

Estimation of Dressing Effects in Chemical Mechanical Polishing with Sorts of Diamond Dressers

Kuen-Ren Chen^{1, a} and Hong-Tsu Young^{1, b}

¹Room 138, Engineering Building, No.1, Sec.4, Roosevelt Rd., Taipei 10617, Taiwan

^achen.kuenren@gmail.com, ^bhyoung@ntu.edu.tw

Keywords: Chemical Mechanical Polishing, Diamond Dresser, Dressing, Trajectory Analysis.

Abstract. Chemical mechanical polishing (CMP) is a semiconductor process that is used to achieve global planarization. To maintain the stability and the throughput of CMP, the polishing pad needs to be dressed by a diamond dresser. By calculating the distribution of the scratch numbers of diamond grits in the pad, the effect of dressing can be estimated and the pad profile can be predicted. All the parameters of dressing motion were taken into account. The target equipment of this study is Mirra Mesa, which is manufactured by Applied Materials. Different types of dressers, including ring-types and full-type, different pitches and arrangements are used in CMP process. Each of them was simulated in this study. The diamond pitch is relative to the number of diamonds and the scale of the total scratch time, but does not affect the profile of the scratch time. Regular or random arrangement makes no difference in the profile of the scratch times. The width of the ring influences the number of diamonds and the scale of the total counts while the profiles remain similar.

Introduction

It has long been noted that material removal rate (MRR) drops rapidly over time without conditioning or dressing during the CMP process [1]. It is undesirable because it will decrease the throughput and increase the cost of CMP process. As the pad wears out and its surface becomes glazed during the CMP process, the pad surface increases, and the average contact pressure decreases. These and other changes to the pad lead to the decrease in the material removal rate. The use of a dresser is to keep the pad surface condition consistent. Hence, dressing is a critical part in obtaining and maintaining a good polishing performance.

Numerous studies focused on the removal rate of CMP, while fewer on the dressing rate. However, certain theorems and principles were adopted in both the removal rate and the dressing rate. The Preston's equation $MRR = K_p PV$ is the first and most commonly used form for the CMP removal rate and also is used in dressing rate later [2]. It is an empirical equation, which states that the average MRR is linearly proportional to the product of the applied down force (pressure) and the relative velocity between wafer and pad surface. However, in practical application, the relationship between MRR and down force is usually nonlinear. So, later studies proposed some nonlinear models, like Zhang et al. suggested an equation $MRR = K_p \sqrt{PV}$ [3] and Shi provided $MRR = K_p P^{2/3} V$ [4]. Another MRR model based on statistical distribution of abrasive particle sizes and considering uniform pad surface roughness was proposed by Luo and Dornfeld [5]. There are still many researches considering the effect of pad properties, like pad roughness, and other polishing conditions on the polishing removal rate.

Kim and Jeong defined the l as the sliding distance and stated that the sliding distance was positively relative to the wear thickness of the pad [6]. The sliding distance is referred to Preston's equation. Chang et al. further applied the sliding distance to predict the pad profile [7]. On the other hand, Borucki derived the asperity height probability distribution function (PDF) of pad asperities to calculate the wear rate of pad [8]. The PDF is used in both the removal rate of wafer and the wear rate of pad [9]. Kasai conducted a series of kinematic analyses and also took into account the downward force [10], but his model analyzed the motion of a disk and its trajectory. All of them discussed the velocity, but only Tyan's model touched upon the parameters of dresser and none of them concerned

the sweep motion. Though in Tyan's study, sweep had been mentioned, it was assumed to be very slow and its effect was neglected. In this study, all the parameters of dressing kinematics were taken into account, including the sweep motion and the model was applied in different kinds of dresser.

Method

The basics intent of the estimation model are to calculate the scratch numbers for a given cross section of a pad. In this study, we focused on discussion of different kinds of dresser, while other parameter has been examined in another study [11].

At first, one must choose the diamond pattern and set the dressing condition, and then the trajectories will be a set of equations of the diamond pattern. For computing convenience, the coordinate is fixed on the polishing pad and the X-axis is selected for the given cross section. The scratch numbers can be obtained by solving the trajectory function values of $Y=0$. Finally, the cross section is divided into many small sections and scratch time for each section is calculated.

Therefore, the trajectory of the dresser should be defined, so the x-y coordinate is fixed on the pad, which is adhered to the platen. Then, the coordinate system will rotate with the platen. The variables are also demonstrated in Fig. 1, where R_p is the distance from the axis of dresser arm to the center of pad, R_d is the length of dresser arm, ω_1 is the angular velocity of pad (platen), ω_2 is the angular velocity of dresser arm, α is the initial phase angle of pad, β is the initial phase angle of dresser arm, t is time, x_d is the horizontal distance from the i th diamond to the center of dresser and y_d is the vertical distance from the i th diamond to the center of pad.

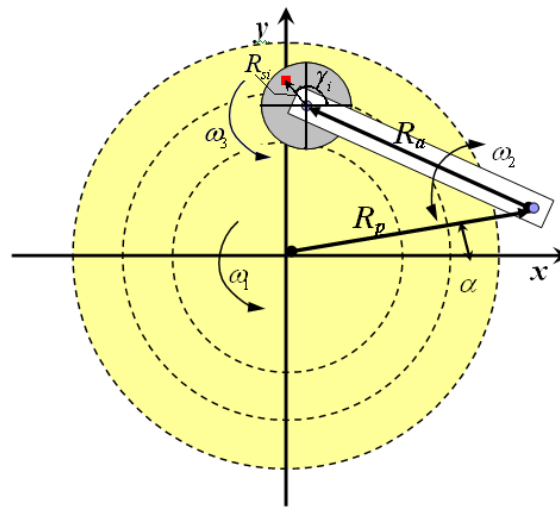


Fig. 1. Schematic of variables in trajectory equation.

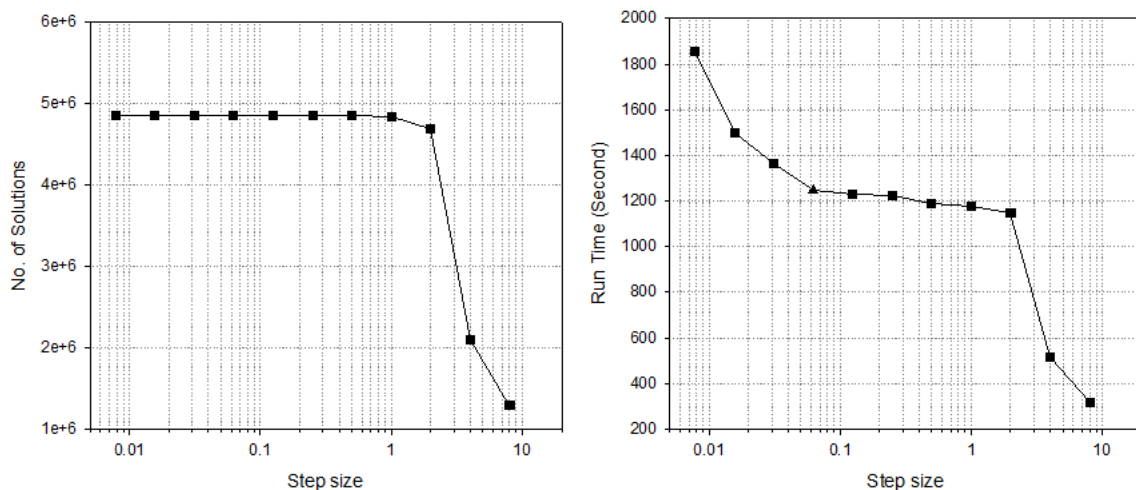
Both of R_p and R_d are constants since they are two dimensions of the CMP equipment. They could be directly measured from the practical equipment. The ω_1 can be adjusted directly by the control system of Mirra Mesa and their units are in rounds per minute (rpm). It is relatively more complex to determine the exact value of ω_2 . Three variables will influence ω_2 . The ω_2 is determined by its sweep frequency and amplitude, and the amplitude is relative to the start point, the nearest point between the center of the dresser and the center of the pad, while the end point is the farthest one. So, sweep frequency, start position and end position are the three variables that influence ω_2 and all the three can be directly modified through the control system. Though the amplitude is the dependent variable of the two points, practically, it needs to be transformed to a radian system to calculate its angular velocity. Furthermore, ω_2 should be converted to be in the same unit as ω_1 and ω_3 . Accordingly, the start point indicates a start angle while the end point indicates an end angle of sweep. The difference between start and end angles is the amplitude.

The aim of this model is to observe the scratch time distribution over a given cross section. For calculating convenience, this cross section is selected for the x-axis. That is, for all points that their

trajectory equations, $y(t)$ are zero, those are scratching the cross section, x -axis. Hence, the objective is to solve those equations for $y(t)=0$. This is a root searching problem for solutions of the equation in one variable, t . Bisection method [12, 13] was adopted in the program to solve the equation for y equals to zero. The first problem that accompanies the bisection method is how to find the two adjacent points that are in different signs. Forwarding with a small step is used to scan the whole range of time. Once the two adjacent points are in opposite signs, and then the bisection method can be carried out to solve the approximate solution. Another problem will continuously emerge, regarding how small the step size is adequate. If the step size is too big, it might miss some solutions as soon as the step crosses them. A convergence test was implemented to decide the step size. Fig.2 (a) illustrates the results of the convergence test for a full-disk dresser with smallest pitch which means the most number of diamonds. It converged when step sizes were smaller than 1.

Nevertheless, it is not efficient to make the step size as small as possible. The run time of the solving program will increase dramatically as the step size narrowing down. Hence, the run time should be taken into account at the same time. While the number of solutions run up to the degree of 10^5 or 10^6 , missing tens or hundreds of solutions is still acceptable. It is not required to find all of these solutions precise. So, a balance between number of solutions and the run time should be poised. Fig. 2 (b) compares the step sizes and the run time. It can be discovered that, from large steps to small ones, the run time increases rapidly in the beginning, but increases smoothly later, and then increases quickly again. The reason is that the run time of the program is dominated by two loops, one is scanning the whole range and the other is root-searching with the repeated bisection method. At the start, the number of solutions increases speedily, and the run time does too because the more solutions are found, the more repeat times of bisection method are needed. While the number of solutions is converged, the time of the performing bisection method is increasing, so the smoothly increasing time is due to the diminishing of the step size. So far, the run time is still dominated by root-searching, but when the step size was continuously decreases, the run time will be led by the scanning process and eventually increases dramatically.

Finally, the step size was chosen as 0.0625, which is the smallest step size before the run time was dominated by step size. Therefore, the step size could ensure the convergence of number of solutions without scanning inefficiently with too small a step size.



(a) Step size and number of solution.

(b) Step size and run time.

Fig.2. Convergence test of program.

Regardless of the distribution area, the diamonds are located in the area within a radius of 51 mm. The geometry of a disk-type dresser can be separated into three aspects, while other factors will not affect the scratch trajectory. The three aspects are diamond pitch, arrangement and distribution area, which will be discussed below.

1. Diamond Pitch

The diamond pitch is the interval of two adjacent diamonds. When diamonds distribute regularly over the whole dresser, the number of diamonds have different pitches. The number is the amount that appears in the simulation diamond pattern. The number of diamonds is inversely proportional to the square of pitch. The pitches of 600 and 450 μm can be found in real products of dresser.

2. Arrangement

The arrangement of diamonds depends on the manufacturing process of the dresser. It would be arranged in a grid or at random. To investigate the effect of the arrangement of diamonds, different ways of arrangement are made in the diamond patterns. The diamonds might be arranged in grid or different degrees of irregularities. The degree of irregularity is determined by the range of random interval. For example, while the random interval is 50% of the diamond pitch with 600 μm , the range of random number would be from -300 to +300. The algorithm would pick two random numbers from this interval, and one is for horizontal coordinate while the other is for vertical coordinate. Assuming that the two random numbers are -40 and 280, then this diamond would be located in (-40, 280) relative to the grid one. The ranges of random intervals vary from 15%, 30% to 50% of diamond pitch. These diamond patterns with different random ranges are shown in Fig. 3.

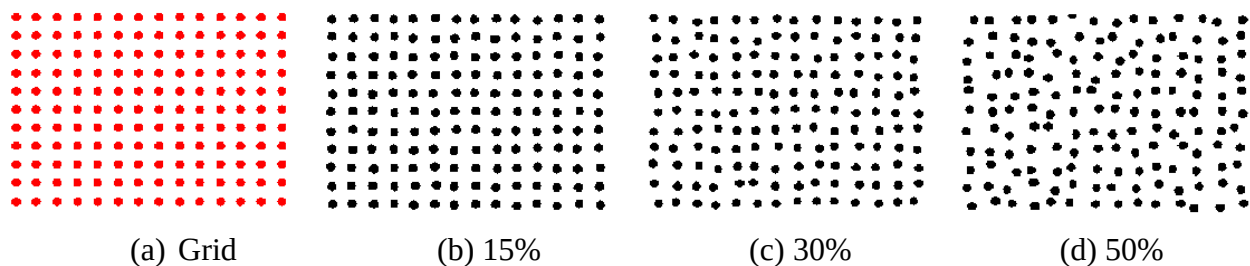


Fig. 3. Diamond distributions with different random ranges.

3. Distribution Area of Diamond

There are various types of dressers in the market. Considering the distribution area of diamonds, some of them are ring types and others are full dressers. The width of the ring can be determined by the outer radius and the inner radius of the ring. All dressers come in the same outer radius regardless of full or ring types, but differ in the inner radius. The full type dresser can be seen as the ring type dresser with an inner radius equal to zero. That is, the full type dresser is the one with its width equal to its outer radius. Therefore, the distribution area of diamonds can be characterized by the width of the ring. Fig 4. illustrates the diamond patterns with different widths of ring.

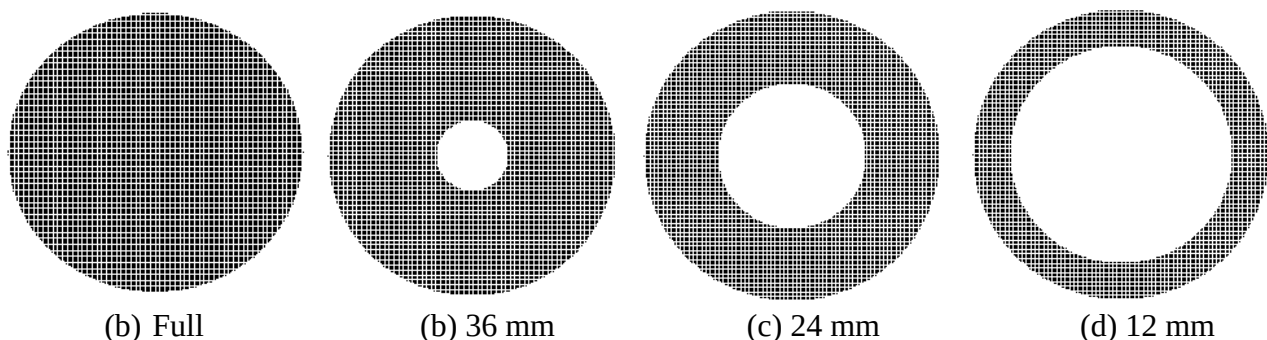


Fig. 4. Diamond patterns with different widths of ring.

Results and Discussion

In the following simulations, parameters are targeted to Applied Materials, Mirra Mesa. The parameters include the geometry of platen (pad), arm and other dressing conditions. The dressing conditions that these simulations were referring to are listed in Table 4, except those simulations of different dressing conditions. This set of dressing conditions were practical ones from the industry. Basically, only one dressing condition would be changed to examine its effect at one time. The scratch times would be calculated for the time period that platen rotated 100 rounds. It was almost one minute under the rotational speed of platen of 109 rpm.

Dressers with pitches of 450, 600, 1200 and 2000 μm were simulated. An example of the results is shown in Fig. 5. Since the number of diamond is inversely proportional to the square of diamond pitch, so does the scale of scratch times. The profiles are similar for all pitches and only vary by scale. Though the scratch times were in different scales, a diamond on the dresser with less diamonds would afford higher down force than that of with more diamonds. So the diamond would take a heavier load and therefore had higher removal effect. Nevertheless, the dressing rates of different pitches might not be ascertained, but the profiles of scratch times, those were the distributions of removal thicknesses, would be much alike.

The effect of diamond arrangement was examined by simulation. The simulation results for ranges of random intervals varied from 15%, 30% of pitch to 50 % of pitch have been made. The three different random arrangements were compared with the regular arrangement. The profiles of scratch times of the four arrangements were almost the same. It can be concluded that, once the number of diamond was fixed, the scratch times would be the same regardless of how it was arranged.

It was stated that there were different types of dresser with various diamond distributions. The effect of diamond distribution was simulated in this section and was demonstrated in Fig. 6. It can be found that all the profiles were similar but in different scales. The difference scratch times between full disk and width of 36 mm was caused by the diamonds in the central part of dresser, and the difference between widths of 36 mm and 24 mm was due to the difference of width of ring, so did the discrepancy of widths of 24 mm and 12mm. On the other hand, assuming the down force on the dresser distribute as a function of radius, then the dressing effect considering the down force may be seen as a linear combination a set of the results from different width of ring. Nevertheless, the result of a linear combination of them may remain the same profile of scratch times, since all the components are in similar profile. It can be also inferred that the more diamond on the dresser, the less loading is distributed to each diamond. Therefore, the life time of dresser may be extended by increasing the number of diamonds, but this inference still need to be experimentally verified.

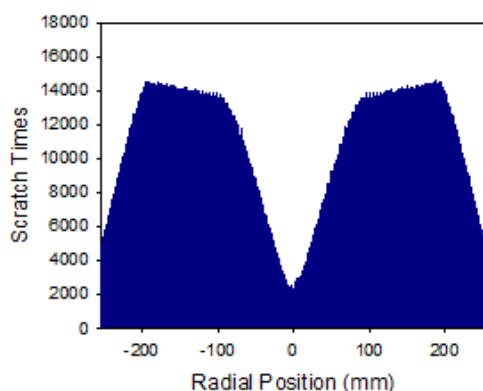


Fig. 5. Result for pitch 450 μm .

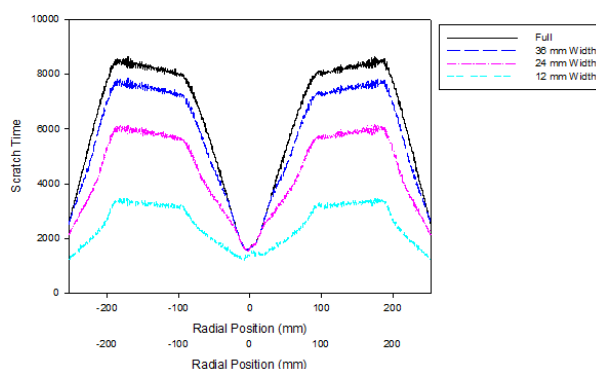


Fig. 6. Results of different widths of rings.

Conclusions

The model is suitable for various parameters including types of dresser, the rotational speed of the dresser and the pad, the sweep frequency and the sweep amplitude. Those parameters in the simulation model directly target the dressing conditions, users can simply adjust the parameters that they would like to change. No more calculation is need for users when applying this model and no parameters of dressing motion was neglected.

The effects of the diamond distributions are examined. The diamond pitch, arrangement and distribution were studied. The diamond pitch directly influences the number of diamond, so does the scale of scratch times profile, but has no effect on the shape of scratch times. The arrangement does not play any role in the scratch times. It makes no difference on scratch times from regular to random arrangement. In addition, the distribution area of diamond is relative to the number of diamonds and then will affect the scale of scratch times. However, the profiles of scratch times are similar for different widths of ring, so the distribution of down force on the dresser can be considered trivial for the scratch time profile. To sum up, diamond distribution will simply influence the number of diamond but does not change the shape of scratch times.

References

- [1] Stein, D., Hetherington, D., Dugger, M. and Stout, T.: J. Electron. Mater. Vol. 25(10) (1996), p. 1623.
- [2] Preston, F.W.: J. Soc. Glass Tech., Vol. XI (1927), p. 214.
- [3] Zhang, F., Busnaina, A.A. and Ahmadi, G.: J. Electrochem. Soc. Vol. 146 (1999), p. 2665.
- [4] Shi, F.G. and Zhao, B.: Appl. Phys. A: Mater. Sci. Process. Vol. 67(2) (1998), p. 249.
- [5] Jianfeng, L. and Dornfeld, D.A.: Semiconduct. Manuf, Vol. 14(2) (2001), p. 112.
- [6] Kim, H. and Jeong, H.: J. Electron. Mater. Vol. 33(1) (2004), p. 53.
- [7] Chang, O., et al.: Microelectron. Eng. Vol. 84(4) (2007), p. 577.
- [8] Borucki, L.J., et al.: J. Eng. Math. Vol. 50(1) (2004), p. 1.
- [9] Borucki, L.: J. Eng. Math. Vol. 43(2-4) (2002), p. 105.
- [10] Kasai, T.: Tribol. Int. Vol. 41(2) (2008), p. 111.
- [11] Chen, K.-R. and Young, H.-T.: ECS Transactions, Vol. 19(7) (2009), p. 61.
- [12] Faires, J.D. and Burden, R.L.: *Numerical Methods*. 1993, Boston: PWS-Kent Pub. Co.
- [13] Burden, R.L. and J.D. Faires., *Numerical Analysis*. 2005, Belmont, CA: Thomson Brooks.