Run to run control in tungsten chemical vapor deposition using H_2/WF_6 at low pressures

Ramaswamy Sreenivasan
Institute for Systems Research and Department of Chemical Engineering, University of Maryland, College Park, Maryland 20742

Theodosia Gougousi, Yiheng Xu, and John Kidder, Jr.
Institute for Systems Research and Department of Materials and Nuclear Engineering, University of Maryland, College Park, Maryland 20742

Evanghelos Zafiriou
Institute for Systems Research and Department of Chemical Engineering, University of Maryland, College Park, Maryland 20742

Gary W. Rubloff[a)]
Institute for Systems Research and Department of Materials and Nuclear Engineering, University of Maryland, College Park, Maryland 20742

(Received 19 September 2000; accepted 6 August 2001)

Run to run control with an Internal Model Control (IMC) approach has been used for wafer state (thickness) control in the tungsten chemical vapor deposition (CVD) process. The control implementation was preceded by establishing a stable wafer state thickness metrology using *in situ* mass spectrometry. Direct reactor sampling was achieved from an Ulvac ERA-1000 cluster tool module during the H_2/WF_6 W CVD process at 0.5 Torr for temperatures 350–400 °C using a 300 amu closed ion source Inficon Transpector system. Signals from HF product generation were used for in-process thickness metrology and compared to *ex situ*, postprocess thickness measurements obtained by microbalance mass measurements, providing a metrology accuracy of about 7% (limited primarily by the very low conversion efficiency of the process used, ~2%–3%). A deliberate systematic process drift was introduced as a −5 °C temperature change for each successive wafer, which would have led to a major (~50%) thickness decrease over ten wafers in an open loop system. A robust run to run (RtR) control algorithm was used to alter the process time in order to maintain constant HF sensing signal obtained from the mass spectrometer, resulting in thickness control comparable to the metrology accuracy. The efficacy of the control algorithm was also corroborated by additional experiments that utilized direct film weight measurements through the use of the microbalance. A set of simulations in Matlab® preceded the control implementation and helped in tuning the controller parameters. These results suggest that *in situ* chemical sensing, and particularly mass spectrometry, provide the basis for wafer state metrology as needed to achieve RtR control. Furthermore, since the control was consistent with the metrology accuracy, we anticipate significant improvements for processes used in manufacturing, where conversion rates are much higher (40%–50%) and corresponding signals for metrology will be much larger. © *2001 American Vacuum Society.* [DOI: 10.1116/1.1406159]

I. INTRODUCTION

A. Overall goal

The implementation of advanced process control techniques in the semiconductor manufacturing industry is an ongoing challenge.[1–3] Transferring process control technology into semiconductor manufacturing is difficult from a metrology perspective more than from a control theory perspective. There are primarily two reasons for this difficulty.[2] One is the absence of accurate models of semiconductor processes, both first principle and empirical. The lack of first principle models can be attributed to the complicated transport phenomena and reaction kinetics, either of which may dominate depending on the operating range of the process parameters. This complexity makes first principle modeling quite difficult, while assumptions made to simplify the process deprive the model of the accuracy required for conventional feedback control. Empirical models often are useful within the operating range but have to rely on model updating for maintaining accuracy. The lack of availability of good sensor-based metrology is the second reason. Certainly control systems, e.g., in the chemical industry, do require accurate models and good sensor metrology. Sensors are required to feed measurement data to a robust control algorithm. The algorithm, in turn has to operate on manipulated variables (in other words to turn the right knobs). The mass spectrometer is the sensor that this article addresses. The mass spectrometer is a very versatile tool capable of simultaneous spectral analysis of several species in a gas. This article comes out with the message that the mass spectrometer can be used to

[a)]Electronic mail: rubloff@isr.umd.edu

establish a real-time, noninvasive metrology for supporting robust run to run (RtR) control strategies. We have a semiconductor process and a cluster tool. We have an *in situ* sensor namely the mass spectrometer. We start out developing the empirical model, then develop a robust control algorithm, tune the controller through simulations, and implement the control system.

B. Background

Early work justified the need for advanced process control in the semiconductor industry.[4,5] The main factors initially necessitating the need were an improved yield ramp and consistent product quality. Research was done in metrology and methods to integrate them into the processes.[6–10] Simultaneous work went into process modeling.[11] Researchers carried out design of experiments for model identification for various processes.[12] Control architecture was another topic of research.[13] With the preliminary foundations established, control theory for the semiconductor industry was developed, and algorithms mainly incorporating a RtR approach were implemented.[14–17] By run to run control we mean a form of discrete process and equipment control in which the product recipe is modified between runs to keep the controlled variables on target. However no corrective action is performed by the controller at an *in situ* level, i.e., during a run.

C. Benefits of control

Disturbance rejection, setpoint tracking (keeping the control variable on target) and noise rejection are the three primary contributions of a control system. Effectively, this implementation translates into increased productivity, decreased manufacturing cycle time, and improved flexibility. Also in general, integrating the control hardware with the appropriate software capable of efficient data acquisition systems with alarming, trending, and data archiving, rends the control system fault detection capabilities and decreases down time. With this project in mind, the control system is conducive to keeping the controlled variable on target in spite of unmodeled drifts in the process. It is relevant to mention that this control application addresses a very challenging case, in which sensor signals indicative of wafer state metrology were sharply limited by the low conversion rate of the process (2%–3%) and correspondingly small sensor signals associated with reaction production generation and reactant depletion. This is explained in Sec. II C.

II. EXPERIMENTAL APPROACH

A. Cluster tool and chemical vapor deposition (CVD) process

Our research focuses on RtR control on a sophisticated tungsten deposition industrial cluster tool the ULVAC ERA-1000 using the mass spectrometer as the *in situ* sensor. The process in brief (for definition from a controls perspective) is as follows: Reactant gases H_2 and WF_6 are fed to the reactor to accomplish a blanket W CVD process as a result of which tungsten is deposited on the silicon wafer placed in the reactor on a rotating susceptor. The inputs for this process thus are:

(i) WF_6 flow rate of 40 sccm,
(ii) H_2 flow rate of 10 sccm,
(iii) wafer temperature of 400 °C,
(iv) reactor pressure of 40 Torr, and
(v) deposition time of 2–10 min.

The parameter settings mentioned above would be an example of a typical recipe input into the Ulvac panel. The result is a coating of tungsten on the wafer in the reactor.

Tungsten can be deposited on Si surfaces by either H_2 or SiH_4 reduction of WF_6. We are using H_2 reduction in this research work. The reaction[18] takes place in two main steps. The nucleation step: The Si surface on the wafer provides the initial reaction site for WF_6 to form a seed W layer

$$3\mathrm{Si}+2\mathrm{WF_6}\rightarrow 2\mathrm{W}(s)+3\,\mathrm{SiF_4}\uparrow \tag{1}$$

and

$$\mathrm{Si}+\mathrm{WF_6}+\mathrm{H_2}\rightarrow 2\,\mathrm{SiHF_3}\uparrow+\mathrm{W}(s). \tag{2}$$

The formation of this seed layer of W on the entire Si surface initiates further reduction of WF_6 by providing activated surface sites as follows:

$$\mathrm{W^*}(s)+\mathrm{WF_6}\rightarrow \mathrm{WF_6^*}+\mathrm{W}(s), \tag{3}$$

$$\mathrm{WF_6^*}+3\,\mathrm{H_2}\rightarrow \mathrm{W}(s)+6\mathrm{HF}\uparrow, \tag{4}$$

where * denotes an activated surface site. From the stoichiometry it is evident that for each W atom deposited on the surface, three molecules of H_2 and one molecule of WF_6 are consumed, and six atoms of HF are produced. This information forms the basis of our measurement for control. By measuring the evolution of the product HF with a quadrupole mass spectrometer with time using an appropriate metric one can estimate the amount of W deposited using an empirical model between the metric and the weight of W deposited. This information can then be effectively used for run to run control using a robust algorithm.

B. Control problem

The main objective is to keep the weight of the deposited tungsten film as close as possible to a target value in spite of process disturbances in the form of sudden shifts (step changes) or drifts in the input variables. The approach to control starts with measuring the weight (the primary controlled variable) of the film after one run using an electronic balance or by measuring a variable (the secondary controlled variable) which is related to the weight by a well-defined model measured by an appropriate sensor. Then a controller measures the difference between the target and actual value and adjusts a manipulated variable or variables and thus defines the recipe for the next run. The problem definition is as follows:

(i) the primary controlled variable: weight of the tungsten film (W);

(ii) the secondary controlled variable: the normalized integrated mass spec signal; and
(iii) the manipulated variable: deposition time.

Thus we are working with a single input single output system.

It is pertinent to mention that the flow rates of the gases, namely H_2 and WF_6, were considered as choices for potential manipulated variables but were deliberately left out because they could not be varied enough to significantly change the conversion rate. The same can be said of the temperature. The maximum wafer temperature setting allowed for this process is 500 °C (this temperature is different from the true wafer temperature as it is the temperature sensed by a thermocouple placed close to the heating lamps) and settings lower than 450 °C result in very low conversion rates. This 50 °C range however proved useful as an artificial disturbance deliberately introduced into the process to test the control system.

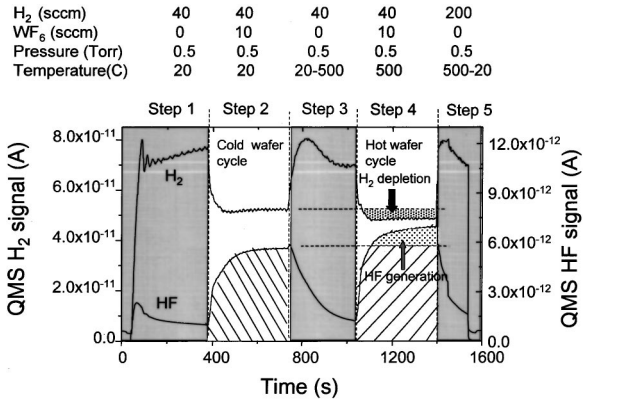

FIG. 1. Typical plot of the mass spectrometer signals for one complete wafer cycle. The sensor metrology metric [Eq. (5)] is derived as a normalized signal related to the HF generation or H_2 depletion areas and takes into account a background in the hot cycle which can be gauged from the cold cycle.

C. Tool constraints

The Ulvac cluster tool was designed for selective W deposition. By ''selective'' we mean a process where W is selectively deposited only on Si and metal surfaces, but not on insulator surfaces such as SiO_2 which is pervasively used in silicon technology. Selective deposition requires low temperatures and pressures. The Ulvac tool was designed for pressures below 1 Torr, using SiH_4 reduction of WF_6.

In contrast, the industry has generally adopted blanket as opposed to selective W CVD processes, which use pressures in the range 40–100 Torr or above and a two-step process sequence comprising initial SiH_4 reduction for the nucleation of a W seed layer followed by H_2 reduction to form most of the W film. With limitations of the tool to pressures below 1 Torr, we achieved only very low conversion rates for the H_2 reduction of the WF_6 process. By conversion rate, we refer to the fraction of reactants which undergo chemical reaction to form new species as products, e.g., conversion of WF_6 to deposited W metal. This low conversion rate limited the accuracy of the *in situ* metrology, which depends directly on conversion of reactants to products.

D. Sensor-based metrology

A quadrupole mass spectrometer has been used as the *in situ* sensor for measuring the extent of deposition. Since product generation and reactant depletion directly reflect the extent of deposition, their measurement in the gas phase can be exploited for sensor based metrology.[10] Whether this sensor is good enough for controls will be further corroborated by the results of the control experiments.

The mass spectrometer is connected directly to the reactor through a gas sampling system, which ensures that a small amount of process gas downstream of the wafer is drawn into the chamber of the mass spectrometer. The mass spectrometer generates a signal which is proportional to the amount of HF in the chamber. Ideally this measurement could be translated into the weight of W deposited through the empirical model developed prior to the control experiments.

Figure 1 shows typical signals. This configuration exhibited a response time of 3–4 s to process gas composition changes in the reactor, which is small compared to the process time of about 360 s.

Note that there is always some residual HF seen by the mass spectrometer, and this HF background increases substantially during the cold wafer cycle. We attribute this HF to wall reactions which occur during flow of the WF_6 and H_2 reactants. Hence using this ''raw'' HF signal would be erroneous. Hence before every reaction it is important to obtain the signals from the chamber and subtract this background signal from the cumulative signal produced during the reaction. This is achieved by subjecting the wafer to a cold cycle where the wafer is not heated and the reagent mixture (H_2/WF_6) is flowed at process pressure over the cold wafer. Hence no reaction takes place, thus simulating the conditions required to obtain the background HF signal.

After the completion of this cycle, WF_6 is diverted through a bypass line and the wafer is heated with heating lamps. Then WF_6 is redirected back into the reactor and the reaction takes place. The mass spectrometer now captures the cumulative HF signal. In our experiments the processing times for both cycles are chosen to be the same. The sensor metrology metric is defined as follows:

$$S_{\mathrm{HF}}=\frac{A_{\mathrm{HW}}(\mathrm{HF})-A_{\mathrm{CW}}(\mathrm{HF})}{A_{\mathrm{CW}}(\mathrm{HF})}*T, \tag{5}$$

where A_{HW} represents the integrated mass spectrometer signal for the hot wafer cycle, A_{CW} reflects the corresponding integrated mass spectrometer signal for the cold wafer cycle, and T represents the deposition time. The deposition time is included here in order to recognize that the thickness (presumed linear in the metrology metric) will depend on time if the rate is constant, and because experiments were done with varying deposition time as well to investigate the accuracy of the metrology. Clearly, the reproducibility of the metrology

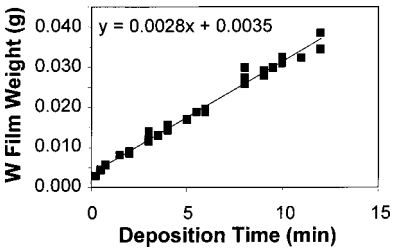

FIG. 2. Linear static model between weight of tungsten deposited and deposition time.

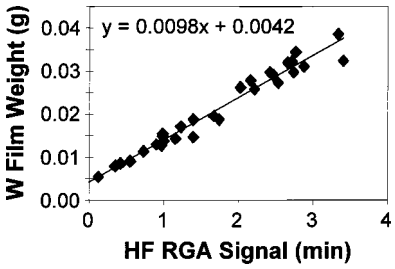

FIG. 4. Linear static model between the normalized integrated mass spectrometer signal and the weight of tungsten deposited. This relationship expresses the extent to which the HF signal provides an accurate metrology signal indicative of the film weight.

metric depends on variation in the relative change in the HF signal, i.e., the fractional term preceding T in Eq. (5).

III. MODELING, CONTROL, AND SIMULATIONS

A. Modeling

The experiments designed for obtaining empirical models were aimed at obtaining three models. 37 wafers were processed for varying deposition times at a wafer temperature of 500 °C. The weight of deposited W was measured with an electronic microbalance as the difference between wafer weights before and after the deposition process. Simultaneous mass spectrometer measurements were also made. The data corresponding to the first wafer at the beginning of each day were omitted to avoid the first wafer effect[10] in the development of the models. The derived models are linear static models.

1. Model between the weight of W and deposition time (Fig. 2)

A linear static model between weight of tungsten deposited and deposition time was obtained. This model was utilized in RtR control experiments to validate the control strategy using *ex situ* process metrology, namely the microbalance.

2. Model between the normalized integrated mass spectrometer signal and deposition time (Fig. 3)

A linear static model between normalized integrated mass spec signal (here the HF generation) and deposition time was developed for use in the control experiments where the mass spectrometer signal is the main controlled output variable. This use of *in situ* metrology for run to run control is the main objective of the paper.

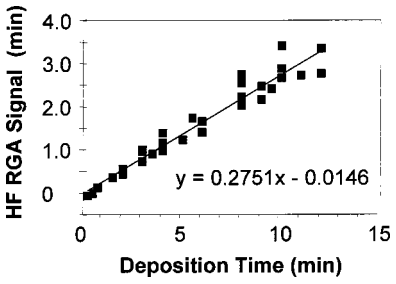

FIG. 3. Linear static model between normalized integrated mass spectrometer signal and deposition time.

3. Model between the weight of W deposited and the normalized integrated mass spectrometer signal (Fig. 4)

A model between the normalized integrated mass spectrometer signal and the weight of tungsten deposited was then constructed from the data obtained for the models mentioned above. This model can be used to check the accuracy of the mass spectrometer. By projecting the mass spectrometer signal value onto the y axis one can obtain the value of the weight of tungsten that should be ideally responsible for the mass spectrometer signal produced. This value of the weight can then be corroborated with the microbalance periodically to ascertain any drift in the mass spectrometer.

B. Control theory

We briefly discuss here the internal model control (IMC) structure for feedback control in order to highlight the control principle and algorithm involved[17] (Fig. 5).

Let the model be represented as

$$\tilde{y}_n = \tilde{\gamma}_{n-1} + \tilde{\beta} x_n, \tag{6}$$

where $\tilde{\gamma}_{n-1}$ denotes the y intercept of the model and $\tilde{\beta}$ denotes the slope of the model. n denotes the run number, and $\tilde{y}_n$ denotes the model output. The models obtained through experiments in the previous section are of the form represented by Eq. (6). These models in Figs. 2 and 3 are used in the control algorithm depending on whether weight is controlled directly or via the HF signal.

Let the process (the real world which in control terminology is referred to as the plant) be represented as

$$y_n = \gamma_n + \beta x_n, \tag{7}$$

where γ_n is the y intercept, β is the slope of the plant, and y_n is the plant output. For the RtR IMC approach, γ_n will be interpreted as the load or disturbance at the output. $\tilde{\beta}$ is fixed but in general different from β due to model uncertainty or error. x_n is the input (recipe variable—the deposition time in this case) which is manipulated to drive y_n to the desired target. $\tilde{\gamma}_{n-1}$ is estimated using the plant data from the past

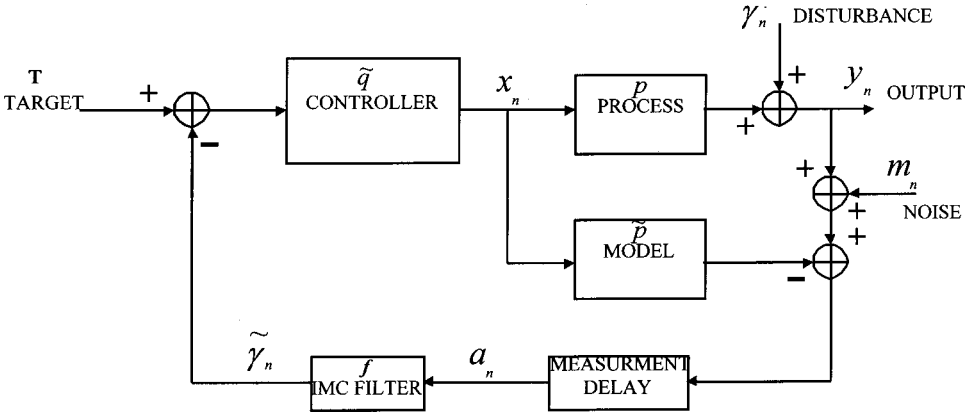

FIG. 5. IMC control structure.

$n-1$ runs. Thus the controller updates the y intercept term of the model after every run. The controller type is chosen to suit the type of disturbance expected.

1. Type 1 exponentially weighted moving average (EWMA)[19,20] controller to handle step-like disturbances

This type is targeted to handling shifts in the process. It is of practical importance to point out that a shift also occurs when the RtR control system is applied to the tool at the very first run because in general the tool is not at steady state, whereas the algorithm assumes the system is at steady state. The controller interprets this initial plant-model output mismatch as a step disturbance in the output and takes corrective action to bring the tool back to steady state. The EWMA controller is based on the following recursive relationships:

$$x_n=\frac{T-\tilde{\gamma}_{n-1}}{\tilde{\beta}}, \tag{8}$$

$$\tilde{\gamma}_n=w(y_n-\tilde{\beta}x_n)+(1-w)\tilde{\gamma}_{n-1};\quad 0<w<1, \tag{9}$$

where T is the target for y_n. w is the weight assigned to the most recent run data and is the only tuning parameter.

Using Eqs. (8) and (9) and by applying z transformation, the algorithm can represented in a z-transfer block diagram. By putting it in a RtR–IMC structure, as shown in Fig. 5, the process and model are represented as

$$p=\beta$$

and

$$\tilde{p}=\tilde{\beta}.$$

The equivalent IMC controller is obtained including a low pass first order filter in the feedback path as shown in Fig. 5.[17] Then we get

$$\tilde{q}=\frac{1}{\tilde{\beta}}\quad \text{(the equivalent IMC controller)} \tag{10}$$

and

$$\tilde{\gamma}_n=\alpha\tilde{\gamma}_{n-1}+a_n(1-\alpha), \tag{11}$$

which physically speaking is an estimate of the drift.

Here we have $\alpha=1-w$. In our algorithm α is used as the tuning parameter and simulations help in fine tuning this value.

2. Type II controller to handle drifts as well as step-like disturbances

The EWMA controller gives no offset for step-like disturbances and step setpoint changes. However in the presence of drifts, it produces an offset resulting in poor performance depending on the modeling error, the amount of drift, and the tuning parameter α. Stefani and Butler[14] used the IMC structure to design RtR predictor corrector controllers, which compensate for the offset, but this produces a lag when measurement delays are present.[17] Note that different design methods can be used with the IMC structures. The IMC design procedure[21] is tailored for use specifically with the IMC structures and is popular in real-time process control applications. Adivikolanu and Zafiriou[17] extended the IMC design procedure to RtR control to develop the type II controller and incorporate the EWMA properties in its design. This results in a controller with zero steady state offset for drifts as well as shifts, whose performance is robust in the present of measurement delays.

The filter equation is then (in the time domain):

$$\tilde{\gamma}_n=\alpha\tilde{\gamma}_{n-1}+\tilde{a}_n(1-\alpha), \tag{12}$$

where

$$\tilde{a}_n=\beta_0 a_n+\beta_1 a_{n-1}+\ldots+\beta_\eta a_{n-\eta}. \tag{13}$$

Solving the filter equation subject to zero steady state offset conditions[17] we get the values for the filter coefficients[17] as

$$\beta_k=\frac{-1}{(1-\alpha)}\cdot\frac{6k}{\eta(\eta+1)(2\eta+1)},\quad k=1,\ldots,\eta \tag{14}$$

and

$$\beta_0=1-\beta_1-\beta_2-\ldots-\beta_\eta. \tag{15}$$

We now have two tuning parameters, namely α and η, which need to be tuned with simulations.

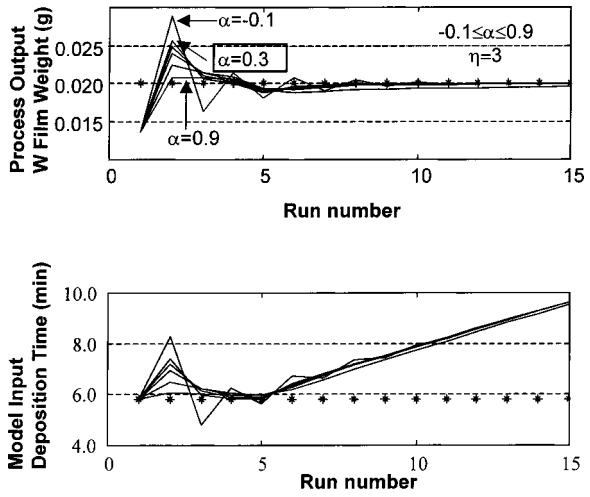

FIG. 6. Simulations used to determine the best value for parameter. A type I controller is used on runs 1–4 to handle the initial process shift. Subsequently, a type II controller is used to handle the process drift introduced at the fifth run. A value of 0.3 for α is finally chosen because it yields the fewest oscillations and a satisfactory response time.

3. An analytical derivation for quantifying the initial shift

Assume we have a target T for the weight of W deposited, for which we go to the model between the weight of W and deposition time and extract the value of the deposition time required for achieving the target. Hence, we use Eq. (8) for $n=1$, with the goal to obtain $\tilde{\gamma}_1=T$. Then we have

$$x_1=\frac{T-\tilde{\gamma}_0}{\tilde{\beta}}, \tag{16}$$

which can be computed from the model parameters in Figs. 6 and 7. When this value of x_1 is fed into the plant (tool), the output is y_1:

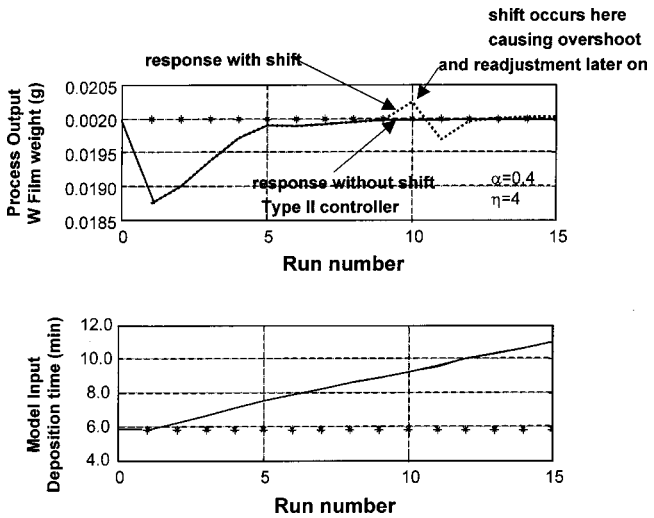

FIG. 7. Simulation showing the type II controller handling a sudden shift in a drifting process. A process shift in the tenth wafer results in a heavier than expected wafer. The controller responds by modifying the deposition time and brings the controlled variable on target after a few runs.

$$y_1=\beta x_1+\gamma, \tag{17}$$

where β and $\gamma=(\gamma_1)$ are the true plant parameters and could correspond to any other linear model within the bounds set by the data in Fig. 2. This output value is different from the value predicted by the model due to plant–model mismatch. This difference is what the controller views like a step-like disturbance at the start, which we call the initial shift.

The shift magnitude is

$$(\tilde{y}_1-y_1)=(\beta-\tilde{\beta})x_1-(\gamma-\tilde{\gamma}_0). \tag{18}$$

Figure 2 can be used to obtain bounds on the possible values of β and γ, which can then be used with Eqs. (18) and (16) to estimate *a priori* the possible magnitude of the initial shift.

C. RtR experiment objectives

As mentioned earlier, the linear static models developed during the modeling stage are used as the models in the IMC structure depending on whether weight is controlled directly or through the HF signal. In all experiments, an artificial drift was introduced into the process to test the control system. This drift is in the form of a temperature drift where the wafer temperature was reduced by a fixed amount after every run. As mentioned earlier, in reality, it is the thermocouple temperature which is measured and decreasing the setpoint for the thermocouple results in decreasing the power supplied to the heating lamps and as a consequence the wafer temperature is also reduced. To the controller this drift is an unmodeled drift. The effect of reducing the wafer temperature is that deposition rate slows down with lower temperatures and without the controller taking the necessary action, the weight of the W or the normalized integrated mass spectrometer signal value would go down with each run.

Two sets of experiments were carried out to assess the control system behavior. The first set of experiments aimed at directly controlling the weight of W deposited to a target value. This meant from the algorithm's point of view that we use the linear static model between the weight of W and the deposition time (Fig. 2) in the algorithm. The input into the algorithm is then the weight from the previous run and the output is the deposition time suggested by the controller for the next run so that the weight of the W deposited moves closer to the target value in spite of the disturbances (drift or shift) present.

The second set of experiments is aimed at controlling the normalized integrated mass spectrometer signal to a target value. This value can be arrived upon in more than one way, the most obvious of which is starting out with the target weight and then calculating the target normalized integrated HF signal from the model between the weight and the HF signal (Fig. 4). Running the tool at the deposition time for this target value of the HF signal would result in a plant–model mismatch and the controller would have to correct for an initial shift. This was avoided by processing a single wafer at a deposition time of 5 min. The value of the HF signal for this first wafer was taken as the setpoint for the control algorithm, thus enabling the controller to ''view'' the tool as

if it were in a steady state already. This set of experiments uses the model between the HF signal and deposition time (Fig. 3) in the IMC algorithm. The input would then be the HF signal from the previous run and the output would be the deposition time suggested by the controller for the next run, so that the HF signal moves closer to the target value in spite of the disturbances (drift or step) present. As a consequence the weight of W deposited also moves closer to its target value.

D. Simulations to tune the controller

The simulations were done using Matlab®[22] for coding the filter equations and the recursive relationships. The goal was to test the controller for various cases of mismatch between the model used in the RtR algorithm and the actual process, and to make sure the response to the systematic drift was neither aggressive nor too sluggish in all cases. These two response characteristics are dependent on the controller parameters α and η and hence simulations were done to select appropriate values for these parameters which gave a favorable response.

From Figs. 2 and 3 left and right hand side bounds on the values of the slopes of the models were obtained as 0.0020/0.0036 for Fig. 2 and 0.2341/0.3061 for Fig. 3. These values were used as the slopes for the actual process in the simulations to constitute the worst case scenario of mismatch with the model. Simulations were also done for several processes inbetween these bounds and for the case when there was no mismatch. This was done with both type 1 and type II controllers. In preparing for the set of experiments, where the objective was to control the film weight directly, simulations helped in determining good values for $\alpha=0.3$ and $\eta=3$. As an example, Fig. 6 shows several simulations for the case with the worst mismatch. $\eta=3$ is used for all cases but varying values of α are simulated.

In order to tune the controller for the next set of experiments, where the objective was to control the mass spectrometer signal to a steady state target value, the model and possible mismatch situations being different, simulations helped in obtaining values for $\alpha=0.4$ and $\eta=4$.

Figure 7 gives an insight into how the controller works. This simulation was carried out to show the response of a type II controller to process anomalies manifested as sudden step or a shift-like disturbance as encountered in process drift. The simulation shows a type II controller compensating for a steady temperature drift by ramping up the deposition time in a calculated manner so that the target is reached at the ninth run. This is shown on the same plot by simulating the exact same controller without the shift. Then the simulation is repeated by introducing a shift in the tenth wafer. The weight of the wafer is unexpectedly heavier than expected for that temperature. The plot shows the controller overshooting the target due to this unexpected shift in the process but recovers quite quickly and is back on target within the next couple of runs. The controller encounters such a situation in one of the RtR control experiments in Sec. IV where the objective is to keep the HF signal on target and it is further discussed there.

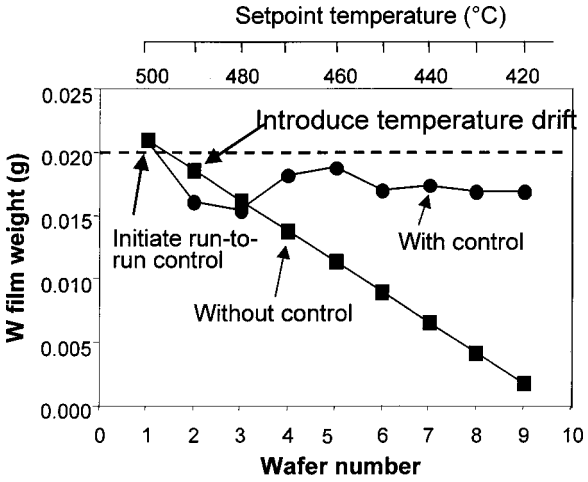

FIG. 8. Weight of tungsten vs run number. Prespecified weight as set point. Type II controller used throughout. A drift of −10 °C per run was introduced starting with the second run. (Experiment to validate control system.)

IV. EXPERIMENTAL RtR CONTROL RESULTS

The experiments included two parts. First, experiments were done (Sec. IV A) to validate the control methodology, using *ex situ* measurements of the weight of the deposited film as obtained on the microbalance as the controlled variable. Second, experiments were then carried out (Sec. IV B) using the *in situ*, real-time mass spectrometer metrology as the controlled variable. If this metrology is sufficiently indicative of the actual deposited weight, then its control should imply consequent control of the film weight.

A. Validating the control strategy with *ex situ* postprocess metrology

To validate the control methodology, *ex situ*, postprocess metrology was employed in the form of a microbalance used to measure the W film weight. The model in Fig. 2 is used for control. Figures 8 and 9 show the results of two such experiments. In both cases the target was set to 0.02 g of W,

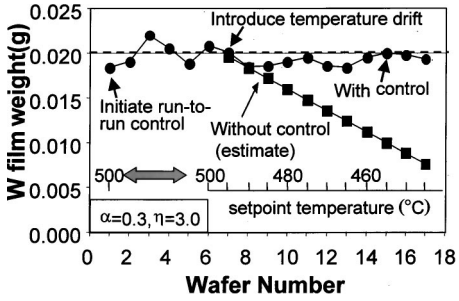

FIG. 9. Weight of tungsten vs run number. Prespecified weight as set point. A drift of −5 °C per run was introduced starting with the second run. Initial process–model mismatch shift handled by type I controller (takes six runs). Type II controller started in the seventh run. (Experiment to validate control system.)

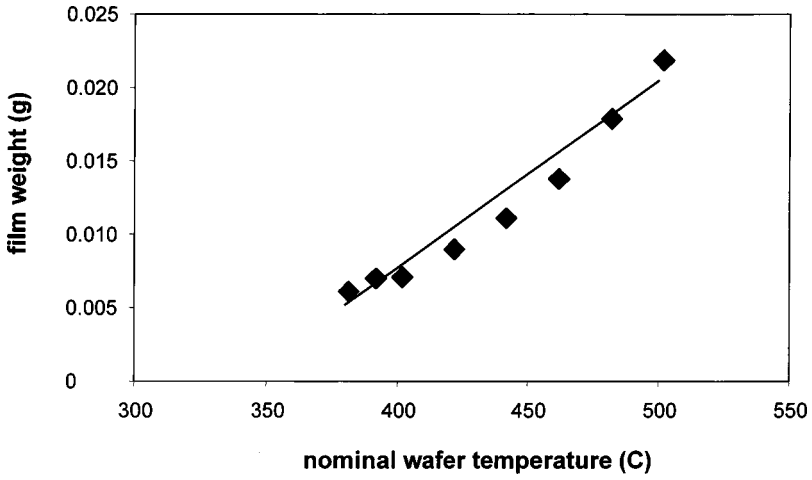

FIG. 10. Model of the process drift obtained by fitting a straight line through data points obtained by running the tool at temperatures lower than 500 °C without implementation of the RtR control system. The deposition time was 6 min.

which meant the deposition time for the first wafer was 5.82 min (from the model between weight of W and deposition time).

Temperature drifts were intentionally introduced as representative of process drift, which should be correctable by run to run control. A model for these drifts is given in Fig. 10, which shows data points which were obtained by running the tool at temperatures lower than 500 °C at a deposition time of 6 min. A straight line was then fitted through the data points. However, this model was used only for estimating the open loop behavior and comparing it to the controller's performance to show the performance of the controller. It was not used in the controller design nor by the control algorithm during implementation.

The performance of the controller can be compared to the estimate of the behavior in the absence of control shown by the line labeled ''w/o control (estimate)'' in Figs. 8 and 9. The points on this line are derived by linear interpolation from the model depicted in Fig. 10.

Run to run control was initiated at the very first run. In Fig. 8, the first run resulted in a weight difference from 0.02 g due to mismatch between the model and the actual process, causing the controller to take corrective measures. The type II controller used here overcorrects for this initial shift. The temperature drift of −10 °C per run started from the second run. The controller recognized the drift within the next two runs and took corrective action, preventing the weight from going down in spite of very low substrate temperatures. The offset may be attributed to the very low temperatures at which even very long deposition times do little to increase the deposition of W due to very slow reaction kinetics. Thus −10 °C per wafer was too large a drift for optimal control.

Figure 9 shows another series of experiments. The drift size was reduced to −5 °C per run. It was decided to have a type I controller (which is more adept at handling shifts) to handle the initial shift due to plant–model mismatch at the first run. The type I controller brought the tool back to steady state after six runs and the drift was introduced at the seventh run and the type II controller started. The controller performed very well and handled the drift keeping the film weight almost to within 3% of the target of 0.02 g, which could very well be attributed to the uncertainty in the electronic balance.

The experiments performed above should not mislead the reader into thinking that it is important to know when the drift occurs for the implementation of control. This is definitely not the case. The two experiments described above were performed with the aim of testing the control strategy and hence we exploited the fact that we could introduce the drift artificially at any run. In the actual implementation phase (Sec. IV B), we implement the control system to a tool at steady state and use type II controller right from the beginning. This implementation represents the realistic case where the drift may occur at any time. Another realistic implementation would be using a type I controller to handle the initial shift for a tool not at steady state (which means the target is a prespecified setpoint) for a few runs and then switching over to a type II controller. Again the drift may occur at any time.

B. Applying run to run control with *in situ* process metrology

In situ process metrology based on the real-time HF product signal generated in the reaction was used as the (secondary) controlled variable for run to run control, with the expectation that control of the HF signal would also mean control of the weight of the deposited film (primary controlled variable). In order to give the controller a plant at steady state, the tool was run to process a wafer and the signal measured. This signal was taken as the setpoint and fed to the control algorithm. The controller started taking action from the next run onwards. Only the type II controller was used for these experiments. The drift size was −5 °C per wafer and was started at the second wafer.

Figure 11 shows the result of such an experiment. The target normalized integrated mass spectrometer signal was obtained after the first run as 1.2735 min. The controller did a good job driving the signal back to the target, even though the substantial process (i.e., temperature) drift occurred during the sequence of wafers processed.

The performance of the run to run control system is evi-

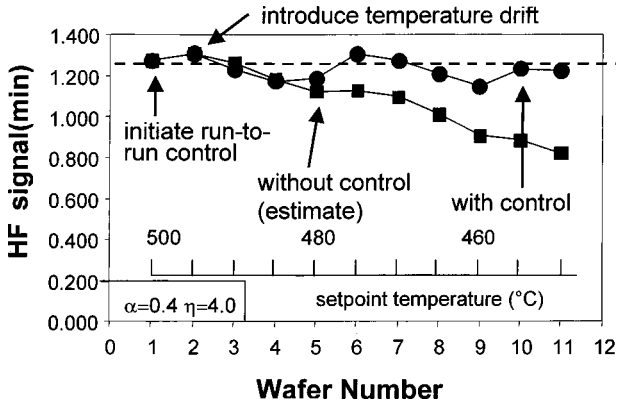

FIG. 11. Normalized integrated HF signal vs run number. Type II controller used throughout.

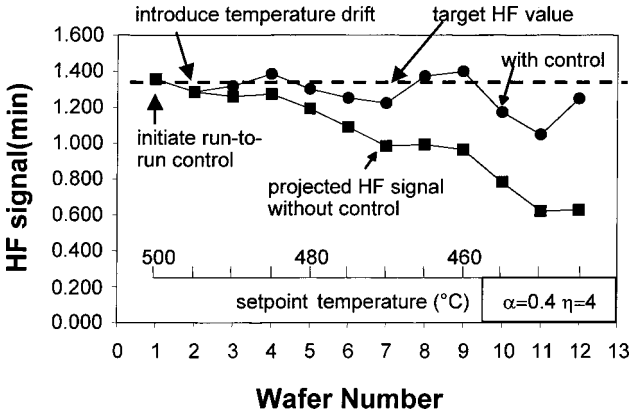

FIG. 13. Normalized integrated HF signal vs run number. Type II controller used throughout. The variable under direct control is the HF signal.

dent in comparing it to the expected signal (''w/o control (estimate)'') which would have occurred without control. The latter, as shown in Fig. 11, was estimated from the HF signal for 5 min of deposition at 500 °C normalized to the actual deposition time used (where deposition time was the control variable). Analytically the estimated value of the HF signal at the nth run would be

$$[S_{\mathrm{HF}}]\text{estimate at } n\text{th run} = \left\{ \frac{[A_{\mathrm{HW}}(\mathrm{HF})]_n - [A_{\mathrm{CW}}(\mathrm{HF})]_n}{[A_{\mathrm{CW}}(\mathrm{HF})]_n} * T_n \right\} * \frac{T_{500\,^\circ\mathrm{C}}}{T_n}. \quad (19)$$

During the sixth run a process anomaly temporarily slowed down the drift (as indicated by the square markers for estimated behavior without control) and caused the controller to overshoot the target. The controller once again adjusted the deposition time to bring back the signal back on target by the tenth run. The simulation plot shown earlier in Fig. 7 provides a good insight to controller behavior when handling sudden process shift-like anomalies during a steady drift.

Figure 12 shows that the resulting weight of deposited W was reasonably controlled after the anomaly during the sixth run, but this anomaly apparently invalidated the original model relating the mass spectrometer signal to the deposited film weight. Clearly, unmodeled shifts in metrology signals, which are intended to indicate desired performance characteristics, can present a problem in the application of run to run control.

Figure 13 shows the results for a similar set of experiments, in which such anomalies were smaller, if present at all. The resulting control on film weight, as seen in Fig. 14, was notably improved in comparison to the results in Fig. 12. The variation of film weight in Fig. 14 appears comparable to the variation in HF signal in Fig. 13, indicating that the run to run control methodology was effective, and constrained primarily by the quality of the thickness metrology available, i.e., the accuracy of the thickness versus HF signal relationship represented by the model in Fig. 4. A similar conclusion could be reached from the experiments in Figs. 11 and 12, except perhaps for the complexity introduced by the anomaly at wafer 6.

V. DISCUSSION

Although the run to run variation in controlled variables (either film weight in Figs. 8 and 9 or HF signal in Figs. 11 and 13) was significant and larger than one would desire for manufacturing applications, it is clear by comparison to the

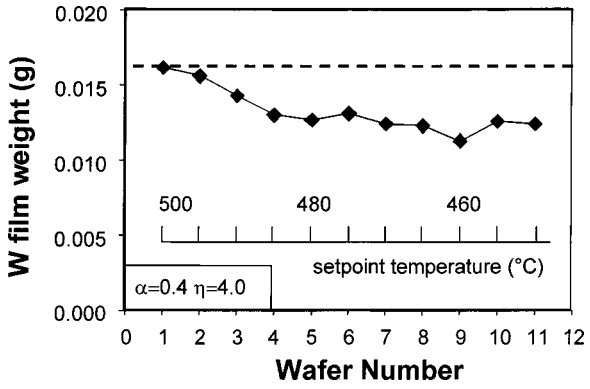

FIG. 12. Weight of tungsten (not under direct control) vs run number for the same experiment (Fig. 11) where the variable under direct control is the HF signal.

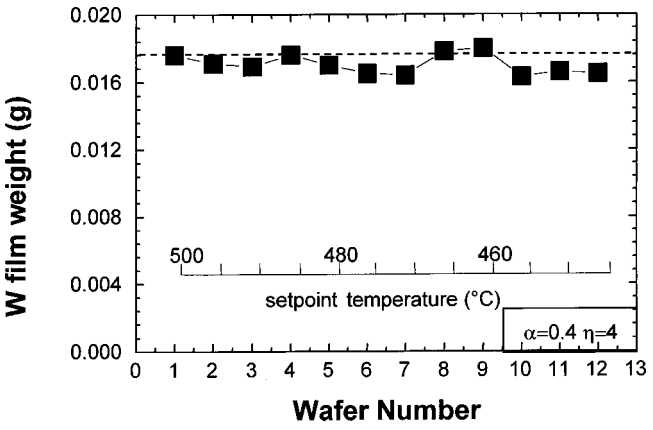

FIG. 14. Weight of tungsten (not under direct control) vs run number for the same experiment (Fig. 13) where the variable under direct control is the HF signal.

''w/o control'' curves that the run to run approach is beneficial in compensating for system drift: the ''with control'' cases all exhibited significantly better stability than the ''w/o control'' cases. This underscores the fact that the run to run control methodology has value in applications where process drift is a major concern.

Indeed, one limitation in this work is the quality of the mass spectrometer metrology, which in turn was constrained substantially by the low conversion efficiency of the process used. As mentioned above, the reactant conversion rates were on the order of 2%–3%. Since the mass spectrometer metrology is totally based on quantitation of product generation or reactant depletion, such low conversion rates mean less-than-adequate *in situ* metrology. Further work will address processes with much higher conversion rates (30%–50%), as typically employed in manufacturing, e.g., for blanket W CVD (as opposed to selective W CVD conditions studied here).

It is also evident from Figs. 8 and 9 that even with the direct film weight measurement (*ex situ*, postprocess) as the controlled variable, there remains notable variation in this variable. We attribute this variation to random/stochastic process drift, as contrasted with systematic, monotonic, and smoothly varying drift of the kind we attempted to introduce by the $-5\,^{\circ}C$ wafer temperature drift.

For the exercise of run to run control, our methodology adjusted the *y* intercept of the model (Fig. 2 or 3) after every run. One might argue that it would be more natural to adjust the slope, rather than the intercept, because the chemistry of the process is likely to cause drift in the model to reflect the rate at which deposition occurs, or at which the sensor sensitivity drifts. However, the purpose of process control research here has been to develop robust control strategies, which are insensitive to errors either in the accuracy of the model or arising from noise in experimental measurements. Therefore, we chose not to exploit additional physical insights in making the control model. Nevertheless, reasonable control performance was achieved, as measured against the intrinsic accuracy of the sensor metrology in use.

Finally, it is important to distinguish between the application of run to run control for compensation of systematic, steady process drift as opposed to random, stochastic changes in process behavior. Run to run control is intended to compensate for steady, systematic process drifts because it exploits measurement of a controlled variable over time and modeling to predict where the drift is headed. In contrast, run to run control for batch processes such as this relies on only one measurement for each wafer/batch: if the measurement has a substantial random component due to stochastic sources, it becomes unreliable as a controlled variable to achieve run to run process control.

VI. CONCLUSIONS

These investigations have shown that *in situ* chemical sensing and robust run to run control can be successfully exercised in a difficult system. The chemistry of the process as practiced limited reactant conversion rates to 2%–3%, severely limiting the signal available in the form of reaction product generation (HF) or reactant depletion (H_2 or WF_6), and this resulting in a thickness metrology based on mass spectrometry sensing, which was reproducible only to about $\pm 7\%$.[10] Nevertheless, both the HF product signal measured by mass spectrometry (the secondary controlled variable) and the film weight (the primary controlled variable) were stabilized to about $\pm 10\%$, i.e., nearly as good as could be expected from the reproducibility of the thickness metrology.

The run to run control methodology is aimed at compensating for systematic drifts in process behavior. Because of these metrology limitations, testing of the control methodology could only take place if significant process drift were to occur, larger than should be expected in manufacturing tools. Therefore we introduced an intentional process drift in the form of a change in wafer temperature with each successive wafer, $-5\,^{\circ}C$ per wafer, which was sufficient to substantially alter the deposition rate compared to the reproducibility specification of the mass spectrometer metrology. The control system was successful in largely compensating for this drift, demonstrating that such control could be of value in the presence of gradual process drift.

Applications in manufacturing typically do not exhibit drifts of this magnitude. However, reactant conversion rates in manufacturing are much higher than those achieved here, on the order of 30%–50%. Fortunately, this means that the *in situ*, mass spectrometry based metrology should be considerably more reproducible than the $\pm 7\%$ characteristic seen here. The prognosis, then, is that controllability will improve with the reproducibility of the metrology, and in this sense further experimentation in processes with much higher conversion rates are a high priority for research.

Another important issue for process control is the distinction between systematic drift and random variability in process behavior. Chemical processes have many complexities, such as the presence of wall reactions and the dynamic behavior of complex equipment components, leading to both systematic drifts and random variability. Run to run control can in principle succeed in compensating for systematic, smooth drifts which are slow compared to the measurement frequency, which for run to run control is the time from one wafer to the next. The temperature drift intentionally introduced here met this requirement, so that control could be demonstrated.

However, random variability in process behavior, or more rapid systematic drift, would not be expected to be correctable using run to run control. Therefore, in pursuing the application of *in situ*, real-time mass spectrometer sensing for wafer state metrology, it would be important to investigate real-time control strategies as well as run to run. Because the mass spectrometer senses real-time changes in the process environment, it may provide the capability for real-time process control or process end-point control, provided reaction signals are sufficiently large (product generation or reactant depletion), and that signal processing and control calculations can be done fast enough. Future research will explore these directions.

[1]S. W. Butler, J. Vac. Sci. Technol. B **13**, 1917 (1995).
[2]N. A. Chaudry, R. Telfeyan, B. Moore, H. Etemad, and J. Moyne, Proceedings of the Semiconductor Research Corporation Technical Conference, Atlanta, GA, September 1993.
[3]G. Barna and S. W. Butler, *Semiconductor Fabtech*, 4th ed. (ICG, London, 1996), p. 59.
[4]M. Liehr and G. W. Rubloff, J. Vac. Sci. Technol. B **12**, 2727 (1994).
[5]S. W. Butler, Proceedings Semicon/West San Francisco, CA, July 1994, p. 82.
[6]S. W. Butler, AIP Conf. Proc., 1998, p. 47.
[7]L. L. Tedder, G. W. Rubloff, B. F. Conaghan, and G. N. Parsons, J. Vac. Sci. Technol. A **14**, 267 (1996).
[8]L. L. Tedder, G. W. Rubloff, I. Shareef, M. Anderle, D. H. Kim, and G. N. Parsons, J. Vac. Sci. Technol. B **13**, 1924 (1995).
[9]D. W. Greve, T. J. Knight, X. Cheng, B. H. Krogh, and M. A. Gibson, J. Vac. Sci. Technol. B **14**, 489 (1996).
[10]T. Gougousi, Y. Xu, J. N. Kidder, Jr., G. W. Rubloff, and C. Tilford, J. Vac. Sci. Technol. B **18**, 1352 (2000).
[11]S. W. Butler, J. Stefani, M. Sullivan, S. Maung, G. Barna, and S. Henck, J. Vac. Sci. Technol. A **12**, 1984 (1994).
[12]G. S. May, IEEE Trans. Semicond. Manuf. **4**, 83 (1991).
[13]M. Sullivan, S. W. Butler, J. Hirsch, and C. J. Wang, IEEE Trans. Semicond. Manuf. **7**, 134 (1994).
[14]S. W. Butler, J. Stefani, and G. Barna, *Proceedings of the 1993 American Control Conference* (IEEE, Piscataway, NJ, 1993), p. 3003.
[15]A. Hu, E. Sachs, and A. Ingolfsson, 1992 IEEE/SEMI Intl. Semiconductor Manufacturing Science Symposium, p. 73.
[16]E. Sachs, A. Hu, and A. Ingolfsson, IEEE Trans. Semicond. Manuf. **8**, 26 (1995).
[17]S. Adivikolanu and E. Zafiriou, IEEE Trans. Electronics Packaging Manuf. **23** (2000).
[18]J. J. Hsieh, J. Vac. Sci. Technol. A **11**, 3040 (1993).
[19]J. S. Hunter, J. Quality Technol. **18**, 203 (1986).
[20]A. Ingolfsson and E. Sachs, J. Quality Technol. **25**, 271 (1993).
[21]M. Morari and E. Zafiriou, *Robust Process Control* (Prentice-Hall, Englewood Cliffs, NJ, 1989).
[22]MatLab is commercially available from the Mathworks (www.mathworks.com).