CHAPTER 6

IN SITU SENSING AND ENDPOINT DETECTION IN COPPER CMP

This chapter presents an *in situ* sensing technique for monitoring the polishing state and for detecting process endpoint for Cu CMP. An optical reflectance sensor with a spectral distribution ranging from 775 nm to 990 nm is employed on a rotary polishing tool to measure the surface reflectance of patterned Cu wafers. Theories of reflectance related to surface topography and surface composed of various coating materials are proposed. Sensor kinematics is studied to design the loci and to map the sampled data onto the wafer surface to gain both global and local information of the surface. Additionally, statistical models are developed to determine the surface condition based on the characteristics of the sample reflectance data. A nested variance model is proposed to decouple the variance components of within-wafer and within-die nonuniformity. This information may be employed to control the process parameters to improve the material removal rate and the within-wafer uniformity. Algorithms based on the changes in mean, variance, range, and distribution of the surface reflectance are examined for their effectiveness in process monitoring and endpoint detection. Experiments are also performed to validate these schemes.

6.1 Introduction

The copper damascene process is emerging as a crucial technology to produce high-speed, high-performance, and low energy-consuming Ultra-Large-Scale Integrated (ULSI) circuits. In Cu damascene, the CMP process removes excess Cu and barrier materials (typically Ta, Ti, TaN or TiN) to form interconnects inside the trenches in the inter-layer dielectric (ILD, typically SiO_2 or polymers). The MRR of Cu strongly depends on the pattern geometry (Steigerwald et al., 1994, Park et al., 1999). The nonuniform pattern layout usually causes

nonuniform polishing across the die area. It results in partial overpolishing of the region with higher Cu area fraction and dishing of the soft Cu lines. The Cu loss and surface nonuniformity due to overpolishing and dishing may affect the reliability and performance of interconnects and so must be minimized. Additionally the spatial variation of the process parameters (velocity, pressure, slurry transport, etc.) and the process random variation will increase the within-wafer and within-lot nonuniformity (Smith et al., 1996; Litvak and Tzeng, 1996). These will vary the cycle time, or the endpoint, of the Cu CMP and affect the process yield. In order to reduce the variance of polishing outputs (uniformity, overpolishing and dishing), it is necessary to integrate an *in situ* sensing and endpoint detection technique with the process optimization schemes to improve process performance.

The ideal wafer-level endpoint for copper CMP process may be defined as the time when the excess Cu and barrier layers are fully cleared on all dies of a wafer. Due to polishing nonuniformity, however, all the dies on a wafer generally will not reach the endpoint at the same moment. Some of the dies may be overpolished. Thus the endpoint of CMP can be a representation of the optimal polishing time when the number of out-of-specification dies (either under- or over-polished) reaches a minimum and the process yield is maximized. However, the remaining Cu thickness in each die area is difficult to measure in real-time to determine the endpoint. Many of the present *in situ* sensing techniques rely on indirect methods to detect the moment of Cu/barrier clear-up, such as the changes in the friction force, the ion concentrations of the Cu/barrier materials, and the electrical impedance of the surface (Sandhu et al., 1991; Schultz, 1992; Leach et al., 1993; Litvak et al., 1996; Murarka et al., 1997; Bibby and Holland, 1998; Beckage et al., 1999; Stien and Hetherington, 1999). However, these methods are limited in practical applications by lack of reliability and high noise-to-signal ratio. Moreover, these techniques provide only average information over a relative large area (usually wafer-level). They lack the capability of sensing within-wafer and die-level uniformity. These methods can best be used as supplementary methods in ensuring the detection of endpoint.

A photoacoustic technique for thickness measurement of multi-layer stacked films has recently been proposed (Rogers et al., 1997; Banet et al., 1998; Joffe et al., 1999). Two optical excitation pulses overlap on the surface of the coating to form an interference pattern.

Absorption of light by the film generates counter-propagating acoustic waves. The film thickness can be calculated by measuring the acoustic frequency. However, this method is limited to a blanket area with the dimensions much larger than the beam size. It is difficult to model the generation and the propagation of the acoustic wave in thin Cu film on the patterned area. Hence, this method is currently used for the measurements for blanket wafers or large patterns which can be simulated as blanket areas.

Among all the available endpoint detection techniques, the optical sensing techniques of interferometry and reflectance measurement provide the localized means for detecting the endpoint (Bakin et al., 1998; Chan et al., 1998; Fang et al., 1998; Bibby et al., 1999, Nam, 2000). The interferometry technique measures the film thickness based on the interference of light from the surface and the underlying layers. This is suitable for measuring transparent films such as dielectric layers, but not effective for opaque metal films. By contrast, the reflectance measurement is ideal for detecting the surface topography and the metal area fraction on the surface. The surface condition in polishing and the process endpoint may be determined by analyzing the measured data. Polishing through the Cu layer sharply reduces the intensity of reflecting light due to a substantial difference in the reflectance of Cu and SiO_2. The extent of Cu remaining on the detected subdie area can be determined and mapped by moving a small beam light source across the wafer. Moreover, it may also be possible to gain information on surface planarity and dishing by this metrology because the reflectance of patterned surface is influenced by the topography of the pattern (Rayleigh, 1907; Eckart, 1933; Beckmann and Spizzichino, 1963; Uretsky, 1965; Zipin, 1966; Desanto, 1975 and 1981; Bahar, 1981).

Several issues must be resolved to fully utilize the reflectance measurement technique. The signal from the sensor cannot be unambiguously interpreted because it contains both systematic and random attributes of the surface condition of the polished surface and the polishing environment. The reflectance varies across the wafer due to the nonuniform pattern layout and within-wafer nonuniform polishing. These effects must be decoupled to gain information for polishing uniformity control. The effects of light scattering due to slurry particles and the increase of surface roughness and the diffraction effect from the submicron features must also be understood. That will help to analyze the measured data and decompose

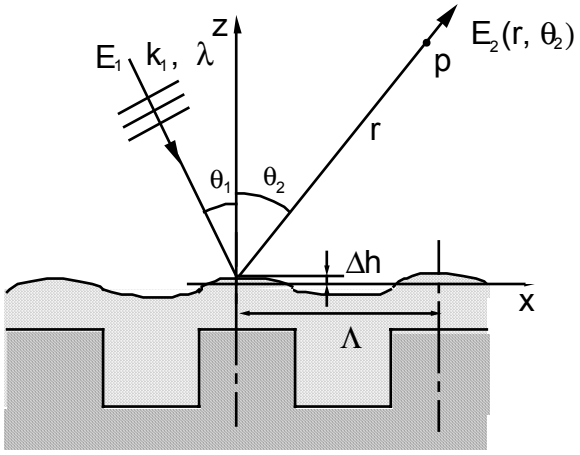
the " noise-like" signal to reveal the true process endpoint. In addition to the interpretation of the signal, the sensor kinematics must be studied to map the data onto the wafer surface to gain both wafer- and die-level information. An adequate spatial resolution for indicating local conditions can be achieved through sensor trajectory design. Such information will increase the confidence level of interference about the wafer surface condition, and will lead to a more accurate and repeatable decision-making for the endpoint detection.

This chapter employs the reflectance measurement technique on both blanket and patterned Cu wafers. It examines the feasibility of *in situ* sensing in Cu CMP. Theories relating the reflectance to both surface topography and the existence of materials with different reflection coefficients are proposed and verified by experiments. Sensor kinematics is studied to design the loci and sampling plan for the surface condition. A statistical model is developed to decouple the variation of the surface reflectance and to estimate the variance components contributed by nonuniform polishing due to pattern variation and within-wafer nonuniform polishing. Algorithms for utilizing those two techniques in process endpoint monitoring and control are also discussed.

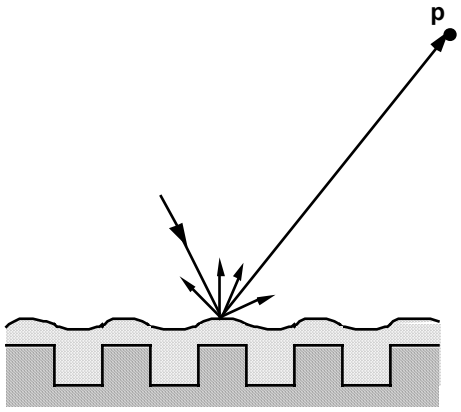
6.2 Theory

6.2.1 Light Scattering from a Periodic Surface Structure. The scattering of light by a periodic wavy surface has been investigated by many researchers (Rayleigh, 1907; Eckart, 1933; Beckmann and Spizzichino, 1963; Uretsky, 1965; Desanto, 1975 and 1981). This section reviews important formulations and their solutions to understand the effects of pattern geometry on surface reflectance by scattering. Consider the problem of plane wave being scattering by a periodic surface S, where $z = h(x)$, as shown in Fig. 6.1 (a). Let E_1 and E_2 denote the incident and scattered fields. The incident light (electric) field E_1, assumed to be unit amplitude, can be expressed as:

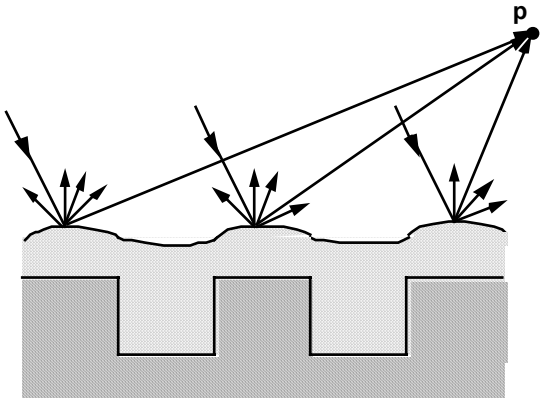
$$E_1 = \exp[i(k_1 \sin\theta_1 x - k_1 \cos\theta_1 z) - i\omega t]; \tag{6.1}$$



(a)



(b)



(c)

Figure 6.1 Schematics of light scattering on a patterned Cu surface.

where k_1 is the wave number of the incident light wave ($k_1 = 2\pi/\lambda$), θ_1 the incident angle, ω the angular frequency ($\omega = 2\pi f$), and t the time. If just the scattered field at a fixed time is of interest, the $\exp(-i\omega t)$ can be suppressed. The scattered field E_2 at any point P of observation above the surface is given by the Holmholtz integral (Beckmann, 1963):

$$E_2(P) = \frac{1}{4\pi} \iint_s \left(E \frac{\partial \psi}{\partial n} - \psi \frac{\partial E}{\partial n} \right) ds \tag{6.2}$$

with:

$$\psi = \exp(ik_2 r) / r \tag{6.3}$$

where r is the distance between the given observation point P and any point on the surface $(x, h(x))$, and k_2 is the wave number of the scattered wave ($k_2 = k_1 = 2\pi/\lambda$). The point P is assumed in the Fraunhofer zone, i.e. $r \to \infty$, to focus on the plane scattered waves rather than spherical ones. In order to solve the scattered field E_s in Eq. (6.2), the total field E and its normal derivative $\partial E/\partial n$ on the boundary surface can be approximated as ("Kirchhoff's method"):

$$(E)_S = (1+\gamma) E_1 \tag{6.4}$$

and:

$$\begin{aligned} \left(\frac{\partial E}{\partial n} \right)_S &= (1-\gamma) E_1 (\mathbf{k}_1 \cdot \mathbf{n}) \\ &= (1-\gamma) E_1 \left(k_1 \sin\theta_1 \frac{\partial h(x)}{\partial x} - k_1 \cos\theta_1 \right) \end{aligned} \tag{6.5}$$

where γ is the reflection coefficient of a planar surface and $\mathbf{n}$ is the unit vector normal to the surface at the point of interest. The reflection coefficient γ depends not only on the local angle of incidence and the electrical properties of the surface material but also on the polarization of

the incident wave. For simplicity, the surface is assumed to be perfectly conducting, i.e., γ = -1 for a horizontal polarization (electric vector perpendicular to the incident plane) for the following analysis.

Equation (6.2) can be integrated over a specified periodic surface profile, such as a sinusoidal surface pattern:

$$z = h(x) = (\Delta h)\cos(2\pi x / \Lambda) \tag{6.6}$$

where Δh is the half step-height and Λ the pitch of the features. The scattered field will also follow the same period Λ along the x direction, which simplifies the integral in Eq. (6.2) by integrating within one period instead of over the entire surface. Moreover, the periodicity of the surface topography implies that the scattered field can be written as a superposition of the Fourier series representing the plane waves at different modes, in which the reflective (scattering) angle of each mode θ_{2m} follows the relation (grating equation).

$$\sin\theta_{2m} = \sin\theta_1 + m\lambda / \Lambda \qquad (m = 0, \pm 1, \pm 2, ...) \tag{6.7}$$

The zero mode represents the condition of specular reflection, in which $\theta_2 = \theta_1$, and the direction of the scattered plane wave will be away from the specular angle for larger m. The solution for the scattered field in the primary direction of each mode θ_{2m} at the far field can be obtained by applying Eqs. (6.3), (6.4), (6.5), (6,6), and (6.7) into Eq. (6.2) and performing integration over the surface ($-L \le x \le L$). The reflection coefficient γ is written as a function of optical properties of the coating and the local angle of incidence to calculate the integration. The result can be normalized by the field reflected on a specular planar surface E_{20}, which defines the scattering coefficient ϕ ($=E_2/E_{20}$), and can be written as (Beckmann, 1963):

$$\phi(\theta_1, \theta_{2m}) = -\sec\theta_1 \frac{1 + \cos(\theta_1 + \theta_{2m})}{\cos\theta_1 + \cos\theta_{2m}} (-i)^m J_m(s) + C_1(n_1) \tag{6.8}$$

where J is the Bessel function, $s = 2\pi\Delta h / \lambda(\cos\theta_1 + \cos\theta_2)$, and n_1 the residual parts of the ratio L/Λ. Equation (6.8) only gives the scattering coefficient at the primary scattering angle of each mode. For θ_2, the result is given as

$$\phi(\theta_1,\theta_2) = -\frac{\sin 2np\pi}{2n\sin p\pi}\sec\theta_1 \frac{1+\cos(\theta_1+\theta_2)}{\cos\theta_1+\cos\theta_2} e^{-ip\pi}\left[J_{-p}(s) + \frac{\sin p\pi}{\pi}\int_0^\infty e^{pt - s\sinh t}dt\right] + C_2(n_1) \tag{6.9}$$

where $p = (L/\lambda)(\sin\theta_1 - \sin\theta_2)$, $s = 2\pi\Delta h / \lambda(\cos\theta_1 + \cos\theta_2)$ and n the integer parts of the ratio L/Λ. In the far field (Fraunhofer zone, i.e. $r \to \infty$), only one mode of scattered plane wave can be observed at the given point P (in the direction of θ_2), as shown in Fig. 6.1 (b). As shown in Fig. 6.1 (c) in the near field, or the Fresnel zone, the total scattered field at P, normalized by E_{20}, is given by superposing all the scattering modes contributed from the neighboring periodic surface. Both the amplitude and the direction of each mode, given by Eqs. (6.8) and (6.9) and the phase difference between each mode, must be considered in calculating the total scattered field. In practice, the calculation of the total scattered field may be complicated and may need to be performed numerically for the sensor located near the measured surface. Diffusion scattering takes place when the $\Delta h/\lambda$ ratio increases with constant pitch Λ (Brekhovskikh, 1952). Light will be scattered away from the direction of specular reflection, i.e., light is reflected into the direction of higher scattered modes (larger m) and will not be received by the sensor. Therefore the surface reflectance, which is proportional to the square of amplitude of the reflecting field, decreases with the step-height of the feature $2\Delta h$ with Δh comparable or larger than the wavelength of incident light. On the contrary, when the surface is planar, i.e. $\Delta h \approx 0$, the surface reflectance will be close to that of a specular surface. Moreover, based on the law of energy conservation, the overall scattering coefficient ϕ should be always equal or less than unity.

The number of possible modes m for the scattering field is limited by the condition that $\alpha_n = \sin\theta_n$ is less than unity. If $2\pi/kL$ (or λ/L) is close to unity, i.e., the wavelength of the incident light is close to the waviness of the pattern, there will be only one mode and the surface will reflect specularly regardless of the surface topography. For the submicron Cu

patterns employed in the current design, the reflectance measured at the onset of the process endpoint by a light source with comparable or larger wavelength will essentially indicate the Cu area fraction only. The slight surface topography due to overpolishing and dishing will not significantly affect the reflectance. As shown in Fig. 6.2, the surface reflectance *R,* proportional to the square of the reflection coefficient, of the composite surface at the onset of endpoint therefore can be written as:

$$R = A_f R_{Cu} + (1 - A_f) R_{Oxide} \tag{6.10}$$

where A_f is the area fraction of Cu interconnects, and R_{cu} and R_{oxide} the reflectances of Cu and TEOS, respectively, in specular reflection.

6.2.2 Sensor Kinematics. The sensor loci on the rotating wafer surface can be determined by the relative velocity of the sensor to the wafer and the initial position of the sensor, as shown in Fig. 6.3. The relative velocity of the sensor on the rotating wafer can be obtained by two steps: find the relative velocity of the sensor to the stationary *X, Y* coordinates fixed at the center of the wafer and then perform a coordinate transformation with respect to the wafer rotation. The velocity components for the sensor, $v_{X,s}$, and $v_{Y,s}$, and the wafer, $v_{X,w}$, and $v_{Y,w}$, in *X, Y* coordinates can be expressed as:

$$v_{X,s} = -r_s \omega_p \sin(\omega_p t + \theta_0) - \dot{r}_{cc} \tag{6.11a}$$

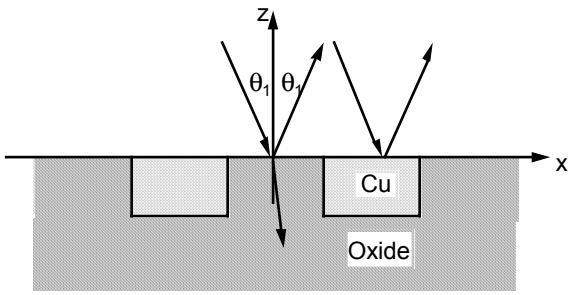
$$v_{Y,s} = r_s \omega_p \cos(\omega_p t + \theta_0) \tag{6.11b}$$

and:

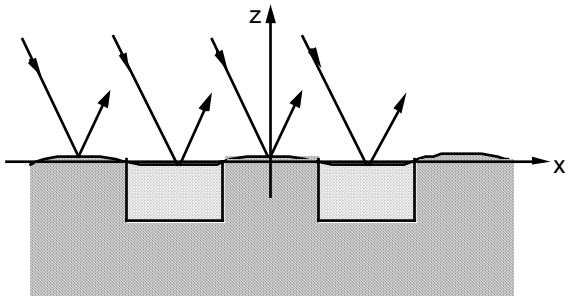
$$v_{X,w} = -r_s \omega_w \sin\theta \tag{6.12a}$$

$$v_{Y,w} = \omega_w (r_s \cos\theta - r_{cc}) \tag{6.12b}$$

where r_s is the offset of the sensor from the center of the platen, r_{cc} the offset of the centers of the wafer and the platen, ω_w and ω_p are the angular velocity of the wafer and the platen, and θ the angle of the sensor with respect to the *X* coordinate. In addition to wafer rotation, the wafer may translate relative to the platen center, so called sweeping, with a velocity $\dot{r}_{cc}$ to



(a)



(b)

Figure 6.2 Schematics of light scattering from (a) a planar composite surface, and (b) a wavy composite surface.

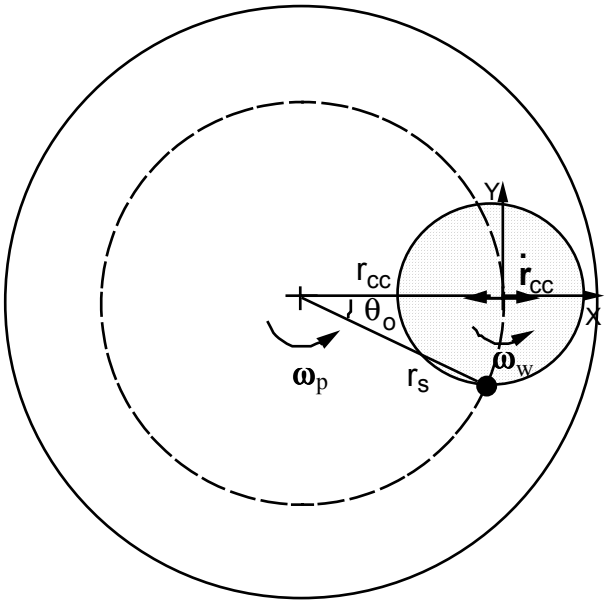


Figure 6.3 Sensor kinematics.

utilize the entire pad surface. For simplicity, the sweeping is assumed to be along the X coordinate only. Therefore, the components of the relative velocity of the sensor to the wafer, $v_{X,R}$ and $v_{Y,R}$, in X, Y coordinates can be written as:

$$\begin{aligned} v_{X,R} &= v_{X,s} - v_{X,w} = [-r_s\omega_p \sin(\omega_p t + \theta_0) - \dot{r}_{cc}] + r_s\omega_w \sin\theta \\ &= -r_s(\omega_p - \omega_w)\sin(\omega_p t + \theta_0) - \dot{r}_{cc} \end{aligned} \tag{6.13a}$$

and:

$$\begin{aligned} v_{Y,R} &= v_{Y,s} - v_{Y,w} = r_s\omega_p \cos(\omega_p t + \theta_0) - \omega_w(r_s\cos\theta - r_{cc}) \\ &= r_s(\omega_p - \omega_w)\cos(\omega_p t + \theta_0) + \omega_w r_{cc} \end{aligned} \tag{6.13b}$$

These velocity components can also be represented in terms of a rotating coordinate system *(x, y),* with the origin located at the center of the wafer and rotating at the same angular velocity ω_w as the wafer. The velocity components on the rotating coordinates, $v_{x,R}$ and $v_{y,R}$, are given by the coordinate transformation rule:

$$\begin{bmatrix} v_{x,R} \\ v_{y,R} \end{bmatrix} = \begin{bmatrix} \cos\omega_w t & \sin w_w t \\ -\sin\omega_w t & \cos w_w t \end{bmatrix} \begin{bmatrix} v_{X,R} \\ v_{Y,R} \end{bmatrix} \tag{6.14}$$

and can be written as:

$$v_{x,R} = -r_s(\omega_p - \omega_w)\sin((\omega_p - \omega_w)t + \theta_0) + r_{cc}\omega_w \sin\omega_w t - \dot{r}_{cc}\cos\omega_w t \tag{6.15a}$$

$$v_{y,R} = r_s(\omega_p - \omega_w)\cos((\omega_p - \omega_w)t + \theta_0) + r_{cc}\omega_w \cos\omega_w t + \dot{r}_{cc}\sin\omega_w t \tag{6.15b}$$

Therefore, the displacement of the sensor on the wafer with respect to the rotating x, y coordinates is given by integrating the velocity in Eqs. (6.15a) and (6.15b):

$$\begin{aligned} x &= \int v_{x,R}\,dt \\ &= -r_s(\omega_p - \omega_w)\int \sin[(\omega_p - \omega_w)t + \theta_0]dt + \omega_w \int r_{cc}\sin\omega_w t dt - \int \dot{r}_{cc}\cos\omega_w t dt \end{aligned} \tag{6.16a}$$

$$y = \int v_{y,R} dt$$

$$= r_s(\omega_p - \omega_w) \int \cos[(\omega_p - \omega_w)t + \theta_0]dt + \omega_w \int r_{cc} \cos \omega_w t dt + \int \dot{r}_{cc} \sin \omega_w t dt \qquad (6.16b)$$

The initial condition must be prescribed to solve Eqs. (6.16a) and (6.16b) for the position of the sensor on the wafer surface at a given time. It is convenient to assume that the sensor is initially located at the edge the wafer, with a initial angle θ_0 with respect to the fixed X coordinate. For simplicity, it is also assumed that no sweeping motion occurs in polishing, i.e., $\dot{r}_{cc} = 0$. In practice, the effect of sweeping motion on the sensor trajectory across the wafer can be neglected if the sweeping velocity is much lower than the liner velocities of the wafer relative to the pad. With those assumptions, the position of the sensor on the wafer can be expressed as:

$$x = r_s \cos[(\omega_p - \omega_w)t + \theta_0] - r_{cc} \cos \omega_w t \qquad (6.17a)$$

$$y = r_s \sin[(\omega_p - \omega_w)t + \theta_0] + r_{cc} \sin \omega_w t \qquad (6.17b)$$

As long as the condition $x^2+y^2 < r_w$ (where r_w is the radius of the wafer) is satisfied, the sensor is located inside the wafer/pad contact interface. Because the wafer is against the platen in polishing, the sensor trajectory given in Eqs. (6.16) and (6.17) is observed from the wafer backside. The trajectory on the front surface is symmetric to the results from Eqs. (6.16) and (6.17) with respect to the y axis.

When the angular velocities of the wafer and the platen are equal, i.e. $\omega_w = \omega_p$, Eqs. (6.17a) and (6.17b) can be further simplified and the locus of the sensor is an arc with the radius equals to r_{cc} and centered at ($r_s cos\theta$, $r_s sin\theta$,) relative to the rotating x, y coordinates:

$$(x - r_s \cos\theta_0)^2 + (y - r_s \sin\theta_0)^2 = {r_{cc}}^2 \qquad (6.18)$$

When the angular velocities of the wafer and the platen are the same, the sensor enters the wafer/pad interface at the same point on the periphery of the wafer. It always produces the same locus on the wafer surface, as shown in Fig. 6.4. In practice, the angular velocity of the

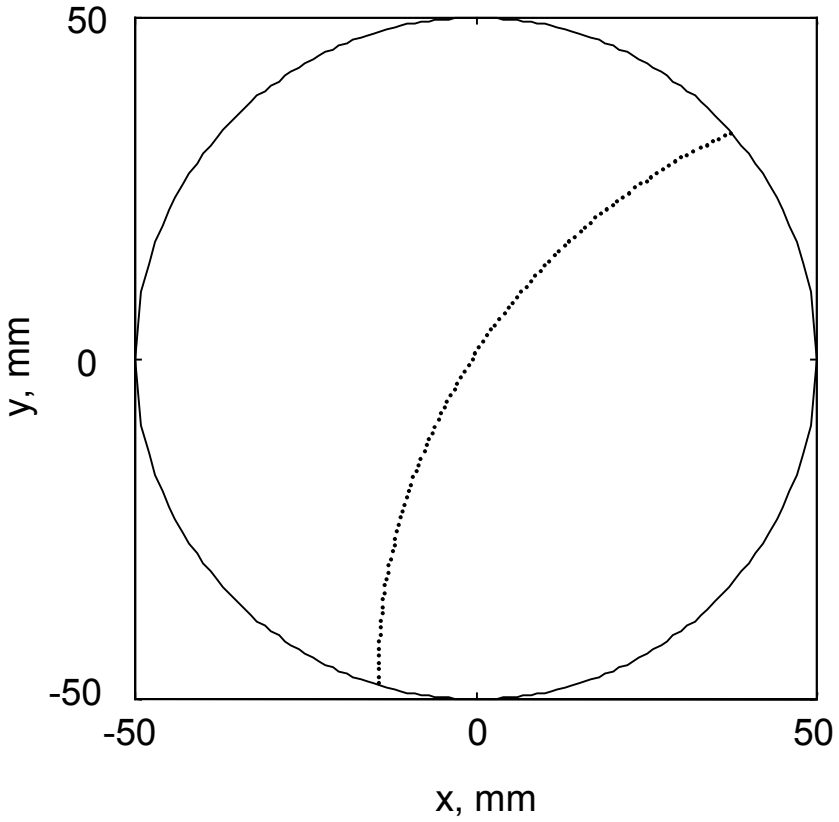


Figure 6.4 The simulated locus for the reflectance sensor across the wafer at the condition $\omega_w = \omega_p$ and $r_s = r_{cc}$.

wafer must be slightly offset from the platen so that the sensor can scan over the entire wafer surface in different radial directions. Figure 6.5 shows the sensor loci for the conditions ω_w = 1.05 ω_p and $r_s = r_{cc}$, in which 20 identical loci start from 20 equally spaced points on the periphery of the wafer edge, respectively and repeatedly, if no wafer slippage occurs. As illustrated, the sampling density will be much higher at the center of the wafer, and lower at the edge where more dies are located. The lower sampling density on the edge dies might result in bias inference for the overall surface condition. How to design sensor loci to sample enough data on a desired surface area will be discussed later in detail.

6.2.3 Statistical Treatment of the Sampled Data. The surface conditions of the wafer during polishing can be extracted from the real-time reflectance data. The statistics used to infer the surface conditions include the maximum and minimum reflectance values, the range, the mean value, the variation, the shape of the distribution of the reflectance data, etc. Three levels, wafer-, die- and device- or subdie-level, of information can be obtained from the dataset. The spot size of the sensor is chosen so that it is comparable to or smaller than the subdie area but still much larger than the dimensions of interconnects. Therefore, an individual measurement represents the reflectance on the specific device or pattern area on the wafer, from which the surface topography and Cu area fraction can be inferred. It is difficult to map the measurement results onto the exact location of a particular device or pattern because of wafer slippage inside the carrier. The individual datum can only be mapped onto the surface within a grossly defined area. Similarly, the die-level information may be obtained along a specific segment(s) corresponding to the die location on the loci. However, it may only roughly represent the surface condition within the vicinity of the interested die region. Polishing results for the dies at the same radial distance from the wafer center very often exhibit similar trends. Hence, data from within adjacent dies at the same radius sometimes may be combined to increase the sample size for the die at a particular radius to elucidate the spatial dependence of material removal in the radial direction.

Moreover, wafer-level information can be retrieved either from a single or multiple scans across the wafers. In endpoint detection, it is preferable to take samples from multiple loci so that the surface condition over a specific region or the entire wafer surface can be determined

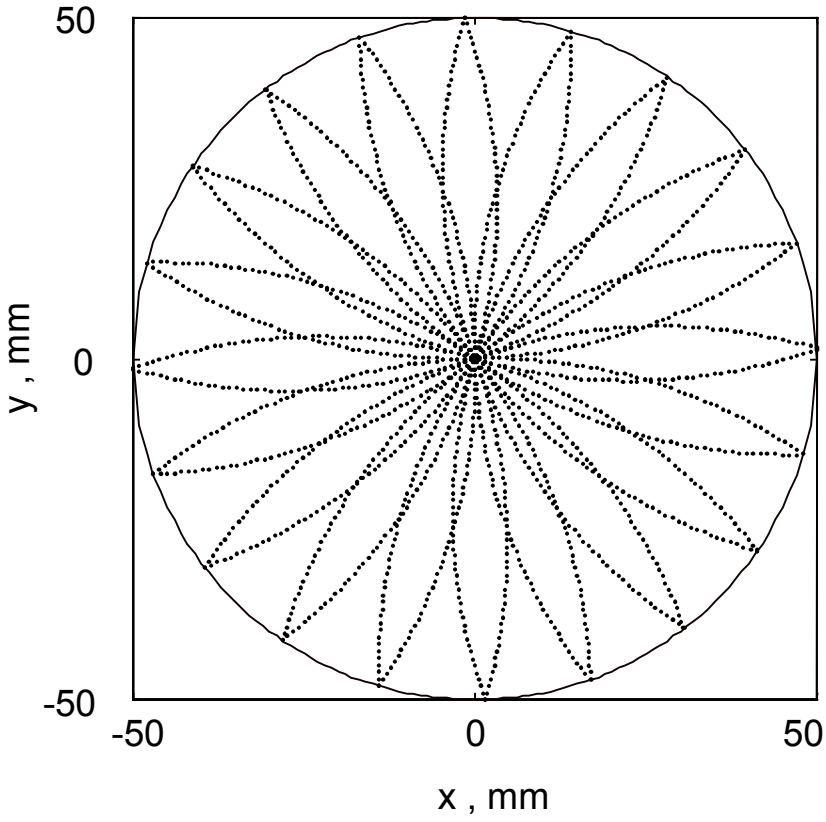


Figure 6.5 The simulated locus for the reflectance sensor across the wafer at the condition $\omega_w = 1.05\omega_p$ and $r_s = r_{cc}$.

from the pooled dataset. The more loci employed, the larger, more uniform samples can be taken. Therefore, a higher level of inference can be achieved. However, there is a possibility that the surface condition may change significantly during the long sampling period needed for multiple scans. This may affect the reliability of the inference and will delay decision making and feedback control. The moving average method is used to estimate the average reflectance on the surface to eliminate this problem. The sensor scans the wafer surface once per platen revolution. If the reflectance sampled at the *j-th* point along the locus at the *i-th* time period, each time period is equal to the duration for one revolution of the platen, denoted as x_{ij}. If total n points are taken along each locus, the mean reflectance along the locus at the *i-th* period, $\bar{x}_i$, is:

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \tag{6.19}$$

Suppose the number of loci to cover the entire wafer surface or a area of interest is w, the moving average of the sampling reflectance at the *i-th* period, M_i, is defined as:

$$M_i = \frac{\bar{x}_i + \bar{x}_{i-1} + \cdots + \bar{x}_{i-w+1}}{w} \tag{6.20}$$

That is, at the *i-th* time period, the observations from the newest one scan and the previous *(w-1)* scans are employed to estimate the mean reflectance of the entire wafer or the surface of interest. Thus, the surface condition as inferred from the reflectance measurements can be updated every scan. For example, the sensor takes about 10 scans to cover the wafer at the condition of $\omega_w = 1.05\omega_p$. If the platen runs at 75 rpm, it takes 8 seconds to scan over the entire wafer surface, in which the locus rotates 180° relative to the wafer, and 16 seconds to rotate back to the first locus. The moving average can capture the change of surface reflectance due to both the change of surface topography and the change of Cu area fraction within a short period, in this case in less than one second. However, it may still smooth over the rapid change due to the partial oxide exposure on small portions of the wafer surface near the onset of endpoint by averaging the current data with the previous data (which is taken in 8 seconds in the example).

The total variance of the surface reflectance at *i-th* time period, S_i^2, can be estimated based on the same pooled dataset employed in the moving average:

$$S_i^2 = \frac{\sum_{i-w+1}^{i}\sum_{j=1}^{n}(x_{ij} - M_i)^2}{N-1} \tag{6.21}$$

where N is the total number of samples in the moving average subset ($N = wn$). The total variance is calculated based on the deviation of the reflectance at each sampling point relative to the total estimated mean of the surface of interest, as estimated by the moving average. In addition to the total variance, the variance along each locus, the range of data, and their maximum and minimum must be tracked to identify the rapid change of surface reflectance at the moment when the barrier or oxide layer is exposed. The variance can determine the percent of overpolished area on the wafer surface at the end of the process. Additionally, the distribution of the data can determine the regime of polishing. For example, the skewness of the data distribution in polishing can be compared with the theoretical value at endpoint, which can be estimated based on the given pattern layout and sensor kinematics. The definition of skewness β can be found in many statistics texts. It may be defined as (Sachs, 1982):

$$\beta = \frac{3(\bar{x} - \tilde{x})}{S} \tag{6.22}$$

where $\bar{x}$ is the mean, $\tilde{x}$ the median and S the sample standard deviation of the selected dataset, which can be estimated from one locus or multiple loci, and calculated from Eqs. (6.19), (6.20), and (6.21). These statistics can apply to the die-level estimation of surface condition. For instance, data taken within a specific range of radius (an annulus) can be combined. The same statistical methods can estimate the surface reflectance over the specific area. The effectiveness of each of these methods on endpoint detection will be examined in Section 6.5.3.

6.3 Experimental

An optical sensor unit (Philtec D64) comprising of light-emitting diodes (LEDs), bundled glass fibers for light transmission and receiving, and an amplifier detected the conditions of the wafer surface based on surface reflectance. Table 6.1 lists the specifications of the sensor. As shown in Fig. 6.6, the wavelength of the LED light source ranges from 775 nm to 990 nm, with a mean around 880 nm and standard deviation about 60 nm. At the sensor tip, the uncollimated light rays diverge outward from the transmit fibers, and only the reflected light within the area with the same diameter, about 1.6 mm, of fiber bundle is received. The spot size was chosen so that it is small enough to detect different surface conditions on different patterns (sub-die areas) on the wafer. However, it is larger than the individual line or feature to even out the small variation of reflectance due to local (sub-device level) randomness of material removal. Because of the divergence of the light beam, the sensor is sensitive to the gap between the tip and the targeted surface. Figure 6.7 shows the characteristic of the sensor output (reflectance) on a copper mirror surface corresponding to the gap distance. In practice, the sensor was operated at a distance of around 5 mm so that the sensor response is less sensitive to the slight change of gap distance during polishing or the surface waviness of the wafer.

The sensor unit was installed on the platen base with the tip embedded inside a holder through the platen. A translucent plastic window (Rodel JR111) allowed the sensor view the wafer surface through the polyurethane polishing pad stacked on the platen. The material of the window has similar wear properties as those of the pad so that the surface of the window remained at the same level of the rest of the pad surface and did not affect the sensor measurement or polishing uniformity. The sensor was linked to a power supply and a data acquisition system via a rotary coupling. The output signal was amplified before the coupling to enhance the signal-to-noise ratio. Additionally, an off-line set-up measured the surface reflectance of the polished wafer. Two rotary stages with angle readings mimicked the kinematics due to the rotation motion of the wafer carrier and the platen. The position of the sensor on the wafer was determined from the angles of both the rotation of the wafer and the sensor arm and the distance between the two centers of the rotary stages. Comparing the

Table 6.1: Specifications of the reflectance sensor.

Item	Specification
Light Source	High Intensity LED
Wavelength (nm)	780-990 (μ=880, σ=50)
Spot Diameter (mm)	1.6
Light Beam Spread (°)	30
Operation Distance (mm)	0-6.35
Stability (%)	<0.1% full scale
Frequency Response (kHz)	< 20

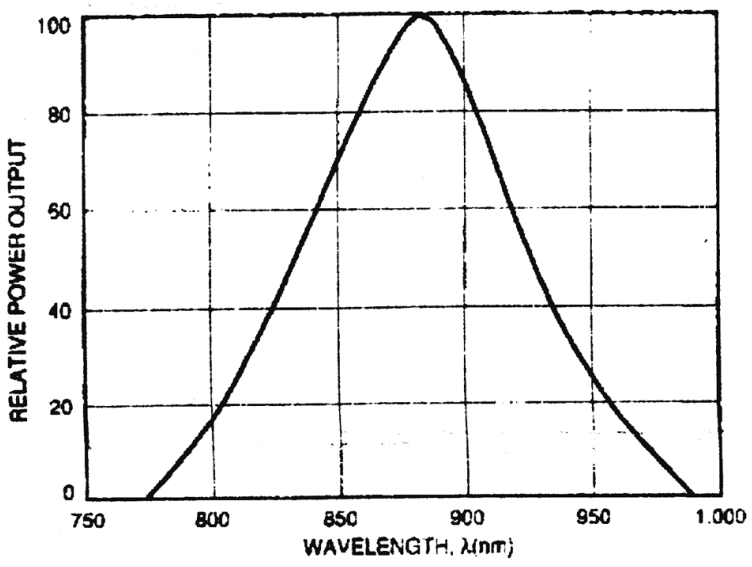


Figure 6.6 The spectrum of the LED light source of the reflectance sensing system.

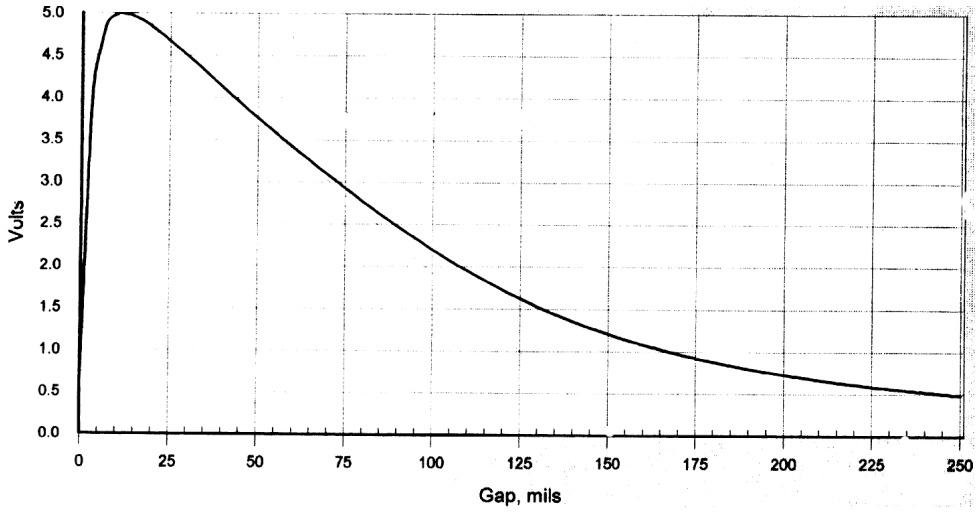


Figure 6.7 The effect of the gap distance between the wafer surface and the fiber optics tip on the sensor tip.

measurements from this set-up with those from *in situ* sensing identified the effects of slurry and wafer slippage on the reflectance sensing.

Both blanket and patterned Cu wafers were employed for experiments to verify the capability of the sensor and to determine the detection schemes. The blanket Cu wafer was composed of a 20 nm TaN barrier layer followed by a 1 μm thick PVD Cu coating on a Si substrate. For the patterned wafer, a tested damascene structure was employed, which was comprised of an array of line-spacing structures with different linewidths and pitches. A detailed layout of the pattern can be found in Chapter 4. This pattern is transferred into a 1.5 μm thick TEOS coating with trenches etched to a depth of 1 μm on a 100 mm silicon substrate. A 20 nm Ta layer followed by a 1 μm thick PVD Cu layer was deposited on the top of the patterned oxide surface. Table 6.2 lists the experimental conditions.

6.4 Results

This section examines the experimental results of blanket and patterned Cu wafers to investigate the characteristics of the reflectance sensing technique. The reflectance of a planar Cu area measured *in situ* in polishing may deviate from its theoretical value due to surface roughness, slurry particles, variation of gap between the wafer and the sensor in polishing, and random noise. Variation of surface reflectance due to these effects is studied based on the measurements in blanket wafer polishing. Additionally, the surface reflection in patterned wafer polishing is affected by the surface topography in the planarization regime and by the area fraction in the polishing regime, which significantly contributes to the variation of measurements. Both off-line and *in situ* measurements were conducted to study the effects of pattern geometry and Cu area fraction on the reflectance. These results are compared with the reflectance from light scattering theory with the assumptions of single wavelength, plane incident wave, and periodic surface structure. The characteristics of reflectance across a desired area during polishing are examined to correlate the measurements with different regimes of Cu CMP. These will help establish different schemes for *in situ* sensing and endpoint detection.

Table 6.2: Experimental conditions.

Experimental Parameters	Experimental Conditions
Diameter of Wafer (mm)	100
Normal Load (N)	391
Normal Pressure (kPa)	48
Rotational Speed (rpm)	75
Linear Velocity (m/s)	0.70
Duration (min)	1 - 6
Sliding Distance (m)	42 - 252
Slurry Flow Rate (ml/min)	150
Abrasive	α-Al_2O_3
Abrasive Size (nm)	300
pH	7

6.4.1 Tests on the Blanket Wafers. Figure 6.8 shows a typical result of surface reflectance on a blanket Cu wafer during polishing. To elucidate the effects of slurry and scratching, the normalized mean reflectance is defined as the average reflectance over ten passes across the wafer divided by the reflectance on a scratch-free Cu wafer under the same pressure condition (at the same gap between the wafer surface and the sensor). At the initial stage, the reflectance was about 30% less than that without slurry. The light scattering from slurry particles caused the reduction. Because the sensor was operated in the range within which it is less sensitive to the change of gap distance, the decrease of reflectance was mainly due to the particle scattering. The normalized mean reflectance gradually dropped by 0.1 to about 0.6 after two minutes of polishing and the standard deviation increased slightly to about 0.02 from an initial small value. These observations suggest that the surface was roughened due to abrasion. After two minutes, the variance of the surface reflectance increased without significant change of the mean. Inspection of wafer surface at this stage indicated that a small portion of the Cu was cleared and the less reflective TaN was exposed. Since the majority of the surface was still covered with Cu, the mean did not drop significantly. The mean then started to drop and the variance kept increasing with the increase of Cu clearing at about two minutes and 30 seconds. The standard deviation started to decreasing and the mean gradually reached a lower level until all the Cu was cleared, about six minutes. The harder TaN barrier acted as a polishing stop and retained a low level of variance of surface reflectance after all the Cu was removed. After one more minutes of overpolishing, the TaN was polished through and the normalized mean reflectance decreased further to about 0.2 with the exposure of Si substrate.

6.4.2 Off-Line Measurements on Patterned Wafers. Figures 6.9 and 6.10 show the effects of surface topography on reflectance. These data were obtained off-line on patterns at the center die with various linewidths between 0.5 and 100 μm and constant Cu area fractions of 0.5 and 0.01, respectively. The normalized reflectance is defined as the ratio of the measured reflectance on each sub-die to the reflectance of the unpolished blanket Cu surface. Figure 6.11 shows the corresponding evolution of step-heights (= $2\Delta h$) for these damascene structures (sub-dies). Lower nominal pressure (28 kPa) and relative velocity (0.46 m/s) were applied than those of the industrial practice. By six minutes, most of the high features were removed and the surface was planarized. A thin, smooth Cu layer was left on the ILD

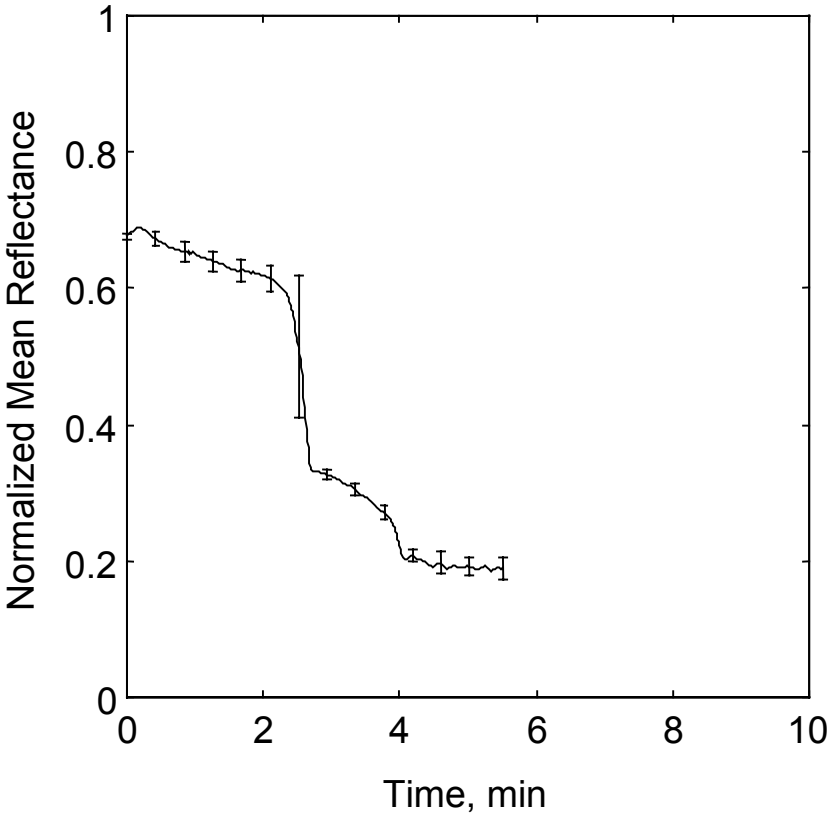


Figure 6.8 The evolution of surface reflectance for the blanket wafer polishing.

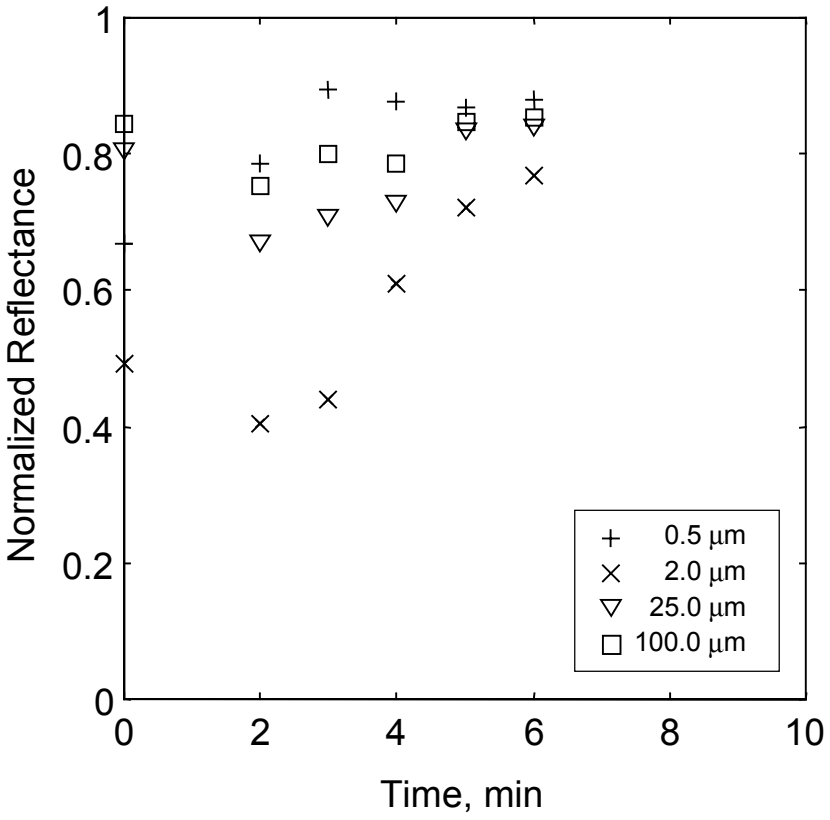


Figure 6.9 The results of off-line measurements at the Cu planarization regimes on the pattern with 0.5 area fraction ($w/\lambda = 0.5$).

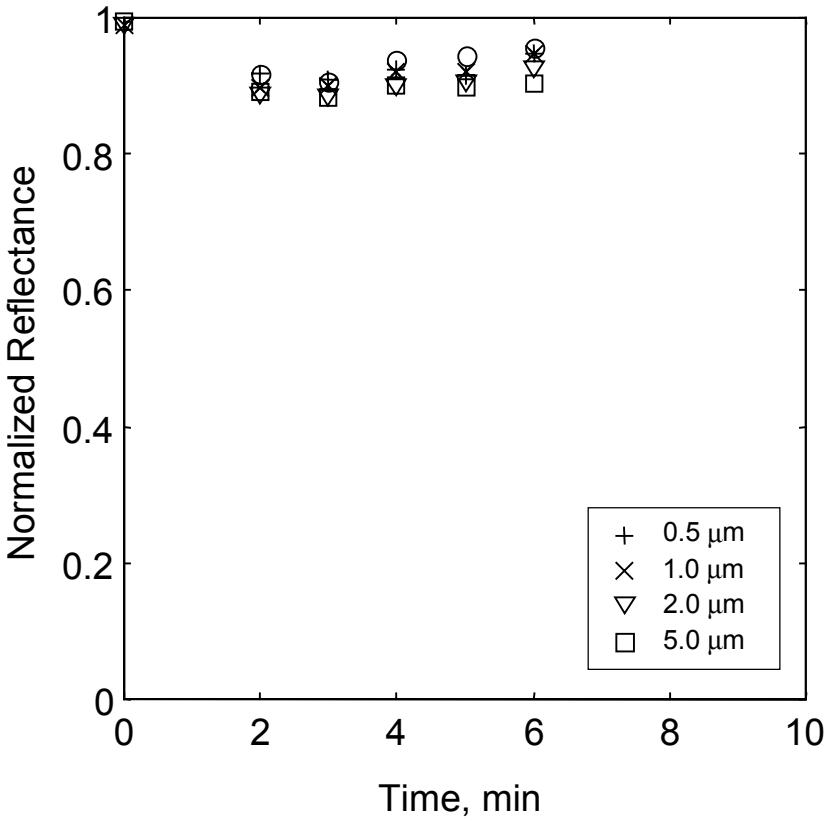


Figure 6.10 The results of off-line measurements at the Cu planarization regime on the patterns with 0.01 area fraction ($w/\lambda = 0.01$).

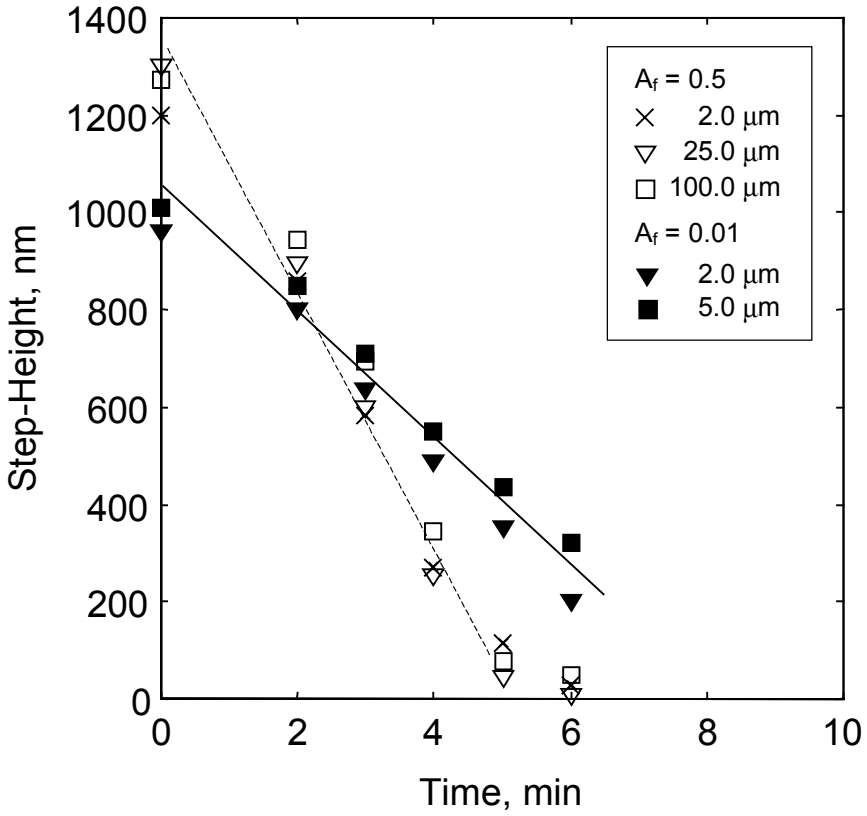


Figure 6.11 Time evolution of step-heights for patterns with constant area fractions 0.5 and 0.01.

without being polished through. For patterns of 0.5 Cu area fraction, the initial variance of the reflectance resulted from the variation of step-height and pitch on the surfaces of the different sub-dies. The reflectance is mainly affected by the pitch (or linewidth) of the pattern because the initial step-heights are close for the patterns with linewidth 2, 25 and 100 μm, except that of the 0.5 μm structures. The smaller the pitch, the more light scatters on the surface and reduces reflectance. This can be explained by the less reflective Cu surfaces on low features due to the roughness from the deposition process. After two minutes of polishing, the normalized reflectance decreased about 0.1. The surface roughness increased by particle abrasion to contribute to the overall reduction of the surface reflectance. The reflectance of the 0.5 μm line area, however, increased because the surface was mostly planarized before two minutes.

The reflectance of each pattern increased gradually after the initial drop and then finally reached a steady value due to the planarization of high features. This trend was prdicted by Eq. (6.9) in Section 6.2.1. It can be numerically shown that the reflectance coefficients in Eq. (6.9) at higher scattering mode ($m >> 1$) increase rapidly when $\Delta h/\lambda$ ratio increases (or the s value increases). The detail scattering patterns calculated by this equation for $2\pi\Delta h/\lambda = 0.03$, 0.1, 0.3, 1, 3, and 10 at various incident angles of light can be found in the literature (Beckmann, 1963). Diffusion scattering takes place when the $\Delta h/\lambda$ ratio increases and less reflecting light is received by the sensor. Light is more likely to scatter into the direction of specular reflection to be received by the adjacent fibers when the step-height, $2\Delta h$, decreases. Therefore, the surface reflectance increases. As shown in Figs. 6.9 and 6.11, the step-heights for various features were less than 100 nm after polishing for 5 minutes, and the normalized surface reflectance for various features keeps at a similar steady level, about 0.85, on the tested wafers. This implies that the optical sensing technique is less sensitive to the small surface topography. The reflectance for the patterns of 0.01 Cu area fraction also dropped to about 0.1 due to the increase of surface roughness. It then remained at the same level of 0.9 until the surface was planarized. Since the area fractions of low features of these patterns are small, the surface reflectance is not significantly affected by the evolution of the pattern topography. The measurements are similar to those on a blanket Cu surface.

Figures 6.12 and 6.13 show the trend of surface reflectance of various patterns, with 0.5 and 0.01 Cu area fractions, in different process regimes -- planarization, polishing and overpolishing. Figures 6.14 and 6.15 show the corresponding evolution of dishing. The pressure and the velocity applied were close to the industrial practice of 48 kPa and 0.79 m/s. In planarization regime, the surface topography of Cu layer was smoothed down quickly due to the contact between the high features and the pad. Experiments showed that most of the patterns on test wafers were planarized after one minute of polishing and the normalized reflectance reached a similar level, about 0.9, for all patterns measured. Between one and three minutes, the planar Cu layer was removed as that in blanket Cu polishing. In this polishing regime, the normalized reflectance stayed the same, constant about 0.9, independent of original pattern geometry. After about three minutes, the reflectance dropped significantly and sharply because the Cu layer had been polished through and the less reflective underlying oxide was partially exposed. The sub-die areas with higher Cu area fraction may have been polished through faster because the planarization rate is dependent on the pattern geometry. Figures. 6.12 and 6.13 show that the sub-die with high area fraction of 0.5 was polished through first and the Ta barrier exposed after about two minutes. Concurrently, the reflectance started to drop to about 0.8 when the Ta started to be exposed and then further down to about 0.5 when the oxide surface was exposed at three minutes. Nevertheless, all tested patterns seemed to reach the onset of oxide exposure between two and three minutes.

Oxide overpolishing occurred after the onset of oxide exposure. The reflectance kept decreasing until all the excess Cu and barrier (Ta) materials were removed (i.e., process endpoint), after about four minutes of polishing. After the endpoint, the reflectance remained constant, regardless of the slight nonuniformity of topography due to dishing of the soft Cu lines and rounding and overpolishing of the adjacent oxide regions. This again agrees with the earlier results in that the employed sensing technique is insensitive to the small variation of the step-height. Hence, the variation of the reflectance in this regime was mainly due to the different area fraction of Cu interconnects. The areas with higher Cu area fraction are more reflective. However, the experimental values were lower than the theoretical prediction of reflectance for all patterns, especially for those with high area fractions. Equation (6.10) predicts that the (normalized) reflectance is about 0.62 and 0.24 for the patterns with area fractions of 0.5 and 0.01, respectively, in which the R_{TEOS}/R_{Cu} ratio of 0.23 is assumed based

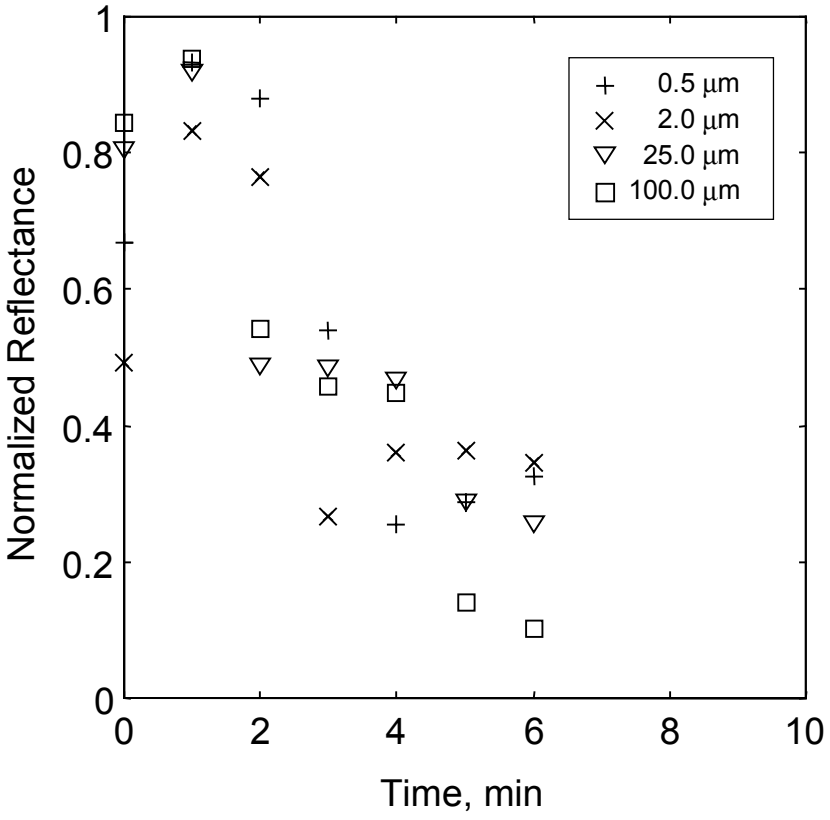


Figure 6.12 The results of off-line measurements at various process regimes on the pattern with 0.5 area fraction.

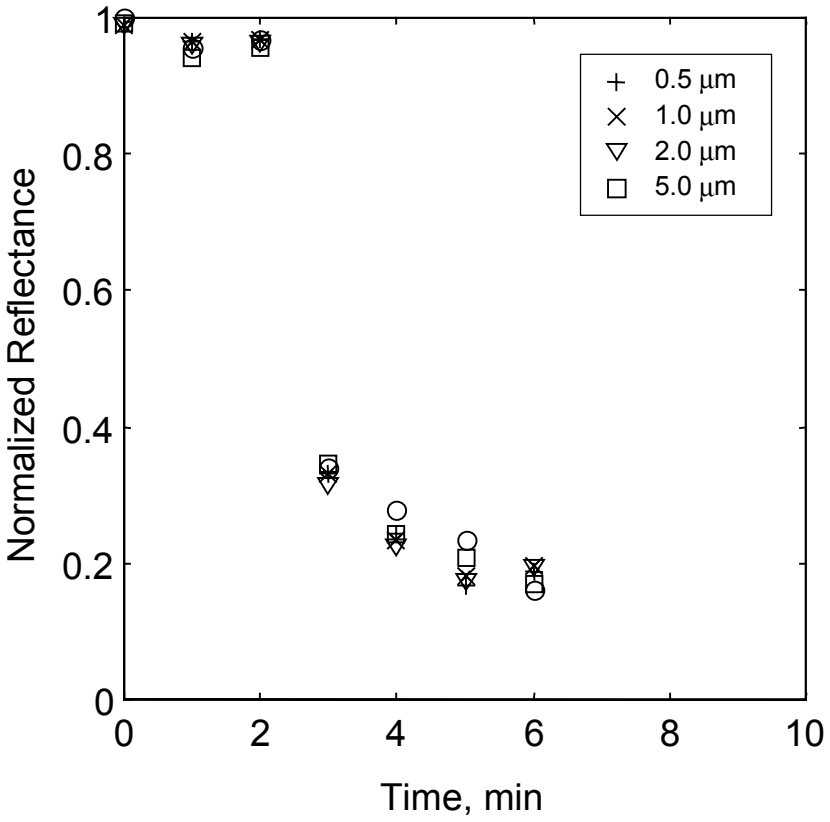


Figure 6.13 The results of off-line measurements at various process regimes on the patterns with 0.01 area fraction.

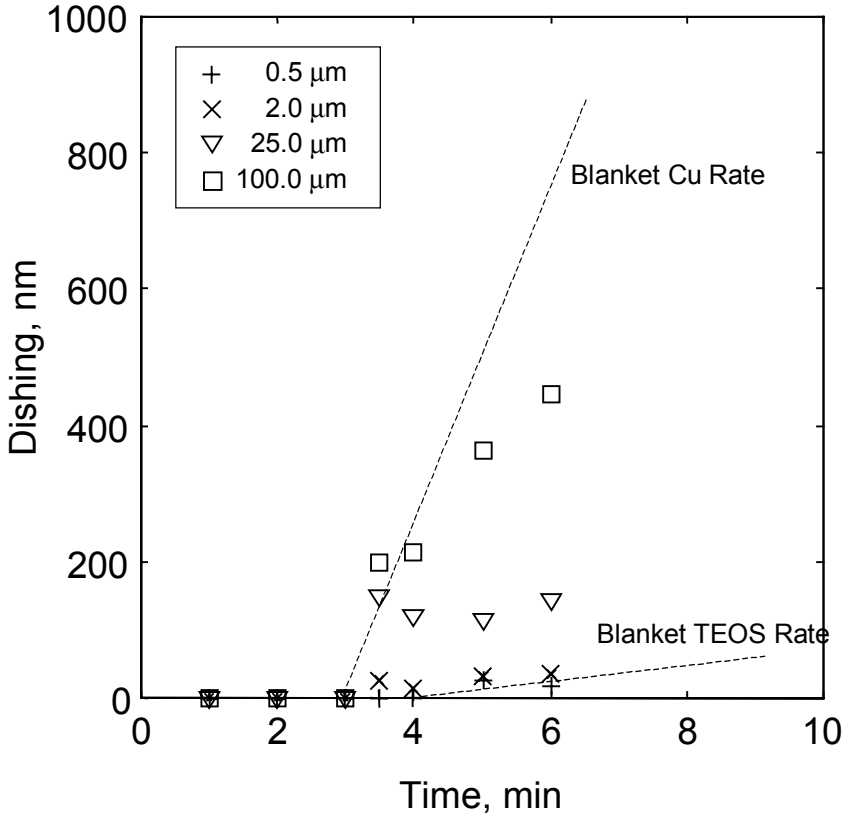


Figure 6.14 Time evolution of Cu dishing for patterns with constant area fraction 0.5 and various linewidths.

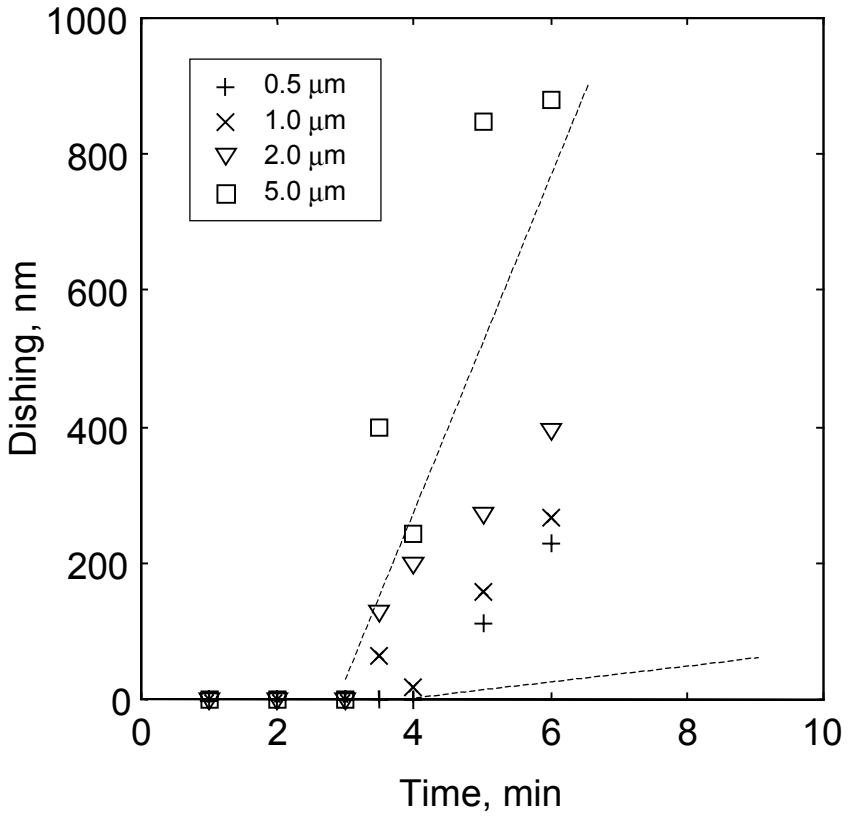


Figure 6.15 Time evolution of Cu dishing for patterns with constant area fraction 0.01 and various linewidths.

on the experimental measurement on blanket films. In reality, the light transmitted through the oxide and reflected from the underlying Si substrate may be blocked by the Cu lines, which decreases the intensity of reflected light from the oxide surface and reduces overall reflectance of the sub-die. Additionally, scratches and less reflective Cu oxides (due to corrosion) were found on the surfaces of Cu lines. These also reduced surface reflectance, especially for the pattern with more Cu area fraction.

6.4.3 Off-Line Measurements Along the Sensor Loci. Figure 6.16 plots the off-line measurements along different sensor loci in terms of the mean value and the standard deviation. The wafer employed is the one shown in Section 6.4.2, polished for four minutes at normal conditions, at which time the majority of dies were polished to the endpoint and some may have been overpolished slightly. The loci follow the sensor trajectories in polishing at the conditions of $\omega_w = \omega_p$ and $r_s = r_{cc}$, in which the sensor travels along an arc of a radius r_{cc}. Loci across different radial directions were employed to elucidate the effects of different loci on the statistics of the surface reflectance of the patterned wafer. The mean and the variance of reflectance data across wafer varied with the orientation of the locus. The mean value varied from 0.24 to 0.26 among the selected loci, compared with the average reflectance about 0.25 of the center die. The standard deviation varied between 1 and 1.2, compared with 1.8 in the center die. The variations of the mean and standard deviation mainly resulted from the different sensor loci due to the non-axial-symmetric pattern layout and from the within-wafer nonuniform polishing. It is not uncommon that the within-wafer nonuniform polishing often exhibits an axial symmetry, such as the "bull's eye effect" (Stine, 1997). Therefore, the variations of reflectance between loci due to wafer-level nonuniformity may be comparable to that of the pattern layout.

Figure 6.17 shows the mean and standard deviation of the reflectance of the center die and across the wafer measured off-line at different polishing stages. The effect of different loci is minimized by combining data from several loci, for instance from five loci evenly spaced across the wafer in this case. The effect of within-wafer nonuniform polishing on the variation of surface reflectance can be determined by comparing the difference between those two data sets. Before polishing, the mean reflectance across the wafer is higher than that of the center die because of the nonuniform coating from the Cu PVD process. The step-heights

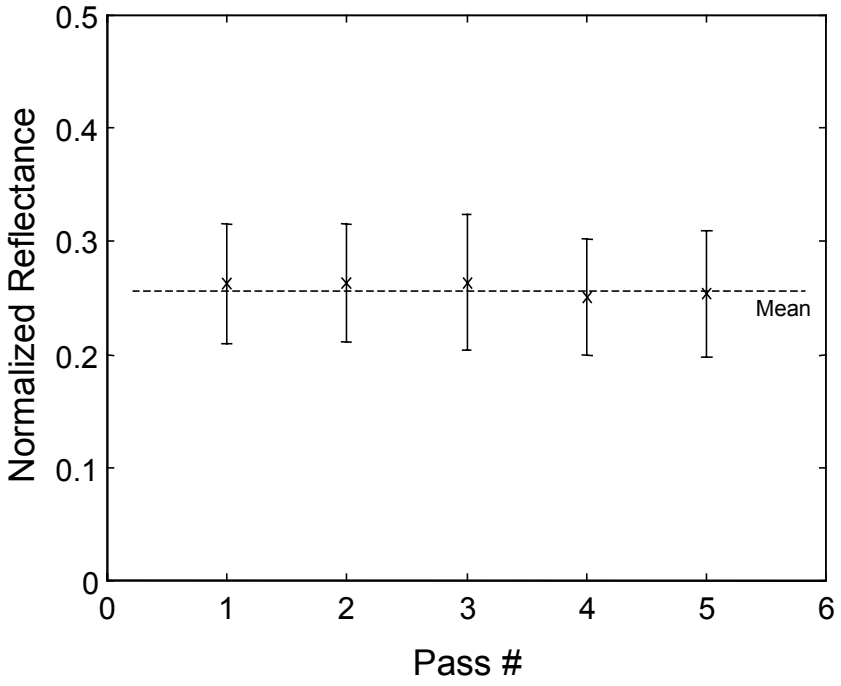


Figure 6.16 Off-line measurements of the mean and standard deviation of surface reflectance along different loci across the wafer at the onset of endpoint.

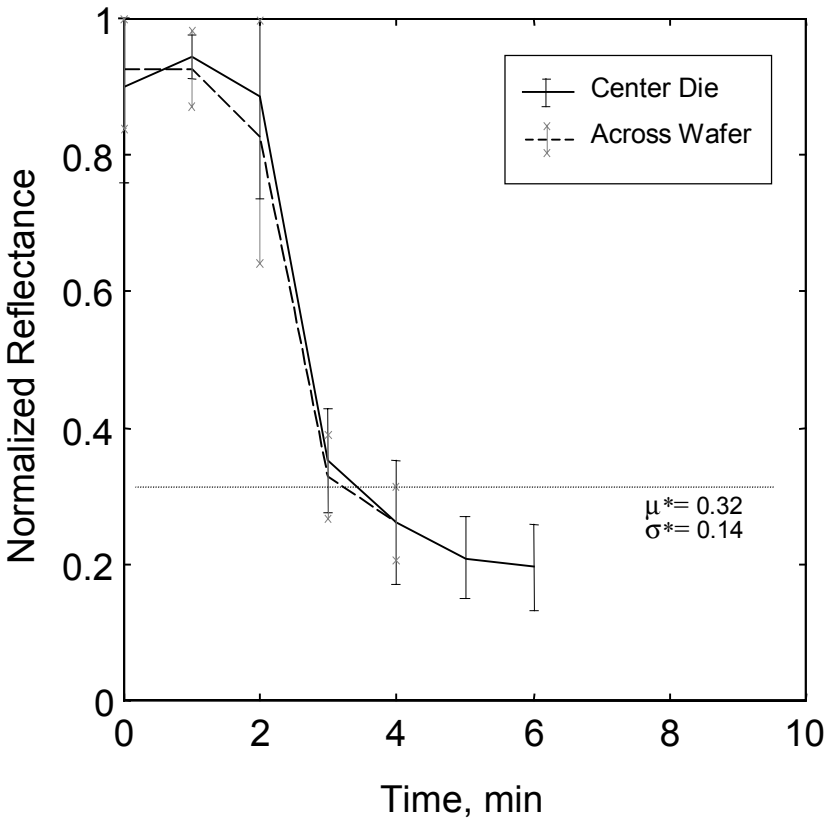


Figure 6.17 Comparison of the off-line measurements (mean and standard deviation) on the center die and across wafer at various polishing stages. The across-wafer data is calculated based on the measurements along five loci.

of patterns are smaller at the edge dies and thus the average reflectance on the edge dies will be higher than that of the center die. Thus the overall mean reflectance is smaller than that of the center die. Similarly, the standard deviation of the edge dies is generally smaller because trenches are shallower due to the nonuniform Cu deposition. After polishing for a short duration, the overall mean became less than the average reflectance of the center die. This is because the polishing rate at the edge was faster than at the center and the less reflective barrier and/or oxide layers were exposed at the wafer edge. The standard deviation of the reflectance across the wafer was also greater than that of the center with the increase of surface nonuniformity. More barrier and oxide layers were exposed and progressed from the edge toward the center with the increase of time. With the increase of wafer-level nonuniformity, the difference between the two means and the standard deviations increased continuously. When the majority of the dies reach the end-point, the mean surface reflectances across the wafer and at the center return to similar levels because the hard oxide layers retains surface uniformity even with a slight overpolishing. The small dishing will not affect the reflectance significantly. The variance of the reflectance of the center die of the four-minute sample is greater due to the remaining patches of Cu/barrier materials. In practice, the overall mean and variance of the reflectance may be compared with those on different surface areas (die-level zones) to determine the process endpoint.

6.4.4 *In Situ* Measurements on Patterned Wafers. Figure 6.18 is an example of *in situ* measurement on a patterned Cu wafer. The y-axis represents the normalized reflectance, which is defined as the reflectance measured divided by the reflectance on blanket Cu wafers before polishing. In the experiments, the angular velocity of the wafer was offset from the angular velocity of the platen by 5% ($\omega_w = 1.05\ \omega_p$) so that the loci covered the entire wafer surface. Figure 6.19 shows the moving average of the reflectance for ten passes and the standard deviation based on the pooled data from those passes. Compared with the off-line measurements, the reflectance measured in polishing was lower because of light scattering by the slurry. It dropped approximately 20% to 25% in the planarization regime, but less significantly in the overpolishing regime. The mean decreased slightly right after polishing because of surface roughening. It then started to increase until reaching a constant level around one minute after the surface had been planarized, as discussed in earlier. After two minutes, the mean dropped again because of the removal of Cu on the surface. Because the

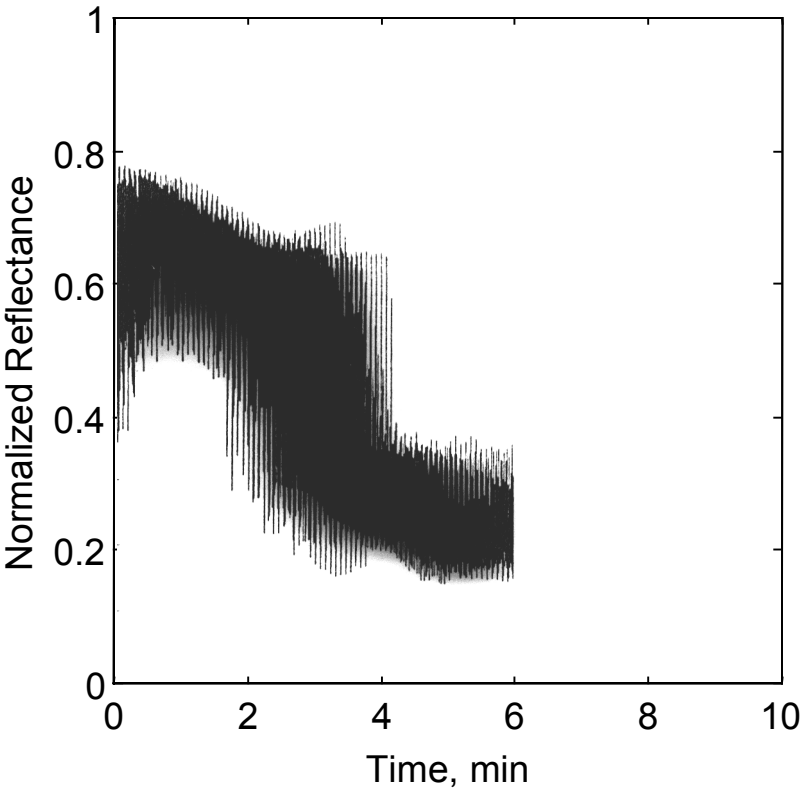


Figure 6.18 Raw data from the in-situ reflectance measurement.

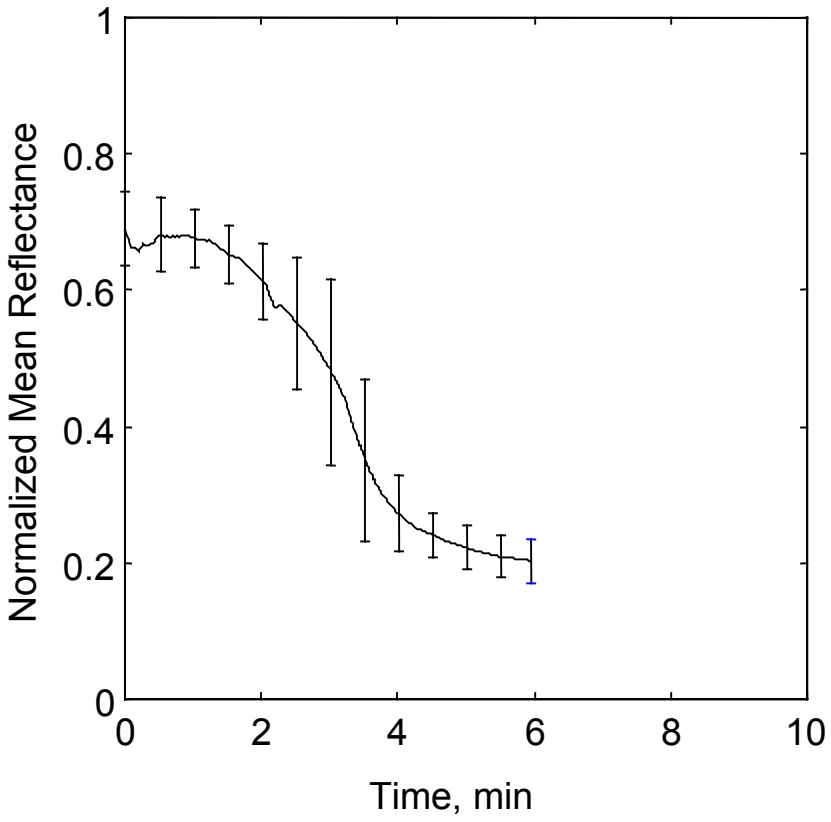


Figure 6.19 Results of in-situ measurements of the moving average and standard deviation of wafer-level surface reflectance.

Cu was removed nonuniformly due to the initial pattern layout and the variation of the coating thickness, the underlying oxide was gradually exposed on the surface. The mean dropped less steeply compared with data on a specific die, such as the center die in the earlier example. The onset of wafer-level endpoint was about four minutes in this experiment. The mean kept decreasing, but at a slower rate, after the endpoint with the gradual increase of surface roughness due to overpolishing and dishing. Due to the effect of slurry and the variation of wafer-to-wafer uniformity, the mean at the onset of endpoint varies in each run. Besides, the mean does not drop rapidly at the onset of endpoint because the Cu is removed nonuniformly in different sub-dies and not all the dies reach the endpoint at the same time. A clear sign of change in mean is lacked for endpoint indication. The mean serves only as a rough indication of the onset of process endpoint.

Figure 6.20 plots the standard deviation of the pooled data in the moving sampling set over 10 passes versus time. Because the variation of the reflectance is mostly due to the pattern geometry and Cu area fraction, the distribution is generally not normal. Figure 6.21 shows the distribution of the normalized reflectance in terms of relative frequency. The initial shape of the distribution represents the initial surface topography of the wafer. There were two peaks of the standard deviation during polishing. The first peak occurred at the beginning of the process corresponding to the minimum mean reflectance in the Cu planarization regime. This resulted from the initial surface topography and surface roughening. The standard deviation in the planarization regime reached a minimum when the majority of the pattern had been smoothed down and the mean reached a maximum. The surface condition at this stage is similar to that of a blanket wafer. The variation of the surface reflectance is affected by the surface roughness, slurry scattering, and random error of measurement and thus represents a normal distribution in Figs. 6.21 (b) and 6.21 (c). The maximum reflectance occurred in the middle of Cu clearing regime, at about one minutes of polishing in this case (Fig. 6.21 (b)). The low end of reflectance distribution, between 0.4 to 0.5, in Fig. 6.21 (c) indicates the presence of Ta barrier layer. After three minutes of polishing, a broad distribution with two peaks is observed in Fig. 6.21 (d). The subgroup of surface reflectance centered at a lower value represents the sub-die area on which the oxide is exposed. The other subgroup, with the mean close to the rough blanket surface, indicates that

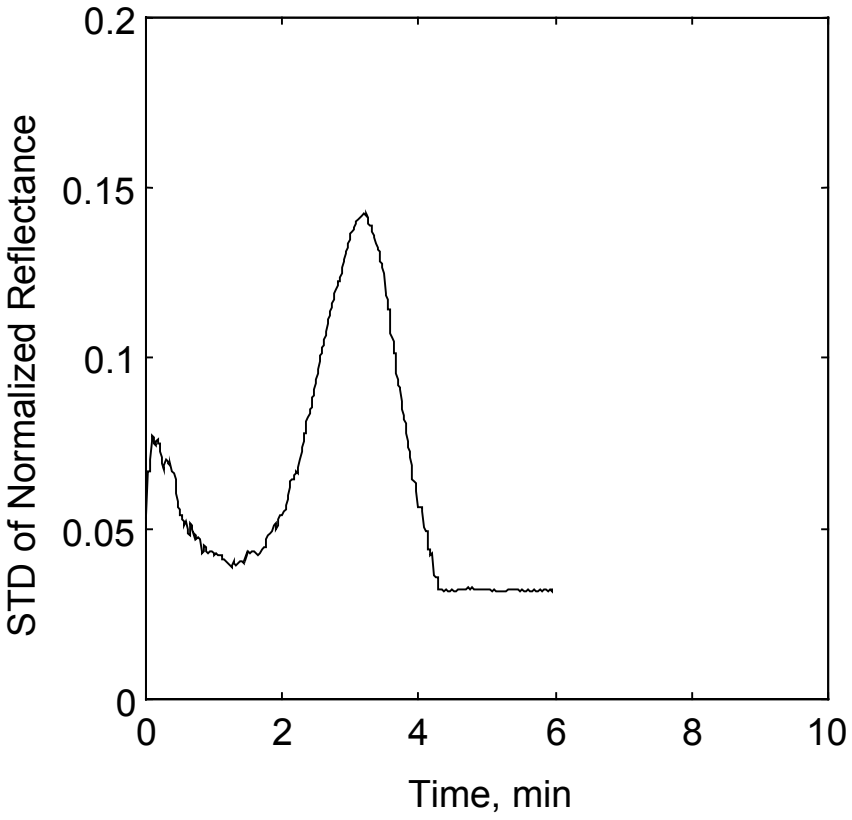


Figure 6.20 Results of in-situ measurements of the standard deviation of wafer-level surface reflectance.

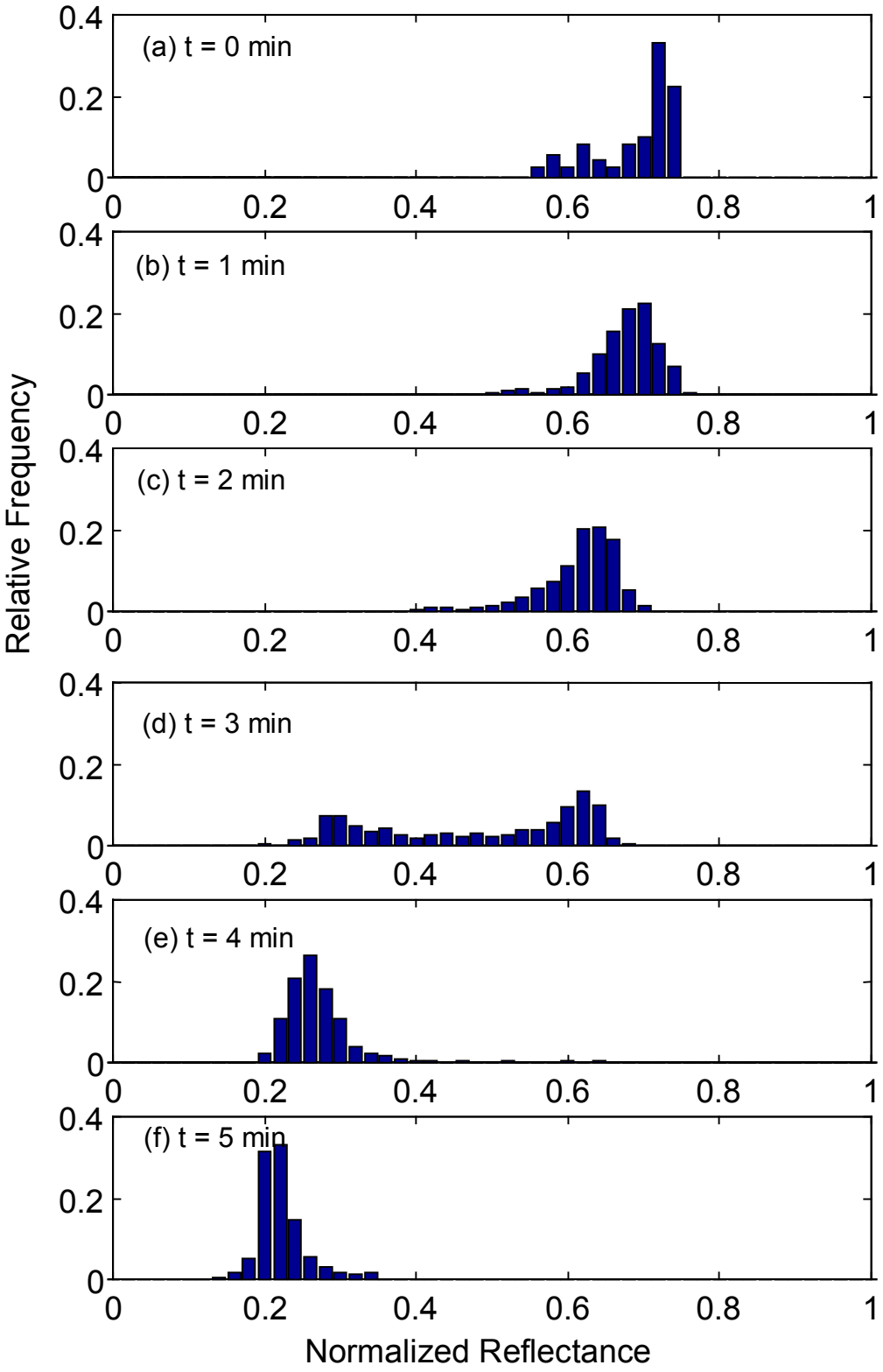


Figure 6.21 Distribution of surface reflectance versus polishing time from the in-situ measurements.

the high reflective Cu and/or Ta barrier layer still partially covers the surface. After the maximum, the standard deviation decreased quickly with the increase of area of oxide exposure. At the onset of endpoint, the standard deviation reached a sharp turning point and then remained at a low constant level. As observed in the previous off-line measurement, the variation of the surface reflectance reached a minimum when the highly reflective Cu was cleared. However, because the resolution of the sensor is limited by the spot size, it may not effectively detect small patches of metal on the surface. In practice, a short period of overpolishing may be necessary to ensure that all Cu and barrier material is removed. After the endpoint, the standard deviation is determined by the designed pattern layout (local Cu area fraction as shown in Eq. (6.10)) which affects the skewness of the distribution, as shown in Figs. 6.21 (e) and 6.21 (f). It is shown that all the Cu on the ILD has been removed after four minutes of polishing by the fact that the height-reflectance peak has vanished. The high end of reflectance distribution, around 0.4, indicates that some Ta was left on part of the surface. Figure 6.21 (f) shows that the variation of surface reflectance will not change significantly with the small variation of surface topography from overpolishing and dishing.

6.5 Discussion

6.5.1 Locus Design and Sampling Plan. The sampling scheme relies greatly on the sensor loci and sampling frequency for reliable spatial and temporal information on surface reflectance. At the die-level, many loci must be taken on the die of interest to detect the variation of reflectance due to the nonuniform topography, Cu area fraction, and the non-symmetric layout. Based on the kinematics, the sensor locus is determined by the parameters of ω_w, ω_p, r_s, and r_{cc}. For some conditions, such as the example in Fig. 6.5 with $\omega_w = 1.05\ \omega_p$ and $r_s = r_{cc}$, the sensor can cover the center die many times, but may pass over the edge die once or never. One way to improve the sampling density on the edge die is to increase the number of loci on the wafer by reducing the difference between ω_w and ω_p. However, this will increase the time period to scan over the entire wafer surface and thus may delay the detection of rapid changes of reflectance of a local area. The wafer slippage, both rotational and translational inside the recess, will also make the control of velocity offset within a small

range very difficult. The smallest offset of the wafer and the platen velocity is about 3 to 5%, typically.

On the other hand, the distance between centers of the wafer and the platen r_{cc} may be changed during the polishing. This "sweeping motion" may help cover a desired region on the wafer surface. Figure 6.22 shows an example at $r_s = 1.25\, r_{cc}$ with $\omega_w = 1.05\omega_p$ and $\dot{r}_{cc} = 0$, in which only the outer area is sampled. Compared with the high sampling density at the center in Fig. 6.5, the sampling density is much higher and now uniform around the edge. In practice, the entire wafer may be scanned first to roughly determine the overall surface condition. The area at a particular radius of interest can be scanned with a higher sampling density for a better inference of the local condition. Moreover, two or more sensors can be installed at different radii r_s and different angles (phase) on the same platen. The combined loci will give a higher and more uniform distributed sampling density of both the center and the edge region.

Another important parameter for designing the sampling plan is the sampling frequency. At least one measurement must be taken from each subdie along the sensor locus to detect the variation of reflectance between the different sub-die areas and different dies. It is preferable to have one or more replicants on each pattern to reduce the error due to random variation in measurement. For the 100 mm patterned wafer, about 40 subdies are located along a locus (10 dies along a locus with four subdies across each die diagonal). With at least one replication on each subdie area, about 100 points are required in the test, which corresponding to a 100 Hz sampling rate at the typical wafer rotational speed of 60 rpm. Thus, the sampling rate should be greater than 100 Hz, and more replicants can be taken to even out the effect of random error.

6.5.2 Variance Components of the Surface Reflectance. The surface reflectance of a patterned wafer varies with the surface roughness, pattern topography and area fraction, and optical properties of the coatings. Due to the within-wafer nonuniform material removal, the surface topography and the remaining fraction of Cu during polishing may vary among different dies across the wafer. The within-wafer nonuniform polishing usually results from certain systematic sources, such as nonuniform velocity distribution, pressure distribution,

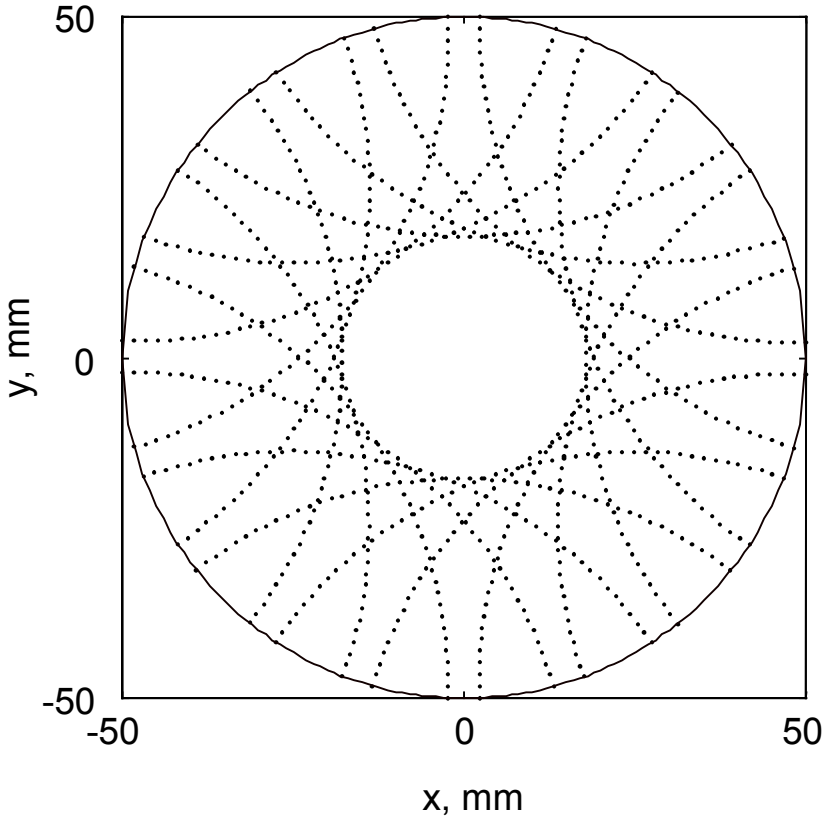


Figure 6.22 The simulated loci for the reflectance sensor across the wafer at the condition $\omega_w = 1.05\omega_p$ and $r_s = 1.25r_{cc}$.

interfacial temperature distribution, slurry flow, and contact conditions (Stine, 1998). Its effect on polishing always follows a systematic pattern which tends to be repeatable between wafers of the same lot. The wafer-level nonuniformity affects the pattern evolution on the same die with a similar trend. The relative rates of material removal between different patterns on a die will remain similar to another die at different location because the factors that affect wafer-level nonuniformity will have less interaction with the die- or device-level polishing behavior. For instance, die-level polishing is mostly affected by the pattern geometry, such as linewidth and area fraction. Therefore, the variation of reflectance measurements on a die tends to follow the same distribution and is nested within the die. Based on this assumption, a two-level nesting variance structure (Drain, 1997) is employed to decompose the effects of within-wafer and die-level nonuniform polishing. Assuming that the reflectance at each level is normally distributed, the reflectance at location j of die i on the wafer, R_{ij}, can be written as:

$$R_{ij} = \mu + W_i + D_{j(i)} \tag{6.23}$$

where μ is the average reflectance within the wafer from multiple loci, W_i the die-to die (or within-wafer) effect on die i, and $D_{j(i)}$ the within-die effect at location j on die i. The total, within-wafer and within-die variances of surface reflectance are expressed as σ_T^2, σ_W^2, σ_D^2 respectively. Additionally, the within-die effect, $D_{j(i)}$, is assumed to be normal and the two-level variance components are assumed to be independent to each other. Therefore, the total variance of reflectance, σ_T^2, can be written as:

$$\sigma_T^2 = \sigma_W^2 + \sigma_D^2 \tag{6.24}$$

Appendix 6A gives the detailed procedures to estimate the variance components. The results of decomposition of estimated variance components, S_W^2 and S_D^2 with respect to the *in situ* measured data are plotted in Fig. 6.23. Table 6.3 lists the value of each component and the F ratio, defined as S_W^2/S_D^2, for every 30 seconds to examine the significance of within-wafer nonuniformity on the variation of surface reflectance. Additionally, the polishing results for all dies at the same radius are assumed to be similar and are combined into a subset

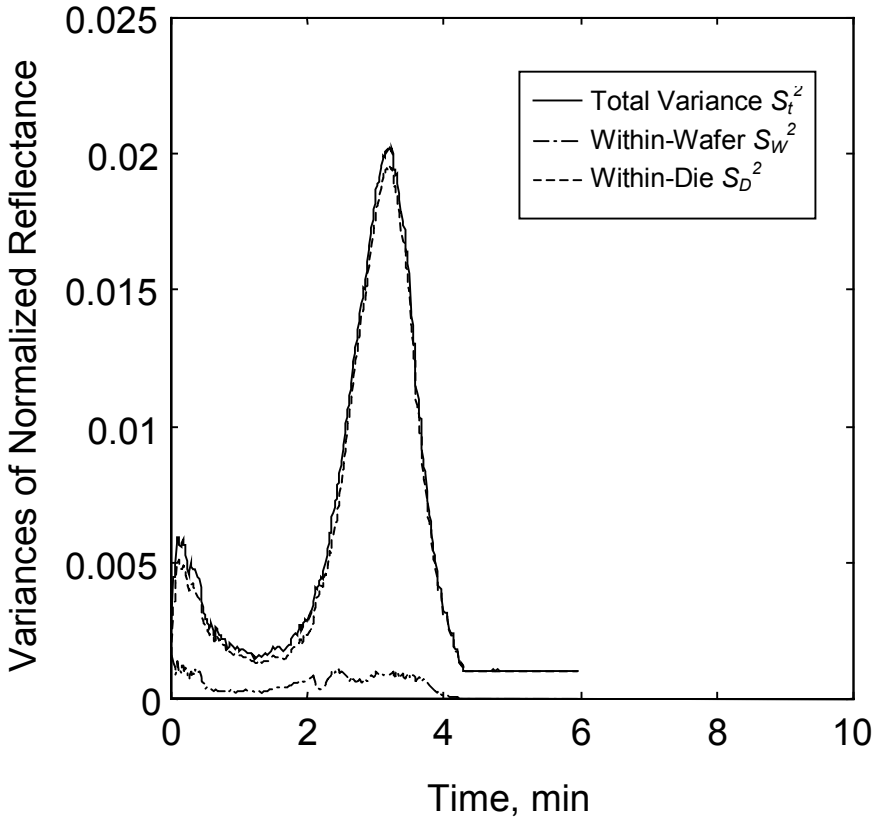


Figure 6.23 Decomposition of the within-wafer and within-die variance for the in-situ measurements.

Table 6.3: Analysis of variance of two-level nested model for surface reflectance.

Time (Minutes)	Within-Wafer Variance, S_W^2	Within-Die Variance, S_D^2	F Ratio (S_W^2/S_D^2)	Pr(F)
0	15.94×10^{-4}	1.64×10^{-3}	0.965	0.59
0.5	3.89	2.62	0.149	0.07
1.0	2.62	1.58	0.166	0.08
1.5	3.88	1.54	0.252	0.14
2.0	7.49	2.51	0.299	0.17
2.5	9.30	8.45	0.110	0.05
3.0	9.22	18.11	0.051	0.02
3.5	7.24	13.67	0.053	0.02
4.0	1.39	3.08	0.045	0.01
4.5	0.15	1.01	0.015	~0
5.0	0.01	1.04	0.001	~0

for estimation of the die-level variation. The high F ratio on the wafer before polishing indicates that the within-die means at different radii are different and the probability of mean difference between dies, Pr(F) (which implies the existence of within-wafer nonuniformity), is about 0.6. This is due to the variation of initial step-height from the deposition process. The within-wafer nonuniformity decreases after polishing starts, and remains at a low level with respect to the total variation. The confidence level (probability) of the hypothesis that there is a mean difference between the dies, is less than 20%. This suggests that the surface is planarized uniformly (or topography becomes more uniform across the wafer) through the CMP process. The within-wafer variance and the F ratio even drop to very low levels, (Pr(F) ~ 0), after the wafer-level endpoint is reached. This is because the underlying oxide surface is harder than Cu and can retain the surface planarity and the wafer-level polishing uniformity. On the other hand, the within-die effect contributes significantly to the total variance of surface reflectance throughout the process. The process endpoint can be determined based on the change of within-die variance component as a result of the drastic change of Cu area fraction. In practice, the total variance might be employed to approximate the within-die variance to determine the process endpoint. The small effect of within-wafer nonuniformity will not affect the accuracy of detection.

Moreover, within-wafer variance is just an indication of the nonuniform reflectance of the surface. It may not directly correlate with the uniformity of the remaining Cu thickness. However, it directly represents the uniformity of surface condition. This information can be used to monitor the across-wafer surface condition and uniformity. It may also be employed in a feedback control loop to adjust the process parameters, such as pressure distribution and velocities of wafer carrier and platen, to improve the uniformity of polishing.

6.5.3 Endpoint Detection Algorithms. In the previous sections, the characteristics of surface reflectance at endpoint and other stages of Cu polishing, in terms of moving average, distribution, and the variation of the reflectance across the wafer, were discussed. These characteristics can be employed to design the endpoint detection algorithm(s). The moving average can be used to detect the moment that the surface reflectance drops under a certain threshold, as shown in Fig 6.19. The threshold is determined by the average area fraction of Cu and the optical properties of surface materials with respect to the wavelength(s) employed.

Because of the random effect of slurry scattering, surface roughness, and random error, the threshold usually will deviate from the theoretical mean reflectance presented in Section 6.2.1 and so must be determined based on observations from a few preliminary tests. Moreover, the sampled reflectance corresponding to the "true" wafer-level endpoint will fall into a statistical distribution related to the variation in initial coating uniformity, the variation of process parameters, and the random error from sampling and sensing. Accordingly, a hypothesis test must be conducted to ensure that the moving average M falls within a given interval with respect to an acceptable confidence level. Because the true variance of the surface reflectance is unknown, the 100(1-α) confidence interval is determined using the appropriate Student t sampling distribution for the sample standard deviation S (Montgomery, 1996):

$$\left(M - t_{\alpha/2,N-1} \cdot \frac{S}{\sqrt{N}}\right) \le \mu \le \left(M + t_{\alpha/2,N-1} \cdot \frac{S}{\sqrt{N}}\right) \tag{6.20}$$

Figure 6.24 shows the results of the moving average of the surface reflectance versus time with an estimated interval at 99.5% confidence level ($\alpha = 0.005$). Since the sample size N is very large, the estimated true mean is confined to a small interval. Moreover, the threshold may also have its underlying distribution from the historical data. It may be ambiguous sometimes to determine the endpoint from the overlapping of the two confidence intervals. The threshold also varies with different chip layout and design. It may be time-consuming to develop a new endpoint detection recipe for every change or new chip design.

Compared with the moving average, the variance (or standard deviation) of surface reflectance provides a more robust means to detect the endpoint. Figure 6.20 shows a clear change in variance at the onset of endpoint. The endpoint can be determined based on both the slope and the threshold level of the variance curve. Because of the high reflectance difference between Cu and oxide, the change of variance with time is usually drastic right before the endpoint for any chip design. The variance of surface also remains at a low level after the endpoint because the hard oxide retains the surface uniformity. Similarly, the variance can be estimated from the measurements based on a desired confidence interval.

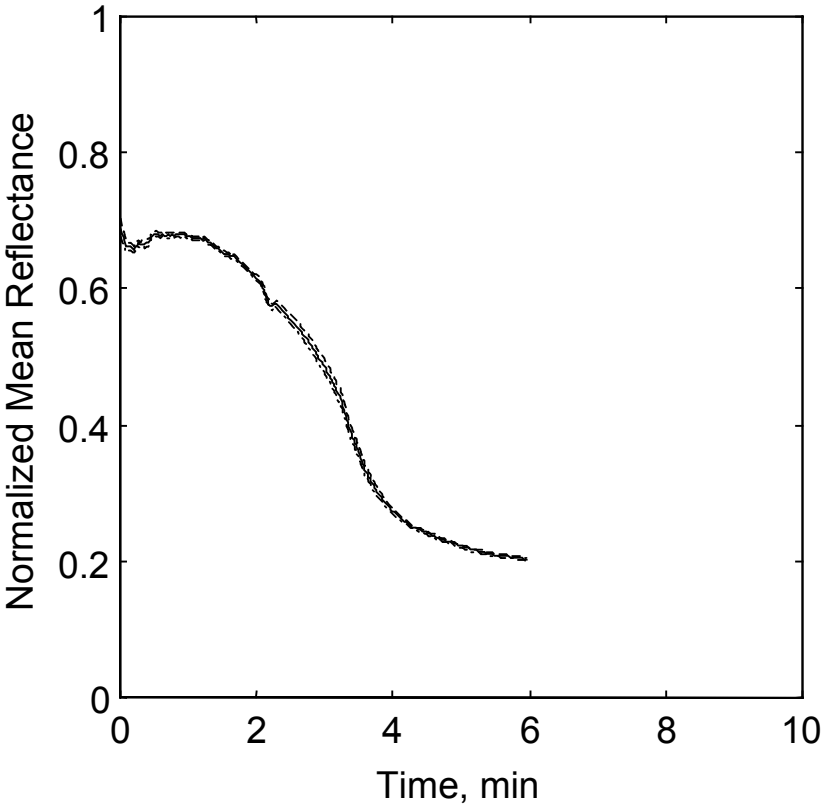


Figure 6.24 Results of the sampled moving average versus time with estimated interval at 99.5% confidence interval.

Without knowing the true variance of the surface reflectance σ^2, the variance interval with $100(1-\alpha)$ confidence level is given based on the Chi-square (χ^2) distribution:

$$\frac{(N-1)S^2}{\chi^2_{\alpha/2,N-1}} \le \sigma^2 \le \frac{(N-1)S^2}{\chi^2_{1-\alpha/2,N-1}} \tag{6.21}$$

The estimated variance does not vary significantly within a short period of overpolishing. The threshold of variance will also remain approximately a constant between runs for a given pattern design. Therefore, the endpoint is more easily determined based on the variance information than from the mean (moving average). In practice, the ratio of standard deviation to mean reflectance can be employed to incorporate the characteristics of mean and variance of reflectance for endpoint detection, as shown in Fig. 6.25. The endpoint is indicated as a local minimum and can be determined without the complexity of calculating the slope and the confidence intervals.

In addition to the wafer-level endpoint, the onset of endpoint on the dies can also be determined based on mapping of sampling loci onto the wafer surface. The surface conditions on different zones, such as "rings" at different radii, can be determined based on the same techniques used in the wafer-level endpoint detection. The sampling loci can be designed as described in the earlier section to select the sensing area and resolution. Moreover, the mean, variance, and distribution of the surface reflectance also provide information for different stages in the polishing process. The variance and the variance-to-mean ratio reach a minimum, and the distribution becomes normal when the Cu pattern is planarized. Figure 6.26 shows that the range of the reflectance increases drastically when the underlying oxide starts to expose. The variance to mean ratio reaches a maximum when the majority of the excess Cu on the surface is cleared. This information can be integrated as part of the *in situ* sensing technique to determine progress of the CMP process. For multi-step polishing processes, this information can also be used to determine the endpoints of each step and increase the capability of process control.

An experiment was conducted to validate the effectiveness of various endpoint detection schemes with the same process condition listed in Table 6.2. Polishing was stopped as soon

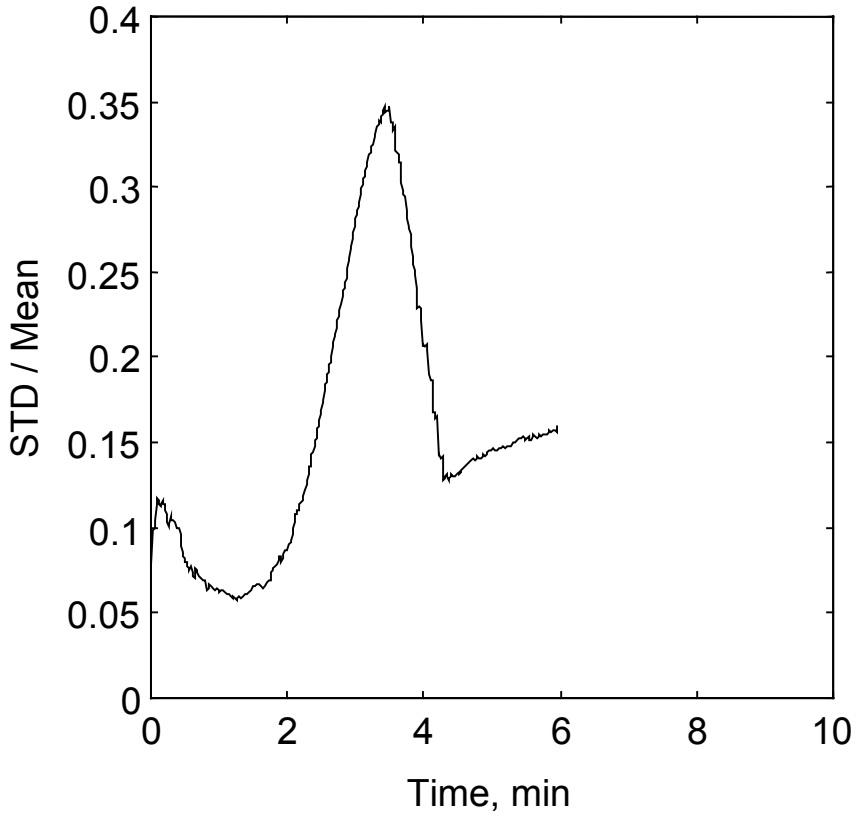


Figure 6.25 Results of in-situ measurements of the ratio of the standard deviation to the mean reflectance (wafer-level).

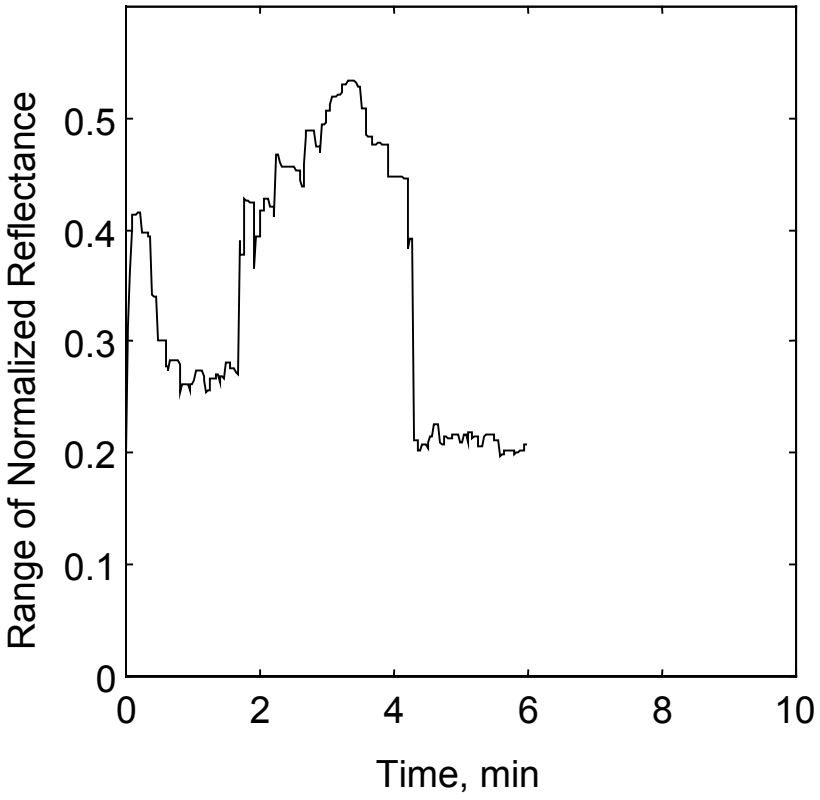


Figure 6.26 Results of the rang of surface reflectance versus polishing time (wafer-level).

as the standard deviation, the standard deviation-to-mean ratio, and the range indicated the onset of (wafer-level) endpoint, as shown in Fig. 6.27. Figure 6.28 shows the picture of the wafer surface, which agrees with the result from the sensing system. The Cu is cleared from the surface. Although it is hard to identify from observation, an ultra-thin Ta barrier, which is more transparent to the light than the thick layer, may still remain on the surface and may not be detected by the optical sensor. In practice, a short period of overpolishing may be applied after the sensor detects the endpoint to ensure that all the metals are completely removed. A robust process recipe as described in Chapter 5, which uses a slurry with high selectivity, can also retard oxide overpolishing in this period.

6.6 Conclusions

A novel *in situ* reflectance sensing technique detects the polishing uniformity and endpoint of the Cu CMP process. An analytical model of the effects of surface topography and the area fraction of Cu on the reflectance of a patterned surface was developed with the assumptions of a monochromatic light source and plane incident light wave. The model is limited to features with linewidths much larger than the wavelength of light to avoid the complexity of diffraction. In practice, however, the size of the features of the state-of-art chip design at the first few levels of interconnects may be comparable or even less than the wavelengths of applicable light sources. A broadband optical sensor (uncollimated) correlated the measured reflectance with the geometry of surface pattern. The reflectance follows the trend predicted by the model with respect to the change of the surface topography and the Cu area fraction in polishing. When the step-height of the pattern is less than 50 nm, however, the reflectance is less sensitive to the evolution of surface topography. The surface waviness remains at a low level for a short period of overpolishing. Thus the reflectance sensor only detects the change of Cu area fraction at the onset of endpoint. Experiments on both blanket and patterned Cu wafers studied the effects of light scattering result due to surface roughness and the slurry particles on surface reflectance. A better decision for determining the endpoint can be made based on the understanding of these systematic sources of variation.

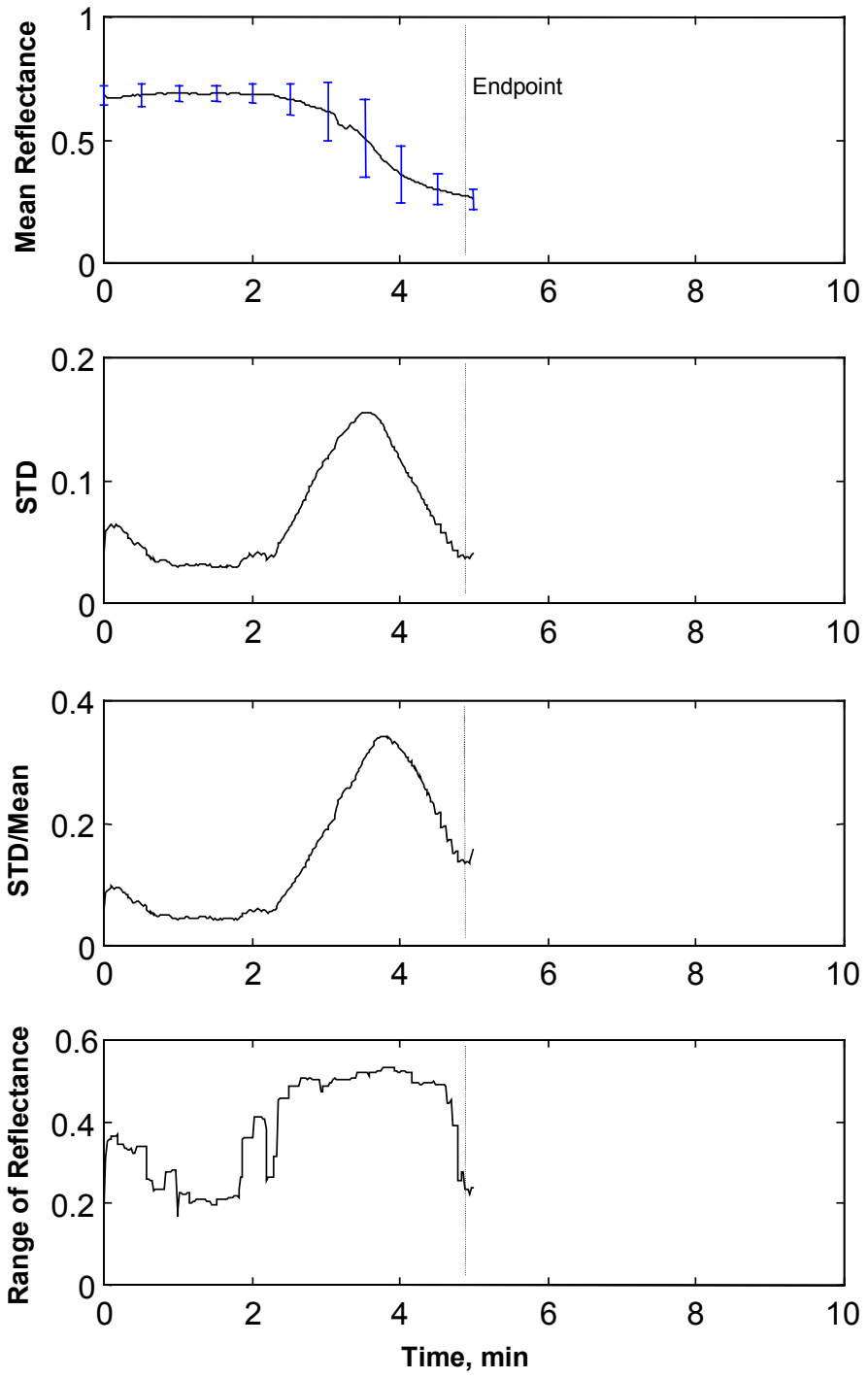
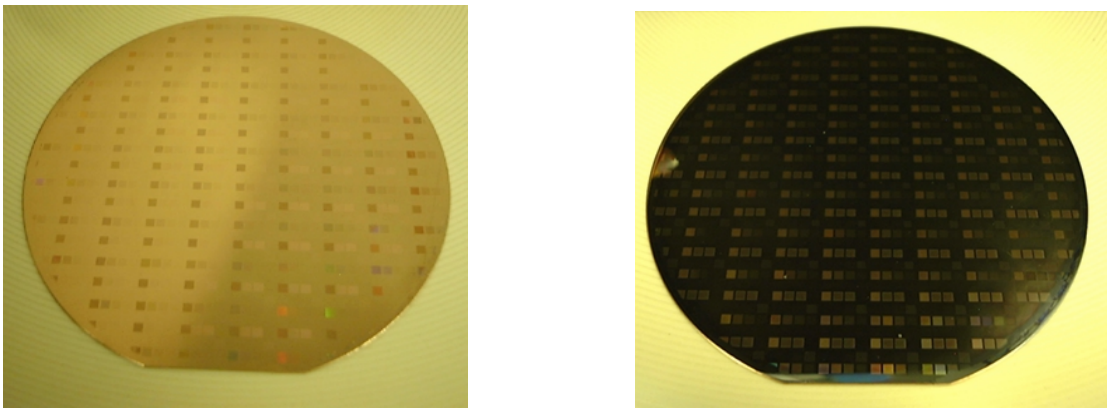


Figure 6.27 Experimental validation for various in-situ sensing and endpoint detection schemes.



(a) (b)

Figure 6.28 Pictures of patterned wafer surface (a) before polishing, and (b) at the onset of process endpoint (wafer-level) determined by the proposed detection schemes.

The kinematics of the sensor was studied and its loci on the wafer were designed to cover the desired sampling area. Based on mapping the reflectance onto the wafers, a more local means of detecting the surface condition in polishing was developed. Compared with other global detection techniques, this method provides better spatial resolution and reliability in detecting the endpoint. Additionally, information regarding the within-wafer nonuniform polishing was determined based on localized measurements and statistical analyses. This information may serve as feedback to control process parameters in real time to improve polishing uniformity.

Moreover, metrics for detecting process endpoint, such as the mean, the variation, the range and the distribution of the surface reflectance, were developed and their effectiveness examined by experiments. The ratio σ/μ, the standard deviation divided by the mean reflectance, can effectively determine both the global and the local endpoints. The ratio reaches a minimum when the endpoint is achieved no matter what the pattern design is. The characteristics of various schemes with respect to different stages of polishing were also discussed. These will provide additional information for process monitoring and control for improving the quality and performance of the Cu CMP process.

Nomenclature

A_f = area fraction of metal pattern
C_1, C_2 = dimensionless integration constants
E_1, E_2 = electric fields of the incident and scattered light (V)
f = frequency of the incident light (Hz)
h = thickness of the coating (m)
Δh = half step-height of the pattern (m)
k_1, k_2 = wave numbers of the incident and scattered light waves (m^{-1})
L = length of the illuminated surface area(m)
M_i = moving average of the surface reflectance at the *i-th* period
m = mode of scattered light wave
N = number of samples including in the moving average
$\mathbf{n}$ = unit vector normal to the wafer surface
n = number of samples taken along a sensor locus
p = $(L/\lambda)(\sin\theta_1 - \sin\theta_2)$
R, R_{Cu}, R_{Oxide} = reflectances of the pattern, blanket Cu and blanket oxide surfaces
r, θ = polar coordinates
r_{cc}, r_s = distances between the centers of the wafer and the pad, and the centers of the sensor and the pad (m)
S_i = variance of the surface reflectance across the wafer or the area of interest at the *i-th* period
t = duration of polishing (s)
v_R = magnitude of the relative velocity (m/s)
w = number of loci employed in the moving average to cover the wafer surface
X, Y, Z = stationary Cartesian coordinates (m)
x, y, z = rotating Cartesian coordinates (m)
x_{ij} = surface reflectance sampled at the *j-th* point along the *i-th* locus
$\bar{x}_i$ = mean reflectance along the *i-th* locus
α = confidence coefficient
β = skewness of the distribution of surface reflectance
ϕ = scattering coefficient
γ = reflection coefficient
Λ = pitch of the pattern structure (m)
λ = wavelength of the incident light (m)
θ_1, θ_2 = incident and scattering angles
σ_T, σ_W, σ_D = total, within-wafer and within-die standard deviations of the surface reflectance
ω = angular frequency of the incident light wave (rad/s)
ω_p, ω_w = angular velocities of the pad and the wafer (rad/s)
ψ = auxiliary function

References

Bahar, E., 1981, "Scattering Cross Sections for Composite Random Surfaces: Full Wave Analysis," *Radio Sci.*, Vol. 16, pp. 1327-1335.

Bakin, D.V., Glen, D.E., and Sun, M.H., 1998, "Application of Backside Fiber-Optic System for *In Situ* CMP Endpoint Detection on Shallow Trench Isolation Wafers," *Proc. of SPIE*, Vol. 3507, pp. 210-207.

Banet, M.J., Fuchs, M., Rogers, J.A., Reinold, J.H., Knecht, J.M., Rothschild, M., Logan, R., Maznev, A.A., and Nelson, K.A., 1998, "High-Precision Film Thickness Determination Using a Laser-Based Ultrasonic Technique," *Appl. Phys. Lett.*, Vol. 73, pp. 169-171.

Beckage, P.J., Lukner, R., Cho, W., Edwards, K., Jester, M., and Shaw, S, 1999, "Improved Metal CMP Endpoint Control by Monitoring Carrier Speed Controller Output or Pad Temperature," *Proc. of SPIE*, Vol. 3882, pp. 118-125.

Beckmann, P. and Spizzichino, A., 1963, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon Press, Oxford.

Bibby, T. and Holland, K., 1998, "Endpoint Detection in CMP," *J. Electronic Materials*, Vol. 27, pp. 1073-1081.

Bibby, T., Adams, J.A., and Holland, K., 1999, "Optical Endpoint Detection for Chemical Mechanical Planarization," *J. Vac. Sci. Technol. B*, Vol. 17, pp. 2378-2384.

Brekhovskikh, L. M., 1952, "The Diffraction of Waves by a Rough Surface, Parts I and II," (in Russian) *Zh. Eksper. i Teor. Fiz.*, Vol. 23, pp. 275-304.

Chan, D.A., Swedek, B., Wiswesser A., and Birang, M., 1998, "Process Control and Monitoring with Laser Interferometry Based Endpoint Detection in Chemical Mechanical Planarization," *1998 IEEE/SEMI Advanced Semiconductor Mfg. Conf. and Workshop*, pp. 377-384.

Desanto, J.A., 1975, "Scattering from a Perfectly Reflecting Arbitrary Periodic Surface: An Exact Theory," *Radio Sci.*, Vol. 16, pp. 1315-1326.

Desanto, J.A., 1981, "Scattering from a Sinusoid: Derivation of Linear Equations for the Field Amplitudes," *J. Acoustical Soc. Am.*, Vol. 57, pp. 1195-1197.

Drain, D., 1997, *Statistical Methods for Industrial Process Control*, Chapman and Hall, New York.

Eckart, C., 1933, "A general Derivation of the Formula for the Diffraction by a Perfect Grating," *Physical Review*, Vol. 44, pp. 12-14.

Fang, S.J., Barda, A., Janecko, T., Little, W., Outley, D., Hempel, G., Joshi, S., Morrison, B., Shinn, G.B., and Birang, M., 1998, "Control of Dielectric Chemical Mechanical Polishing (CMP) Using an Interferometry Based Endpoint Sensor," *Proc. IEEE 1998 International Interconnect Technology Conf.*, pp. 76-78.

Joffe, M.A., Yeung, H., Fuchs, M., Banet, M.J., and Hymes, S., 1999, "Novel Thin-Film Metrology for CMP Applications," *Proc. 1999 CMP-MIC Conf.*, pp. 73-76.

Leach, M.A., Machesney, B.J., and Nowak, E.J., U.S. Patent #5,213,655, May 25, 1993.

Litvak, H.E. and Tzeng, H.-M., 1996, "Implementing Real-Time Endpoint Control in CMP," *Semiconductor International*, Vol. 19, pp. 259-264.

Montgomery, D.C., 1996, *Introduction to Statistical Quality Control*, 3rd ed., John Wiley & Sons., Inc., New York, pp. 101-111.

Murarka, S., Gutmann, R., Duquette, D., and Steigerwald, J, U.S. Patent #5,637,185, June 10, 1997.

Nam, J., 2000, "Optical Endpoint Detection for the Chemical Mechanical Polishing Process", Master's Thesis, Massachusetts Institute of Technology.

Lord Rayleigh, 1907, "On the Dynamical Theory of Gratings," *Proc. Roy. Soc.*, *A*, Vol. 79, pp. 399-416.

Park, T., Tugbawa, T., Boning, D., Chung, J., Hymes, S., Muralidhar, R., Wilks, B., Smekalin, K., Bersuker, G., 1999, "Electrical Characterization of Copper Chemical Mechanical Polishing," *Proc. 1999 CMP-MIC Conf.*, pp. 184-191.

Rogers, J.A., Fuchs, M., Banet, M.J., Hanselman, J.B., Logan, R., and Nelson, K.A., 1997, "Optical System for Rapid Materials Characterization with Transient Grating Technique: Application to Nondestructive Evaluation of Thin Films Used in Microelectronics," *Appl. Phys. Lett.*, Vol. 71(2), pp. 225-227.

Sachs, L., *Applied Statistics: A Handbook of Techniques*, translated by Reynarowych, Z., Springer-Verlag, New York.

Sandhu, G., Schultz, L., and Doan, T., U.S. Patent #5,036,015, July 30, 1991.

Schultz, L., U.S. Patent #5,081,796, Jan. 21, 1992.

Smith, W.L., Kruse, K., Holland, K., and Harwood, R., 1996, "Film Thickness Measurements for Chemical Mechanical Planarization," *Solid State Technology*., Vol. 39, pp. 77-86.

Steigerwald, J.M., Zirpoli, R., Murarka, S.P., Price, D. and Gutmann, R.J., 1994, "Pattern Geometry Effects in the Chemical-Mechanical Polishing of Inlaid Copper Structures," *J. Electrochem. Soc.*, Vol. 141, pp. 2842-2848.

Stine, B.E., 1997, "A General Methodology for Acessing and Characterizing Variation in Semiconductor Manufacturing", Ph.D. Thesis , Massachusetts Institute of Technology.

Stien, D.J. and Hetherington, D.L., 1999, "Prediction of Tungsten CMP Pad Life Using Blanket Romoval Rate Data and Endpoint Data Obtained from Process Temperature and Carrier Motor Current Measurements," *Proc. of SPIE*, Vol. 3743, pp. 112-119.

Uretsky, J.L., 1965, "The Scattering of Plane Waves from Periodic Surfaces," *Annals of Phys.*, Vol. 33, pp. 400-427.

Zeidler, D., Plotner, M., and Drescher, K., 2000, "Endpoint Detection Method for CMP of Copper," *Microelectronic Engineering*, Vol. 50, pp. 411-416.

Zipin, R.B, 1966, “A Preliminary Investigation of Bidirectional Spectral Reflectance of V-Grooved Surfaces,” *Appl. Optics*, Vol. 5, pp. 1954-1957.

Appendix 6A Decomposition of Variance Components

The reflectance data on the dies at the same radius can be combined to estimate the die-level variance. The wafer-level variation is given by measurements in "annuli" at different radii. The total, with-wafer, and within-die variances are denoted by σ_T^2, σ_W^2, and σ_D^2, respectively. The within-die variance, σ_D^2, is assumed to be normal, though in practice it will contain components both from systematic variation due to the effect of pattern geometry and from process random variation. The above simplification can be justified by the analysis of variance (ANOVA) technique (Drain, 1997).

Assuming that the variances are independent to each other, the total variance, σ_T^2, can be expressed as:

$$\sigma_T^2 = \sigma_W^2 + \sigma_D^2 \tag{6A.1}$$

If the sampling plan is balanced, i.e., same number of measurements are taken at each level on each sample, the variances of each level can be estimated by the following analysis. The variance component for within-die effect, S_D^2, can be estimated by:

$$S_D^2 = \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{(x_{ij} - \bar{x}_i)^2}{n-1} \tag{6A.2}$$

where m is the number of dies (or annuli) sampled, n the sample size at each die (or annulus), $\bar{x}_i$ the mean of reflectance at die i (or annulus i). One the other hand, the within-wafer variance σ_W^2 can be obtained by using the relation between the variance of die means, $\sigma_{\bar{W}}^2$, and the within-die variance:

$$\sigma_{\bar{W}}^2 = \sigma_W^2 + \sigma_D^2 / n \tag{6A.3}$$

The variance of die means $\sigma_{\bar{W}}^2$ can be estimated by $S_{\bar{W}}^2$:

$$S_{\overline{W}}^2 = \sum_{i=1}^{m} \frac{(\overline{x}_i - \hat{x})^2}{m-1} \tag{6A.4}$$

where $\hat{x}$ is the grand mean of the measurements, the moving average of multiple scans over a given span in this case.