exposure). The activation energy obtained from the SIMS data in Fig. 3.6 is about 3.04 kcal/mol.

Previous work[52] under molecular flow conditions showed that the WF_6 sticking coefficient on a Si_2H_6 saturated surface is dependent on surface temperature, with an activation energy of 3.1 kcal/mol in the temperature regime 200-300ºC. It has also been found that when Si_2H_6 exposure is large (greater than 300L, where 1 L = $1x10^{-6}$ torr-liter), the Si_2H_6 adsorption rate on the WF_6-saturated surface increases with increasing process temperature and the activation energy for Si_2H_6 adsorption is 2.6 kcal/mol. These results for temperature dependence of growth rate are generally consistent with our results.

On the other hand, more complex explanations may not be ruled out at this point: e.g., if the reaction half-cycles entail some multilayer adsorption, higher temperature

could volatilize the additional layers and recover chemical reactivity for a larger fraction of the surface, thereby enhancing deposition rates. If one were to presume the kinetics for SiH_4 and Si_2H_6 adsorption on a WF_6-saturated surface to be similar, then it may be tempting to attribute the QMS-derived activation energies to multilayer adsorption/reaction in support of this explanation.

3.3 Discussion

3.3.1 ALD process dynamics and kinetics

The novel UHV-ALD reactor employed here, and particularly its incorporation of real-time, in-situ chemical sensing by QMS, provide a revealing and fairly direct picture of ALD process dynamics. The surface reaction saturation phenomena for which ALD is so strongly pursued are revealed directly, allowing process recipe optimization as well as a number of other possibilities, e.g. advanced process control. Depletion effects across the wafer for limited exposure dose situations confirm the depletion phenomena which should be expected if each exposure half-cycle starves the surface with precursor supply short of saturation amounts.

3.3.2 Temperature-dependence

The temperature-dependence of ALD deposition rates observed here (Figs. 3.6, 3.7) is most readily attributed to activation energies for half-cycle adsorption/reaction. While this may be managed in manufacturing process equipment through highly uniform wafer temperatures, thermal non-uniformities will compromise the uniformity expected from an ideal ALD process. Indeed, the simplest picture of the ALD process envisions self-limiting surface adsorption/reaction which is fully transport limited, and where

adequate exposure times need simply be chosen, so that uniformity is fully assured by self-limiting surface adsorption/reaction. In other words, the temperature-dependence of ALD deposition demonstrates that while ALD offers many profound advantages, it is not a perfectly robust solution for materials deposition.

The activation energy during each ALD half cycle contributes differently to the total film growth activation energy. The total film growth during one ALD cycle can be approximately considered as the sum of the materials deposited during each of the two ALD half-cycles. If the surface reaction time is similar under different process temperatures, then the total film growth rate is influenced by both growth rates during each ALD half cycle. Assume the activation energy during the two reactant exposures are E_{a1} and E_{a2} correspondingly, then the total activation energy for film growth E_{a3} should have a value between E_{a1} and E_{a2}, and the total film growth rate is mostly affected by the slowest half-reaction.

3.3.3 Exposure dose dependence

The indications of dose-dependence and interaction between the complementary half-cycle doses (Figs. 3.5) represent another complexity of ALD. While some chemisorption systems are nearly ideal in achieving self-limiting adsorption, ALD chemistries may well involve multilayer adsorption and consequent reaction complexity with manifestations in dose interactions between the two exposure half-cycles. This is certainly a challenge to disentangle, and also a potential limitation in ALD manufacturability. On the other hand, the stochastic nature of reactant impingement guarantees that possible formation of multilayers will lead to such complexity: to achieve

complete coverage or reaction of one precursor, exposure must be carried out until every surface site has been reacted, or material defects (e.g., unoxidized metal atom) may result. The dependence of ALD results on exposure dose and on temperature represent potential limitations to realizing the full benefit of ALD, and thus deserve further detailed study.

3.3.4 Chamber wall reactions

The ALD process deposits material on internal surfaces of the reactor, as well as on the wafer. This is evident at the surfaces maintained at or near the wafer deposition temperature. These surfaces include areas of the substrate heater and the cap assembly which surrounds wafer to create an ALD mini-reactor environment. Since the QMS measures product species associated with half-cycle reactions wherever they occur in the reactor, it actually monitors ALD processes on both the wafer surface and on neighboring hot surfaces. To the extent that chamber wall surfaces thus participate in the ALD process, they simply enlarge the active area of the wafer insofar as QMS results are concerned, with two caveats. First, if the temperature of the parasitic surfaces differs substantially from the wafer temperature, the QMS signal intensity, dynamics, and perhaps even component distribution could be affected. Second, since a load-locked system is employed in our reactor, each wafer has a starting surface different from the steady-state surface conditions achieved after ALD cycling, while these additional surfaces remain in steady through a sequence of wafers.

To first order it appears that reactions associated with heated internal surfaces of the reactor behave similarly to those on the wafer surface. ALD dynamics as seen by the QMS are well behaved and as expected, without e.g. long tails of products after the end

of a half cycle which might have suggested different reaction kinetics on internal, non-wafer surfaces.

Details in Figures 3.3 and 3.4 confirm that the reaction rates on wafer surface and heated reactor internal surface are comparable, and that the process dynamics seen by QMS depicts wafer surface reaction with adequate accuracy. Results in Fig. 3.3, and the accompanying discussion, show that the QMS signal did not differ between the [15sec:20sec] and [10sec:8sec] cases, and the corresponding across-wafer thickness profiles in Fig. 3.4 are consistent with this. However, upon further reducing exposures to [8sec:6sec] and [5sec:4sec] we found that the QMS signals decreased notably, and the effects of this are thus seen in Fig. 3.4. Thus the QMS indicates when changes in deposition occur, even in the presence of additional reacting surface areas; of course the location of deposition rate changes cannot be inferred simply from the QMS data, but are found directly in the post-process thickness maps. In summary, the additional heated surfaces where ALD occurs behave to first order as a real extensions of the wafer surface, and they do not adversely affect our ability to exploit QMS for real-time ALD sensing and metrology.

Step	Time (secs)	WF_6 flow rate (sccm)	SiH_4 flow rate (sccm)	Ar flow rate (sccm)	N_2 flow rate (sccm)	Pressure (Torr)
Exposure	15	1	0	99	0	3
Purge	20	0	0	100	0	3
Exposure	20	0	1	0	99	3
Purge	20	0	0	100	0	3

Table 3-1 Default W ALD process conditions. Unless otherwise specified, the process temperature is 325ºC, corresponding to an actual wafer temperature ≈350ºC

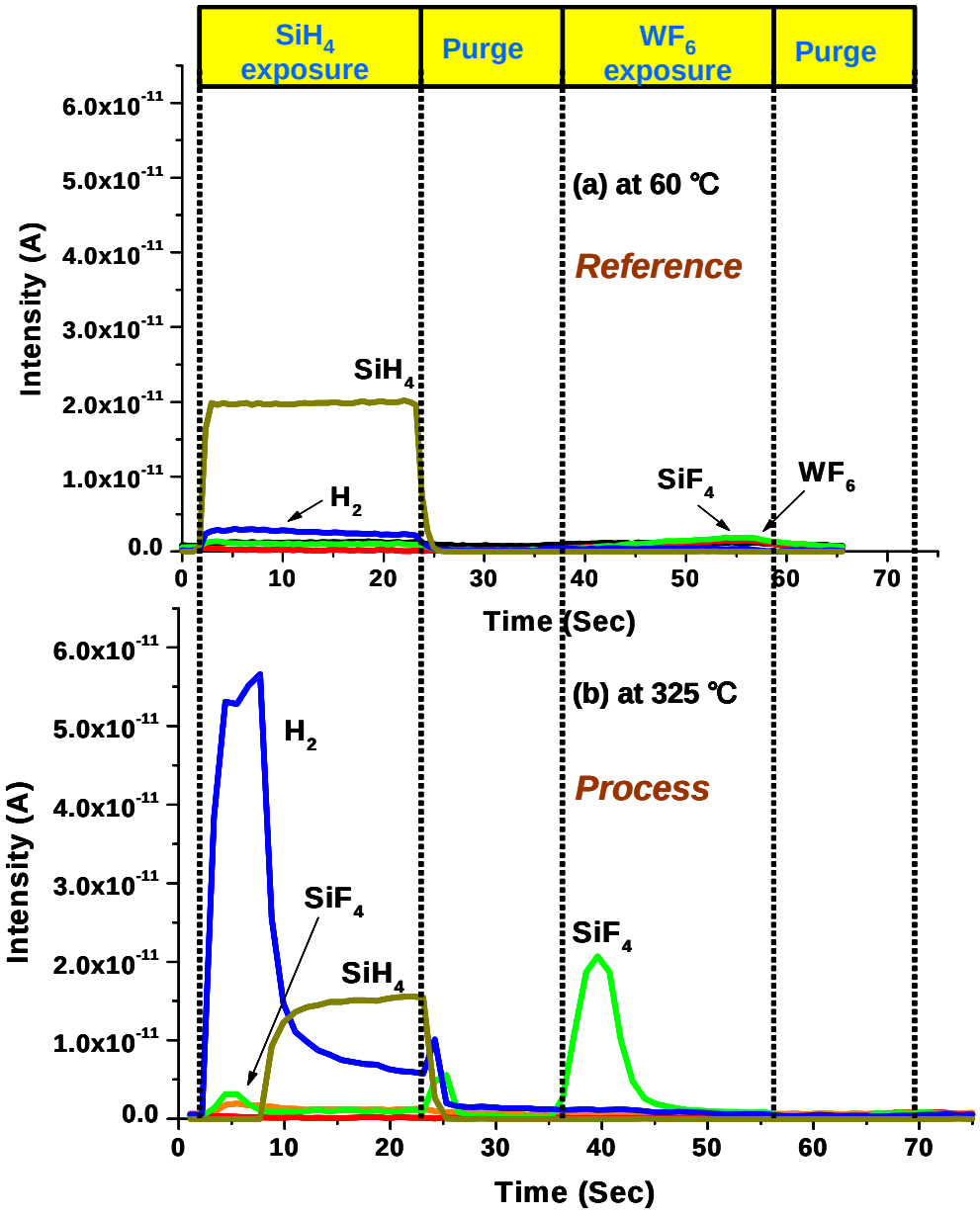
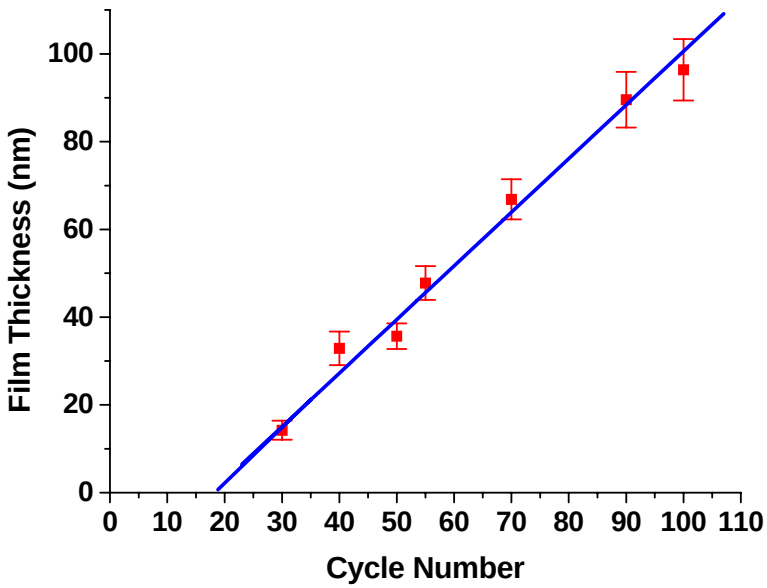
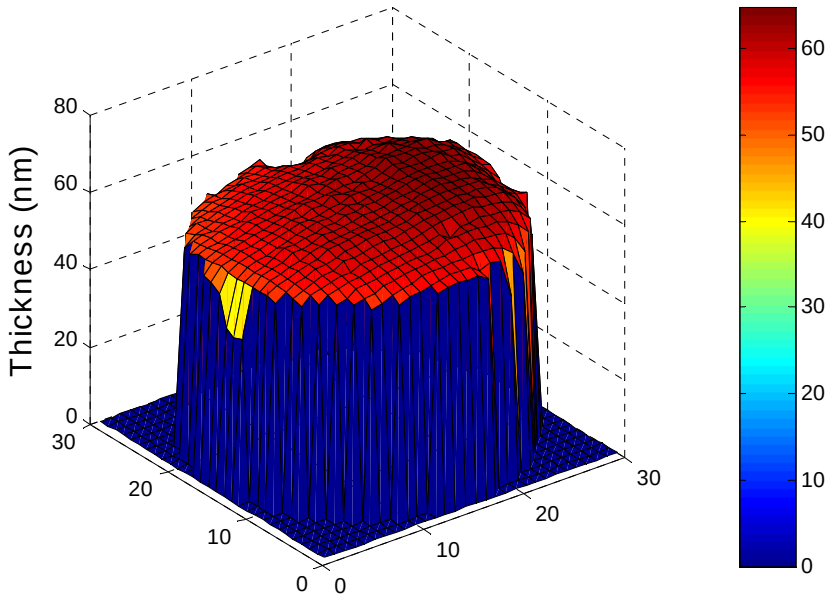


Figure 3.1 Dynamic QMS sensing through one ALD process cycle at (a) low temperature and (b) reaction temperature



(a)



(b)

Figure 3.2 W ALD film characterization: (a) film thickness vs cycle number; (b) film thickness map (70 ALD cycles)

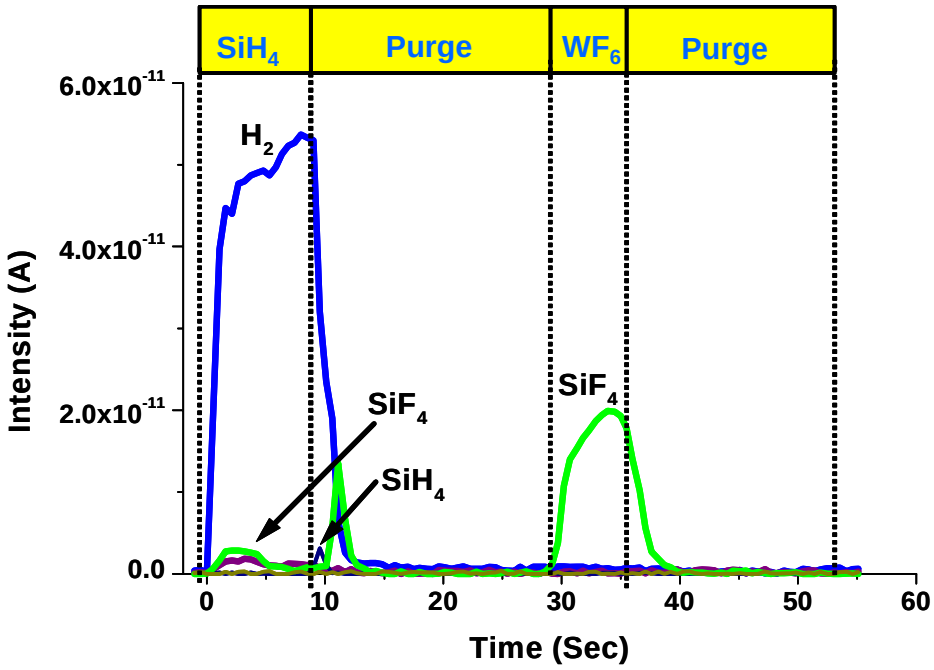


Figure 3.3 QMS pattern after exposure is reduced to surface reaction time

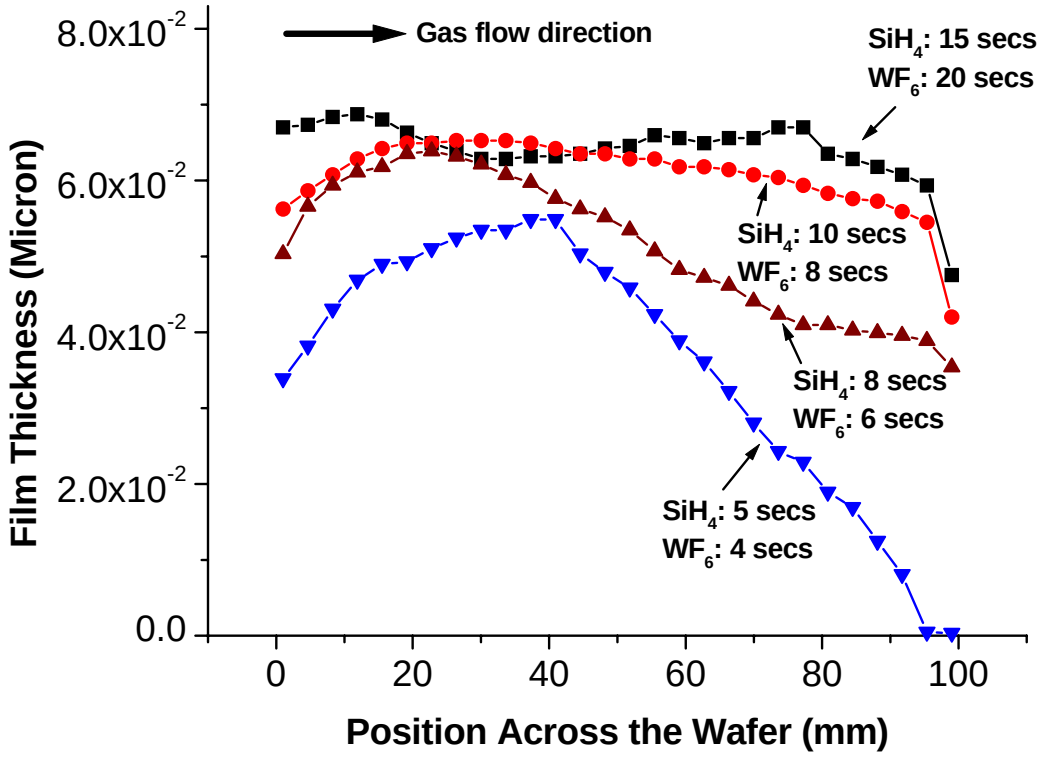


Figure 3.4 W ALD film thickness profile across wafer under different exposure time

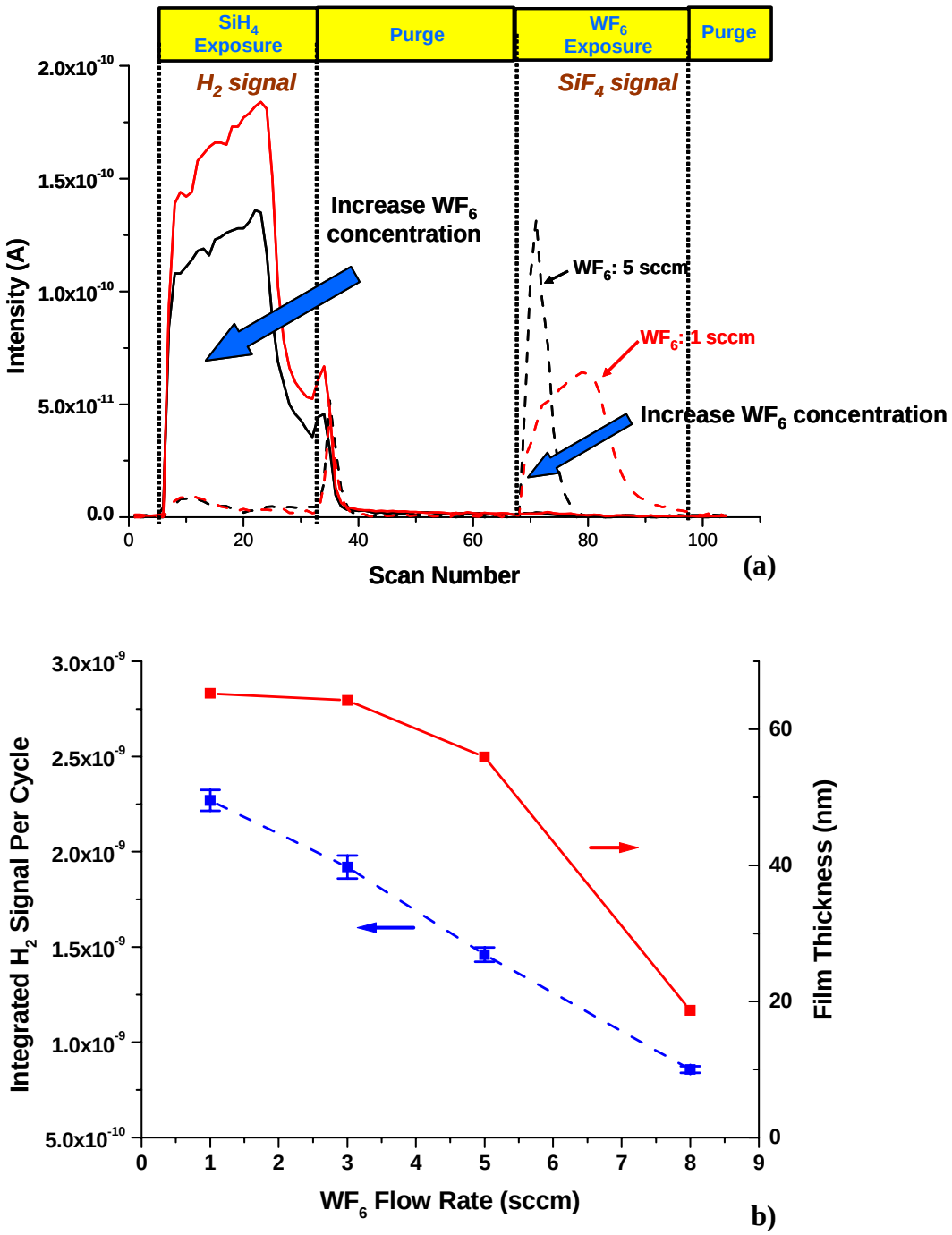


Figure 3.5 (a) QMS pattern changes; (b) W ALD film thickness changes under different WF_6 flow rate (all other conditions remain the same)

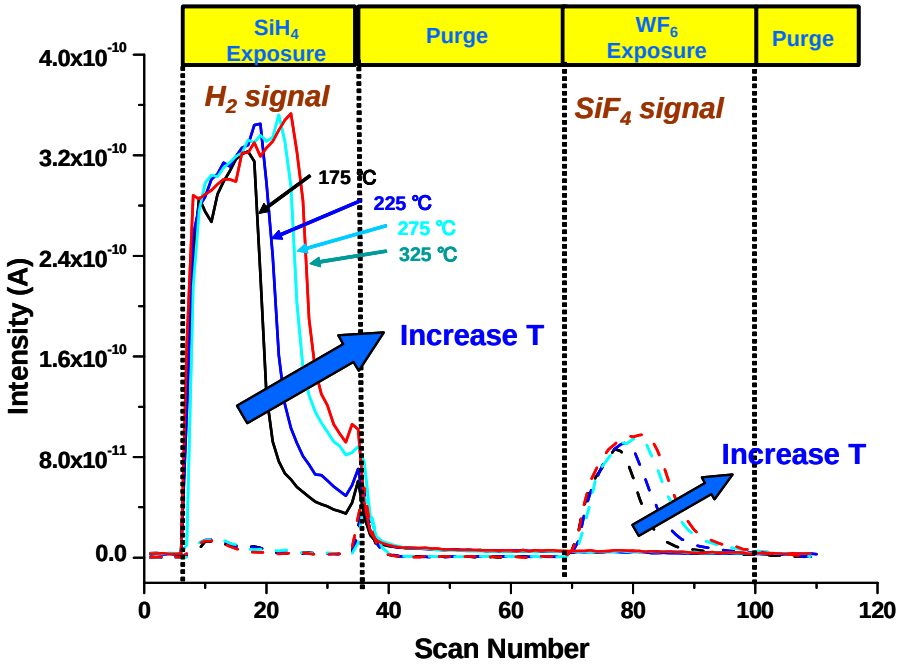


Figure 3.6 QMS pattern changes under different process temperatures

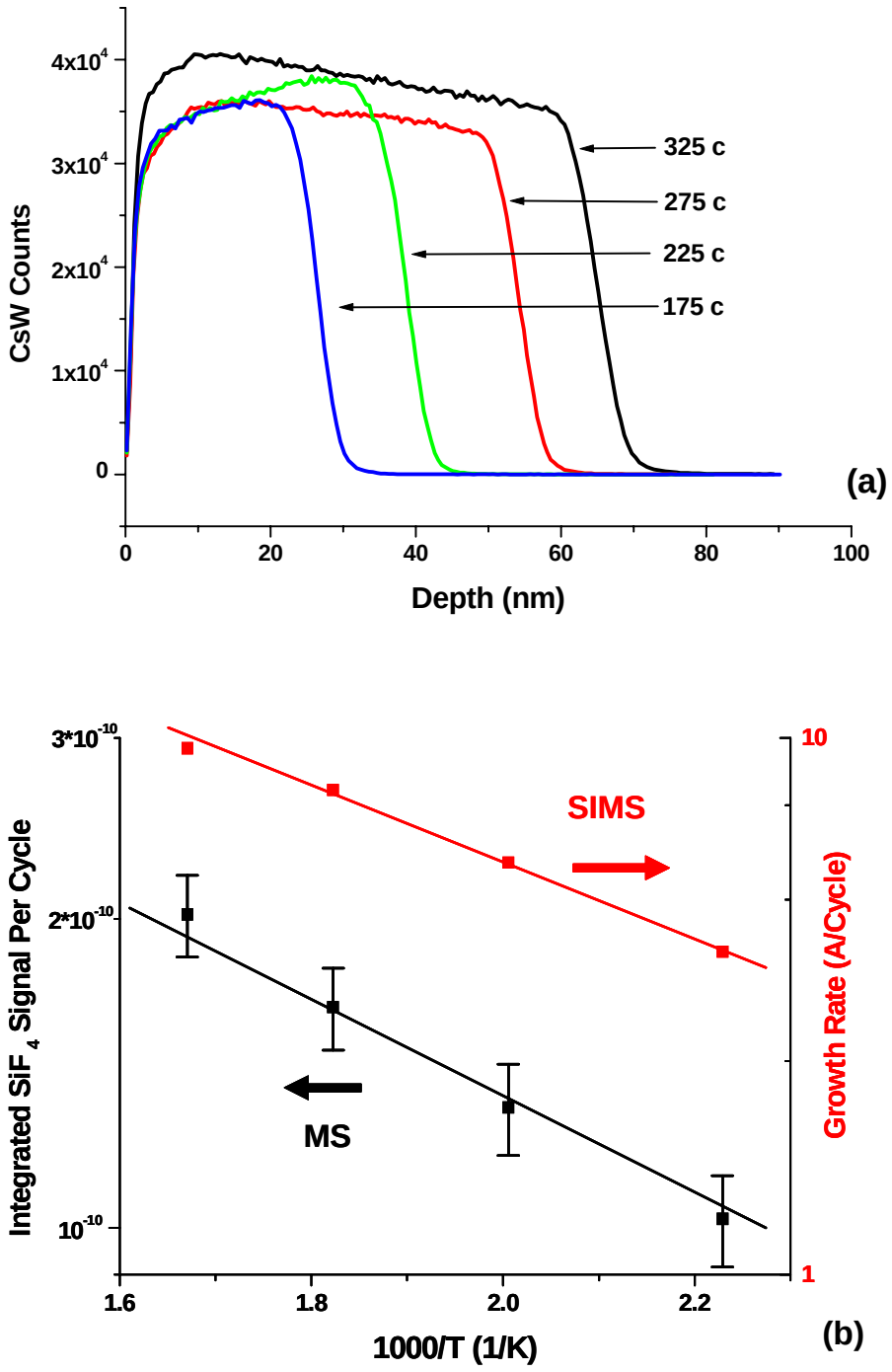


Figure 3.7 (a) SIMS profile of W ALD films; (b) Arrhenius plot

Chapter 4 Real time W ALD film growth observation and film thickness metrology

As discussed in chapter 3, in-situ QMS is able to directly reveal process dynamics of W ALD surface reactions, which in turn facilitates the understanding and optimization of ALD process.[53] The result indicates that the dynamic QMS signal is strongly related to the ALD film growth, so the QMS signal can be potentially used to observe film growth in real time and develop film thickness metrology in ALD processes.

In this chapter, results demonstrating QMS based film growth observation and thickness metrology in W ALD process are presented. ALD film growth kinetics is directly observed by integrating the byproduct QMS signal for each process cycle as a measurement of deposition during that cycle, and plotting them against cycle number. For continuous ALD runs, the summations of integrated byproduct QMS value over each multi-cycle run present a good linear correlation with the corresponding film thickness, so QMS measurements can be used for predicting ALD film thickness in real time. This provides promise for a real-time thickness metrology sufficient to support robust process control strategies.

4.1 W ALD film growth kinetics observation

QMS can clearly reveal the detailed W ALD process dynamics, and the intensity of byproduct QMS signal strongly depends on the process conditions. The QMS result reflects the ALD reaction dynamics and consequently real time ALD film growth. Therefore it is expected that the integrated byproduct QMS signal over one ALD cycle is correlated to the film growth during that cycle.

Figure 4.1 is the "integrated byproduct QMS signal *vs.* cycle number" curve (called as the integrated byproduct curve) of one complete ALD run comprising 70 process cycles. For the SiF_4 byproduct, the QMS signal is integrated through each WF_6 exposure since it is the primary byproduct during WF_6 exposures. Correspondingly, H_2 QMS signal is integrated through each SiH_4 exposure. There are two different regimes in Figure 4.1: the integrated byproduct QMS value first increases as the ALD process proceeds and then saturates at a certain level. These two different regimes represent respectively the nucleation phase and the steady state growth phase of ALD film growth. For example, ALD film growth may start from most reactive surface sites and then form isolated islands on the substrate. As the process proceeds, the islands expand their size with more reactive surface groups available. More reactant therefore will adsorb on the surface in one exposure step, and more byproduct will be released. So the integrated byproduct QMS value (and the amount of W deposited) in the nucleation region will increase as the process cycle number increases. After the film growth completely covers the surface, the deposition rate per cycle and the byproduct generation in each exposure should become essentially constant, as observed.

The integrated byproduct curves for both byproducts are consistent with each other. Despite differences in signal intensity, they demonstrate the same nucleation region length and the same curve shape. For simplicity, in the following discussions, only one of the curves will be shown. The nucleation region length (~20 cycles) given by both curves in figure 4.1 is consistent with the ex-situ film measurements ("film thickness *vs.* cycle number" curve) shown in figure 4.2, where the film thickness is measured by 4 point probe measurement. Because the W ALD film is likely not continuous in the

thickness regime (<20 nm) for the nucleation region, we expect 4 point probe resistivity measurement to indicate zero thickness in the nucleation region. As shown in figure 4.2, both the QMS observation and the post-process measurement give the same result.

Therefore ALD film growth kinetics is revealed in real time by the integrated byproduct QMS curves. It is demonstrated here for the first time in a wafer scale prototype reactor. Such nucleation behavior has already been investigated by a variety of methods such as RBS and AES,[6,15,54-57] but they generally either are not compatible with manufacturing tools or require several ALD runs to build the film growth curve. For example, the "4 point probe measured film thickness curve" in figure 4.2 provides only 6 points, while requesting 6 separate ALD runs and post-process measurement for each run. In contrast, the complete picture of ALD film nucleation and steady state growth is revealed in just one ALD run using the in-situ QMS sensing result in figure 4.2.

As shown in figure 4.3, the integrated byproduct QMS curves, which affect the ALD growth rate, can be well described by the exponential rise equation: $y = A(1 - \frac{1}{\exp[B(x-1)]})$, where A is the integrated byproduct QMS value in the steady state growth region and B is the fitting parameter that decides the curve shape in the nucleation region length (B is 0.13 in the case of figure 4.1).

For a sequence of ALD runs, the integrated byproduct curves are different between the first wafer (defined as the first ALD run after the reactor has been idle) and the non-1st wafer cases even though the process recipes remain the same. As shown in figure 4.4(a), the nucleation for the first wafer is retarded of that for subsequent wafers, while the steady state growth rates are comparable for all wafers. The nucleation

difference only exists between the first wafer and non-1st wafer cases, and the integrated curves are similar for all non-1st wafer cases, as shown in figure 4.4(b).

We attribute the "first wafer effect" in the nucleation region to the fact that the reactor wall surface condition is different during the processing of the first wafer and the second wafer after the system has been idle. ALD film growth happens not only on the wafer surface but also on parts of the internal surfaces of the mini-reactor wall, where the temperature is high enough for ALD reactions. In our UHV-ALD reactor, these surfaces include areas of the substrate heater and the cap assembly which surrounds wafer to create the ALD mini-reactor. When the first wafer is processed, the reactor wall will be covered with ALD film, and the reactor wall surface groups will also change from the original ones to the more reactive ones as ALD process proceeds. Therefore the reactor wall surface condition starts from an "idle" state and transforms to an "active reaction state" as ALD process continues. This transformation is completed in the first wafer run. When the second (or subsequent) wafer is introduced through the load-lock, its surface requires nucleation. However, the surfaces of the mini-reactor walls (heater holder, cap, etc.) have passed the nucleation stage and remain characteristic of the steady state ALD surface conditions, so these areas— larger than the wafer surface itself —do not contribute to the nucleation region as seen by QMS. Therefore, subsequent wafers show a more rapid rise to steady state as seen by the in-situ QMS sensor (Figure 4.4b). A new "first wafer effect" will only be observed after sufficient idle of the reactor (or of course, air exposure) degrades the clearness of the W ALD film on the mini-reactor walls.

The effect of reactor wall conditions on the QMS signal proves the importance of high "wafer surface area to reactor surface area ratio". With a higher "wafer to reactor

surface ratio", the QMS signal is less affected by ALD reactions on the reactor wall and the QMS sensing result will better reflect ALD reactions on the wafer. In our UHV-ALD reactor, the compact assembly of embedded mini-reactor not only provides a small reactor volume, but also reduces the "wafer surface area to reactor surface area ratio" to about $1/3$.

Figure 4.2 demonstrates that the QMS data is consistent with the post-process measurement, even though we know that the existence of reactor wall reactions complicates the QMS result. The influence of wall effects is minimized when ultra-clean conditions are maintained between ALD runs, e.g. by load-locking and UHV capability. Furthermore, since the mini-reactor wall are maintained near the wafer temperature. To first order, this allows us to use QMS measurements to investigate the dependence of ALD film growth kinetics on parameters such as temperature and initial wafer surface conditions.

4.2 W ALD film thickness metrology

To establish a film thickness metrology using in-situ QMS, the QMS signal must be related to the film thickness as measured post-process. As discussed above, the integrated byproduct QMS value in each ALD cycle well represents the ALD film growth in that individual cycle. So the integrated QMS data throughout one ALD run should correlate with ALD film thickness.

The relationship between W ALD film thickness and the total integrated byproduct QMS value in each ALD run is shown in figure 4.5. It shows data from 5 continuous ALD runs and the film thickness is measured by post-process 4 point probe

measurement. Different process temperatures and different ALD cycle numbers are intentionally used in the 5 ALD runs. The process temperature and ALD cycle number for each ALD run are listed in table 5.1. Clearly, there is a linear relationship between the film thickness and the total integrated value no matter which byproduct is used. This relationship can then be utilized to predict W film thickness in real time using QMS.

Unfortunately the first wafer data doesn't fit the linear relationship. This is again because of the reactor wall effect. During continuous ALD runs, the reactor wall condition is different for the first ALD run and the following ALD runs. The reactor wall condition difference then leads to differences in the QMS signal and causes the first wafer data to deviate from the linear relationship.

The existence of reactor wall reactions is evidenced by the positive X axis intercept in figure 4.5. The positive X axis intercept indicates that even when there is no film growth on the wafer, there is still a positive total integrated byproduct QMS value, which is attributed to reactor wall reactions.

Since the deviation of the first wafer data is caused by the different reactor wall condition when processing the first wafer. One can imagine that the first wafer data deviation can be eliminated by properly treating reactor wall before ALD processes to make sure the reactor wall condition is the same throughout the whole continuous ALD runs. This is confirmed by figure 4.6. After 30 ALD process cycles are used to pre-condition the reactor wall before the first ALD run, the first wafer data is compliant with the linear relationship. And the linear fit to the data (figure 4.6) shows an intercept comparable to the intercept in figure 4.5.

4.4 Discussion

ALD process deposits film on the wafer, as well as on any internal surfaces of the reactor which are maintained at or near the wafer deposition temperature. Since QMS measures byproduct species associated with reactions wherever they occur in the reactor, it actually monitors ALD processes on both the wafer surface and on neighboring hot areas. As shown in figure 4.7, to the extent that reactor wall surfaces participate in the ALD reaction, the effect can be considered as an additional virtual wafer surface area which contributes to the QMS signal.

The effect of reactor wall reactions on QMS signal can be understood by treating the reactor wall as an additional virtual wafer. Suppose one can separate ALD reactions on the real wafer from ALD reactions on the virtual wafer, and obtain the corresponding QMS signals separately. For the first wafer case, film growth on both wafers starts from the nucleation region because after the reactor has been idle, the virtual wafer has a "fresh" surface condition just like the real wafer does. Since the experimental data can be well described by the exponential rise equation: $y = A(1 - \frac{1}{\exp[B(x-1)]})$, the equation can also be used to simulate the "integrated byproduct QMS curve" for the first wafer case, as shown in figure 4.8. Figure 4.8(a) is the simulated curve of the real wafer reaction, and figure 4.8(b) is the simulated curve of the virtual wafer reaction shown with larger QMS intensity corresponding to the larger surface area. The A, B value for both curves may be different depending on the surface area and the nucleation region length of ALD reactions. The relative QMS signal intensity of both curves will primarily depend on the surface area ratio if both wafer reactions have similar process conditions. The

length of the nucleation region (L1, L2) depends on the surface status as well as process conditions, and can be different for the two reactions.

Since ALD reactions on both the real and virtual wafers contribute to the byproduct generation, the resulting integrated byproduct QMS curve, which is simulated in figure 4.8(c), is actually the combination of figure 4.8(a) and 4.8(b). In figure 4.8(c), the QMS signal intensity is simply the sum of figure 4.8(a) and 4.8(b), and the nucleation region length is determined primarily by the reaction that has a longer nucleation region. Assume the A, B parameters used for figure 4.8(a) and figure 4.8(b) are A_1, B_1 and A_2, B_2 individually, the simulated curve in figure 4.8(c) is described by:

$$y = A_1(1 - \frac{1}{\exp[B_1(x-1)]}) + A_2(1 - \frac{1}{\exp[B_2(x-1)]})$$

For non-1st wafer cases in continuous ALD runs, the simulated "integrated byproduct QMS curve" of the wafer reaction (figure 4.9(a)) is still the same as the first wafer case, but the curve of the wall reaction (figure 4.9(b)) is different. As shown in figure 4.9(b), ALD film growth on the virtual wafer (reactor wall) goes directly into the steady state region because the reactor wall surface condition has already been changed during previous ALD runs, and therefore no nucleation region exists.

Just like the first wafer case, the combination of figure 4.9(a) and (b) also gives the simulated as-seen curve (figure 4.9(c)). The integrated curve of non-1st wafer cases starts from a higher level than the first wafer case, and both curves saturate at the same level. Even though ALD reactions on the virtual wafer contribute to QMS signal, the nucleation region length in non-1st wafer cases is solely determined by ALD reactions on

the real wafer. Therefore the curve in figure 4.9(c) can be described by: $y = A_1(1 - \frac{1}{\exp[B_1(x-1)]}) + A_2$.

The experimental data in figure 4.4 shows that the nucleation region length is the same for both the first wafer case and the non-1st wafer case, therefore suggesting B_1 and B_2 may have a similar value. Actually, in figure 4.3 where the first wafer data is used, the experimental data is well described by the exponential rise equation using the same value (0.13) for both B_1 and B_2. In addition to that, as shown in figure 4.10, the experimental data of non-1st wafer cases can also be well fitted using $y = A_1(1 - \frac{1}{\exp[B_1(x-1)]}) + A_2$, where 0.13 is used for B_1 and A_1, A_2 are adjusted according to the experimental data. The same B value for both wafer ALD reactions and reactor wall ALD reactions suggests that the growth kinetics, which includes the nucleation region length, is similar for both cases. This is not surprising: since the mini-reactor in the UHV-ALD reactor has a compact assembly and a small volume, it can be expected that both the virtual wafer reaction and the real wafer reaction have similar process conditions (temperature, flow conditions etc.). The consistency between QMS result and post process measurement as shown in figure 4.2 also proves this.

The integrated byproduct QMS signal curve for non-1st wafer cases can also be used to calculate the relative contribution of ALD reactions on the wafer and ALD reactions on the reactor wall to QMS signal. As shown in figure 4.9(c), the relative contribution of ALD wafer reactions and ALD wall reactions can be estimated as b/a.

The least square fit of the integrated byproduct QMS curves for continuous ALD runs demonstrates that the experimental data can be well simulated using the first order exponential rise equation: $y = y_0 + A_1 \exp\left[-\frac{(x - x_0)}{t}\right]$. This behavior is indicative of a situation in which the growth rate is linear to the available surface site density, as described by $\frac{d\theta}{dt} \propto (1-\theta)$, where θ is the surface nucleation site density. As shown in figure 4.11, parameter t, which determines the nucleation region length, is 10.92, 11.12 and 14.4 for the first wafer, second wafer and third wafer cases correspondingly, indicating about 30% variation among them. However the data shows significant noise, where the deviation of parameter t is 0.8, 1.83 and 1.8 for the 3 cases, and we could not expect to find the nucleation region length to be more precisely determined, therefore, the experimental data is well fitted by the first order exponential rise equation, and the nucleation region lengths are similar in all the ALD runs.

If the exposure time is significantly longer than the surface reaction time needed to completely saturate the reaction, then the byproduct background (e.g., H_2 signal from SiH_4 fragmentation inside the QMS sensor during SiH_4 exposure after the surface reaction saturates) in the exposure step may constitute a significant portion of the integrated byproduct value. However, since this background signal intensity and the exposure time are constant throughout the entire ALD run, the increased byproduct background will only change the intensity of the integrated byproduct QMS value, but not the shape changes of the integrated curves. And therefore the observation of film growth using the integrated curves will not be affected.

Wafer sequence	Process Temperature (°C)	Process Cycle Number
1st	175	70
2nd	175	50
3rd	225	70
4th	275	50
5th	325	50

Table 4-1 Process temperature and cycle number for ALD runs shown in figure 4.5

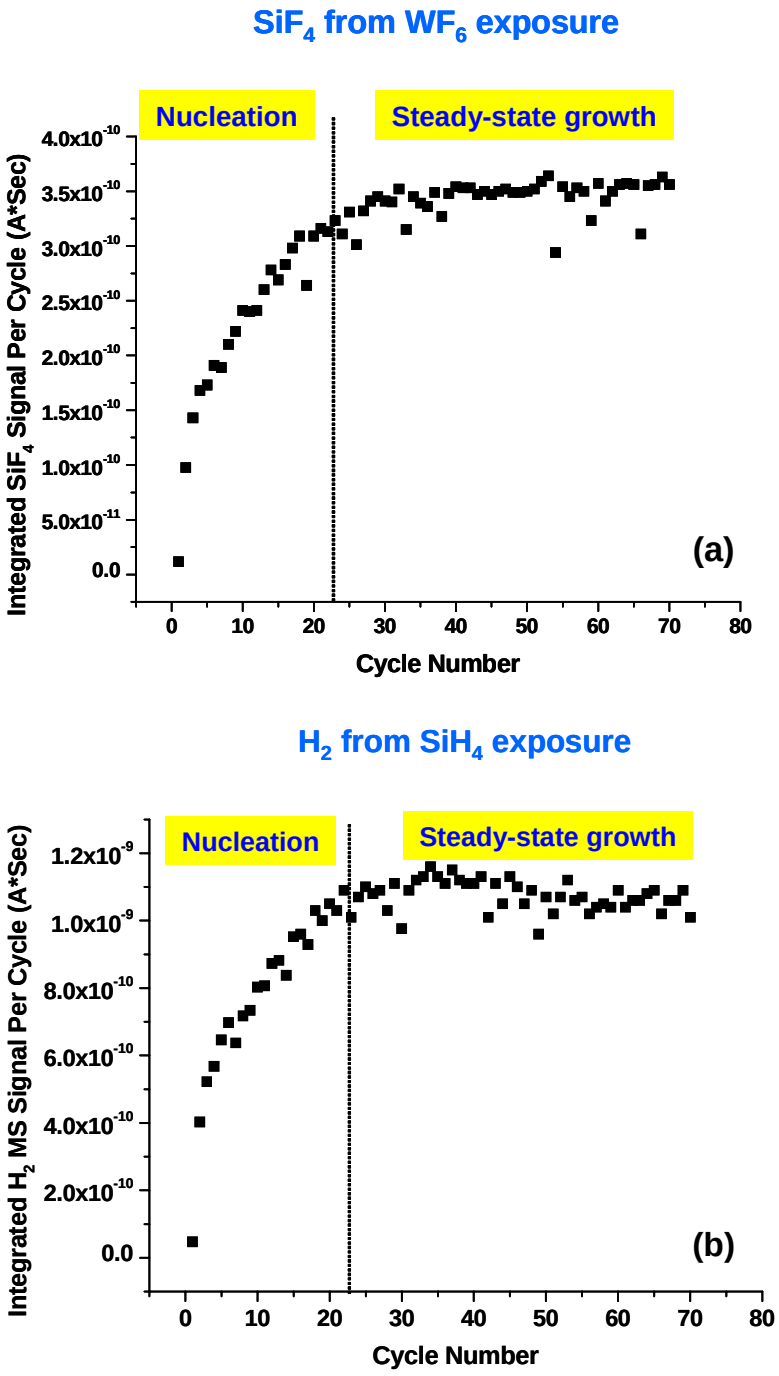


Figure 4.1 Direction observation of W ALD film growth kinetics using QMS: (a) SiF_4 from WF_6 exposure; (b) H_2 from SiH_4 exposure

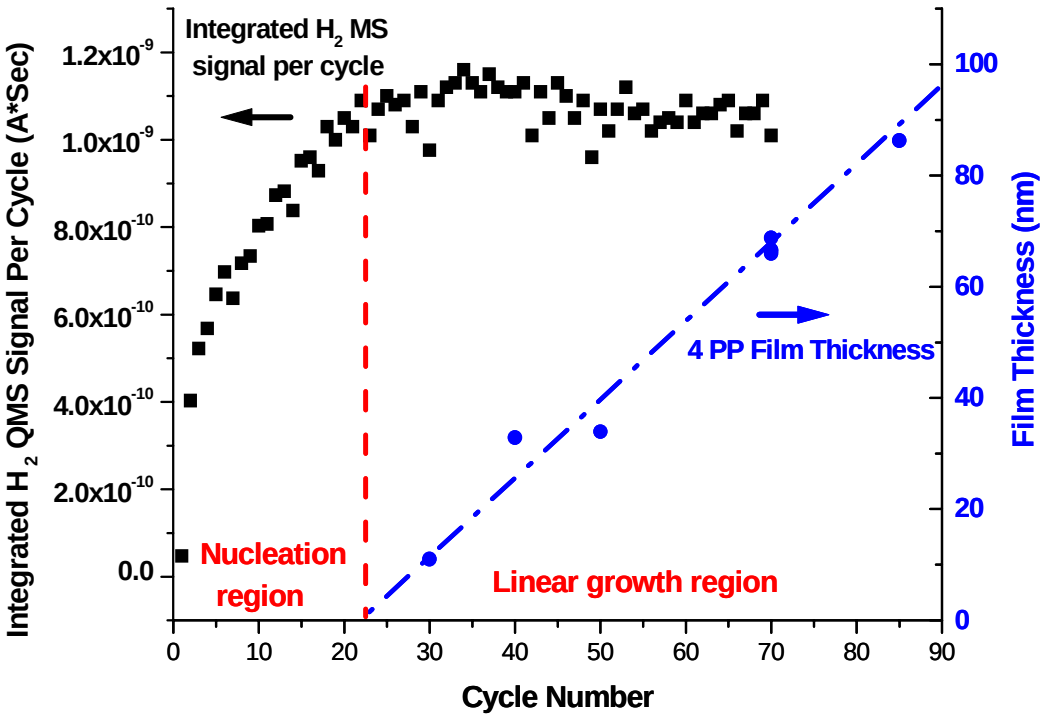


Figure 4.2 Comparison of QMS measurement with ex-situ characterization

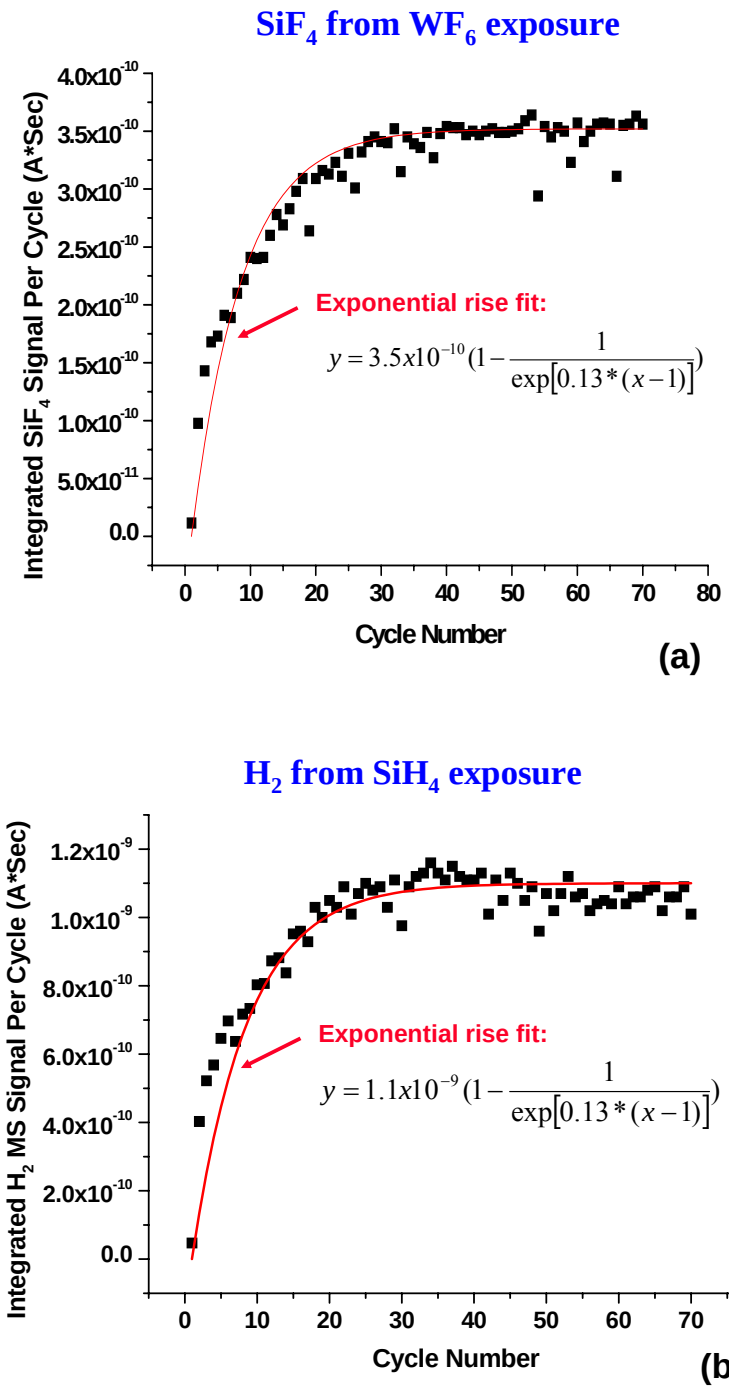


Figure 4.3 Exponential rise fit of the integrated byproduct QMS curves: (a) SiF_4 signal; (b) H_2 signal

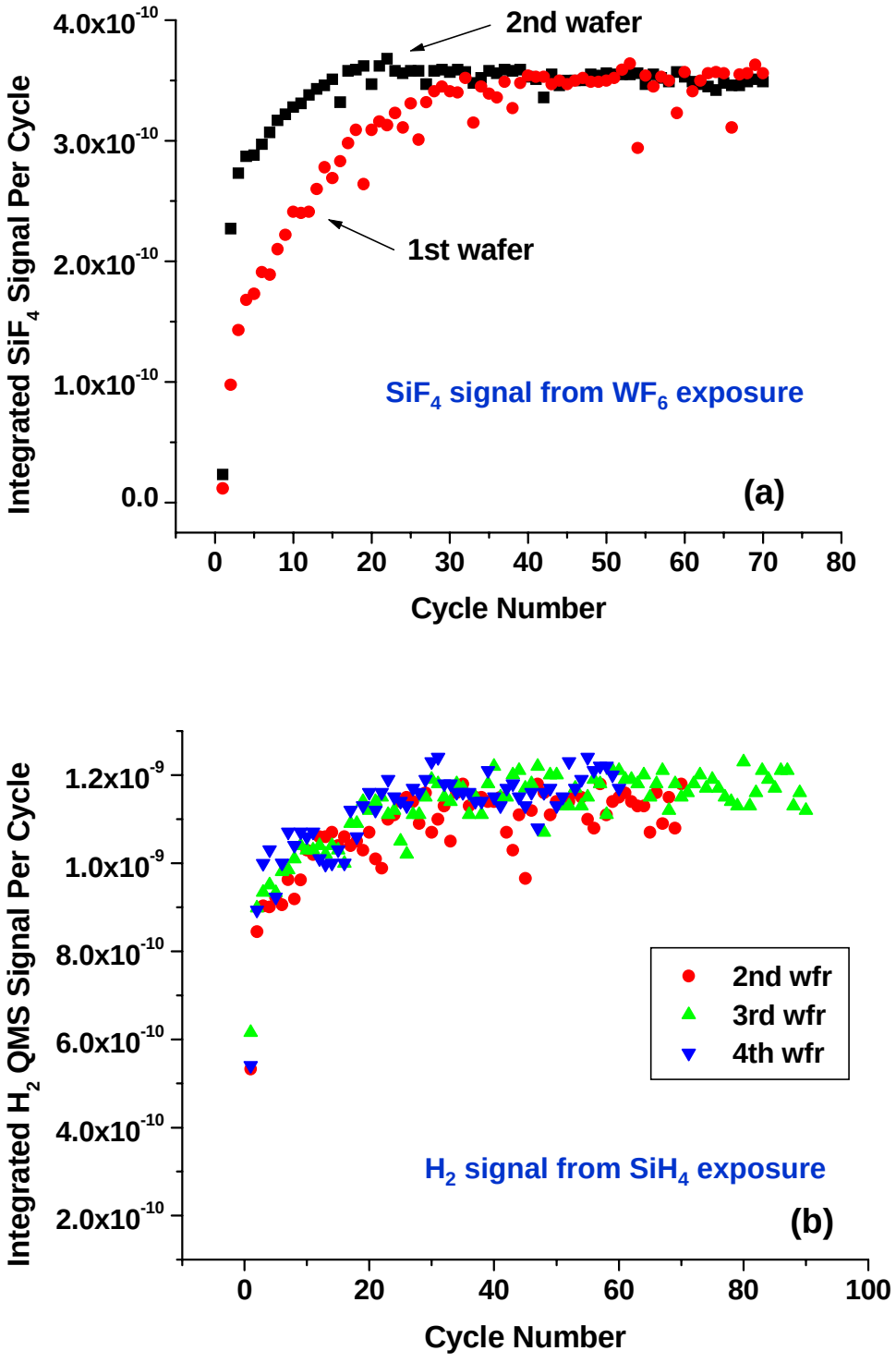


Figure 4.4 QMS observation of ALD film growth kinetics for continuous ALD runs: (a) 1st and 2nd wafer cas es; (s) non-1st wafer cases

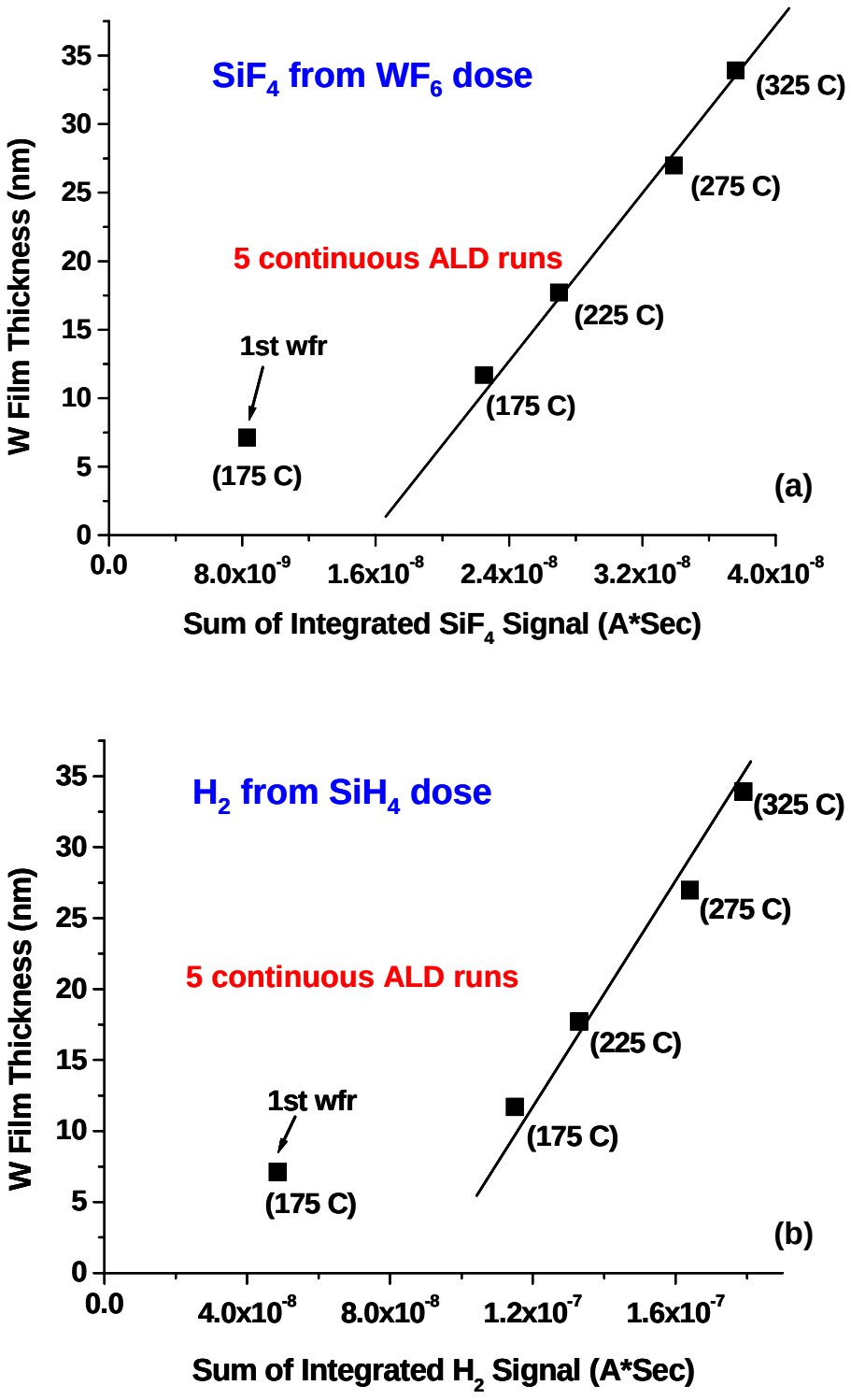


Figure 4.5 ALD film thickness metrology using QMS: (a) SiF_4 QMS signal; (b) H_2 QMS signal

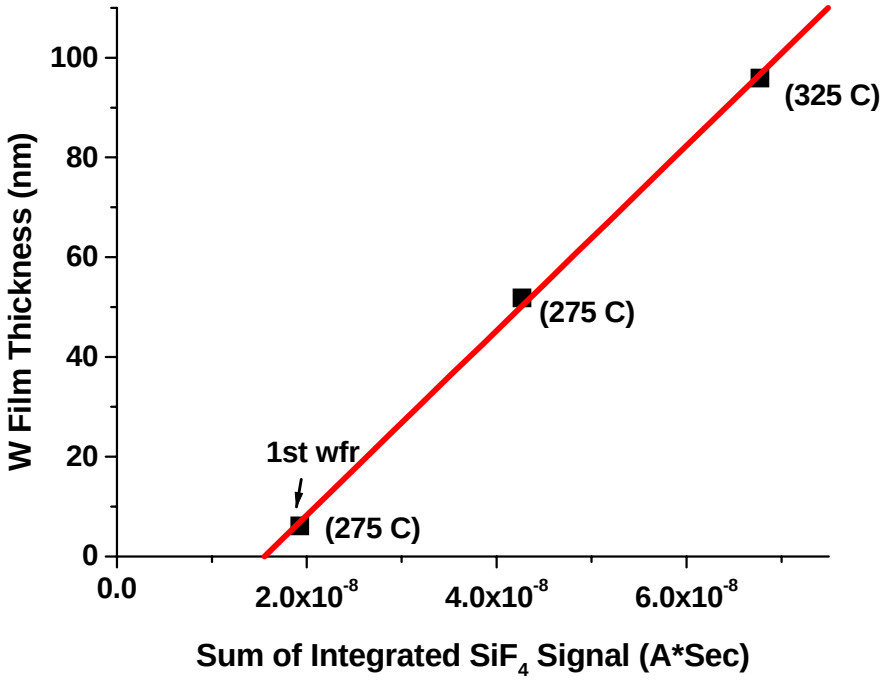


Figure 4.6 ALD film thickness metrology using QMS with pre-process chamber treatment

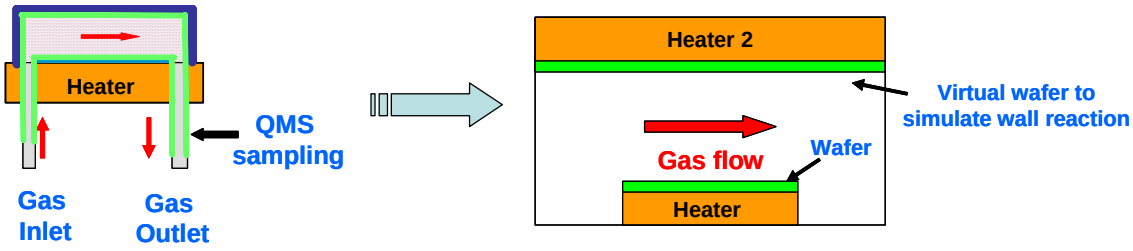


Figure 4.7 Schematic of wafer and wall reactions during W ALD process

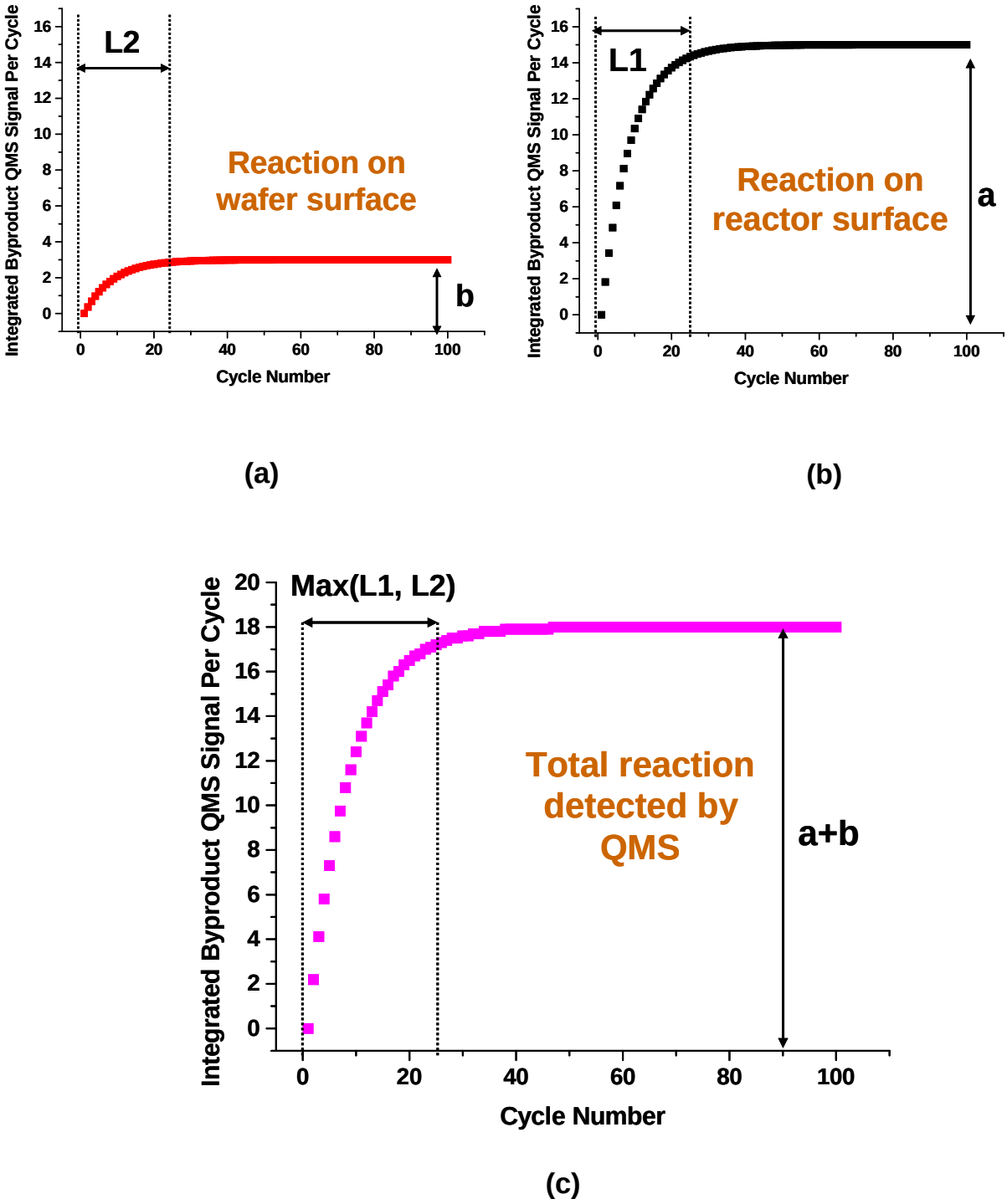


Figure 4.8 Simulated "integrated byproduct QMS signal curve" for the first wafer case: (a) from wafer reaction; (b) from wall reaction; (c) the as-seen curve

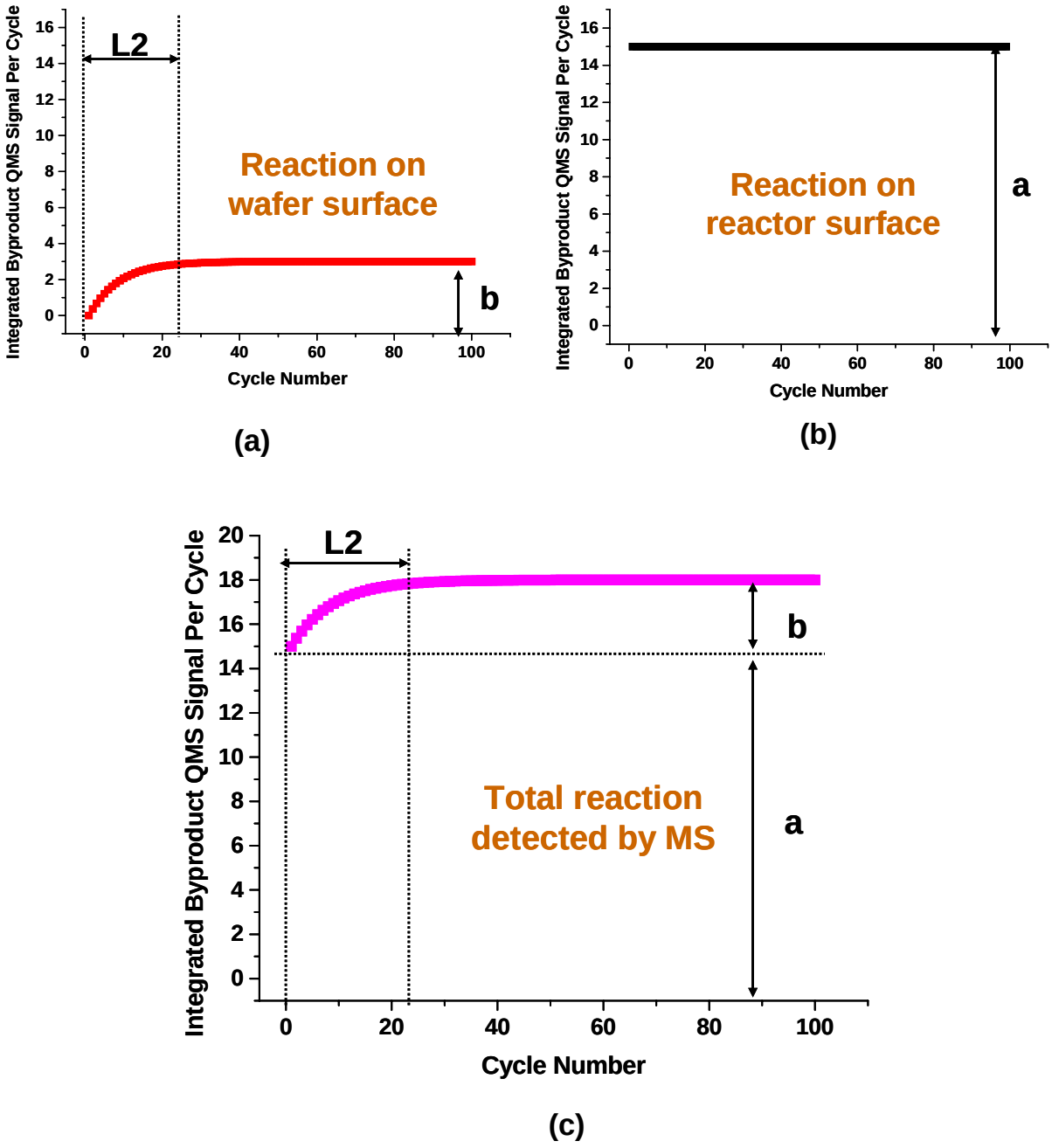


Figure 4.9 Simulated "integrated byproduct QMS signal curve" for non-1st wafer cases: (a) from wafer reaction; (b) from wall reaction; (c) the as-seen curve

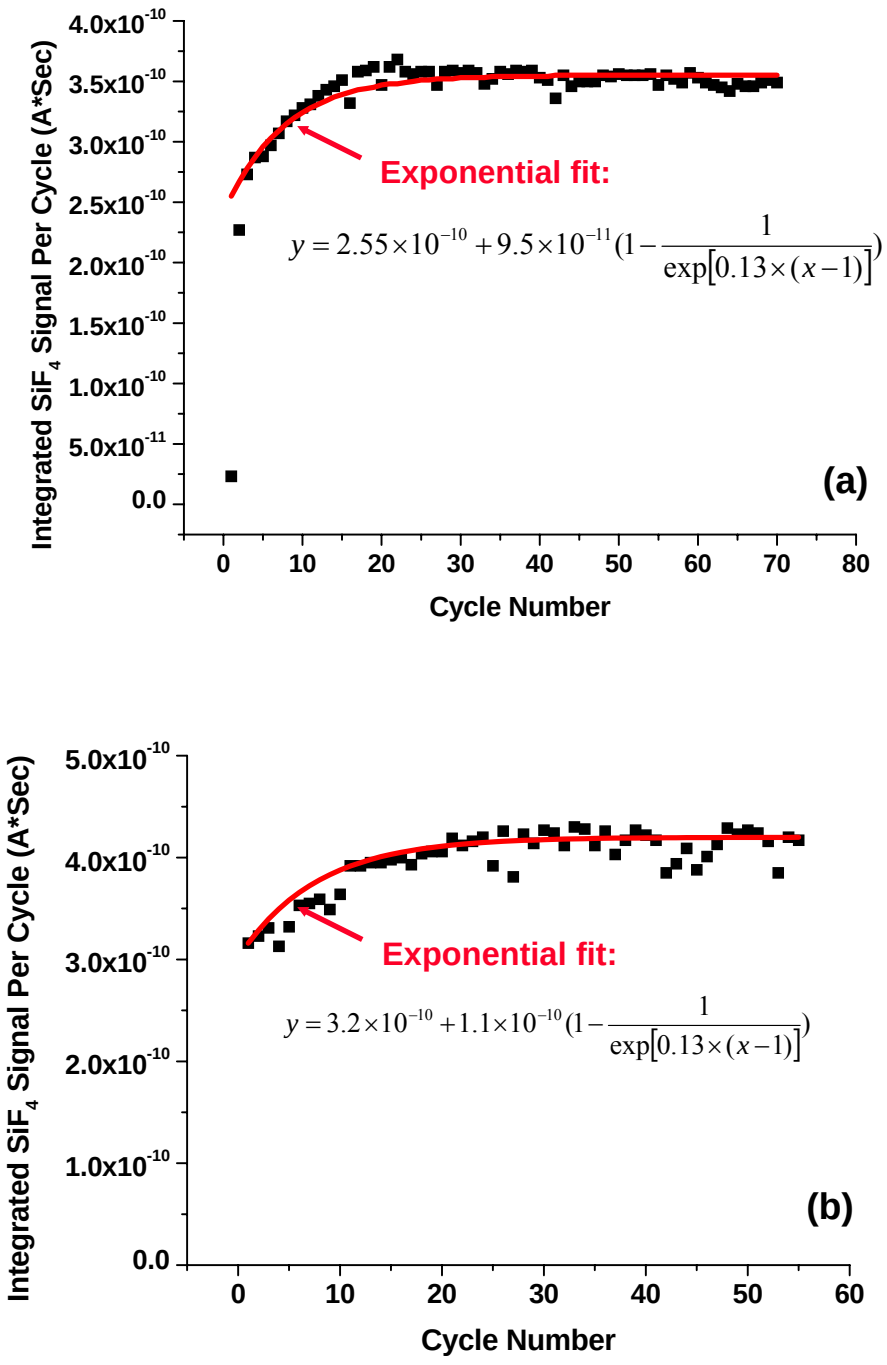


Figure 4.10 Exponential rise fit of the integrated byproduct QMS curves for non-1st wafer cases: (a) SiF_4 used in the first exposure; (b) WF_6 used in the first exposure

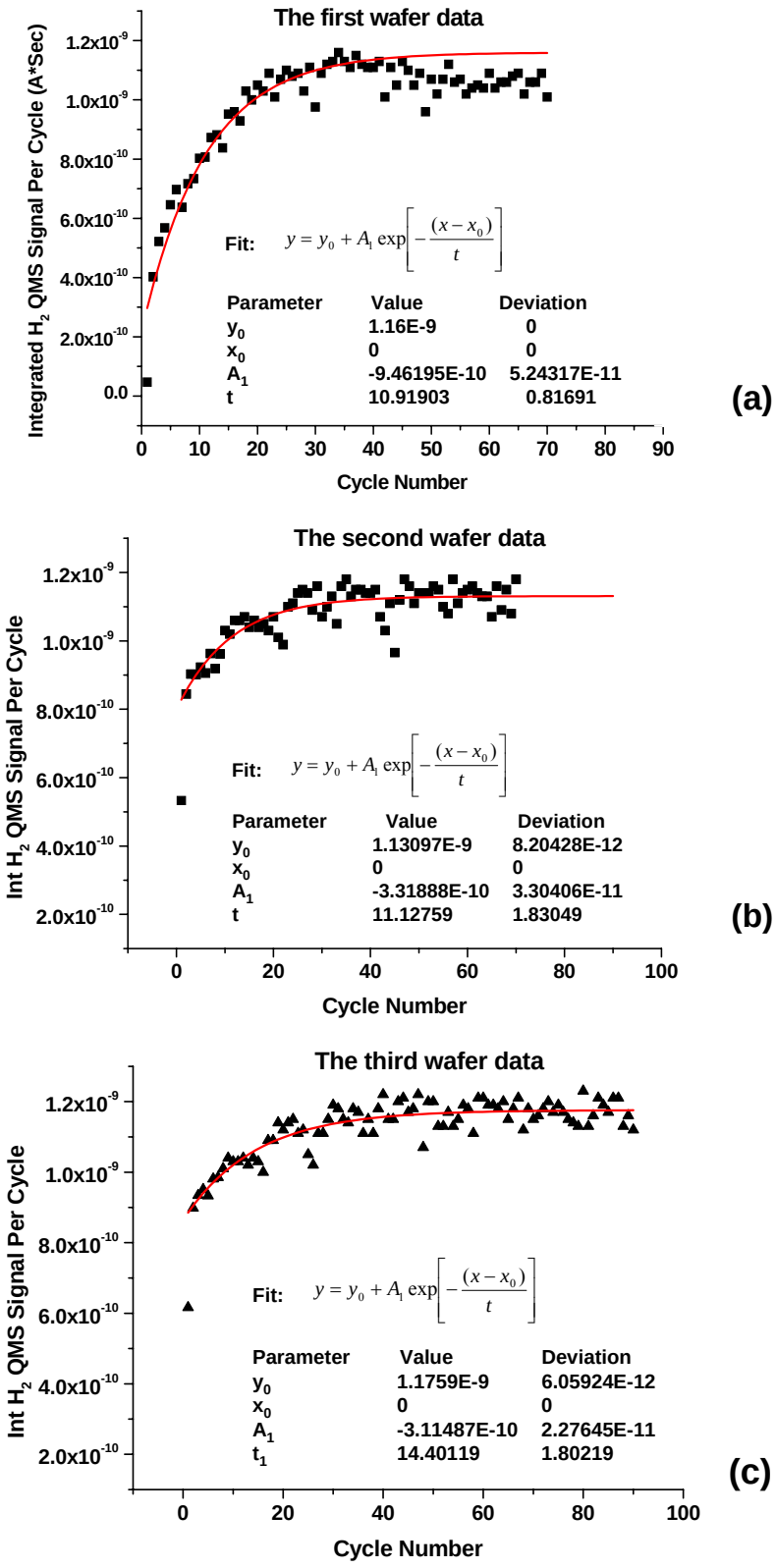


Figure 4.11 The least square fit of integrated byproduct QMS curves: (a) the first wafer case; (b) the second wafer case; (c) the third wafer case

Chapter 5 Real time observation of HfO_2 ALD process

Another important application for ALD process is to ultra-thin high K dielectrics deposition. In high K dielectric ALD processes, metal-organic precursors are often used and the complicated reaction mechanism poses challenges for in-situ process sensing. In addition to that, the structure and characteristics of furnace type hot wall ALD reactors, which are widely used in ALD university research environment, are quite different from substrate heated reactors, and the application of in-situ process sensors on furnace type reactors can also be different. Therefore, in this chapter, in-situ process sensing is applied to HfO_2 ALD process, using a metal-organic precursor and in a furnace type ALD reactor. This work, together with previous in-situ process sensing work on W ALD process, enhanced the in-situ process sensing study by investigating it in different ALD processes and on different ALD reactors.

5.1 HfO_2 ALD process

Replacing SiO_2 gate oxides by high K dielectrics is considered an important step to reduce the effective oxide thickness (EOT) required for further scaling of metal oxide semiconductor field-effect transistors (FET). The high permittivity of high K dielectrics makes it possible to reduce gate leakage current while maximizing the gate capacitance required for maintaining the drive current. The same properties of high K dielectrics also make them highly desired for the insulating dielectrics in the capacitive elements in many memory devices. There are numerous insulators with a higher permittivity than SiO_2,[58] among them HfO_2 has been highlighted due to its high dielectric constant (>20), large bandgap (5.68 ev) and stability with the Si.[6,54] ALD appears to be one of the most

attractive deposition methods for gate dielectrics growth because of its excellent process controllability, even for extremely thin high-K films.[59]

There are two kinds of hafnium precursor in general: halides and metal-organic compounds.[60] Halide precursors are used because they are readily available and they have relatively higher vapor pressure compared to metal-organic precursors. However, the quality of HfO_2 capacitors grown from solid halide sources may sometimes suffer from particles transported from the solid hafnium precursor or from halide impurities in the film, a major source for leakage current.[61] HfO_2 ALD films from metal-organic precursors have high dielectric constant (13~17)[60], but can be problematic because of the low vapor pressure and poor stability. The most widely used oxidant is H_2O, though Ozone (O_3) is attractive as a much stronger oxidation agent to reduce impurity levels in HfO_2 ALD film.[62] However, the strong oxidation ability also makes it problematic for gate dielectric applications due to extensive oxidation of the underlying silicon.[63]

Tetrakis(dimethylamido)hafnium (TDMAH, formula: $Hf(N(CH_3)_2)_4$) and H_2O are used in our HfO_2 ALD process. TDMAH is chosen primarily because it has a reasonable high vapor pressure (1 torr at 70ºC). It has also been used together with H_2O to grow HfO_2 film in ALD mode with a growth rate of 0.1 nm per cycle and low impurity levels.[62,64] Previous work proposed $H\text{-}N(CH_3)_2$ as the only gas phase byproduct, and the ALD reaction path can be described as figure 5.1.

5.2 HfO_2 ALD Reactor

A furnace type hot wall ALD reactor (called as furnace-ALD reactor) is constructed for HfO_2 ALD process. Figure 5.2 shows the schematic and the photo of this reactor. A 3 inch diameter stainless steel tube is located in a furnace heater and used as

the reactor chamber. The reactor diameter on the gas inlet side is slowly increased to reduce the chances of turbulent flow inside the reactor. This reactor is supported by a mechanical pump during ALD process. A turbopump is used when the reactor is in idle status.

This reactor uses the "fixed volume container" concept as proposed by R.G. Gordon[64,65] to control the amount of reactant during each exposure. As shown in figure 5.3 (a), a fixed volume container is used to accumulate and store reactants before releasing them into the reactor. The upstream valve to the fixed volume containers (V1) is first opened before reactant exposure step so the container is filled with reactant vapor. During the exposure step, valves V1 is closed and the downstream valve V2 is opened so only the reactant in the container is supplied to the ALD reactor. Since reactant vapor pressure is fixed at fixed temperature, the amount of reactant supplied in each exposure (reactant dose) is also fixed. This method avoids the use of expensive and sometime unstable mass flow controllers, and the gas delivery system is also simple.

There are still some drawbacks with the proposed gas delivery system configuration. In the "fixed volume containers" setup, the upstream valve V1 is normally open except during exposure steps, and it is found that the H_2O dose in each exposure is actually not constant. As shown in figure 5.3(c), although the open time of valve V1 (or water filling time) is fixed, the water pressure in the fixed volume container just before water exposure decreases as ALD process proceeds, and therefore H_2O dose in each exposure also decreases. We attribute this to evaporative cooling of the water source, leading to H_2O vapor pressure decrease. Indeed, the water vapor pressure is very sensitive to the water tank temperature: putting one hand on the water container (made of

stainless steel) causes the water pressure in the fixed volume container to increase immediately. Another drawback of this delivery system is that it doesn't have flexible dose control ability. If one wants to change the reactant dose, another container with different volume must be used.

Therefore we modified the delivery system to have reactant dose end-point control ability. The idea is to use a pressure gauge to monitor the pressure inside the fixed volume container, and close the upstream valve V1 when the target pressure is reached. Therefore, the amount of reactant dose can be conveniently controlled by adjusting the targeted pressure. A needle valve is added just downstream to V1 to flexibly adjust the H_2O vapor filling rate, as shown in figure 5.3(b). Similar control is now being implemented for the TDMAH source. Figure 5.3 also shows the changes of water pressure inside the fixed volume container just before exposure after the reactant dose end-point control is used. With the reactant dose end-point control, H_2O dose in every exposure is constant.

Multiple in-situ process sensors are integrated to the reactor, including QMS, QCM and Fourier Transform Infrared Spectroscopy (FTIR). QMS is located downstream to the wafer with an orifice for pressure reduction. The orifice position is very close to the exhaust line to minimize the dead volume. The QCM sensor is located deep inside the reactor and very close to the wafer, so the temperature and flow conditions are similar for both QCM sensor and the wafer.

5.3 Experimental conditions

HfO_2 ALD experiments were performed on small wafer pieces (about 1 inch x 1 inch) cut from 4 inch (100mm) diameter Si(100) wafers with 500 Å thermally grown SiO_2. The wafers were cleaned in acetone and methanol before experiments.

TDMAH is solid at room temperature (its melting point is 30ºC). It is kept in an oil bath at 75ºC in our setup. TDMAH vapor pressure is about 1 torr at this temperature. Higher temperature is not used because it would risk decomposition of TDMAH. The whole TDMAH delivery line is heated at about 80ºC. The water tank and its gas delivery line are at room temperature (25ºC) with water vapor pressure at about 25 torr. N_2 is used as the carrier/purge gas and is constantly flowing during ALD processes. A constant process pressure of about 115 mTorr is reached with 33 sccm N_2 gas flow. The reactor pressure is abruptly changed by reactant exposure but will go back to the base pressure immediately. The wafer temperature is directly measured by attaching a thermal couple on the wafer holder. The default process conditions are listed in table 5.1. Unless otherwise specified, all the experiments reported here were conducted under these default process conditions.

The thickness of HfO_2 film is measured in-situ by QCM and ex-situ by an ellipsometry using 2.1 as the reflex index of HfO_2. Both measurements agree with each other. Film composition is measured using X-ray Photoelectron Spectroscopy (XPS), Rutherford Back-scattering Spectroscopy (RBS) and SIMS.

5.4 Results

5.4.1 In-situ sensing results

Figure 5.4 is the QCM result during one HfO_2 ALD experiment, the mass increase on QCM is converted to film thickness by using 9.68g/cm^3 as HfO_2 film density. The grow rate increases at the beginning of the process and saturates when the process is in steady growth region. At 200ºC, the growth rate is about 1.1 Å per cycle, very close to previous report.[64]

QCM is also able to reveal HfO_2 ALD process details. Figure 5.4 (b) is a closer look of the QCM result during linear growth region as shown in figure 5.4 (a). HfO_2 ALD film thickness, which is converted from the film weight, increases during TDMAH exposure because TDMAH molecules react on the film surface and the film weight is increased. Although some of the -$N(CH_3)_2$ ligands are released into the gas phase as shown in figure 5.1, the majority part of TDMAH molecules stay on the surface and overall the film mass increases. On the other hand, QCM result reveals that the film thickness decreases during water exposure. This is because during H_2O exposure, light weight H_2O molecules react on the surface and heavier H-$N(CH_3)_2$ molecules are released. Therefore the overall film weight decreases during H_2O exposure.

The gas phase changes in HfO_2 ALD processes are monitored by the downstream QMS. Based on the NIST database, mass 44 is used for byproduct $HN(CH_3)_2$ and the byproduct is clearly observed.

As shown in figure 5.5, QMS clearly observes the rise of byproduct $HN(CH_3)_2$ signal during both reactant exposures, consistent with the proposed reaction chemistry. Unexpected gas species like H_2 and O_2 are found rising during exposures, they are also

detected during continuous H_2O or TDMAH exposures (different from the alternating exposures as in ALD process case), which proves that these unexpected gas species are from the reactant exposure background. The integrated byproduct QMS signal can't be used to reveal the film growth kinetics on the wafer because: first, the reactor wall reaction is dominant in the furnace reactor because the reactor wall surface area is much larger than the small wafer (~ 1 inch^2); second, the reactor is not load-locked, so the reactor wall surface condition is changed during wafer loading/unloading, causing problems for the integrated byproduct QMS signal curves.

5.4.2 Process mechanism investigation

The reaction mechanism shown in figure 5.1 is proposed based on simple bond-energy considerations and the acid-base theory[64], and it is possible that the ALD reaction may take other reaction paths and release byproduct $HN(CH_3)_2$. The details of HfO_2 ALD reactions can be investigated by QCM and QMS in-situ sensing. Assume the number of -$N(CH_3)_2$ released from one TDMAH molecule during TDMAH exposures is n, then when one TDMAH molecule is reacted during TDMAH exposure, the number of $HN(CH_3)_2$ molecules released is n, and the remaining $–N(CH_3)_2$ is released during the following H_2O exposure, with $(4-n)$ H_2O molecules participating in the reaction. Therefore the ratio of mass changes during TDMAH exposure and water exposure becomes:

$$\frac{mass(TDMAH)-n\times mass(HN(CH_3)_2)}{(4-n)\times mass(HN(CH_3)_2)-(4-n)\times mass(H_2O)}.$$

From the QCM measurement as shown in figure 5.4(b), the thickness increase during one TDMAH exposure is about 1.26 Å while the thickness decrease during one

H_2O exposure is about 0.25 Å, so the ratio of mass changes during TDMAH exposure and water exposure is about 5, which gives a *n* value of 2.02, consistent with the proposed reaction mechanism.

The proposed reaction mechanism indicates that both TDMAH exposure and H_2O exposure release the same amount of byproduct $HN(CH_3)_2$, and it can also be verified using QMS sensing results. The ratio of byproduct $HN(CH_3)_2$ generated during TDMAH exposure and during water exposure can be calculated by integrating $HN(CH_3)_2$ QMS signal through each TDMAH and H_2O half cycle, and subtracting the corresponding $HN(CH_3)_2$ background signal. The calculated $HN(CH_3)_2$ ratio during TDMAH and H_2O half cycles is very close to 1, again consistent with the QCM result.

5.4.3 Complexity of ALD reaction

The relationship between ALD growth rate and reactant dose (with 20 seconds purge time) is shown in figure 5.6. The growth rate is measured by QCM. In figure 5.6(a), the water pressure in the fixed volume container just before water exposure actually represents the amount of water exposure dose. The higher the water pressure, the more water is supplied to the ALD reactor during exposure. The growth rate first increases as the amount of water dose increases, then it turns to saturation after certain point, demonstrating self-limiting reaction behaviors. But a close examination tells that the self-limiting reaction is actually complicated by other factors, as evidenced by the continuously slight increase of growth rate. Figure 5.6(b) is the relationship between growth rate and water purge time at 124ºC. When the water purge time is reduced, the ALD growth rate increases, again contradictory to the simple self-limiting reaction explanation. The dependence of growth rate on reactant dose and purge time indicates

that the self-limiting behavior is not perfect in HfO_2 ALD process, probably due to precursor physisorption that happens in parallel to chemisorption.

There will be less physisorbed water molecules at higher process temperature, but high process temperature can also cause other problems. Figure 5.7 is the SIMS result of two HfO_2 samples which are grown under the same process conditions except different process temperatures. Sample 1 is grown under default process conditions (process temperature is 200 ºC) for 100 cycles, sample 2 is also grown for 100 cycles but at 300 ºC. The Hf atom profiles are similar for both samples, as shown in figure 5.7 (a). But the C impurity profile as shown in figure 5.7 (b) shows that C impurity level is higher at higher process temperature, probably due to the decomposition of precursor TDMAH inside the reactor. With longer purge time, there will be less physisorbed precursor molecules left. But the longer purge time also increases the chance of TDMAH decomposition on the substrate surface before it is reacted in the following water exposure.

5.4 Discussion

The in-situ process sensing study on HfO_2 ALD process once again demonstrates the complexity of ALD processes. Perfect self-limiting behavior is hardly observed, as proved in the growth rate dependence on reactant dose, purge time and process temperature. Physisorbed reactant molecules play an important role in those effects. Although there are ways to reduce the amount of physisorbed reactant molecules, they might affect other aspects of the ALD process.

The complexity of ALD process justifies the need for in-situ process sensing. Both QCM and QMS sensing in the HfO_2 ALD process are consistent with each other

and provide detailed real time process information, therefore accelerating process learning and optimization.

Step	Time (Sec)	N_2 flow rate (sccm)	Water pressure before exposure (Torr)	Pressure (Torr)	Temperature (℃)
Water Exposure	1	33	7	115	200
Purge	10	33	NA	115	200
TDMAH Exposure	1	33	NA	115	200
Purge	10	33	NA	115	200

Table 5-1 Default HfO2 ALD process conditions

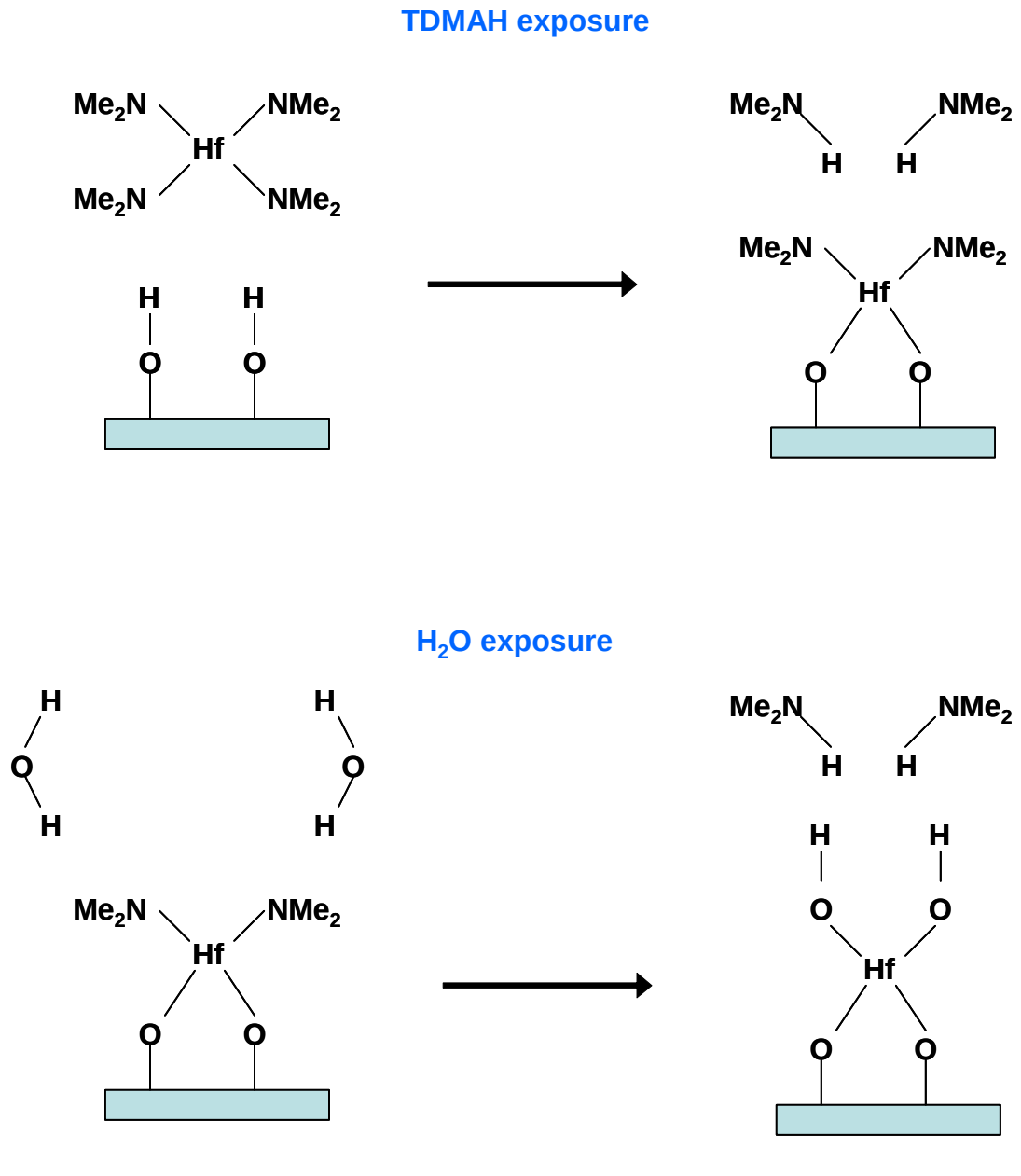
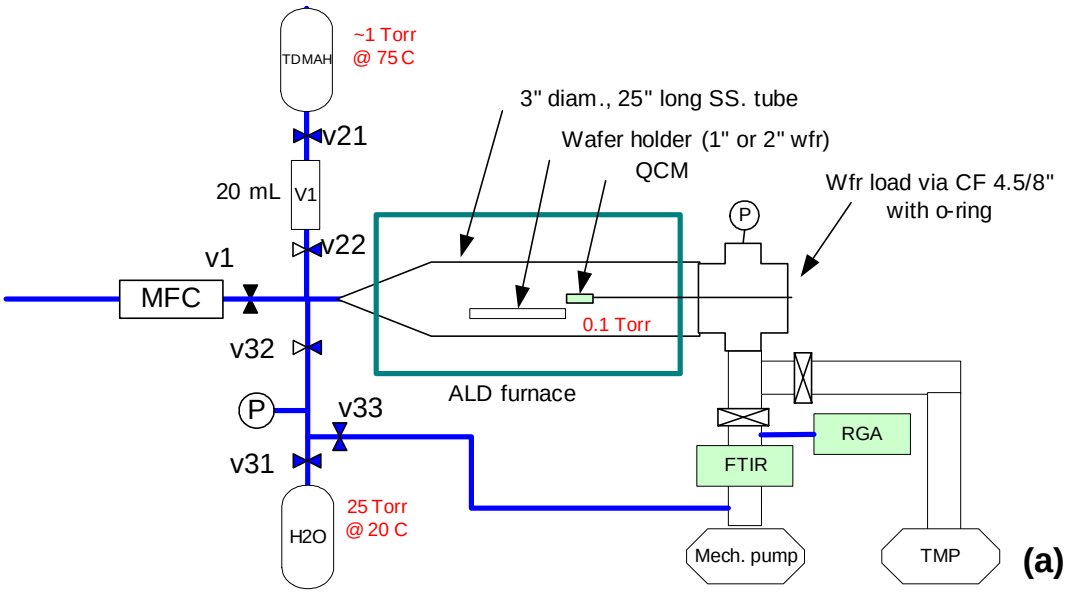


Figure 5.1 Proposed HfO2 ALD reaction



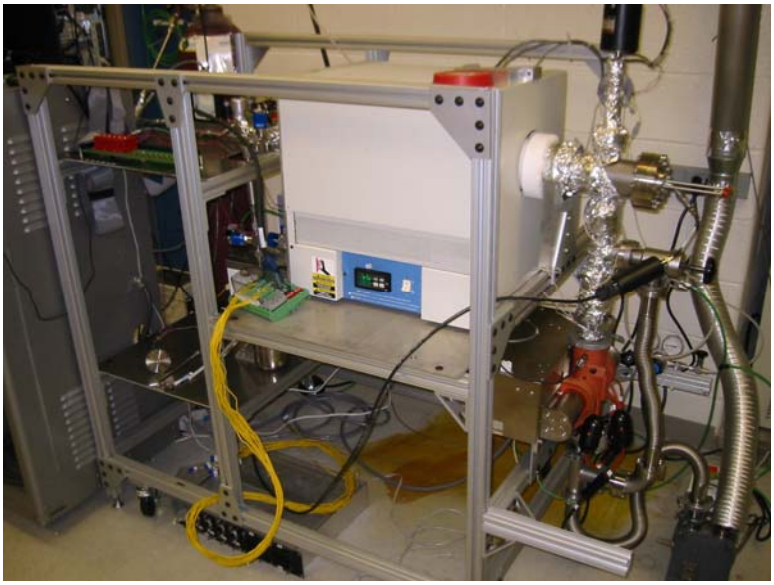


Figure 5.2 Furnace type hot wall ALD reactor: (a) schematic; (b) picture

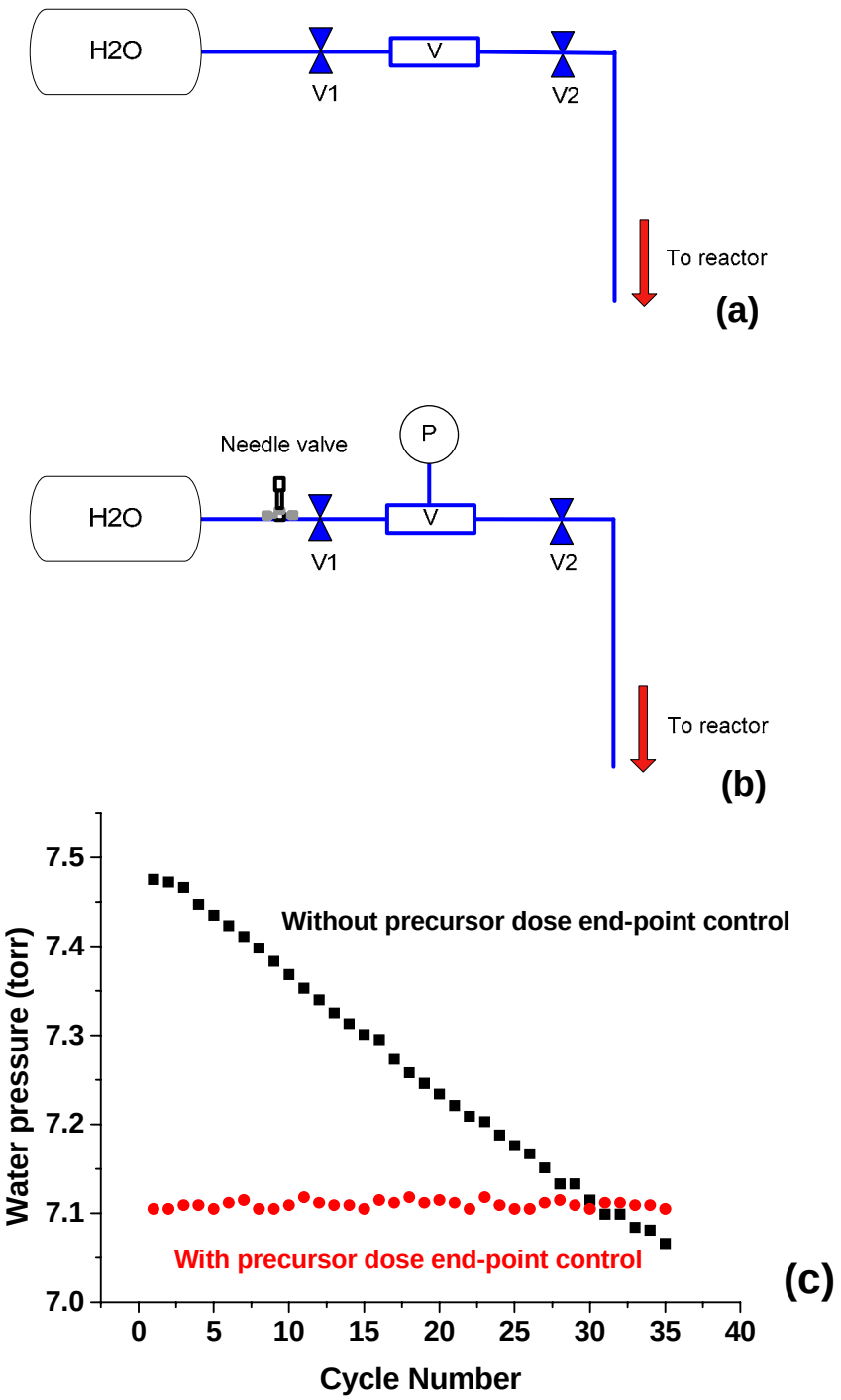


Figure 5.3 (a) "Fixed volume container" design; (b) "Fixed volume container" design with dose end-point control; (c) water pressure in the fixed volume container before H_2O exposure

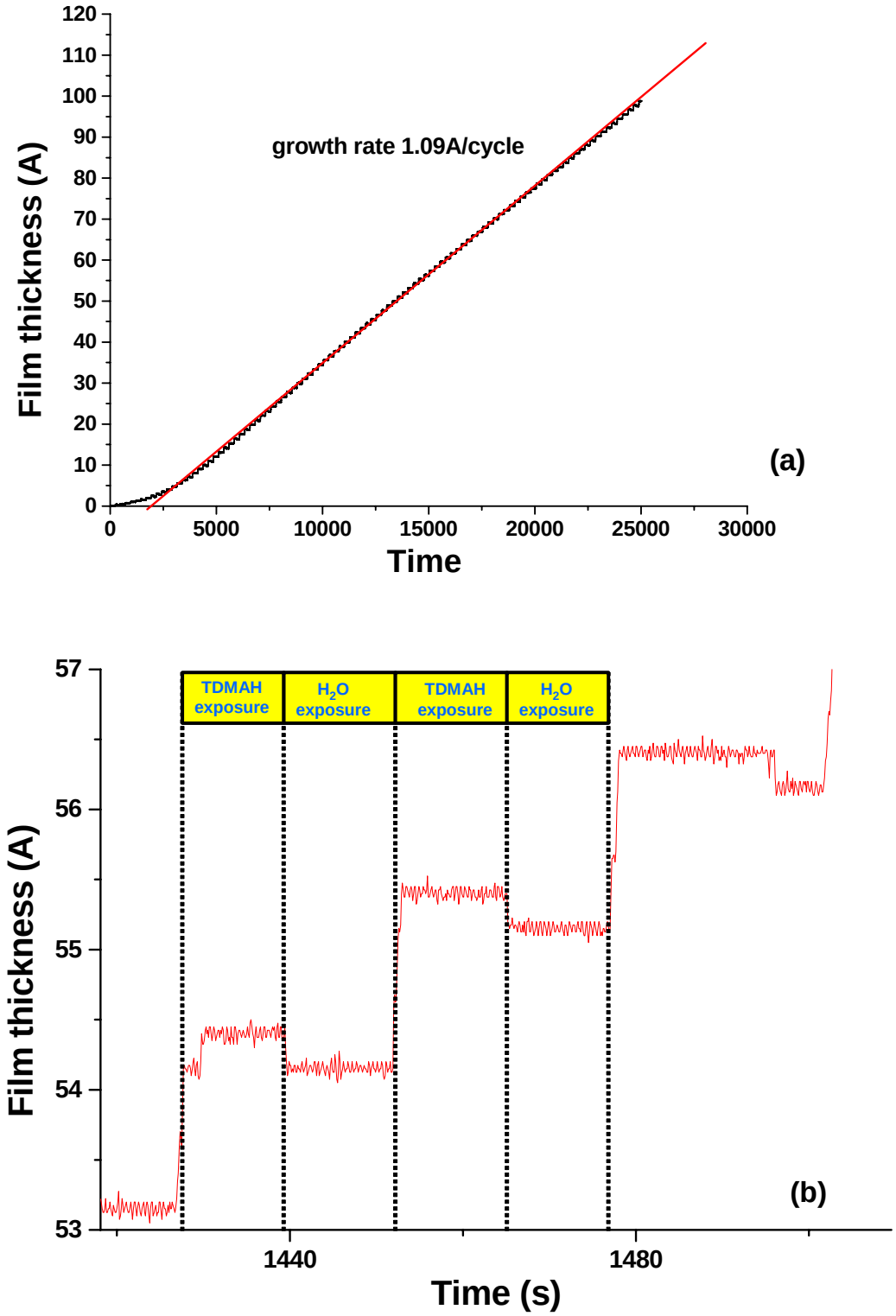


Figure 5.4 In-situ QCM sensing of HfO2 ALD process: (a) overview; (b) details

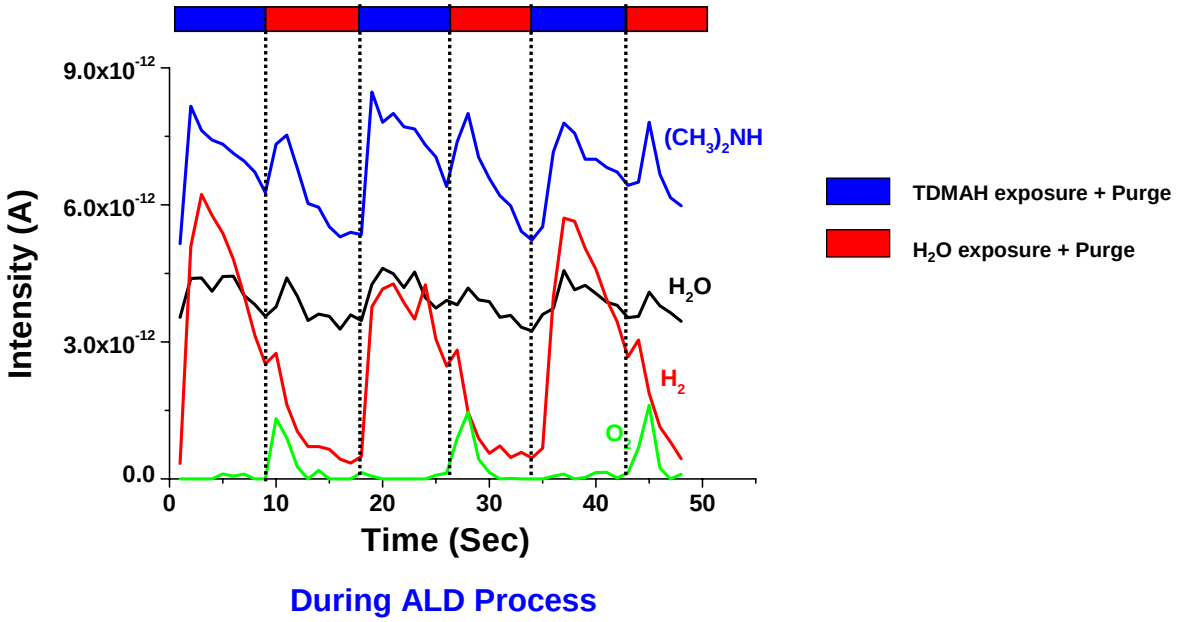


Figure 5.5 Dynamic QMS signal in HfO_2 ALD process

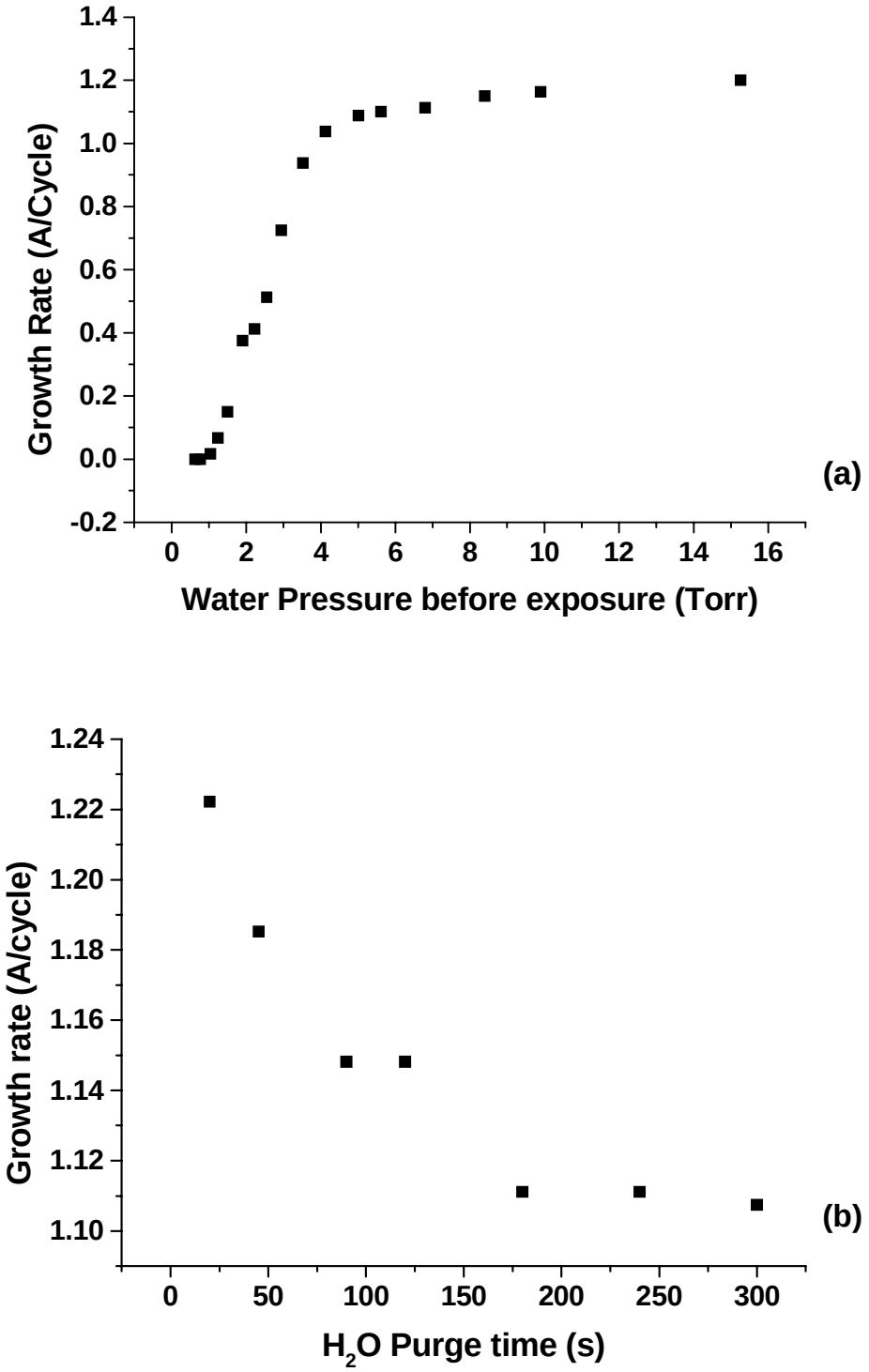


Figure 5.6 Growth rate change during HfO_2 ALD process: (a) H_2O exposure; (b) TDMAH exposure

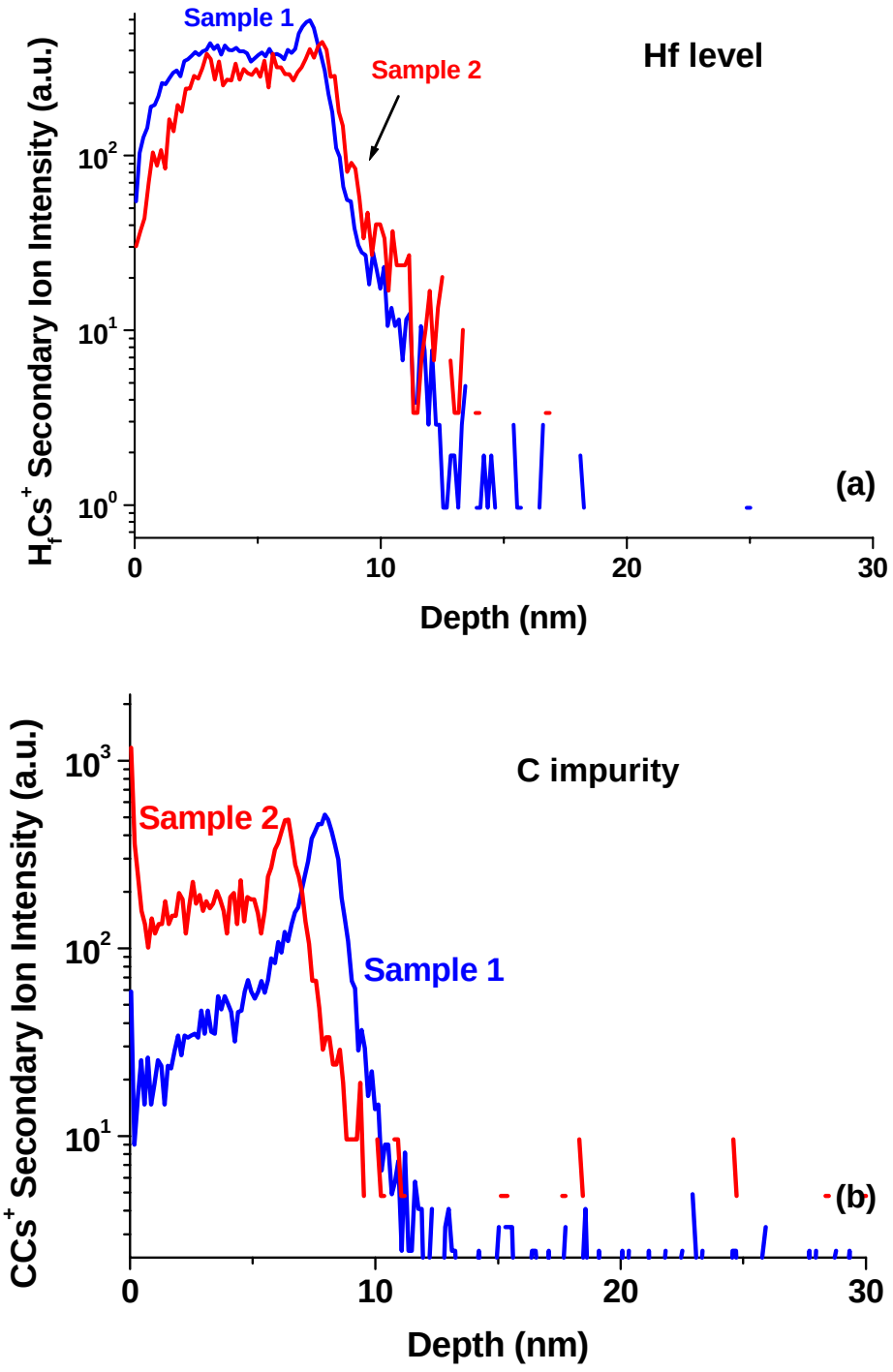


Figure 5.7 SIMS profile of HfO_2 film: (a) Hf level; (b) C level

Chapter 6 Conclusions

A novel wafer scale UHV ALD reactor is constructed to simulate ALD manufacturing environment. The UHV character of the reactor presents an attractive means for elucidating process chemistry in further detail. And the flexibility of the mini-reactor design, based on a movable cap enclosing the wafer in a small volume, enables exploration of various modes of equipment and process cycle operation, including use of the vacuum ballast outside the ALD cap region for enhancing gas pumping in some process cycles, and comparison of ALD processes at different pressures, from viscous to molecular flow. An in-situ QMS is integrated to the ALD reactor, its setup is optimized to achieve real-time ALD process observation.

This reactor is successfully used to investigate wafer scale W ALD process. The details of W ALD process dynamics are directly observed through real-time QMS process sensing. The in-situ sensing results are consistent with ex-situ post-process thickness and uniformity measurements, thus confirming the value of QMS as a quantitative predictor of wafer state, as needed for advanced process control.

This real-time QMS sensing enhances process learning in two ways. First, direct observation of surface saturation for each half-cycle enables the optimization of exposure and purge times in order to minimize ALD cycle time, a critical requirement for ALD manufacturability. Second, the QMS sensing provides a means to assess more subtle interactions which occur in ALD, such as imperfect surface saturation and interactions between exposure-dose recipes for the two reactant species used.

In addition, these results point the way to further accomplishment in understanding and developing ALD processes. Real-time QMS sensing, shown to be a useful indicator of wafer state (e.g., thickness, deposition rate), provides a basis for metrology and process control applications. W ALD film growth can be directly observed by integrating byproduct QMS signal through each precursor exposure and plotting it against ALD cycle number. In-situ QMS sensing results indicate W ALD film growth kinetics strongly depends on wafer state, process temperature and reactor history, as verified by post-process measurement.

QMS measurement is strongly dependent on reactor status, causing a clear first wafer effect after the system is left idle for some time. Such behavior is attributed to competing reactions on the reactor walls where different gas exposure history is encountered. By treating the reactor wall as an additional "virtual" wafer inside the reactor, we are able to well explain the effect of reactor wall on QMS measurement.

The integrated byproduct QMS signal can be also used for ALD film thickness metrology. Despite of the first wafer effect that can be reduced by pre-process reactor treatment, a linear correlation exists between QMS measurement and ALD film thickness for continuous ALD runs, suggesting a promising start of QMS based ALD film metrology.

In-situ process sensing using QCM and QMS is also applied to HfO_2 ALD process that uses metal-organic precursors and is conducted in a furnace type hot wall ALD reactor. The in-situ process sensing results once again demonstrate the complexity of ALD processes. Perfect self-limiting behavior is hardly observed, as proved in the growth rate dependence on precursor dose, purge time and process temperature. The

complexity of ALD process in turn justifies the need for in-situ process sensing to accelerate process learning and enhance manufacturability.

Chapter 7 Future work

7.1 Sensor based advanced process control (APC) for ALD

In-situ downstream process sensing for ALD processes has been demonstrated in this work. Both QMS and QCM are able to reveal essential process dynamics in ALD process, and the integrated byproduct QMS signal can be used to observe film growth in real time. The linear relationship between integrated QMS signal and film thickness also suggests that in-situ sensors can be used for ALD film thickness metrology.

It is also important to develop a sensor based upstream control scheme to monitor reactant state and achieve stable reactant delivery, especially for applications where low vapor pressure metal-organic precursors are used. Our group has previously used an acoustic sensor to control precursor delivery by monitoring the composition of a binary gas mixture, i.e., precursor and carrier gas.[66] The closed-loop feedback from the acoustic sensor to the delivery devices allowed for accurate maintaining of the composition on target, even in the presence of up to 50% fluctuation of the source vapor pressure. The HfO_2 ALD work as discussed earlier also demonstrated that precise and stable precursor control can be achieved via precursor end-point control.

The in-situ process sensing can be further developed by integrating both downstream and upstream control scheme to achieve sensor based advanced process control (APC) for ALD processes in a system level. Take the ternary $HfSi_xO_y$ ALD process for example (Figure 7-1); in-situ QMS, FTIR and QCM sensors can be used downstream for process fault detection, real time film thickness metrology and run to run control. It is also worth trying to correlate in-situ process sensing result with ex-situ film

characterization to predict film quality in real time. For upstream gas delivery, upstream sensors like acoustic sensor can be used to monitor precursor status and sensor based gas delivery system with end-point control can also be used to maintain constant precursor dose. This system can be used as a test-bed for APC in ALD processes by integrating both upstream sensors and downstream sensors in a system level.

7.2 ALD film growth investigation in the nucleation region

Continuous ALD film can only be available after the film growth has passed the nucleation region, therefore limiting the application of ALD processes, especially in the deposition of ultra-thin film applications where thickness of only few atomic layers is needed. In addition to that, ALD film roughness is mostly determined by the nucleation region length. If the spacing between ALD film islands is larger, it will take more time for ALD film to become continuous and the island size (correspondingly the film roughness) will also be bigger. So in order to fully utilize the inherent advantages of ALD processes, it is very important to understand ALD film growth in the nucleation region and consequently know how to reduce or even eliminate the island growth region.

Previous research has used AES, medium energy ion scattering (MEIS) and RBS to investigate ALD film growth in the nucleation region. These approaches are not the direct observation of individual ALD island growth; instead they measure the average ALD film on the substrate and use the film increase per cycle to investigate the island growth without revealing enough microscopic information. If one can directly observe individual ALD island growth (e.g., island size, spacing between islands) as a function of process cycle number, then a clear physical picture of ALD film island growth can be obtained. Techniques such as atomic force microscopy (AFM) and cross-section

transmission electron microscopy (TEM) are potential candidates for the direct observation of ALD film growth.

The length of the nucleation region is mainly determined by the substrate surface status before ALD process. If the substrate surface has more reactive surface groups that are evenly distributed on the surface, ALD film then will become continuous after a shorter island growth region. With shorter island growth region, it is then possible to grow continuous ALD film with smaller film thickness. And the film roughness and the number of defects may also be reduced. So it is also important to investigate the effect of different surface treatment schemes on the ALD film growth, especially in the nucleation region. And the complete physical picture of ALD island growth as obtained by AFM or TEM will assist in the understanding of film growth and the selection of proper surface treatments.

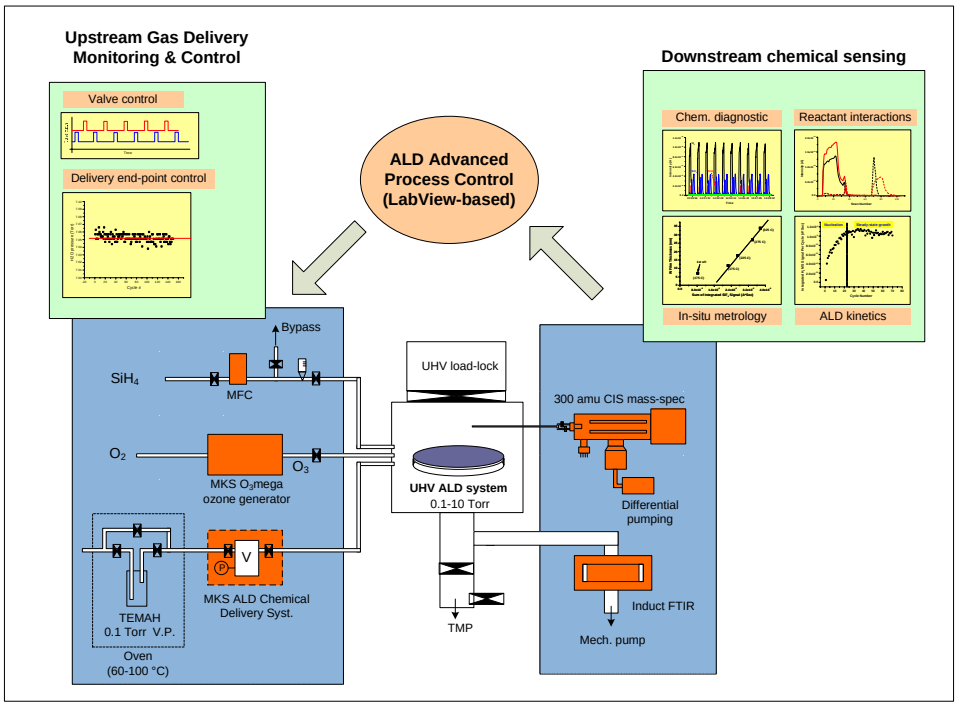


Figure 7.1 Sensor based ALD test-bed for advanced process control (APC)

Bibliography

[1] M. Leskela and M. Ritala, Thin Solid Films **409** (1), 138 (2002).

[2] H. Kim, Journal of Vacuum Science & Technology B **21** (6), 2231 (2003).

[3] L. Niinisto, J. Paivasaari, J. Niinisto et al., Physica Status Solidi a-Applied Research **201** (7), 1443 (2004).

[4] N. D. Hoivik, J. W. Elam, R. J. Linderman et al., Sensors and Actuators a-Physical **103** (1-2), 100 (2003).

[5] M. D. Groner, F. H. Fabreguette, J. W. Elam et al., Chemistry of Materials **16** (4), 639 (2004).

[6] M. L. Green, M. Y. Ho, B. Busch et al., Journal of Applied Physics **92** (12), 7168 (2002).

[7] T. Yamamoto, Y. Izumi, N. Sugiyama et al., Ieice Transactions on Electronics **E87C** (1), 17 (2004).

[8] C. C. Finstad and A. J. Muscat, Ultra Clean Processing of Silicon Surfaces Vii **103-104**, 7 (2005).

[9] T. Seidel, G. Y. Kim, A. Srivastava et al., Solid State Technology **48** (2), 45 (2005).

[10] O. Sneh, R. B. Clark-Phelps, A. R. Londergan et al., Thin Solid Films **402** (1-2), 248 (2002).

[11] M. Ylilammi, Thin Solid Films **279** (1-2), 124 (1996).

[12] M. A. Alam and M. L. Green, Journal of Applied Physics **94** (5), 3403 (2003).

[13] K. Kukli, M. Ritala, T. Sajavaara et al., Advanced Materials **14** (17), A199 (2002).

[14] A. M. Hoyas, J. Schuhmacher, D. Shamiryan et al., Journal of Applied Physics **95** (1), 381 (2004).

[15] A. Satta, A. Vantomme, J. Schuhmacher et al., Applied Physics Letters **84** (22), 4571 (2004).

[16] A. Rautiainen, M. Lindblad, L. B. Backman et al., Physical Chemistry Chemical Physics **4** (11), 2466 (2002).

[17] H. Kim and S. M. Rossnagel, Journal of Vacuum Science & Technology a-Vacuum Surfaces and Films **20** (3), 802 (2002).

[18] J. W. Elam, M. D. Groner, and S. M. George, Review of Scientific Instruments **73** (8), 2981 (2002).

[19] J. W. Klaus, S. J. Ferro, and S. M. George, Thin Solid Films **360** (1-2), 145 (2000).

[20] J. W. Elam, C. E. Nelson, R. K. Grubbs et al., Thin Solid Films **386** (1), 41 (2001).

[21] Y. Luo, D. Slater, M. Han et al., Applied Physics Letters **71** (26), 3799 (1997).

[22] R. Ares, S. P. Watkins, P. Yeo et al., Journal of Applied Physics **83** (6), 3390 (1998).

[23] A. Rosental, P. Adamson, A. Gerst et al., Applied Surface Science **107**, 178 (1996).

[24] M. Schuisky, J. W. Elam, and S. M. George, Applied Physics Letters **81** (1), 180 (2002).

[25] Vincent Stroobant Edmond de Hoffmann, *Mass Spectrometry: Principles and Applications*, 2nd ed. (John Wiley & Son Ltd., West Sussex, 2001).

[26] T. Gougousi, Y. H. Xu, J. N. Kidder et al., Journal of Vacuum Science & Technology B **18** (3), 1352 (2000).

[27] Y. Xu, T. Gougousi, L. Henn-Lecordier et al., Journal of Vacuum Science & Technology B **20** (6), 2351 (2002).

[28] L. L. Tedder, G. W. Rubloff, B. F. Conaghan et al., Journal of Vacuum Science & Technology a-Vacuum Surfaces and Films **14** (4), 2680 (1996).

[29] M. Juppo, A. Rahtu, M. Ritala et al., Langmuir **16** (8), 4034 (2000).

[30] A. Rahtu, T. Alaranta, and M. Ritala, Langmuir **17** (21), 6506 (2001).

[31] M. Juppo, A. Rahtu, and M. Ritala, Chemistry of Materials **14** (1), 281 (2002).

[32] A. Rahtu and M. Ritala, Chemical Vapor Deposition **8** (1), 21 (2002).

[33] R. Matero, A. Rahtu, and M. Ritala, Chemistry of Materials **13** (12), 4506 (2001).

[34] A. Rahtu, K. Kukli, and M. Ritala, Chemistry of Materials **13** (3), 817 (2001).

[35] R. K. Grubbs, N. J. Steinmetz, and S. M. George, Journal of Vacuum Science & Technology B **22** (4), 1811 (2004).

[36] A. Rahtu and M. Ritala, Journal of Materials Chemistry **12** (5), 1484 (2002).

[37] E. Hasunuma, S. Sugahara, S. Hoshino et al., Journal of Vacuum Science & Technology A **16** (2), 679 (1998).

[38] S. M. Rossnagel, A. Sherman, and F. Turner, Journal of Vacuum Science & Technology B **18** (4), 2016 (2000).

[39] J. S. Becker and R. G. Gordon, Applied Physics Letters **82** (14), 2239 (2003).

[40] M. Ritala, M. Juppo, K. Kukli et al., Journal De Physique Iv **9** (P8), 1021 (1999).

[41] A. I. Chowdhury, W. W. Read, G. W. Rubloff et al., Journal of Vacuum Science & Technology B **15** (1), 127 (1997).

[42] G. Q. Lu, L. L. Tedder, and G. W. Rubloff, Journal of Vacuum Science & Technology B **17** (4), 1417 (1999).

[43] S. Cho, L. Henn-Lecordier, Y. J. Liu et al., Journal of Vacuum Science & Technology B **22** (3), 880 (2004).

[44] S. Cho, G. W. Rubloff, M. E. Aumer et al., Journal of Vacuum Science & Technology B **23** (4), 1386 (2005).

[45] Y. Saito and T. Takagi, Journal of the Electrochemical Society **143** (11), 3670 (1996).

[46] K. R. Williams and R. S. Muller, Journal of Microelectromechanical Systems **5** (4), 256 (1996).

[47] R. Canteri and M. Anderle L. Moro, Surf. Interf. Anal **18**, 765 (1992).

[48] M.Fedrizzi M. Bersani, M.Sbetti and M.Anderle, presented at the AIP Conference Proceedings, Woodbury, N.Y., 1998 (unpublished).

[49] M. L. Yu, K. Y. Ahn, and R. V. Joshi, Ibm Journal of Research and Development **34** (6), 875 (1990).

[50] M. L. Yu and B. N. Eldridge, Journal of Vacuum Science & Technology a-Vacuum Surfaces and Films **7** (3), 625 (1989).

[51] S. H. Kim, E. S. Hwang, B. M. Kim et al., Electrochemical and Solid State Letters **8** (10), C155 (2005).

[52] J. W. Elam, C. E. Nelson, R. K. Grubbs et al., Surface Science **479** (1-3), 121 (2001).

[53] W. Lei, L Henn-Lecordier, M. Anderle et al., Journal of Vacuum Science & Technology (Submitted).

[54] J. F. Damlencourt, O. Renault, F. Martin et al., Applied Physics Letters **86** (14) (2005).

[55] H. S. Chang, H. Hwang, M. H. Cho et al., Applied Physics Letters **86** (3) (2005).

[56] D. H. Triyoso, R. I. Hegde, J. Grant et al., Journal of Vacuum Science & Technology B **22** (4), 2121 (2004).

[57] R. K. Grubbs, C. E. Nelson, N. J. Steinmetz et al., Thin Solid Films **467** (1-2), 16 (2004).

[58] G. D. Wilk, R. M. Wallace, and J. M. Anthony, Journal of Applied Physics **89** (10), 5243 (2001).

[59] K. Forsgren, A. Harsta, J. Aarik et al., Journal of the Electrochemical Society **149** (10), F139 (2002).

[60] K. Kukli, M. Ritala, M. Leskela et al., Chemistry of Materials **15** (8), 1722 (2003).

[61] D. Triyoso, R. Liu, D. Roan et al., Journal of the Electrochemical Society **151** (10), F220 (2004).

[62] M. J. Cho, D. S. Jeong, J. Park et al., Applied Physics Letters **85** (24), 5953 (2004).

[63] E. P. Gusev, E. Cartier, D. A. Buchanan et al., Microelectronic Engineering **59** (1-4), 341 (2001).

[64] D. M. Hausmann, E. Kim, J. Becker et al., Chemistry of Materials **14** (10), 4350 (2002).

[65] J. S. Becker, S. Suh, S. L. Wang et al., Chemistry of Materials **15** (15), 2969 (2003).

[66] L. Henn-Lecordier, J. N. Kidder, and G. W. Rubloff, Journal of Vacuum Science & Technology A **22** (5), 1984 (2004).