In Situ Endpoint Detection by Acoustic Emissions in Chemical–Mechanical Polishing of Metal Overlay

Hong Hocheng and Yun-Liang Huang

Abstract—Chemical–mechanical polishing (CMP) has been recognized indispensable to achieve the global planarity in removal of metal overlay across the wafer, when the number of interconnecting metal layers and the size of wafer increase and the line rule reduces to nano scale. CMP has to stop at the endpoint when the overlay metal has been removed, or dishing will occur which affects the subsequent lithography in IC manufacturing. An effective in situ endpoint monitoring method essentially improves the yield rate and throughput of CMP. One notices the coefficient of friction between the pad and dielectric layer is distinguishably lower than the one between the pad and the copper overlay, based on that an endpoint monitoring method using acoustic emissions during the chemical mechanical polishing is feasible. The proposed method is tested in the experiment, and the comparison with the previous thermal monitoring technique shows consistent results.

Index Terms—Acoustic emission, chemical–mechanical polishing (CMP), copper overlay, endpoint detection, pad, thermal monitoring.

I. INTRODUCTION

CHEMICAL–MECHANICAL polishing (CMP) has emerged recently as an indispensable processing technique for planarization in submicron multilevel very large-scale integration (VLSI). Table I shows the number of interconnecting metal layers and the size of wafer increase constantly. With the shrinking device scale and higher performance requirements of integrated circuits, the utilization of the advanced photolithography is essential. The increasing number of interconnection layers requires a highly planar surface on every layer to ensure that the accumulated nonplanarity of the layers is less than the depth of focus (DOF) of the lithographic stepper. The copper wiring was introduced to the microprocessor units (MPUs) where the performance had been limited by the conductivity of the aluminum alloy. Copper possesses lower resistance and higher lifetime against electro migration. Copper has replaced aluminum as the mainstream metallization for ICs.

In the metal CMP process, the previously deposited metal layer is removed, leaving the metal line and the metal in the contact vias or plugs. Fig. 1 shows a schematic diagram of the planarized layers in integrated circuits.

An evenly polished surface produces the good metal step coverage and provides a field within the lithography depth of focus, that contact vias and metal wires can be well patterned [1]. The accumulated variation in film thickness will impair the process. CMP has become a key process for the Cu interconnections, low-κ materials, and dual damascene process.

During a CMP process, as shown in Fig. 2, the wafer is held by a rotating carrier that brings the wafer against a polishing pad. The pad is mostly in circular motion. A down force to the wafer and a relative velocity between the wafer and the pad will provide the desired removal rate. The slurry made of nano-sized abrasive particles suspended in a chemical solution is dispensed on the pad. In order to ensure the wafers are properly planarized and to control the constant yield of CMP, the online monitoring of the process endpoint is needed.

Since the wafer surface is in full contact with the polishing pad, detecting its endpoint is a challenge. There are patented methods to answer the need for in situ monitoring and endpoint detection. The reported working principles can be categorized into a list of seven: optical, electrical, acoustical/dynamic, thermal, frictional, chemical/electrochemical methods, and others [2]. The optical method is among the most often employed methods in monitoring CMP processes. For a patterned wafer, however, the existence of the metal lines often interferes with the optical measurement. Besides, the number of points

![Fig. 1. Schematic diagrams of planarized layers in integrated circuits.](image-url)
on the wafer of the measurement may also affect the correctness of the information collected from the wafer. The electrical methods mostly use the embedded electrodes in the pad or carrier to measure the changes in the electrical conductivity, capacity, impedance, or resistance during CMP processes to detect the process endpoint. These installations require a modified carrier and platen to accommodate the electrodes and a complex wiring route, which seriously limit the acceptance in practice. As for the acoustic emissions (AE), the merit is that the inherent high-frequency feature can keep the acquired signals away from the low-frequency noise. The temperature on pad has been used to monitor metal CMP process. Quite a few CMP machines are equipped with a platen temperature controller or a pad temperature-measuring module to ensure stable chemical reaction during polishing. It can perform the temperature measurement at a distance from the pad without attaching sensors with the entangled wires. The thermal method is limited for metal CMP due to the low sensitivity of the temperature change of the oxide layer on wafer. The frictional methods measure the change of the torque, the motor current, or the motor voltage applied to the carrier or the platen to monitor the endpoint of removal of the metal layer. However, the acquired signals often contain considerable noise, or the signal is too weak, that limits its applications. The chemical methods mainly measure the particle size, the presence of certain vapor, the electrochemical potential, zeta potential, or conductivity to detect the endpoint. The disadvantage is the probes have to contact the moving wafer. The response time is also a concern.

Many of these methods require the rearrangement of the machine structure and wires, some can only be implemented on certain types of machines. Besides, each monitoring method has its advantages and disadvantages; a multisensor system compensating the individual drawbacks is considered a good approach to the accurate in situ endpoint detection. In fact, some CMP machines now are equipped with optical as well as thermal or frictional sensors. Hence, the exploration of various monitoring techniques is essential for the advancement toward accurate CMP endpoint detection. The purpose of the current study is to introduce the thermal and acoustic emissions as monitoring signals for CMP endpoint. The result of this study is expected to be feasible for an effective multisensor monitoring system on different CMP machines without considerable machine modifications.

II. PRINCIPLES OF ENDPOINT DETECTION

A. Acoustic Monitoring

Acoustic emission has been used for tool condition monitoring since the late 1970s [3]–[6]. In metal cutting, the common AE sources are attributed to plastic deformation in the shear zone and the tool/chip interface, rubbing of the tool on the tool/chip interface and the machined workpiece surface, chip breakage and entanglement, and chipping and breakage of cutting tool [7], [8]. AE can also be used to investigate the abrasive process [9]. The frequency spectrum and the signal magnitude are obtained from acoustic sensors attached to the polishing equipment for further analysis [10]–[16], [2]. Tang et al. used AE sensors to monitor a dielectric CMP process. The root-mean square (RMS) signals were found closely related to the average material removal rate. Microscratching during CMP can be also detected [11], [12]. The results are encouraging for the current research on the metal CMP process, which uses different acoustic emissions from polishing the metal and dielectric layers for endpoint detection.

One notices the coefficient of friction between the pad and dielectric layer, 0.35, is distinguishably lower than that between the pad and the copper overlay, 0.7 [17], [18]. Secondly, the metal layer on the wafer surface is scratched by the abrasive particles in the slurry giving out acoustic emission signals. When the upper copper layer is completely removed, namely the process endpoint, as shown in Fig. 3, the slurry particles start to scratch the dielectric material beneath, which will send out different acoustic emission signals. Thus, the emitted acoustic signals during CMP of metal overlay can be used to differentiate the borderline.

B. Thermal Monitoring

Most CMP machines are equipped with a pad temperature sensor to monitor the stability of the chemical reaction during polishing. The pad temperature has also been used to monitor the endpoint of a metal CMP process with little analytical approach except a kinematic view point [19]. During a metal CMP process, the excess metal layer on dielectric film will be removed. At the beginning of the process, the area of the metal film being polished by the pad equals the wafer surface area. When the area of the metal film is reduced to that of metal via, the endpoint is reached. Meanwhile, the polish arm exerts a downward pressure onto the wafer. In association with the coefficients of friction between the pad and the metal layer and the dielectric layer, the frictional forces are produced. Multiplied by the polishing speed, the kinematic polishing power is required for the CMP process. The kinematic power is consumed for polishing work and eventually becomes measurable thermal emissions. A correlation between the kinematic parameters and the pad temperature with the derived monitoring technique has been proposed [19]–[21]. One can see that the pad temperature rise is determined by the down force, coefficients of friction, polishing speed, and the area of metal and dielectric layers. During the process, the down force and the kinematic parameters (such as the platen speed) are constant, while the area of the metal layer (or the dielectric layer) and the associated coefficients of friction vary. At the beginning of the process, the pad polishes the metal
layer only, as shown in Fig. 3, until the endpoint of the CMP process when the pad polishes the dielectric layer with a relatively small portion of metal via (less than 1% of the whole dielectric area). As mentioned, the coefficient of friction between the pad and dielectric layer is distinguishably lower than that between the pad and the metal layer; therefore, the pad temperature increases more mildly than polishing the metal layer. In use of a mathematic regression applied to the measured temperature rise on the pad, the transition from polishing the copper layer to the dielectric layer in the process can be detected. The first derivative of the thermal emissions keeps constant at the beginning phase when the metal overlay is being polished. It decreases in the transition when metal layer and dielectric layer are both being polished simultaneously across the wafer and becomes constant again when the dielectric layer is being polished eventually. The polishing endpoint is thereby identified.

The following assumptions were made for the thermal monitoring method: 1) the amount of heat produced from the exothermic chemical reaction was insignificant compared to the heat generated by mechanical friction, since the amount of polished copper material is very small; 2) the wafer was in full contact with the pad during the process; and 3) the pressure exerted by the carrier and the rotational rates of both platen and carrier was constant through the process.

III. EXPERIMENTAL SETUP

The CMP machine in the experiment is an IPEC 372 M wafer polisher carrying a 150 mm wafer with 1 μm of oxide dielectric layer and 1500-Å copper overlay film. The polishing consumables include a grooved pad, the slurry abrasive of 0.1 M 

\[ \text{HNO}_3 \] solution, and \( \text{Al}_2 \text{O}_3 \) particles of 0.1 to 0.3 μm in size and weight percentage of 3%–5%. The down pressure and the backside pressure were set 4.0 and 0.2 psi, respectively. The slurry flow rate was 100 ml/min. The polishing conditions are shown in Table II.

<table>
<thead>
<tr>
<th>Process Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down Pressure</td>
<td>4.0 psi</td>
<td>0.2 psi</td>
<td>6.2 psi</td>
<td>4.0 psi</td>
</tr>
<tr>
<td>Backside Pressure</td>
<td>100 ml/min</td>
<td>70 rpm</td>
<td>50 rpm</td>
<td>70 rpm</td>
</tr>
<tr>
<td>Slurry Flow Rate</td>
<td>100 ml/min</td>
<td>50 rpm</td>
<td>30 rpm</td>
<td>50 rpm</td>
</tr>
</tbody>
</table>

The AE sensor is required to be placed as close to the acoustic source, i.e., the polished wafer, as possible. In the current study, a Physical Acoustics Corporation (PAC) wide-band AE sensor S9208 is placed at the backside of the carrier in a hole drilled on the plastic cover of the carrier assembly, thus the AE sensor can be mounted on the metal part of the carrier. A slip ring is specially designed, as shown in Fig. 4, and put in the carrier assembly so that the tongues inside the slip ring transfer the AE signal from the inner ring on rotating carrier to the stationary outer ring. The inner and outer rings are kept connected all the time. The inherent high-frequency feature of AE signals can separate with the low-frequency noise. The AE noise caused by the friction of the slip rings is about 50–80 kHz, which is well below the higher-than-200-kHz signals produced in polishing wafer. The O-ring between the plastic cover and the metal part of the carrier can further isolate the undesired noise. The AE signals acquired by the sensor are transferred to the 49-dB preamplifier with a built-in bandpass filter of 95 kHz to 1 MHz. The sampling rate is five mega samples per second. A threshold level was set according to an online protest to cancel the background noise and to serve as a reference for the waveform properties. Two preliminary calibration tests verify that the AE signals are sent from the wafer surface and immediately correctly received by the AE sensor at the backside of the carrier. The acoustic emissions do not transfer via the pad and slurry.

In the experiment of monitoring using thermal emissions, the ADEMA Thermostream® 900 System is installed by the wafer polisher to measure the pad temperature during CMP processes. This image system captures the two-dimensional thermal emissions on the polishing pad with the temperature resolution of 0.1 K and 272 × 136 pixels in each image. The measured thermal image is calibrated before each experiment to ensure the coherence of the measurement [19], [21].

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

The energy index and RMS value of the acoustic emissions are used. They are calculated as in the following:

\[
\text{Energy Index} = \frac{1}{R} \int_0^\infty V^2(t) \, dt
\]

\[
\text{RMS} = \sqrt{\frac{1}{\Delta t} \int_0^{\Delta t} V^2(t) \, dt}
\]

where \( t \) is the integration time, \( R \) is the electrical resistance of the measuring circuit, and \( V(t) \) is the level of measured acoustic signal. The calculated AE energy is proportional to the true AE...
energy in the event [22]. It is later normalized in the experimental analysis.

Fig. 4 shows the energy index versus the polishing time. The metal film is first reacted with the chemical in the slurry, followed by the removal of copper by the slurry abrasives sending out significant elastic waves. It took a few seconds to pick up the preset down force and backside pressure at the beginning. When the metal film is fully removed, the chemical in the slurry does not react with the exposed dielectric oxide layer. The particles in the slurry scratching the surface remove little of the amount of the hard oxide material, and the AE signals decrease correspondingly to a lower level. At this moment, the process endpoint is called. The polished wafer is unloaded and examined for the removal of the copper overlay.

As indicated in Fig. 5, the energy index can identify the endpoint at varying polishing conditions of rotational rates of platen and carrier. The four condition sets well represent the polishing recipe in practice. The time of the endpoint is determined by these conditions. The correlation between the processing parameters and the time of endpoint involves both mechanical and chemical aspects and remains an empirical observation and practice. In general, slower polishing speed (at lower rotational rates of platen and carrier) needs longer time due to lower removal rate, based on the widely accepted empirical Preston equation [23]. Hence, the case (d) of platen speed = 50 r/min and carrier speed = 30 r/min shows the latest endpoint. The numerical normalization of the signal values was performed by dividing the difference between each measurement and the minimum measured value in the four experimental conditions with the difference between the maximum and the minimum of the measured values. The RMS of the AE signal shows the same results, as shown in Fig. 6. For the patterned wafer, the overlay copper film on the dielectric layer will be removed, left with the copper filled in vias. The AE signals can nevertheless tell the endpoint in time because they are insensitive to the quite small area of the exposed copper on dielectric layer (less than 1% of the whole dielectric area). Polishing the metal film sends out more acoustic emissions than the dielectric film in Fig. 4, one further notices the ratio is about 2. The number is found to correspond to the ratio of the coefficients of friction between the polishing pad and these films, 0.7 and 0.35, respectively.

The temperature rise can tell the progress of metal film removal, since the temperature rise is a function of the coefficients of friction between the pad and the metal film as well as the dielectric film [21]. In the beginning phase, the pad polishes the metal layer only, thus the temperature rises steadily with time. When part of the metal film is removed in the transition phase, the pad polishes both metal and dielectric film. Since the coefficient of friction for dielectric film is lower than copper, the temperature rises less rapidly than in the previous phase and varies with the relative amount of metal film left on wafer surface. When the metal film on wafer surface is fully removed in the end phase, the temperature rises steadily and more mildly as compared to polishing metal film in the beginning period. The second-order curve can happen when a very thick Cu film is polished. However, in most cases of CMP processes (including the currently investigated case), the Cu film is so thin that the temperature increases approximately linearly until the underlying oxide emerges.
A monitoring strategy can be developed in use of the regression technique based on this thermal behavior [21]. When monitoring, the pad temperature is constantly measured at several points on the scanning path. The least square regressions of the first and second order are continuously carried out for all existing temperature measurement. The errors of two regressions are compared to the other continuously. In the beginning period, the linear regression indeed shows less error than the second-order regression until the transition phase when both metal and dielectric films are polished and the behavior of temperature rise changes. After the error of linear regression is found larger than the second-order regression for a period of five continuous measurements, the beginning of transition is identified. Similarly, after the error of the linear regression is found again less than the second-order regression for a period of five continuous measurements, the end of transition is identified. Namely, the metal film is fully removed, or the endpoint is reached. The error of all regression is less than 5%. Fig. 7 shows the monitoring strategy schematically. Fig. 8 shows the history of temperature rise on the pad at the location of radial distance 110 mm from the pad center at different polishing conditions. The temperature rise at these points was found to follow the characteristic pattern described above. At different locations, the time of reaching the endpoint was found to be a little different. Monitoring of more locations on the pad is necessary, so that the actual endpoint of the process is identified as the latest one.

The results of monitoring using acoustic and thermal emissions are compared to each other under the same process conditions, as shown in Table III, [19], [21]. The slight delay of 0%–15% in use of the thermal method is attributed to the less agile response of the heat transfer in the system (including the slurry and polymer pad) than the acoustic wave propagation, as well as the required check for the regression errors along BC and beyond (in Fig. 6) as mentioned previously. The experimental results demonstrate the endpoint in copper CMP can be monitored in situ by both methods used individually or in a fusion manner for the nano-scale integrated circuits manufacturing.

V. CONCLUSION

CMP has become an indispensable process in deep submicron semiconductor manufacturing. In a metal CMP process, the endpoint is reached when the overlay metal layer is removed. An effective in situ endpoint monitoring method is needed to improve the yield rate and throughput of CMP. Both the acoustic and thermal emissions during polishing the metal layer and the
dielectric layer are different due to the inherent difference in the tribological behavior. In fact, the coefficient of friction between the polishing pad and the dielectrics is lower than that between the pad and the metal. The acoustic emissions decrease to a lower level when the metal film is removed. The measured RMS signal as well as the energy index can be used to indicate the process endpoint agily. Since the frictional forces between the pad and the metal layer and the dielectric layer consume the polishing power and cause the pad temperature to rise, the same fact of the difference in the coefficients of friction causes the pad temperature to increase mildly compared to polishing the metal layer. By use of a simple mathematic regression, the endpoint of the process can be identified from the change of temperature rise. The experimental results can be applied to supplement the current plant monitoring practice at no significant costs for modifications of existing CMP equipment to improve the yield rate of nano-scale IC manufacturing.

REFERENCES


Hong Hoang was born in 1958. He received the B. Sc. degree from National Taiwan University, Taiwan, in 1980, the Diplom-Ingenieur from Technische Hochschule Aachen, Germany, and the Ph.D. degree from the University of California, Berkeley, in 1988.

Presently, he is a Professor at National Tsing Hua University, Taiwan, R.O.C. His fields of interest are innovative manufacturing processes.

Yun-Liang Huang was born in 1973. He received the B. Sc. degree in mechanical engineering from National Chiao Tung University, Taiwan, R.O.C., in 1995, and the Ph.D. degree from National Tsing Hua University, Taiwan, in 2002.

Presently, he is a Senior Engineer at AU Optronics Corporation. His major research field is chemical mechanical planarization and intellectual property.