Thermal Uniformity of 12-in Silicon Wafer in Linearly Ramped-Temperature Transient Rapid Thermal Processing

Senpuu Lin and Hsin-Sen Chu

Abstract—**This paper presents a systematic method for estimating the dynamic incident-heat-flux profiles required to achieve thermal uniformity in 12-in silicon wafers during linearly ramped-temperature transient rapid thermal processing using the inverse heat-transfer method. A two-dimensional thermal model and temperature-dependent silicon wafer thermal properties are adopted in this study. The results show that thermal nonuniformities on the wafer surfaces occur during ramped increases in direct proportion to the ramp-up rate. The maximum temperature differences in the present study are 0.835 °C, 1.174 °C, and 1.516 °C, respectively, for linear 100 °C/s, 200 °C/s, and 300 °C/s ramp-up rates. Although a linear ramp-up rate of 300 °C/s was used and measurement errors did reach 3.864 °C, the surface temperature was maintained within 1.6 °C of the center of the wafer surface when the incident-heat-flux profiles were dynamically controlled according to the inverse-method approach. These thermal nonuniforimities could be acceptable in rapid thermal processing systems.**

Index Terms—**Inverse heat-transfer method, linear ramp-up rate, rapid thermal processing, thermal uniformity, 12-in silicon wafer.**

I. INTRODUCTION

AS DEVICE dimensions shrinks to the submicrometer range, reduction of the thermal budget during microelectronic processing is becoming a critical issue. Single-wafer rapid thermal processing (RTP) has become an alternative to conventional furnace-based batch processing in many processes [1]. To obtain uniform processing across wafers and to prevent creation of slip defects due to thermal stresses, temperatures must be nearly uniform across wafers throughout the process cycle [2].

It is known that incident-heat-flux profiles (energy distributions) in RTP systems must be nonuniform across wafers to ensure temperature uniformity at all times because of heat losses at wafer edges. Hill and Jones [3] investigated the temperature uniformity of a 150-mm (6-in) wafer in which the intensity was linearly enhanced to a maximum of 8% over the last 15 mm of the wafer. Kakoschke *et al.* [4] presented a wafer-heating theory for estimating the edge-heating compensation required vertically and laterally to ensure temperature

uniformity during process ramp-up and steady-state phases. Gyurcsik *et al.* [5] used a two-step procedure to solve an inverse optimal-lamp-contour problem that involved achieving wafer temperature uniformity in the steady state. Sorrel *et al.* [6] applied power-law (first-, second-, and seventh-degree) irradiation profiles to study the increases required in perimeter radiation to maintain a wafer at an approximately uniform temperature. Riley and Gyurcsik [7] presented a wafer-edge nodal analysis to determine the amount of lateral heating needed to counteract edge cooling. Cho *et al.* [8] presented a method for optimizing incident-heat-flux profiles by studying wafer heat-loss profiles. Following the work of Riley and Gyurcsik [7], Perkins *et al.* [9] used wafer-edge node analysis to determine the idealized intensity profiles required for maintaining thermal uniformity both during transient and steady states.

The works mentioned above describe quantifying incident heat fluxes over wafers to achieve the necessary thermal uniformity requirements during RTP. Some of these approaches have been largely trial-and-error, which can be quite expensive and time-consuming. Some approaches, such as those of Cho *et al.* [8] and Perkins *et al.* [9], are systematic design methods for knowing whether a design satisfying given specifications exists and whether a given approach has an optimal design that satisfies given specifications. There may be more efficient systematic methods for determining the incident heat-flux profiles required over a wafer to ensure thermal uniformity. To the best of our knowledge, no inverse heat-transfer methods for use in determining how to achieve thermal uniformity during linear ramp-up transient RTP have been published to date. The inverse heat-transfer method deals with determining crucial parameters in analyses such as those for internal energy sources, surface heat fluxes, thermal properties, etc., and has been widely applied to many design and manufacturing problems [10]–[13]. Recently, we [14], [15] applied a one-dimensional thermal model to study the thermal uniformity of 12-in silicon wafers subjected to a uniformly distributed heat flux during RTP using the inverse-source method. It was discovered that the resulting maximum temperature differences were only 0.326 °C during RTP. However, many researchers have adopted the two-dimensional (2-D) (radial and axial) thermal model to study the RTP thermal uniformity problem. The inverse boundary heat-flux method [13] is practical for use in determining how to achieve thermal uniformity with 2-D RTP systems in which the surface heat-flux strengths required to achieve temperature uniformity are unknown.

In the present paper, a finite-difference-method formulation of a 2-D (radial and axial) thermal model in which both
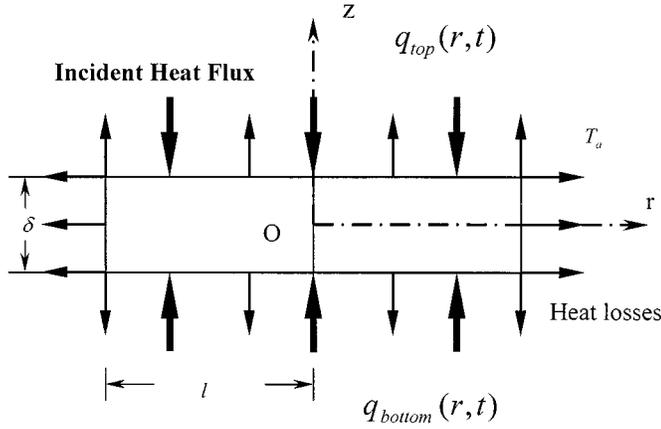
Fig. 1. Schematic representation of energy flux in a silicon wafer under two-sided incident radiation and radiant loss emitted from all surfaces.

surfaces of a 300-mm-diameter 0.775-mm-thick silicon wafer are heated is studied for application to RTP systems. The temperature-dependent thermal properties of the silicon wafer are considered. Then, the incident-heat-flux profiles over the wafer surface calculated using the inverse boundary heat-flux method for tracking uniform temperature trajectories during both ramp-up and steady-state RTP phases are examined. We also discuss measurement-error effects on the wafer thermal uniformity.

II. THERMAL MODEL

Consider the thin circular silicon wafer shown in Fig. 1. Let l and δ be the radius and thickness, respectively. T_0 is the initial uniform wafer temperature, and the ambient temperature is T_a. Symmetric heating on both sides of the wafer is adopted. The total incident heat fluxes on the top and the bottom surfaces of the wafer are denoted by q_{top} and q_{bottom}, respectively. We may assume without loss of generality that the incident heat fluxes on both sides during processing are equal, i.e., $q_{\text{top}}(r,t) = q_{\text{bottom}}(r,t) = q(r,t)$. Radiant heat losses occur at all surfaces. The process is considered to operate in a vacuum so heat transfer due to convection can be ignored.

The governing equation for an axially symmetric cylindrical coordinate system with its origin at the wafer center is

$$\rho c(T)\frac{\partial T}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[k(T)r\frac{\partial T}{\partial r}\right] + \frac{\partial}{\partial z}\left[k(T)\frac{\partial T}{\partial z}\right] \quad (1)$$

where wafer temperature T is a function of the radius r, thickness z, and time t; ρ, $k(T)$, and $c(T)$ are the wafer density, thermal conductivity, and specific heat capacity, respectively. Here, because of the large temperature variations that occur during processing, the temperature dependence of wafer thermal conductivity as well as specific heat capacity must be considered as follows [1]:

$$k(T) = 802.99T^{-1.12} \quad (\text{W}\cdot\text{cm}^{-1}\cdot\text{K}^{-1}) \quad 300\text{–}1683\text{ K} \quad (2a)$$
$$c(T) = 0.641 + 2.473 \times 10^{-4}T \quad (\text{J}\cdot\text{g}^{-1}\cdot\text{K}^{-1}) \quad > 300\text{ K}, \quad (2b)$$

while the wafer density is assumed to be constant and equal to 2.33 g·cm^{-3}. Since the silicon wafer is considered to be homogeneous in the present study, the dependence of $k(T)$ on spatial position is introduced only implicitly by the spatial depen-

dence of the temperature. Because $k(T)$ is weakly dependent on temperature, instantaneous spatial temperature variations across wafers at given times are expected to be small enough (≤ 200 K) so that spatial variations in thermal conductivity may be ignored [4]. Thus, (1) may be reduced to

$$\rho c(T)\frac{\partial T}{\partial t} = k(T)\left[\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial z^2}\right]. \quad (3)$$

The initial and boundary conditions for the system described above are

$$T(r,z,t) = T_0, \qquad \text{at } t = 0 \quad (4)$$

$$\frac{\partial T}{\partial r} = 0, \qquad \text{at } r = 0 \quad (5)$$

$$-k(T)\frac{\partial T}{\partial r} = \varepsilon_{\text{edge}}\sigma_s(T^4 - T_a^4), \qquad \text{at } r = l \quad (6)$$

$$\frac{\partial T}{\partial z} = 0, \qquad \text{at } z = 0 \quad (7)$$

$$-k(T)\frac{\partial T}{\partial z} = -\alpha_{\text{surface}}q(r,t) + \varepsilon_{\text{surface}}\sigma_s(T^4 - T_a^4),$$
$$\text{at } z = \frac{\delta}{2} \quad (8)$$

where α_{surface} is the wafer surface absorptivity, $\varepsilon_{\text{surface}}$ is the wafer surface emissivity, $\varepsilon_{\text{edge}}$ is the emissivity for the radiant heat losses at the wafer edges, and $\sigma_s = 5.672 \times 10^{-12}$ W·cm^{-2} K^{-4} is the Stefan–Boltzmann constant. Note that absorptivity and emissivity may depend on wafer temperature, position, and radiant spectral wavelength [16], [17]. In this work, we may assume without loss of generality that the absorptivity of all wafer surfaces equals the emissivity of those surfaces. For simplicity, the emissivity of all surfaces is assumed to be the same and simply temperature-dependent as described by Virzi [18]:

$$\alpha_{\text{surface}} = \varepsilon_{\text{surface}} = \varepsilon_{\text{edge}} = \varepsilon(T)$$
$$= 0.2662 + 1.8591T^{-0.1996}e^{-(1.0359 \times 10^{25}/T^{8.8328})}. \quad (9)$$

Thus, (6) and (8) may be rewritten, respectively, as

$$-k(T)\frac{\partial T}{\partial r} = \varepsilon(T)\sigma_s(T^4 - T_a^4), \qquad \text{at } r = l \quad (10)$$

and

$$-k(T)\frac{\partial T}{\partial z} = -\varepsilon(T)q(r,t) + \varepsilon(T)\sigma_s(T^4 - T_a^4),$$
$$\text{at } z = \frac{\delta}{2}. \quad (11)$$

Defining wafer thermal diffusivity as $\kappa(T) = (k(T)/\rho c(T))$ and introducing the dimensionless temperature θ, radial position R, axial position Z, time τ, incident heat flux Q, thick-to-radius ratio η, thermal conductivity $K(\theta)$, specific heat capacity $C(\theta)$, thermal diffusivity $D(\theta)$, and emissivity $\varepsilon(\theta)$ as

$$\theta = \frac{T}{T_a}, \quad R = \frac{r}{l}, \quad Z = \frac{z}{l}, \quad \tau = \frac{k_a t}{\rho c_a l^2},$$
$$Q(R,\tau) = \frac{l}{k_a T_a}q(r,t), \quad \eta = \frac{\delta}{l}, \quad K(\theta) = \frac{k}{k_a},$$
$$C(\theta) = \frac{c}{c_a}, \quad D(\theta) = \frac{K(\theta)}{C(\theta)}, \quad \varepsilon(\theta) = 0.2662$$
$$+ 1.8591(T_a\theta)^{-0.1996}e^{-(1.0359 \times 10^{25}/(T_a\theta)^{8.8328})}$$
$$(12)$$

where $k_a = k(T_a)$ and $c_a = c(T_a)$ are the thermal conductivity and specific heat capacity of the wafer at the ambient temperature T_a, respectively. The energy equation (3) then becomes

$$\frac{\partial \theta}{\partial \tau} = D(\theta)\left(\frac{\partial^2 \theta}{\partial R^2} + \frac{1}{R}\frac{\partial \theta}{\partial R} + \frac{\partial^2 \theta}{\partial Z^2}\right),$$
$$0 < R < 1, 0 < Z < \frac{\eta}{2}. \quad (13)$$

The initial condition and boundary conditions become, respectively,

$$\theta(R, Z, \tau) = \theta_0, \quad \text{at } \tau = 0 \quad (14)$$

and

$$\frac{\partial \theta}{\partial R} = 0, \quad \text{at } R = 0 \quad (15)$$

$$\frac{\partial \theta}{\partial R} = \frac{\varepsilon(\theta)}{K(\theta)}A(1 - \theta^4), \quad \text{at } R = 1 \quad (16)$$

$$\frac{\partial \theta}{\partial Z} = 0, \quad \text{at } Z = 0 \quad (17)$$

$$\frac{\partial \theta}{\partial Z} = \frac{\varepsilon(\theta)}{K(\theta)}[Q(R, \tau) + A(1 - \theta^4)]. \quad \text{at } Z = \frac{\eta}{2}. \quad (18)$$

Here the dimensionless initial temperature θ_0 is T_0/T_a and the dimensionless constant A is $\sigma_s l T_a^3 / k_a$.

The numerical solution techniques used here are from the finite-difference method. A central-difference representation of the space derivative and an implicit backward-difference representation of the time derivative are adopted. We can approximate the energy equation and the initial condition as well as the boundary conditions (13)–(18) using

$$\theta(R, Z, \tau) = \theta((i-1)\Delta R, (j-1)\Delta Z, n\Delta \tau) = \theta_{i,j}^n$$

with the radial coordinate increment $\Delta R = (1/i\max - 1)$, the axial coordinate increment $\Delta Z = (1/j\max - 1)$, and the temporal coordinate increment $\Delta \tau$. After the nonlinear radiant fourth-power terms in (16) and (18) have been simulated using a linear scheme and the SOR-by-lines method [19] has been adopted, the unknowns in the subgroups to be modified simultaneously are set up such that the matrix of coefficients will be tridiagonal in form permitting use of the Thomas algorithm as follows:

$$b_{i,j}^n \theta_{i-1,j}^n + d_{i,j}^n \theta_{i,j}^n + a_{i,j}^n \theta_{i+1,j}^n = c_{i,j}^n. \quad (19)$$

The superscript n is denoted as the index of the temporal grid. The subscripts i and j are denoted, respectively, as the indices of the radial and axial grids. Given the incident heat flux $Q_{i,j\max}^n$,

we can obtain the wafer temperature distributions $\theta_{i,j}^n$. The relative criterion for each iteration k is considered to be

$$\frac{|\theta_{i,j}^{k+1} - \theta_{i,j}^k|}{\max |\theta_{i,j}^{k+1}|} < 10^{-7}$$

and when

$$\frac{|\theta_{i,j}^{n+1} - \theta_{i,j}^n|}{\max |\theta_{i,j}^{n+1}|} < 10^{-5}$$

it may be assumed that processing has reached the steady state.

III. INVERSE HEAT-TRANSFER METHOD

The inverse heat transfer method [13] is practical for use in determining how to achieve thermal uniformity with RTP systems in which the magnitudes of incident heat fluxes required to achieve temperature uniformity are unknown.

The finite-difference scheme in the thermal model above at $\tau = \tau^m = m\Delta \tau$ for each j is used to construct the following matrix equation:

$$[F_j^m]\{\theta_j^m\} = \{\theta_j^{m-1}\} + \{S_j^m\} + [V_j^m]\{\varphi_j^m\} \quad (20)$$

where $[F_j^m]$ is defined in (21), shown at the bottom of the page, and

$$[V_j^m] = \begin{bmatrix} v_{1,j}^m & 0 & \cdots & 0 & 0 \\ 0 & v_{2,j}^m & \cdots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \\ 0 & \cdots & 0 & v_{i\max-1,j}^m & 0 \\ 0 & \cdots & 0 & 0 & v_{i\max,j}^m \end{bmatrix}_{i\max \times i\max} \quad (22)$$

$$\{\theta_j^m\} = \begin{Bmatrix} \theta_{1,j}^m \\ \theta_{2,j}^m \\ \cdots \\ \theta_{i\max-1,j}^m \\ \theta_{i\max,j}^m \end{Bmatrix}_{i\max}, \{\theta_j^{m-1}\} = \begin{Bmatrix} \theta_{1,j}^{m-1} \\ \theta_{2,j}^{m-1} \\ \cdots \\ \theta_{i\max-1,j}^{m-1} \\ \theta_{i\max,j}^{m-1} \end{Bmatrix}_{i\max} \quad (23)$$

$$\{S_j^m\} = \begin{Bmatrix} S_{1,j}^m \\ S_{2,j}^m \\ \cdots \\ S_{i\max-1,j}^m \\ S_{i\max,j}^m \end{Bmatrix}_{i\max}, \{\varphi_j^m\} = \begin{Bmatrix} \varphi_{1,j}^m \\ \varphi_{2,j}^m \\ \cdots \\ \varphi_{i\max-1,j}^m \\ \varphi_{i\max,j}^m \end{Bmatrix}_{i\max}. \quad (24)$$

Here $d_{i,j}^m$, $a_{i,j}^m$, as well as $b_{i,j}^m$ were described in Section II, $S_{i,j}^m$ includes any known variables of the problem, and $v_{i,j}^m$ denotes the coefficient for the unknown variable $\varphi_{i,j}^m$. Note that $\varphi_{i,j}^m = \theta_{i,j+1}^m$ for $j = 1, 2, \cdots, j\max - 1$, while $\varphi_{i,j}^m = Q_{i,j\max}^m$ for $j = j\max$. Once given the temperature distribution of $\theta_{i,j}^m$,

$$[F_j^m] = \begin{bmatrix} d_{1,j}^m & a_{1,j}^m & \cdots & 0 & 0 \\ b_{2,j}^m & d_{2,j}^m & \cdots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \\ 0 & \cdots & b_{i\max-1,j}^m & d_{i\max-1,j}^m & a_{i\max-1,j}^m \\ 0 & \cdots & 0 & b_{i\max,j}^m & d_{i\max,j}^m \end{bmatrix}_{i\max \times i\max} \quad (21)$$

for $j \neq j\max$, the temperature distribution of next axial position $\theta_{i,j+1}^m$ can be obtained from (20). When $j = j\max$, the wafer surface temperature distribution $\theta_{i,j\max}^m$ has been computed from $j = j\max - 1$. The undetermined variables are the incident heat fluxes $Q_{i,j\max}^m$ over the wafer surface. The wafer surface temperature distribution $\{\theta_{j\max}^m\}$ may be rewritten from (20) as follows:

$$\{\theta_{j\max}^m\} = [F_{j\max}^m]^{-1}\{\{\theta_{j\max}^{m-1}\} + \{S_{j\max}^m\}\}$$
$$+ [F_{j\max}^m]^{-1}[V_{j\max}^m]\{\varphi_{j\max}^m\}$$
$$= [M_{j\max}^m]\{\{\theta_{j\max}^{m-1}\} + \{S_{j\max}^m\}\}$$
$$+ [N_{j\max}^m]\{\varphi_{j\max}^m\} \tag{25}$$

where

$$[M_{j\max}^m] = [F_{j\max}^m]^{-1} \quad \text{and} \quad [N_{j\max}^m] = [F_{j\max}^m]^{-1}[V_{j\max}^m].$$

The vector $\{\theta_{j\max}^{m-1}\}$ contains $i\max$ values of the initial distribution or the temperature distribution at the wafer surface for the preceding time step. The $\{S_{j\max}^m\}$ vector includes any known variables of the problem. The vector $\{\varphi_{j\max}^m\}$ consists of the unknown incident heat fluxes $\varphi_{i,j\max}^m$ over the wafer surface for $i = 1, 2, \cdots, i\max$ (i.e., $\{\varphi_{j\max}^m\} = \Sigma_{i=1}^{i\max}\{u_i\}\varphi_{i,j\max}^m$). $\{u_i\}$ is the unit column vector with a unit at the ith component. As well, i is the grid number of the location of the estimated surface heat flux function $\varphi_{i,j\max}$.

For the next time step $m + 1$, we arrive at

$$\{\theta_{j\max}^{m+1}\} = [M_{j\max}^{m+1}]\{\{\theta_{j\max}^m\}$$
$$+ \{S_{j\max}^{m+1}\}\} + [N_{j\max}^{m+1}]\{\varphi_{j\max}^{m+1}\}$$
$$= [M_{j\max}^{m+1}][M_{j\max}^m]\{\{\theta_{j\max}^{m-1}\}$$
$$+ \{S_{j\max}^m\}\} + [M_{j\max}^{m+1}][N_{j\max}^m]\{\varphi_{j\max}^m\}$$
$$+ [M_{j\max}^{m+1}]\{S_{j\max}^{m+1}\}\} + [N_{j\max}^{m+1}]\{\varphi_{j\max}^{m+1}\}. \tag{26}$$

In the same way, the temperature distribution at successive r future times $\tau = \tau^{m+r-1}$ can be represented as follows:

$$\{\theta_{j\max}^{m+r-1}\} = [M_{j\max}^{m+r-1}]\{\{\theta_{j\max}^{m+r-2}\}$$
$$+ \{S_{j\max}^{m+r-1}\}\} + [N_{j\max}^{m+r-1}]\{\varphi_{j\max}^{m+r-1}\}$$
$$= [M_{j\max}^{m+r-1}][M_{j\max}^{m+r-2}]$$
$$\cdots [M_{j\max}^{m+1}][M_{j\max}^m]\{\{\theta_{j\max}^{m-1}\}$$
$$+ \{S_{j\max}^m\}\} + [M_{j\max}^{m+r-1}][M_{j\max}^{m+r-2}]$$
$$\cdots [M_{j\max}^{m+1}][N_{j\max}^m]\{\varphi_{j\max}^m\}$$
$$+ [M_{j\max}^{m+r-1}][M_{j\max}^{m+r-2}]\cdots[M_{j\max}^{m+1}]\{S_{j\max}^{m+1}\}$$
$$+ [M_{j\max}^{m+r-1}][M_{j\max}^{m+r-2}]$$

$$\cdots [M_{j\max}^{m+2}][N_{j\max}^{m+1}]\{\varphi_{j\max}^{m+1}\}$$
$$+ \cdots + [M_{j\max}^{m+r-1}]\{S_{j\max}^{m+r-1}\}$$
$$+ [N_{j\max}^{m+r-1}]\{\varphi_{j\max}^{m+r-1}\}. \tag{27}$$

A sequential in-time procedure is adopted to estimate the unknown incident heat-flux parameters. The time domain is divided into analysis intervals each of length $\tau^{m-1} \leq \tau \leq \tau^{m+r-1}$ [10]. The parameters $\varphi_{i,j\max}^m$ for $i = 1, 2, \cdots, i\max$ are determined simultaneously for each analysis interval. A temporary assumption that the incident heat flux is constant over r future time steps is used in the inverse algorithms

$$\varphi_{i,j\max}^{m+1} = \varphi_{i,j\max}^{m+2} = \cdots = \varphi_{i,j\max}^{m+r-1} = \varphi_{i,j\max}^m,$$
$$\text{for } i = 1, 2, \cdots, i\max. \tag{28}$$

Then, the surface temperatures at each p-radial grid ($p = 1, 2, \cdots, i\max$) for each analysis interval can be expressed as follows:

$$\theta_{p,j\max}^{m+k} = h_{p,j\max}^{m+k,0} + \sum_{i=1}^{i\max} E_{p,i,j\max}^{m+k}\varphi_{i,j\max}^m,$$
$$k = 0, 1, 2, \cdots, r-1 \tag{29}$$

where

$$E_{p,i,j\max}^{m+k} = \sum_{l=0}^{k} h_{p,i,j\max}^{m+k,m+l} \tag{30}$$

where $h_{p,i,j\max}^{m+k,m+l}$ is defined in (31), shown at the bottom of the page, and $h_{p,j\max}^{m+k,0}$ is defined in (32), shown at the bottom of the next page, and $[u_p]$ denotes a unit row vector (i.e., a unit at the p-component). The subscripts p and i are the radial grid number of the temperature distribution location and the unknown incident heat flux location, respectively. $j\max$ denotes the axial grid number of the wafer surface.

When $\tau = \tau^m$, the estimated parameter vectors $\{\varphi_{j\max}^1\}$, $\{\varphi_{j\max}^2\}$, $\cdots$, and $\{\varphi_{j\max}^{m-1}\}$ have been evaluated and the task is now to determine the unknown incident heat flux vector $\{\varphi_{j\max}^m\}$. We can construct the following matrix equation:

$$\{\vartheta\}_{r \cdot i\max \times 1} = \{\Phi\}_{r \cdot i\max \times i\max}\{\varphi_{j\max}^m\}_{i\max \times 1}. \tag{33}$$

After the known surface temperature distributions have been substituted into vector ϑ, the components of vector $\{\varphi_{j\max}^m\}$ can be found using the linear least-squares-error method. The result is

$$\{\varphi_{j\max}^m\} = (\Phi^T\Phi)^{-1}\Phi^T\vartheta. \tag{34}$$

$$h_{p,i,j\max}^{m+k,m+l} = \begin{bmatrix} [u_p][N_{j\max}^{m+k}]\{u_i\}, & l = k, & k = 0, 1, 2, \cdots, r-1 \\ [u_p][M_{j\max}^{m+k}][N_{j\max}^{m+k-1}]\{u_i\}, & l = k-1, & k = 1, 2, \cdots, r-1 \\ [u_p][M_{j\max}^{m+k}][M_{j\max}^{m+k-1}][N_{j\max}^{m+k-2}]\{u_i\}, & l = k-2, & k = 2, \cdots, r-2, r-1 \\ \vdots & & \\ [u_p][M_{j\max}^{m+k}][M_{j\max}^{m+k-1}]\cdots[M_{j\max}^{m+2}][N_{j\max}^{m+1}]\{u_i\}, & l = 1, & k = r-2, r-1 \\ [u_p][M_{j\max}^{m+k}][M_{j\max}^{m+k-1}]\cdots[M_{j\max}^{m+1}][N_{j\max}^m]\{u_i\}, & l = 0, & k = r-1, \end{bmatrix} \tag{31}$$
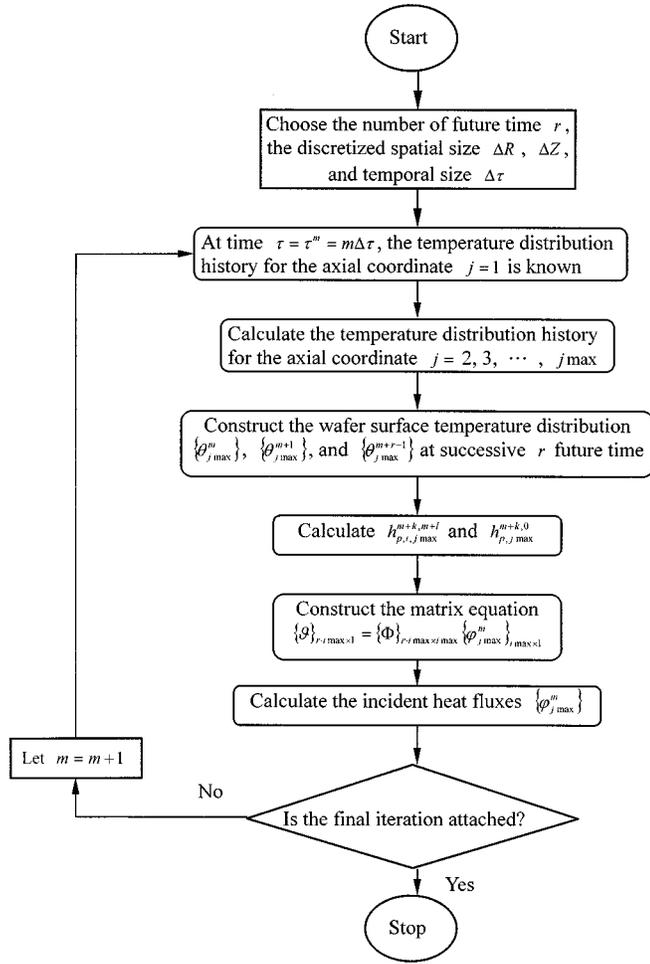
Fig. 2. Flow chart for estimating the dynamic incident-heat-flux profiles required to achieve thermal uniformity using the inverse heat-transfer method.

This equation provides a sequential algorithm flow chart as in Fig. 2, which can be used to determine the unknown incident heat fluxes over the wafer surface by increasing the value of m by one for each time step. Thereafter, the incident heat fluxes can be obtained iteratively along the temporal axis. Thus, the present procedures can estimate the dynamic incident heat-flux profiles on the silicon wafer surface required to achieve thermal uniformity in linearly ramped temperature-transient rapid thermal processing.

IV. RESULTS AND DISCUSSION

The numerical solution techniques described above were used to compute the temperature distributions and the surface incident-heat-flux profiles on a typical 300-mm-diameter 0.775-mm-thick silicon wafer. Numerical simulations ramped linearly from an initial uniform temperature of 27 °C (300 K) to a steady state of 1097 °C (1370 K) at an ambient temperature of $T_a = 27$ °C (300 K) were performed to examine the wafer thermal nonuniformity during RTP as a function of the ramp rate. Since the wafer is so thin that the thermal nonuniformity may be insignificant in axial (z) direction, we concentrated on the wafer surface thermal nonuniformity in radial (r) direction. Plots of the wafer center temperature for three linear ramp-up rates, 100, 200, and 300 °C/s are given in Fig. 3. Random measurement errors were added to the desired temperature trajectories in simulations, as described elsewhere [14]:

$$Y_{1,1}^n = \theta_{1,1}^n + \omega\sigma \tag{35}$$

where the subscript 1 is the grid number of the spatial-coordinate at the wafer center, and the superscript n denotes the grid number of the temporal-coordinate. $\theta_{1,1}^n$ is the dimensionless "exact" desired temperature, $Y_{1,1}^n$ is the dimensionless "measured" temperature, σ is the measurement-error standard deviation, and ω is a random number. The value of ω is calculated using the IMSL subroutine DRNNOR and chosen over the range $-2.576 < \omega < 2.576$, which represents the 99% confidence bound for the measurement temperature. In the present study, the respective dimensional measured temperatures $T_{1,1}^n \pm 0.7728$ °C and $T_{1,1}^n \pm 3.864$ °C were simulated for the cases of $\sigma = 0.001$ and $\sigma = 0.005$.

We set

$$Y_{2,1}^n = Y_{3,1}^n = \cdots = Y_{i\max,1}^n = Y_{1,1}^n \tag{36}$$

as the desired uniform temperature during RTP, for the known temperature distribution used in the inverse heat-transfer method described in Section III to evaluate the unknown incident heat-flux profiles over the wafer surface. After that, the radial temperature distribution across the wafer surface could be computed. Finally, the thermal nonuniformities on the wafer surface during both transient and steady states were investigated.

The inverse wafer-center-temperature trajectory results for three linear ramp-up rates with various measurement errors, $\sigma = 0.0$ (means "exact"), $\sigma = 0.001$ and $\sigma = 0.005$, are shown in Fig. 3. The transient times required to reach the higher steady-state 1097 °C (1370 K) from the initial uniform 27 °C (300 K) were approximately 10.7, 5.35, and 3.57 s for the 100 °C/s, 200 °C/s, and 300 °C/s ramped-up rates, respectively. The accuracy of the proposed method is assessed by comparing the estimated results with the desired temperature trajectories. The re-

$$h_{p,j\max}^{m+k,0} = \begin{cases} [u_p][M_{j\max}^m]\{\{\theta_{j\max}^{m-1}\} + \{S_{j\max}^m\}\}, & k = 0 \\ [u_p][M_{j\max}^{m+1}][M_{j\max}^m]\{\{\theta_{j\max}^{m-1}\} + \{S_{j\max}^m\}\} + [u_p][M_{j\max}^{m+1}]\{S_{j\max}^{m+1}\}, & k = 1 \\ \quad\vdots \\ [u_p][M_{j\max}^{m+r-1}][M_{j\max}^{m+r-2}]\cdots[M_{j\max}^{m+1}][M_{j\max}^m]\{\{\theta_{j\max}^{m-1}\} + \{S_{j\max}^m\}\} \\ \quad + [u_p][M_{j\max}^{m+r-1}][M_{j\max}^{m+r-2}]\cdots[M_{j\max}^{m+1}]\{S_{j\max}^{m+1}\} + \cdots \\ \quad + [u_p][M_{j\max}^{m+r-1}][M_{j\max}^{m+r-2}]\{S_{j\max}^{m+r-2}\} + [u_p][M_{j\max}^{m+r-1}]\{S_{j\max}^{m+r-1}\}, & k = r-1 \end{cases} \tag{32}$$
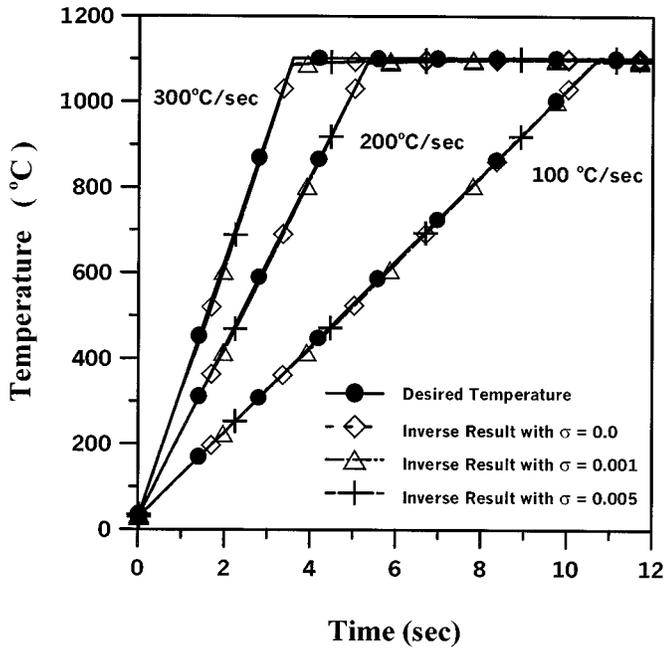
Fig. 3.　Desired uniform temperature trajectories for 100 °C/s, 200 °C/s, and 300 °C/s linear ramp-up rates, and inverse results for measurement errors of $\sigma = 0.0$, $\sigma = 0.001$, and $\sigma = 0.005$.

sults of all test cases have excellent approximations when measurement errors are free ($\sigma = 0.0$). As for the inverse results with measurement errors $\sigma = 0.001$ and $\sigma = 0.005$, it appears that large errors make the estimated results vary from the desired temperature trajectories. For example, the inverse results with measurement errors $\sigma = 0.005$ for linear 300 °C/s ramped-up rates have the maximum deviation from the desired temperature trajectories. The estimation of the inverse analysis is excellent for measurement errors $\sigma = 0.0$. It is noted that the accuracy of the inverse analysis is also good for simulated experimental data containing measurement errors of $\sigma = 0.001$ and $\sigma = 0.005$. Increasing σ from 0.001 to 0.005, the accuracy of the estimated results decreases.

Fig. 4(a)–(c) shows the three-dimensional (3-D) graph of the inverse incident-heat-flux profile results on the wafer surface at a measurement error $\sigma = 0.0$ for uniform temperature tracking of the 100 °C/s, 200 °C/s, and 300 °C/s ramp-up rates, respectively. The axis "Radial Position" shows the distance from the center of the wafer surface in centimeters. The axis "Time" represents the time during this temperature transition. The vertical axis represents the calculated incident-heat-flux profile yielded by the inverse method. The incident heat-flux energy was absorbed by top and bottom surfaces of the wafer, and the heat losses also occurred at all wafer surfaces. Ramping of wafer temperature took place when there was an excess of absorbed energy over heat-loss energy. During the initial transient phase, the wafer temperature increased with the increasing energy absorption, and heat losses also increased as the wafer temperature increased. The initial absorbed energy, required for wafer uniform-temperature tracking, was larger than that during other periods for this temperature transient because, as shown in the energy balance (13), and boundary conditions (16) and (18), the term $\partial\theta/\partial\tau$ on the left-hand side
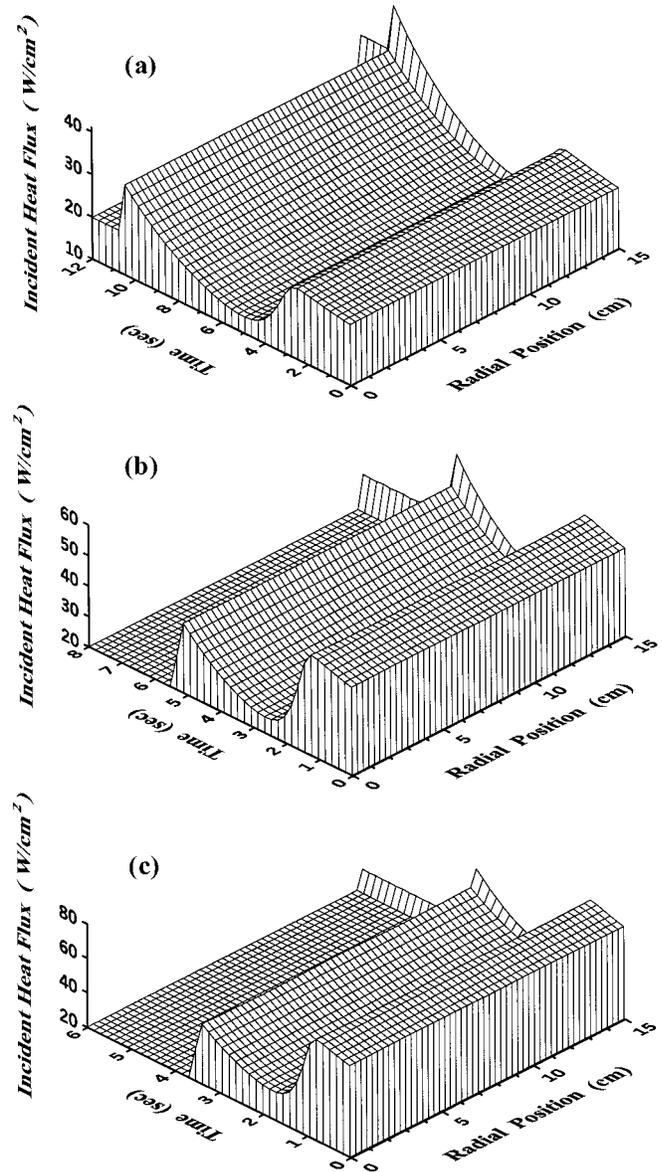

Fig. 4.　Inverse incident-heat-flux profiles results for a measurement error of $\sigma = 0.0$ at linear ramp-up rates of (a) 100 °C/s, (b) 200 °C/s, and (c) 300 °C/s.

of (13), is constant during processing because of the constant temperature ramp-up rate. However, on the right-hand side, the sum of the radial and axial temperature gradient terms, $\partial^2\theta/\partial R^2 + 1/R(\partial\theta/\partial R) + \partial^2/\partial Z^2$, must be equal to the magnitude of $\partial\theta/\partial\tau$. Moreover, according to (16) and (18), the temperature-dependent emissivity $\varepsilon((\theta)$ is 0.3 at the initial lower temperature, and 0.68 from 800 K to 1700 K (see Fig. 2 in [14]). However, the temperature-dependent thermal conductivity $K(\theta)$ decreases with increasing wafer temperature. Thus, the term $\varepsilon(\theta)/K(\theta)$ on the right-hand side of the energy equation at the higher wafer temperature is greater than that at the initial lower wafer temperatures. To balance the energy equation during constant temperature ramp-up processing, the net of incident heat flux and heat losses, $Q(R,\tau)+A(1-\theta^4)$ in (18), is larger in the initial transient phase because of the lower $\varepsilon(\theta)/K(\theta)$. This resulted in larger incident heat fluxes being needed for thermal uniformities in the initial phase. When

the wafer temperature reached about 800 K, the necessary incident heat fluxes dropped due to the larger $\varepsilon(\theta)/K(\theta)$. During higher temperature periods, the heat losses occurring at all surfaces became much greater and greater incident heat fluxes were needed to counteract the heat losses. Thus, the necessary incident heat fluxes increased with the increasing wafer temperature until the wafer reached the higher steady state (the temperature ramp-up rate became zero) and the necessary incident heat fluxes also became steady because the absorbed energy balanced the heat losses. These three different ramp-up rates were all the same at the value of 20 W/cm^2 in this temperature transient during steady-state processing. (after 10.7, 5.35, and 3.57 s for 100 °C/s, 200 °C/s, and 300 °C/s, respectively). Due to the additional heat losses at the wafer edges, more heat compensation was needed at the wafer perimeters during processing. Since the wafer edges were slightly cooler than the center during the initial transient phase, edge-heating compensation was not significant. As ramp-up proceeded, the temperature-dependent thermal conductivity $K(\theta)$ decreased with the increasing wafer temperature and the temperature-dependent emissivity mentioned above; edge heat losses increased as wafer temperature increased from the boundary condition described in (16). Edge-heating compensations were increasingly modulated to meet the requirement of uniform temperature tracking. Thus, the additional amounts of energy directed to the edges to offset edge heat losses became apparent. Finally, as the wafer reached the steady state, the edge heating compensation approached the constant heat-flux scaling factor of 1.26 (25.2 W/cm^2) for a uniform temperature of 1097 °C (1370 K) for these three linear ramp-up rates. These figures show that dynamic individual control of incident heat fluxes is needed for tracking desired uniform-temperature trajectories during RTP. Fig. 5(a) and (b) shows the inverse incident-heat-flux profiles resulting from tracking the desired uniform-temperature trajectory of the 300 °C/s ramp-up rate with measurement errors of $\sigma = 0.001$ and $\sigma = 0.005$, respectively. These results are almost the same as those shown in Fig. 4(c). However, the incident heat flux profiles had to be dynamically modulated according to the measurement-error effects to maintain thermal uniformity during both transient and steady state phases.

Ideally, if it were not for the wafer edges, thermal uniformity could be achieved by applying uniform heat-flux profiles of varying strengths to the top and bottom surfaces of a wafer to achieve uniform temperature-trajectory tracking. However, thermal distortions develop near the wafer edges during processing. Many rapid thermal processes [9] direct additional amounts of energy toward the edges to counteract these thermal nonuniformities occurring at the wafer edges, thus achieving results similar to our inverse incident-heat-flux results shown in Fig. 4. Inverse dynamic incident-heat-flux profile results for wafer surface thermal nonuniformities at three linear ramp-up rates with a measurement error of $\sigma = 0.0$ are shown in Fig. 6(a)–(c), respectively. The vertical axes represent thermal nonuniformities graphed according to temperature differences between points on the wafer surface and the wafer surface center. These figures show that when incident heat-flux profiles are controlled as in our inverse-method approach, temperature
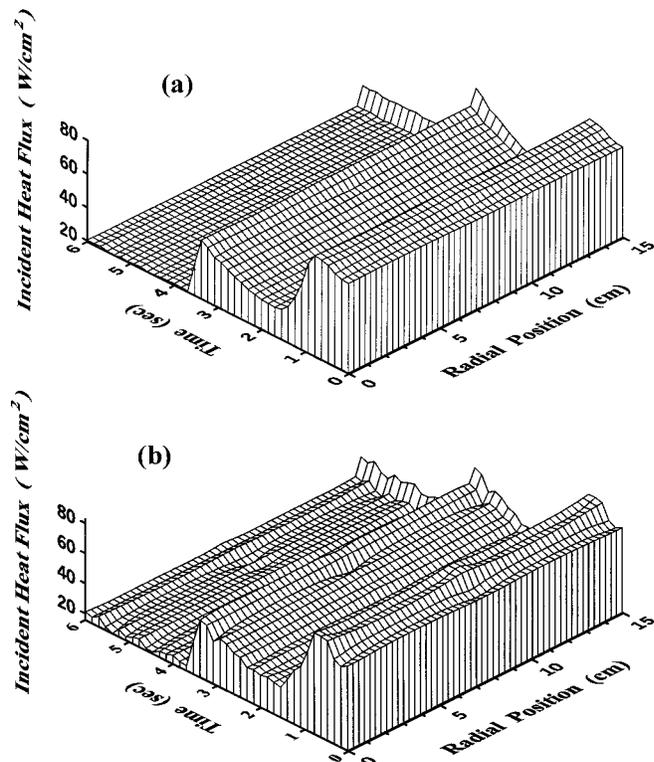

Fig. 5. Inverse incident-heat-flux profiles results at a linear 300 °C/s ramp-up rate for measurement errors of: (a) $\sigma = 0.001$ and (b) $\sigma = 0.005$.

differences develop at the edges. Initially, the temperature difference (thermal nonuniformity) is not significant, however, as the ramp-up proceeds, the thermal nonuniformity developed at the edge increases with increasing edge-heating compensation, as shown in Fig. 4. When the wafer reaches the higher steady state, the incident-heat-flux profile changes from the transient stage to the steady stage, the thermal nonuniformity drops gradually and approaches the steady-state. Thus, edge-heating compensation has an overheating effect on thermal uniformity during processing.

Generally, the temperature over the wafer must be maintained within 2 °C of the wafer center during rapid thermal processing [6]. Fig. 6(a)–(c) shows that the thermal nonuniformity was not significant in the present inverse incident-heat-flux profiles, even though during transient periods, resulting maximum temperature differences were 0.835 °C, 1.174 °C, and 1.516 °C for the 100 °C/s, 200 °C/s, and 300 °C/s ramp-up rates, respectively. It was found that thermal nonuniformities occurring during ramp-up increased with the ramp-up rate, but remained within 1.6 °C during processing. Fig. 7(a) and (b) shows the respective inverse incident-heat-flux results on thermal nonuniformity for the 300 °C/s ramp-up rate when measurement errors of $\sigma = 0.001$ and $\sigma = 0.005$ were introduced. The resulting maximum temperature differences were 1.523 °C and 1.493 °C for measurement errors $\sigma = 0.001$ and $\sigma = 0.005$, respectively. The thermal nonuniformity decreased as the measurement error σ was increased from 0.001 to 0.005. The maximum temperature difference was less than 1.6 °C even though the measurement error did reach 3.864 °C (in the case of $\sigma = 0.005$). The thermal nonuniformity could be acceptable during RTP.
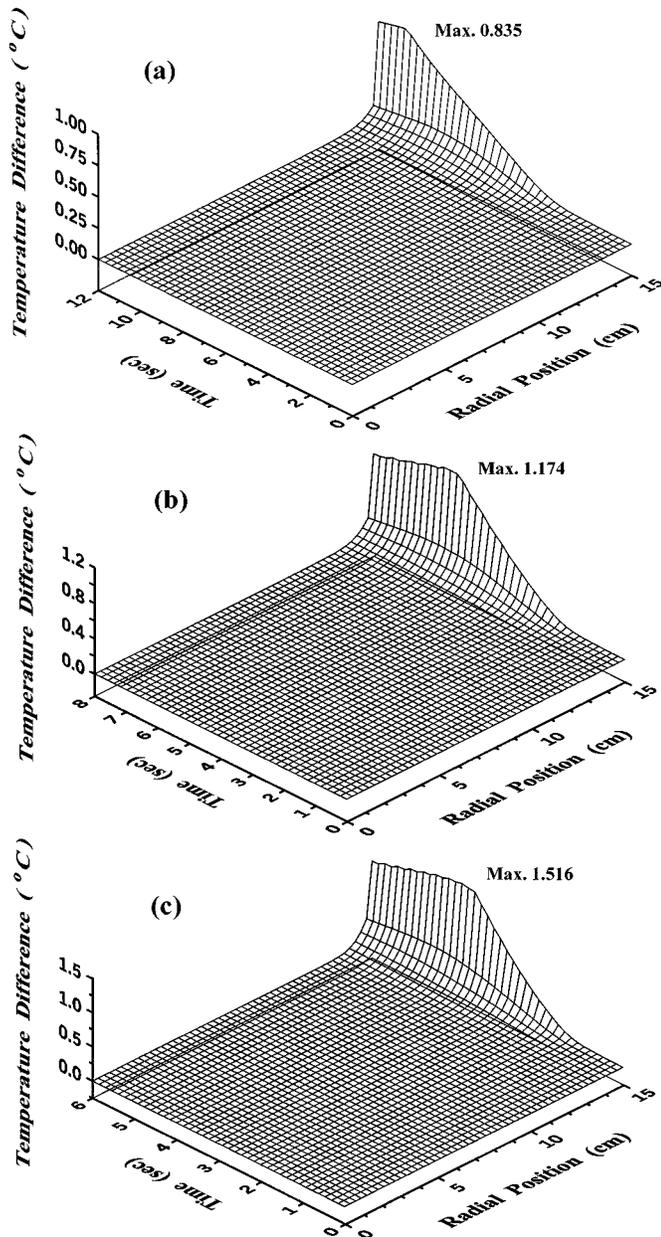
Fig. 6. Inverse incident-heat-flux thermal nonuniformity profile results for a measurement error of $\sigma = 0.0$ at linear ramp-up rates of: (a) 100 °C/s, (b) 200 °C/s, and (c) 300 °C/s.
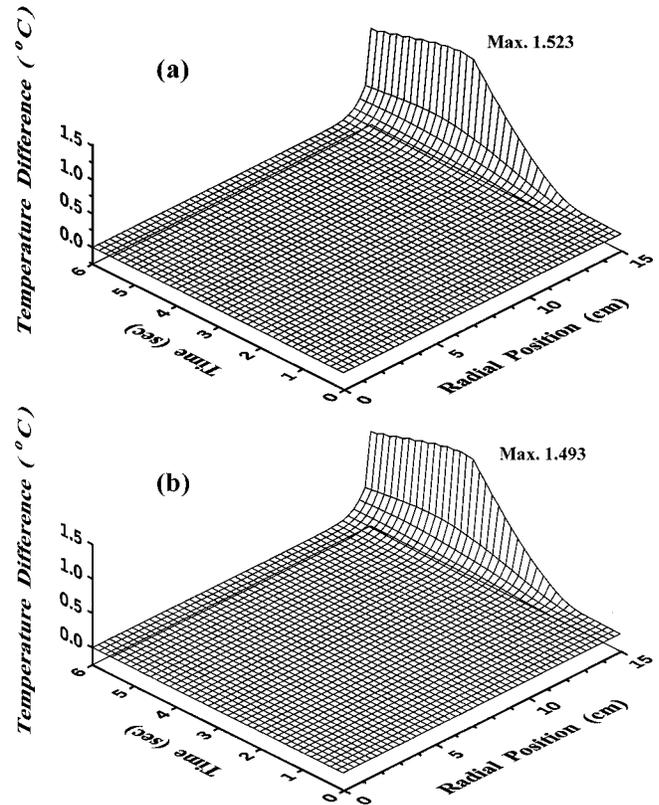

Fig. 7. Inverse incident-heat-flux thermal nonuniformity profile results at a linear 300 °C/s ramp-up rate for measurement errors of: (a) $\sigma = 0.001$ and (b) $\sigma = 0.005$.
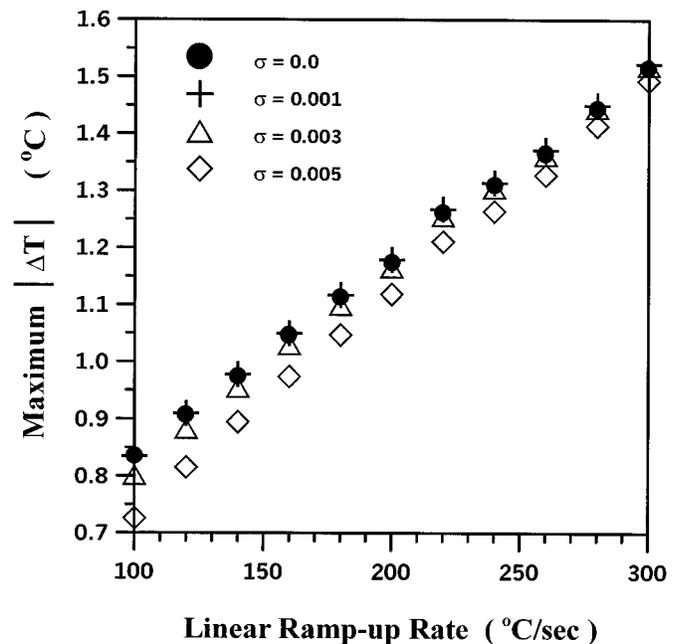

Fig. 8. Resulting maximum temperature difference for measurement errors of $\sigma = 0.0$, $\sigma = 0.001$, $\sigma = 0.003$ and $\sigma = 0.005$ as a function of the linear ramp-up rate.

Fig. 8 illustrates the resulting maximum temperature differences ($|\Delta T|$, the absolute value of temperature difference between the edge and the center of the wafer surface) during transients as a function of the desired linear ramp-up rates for measurement errors of $\sigma = 0.0$, 0.001, 0.003, and 0.005, respectively. Our present results show that the maximum temperature differences occurring during ramp-up increased with the ramp-up rate. Increasing measurement error σ from 0.001 to 0.005, the thermal nonuniformity of the inverse results decreases. Although a linear ramp-up rate of 300 °C/s was used and measurement errors did reach 3.864 °C (in the case of $\sigma = 0.005$), the surface temperature was maintained within 1.6 °C of the center of the wafer surface when the incident-heat-flux

profiles were dynamically controlled according to the inverse-method approach. These thermal nonuniformities could be acceptable in RTP systems.

V. CONCLUSION

This paper presents a systematic method for estimating incident heat flux on a 300-mm-diameter 0.775-mm-thick silicon wafer to achieve uniform temperature tracking at several linear ramp-up rates during rapid thermal processing using the inverse heat-transfer method. Temperature-dependent thermal properties of the silicon wafer were considered in this study. Using a 2-D thermal model, temperature solutions for the inverse-method matrices can be constructed by applying the finite-difference scheme to calculate the desired incident-heat-flux profiles required for uniform temperature tracking. In the present study, the wafer was ramped-up from an initially uniform 27 °C temperature to a steady state of 1097 °C via simulation at several linear ramp-up rates. The resulting maximum temperature differences as a function of the desired linear ramp-up rates using several measurement errors were investigated. The maximum temperature differences in our present study were 0.835 °C, 1.174 °C, and 1.516 °C, respectively, for the 100 °C/s, 200 °C/s, and 300 °C/s ramp-up rates when the incident heat fluxes on the wafer could be dynamically controlled according to the inverse-method approach. Thermal nonuniformities occurring during ramp-up increased with the ramp-up rate. Although a linear 300 °C/s ramp-up rate was used and the dimensional measurement error reached 3.864 °C ($\sigma = 0.005$), the resulting maximum temperature differences were not significant and remained under 1.6 °C when the incident-heat-flux profiles were dynamically controlled according to the inverse-method approach. These thermal nonuniforimities could be acceptable in RTP systems.

REFERENCES

[1] V. E. Borisenko and P. J. Hesketh, *Rapid Thermal Processing of Semiconductors*.   New York: Plenum, 1997.
[2] H. A. Lord, "Thermal and stress analysis of semiconductor wafers in a rapid thermal processing oven," *IEEE Trans. Semiconduct. Manufact.*, vol. 1, pp. 105–114, Aug. 1988.
[3] C. Hill, S. Jones, and D. Boys, "Rapid thermal annealing—Theory and practice," *Reduced Thermal Procesing for ULSI, NATO ASI Series B: Physics*, pp. 143–180, 1989.
[4] R. Kakoschke, E. Bubmann, and H. Föll, "Modeling of wafer heating during rapid thermal processing," *Appl. Phys. A*, vol. 50, pp. 141–150, 1990.
[5] R. S. Gyurcsik, T. J. Riley, and F. Y. Sorrell, "A model for rapid thermal processing: Achieving uniformity through lamp control," *IEEE Trans. Semiconduct. Manufact.*, vol. 4, pp. 9–13, Feb. 1991.
[6] F. Y. Sorrell, M. J. Fordham, M. C. Öztürk, and J. J. Wortman, "Temperature uniformity in RTP furnaces," *IEEE Trans. Electron Devices*, vol. 39, pp. 75–79, Jan. 1992.
[7] T. J. Riley and R. S. Gyurcsik, "Rapid thermal procesor modeling, control, and design for temperature uniformity," in *Mat. Res. Soc. Symp. Proc.*, vol. 303, 1993, pp. 223–229.
[8] Y. M. Cho, A. Paulraj, T. Kailath, and G. Xu, "A contribution to optimal lamp design in rapid thermal procesing," *IEEE Trans. Semiconduct. Manufact.*, vol. 7, pp. 34–41, Feb. 1994.
[9] R. H. Perkins, T. J. Riley, and R. S. Gyurcsik, "Thermal uniformity and stress minimization during rapid thermal processes," *IEEE Trans. Semiconduct. Manufact.*, vol. 8, pp. 272–279, Aug. 1995.
[10] J. V. Beck, B. Blackwell, and C. R. St. Clair, *Inverse Heat Conduction—III-posed Problem*.   New York: Wiley, 1985.
[11] E. Hensel, *Inverse Theory and Applications for Engineers*.   Englewood Cliffs, NJ: Prentice-Hall, 1991.
[12] O. M. Alifanov, *Inverse Heat Transfer Problems*.   Berlin, Germany: Springer-Verlag, 1994.
[13] K. Kurpisz and A. J. Nowak, *Inverse Thermal Problems*.   Boston, MA: Computational Mechanics Publications, 1995, pp. 106–108.
[14] S. Lin and H. S. Chu, "Thermal uniformity of 12-inch silicon wafer during rapid thermal processing by inverse heat transfer method," *IEEE Trans. Semiconduct. Manufact.*, to be published.
[15] ——, "Using inverse modeling to estimate the incident heat flux required to achieve temperature uniformity across circular disk," Microscale Thermophys. Eng., 2000, to be published.
[16] J. P. Hebb and K. F. Jensen, "The effect of multilayer patterns on temperature uniformity during rapid thermal processing," *J. Electrochem. Soc.*, vol. 143, no. 3, pp. 1142–1151, Mar. 1996.
[17] P. Y. Wong, I. N. Miaoulis, and C. G. Madras, "Transient and spatial radiative properties of patterned wafers during thermal processing," in *Mat. Res. Soc. Symp. Proc.*, vol. 387, 1995, pp. 15–20.
[18] A. Virzi, "Computer modeling of heat transfer in Czochralski silicon crystal growth," *J. Cryst. Growth*, vol. 112, pp. 699–722, 1991.
[19] D. A. Anderson, J. C. Tannehill, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*.   New York: McGraw-Hill, 1984, pp. 134–135.

Senpuu Lin received the B.S. degree from National Cheng Kung University, Tainan, Taiwan, R.O.C., and the M.S. degree from Lamar University, Beaumont, TX, both in mechanical engineering, in 1979 and 1985, respectively. He is currently pursuing the Ph.D. degree in mechanical engineering at National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

Since 1986, he has been an instructor in National Lien-Ho Institute of Technology, Miaoli, Taiwan, R.O.C. His research interests include heat transfer problem by inverse methods and semiconductor manufacturing.

Hsin-Sen Chu was born in Hsinchu, Taiwan, on June 12, 1952. He received the B.S., M.S., and Ph.D. degrees from National Cheng Kung University, Tainan, Taiwan, R.O.C., all in mechanical engineering, in 1974, 1977 and 1982, respectively.

Presently, he is a Professor in the Department of Mechanical Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C.. He joined the NCTU faculty in 1984. From 1985 to 1986, he was a Visiting Scholar at University of California, Berkeley. His current research interests are concerned with the heat transfer in semiconductor manufacturing, microscale heat transfer, and energy engineering.

Dr. Chu was the recipient of the Outstanding Teaching Award sponsored by the Ministry of Education, R.O.C., in 1992 and the Outstanding Research Award sponsored by the National Science Council, R.O.C. in 1999.