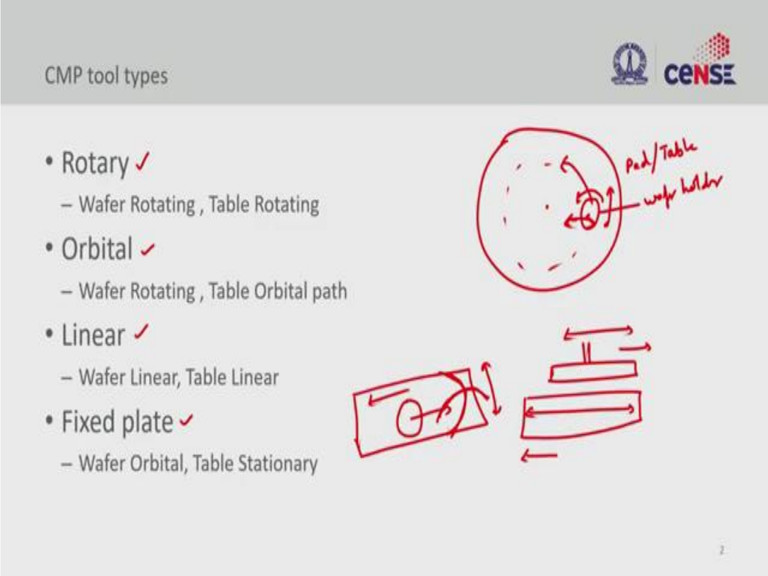
Fundamentals of Micro and Nanofabrication
Prof. Shankar Kumar Selvaraja
Center for Nano Science and Engineering
Indian Institute of Science, Bengaluru

Lecture – 52
Chemical Mechanical Polishing (CMP): Tool and Process

So, in this session on Chemical Mechanical Polishing, we will continue looking at how the CMP process is done. We will look at the tool configuration to start with and then look at the critical issues that one may face, and then how we are going to address those issues is what we are going to discuss in this lecture.

(Refer Slide Time: 00:56)



CMP tool types

- Rotary ✓
  - Wafer Rotating , Table Rotating
- Orbital ✓
  - Wafer Rotating , Table Orbital path
- Linear ✓
  - Wafer Linear, Table Linear
- Fixed plate ✓
  - Wafer Orbital, Table Stationary

So, let us look at the tool types and tool configuration. So, in the last lecture on CMP, we discussed the four essential elements, and among the four elements, the two are the mechanically important ones. We see a top-down image here. There is a pad, and then we have a wafer holder. So, these two become a significant factor in addition to the conditioner and slurry dispensed.

So, the important mechanism here is an abrasive action between the pad and the wafer. The pad and wafer must be in constant motion, and the motion can be of different kinds; we have two elements, and we can make both of them rotate or just one of them rotate, and also based on the way they rotate. So, depending on the motion, there are different tool configurations, out of which we have listed four possible configurations. However, there could be many based on application.

The first configuration is rotary. Here, both the pad and wafer are rotating with respect to the central axis. The next type is orbital, in which the wafer is rotating while the table is taking an orbital path. So, instead of rotation, we will have an orbital motion of the table or, in this case, the pad.

The next configuration is linear. So, linear is pretty straightforward where both the wafer and the table are moving linearly. Either they are orthogonal to each other or they are moving in opposite direction. For instance, if we take a cross-section, the pad can be moving left and right, and the wafer is also moving left and right. But they are always moving in opposite directions; when this pad moves to the left, the wafer will move right, and this is how friction is generated. If both are moving in the same direction, then the force applied will be minimal.

The last configuration is a fixed plate geometry. So, the pad is stationary while the wafer is moving. In this case, the wafer can make an orbital motion. So, it is going to move along an axis while the pad is stationary.

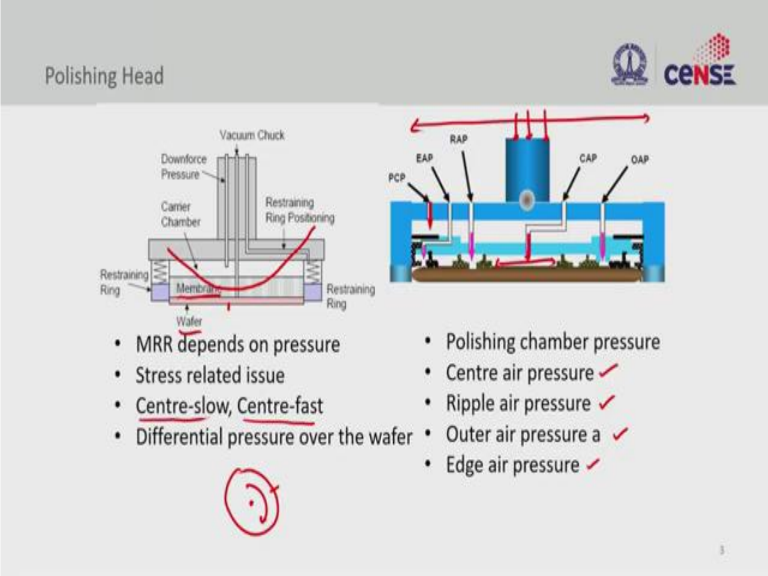
So, this is how we can implement a rotary, orbital, linear, or fixed plate configuration. It is up to the tool manufacturer to find what is the best way of removing the material. So, we want to apply maximum pressure at the same time we do not want to spend a lot of energy rotating these objects. Also, we want to get the best uniformity.

To give an example, if we have just linear motion, i.e., both wafer and plate moving in left and right directions, then we will have a non-uniformity that is going to form a profile as explained in the slide. This is because we have movement in only one direction.

So, what we could do is to have movement in the vertical direction also. This motion will again give a non-uniformity orthogonal to the previous direction. So, we will have some sort of Gaussian-like shape. But, when we start rotating and also move the wafer left and right, the non-uniformity will be distributed, called a systematic non-uniformity.

This is the best way of getting an excellent uniformity because we are not relying on a unidirectional approach. We are distributing the force across the wafer location, not just at the center or the edge. So, all these things become very important while choosing the system that should give highly uniform flatness.

(Refer Slide Time: 06:58)



Polishing Head

- MRR depends on pressure
- Stress related issue
- Centre-slow, Centre-fast
- Differential pressure over the wafer

- Polishing chamber pressure
- Centre air pressure ✓
- Ripple air pressure ✓
- Outer air pressure a ✓
- Edge air pressure ✓

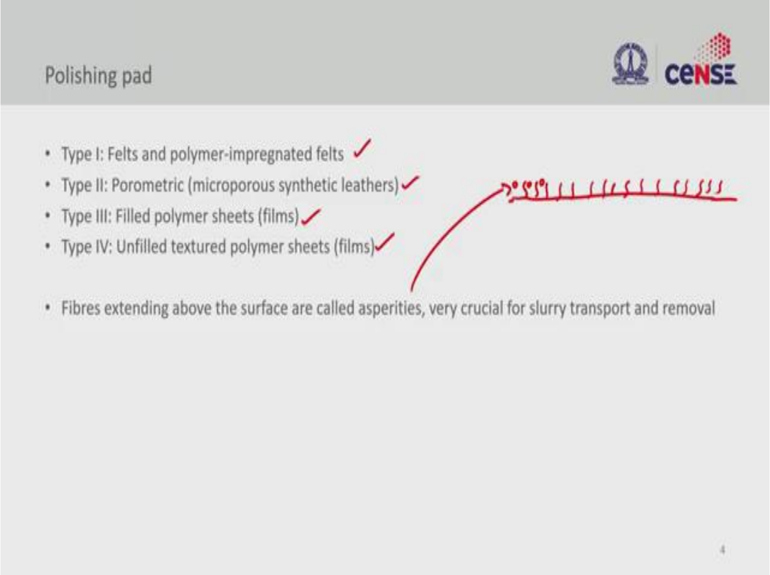
Now we look at how to achieve flatness. It is not just about getting the wafer into contact with the pad but about the applied pressure; pressure should be uniform across the wafer. So, if we take a tiny sample, let us say a 1 inch by 1 inch, then we can reasonably get a uniform force distribution. But here we are talking about very large wafers like 4 inches, 6 inches, 8 inches, 12 inches in diameter. So, if the wafers become larger and larger and the wafer holder is also getting larger and larger. Though we have a larger wafer, we are going to apply pressure starting from the middle. The spindle will apply pressure on this wafer, but we want it to be equally distributed across the wafer.

So, in recent days the polishing head technology has improved a lot where we can change the pressure locally. We can apply different pressure at different positions. For instance, in this case, the central location has a different pressure valve than the outer ring right or the outer edge. So, by applying various air pressures, we are going to control the pressure that is applied between the wafer and the pad, and wafer flatness is achieved. If we do not control it locally, we will have a variation discussed in the earlier slide.

So, if we look at the polishing head, we will have a very simple membrane, and then we have a wafer held by the restraining rings. So, when we apply a force downwards, the force at the center will always be high compared to the edges. So, because of that, pressure or the force profile will be as shown in the schematic. So, we will always remove more material from the center compared to the edge.

So, this is how we get a distributed pressure, but at the same time we should also notice the velocity difference that we get. So, if we take a circular geometry and rotate about its axis, the velocity at the edge will be much higher than the velocity at the centre. At the centre, it is stationary. As we saw in the Preston equation, the removal rate is directly proportional to velocity. So, that means a large pressure is applied at the centre, but the removal rate is very low. At the edges, the pressure is low, but the velocity is very high, and we do expect a higher removal rate. However, whether these two will compensate each other or not is the question. So, that depends on the topography or the pressure differential that we have. In order to control that differential, one needs to control the pressure locally, from the center to the edge. This is where the advanced polishing head holders come into picture where we create different chambers that could help us achieve a better uniformity.

(Refer Slide Time: 11:09)



Polishing pad

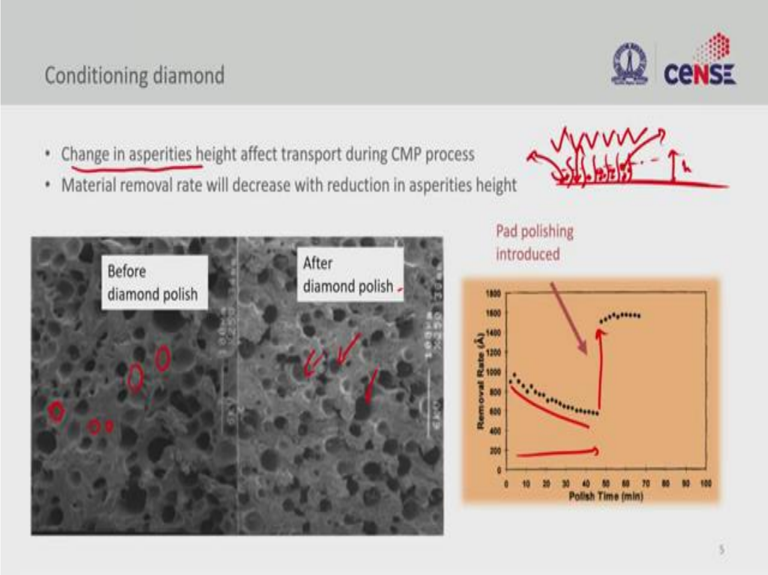
- Type I: Felts and polymer-impregnated felts ✓
- Type II: Porometric (microporous synthetic leathers) ✓
- Type III: Filled polymer sheets (films) ✓
- Type IV: Unfilled textured polymer sheets (films) ✓

- Fibres extending above the surface are called asperities, very crucial for slurry transport and removal

Now we will look at the pad part. So, there are different types of pads one can get. A pad is a soft polymer material on which we can put the abrasives. So, there is a type I that is nothing, but polymer-type material could have felts that can help transport the abrasives. The next thing is porous type a leather material; that can help you transport this abrasive material. Another one is just simple filled polymer sheets or unfilled textured polymer sheets.

So, all these materials have some texturing on the surface. This texturing helps to transport the abrasive slurry that is dispensed and also helps to remove the products formed as a result of abrasive action between the slurry and the wafer surface. So, all we need is a surface and some elements or soft corrugations, called asperities, on the surface. These soft corrugations will help transport this slurry and help in applying this pressure onto the wafer surface.

(Refer Slide Time: 12:54)



Conditioning diamond

- Change in asperities height affect transport during CMP process
- Material removal rate will decrease with reduction in asperities height

The next thing is a conditioning diamond. This is the third element that we saw as one of the important components in the CMP system in addition to pads and wafer head. So, the surface of the pad will get smoothened after a while. The reason for that is abrasives are going to come and deposit. Initially, we have a very rough surface. Then, slurry makes the surface smooth, and also, the removed etch products will come and deposit onto the pad. Because of that deposition, we will have a smoothened surface. The deposition fills up the corrugations, and their effective height reduces. The smaller the height, the less efficient in transporting the slurry and also the product. The conditioners are used to rejuvenate the pad. So, these conditioners have abrasive diamond tips. They will remove these products out the pad becomes abrasive again.

The SEM picture on the left side shows the one that is collecting all these slurry and then becoming smooth. We can see that the number of holes is less, and some of the holes are filled. After the conditioning process, we see all these holes open up and this is how we increase the abrasive nature, which is evident in the plot on the right. So, as the polishing time increases, the removal rate decreases because the abrasive nature of the pad is getting reduced. Once the polishing step is introduced, the removal rate improves because we are making this texturing height higher so that slurry transport becomes effective. The pad is not smooth anymore. We can use the analogy of sandpaper. So, when we are polishing or removing any rust from the iron pipe, once we start sanding, we will start accumulating t the products from the rust, then we would not be able to remove them anymore. After tapping the sandpaper a few times to remove all those deposits, we have the sand ready for abrasive action again. So, this is how the abrasive pad is conditioned to increase the removal rate.

(Refer Slide Time: 16:36)

Slurry

- Slurry choice is material dependent
- slurry composition: abrasive particles and solution.
- Abrasive particle
  - Size, type and concentration (distribution)
  - 10-300 nm
  - SiO2 and Alumina; most popular
  - CeO2, TiO2, MgO, ZrO
  - They after MRR and surface damage (hardness of the particle)
  - Concentration 1-30%
- Solution
  - Choose appropriate PH/zeta potential
  - Bulk and surface PH
  - Addition of surfactants

The next important component is slurry that enables abrasive action. So, what is the constituent of this slurry? The slurry consists of abrasive particles in a solution. So, we have some particles in some solution, and then we dispense this solution on the pad so that these particles could go and do abrasive action on the wafer.

The abrasive particle has different characteristics like size type and also concentration right. So, whether we have many particles or a very small number of particles in a given volume, the concentration is important. The size of the particle can range from 10 nanometres to 300 nanometres. It depends on the type of polishing we are trying to do; the final smoothness strongly depends on the abrasive particle size.

The particles primarily used are dielectrics like silicon dioxide or silica particles or alumina particles as they have high yield strengths. We can apply enormous forces, yet these particles will not disintegrate. Other particles like cesium oxide, titanium, titanium dioxide, magnesium oxide, and zirconia can be used for specialized purposes.

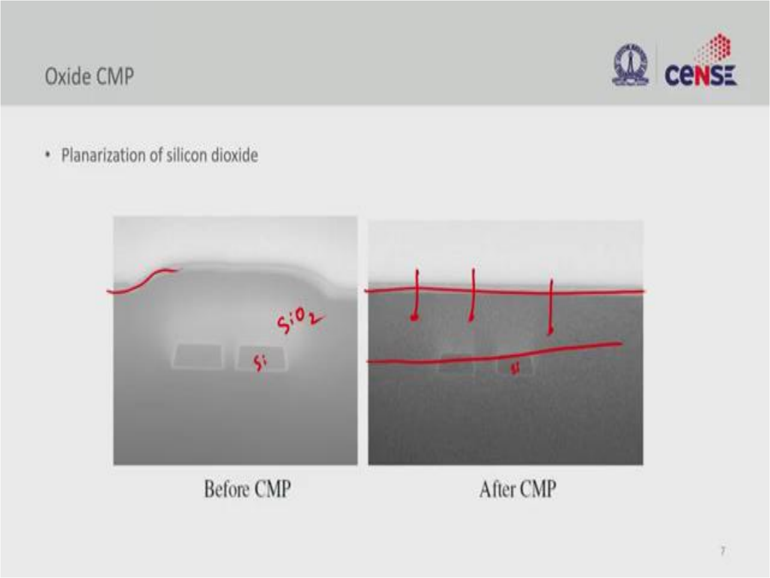
The hardness of the particle is something that one should be very careful about. Though we say these particles should withstand high pressure, we do not want them to be too hard to result in surface damage. They should not create scratches. The whole idea is to polish; we do not want these particles to make the surface rough. So, that is something that one should be careful about while choosing the correct abrasive particle.

The concentration can vary from 1 percentage to 30 percent. The concentration helps in the material removal rate. So, when we have a higher concentration, the removal rate is very high because we have more particles available for abrasive action. But, at the same time, if we want to have a smoothening process, we want to have a low removal rate. If we do not want topography to change and we just want to polish to make it smooth, then concentration and hence the removal rates can be lower to avoid material loss.

The solution that can be used strongly depends on particle nature. So, the type of particle used and the zeta potential between the particle and solution dictates the solution choice. One needs to be also aware of the pH. So, that is the main thing. Also, whether the solution has wettability with respect to the pad or not. If the pad is hydrophobic to the solution, we can not dispense the particle uniformly. So, we could have an accumulation of these particles. So, one needs to be careful about choosing the proper pH for the solution, and also, it should be in concurrence with the type of particle that we are going to use. We do not want to accumulate these particles within the solution and want a good distribution of particles

So, one needs to apply some thought to the solution selection after choosing the abrasive particle.

(Refer Slide Time: 21:08)



Oxide CMP

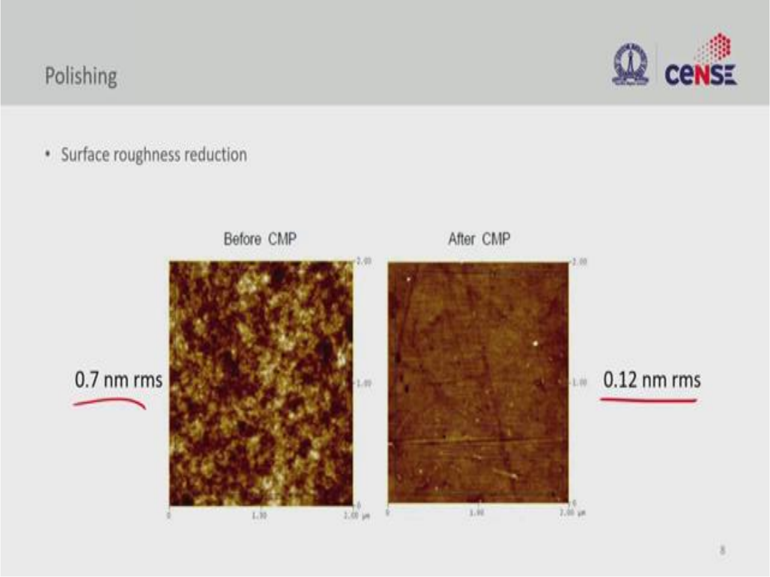
- Planarization of silicon dioxide

Before CMP After CMP

So, there are two reasons we use CMP. One is to make the surface flat. Cross-section SEM shows how one can use CMP to remove the topography; the left-hand image shows after a conformal deposition where we see the topography is still there, but after polishing, we see a very flat surface.

If we want to expose the underneath structure, we can keep on doing the polishing process and reach silicon. In this process, the CMP was stopped right after getting a flat surface.

(Refer Slide Time: 21:53)



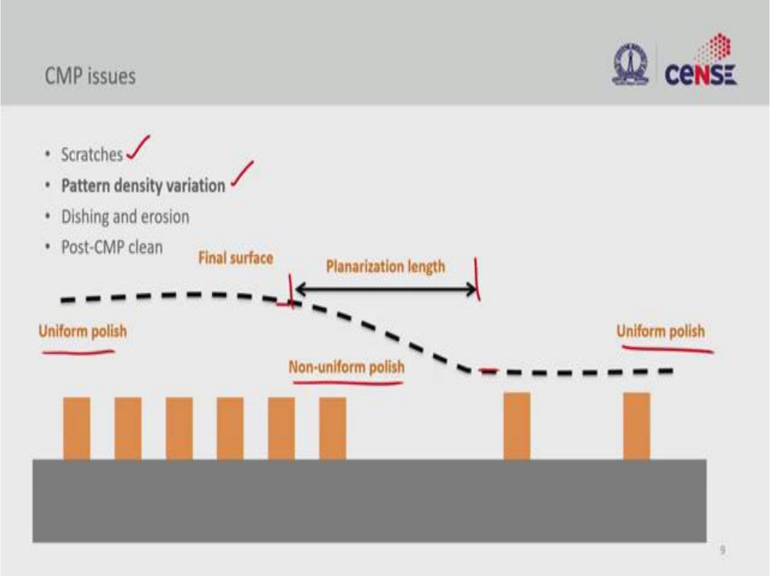
Polishing

- Surface roughness reduction

Before CMP After CMP

The other reason we use CMP is to make the surface smooth. So, the as deposited oxide layer, in this case, has 0.7 nanometer rms roughness and then after polishing we get 0.12 nanometer. So, making the surface smoother is also an important application of CMP process.

(Refer Slide Time: 22:20)



CMP issues

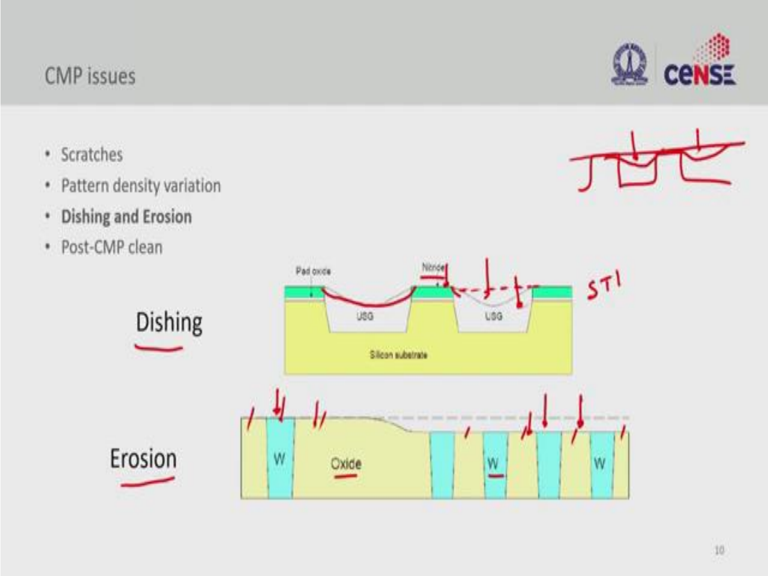
- Scratches
- **Pattern density variation**
- Dishing and erosion
- Post-CMP clean

In the following few slides, we are going to look at the issues related to CMP. The first issue we already saw earlier about scratch. One should be very careful about particle size and the pad's nature to create no scratches. So, we need a very soft surface on the pad in order to get a smooth surface.

The next thing is pattern density-dependent flatness. This is very typical of a density-dependent CMP process. So, we have a large thickness variation after polishing, which comes as an intermediate stage in the CMP process that we call local planarization. In this process, we have a uniform polish on both sides, but the only problem is that we have a non-uniform part between the two uniform layers. The horizontal distance between the two locally planarized surfaces is called planarization length, and this strongly depends on pattern density.

So, when we are doing CMP, make sure that pattern density is uniform. If we have an isolated line and then a very high dense pattern, we cannot expect both the regions to become flat right without losing material from one side or the other side. So, the isolated line should have dummy structures. It is very common in CMOS process to use dummies which we will see in a design for manufacturability. One of the reasons why we add additional features is to match the pattern density.

(Refer Slide Time: 24:17)



CMP issues

- Scratches
- Pattern density variation
- **Dishing and Erosion**
- Post-CMP clean

The next issue is dishing and erosion. Dishing is nothing but a prolonged etch beyond necessary. So, ideally, we want to have a flat surface finish, but if we do not stop it and keep on polishing the layers, we will have dishing while the other part is unetched.

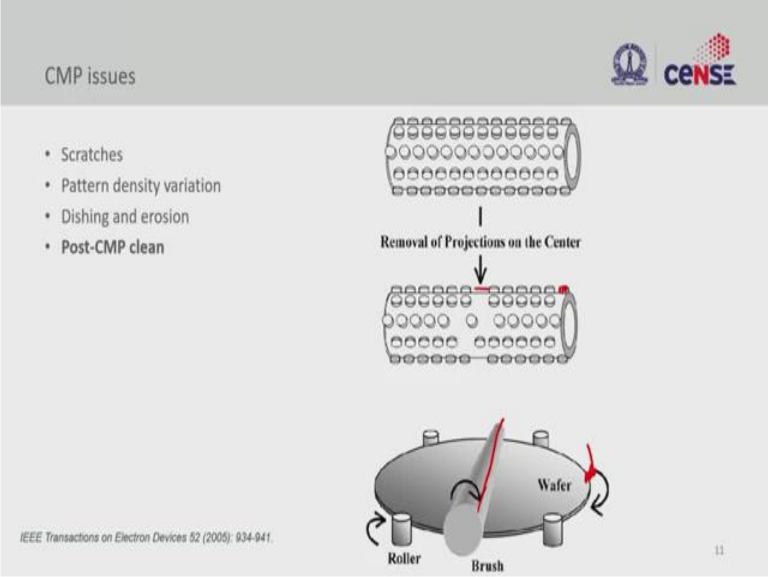
Here we see a Shallow Trench Isolation, an STI process in CMOS where we create an isolated trench using silicon nitride. So, when we are polishing, the soft material will etch quicker than hardened material like nitride. In this case, instead of having a flat layer, we will remove material much quicker in the soft region and, this is what we call dishing. So, do not polish beyond the required limit. So, if we do that, we create this local dishing, which can be a real issue.

The next thing is erosion. So, erosion typically happens when we have density variation. So, the density variation would result in aggressive removal of material in some places and sparse removal in other places. In this case, we look at the CMP of a contact, tungsten plugs in the background of oxide. Tungsten is a smooth material. In a highly dense tungsten-filled region, there will be less silicon dioxide. So, both tungsten and oxide will be removed quickly than in the region containing isolated tungsten contact.

The reason for that is that the fill factor of tungsten in densely filled regions is more than in an isolated region. When we are doing that, we will move the flat profile lower than an isolated feature. This is called erosion, and this erosion can happen when we have different material densities, and one is removing at a much faster rate compared to the other.

So, dishing and erosion are two different issues we face, and this, in turn, results in local and global flatness issues. If we look at dishing, it creates local topography variation, while erosion creates global flatness variation.

(Refer Slide Time: 27:31)



CMP issues

- Scratches
- Pattern density variation
- Dishing and erosion
- **Post-CMP clean**

IEEE Transactions on Electron Devices 52 (2005): 934-941.

The final issue here is about post-CMP cleaning. It is a very crucial step because we use abrasive particles and these abrasive particles need to be removed after doing CMP. If these particles are going to be adsorbed onto the surface, then we will have difficulty in post-processing after CMP. Though we might have a flat and smooth surface, the particles sitting on the surface can create particle contamination for subsequent processes.

The regular practice is to use brush-type cleaners. We have a cylindrical rotating brush, and the wafer is also rotating in the opposite direction. This brush with smooth bristles will physically remove any particles that are sitting on the surface. And, this might look very physical process, but when the particles are adsorbed onto the surface, we need force to take them out. So, this is the best way to do it.

The other way to do this is by using ultrasonic cleaning, which will also shake these particles off the wafer. But the force required to remove these particles could be much higher than what the ultrasonic bath could give. So, one needs to apply this brush cleaning to remove all the particles of the surface fully.

So, with this, we have come to an end of this chemical mechanical polishing lecture where we saw the fundamental nature of the polishing process, the essential components required in polishing, how the process will affect the density difference, and associated issues. Moreover, we saw how to maintain the quality of CMP process by adjusting the nature of the pad, the slurry and so on.

Finally, CMP is a dirty process because we are using large particles. Usually, we want to avoid any kind of particle contamination during our process. However, here, we take help from the abrasive particles to remove the material. But after CMP, we want to remove all those particles off our wafer, which is crucial. So, it is not just polishing but also cleaning; post-CMP clean is equally important, and one needs to take care of this cleaning process.