Wafer back pressure control and optimization in the CMP process*

Men Yanwu(门延武)[†], Zhang Hui(张辉), Zhou Kai(周凯), and Ye Peiqing(叶佩青)

State Key Laboratory of Tribology, Tsinghua University, Beijing 100084, China

Abstract: Chemical mechanical polishing (CMP) is the most effective wafer global planarization technology. The CMP polishing head is one of the most important components, and zone back pressure control technology is used to design a new generation of polishing head. The quality of polishing not only depends on slurry, but also depends on the precise control of polishing pressures. During the CMP polishing process, the set pressure of each chamber is usually not the same and the presence of a flexible elastic diaphragm causes coupling effects. Because of the coupling effects, the identification of multi-chambers and pressure controls becomes complicated. To solve the coupling problem, this paper presents a new method of multi-chamber decoupled control, and then system identification and control parameter tuning are carried out based on the method. Finally, experiments of multi-chambers inflated at the same time are performed. The experimental results show that the presented decoupling control method is feasible and correct.

Key words: CMP; zone pressure control technology; decoupling control DOI: 10.1088/1674-4926/32/12/126002 EEACC: 2502

1. Introduction

With the rapid development of the semiconductor industry, integrated circuit manufacturing equipment and technology, one of the top technologies, is the core of information technology. Chemical mechanical polishing (CMP), photolithography, etching, ion implantation and physical vapor deposition (PVD)/chemical vapor deposition (CVD) is parallel to the core of key technologies in large-scale integrated circuit manufacturing. Nowadays, CMP is widely proved to be the most effective technology of wafer global planarization. With the further improvements of integration level and computing speed, silicon feature sizes further decrease and wiring layers further increase. According to the latest International Technology Roadmap for Semiconductors (ITRS2010^[1]), feature sizes will shrink to 32 nm, the wiring layers will reach 13 and wafer diameters will increase from 300 to 450 mm. Cu has a good conductivity and is used as interconnect lines and low k (k =2.2–4.0)/ultra low k (k < 2.2) dielectric materials in the semiconductor industry. As the elastic modulus of copper and low k is greatly different, the ratio of Cu and low k is 15: 1-60:1, while the ratio of Cu and silica is 2:1. Cu polishing under pressure (2–6 psi) in traditional CMP polishers^[2] often results in stripping and low k damage. In addition, the material removal rate (MRR) of traditional CMP is mainly controlled by the larger wafer polishing pressure and chemical characteristics of slurry. Total copper error loss is greater than ± 30 nm, which causes some resistance uncertainty in the circuits and thus cannot guarantee the computing speed and reliability of the chip. Hence it is necessary to decrease polishing pressure and improve the characteristics of slurry, which not only takes copper CMP into the area of low down force, but also brings higher requirements to the CMP process, equipment and control of polishing pressures.

Zone back pressure control technology^[3] is still a hot research issue. The technology involves dividing a wafer into several zones and, by controlling the back pressure of each of the zones, adjusting the regional MRR to guarantee global flatness. Applied Materials have studied the technology before, but there is no public information available. SpeedFam-IPEC have studied a multi-zone pressure compensation algorithm where, first, the wafer surface thickness profile is obtained and, then, the removal profile of the wafer is calculated according to the desired removal thickness and to get the polishing pressure curve by calculating the relationship between the MRR and pressure. Finally, each zone pressure is adjusted to approximate the polishing pressure curve. Sun et al.^[4] established a compensate pressure calculation model by using the Preston equation to analyze the wafer back pressure distribution of three zones and then an 'equi-area method' was used to characterize the desired applied pressure so that the optimum combination of pressures could be determined. Lin^[5] established a 2D axisymmetric quasi-static finite element model with carrier back pressure compensation for the CMP process. The effect of a given carrier back pressure on the stress components and von Mises stress on the wafer surface was analyzed and the effect of different carrier back pressures on von Mises stress and nonuniformity on the wafer surface was investigated. The findings indicated that the axial stress was the dominant factor on the von Mises stress distribution on the wafer surface and during the CMP process it could achieve the purpose of improving the planarization of the wafer surface by compensating for different carrier back pressures. Shiu et al.^[6] regarded multi-zone CMP as a non-square multivariable system. Two control strategies with different degrees of complexity were proposed, one is the ratio control and the other is on-line multivariable control. Results show that achiev-

^{*} Project supported by the National Science Council of the People Republic of China (No. 2008ZX02104) and the State Key Laboratory of Tribology, China (No. SKLT08B08).

[†] Corresponding author. Email: menyanwu@163.com Received 5 May 2011, revised manuscript received 10 October 2011



Fig. 1. Illustration of the CMP mechanism.

Fig. 2. Schematic diagram of a polishing head.



Fig. 3. Pneumatic pressure control system of multi-chambers.

able control performance can be maintained using SVD-based multivariable-controllers and the performance is a little short of the achievable performance by using the ratio control. These studies focused on the effect of magnitude and the distribution of back pressure on the uniformity of the wafer surface polishing. For multi-chambers it is crucial that there are pressureregulating, rapid, stable and accurate responses to the target value, which is the basis of future research. However, little work has been done on how to adjust the pressures of multichambers quickly and accurately in order to achieve the target.

2. Multi-chambers coupling

CMP equipment, as illustrated in Fig. 1, consists of a polishing head, a polishing platen supporting a polishing pad, conditioner, slurry transportation, robot, etc. During the CMP process, a wafer is affixed to the polishing head and pressed facedown onto a polishing pad. Relative motion occurs between the surface of the wafer and the surface of the pad. Slurry with chemical solutions and abrasive particles held in suspension is dripped onto the rotating pad and uniformly dispersed on it due to the centrifugal force during polishing. The chemical solutions oxidize and soften the surface by chemical reaction, the polishing head produces a face-down pressure and mechanical friction is induced from the contact between the wafer's surface and the abrasive particles due to the relative rotation between the platen and the polishing head. Thus the unplanarized thin layer of the wafer surface can be removed^[7]. The function of the conditioner is to remove the remains of polishing pad surface and refresh the pad activity. The polishing head is one of the most important parts of CMP equipment. The new generation of polishing head designs mainly adopts the technology of zone back pressure control. The polishing head mainly consists of multi-chambers, a flexible diaphragm, a retaining ring, etc. Its schematic diagram is shown in Fig. 2.

Polishing head pressure control of multi-chambers includes wafer pressure control and wafer vacuum control. The pneumatic pressure control system of multi-chambers is shown in Fig. 3.



Fig. 4. Chamber 1 and chamber 2 pressure response curves.

The pneumatic pressure control system mainly consists of a controller, digital output (DO) modules, high-precision pressure sensors, electrical pressure proportional valves, a gas source, an oil mist separator, an air filter, valves, digital-toanalog output (D/A) modules and analog-to-digital (A/D) input modules, etc.

As shown in Fig. 2, the part between two chambers and the part contacting with the wafer are flexible diaphragms. When the chamber is inflated, the pressure of the chamber increases, the volume of the chamber will also accordingly increase and the adjacent chamber will be squeezed, so a coupling effect occurs between the two adjacent chambers and even spreads to other chambers. When the inflated pressure increases further, the coupling effect will become apparent. When zone 1 (e.g. the contact part of chamber 1 with wafer) is pressurized by 0.5 psi and 2 psi, separately, the pressure response curves of chamber 1 and chamber 2 are shown in Fig. 4. It can be seen from Fig. 4 that when zone 1 is pressurized, the volume of chamber 1 increases and then the zone 2 is squeezed accordingly, so that the pressure of chamber 2 increases instantaneously. However, as the electrical proportional valve is set to zero in advance, the pressure of chamber 2 will be adjusted to zero finally. In addition, with the zone 1 pressure increasing, the peak pressure of chamber 2 also increases. Figure 5 shows that when zone 3 is pressurized by 0.5 psi and 2 psi, separately, the coupling coefficient of zone 2 and zone 3 or zone 3 and zone 4 is apparently greater than the coupling coefficient of zone 1 and zone 2.

Figures 4 and 5 show that the coupling effect mainly occurs between two adjacent chambers. The inflated single-chamber not only faces the problems of volume variation, time variation and nonlinearity, but also the problem of coupling. Without first solving the problem of coupling, it will be difficult to carry out system identification and parameter estimation accurately for the chambers and execute follow-up optimization control, so it must be performed first in order to decouple the multi-chamber system.



Fig. 5. Chamber 2, chamber 3 and chamber 4 pressure response curves.

3. Multi-chambers decoupling

Known by the viscous Newtonian fluid equations, the single-chamber volume can be expressed as follows^[8]:

$$\tau_{v} \frac{dV_{z,i}}{dt} + (V_{z,i} - V_{z0,i}) = \gamma_{ii} (P_{zi,i} - P_{STP}) + \sum_{i \neq j} \gamma_{ij} (P_{zi,j} - P_{z,j}), \quad (1)$$

where τ_v is the volume expansion/contraction time constant, $V_{z,i}$ is the chamber volume, $V_{z0,i}$ is the initial volume of the chamber under standard temperature and pressure (STP), $P_{zi,i}$ is the chamber pressure, P_{STP} is the initial pressure under STP, γ_{ii} represents the expansion/contraction coefficient and γ_{ij} represents the coupling coefficient between chamber *i* and chamber *j*.

Suppose
$$P_{i,i} = \gamma_{ii}(P_{zi,i} - P_{STP})$$
 and $Q_{i,j} = \sum_{i \neq j} \gamma_{ij}(P_{zi,j} - P_{z,j})$, then Eq. (1) changes to:

$$\tau_{v} \frac{dV_{z,i}}{dt} + (V_{z,i} - V_{z0,i}) = P_{i,i} + Q_{i,j}.$$
(2)

Equation (2) shows that the inflated single-chamber faces the problem of volume and coupling variations. The degree of coupling depends on the coupling coefficient and the difference between two adjacent chamber pressures. So to reduce the adverse effect of coupling, it can make $Q_{i,j} = 0$, and then the only problem remaining is the varying volume. It can be seen from Eq. (2) that when the item $\tau_v \frac{dV_{z,i}}{dt} + (V_{z,i} - V_{z0,i})$ is smaller, the problem of varying volume can be approximated as the constant volume problem. To make $\tau_v \frac{dV_{z,i}}{dt} + (V_{z,i} - V_{z0,i})$ smaller, only decreases the item $P_{i,j} + Q_{i,j}$, however, the $P_{z,i}$ is used as the set pressure, only decreasing the $Q_{i,j}$. In addition, as a different chamber has a different pressure response characteristic before un-tuning control parameters, it is difficult to guarantee $Q_{i,j} = 0$. To solve the above problems we can take the following steps to carry out system identification and control.

Step 1: transform a varying volume problem into a constant volume problem. If a varying volume problem can be



Fig. 6. Chamber 1 pressure response curve with zone 2 (a) un-pressed and (b) pressed 3 psi.

transformed into a constant volume problem, it will facilitate the research of system identification, parameter estimation and control parameter tuning. The method is to make $Q_{i,j} < 0$, e.g. when one chamber is inflated, the adjacent chambers have been inflated in advance. Although the chamber is already squeezed at the beginning of inflation, which makes the volume of investigated chamber smaller, the decreased volume can be ignored in comparison with the total volume. The single-chamber pressure response curve is shown in Fig. 6. Compared with the adjacent unpressurized chamber, it can be seen that when the adjacent chamber is inflated in advance, the coupling effect is reduced greatly and it is easy to get a quick and stable response curve.

Step 2: eliminate the adverse effect of coupling. The method is to make $Q_{i,j} = 0$, e.g. to let adjacent chambers inflate together. As each chamber has its initial tuned control parameters in the above-mentioned step 1, the pressure response speed has been enhanced and the response time has been shortened, which can always maintain $Q_{i,j} = 0$ during the inflation process. At this time it is necessary to optimize the control parameters further after eliminating the adverse effect of coupling.

Step 3: inflate all multi-chambers at the same time. In this step the control parameters obtained in the step 2 will be evaluated and optimized further.

The function of chamber 4 is to control the wafer edge pressure and its volume is the smallest in all chambers. As chamber 4 is the smallest and the initial squeezed volume cannot be ignored any longer, it is impossible to carry out system identification and control parameter tuning accurately by using the above-mentioned 3-steps method and it is necessary to take a different approach to decoupling. The method is to let zone 3 and zone 4 become pressed together, e.g. make $Q_{i,j} = 0$. As previously mentioned, the control parameters are unknown and different at the beginning of un-tuning, so it is hard to make $Q_{i,j} = 0$ and the results are incredible. However, the volumes of chamber 3 and chamber 4 are relatively small and their response time is very short, so it can maintain the $Q_{i,j} = 0$ during the whole inflation process. The pressure response curves of chamber 3 and chamber 4 are shown in Fig. 7.



Fig. 7. Zone 3 and zone 4 pressure response curves.

4. System identification and parameter estimation

The process of system identification involves obtaining the mathematical model by mathematical analysis based on input/output experimental data of the controlled object and determining the model structure (including the form of equations, order of equations and delay state, etc.) in detail. Based on the calculated model, the method used to determine the model parameters is known as parameter estimation^[9]. In this paper the main purpose of system identification is to determine the mathematical model of the pneumatic pressure control system and parameter estimation is used to obtain all the parameters of the mathematical model.

With reference to the above-mentioned decoupling method, the main steps to carry out system identification and parameter estimation are as follows.

Step 1: system identification and parameter estimation are first used for chamber 1. Here, chamber 2 is inflated to 3.5 psi.

Step 2: system identification and parameter estimation are used for chamber 2. Here, chamber 1 and chamber 3 are inflated to 3.5 psi.

Step 3: system identification and parameter estimation are used for chamber 3 and chamber 4. Here, chamber 2 is inflated to 3.5 psi.

The step response method is used to establish the mathematical model of the pneumatic control system in later experiments. As the response time of the electric pressure proportional valve is 0.1 s, so the lag time of the control system can be ignored. Due to the shape of the step response curve transfer function, it is assumed to be first-order inertial function. Its mathematical model is as follows:

$$G = \frac{K}{Ts+1}.$$
 (3)

For the step input voltage signal u_0 , the step response of first-order inertia system is:

$$p(t) = K u_0 (1 - e^{-t/T}).$$
 (4)

When the time t tends to infinity, $p(t)|_{t\to\infty} = p(\infty) = Ku_0$ then

$$K = \frac{p(\infty)}{u_0}.$$
 (5)



Fig. 8. Experimental platform with polishing head.



Fig. 9. The calculated curve compared with the experimental curve.

It is known that the input voltage is 1.5 V, the output pressure value is 15 kPa, so the K = 10, then

$$p(t) = 10u_0(1 - e^{-t/T}).$$
 (6)

Set t equal to T/2, T and 2T separately, then obtain three values of T accordingly based on the information of the curve and take the average of the three values, T can be obtained. For chamber 1, T = 0.3 s, then

$$p(t) = 15(1 - e^{-3.3t}).$$
 (7)

To verify the accuracy of the identification results, the calculated results are compared with the experimental data. The set step pressure value is 20 kPa. It can be seen that the calculated curve is basically consistent with the experimental curve, as shown in Fig. 9, which proves that the above hypothesis of the first-order inertia system is correct. The experimental platform is shown in Fig. 8. In addition, there is a minimum value limit to the response of the solenoid valve, which is mainly due to^[10]: solenoid inertia of the electromagnetic solenoid valve (the lag effect caused by electromagnetic transient phenomena), initial preload of feedback spring, no load motion of moving parts and Kunlun friction.

As regards the remaining chambers, the time constant T can be calculated by the above-mentioned method and they are: $T_2 = 0.26$ s, $T_3 = 0.24$ s, $T_4 = 0.16$ s.

5. System decoupling control

The experimental curve in Fig. 8 is the open-loop outcome without any control algorithm. It can be seen that when pressure reaches the steady-state value, the response time exceeds 2 s. As to a chamber, the response time even exceeds 5 s, which is unacceptable to the high requirements of time and efficiency in the CMP process, so it is necessary to improve the system response characteristics and dynamic performance. During the inflatable experiments, it is inevitable that there is a large deviation pressure value from the set pressure value at the beginning of inflation, which causes an integral accumulation problem of the PID integral part and results in the controlled parameter exceeding the ultimate value of electric proportional valve, so as to cause a large overshoot of the system and even shock. The pneumatic control system takes an integral separate PID algorithm to adjust the pressure of chambers in order to improve the system response characteristic and dynamic performance. The main idea of the algorithm is to cancel the integral role at the beginning of the down-press process to enhance the stability of the control system and introduce the integral action when the pressure value closes to the set value to improve control precision by eliminating steady-state error^[11].

In the inflatable experiments, it was found that if the coupling effects between two adjacent chambers are not considered, the single-chamber control parameters (e.g. PID parameters) obtained are not suitable for multi-chamber simultaneous inflation at all. Therefore it must be done to decouple the multi-chamber pneumatic control system so as to obtain the available control parameters.

The system decoupling control process is mainly divided into three steps:

The first step is to tune the control parameters of single chamber. The second step is to optimize the parameters based on the tuned parameters. The third step is to further optimize the parameters of all chambers. The detailed steps are as follows:

(1) single-chamber system control parameter tuning

According to the previous analysis of decoupling and the method of single-chamber system identification, the main steps to get the single-chamber control parameters are as follows:

Step 1: control parameter tuning of chamber 1. Chamber 2 is inflated to 3.5 psi.

Step 2: control parameter tuning of chamber 2. Chamber 1 and chamber 3 are inflated to 3.5 psi.

Step 3: control parameter tuning of chamber 4 and chamber 3. Chamber 2 is inflated to 3.5 psi.

PID parameter tuning is mainly carried out by using the Matlab simulation and experimental feedback method. Figure 10 is the Matlab simulation model and pressure response curve, and Figure 11 is the experimental response curve after PID tuning.

The controlled parameters of chamber1 after PID tuning are as follows:

At the beginning of inflation: $K_p = 8$, $K_i = 0$.

Close to the set value: $K_p = 8$, $K_i = 5$.

Lots of experiments are carried out to get the pressure response time in different set pressures and steady-state errors after PID parameter tuning. It can be seen from Table1 that



Fig. 10. Matlab simulation model and experimental response curve.



Fig. 11. The experimental response curve after PID tuning.

the response time is under 1.5 s under a low down force and steady-state errors are under $0.02 \text{ psi}^{[12]}$.

The control parameters *PI* of all chambers are shown in Table 2.

Table 1. System response time and steady-state errors under different low down forces.

Set pressure	Steady-state errors	Response time
(psi)	(psi)	(s)
0.5	0.0198	0.544
0.6	0.0160	0.588
0.7	0.0030	0.675
1	0.0184	0.730
1.5	0.0155	0.802
1.8	0.0198	0.769
1.9	0.0063	0.803
2	0.0078	0.910

Table 2. Control	parameters PI	of all chambers.
------------------	---------------	------------------

Chamber	Р	Ι	
Chamber 1	8	5	
Chamber 2	12	6	
Chamber 3	15	6	
Chamber 4	11	6	

(2) Control parameter optimization of a single chamber As the tuned PI parameters were obtained in the case of not eliminating the adverse effect of coupling completely by decreasing the adverse effect of coupling, it is necessary to optimize the experimental results further. The method of op-



Fig. 12. Response curves of all chambers under different pressures. (a) 0.5 psi. (b) 1 psi. (c) 1.5 psi. (d) 2 psi.

Table 3. The optimized parameters of all chambers.

Chamber	Р	Ι	
Chamber 1	5	5	
Chamber 2	11	6	
Chamber 3	12	6	
Chamber 4	11	6	

timization is to make $Q_{i,j} = 0$ of Eq. (2), e.g. to eliminate the adverse effect of coupling completely. The main steps are as follows:

Step 1: control parameter optimization of chamber 1. Chamber 1, chamber 2 are inflated simultaneously.

Step 2: control parameter optimization of chamber 2. Chamber 1, chamber 2, chamber 3 are inflated simultaneously.

Step 3: control parameter optimization of chamber 3 and chamber 4. Chamber 1, chamber 2, chamber 3, chamber 4 are inflated simultaneously.

The optimized parameters after the above-mentioned steps are shown in Table 3.

(3) Parameter tunings of all chambers

The ideal pressure response curves have been obtained after the parameter optimization of a single chamber, and then experiments are carried out for all chambers at the same time and the results evaluated. The experimental curves under pressure of 0.5, 1, 1.5, 2 psi are shown in Figs. 12(a)-12(d).

It can be seen that the above optimized control parameters can meet the requirements perfectly, thus it can also prove that the whole experimental method is feasible and correct.

(4) Pressure distribution of the wafer surface

CMP pressure is between the polishing head and pad, the wafer surface pressure is the ultimate target. The control system implemented above is a semi-closed loop control in the strictest sense. Pressure is transferred from the chambers to the wafer surface. So it is necessary to test the overall pressure effect of all chambers.

At present, a contact area array pressure sensor, which is produced by U.S. production company Tekscan, is used in the experiments to get the distribution of pressure by using dot matrix sensor units. The measuring range of the sensor is 0-5 psi and accuracy is 1%. When all chambers have 1 psi pressure applied, the captured pressure distribution figure is shown in Fig. 13.

From Fig. 13 it can be seen that the silicon surface pressure distribution is more uniform, same with the chambers' pressure 1 psi.

6. Conclusions

In this paper the following conclusions can be drawn:

(1) a decoupling method of multi-chambers is developed for the multi-chambers coupling effects. Three steps are taken to carry out the system identification and control, except for chamber 4. As to chamber 4 another decoupling method is used.

(2) Based on decoupling method, a detailed decoupling process is given to carry out system identifications of all chambers, and then the mathematic models are achieved by parameter estimation. The experimental results prove that the mathematic models obtained are correct.



Fig. 13. Pressure distribution of the wafer surface under 1 psi pressure of all chambers.

(3) According to the mathematic models above, the control system takes the integral separate PI algorithm to adjust the pressures on the wafer back. The parameter tuning of a single chamber is carried out and then optimizations are implemented by a detailed decoupling process. Finally all chambers are inflated at the same time and the decoupling method is checked. The experimental results prove that the decoupling method developed in this paper is feasible and correct. In addition, the system after optimization has higher stability, reliability and accuracy, and can be referenced in the engineering fields.

(4) A wafer surface pressure distribution experiment is carried out to test the overall pressure effect of all chambers. The captured figure shows that the silicon surface pressure distribution is more uniform, as is the chambers' pressure.

References

- International Technology Roadmap for Semiconductors, Interconnect, ITRS, 2010update. http://public.itrs.net. Accessed 30 December 2010
- [2] Xu Chi, Guo Dongming, Jin Zhuji, et al. A signal processing method for the friction-based endpoint detection system of a CMP process. Journal of Semiconductors, 2010, 31(12): 126002
- [3] Sun Yuhui, Kang Renke, Guo Dongming, et al. Zone back pressure control using for vacuum carriers in CMP. Equipment for Electronic Products Manufacturing, 2004, (07): 34 (in Chinese)
- [4] Sun Yuhui, Kang Renke, Guo Dongming. Modeling of back pressure distribution on the wafer loaded in a multi-zone carrier in chemical mechanical polishing. Materials Science Forum, Switzerland: Trans Tech Publications, 2006: 233
- [5] Lin Y Y, Lo S P. Finite element modeling for chemical mechanical polishing process under different back pressures. Journal of Materials Processing Technology, 2003, 140(1–3): 646
- [6] Shiu S J, Yu C C. Multivariable control of multizone chemical mechanical polishing. J Vac Sci Technol, 2004, 22(4): 1679
- [7] Lin Y Y. Influence of a retaining ring on strain and stress in the chemical mechanical polishing process. Materials and Manufacturing Processes, 2007, 22: 871
- [8] Shajii A, Nagarkatti S P. System and method for controlling pressure in remote zones. USA Patent, No. 6986359, 2006
- [9] Yang Sanqing, Wang Renming, Zeng Qingshan. Process control. Wuhan: Huazhong University Press, 2008 (in Chinese)
- [10] Zhong Shengguo, Peng Guangzheng. Modeling and simulation about pneumatic pressure control system based on proportional pressure valve. Machine Tool & Hydraulics, 2004, (10): 90
- [11] Liu Jinkun. MATLAB simulation of advanced PID control. Beijing: Electronic Industry Press, 2004
- [12] Men Yanwu, Zhang Hui, Zhou Kai, et al. Low downforce copper CMP distributed control system. CIMS, 2010, 16(12): 2653