

# IN-SITU FAULT DETECTION OF WAFER WARPAGE IN LITHOGRAPHY<sup>1</sup>

Arthur TAY<sup>2</sup>, Weng Khuen HO<sup>2</sup>, Christopher YAP<sup>3</sup>,  
Chen WEI<sup>3</sup>, and Kuen-Yu TSAI<sup>4</sup>

<sup>2</sup>*Dept of Electrical and Computer Engineering, National University of Singapore (NUS), 10 Kent Ridge Crescent, Singapore (119260)*

<sup>3</sup>*Dept of Mechanical Engineering, NUS*

<sup>4</sup>*Intel Corporation*

**Abstract:** Wafer warpage is common in microelectronics processing. Warped wafers can affect device performance, reliability and linewidth control in various processing steps. We proposed in this paper an in-situ fault detection technique for wafer warpage in lithography. Early detection will minimize cost and processing time. Based on first principle thermal modeling, we are able to detect warpage fault from available temperature measurements. In this paper, the use of advanced process control resulted in very small temperature disturbance making it suitable for industrial implementation. More importantly, the sensitivity for detecting warpage is not compromised even though the temperature signal is small. Experimental results demonstrate the feasibility of the approach. *Copyright © 2005 IFAC.*

**Keywords:** Feedforward Control, Temperature Control, Disturbance Signal, Fault Detection, Models

## 1. INTRODUCTION

Wafer warpage is common in microelectronics processing. Warpage can affect device performance, reliability and linewidth or critical dimension control in various lithographic patterning steps. Figure 1 shows a typical lithography sequence (Quirk and Serda, 2001). In the exposure step a warped wafer results in a non-uniform gap between wafer and mask for a contact aligner and different depth-of-focus at different parts of the wafer for a projection exposure system. The significance of having a wafer of minimal warpage is that it makes possible the reduction of the depth-of-focus to achieve a higher pattern resolution (Quirk and Serda, 2001).

Current techniques of measuring wafer warpage include capacitive measurement probe (Poduje and Balies, 1988) shadow Moire technique (Wei et al., 1998), and pneumatic-electro-mechanical technique (Fauque and Linder, 1998). These are off-line methods where the wafer has to be removed from the processing equipment and placed in the metrology tool resulting in

increase processing steps, time and cost. In this paper, we propose an in-situ fault detection technique for wafer warpage using available temperature signals. We first note that there are a number of low temperature thermal processing steps in the lithography sequence (Schaper et al., 1999) as shown in Figure 1. Thermal processing of semiconductor wafers is commonly performed by placement of the substrate on a heated bakeplate for a given period of time. The heated bakeplate is held at a constant temperature by a feedback controller that adjusts the heater power in response to a temperature sensor embedded in the bakeplate near the surface. The wafers are usually placed on proximity pins of the order of 100 to 200  $\mu\text{m}$  to create an air-gap so that the bakeplate will not contaminate the wafers. As wafers can warp up to 100  $\mu\text{m}$  centre to edge, the percentage change in the air-gap between the wafer and bakeplate can be substantial (see Figure 2) resulting in a non-insignificant variation in the bakeplate temperature.

When a wafer at room temperature is placed on the bakeplate, the temperature of the bakeplate drops at first but recovers gradually because of closed-loop control (see Figure 3). In Figure 3 where a flat wafer was

---

<sup>1</sup> This research was supported by A\*Star under RP. 3979907.

used for Experiment (b) and warped wafers were used for Experiments (c) and (d) also shows that different air-gap sizes gave different magnitudes of temperature disturbances. Hence warpage can be detected from the temperature disturbances. Using temperature to detect warpage is also given in Ho et al. (2004). However, in Ho et al. (2004), the temperature disturbance used as the signal for warpage computation had magnitude of more than a degree Celcius. This may not be implementable in practice as the temperature for some baking steps such as the post-exposure bake have to be controlled to  $\pm 0.1^\circ\text{C}$  (Schaper et al., 1999). In this paper, advanced process control such as feedforward control (Ho et al., 2000) reduced temperature disturbance by 10 times to a tenth of a degree Celcius. We demonstrate that warpage can still be detected and more importantly, there is no loss in sensitivity in the detection algorithm even though the temperature disturbance that provided the signal for detecting warpage has been reduced by about 10 times. Since bakeplate temperature measurements are already available, the proposed method can be implemented without increasing system complexity and equipment cost.

## 2. THERMAL MODEL

In this section, we present the model for the baking process which can be used to estimate wafer warpage. For our purpose, a one-dimensional analysis will be used to characterize the dynamics of heat transfer for a silicon wafer at room temperature placed on a bakeplate maintained at a steady-state temperature. A similar model was derived in Ho et al., (2004). However, radiation was neglected in Ho et al., (2004). In this paper heat transfer through radiation is considered to give a more accurate model.

A schematic of the system under consideration is shown in Figure 2 where 2(a) shows the baking of a flat wafer and 2(b), the baking of a warped wafer. The system consists of 3 basic sections: the bakeplate, the air-gap and the silicon wafer. We assumed that the wafer and bakeplate have the same diameter. The temperature distribution within each section is assumed at any instant to be sufficiently uniform such that it can be considered to be a function of time only i.e. ‘‘lumped’’ model approach. Radiative heat transfer between the bakeplate and ambient, and wafer and ambient are considered. Heat transfer between the bakeplate and wafer due to radiation can be neglected safely as the temperature difference is small. With this approximation, the energy balance of the wafer and bakeplate are given by

$$m_w c_w \dot{T}_w = \frac{k_a A_w}{l_a} (T_p - T_w) - h A_w T_w - \epsilon_w A_w \sigma ((T_w + T_\infty)^4 - T_\infty^4) \quad (1)$$

$$C_p \dot{T}_p = p - \frac{k_a A_w}{l_a} (T_p - T_w) - h A_p T_p - \epsilon_p A_p \sigma ((T_p + T_\infty)^4 - T_\infty^4) \quad (2)$$

where for the 200 mm wafer and the given bakeplate mass of wafer:  $m_w = 55.3 \text{ g}$

specific heat capacity of wafer:  $c_w = 0.73 \text{ kJ/kgK}$   
total heat capacity of bake plate:  $C_p = 750 \text{ J/K}$   
thermal conductivity of air:  $k_a = 0.012 \text{ W/mK}$   
natural convection coefficient:  $h = 10 \text{ W/m}^2\text{K}$   
emittance of wafer:  $\epsilon_w = 0.67$   
emittance of bakeplate:  $\epsilon_p = 0.65$   
Stefan-Boltzmann constant:  $\sigma = 5.67 \times 10^{-8} \text{ W/(m}^2\text{K}^4)$   
surface area of wafer:  $A_w = 0.0314 \text{ m}^2$   
surface area of the side of bake plate:  $A_p = 0.11 \text{ m}^2$   
wafer temperature above the ambient:  $T_w$   
bakeplate temperature above the ambient:  $T_p$   
ambient temperature:  $T_\infty$   
heater power:  $p(t)$   
nominal air-gap thickness:  $l_a$   
bakeplate proximity pin length:  $l_{pp}$

Most thermophysical properties are temperature dependent. However, for the temperature range of interest from  $15^\circ\text{C}$  to  $150^\circ\text{C}$ , it is reasonable to assume that they remain fairly constant as given in the above. They can be obtained from handbooks (Ozisik, 1985; Raznjevic, 1976).

Figure 2 shows the nominal air-gap,  $l_a$ . Notice that for a flat wafer the nominal air-gap,  $l_a$ , is equal to the height of the bakeplate proximity pins,  $l_{pp}$ . We next define  $\lambda_{wp}$  as a measure of the degree of warpage where

$$\lambda_{wp} = l_a - l_{pp}$$

The  $T^4$  dependence of radiant heat transfer complicates calculations. Equations (1) and (2) can be linearized as follows (Ozisik, 1985).

$$\begin{aligned} m_w c_w \dot{T}_w &= \frac{k_a A_w}{l_a} (T_p - T_w) - h A_w T_w \\ &- \epsilon_w A_w \sigma ((T_w + T_\infty)^2 + T_\infty^2) (T_w + 2T_\infty) T_w \\ &\approx \frac{k_a A_w}{l_a} (T_p - T_w) - h A_w T_w \\ &\quad - 4\sigma \epsilon_w A_w T_{mw}^3 T_w \\ &= \frac{k_a A_w}{l_a} (T_p - T_w) - (h + h_{rw}) A_w T_w \quad (3) \end{aligned}$$

where  $h_{rw} = 4\sigma \epsilon_w T_{mw}^3$  and  $T_{mw}$  is the mean of  $T_w + T_\infty$  and  $T_\infty$ .

$$\begin{aligned} C_p \dot{T}_p &= p - \frac{k_a A_w}{l_a} (T_p - T_w) - h A_p T_p \\ &- \epsilon_p A_p \sigma ((T_p + T_\infty)^2 + T_\infty^2) (T_p + 2T_\infty) T_p \\ &\approx p - \frac{k_a A_w}{l_a} (T_p - T_w) - h A_p T_p \\ &\quad - 4\sigma \epsilon_p A_p T_{mp}^3 T_w \\ &= p - \frac{k_a A_w}{l_a} (T_p - T_w) - (h + h_{rp}) A_p T_p \quad (4) \end{aligned}$$

where  $h_{rp} = 4\sigma \epsilon_p T_{mp}^3$  and  $T_{mp}$  is the mean of  $T_p + T_\infty$  and  $T_\infty$ .

Let

$$\begin{aligned} R_a &= \frac{l_a}{k_a A_w} \\ C_w &= m_w c_w \end{aligned} \quad (5)$$

$$R_w = \frac{1}{(h + h_{rw})A_w}$$

$$R_p = \frac{1}{(h + h_{rp})A_p}$$

Equation (3) and (4) can now be expressed as

$$C_w \dot{T}_w(t) = \frac{1}{R_a}(T_p(t) - T_w(t)) - \frac{1}{R_w}T_w(t) \quad (6)$$

$$C_p \dot{T}_p(t) = p(t) - \frac{1}{R_a}(T_p(t) - T_w(t)) - \frac{1}{R_p}T_p(t) \quad (7)$$

At steady state,  $\dot{T}_w(\infty) = \dot{T}_p(\infty) = 0$  and Equations (6) and (7) give

$$T_w(\infty) = \frac{R_w}{R_w + R_a}T_p(\infty)$$

$$T_p(\infty) = \frac{R_p}{R_p + R_a}T_w(\infty) + \frac{R_p R_a}{R_p + R_a}p(\infty)$$

At the beginning of an experiment ( $t = 0$ ), a wafer at ambient temperature ( $T_w(0) = 0$ ) is placed on the bakeplate maintained at a particular steady-state temperature,  $T_p(0)$ . Notice in Figure 3, the bakeplate temperature recovered after the baking process i.e.  $T_p(\infty) = T_p(0)$ .

To simplify our analysis, the following new variables are defined.

$$\theta_w(t) = T_w(t) - T_w(\infty)$$

$$= T_w(t) - \frac{R_w}{R_w + R_a}T_p(\infty)$$

$$\theta_p(t) = T_p(t) - T_p(\infty)$$

$$u(t) = p(t) - p(\infty)$$

Equation (6) and (7) can now be expressed as

$$C_w \dot{\theta}_w(t) = \frac{1}{R_a}(\theta_p(t) - \theta_w(t)) - \frac{1}{R_w}\theta_w(t)$$

$$C_p \dot{\theta}_p(t) = u(t) - \frac{1}{R_a}(\theta_p(t) - \theta_w(t)) - \frac{1}{R_p}\theta_p(t)$$

Laplace transformation gives

$$\Theta_w(s) \left( sC_w + \frac{1}{R_a} + \frac{1}{R_w} \right)$$

$$= C_w \theta_w(0) + \frac{1}{R_a} \Theta_p(s) \quad (8)$$

$$\Theta_p(s) \left( sC_p + \frac{1}{R_a} + \frac{1}{R_p} \right) =$$

$$U(s) + C_p \theta_p(0) + \frac{1}{R_a} \Theta_w(s) \quad (9)$$

where

$$\theta_w(0) = T_w(0) - \frac{R_w}{R_w + R_a}T_p(\infty)$$

$$\theta_p(0) = 0$$

Eliminating  $\Theta_w(s)$  in Equations (8) and (9) gives

$$\Theta_p(s) = \Theta'_p(s) + D(s) \quad (10)$$

$$= \frac{B(s)}{A(s)}U(s) + D(s) \quad (11)$$

where  $\Theta'_p(s) = \frac{B(s)}{A(s)}U(s)$  is the bakeplate temperature due to the heat supplied by the heater to the bakeplate and  $D(s)$  denotes the temperature change resulting from heat removed from the bakeplate by dropping the cold wafer on the bakeplate. Further

$$\frac{B(s)}{A(s)} = \frac{\frac{1}{C_p} \left( s + \frac{1}{C_w} \left( \frac{1}{R_a} + \frac{1}{R_w} \right) \right)}{s^2 + sE_1 + E_0}$$

$$D(s) = \frac{\frac{1}{C_p R_a}}{s^2 + sE_1 + E_0} \theta_w(0)$$

where

$$E_1 = \left( \frac{1}{C_w} + \frac{1}{C_p} \right) \left( \frac{1}{R_a} + \frac{1}{R_w} \right)$$

$$E_0 = \frac{1}{C_w C_p} \left( \frac{1}{R_a R_w} + \frac{1}{R_a R_p} + \frac{1}{R_p R_w} \right)$$

The block diagram is shown in Figure 4 where  $G_c(s)$  is the temperature controller.

### 3. TEMPERATURE CONTROL

We now develop the control strategy using linear programming to compensate for the load disturbance induced by the placement of the wafer at room temperature on the hot plate. This approach has been discussed elsewhere in the literature (Ho et al., 2000) and only the equations necessary for the computation in this paper are given.

From Equation (10), we note that the effect of the disturbance,  $D(s)$ , on the temperature of the bakeplate,  $\Theta_p(s)$ , can be eliminated if,

$$\Theta'_p(s) + D(s) = 0$$

In other words, the input energy flux resulting from the heater balances the output energy flux due to the cold substrate. This can be accomplished without feedback control by adjusting the heater power using Equation (11)

$$\frac{B(s)}{A(s)}U(s) + D(s) = 0$$

$$U(s) = U_{ff}(s) = -\frac{A(s)}{B(s)}D(s) \quad (12)$$

In practice, because of bounds placed on the achievable input power from the heater, the control signal is subjected to saturation within lower and upper bounds, for example  $u \in [U^{min}, U^{max}]$ . In this case,

$$\Theta'_p(s) + D(s) \neq 0$$

if the required input power is outside the achievable bounds. A simple implementation strategy would be to calculate the perfect control move as given by Equation (12), and then truncate at the boundaries. However, for our application we consider the optimal solution. To implement a practical solution, we discretize the problem in sampled data format, denoting the sampling indices as  $k \in \{0, 1, \dots\}$ , and express the goal to minimize the maximum absolute error (Edgar and Himmelblau, 1989) as represented by the objective function

$$\min_{u(k) \in [U^{min}, U^{max}]} \max_{k \in \{0, 1, \dots, N\}} |\theta'_p(k) + d(k)| \quad (13)$$

with the condition that the input is within prespecified limits and the desired output temperature remains equal to its initial steady-state condition over a time horizon  $N$ .

This optimization problem can be solved computationally by the use of the models in Equations (10) and (11). Discretizing for computer implementation gives

$$\begin{aligned} \theta_p(k) &= \theta'_p(k) + d(k) \\ &= \frac{B(q^{-1})}{A(q^{-1})} u(k) + d(k) \end{aligned} \quad (14)$$

where  $q^{-1}$  is the backward shift operator ( $q^{-1}y(k) = y(k-1)$ ) and

$$\begin{aligned} A(q^{-1}) &= 1 + a_1q^{-1} + \dots + a_nq^{-n} \\ B(q^{-1}) &= b_0 + b_1q^{-1} + \dots + b_nq^{-n} \end{aligned}$$

The order of the polynomials,  $n$ , have been assigned to be equal.

This discrete representation can also be expressed in a convolution model at sample time  $k$ ,

$$\theta'_p(k) = \sum_{j=0}^k c_j q^{-j} u(k)$$

where the coefficients are given by

$$c_j = b_j - \sum_{\ell=1}^n a_\ell c_{j-\ell}$$

Over a finite interval,  $N$ , the input and output signals can be represented as finite-dimensional vectors. The solution between the input and output vectors over the interval  $N$  can be expressed as a Toeplitz matrix,

$$\Theta'_p = \Psi \mathbf{U}$$

where

$$\Theta'_p = \begin{bmatrix} \theta'_p(0) \\ \theta'_p(1) \\ \vdots \\ \theta'_p(N) \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N) \end{bmatrix}$$

$$\Psi = \begin{bmatrix} c_0 & 0 & \dots & \dots & 0 \\ c_1 & c_0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_N & c_{N-1} & \dots & c_1 & c_0 \end{bmatrix}$$

The optimization problem of Equation (13) is equivalent to the following linear programming problem (Edgar and Himmelblau, 1989):

Minimize

$$[0 \dots 0 \ 1] \begin{bmatrix} \mathbf{U} \\ e \end{bmatrix} \quad \text{objective function}$$

subject to

$$\begin{bmatrix} \Psi & -\mathbf{1}_I \\ -\Psi & -\mathbf{1}_I \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ e \end{bmatrix} \leq \begin{bmatrix} -\mathbf{D} \\ \mathbf{D} \end{bmatrix} \quad \text{dynamic model}$$

$$\mathbf{U} \leq U^{max} \quad \text{upper control signal saturation}$$

$$\mathbf{U} \geq U^{min} \quad \text{lower control signal saturation}$$

$$\theta'_p(k) + d(k) = 0 \quad k \in [n_f, \dots, N]$$

disturbance to be eliminated from  $n_f$  to  $N$

where  $\mathbf{1}_I$  is a column vector with all  $I$  entries equal to one,  $e(k) = \theta'_p(k) + d(k)$ , and  $\mathbf{D} = [d(0) \ d(1) \ \dots \ d(N)]'$ . For vectors  $\mathbf{v}$  and  $\mathbf{w}$ ,  $\mathbf{v} \leq \mathbf{w}$  means every entry of  $\mathbf{v}$  is less than or equal to the corresponding entry of  $\mathbf{w}$ . The value of  $n_f$  is chosen such that  $n_f$  is feasible and  $n_{f-1}$  is not. This value is then chosen as the minimum time for eliminating the disturbance.

#### 4. EXPERIMENT

The experiments were conducted at  $T_p(0) = T_p(\infty) = 62^\circ\text{C}$  and  $T_w(0) = 0^\circ\text{C}$ . The thermal capacitance of the wafer,  $C_w = 40.4 \text{ J/K}$ , and wafer loss resistance,  $R_w = 2.3 \text{ K/W}$ , can be computed directly from the handbook data. For the given bakeplate, the proximity pin  $l_{pp} = 165 \text{ } \mu\text{m}$ ,  $C_p = 750 \text{ J/K}$  and  $R_p = 0.59 \text{ K/W}$  were expected to be fixed and hence determined beforehand. For  $l_a = 165 \text{ } \mu\text{m}$ , Equation (5) gave  $R_a = 0.43 \text{ K/W}$ . Substituting all the parameters into Equation (11) gave the transfer functions of hot plate and disturbance below.

$$\Theta_p(s) = \frac{7s + 0.5}{5283s^2 + 384s + 1} U(s) - \frac{866}{5283s^2 + 384s + 1}$$

Discretizing with 1 second sampling interval gives Equation (14) as

$$\begin{aligned} \theta_p(k) &= \frac{0.0013q^{-1} - 0.0012q^{-2}}{1 - 1.9297q^{-1} + 0.9299q^{-2}} u(k) \\ &\quad - 2.44 (e^{-0.0027k} - e^{-0.07k}) \end{aligned}$$

Once the model for  $\theta_p(k)$  was known, the optimal feedforward control signal was computed from linear programming with the constraints  $-120 \text{ W} \leq u(k) \leq 80 \text{ W} \ \forall k \in \{0, 1, \dots, N\}$  or equivalently  $0 \text{ W} \leq p(k)$

$\leq 200$  W since  $p(\infty) = 120$  W. The result is shown in Figure 5. A Proportional-Integral Controller of the form

$$G_c(s) = K_p \left( 1 + \frac{1}{sT_i} \right) = 36 \left( 1 + \frac{1}{67s} \right)$$

was used in the experiment and discretization gave

$$G_c(q^{-1}) = \frac{36 - 35.46q^{-1}}{1 - q^{-1}}$$

All the parameters in discrete form of the system block diagram of Figure 4 were known with  $u_{ff}(k)$  given in Figure 5. Computer simulation was run to give the disturbance as shown in Figure 6. Notice that the minimum temperature is  $-0.1^\circ\text{C}$ . By repeating the simulation for different values of  $l_a$ , the minimum temperature can be plotted against  $l_a$  as shown in Figure 7.

The temperature disturbance with feedback control only without feedforward control is given in Experiment (a) of Figure 3. The temperature drop is about  $-1^\circ\text{C}$ . The temperature disturbance with feedback and feedforward control is given in Experiment (b). The temperature drop is about  $-0.1^\circ\text{C}$ . Notice a  $10 \times$  reduction in the magnitude of the temperature disturbance when feedforward control was used. In contrast the temperature disturbance in Ho et al. (2004) like Experiment (a) was larger than  $-1^\circ\text{C}$  because feedforward control was not implemented making it less suitable for industrial implementation. Wafers with center-to-edge warpages of  $55 \mu\text{m}$  and  $110 \mu\text{m}$  were used for Experiments (c) and (d) respectively. The results are given in Table 1. The minimum bakeplate temperatures were  $-0.2^\circ\text{C}$  and  $-0.38^\circ\text{C}$  respectively and the corresponding nominal air-gaps read from the curve in Figure 7 were  $l_a = 149 \mu\text{m}$  and  $l_a = 118 \mu\text{m}$ . The average warpages were  $\lambda_{wp} = -16 \mu\text{m}$  and  $\lambda_{wp} = -47 \mu\text{m}$  respectively.

Consider Experiment (b) and (d) in Table 1. The nominal air-gap,  $l_a$ , changed by about  $50 \mu\text{m}$  from  $167 \mu\text{m}$  to  $118 \mu\text{m}$  while the temperature drop changed by about  $0.25^\circ\text{C}$  from  $-0.12^\circ\text{C}$  to  $-0.38^\circ\text{C}$ . In contrast, the sensitivity in Ho et al., (2004) for detecting  $l_a$  was lower i.e.  $0.1^\circ\text{C}$  change in temperature for a  $50 \mu\text{m}$  change in  $l_a$ . Hence, in this paper, temperature is more sensitive to change in  $l_a$  which is important for detecting warpage from temperature.

Table 1: Warpage Fault Detection (proximity pin of bakeplate,  $l_{pp} = 165 \mu\text{m}$ )

Run No.	Wafer Warpage	Maximum Temperature Drop in $\theta_p$ ( $^\circ\text{C}$ )	Norminal Air-gap $l_a$ ( $\mu\text{m}$ )	Average Warpage $\lambda_{wp} = l_a - l_{pp}$ ( $\mu\text{m}$ )
(b)	flat	-0.12	167	2
(c)	warped ( $55 \mu\text{m}$ )	-0.2	149	-16
(d)	warped ( $110 \mu\text{m}$ )	-0.38	118	-47

Since a flat wafer was used for Run (b),  $l_a$  is expected to be close to the length of the proximity pin,  $l_{pp}$ , and

the average warpage,  $\lambda_{wp} = l_a - l_{pp} = 2 \mu\text{m}$ . Note that  $\lambda_{wp}$  is the deviation of the nominal airgap from the proximity pin in a warped wafer (see Figure 2) and is a good measure of the extent of warpage. If the deviation is large, then warpage is large. A standard approach to fault detection is to define a threshold based on manufacturing requirements and any violation of the threshold is considered a fault.

In this paper, we have demonstrated that it is possible to detect warpage fault from the minimum temperature. We expect accuracy to improve if more sophisticated estimation technique is used. For example, an extension is to consider area under the temperature curve instead of just the minimum point. Yet, a even more elaborate method is to fit the complete temperature trajectory of the bakeplate with the model in the least square sense. These more elaborate techniques can be investigated in future once the concept of relating temperature to warpage is established.

## 5. CONCLUSION

The lithography manufacturing process will continue to be a critical area in semiconductor manufacturing that limits the performance of microelectronics. Enabling advancements by computational, control and signal processing methods are effective in reducing the enormous costs and complexities associated with the lithography sequence. In this paper, advanced process control reduced temperature disturbance to about a tenth of a degree Celcius. We demonstrate that warpage can still be detected and more importantly, there is no loss in sensitivity in the detection algorithm even though the temperature disturbance that provided the signal for detecting warpage is small.

## 6. REFERENCE

- Edgar T.F. and D.M. Himmelblau (1989). *Optimization of Chemical Processes*, New York, McGraw-Hill.
- Fauque, J.A. and R.D. Linder (1998). Extended range and ultra-precision non-contact dimensional gauge, U.S. Patent 5, 789,661.
- Ho, W.K., A. Tay and C. Schaper (2000). "Optimal Predictive Control with Constraints for the Processing of Semiconductor Wafers on Large Thermal-Mass Heating Plates", IEEE Transaction on Semiconductor Manufacturing, Vol. 13, No. 1, pp. 88-96.
- Ho, W.K., A. Tay, Y. Zhou and K. Yang (2004). "In-Situ Fault Detection of Wafer Warpage in Microlithography", IEEE Transactions on Semiconductor Manufacturing, Vol. 17, No. 3, pp. 402-407.
- Ozsisik, M.N. (1985). *Heat Transfer-A Basic Approach*, McGraw-Hill.
- Poduje, N.S. and W.A. Balies (1988). "Wafer Geometry Characterization: An overview I", Microelectronic Manufacturing and Testing, Vol. 11, No. 6, pp. 29-32.

Quirk, M. and J. Serda (2001). "Semiconductor Manufacturing Technology", Prentice Hall.

Raznjevic, K. (1976). Handbook of Thermodynamic Tables and Charts, Hemisphere Publishing Corporation.

Schaper, C., K. El-Awady, A. Tay and T. Kailath (1999). "Control Systems for the Nanolithography Process", 38th IEEE Conference on Decision and Control, USA, pp. 4173-4178, Dec 7-10.

Wei, S., S. Wu, I. Kao and F. P. Chiang (1998). "Measurement of Wafer Surface using Shadow Moire Technique with Talbot Effect", Journal of Electronic Packaging, Vol. 120, No. 2, pp. 166-170.

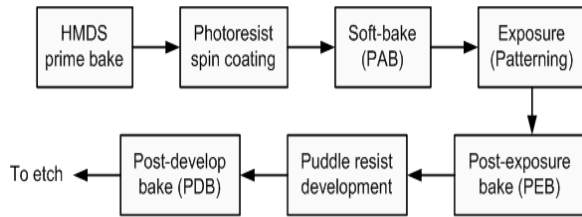


Fig. 1. Steps in the Lithography Process

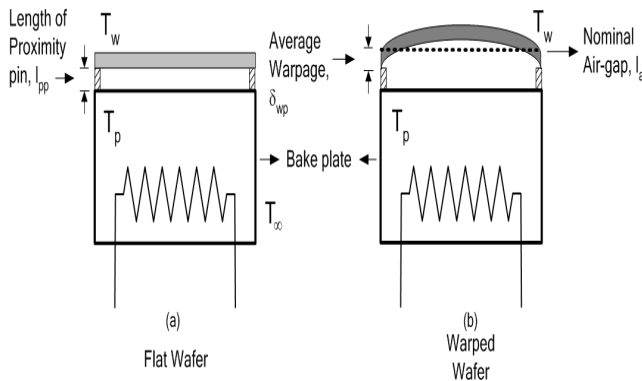


Fig. 2. Baking of a Flat and Warped Wafer

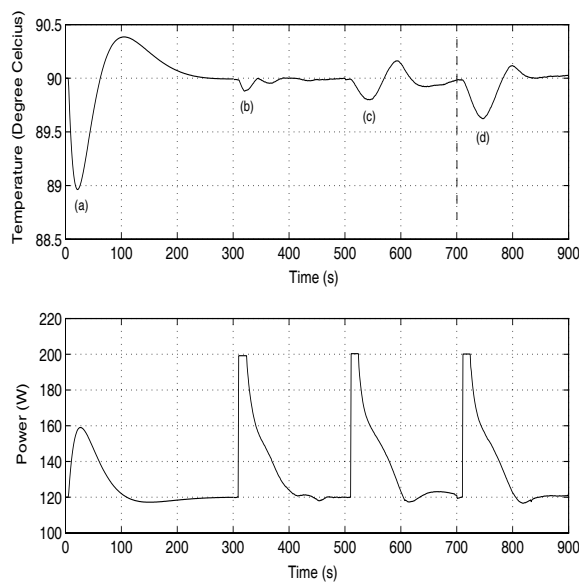


Fig. 3. Experimental Results

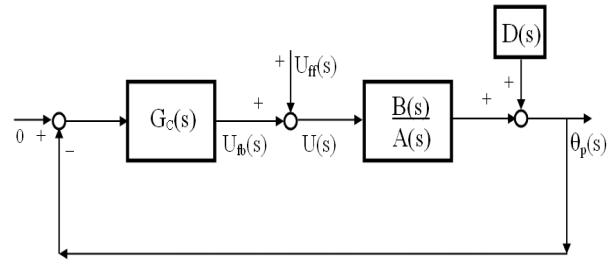


Fig. 4. Block Diagram of the Baking Process

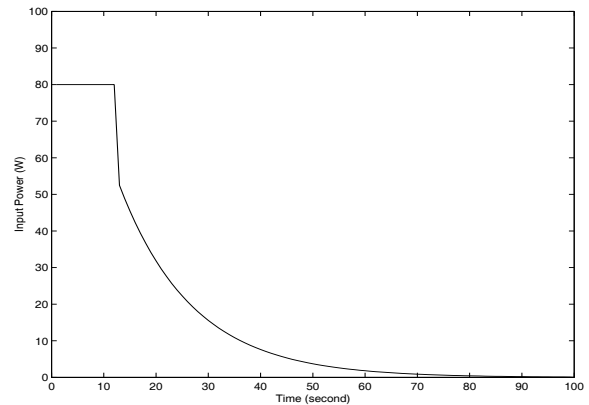


Fig. 5. Feedforward Control Signal

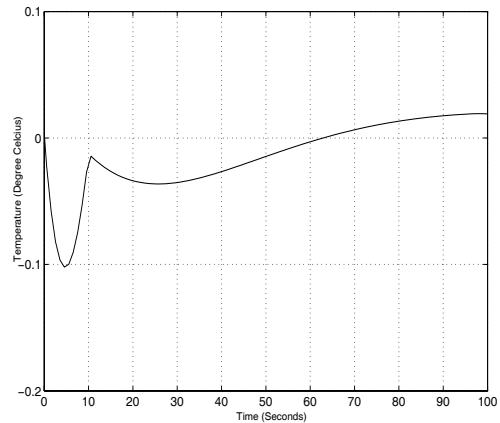


Fig. 6. Temperature Disturbance Simulation

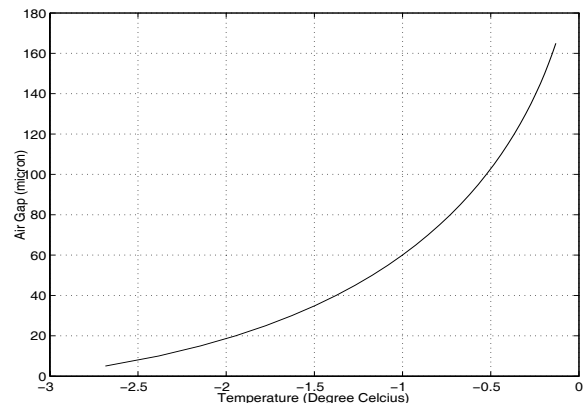


Fig. 7. Norminal Air-Gap,  $l_a$ , versus Temperature Drop