Particle transport and reattachment on a mask surface

Pavel Nesladek\textsuperscript{a}, Steve Osborne\textsuperscript{b}, Christian Kohl\textsuperscript{a}

\textsuperscript{a} Advanced Mask Technology Center, Raehnitzer Allee 9, 01109 Dresden, Germany, 
\textsuperscript{b} Sigmameltec Ltd, 3-37-7 Shimousao, Asao-ku Kawasaki, 215-0022 Japan,

ABSTRACT

The cleaning processes used today for photomasks were developed over decades and optimized to fulfill customer specifications. Some mask procedures were adapted from wafer cleaning technology. A principal technique, megasonic (MS) cleaning, yields high particle removal efficiencies (PRE). However, MS can frequently cause feature damage, and so damage becomes the principle limitation to MS power levels applied to small feature sizes. The use of lower MS power levels can benefit from a better understanding of removal mechanisms. In several publications the effects influencing the mechanisms of particle cleaning were discussed \cite{1}. Particle transfer was investigated experimentally on wafer surfaces using bath tools and was tracked using fluorescent optical microscopy \cite{2}. The goal of our investigation is to test the validity of the aforementioned models for mask cleaning using a spinning mask and a megasonic head mounted on an arm swinging over the mask surface, which is the most common hardware setup used for mask cleaning tools. While this equipment setup provides a useful variability, it also introduces disadvantages \textit{e.g.} non-equal distribution of the megasonic power across the cleaned surface as will be shown. We will focus on some of the main parameters \textit{e.g.} chuck speed, arm swing speed and media flow, which are strongly coupled by the fluid dynamics and cannot be treated separately. All three parameters influence particle-mask decoupling and reattachment during particle transport by the media stream across the mask surface. The approach to estimate the particle removal and reattachment rate is illustrated. The experiments performed allow the conclusion that the reattachment rate on a flat spinning mask surface is lower than previously assumed and the most critical part of the cleaning process is the detachment of the particle from the surface.

Keywords: Particle removal, cleaning mechanism, media flow, spin versus bath processes, particle dot tests

INTRODUCTION

Megasonic technologies lay at the heart of mask cleaning processes. Only a small fraction of test particles are removed by other concurrent treatments in a final clean. Due to the violent nature of MS vibration, mask makers might truly enjoy the thought of leaving megasonics completely behind. However, Figures 1a and 1b show that megasonics can not be entirely left out or left behind without a suitable replacement technology or mitigation protocol to diffuse these potentially destructive vibrations to Sub Resolution Assist Features (SRAFs).

In previous work we demonstrated two such mitigation protocols \cite{3}. The first mitigation protocol resulted in the focal diffusion of MS energies using chuck rotation and arm sweep speeds. Secondly, the occurrence of resonant energies traveling in the plate were demonstrated and evaluated for possible utility. Mask makers remain highly receptive to alternative methods of application or to entirely different technologies such as dry plasma or cryogenic methods or high velocity spray
technologies. No alternative technology has been demonstrated to entirely replace MS. Consequently, improved methods for the application of MS technology remain very interesting despite the breakage deterrents.

Understanding and modeling the mechanisms of particle removal is instrumental to determining improved, gentle and effective MS cleanings. This is a complicated undertaking in which particle reattachment acts as a sort of competitive or reverse reaction. But the extent of this competitive reaction is not well understood or well explored. For example, the mechanical motions of spin processes are integral to the delivery of MS and these processes produce highly interactive parameters as shown in Figure 2 which plots the % PRE against arm swing and chuck rotation. Within this data set the role and extent of particle reattachment remains hidden. If reattachments rates are high, then conceptually speaking, the process time must increase to compensate. When the process time increases, the probability for breakage also increases.

![Figure 2](image)

**Fig.2** – Demonstrates the highly interactive parameters of chuck speed and arm swing speed for a 30s MS exposure. But the extents of particle reattachment remain hidden in the data set. If particle reattachments rates are high then longer exposure times will be needed than for a process in which the particle reattachment rate is low. Reducing MS exposure times is one factor in reducing the probability for feature breakage. Hence, more specific information on reattachment is needed.

New feedback methods are needed to elucidate particle reattachment in spin processes and attain possible benefits. For example, we used particle dots places on clean masks in such a way as to indicate the factors of transport path and reattachment rate. These characteristics become evident when new particles adhere in the free regions of the plate (Figure 3). Plates with a complete particle coating can not yield this type of information.

![Figure 3](image)

**Fig.3** – Particle dots on clean mask will show both the rates of reattachment and the path of the particles off of the plate. Particles removed from the area close to mask center are carried by the media flow toward the mask edge. During transport, the particles may reattach to the surface in areas closer to the mask edge. This has the effect of artificially reducing the PRE in these areas. In order to judge this effect, more knowledge of particle reattachment rates is necessary.

The detailed appearance of the particle dots can also tell us interesting things. Figure 4 demonstrates the tendency of particles to either stay in the fluid or attach to the surface. During drying, the volume of the slurry droplet decreases. For example, if the surface affinity is low, the particles will remain in the fluid longer and produce a high central density. If the surface affinity is high, the particle will attach at their first opportunity and the density will be relatively uniform.
Fig. 4 – a) Distribution of particles deposited from an isolated droplet on the masks surface. The evaporating droplet leaves a fingerprint of the chemical and environmental conditions it experienced during drying. As one can see from the picture, the particles tend to remain in suspension and concentrate during the drying. The droplet drying time is about 10 minutes. Fig. 4b and 4c illustrate two alternative theoretical distributions. Fig. 4b represents the extreme case of high surface affinity where the particles attach to the plate at their first contact. Fig. 4c represents the extreme case of low surface affinity. Particle clustering results.

Fig. 5 – a) Demonstrates the radial decline of cleaning efficiency in: - % PRE vs. Mask Radius; b) - Spatial Distribution of PRE plotted as Counts vs. Radial Location on the Mask. Both plots represent a 30s exposure to MS. The removal process is stopped before completion to better illustrate the radial nature of the removal process and provide the basis for modeling rates and mechanisms.

In order to de-convolute known interactive factors and treat the spatial effect shown if Fig. 5b, the cleaning efficiency is first evaluated as a count density (n/mm²). Then the particle characteristics of hub, ring and rim can be determined with enhanced statistical significance. As discussed above, the radial dependence of cleaning efficiency is easily seen to vary when changing chuck speed and arm swing speed or media flow. The contributions to particle removal by forces other than direct beam exposure become increasingly clear (see Fig. 8). Other contributions to the radial removal pattern observed might be itemized as follows:

- Poor corner removal is due to lower MS power resulting from no direct contact.
- Reattachment of the particles from the central area of the mask.
- Turbulence of the media flow deters the communication of near beam energies (corners and arm reversal points).
- Changes to the MS power delivery can occur due to variation media depth created by hydrodynamics of water flow created by the moving plate and MS arm.

METHODOLOGY

Masks with Complete Particle Coverage
PRE in our work was measured on non structured mask blanks using Si₃N₄ particles with a broad size distribution of about 70nm-5µm with maximum at 105nm (see. Fig. 6). For measurement of defects or particle density, Siemens DFX 70 mask blank inspection tool was used. The defects identified are assigned to one of 6 groups according to their type:

a) Defects chrome side  
b) Chrome pinholes  
c) Chrome inclusions  
d) Defects glass side  
e) Glass inclusions  
f) Removed defects
Only the defects on the front side were considered. In order to minimize the error of PRE estimation, the blanks used were scanned before the application of particles and these defects were marked as potential non-removable defects. This first scan is called the “pre scan” in following text. Subsequently, the particles were applied and next scan is referred to as the “coat scan” and provides information about count and density distribution of particles before the clean. After the cleaning experiment, the third, “post scan”, identifies remaining particles. One can calculate the PRE from the particle count or particle density from these scans in two ways:

\[ \text{PRE}_2 = \frac{(\text{coat} - \text{post})}{\text{coat}} \]  
[1]

\[ \text{PRE}_3 = \frac{(\text{coat} - \text{post})}{(\text{coat} - \text{pre})} \]  
[2]

PRE2 represents the particle removal efficiency calculated with the ideal assumption that all particles identified in the coat scan are removable. If the pre scan defect count is not negligible then the second equation based on the assumption of non-removable particles identified in the pre scan is more useful and provides a better estimate of the cleaning efficiency (PRE3). In a perfect world, the cleaning efficiency would be estimated by tracking each particle separately, and according to the result, assigned to one of the following categories: a) not moved, b) moved to different place on the mask or c) removed from mask surface. Due to the very high effort needed to identify each particle and the inevitable uncertainties of a high particle count, we decided to use the parameters described in Eq.1 and Eq.2.

**RESULTS**

As already mentioned, the PRE depends on several parameters. Due to strong interactions between parameters and nonlinear dependencies, the cleaning efficiency for the whole mask cannot be described by polynomial functions of incoming parameters as one may expect. One result of these interactions is that the cleaning efficiency of a spinning plate decreases as a function of radius (Fig.5a). Some of the well known parameters influencing PRE are the composition of the particle and cleaned surface (zeta potential), the shape of the particle, the distribution of different sizes and the particles density.

In Figure 7, the slope changes which occur to the left of the arm reversal position (r ≤ 70mm) occur in a region that receives direct beam contact. The slope changes which occur to the right of the arm reversal position occur in a region which receives
only near and far beam MS effects. These different beam energies influence both detachment of particles and their rarer reattachment. The change of flow behavior towards turbulent and the affinity of the particle to the plate (electrostatic and Van der Waals forces) are also strong influences. Smaller particles are more resistant to detachment.

Fig. 7 – Variation of radial cleaning efficiency: The distributions of removed particles per mm² are plotted against the radius of arm swing. Slope changes at r = 70mm occur at the arm reversal position. At this place, a peak in the PRE occurs, which is attributed to the time delay at the reversal point (slow down and change of arm direction). The relation between PRE and factors shown here is not linear. This data treatment offers a means of quantifying the collective contributions of: 1) water movement in completing a particles removal from the plate, 2) near and far beam effects of the applied megasonic energy (non-direct beam forces). a) low arm swing speed, b) high arm swing speed, c) low chuck speed, d) high chuck speed, e) medium arm and chuck speeds. Three data sets provide idea of the typical scattering of data discussed.

Radial Distribution of the PRE
Figure 8 compares the observed cleaning against the theoretically modeled direct beam coverage. Figure 8c displays a normalized plot of the difference. The theoretical modeling in Figure 8b uses a constant arm speed and assumes direct exposure to the MS beam to be the only particle removal force. This figure shows a significantly greater accumulated MS exposure at the mask center due to the mechanical motion of the megasonic arm across the mask. Consequently, the PRE falls by 1/r where r is radius from mask center. No direct MS power is delivered left of the reversal point. As shown earlier (Fig. 1), we can omit the impact of the pure media flow without MS power as minor. Three additional forces are largely responsible for the differences seen between the observed and the theoretical: media flow; centrifugal force and indirect MS beam energies.
Fig. 8 – Radial PRE. a) Displays the test results at constant chuck speed of 218rpm and arm swing time 1.9s, 30s process time on mask with uniformly distributed particles. b) Displays the theoretically calculated direct MS beam coverage for the conditions above with a delay time at reversal points of 0.2s and a 4mm beam diameter. The MS beam passes 8mm from the mask center. c) PRE per millisecond of direct MS beam exposure. Assuming MS power to be the only removal force, this function has to be flat. The difference in the PRE efficiency has to be explained by additional effects: e.g. contributions to removal by the media flow, centrifugal force and indirect MS beam.

Fig. 9 – a) Radial PRE function for fixed arm position with MS nozzle placed 4 cm from mask center. The dashed line represents the Gaussian fit; b) spatial distribution of PRE for fixed arm process for particle count per mm².

Figures 9a & 9b show the cleaning response when the megasonic beam is positioned 4cm from the mask center for 30s. If cleaning were uniform on both sides of the beam, then the PRE would follow the Gaussian distribution shown in Figure 8a. The experimental result is distinctly non-Gaussian, with a nearly vertical slope for PRE at r = 33mm. Evidently, the freed particles within the ring are prevented from moving towards the mask edge in the same way as freed particles outside the ring move towards the edge—suggesting that unrestricted movement both within and above the boundary layer plays an important role in reattachment rates. When a particle is prevented from moving, it must reattach when water flow stops.

Fig. 10 – a) Shows the PRE for 3 particle dots applied at different radial locations on the mask as function of the process time. The lines represent the PREs for fully coated masks treated for 30s. b) Shows the positions of the particle dots on the mask.

Also note that the difference between the data shown in the lines versus the data shown as data points of Figure 10a is not meaningful. It is an artifact of comparing a 30s exposure on a fully coated mask to 15s exposure on a mask with isolated particle dots.
Modeling the Kinetics of PRE

As one can see in Figure 10a, the PRE is not linear with respect to time. From the perspective of kinetics, particle removal for a particular type and size of particle can be described by the same mathematical equation as used for first order chemical reactions[4]:

\[
\frac{dn}{dt} = k_n (n_0 - n)
\]  

Here \( \frac{dn}{dt} \) represents the particle removal rate, \( n_0 \) is the initial particle count or particle density per mm\(^2\), \( n \) is removed particle count/density. The constant, \( k_n \), contains the difference between forces of removal and attraction. This coefficient is non zero for particles with removal forces higher than attraction forces and zero for all other combinations. Note please, that the time (\( t \)) in the equation is not the process time, but is the integral removal force. Since the \( k_n \) is different for different groups of particles, we have to treat these groups separately. The more commonly used total removal efficiency is then the sum of integrated effects over all particle sizes on the mask without regard to radial differentiation. According to the hypothesis, the removal rate is smaller for particles with a higher ratio of attraction to removal forces. This effect is observed in Figure 11a.

Figure 11 takes the data of Figure 10 and explodes the particle size information from the central dot. This new curve set also contains a prominent shift at about 180s. This was due to an accidental change in the megasonic power applied. However, the result does demonstrate clearly that when the megasonic power increased by 15%, the \( k_n \) increased significantly, even when the prior cleaning rate was approaching its asymptotic maximum.

Reattachment of Particles

The reattachment rate was estimated for masks with particle dots close to the center. At this central position, both the removal rate and risk of reattachment is advantageously high. Logically, any new particle appearing outside of the dots would have its origin from within the particle dot. The reattachment risk is increased due to the increased transport distance to the mask edge. To avoid mixing of reattached particles from successive cleaning steps, a map subtraction was performed after each 15 second cleaning to perform the needed particle accounting step. In this way, the count, size and placement of the reattached particles was estimated.

Figure 12a shows the removed particle count, reattached particle count and the rate of reattachment (count of reattached particles divided by count of removed particles) as a function of accumulated 15s cleanings. As expected, the reattached particle count decreases as the number of particles removed begins to slow. The reattachment rate is increasing gradually as the mean particle size is reduced (Fig.12c). This increase is consistent with our model showing a progressive concentration of the smallest particles on the plate. The removal process becomes dominated by \( k_n \) for the smallest particles. Even for the smallest particle category, the percent of reattachment in no case ever exceeds 15%. Based on our expectation prior to testing, this seemed low.
Figures 13 and 14 show the results of modifying the particle-to-plate charge interaction (zeta potential) by modifying the pH of the droplet applied. A matrix of particle slurry drops of identical volume and particle source were applied on the mask surface. Since the particle dots originate from the same slurry concentrate (dilution ratio 5000:1), the particle size distribution as well as the shape and composition can be assumed to be identical. Particle counts and PRE were recorded after one 15s cleaning process. The removal efficiencies reveal the relative particle-to-plate attractions. The variability seen in the results is largely attributed to 1) the positions of different pH dots at different radii on the mask and 2) the arm sweep uniformity limitation of a 15s clean. In a general sense, the more basic the slurry droplet, the greater the Si$_3$N$_4$ particle is attracted to the plate. Dots deposited at neutral pH are roughly twice as easily removed as their counterparts at pH 11.

Figure 14 shows the fingerprints of dried droplets over a broad pH range. Although these particle maps are not fully understood, one can say that a particle’s strong affinity to the mask surface becomes apparent by its early attachment to the plate. Secondary influences to the dot fingerprints may be due to the rates of sedimentation during drying. Particles deposited due to a breakdown of the colloid suspension would not necessarily indicate a greater particle-to-plate affinity. The pictures shown in Figure 14 confirm the hypothesis, that at high attraction to surface the particles are distributed more uniformly (see Fig 14a). The root cause for ring like distribution of particles at high pH (Fig.14c & 14d) is not fully understood.

Figure 15 presents the particle size profile for removal and reattachment in a different way than Figures 11 and 12 above. One can see in Figure 15c that removal is very complete and hence the reattachments are only occurring for particles below 150nm. The reattachment rates for 80nm particle are 10 fold higher than for particles greater than 150nm. Reattachment rates over an integrated 180s duration show 80nm particles nearly 100 fold more inclined to reattach than 250nm particles (Fig. 15d).
Fig. 14 – Distribution of particles within the particle dots—‘dot fingerprints’. The pictures illustrate the change in particle density caused by change of pH and so surface energy. The pH range explored represents a 10,000 fold change in concentration of hydronium ions. The droplets differ in their shape due to air flows and vibration during drying as well as the evolving chemical conditions within the droplet. a) pH = 7.1;  b) pH = 8.0;  c) pH = 9.2;  d) pH = 10.4;  e) pH = 10.9

Fig. 15 – Particle size distributions for removals and reattachments:  a) particles removed in the first 15s run;  b) particles reattached in the first 15s run;  c) particles reattached in the 12th run;  d) size distribution for particles reattached within twelve 15s clean process runs. Since the particle counts differ strongly, the distributions are normalized to 1000 particles of the most populated size— in our case 105 nm.
Figure 16 shows a reattachment after stopping the clean process at 15s. Reattachment coincides with the movement of fluid boundary layer on the mask. Our test uses a clockwise rotation. This produces an apparent counter-clockwise motion of the boundary layer. Small particles evidently can not escape the slower moving boundary layer with high probability. Consequently, we see smaller particles form a Coriolis-like spiral in their creep toward the mask’s edge. In Figure 16a, the particles are prevented from moving towards the mask center by centrifugal forces. As the cleaning would proceed unimpeded, the magnitude of reattachments is expected to decrease and their positions on the mask is expected to move gradually towards the edge.

![Graphs showing reattached particles on the mask](image)

**DISCUSSION AND CONCLUSIONS**

A new method of particle application was explored that allows bulk observation of particle reattachments in using MS beam application on a spinning plate. Evaluation of the PRE using masks with particles covering the entire mask surface has a serious disadvantage. The particles may reattach to the surface in areas closer to the mask edge and obscure the PRE in these areas. The method presented here allows quantifying particle reattachment by tracking of a large number of particles of varied size. Custom methods of particle counting and position tracking on the mask were created in R© that allowed the particle density per mm$^2$ to be evaluated. By using single particle dots of ~10mm and ~6000 particles on cleaned surfaces we could correlate removal efficiency with reattachments originating from the slurry dots. A particle’s removal or reattachment can be considered competitive aspects of particle transport.

Although simple proof of particle reattachment is shown in Figures 12 and 16; the rate of reattachment was less than that anticipated. Based on a previous publication [1] wherein fluorescently marked single particle motions were tracked in a bath setup, the frequency of attachment could not be confirmed for our spin processes.

We find the following statements on particle transport consistent with our observations:

Reattachments are linked to a particle’s size and its radial location on the mask. The smallest particle size evaluated was 80nm and they showed the highest rates of reattachment. Particles greater than 150nm were removed with an insignificant number of reattachments when 15s clean processes were applied. In a practical sense, reattachment, with our particular spin cleaning setup, is relevant only to particles less than 150nm which showed a reattachment rate between 3 and 14%. We have to admit, that the cleaning efficiency is deterred by the 15s process interruptions characteristic to our method. Reattachment rates are therefore most probably exaggerated by this means and should be treated as an inequality (i.e. reattachment is <15% for 80nm particles). Imagine a series of 15s cleaning times stops the process of cleaning when some particles are in motion. Stopping the process will force reattachments to occur.

We observed that given a specific MS power; a maximum PRE can be reached at certain process time, which can not be exceeded by further extension of the process time (Fig. 11a). Particle removal can be grossly modeled as the ratio of attractive forces to the forces of removal. As this ratio approaches unity the particle is compelled to move. A particle’s continued motion off the plate is highly dependent on its ability to be entrained in the rapid fluid movements above the boundary layer. Likewise, continued suspension in the media deters reattachment. If the particle remains within the relatively tranquil boundary layer, the probability of reattachment increases. This suggests that an improvement in gentle removals
could be engineered with an MS application providing a more continuous agitation to small particles while maintaining adequate fluid motion to assist particle entrainment in the media above the boundary threshold.

Particle removal is seen to behave as a kinetic interaction and can be described as first order rate reactions. As well, each particle size and its position on the mask have a certain probable reversibility. If one plots the removal for all particle sizes over the entire mask, the shape of the removal curve will be qualitatively the same as that seen in Figure 11a. In contrast, when different particle sizes and radial positions are sequestered, we see that different $k_n$ emerge and can be used to more accurately describe particle behavior. To clarify further, the particle removal rate is greatest for the largest particles and smallest for the smallest particles. Particle reattachment is lowest for large particles and greatest for small particles. Reattachment is inversely proportional to removal. Reattachment decreases with increasing process time and with the reattachment probability for each detached particle.

After a certain process time, the clean process reaches equilibrium between attraction and removal forces. From then on, the cleaning process effectively stops. Any particle detached from the surface, due to fluctuation in removal force (e.g. MS power increase), is with very high probability reattached to the mask surface near its starting position. The effective cleaning rate approaches zero, even when some particles continue to gradually move over the mask surface.

Although smaller particles may bond with lower net charge distributions than big particles, they may still exhibit greater resistance to removal because they associate more intimately within surface structure of the plate. Short range forces dominate in closer association. A small particle’s attachment is enhanced by the opportunities to share in short range dipole and induced dipole forces. Larger particles offer a larger profile against which drag forces can work. A small particles probability to reach conditions where the drag force sufficiently exceeds attraction is small.

Our test of particle adhesion vs. the pH of the applied slurry produced a doubling of adhesive force. This large variation points out an industry wide deficiency. Reliable benchmarking of processes and equipment requires repeatable testing over time and different manufacturing locations. Variability in the chemical nature of particle slurries makes matching processes or specifying tool performance characteristics unreliable. The mask industry would benefit by adopting a standard particle slurry made with a specific particle composition, size profile, pH, complexing agents and suspension method. No standard has yet emerged.

REFERENCES

ACKNOWLEDGEMENT
AMTC is a joint venture of AMD, Qimonda and Toppan Photomasks and gratefully acknowledges the financial support of the Federal German Ministry of Education and Research (BMBF) under Contract No. 13N9675 (ANTARES).