

Reactors, productivity and quality metrics

(2 hour set, 2023)

sami.franssila@aalto.fi

Table of contents

Plasma CVD
Rate limiting step
Other CVD reactors/operating modes
ALD
Plasma-ALD
Noble metal ALD

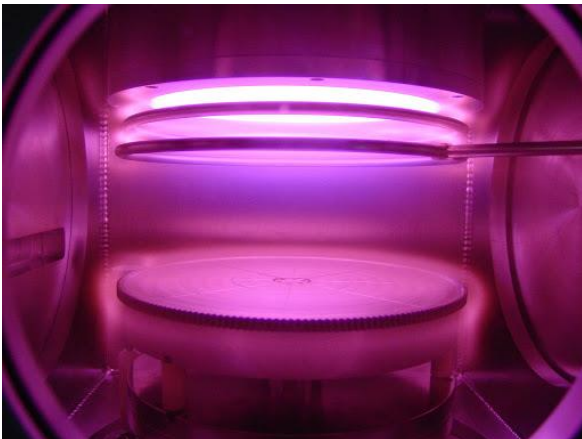
}
} 1st hour
}

Reactor design generic issues
Reactor figures of merit
Film quality measures
Film characterization measurements
Post deposition anneal of films

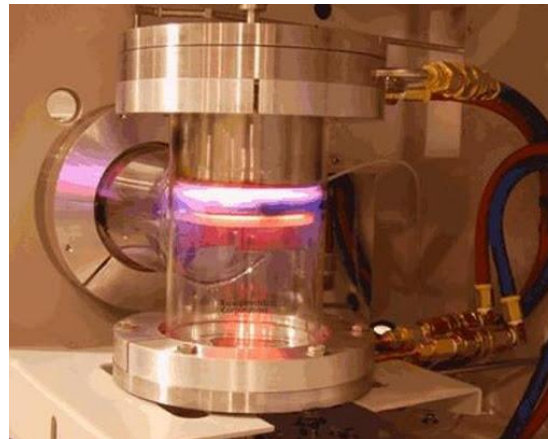
}
} 2nd hour
}

PECVD: Plasma Enhanced CVD

- Plasma creates active specie (radicals, ions)
- No need of high temperatures because of plasma
- More parameters to work with (power, freq, pulses...)
- Usually single wafer reactors
- Need high rates (1-10 nm/s) (thermal 10% of this)

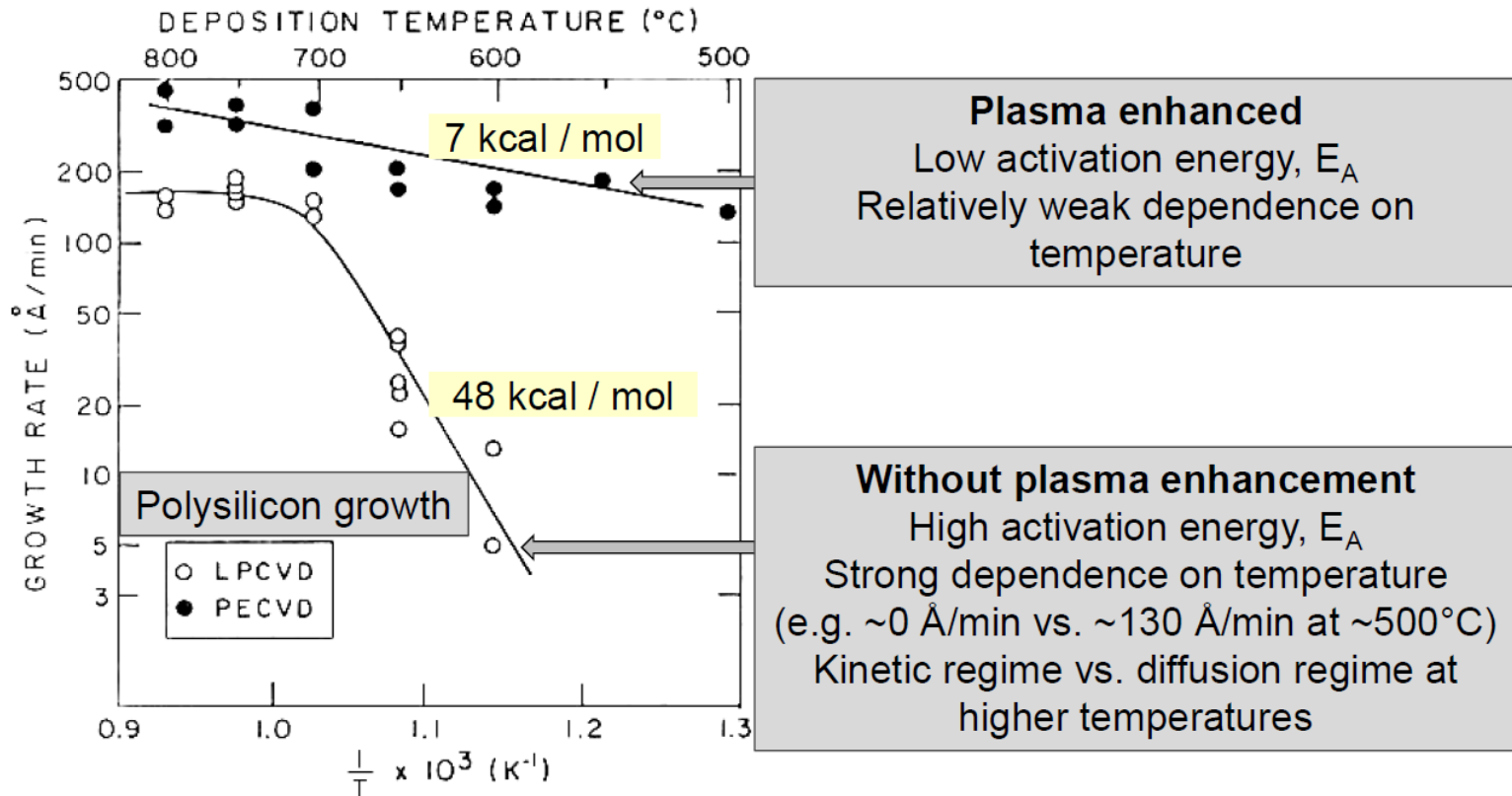


<http://www.nanomaster.com/pecvd.html>



<http://thinfilmscience.com/en-US/PECVD.aspx>

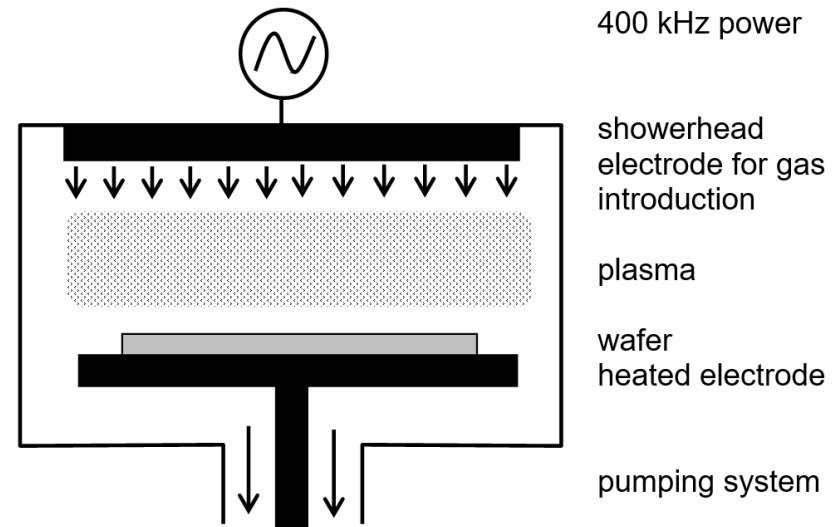
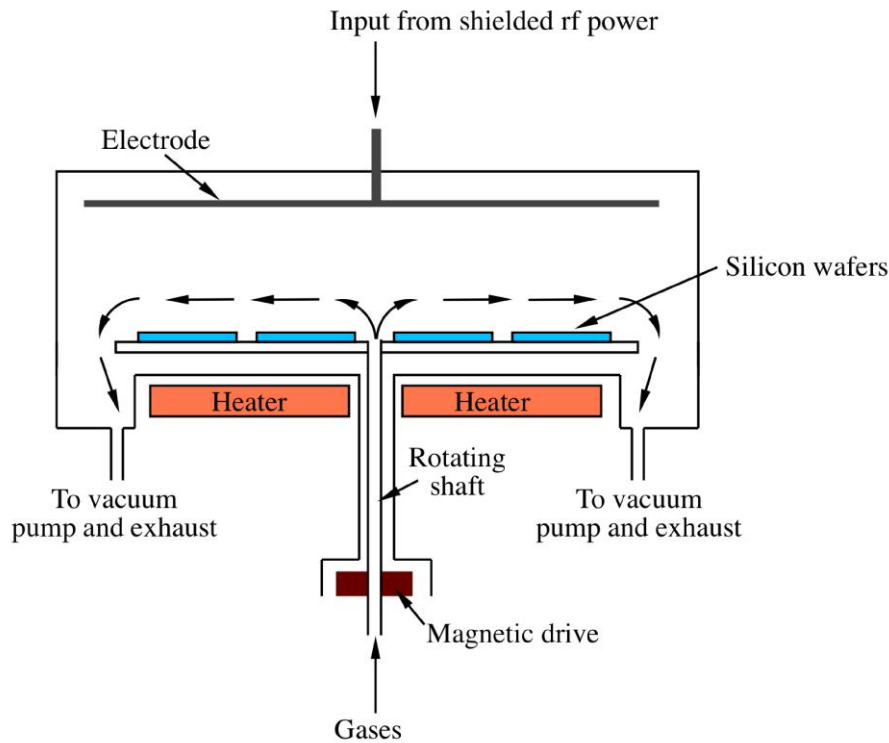
Thermal vs. Plasma-CVD



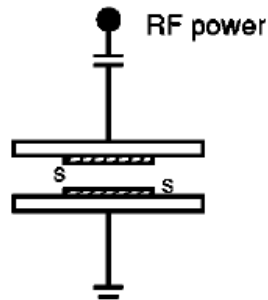
Source: Hajjar et al, J. Electronic Mat., 15, 279 (1986)

Radial flow vs. showerhead

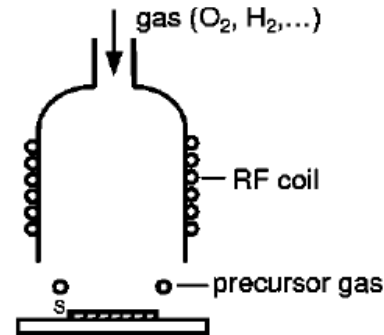
In order to work in mass transport limited regime for fast deposition, gas introduction has to ensure uniform gas distribution.



Plasma pros and cons



(a) Parallel plate RF PECVD



(b) Remote RF PECVD

Increase in power

- More ions and radicals generated
- More ion bombardment
- Atoms kicked off the surface

POSITIVE: loosely bound specie detached

NEGATIVE: depo rate goes down

NEGATIVE: ions may damage already deposited film

Increase in RF coil power

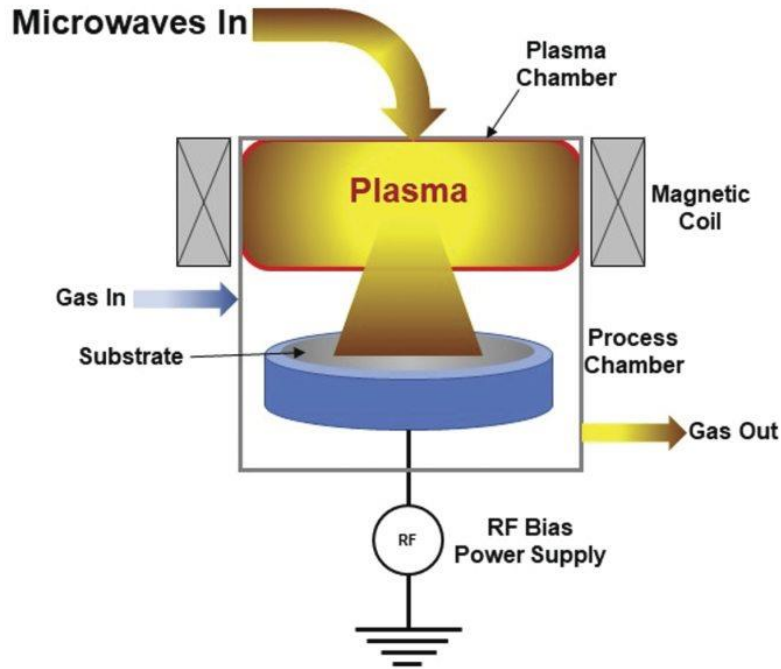
- More ions and radicals
- Depo rate up

No increase in ion bombardment on the wafer, as it is not connected to RF coil power.

Radicals reach wafer by diffusion

- no directionality effect

HDP = High Density Plasma



2 power sources:

GHz-source to generate plasma;
MHz-source to bias the wafer.

Increasing GHz power does not mean more ion bombardment on the wafer.

HDP: high density plasma,
 10^{13} ions/cm³ vs. 10^{10} - 10^{11} /cm³
for RF plasmas vs. 10^{15} /cm³
neutrals @100 Torr pressure

**Maybe 1% of
gas is ionized**

New parameters by plasma

Plasma power:

-more ions and radicals generated → higher depo rate

Ion bombardment

-densification of film

-atoms kicked away → slower deposition rate

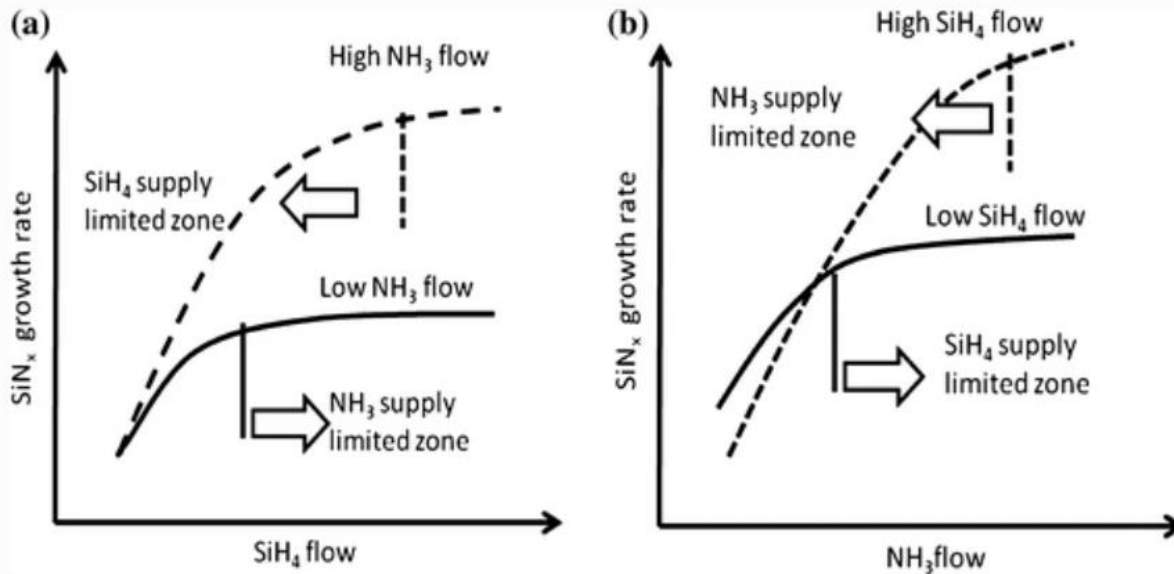
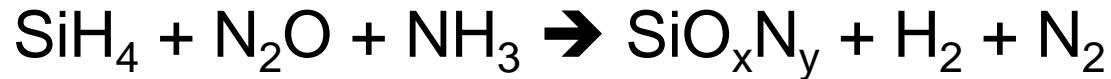
Pulsed operation:

-generation ions & radicals, but less ion bombardment

Frequency:

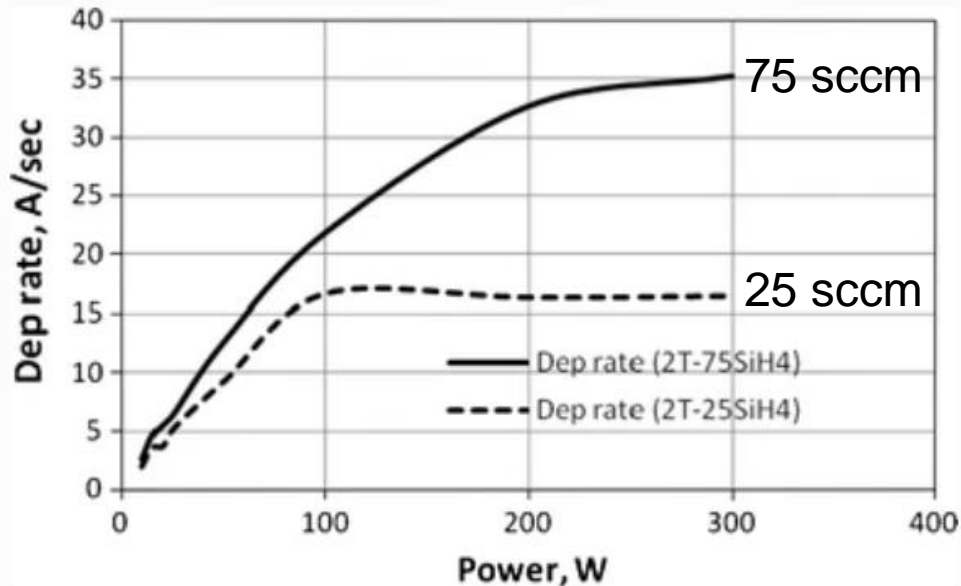
-usually fixed at 13.56 MHz or 2.45 GHz, but in theory...

Rate limiting step



“Increasing the NH₃ supply will further increase the SiN deposition rate until the Si-containing species concentration becomes the limiting factor again.”

Rate limiting step (2)



PECVD SiO_xN_y deposition rate at 2 Torr and 250 °C, 10 sccm N_2O flow, 50 sccm He, and 25 sccm NH_3 , diluted by N_2 (2000 sccm total flow). Two SiH_4 flow rates: 75 sccm (*solid line*) and 25 sccm (*dashed line*) were plotted here

”Assuming 25 sccm of SiH_4 is fully dissociated above 100 W RF power and the diffusion boundary layer is thin enough, SiO_xN_y growth is limited by the lack of a Si-containing precursor supply from the gas phase when O and N are over-saturated.

Thus the growth rate flattens out above 100 W RF power. Increasing the SiH_4 flow from 25 to 75 sccm pushes the growth rate saturation knee to a power level of above 300 W.”

PECVD deposition rate example

If silane (SiH_4) flow in a single wafer PECVD reactor is $5 \text{ cm}^3/\text{min}$ (also known as **sccm**, standard cubic centimeters/minute), under standard temperature and pressure, what is the theoretical maximum deposition rate of amorphous silicon on a 150 mm wafer ?

How many gas molecules flow into reactor, and much solid film thickness grows on wafer?

In the end, explain which processes reduce the deposition rate from the theoretical maximum, and give a guess of a practical deposition rate.

Silicon density 2.2 g/cm^3

Silicon atomic mass is 28 g/mol

Gas mole is 22.4 liters

$$\text{Incoming Si/min} = \frac{V_{\text{Silane}}}{V_{\text{STP}}} * M_{\text{Silicon}}$$

$$= \frac{5 \text{ cm}^3}{22.4 \text{ L/mol}} * 28 \text{ g/mol} = 6.25 * 10^{-4} \text{ g/min}$$

Silane gas flow /min



Silicon weight /min



alpha-Si volume /min



$$\text{Growth rate} = \frac{\alpha\text{-Si volume}}{\text{min}} / \text{Area}$$



$$\alpha\text{-Si volume /min} = \frac{M_{\alpha\text{-Si/min}}}{\rho_{\alpha\text{-Si}}} =$$

$$\frac{6,25 * 10^{-4} \text{ g/min}}{2,2 \text{ g/cm}^3} = 2,84 * 10^{-4} \text{ cm}^3/\text{min}$$



$$\text{Z Rate} = \frac{\alpha\text{-Si volume/min}}{\text{Wafer Area}}$$

$$= \frac{2,84 * 10^{-4} \text{ cm}^3/\text{min}}{\pi * (7,5)^2 \text{ cm}^2} = 160 \text{ nm/min}$$

PECVD gas utilization

In theory, 160 nm/min deposition rate, but:

- some gas is pumped directly from inlet to outlet
- sticking coefficient $\ll 1$ (not all molecules reaching the surface stay at surface)
- some adsorbed molecules desorb before reacting
- unwanted reactions happen, and consume some gas
- some deposition occurs on electrode and on chamber walls, in addition to wafer itself

In practise: ***10 nm/min < rate < 100 nm/min***

Nitride comparison (1)

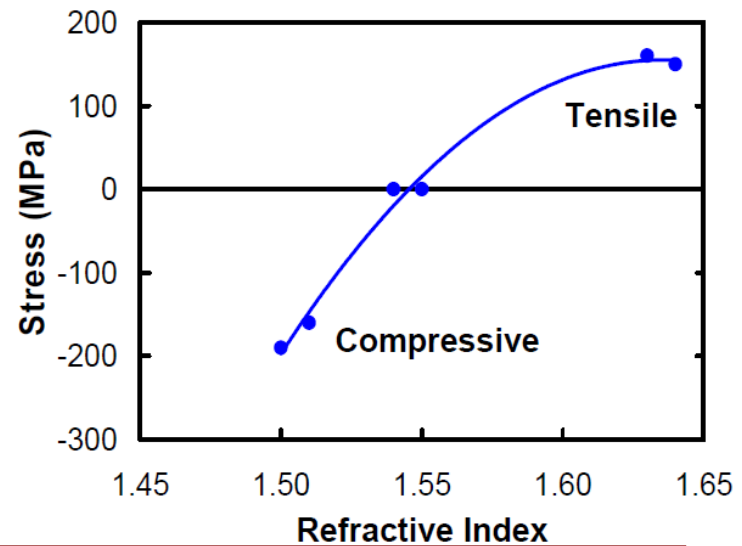
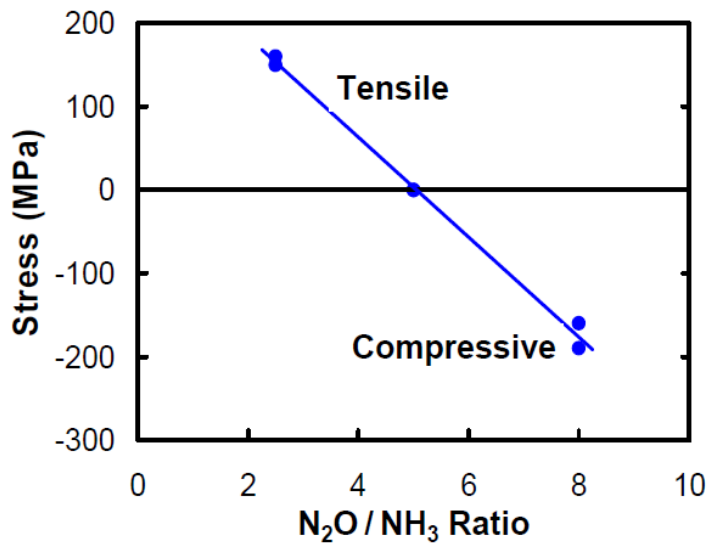
	plasma PECVD	thermal LPCVD	HDP CVD
Deposition rate (Å/min)	1800	300	960
Refractive index	1.985	2.003	1.990
Stress (dyn/cm ²)	-1.5 E 9	+1.0 E 10 ^a	-3.8 E 9
Wet-etch rate in 85% H ₃ PO ₄ (Å/min)	373	57	69
Wet-etch rate in 15:1 BHF (Å/min)	46.0	3.4	2.3

Tradeoff: high deposition rate leads to non-dense film, which is rapidly etched.

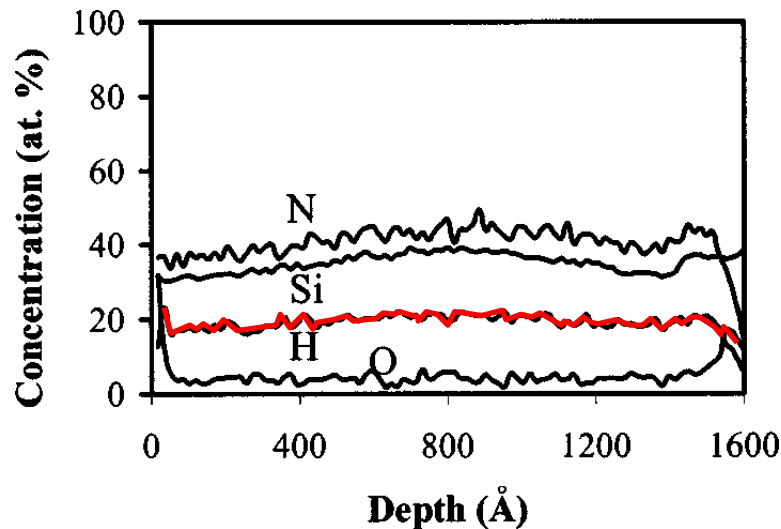
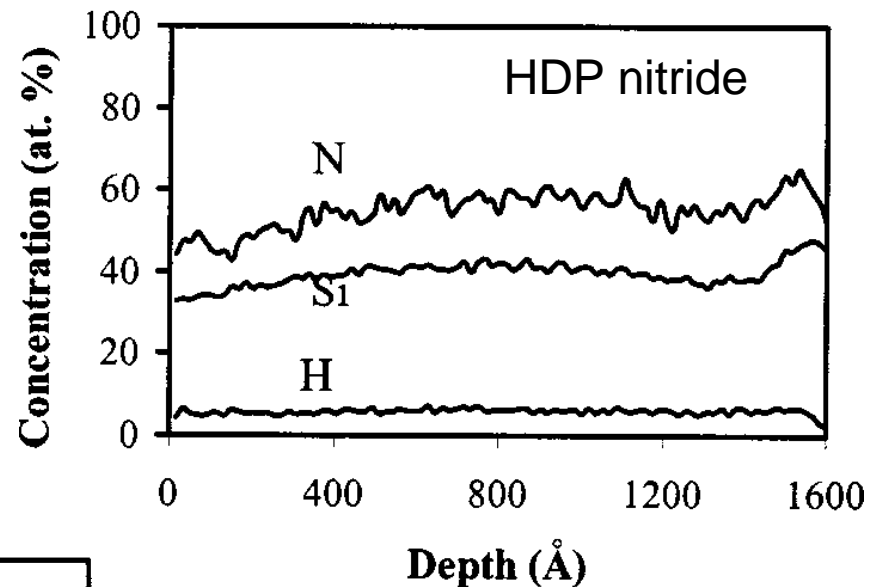
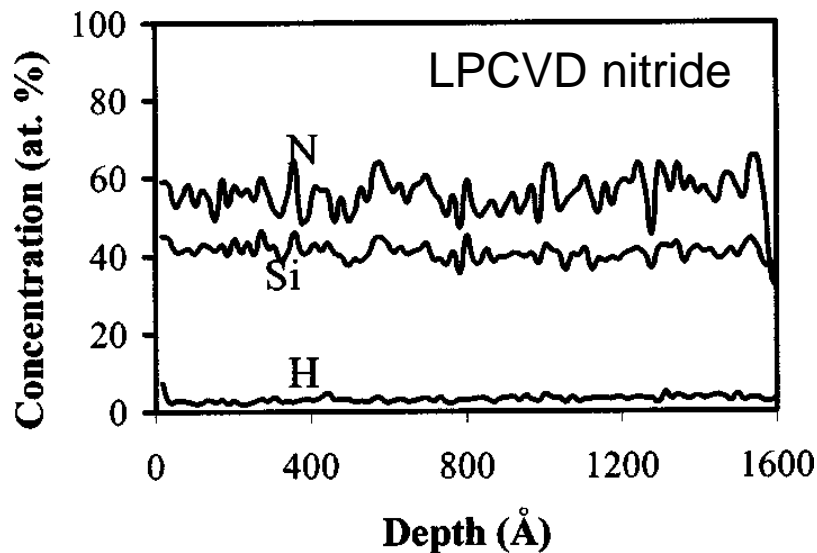
Coupled responses !

- Silicon oxynitride, SiO_yN_x formed from (SiH_4 , NH_3 , N_2O , N_2)
- Zero stress achieved by simple adjustment of $\text{N}_2\text{O}/\text{NH}_3$ ratio

Yes, but refractive index changed, too !

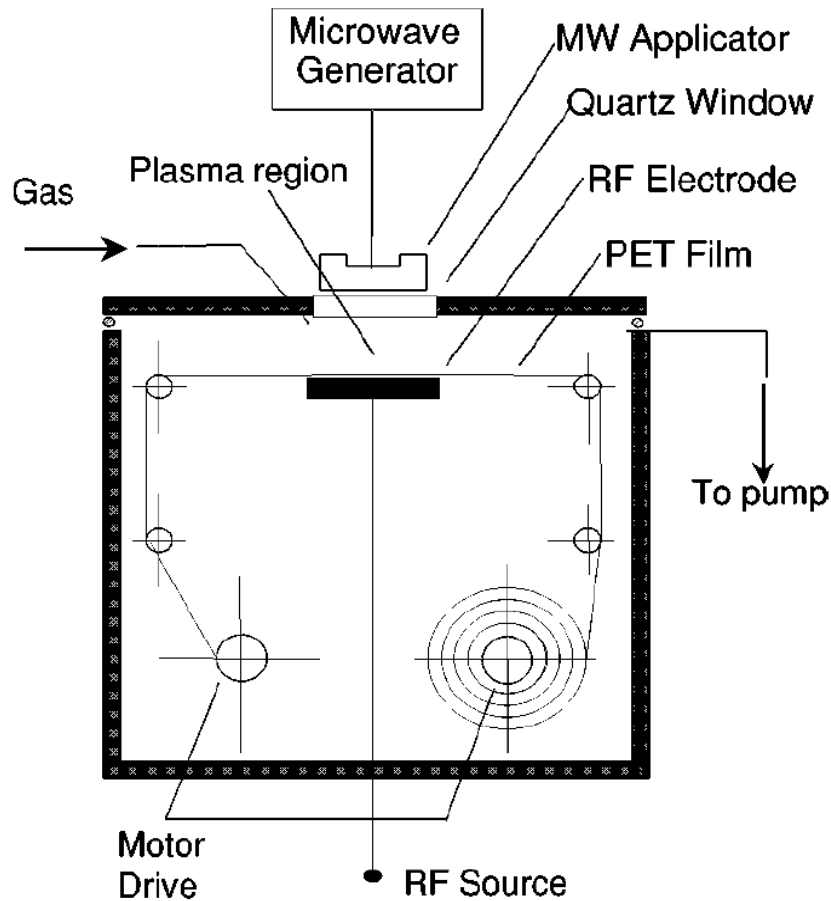


Hydrogen in nitride



20 at% hydrogen

Roll-to-roll PECVD

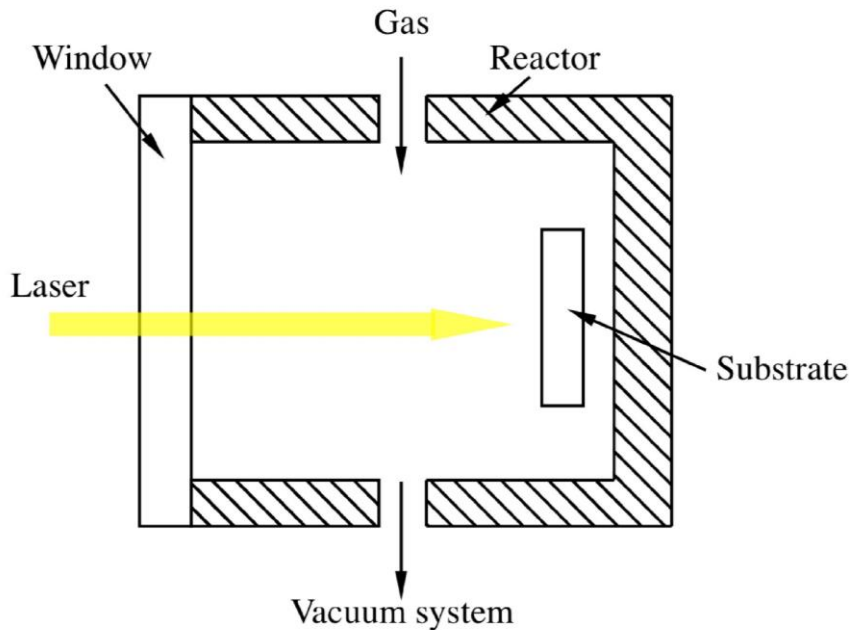


For productivity:

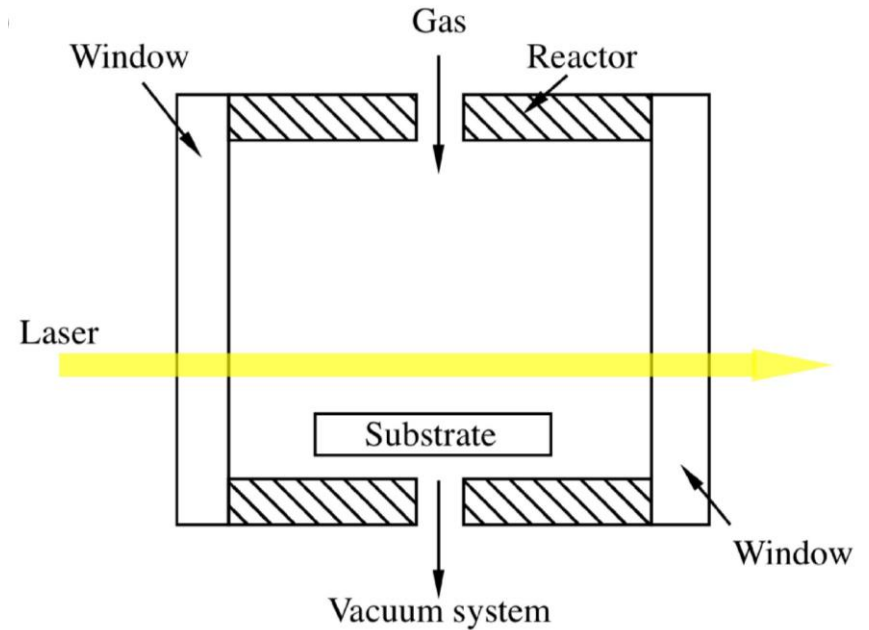
30 cm wide foil coated

Can be 100's of meters long

Laser-assisted CVD/ALD

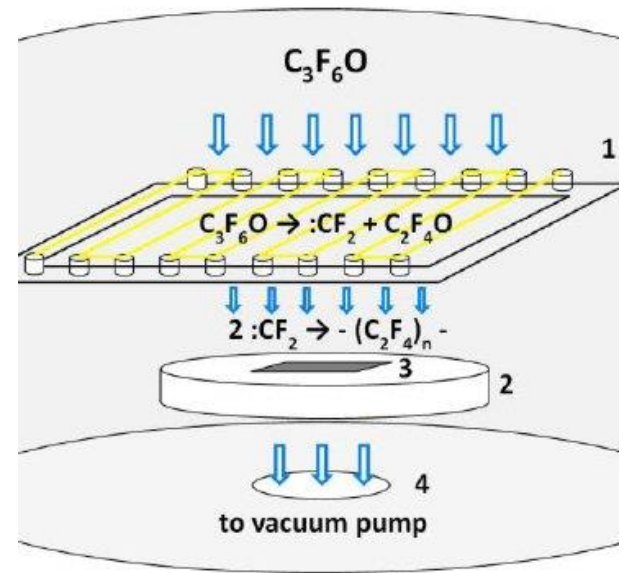
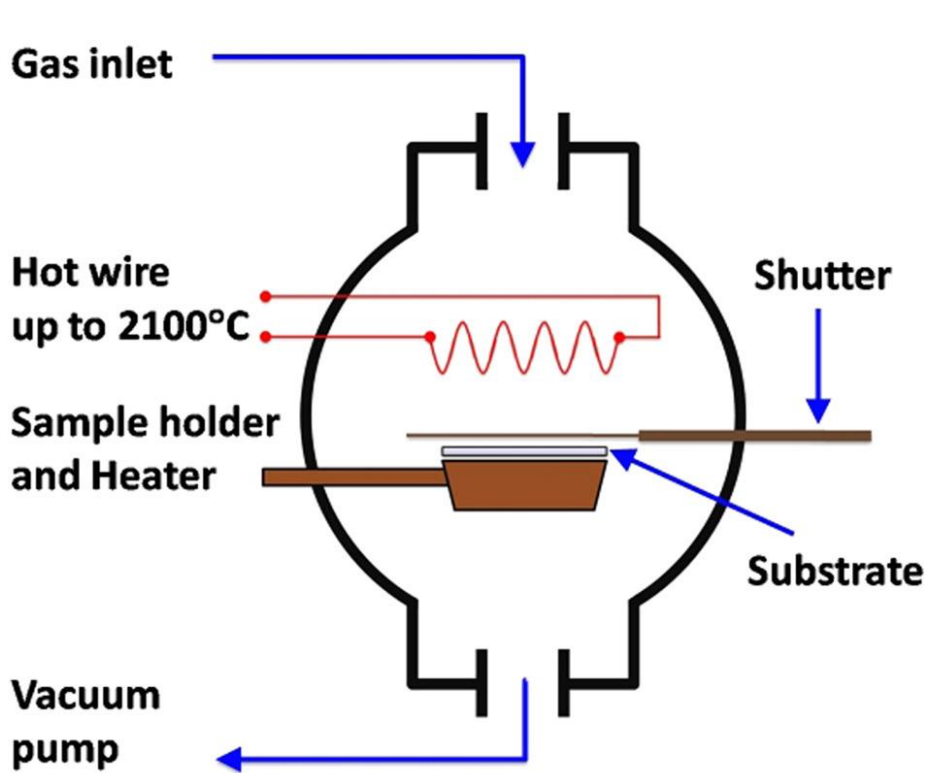


Laser as a heating tool
(photothermal)



Laser as gas excitation tool
(photochemical)

Hot wire CVD



Thermal activation of precursors, but substrate temperature can be kept low.

HW-CVD SiC and diamond

SiC is usually made from $\text{SiH}_4/\text{CH}_4/\text{H}_2$ mixtures.

Mesh temperature 1600°C and the substrate temperature $700\text{--}800^\circ\text{C}$.

At the end of the Si_xC_y composition scale is diamond-like carbon and diamond, which can readily be made by HWCVD using CH_4 and H_2 .

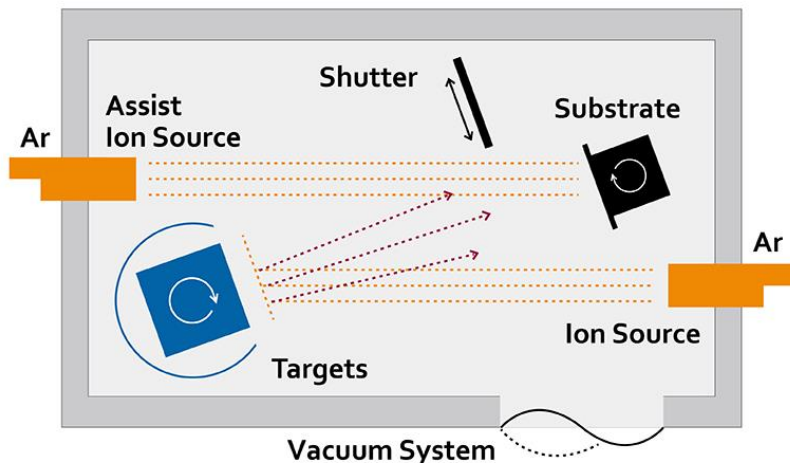
Substrate temperature increased to about 850°C .
The filament temperature was about 2000°C .

Shutter: beam vs. diffusion

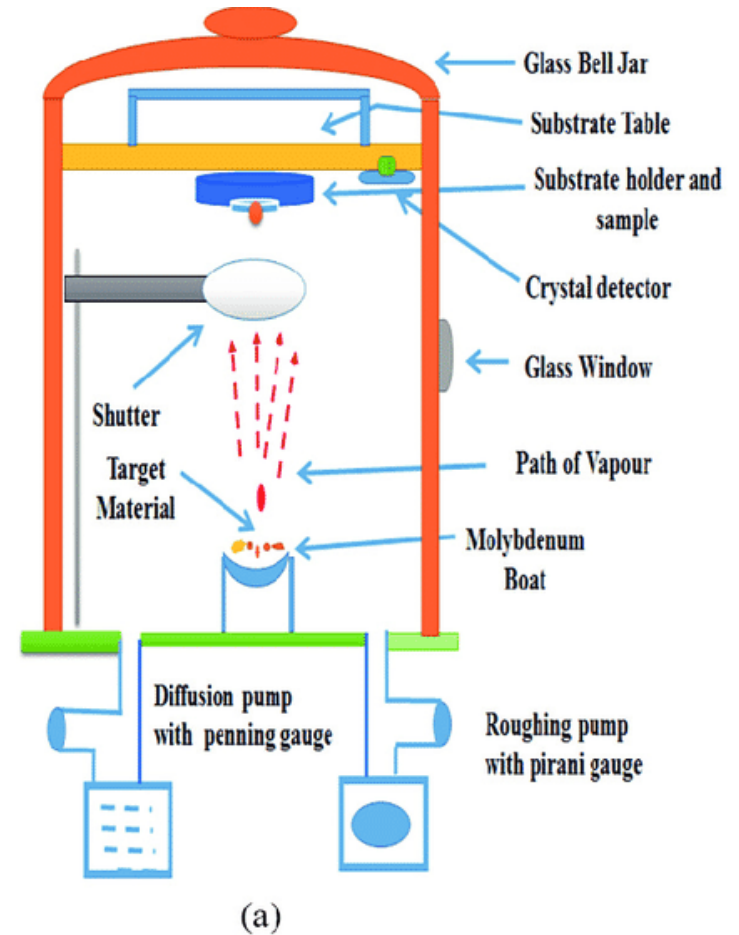
Shutters are generally used to block a beam of atoms; not diffusion of gases.

Shutter is excellent for pre-treatments of substrate or target before actual deposition in PVD.

Shutter acts as heat shield in HW-CVD.

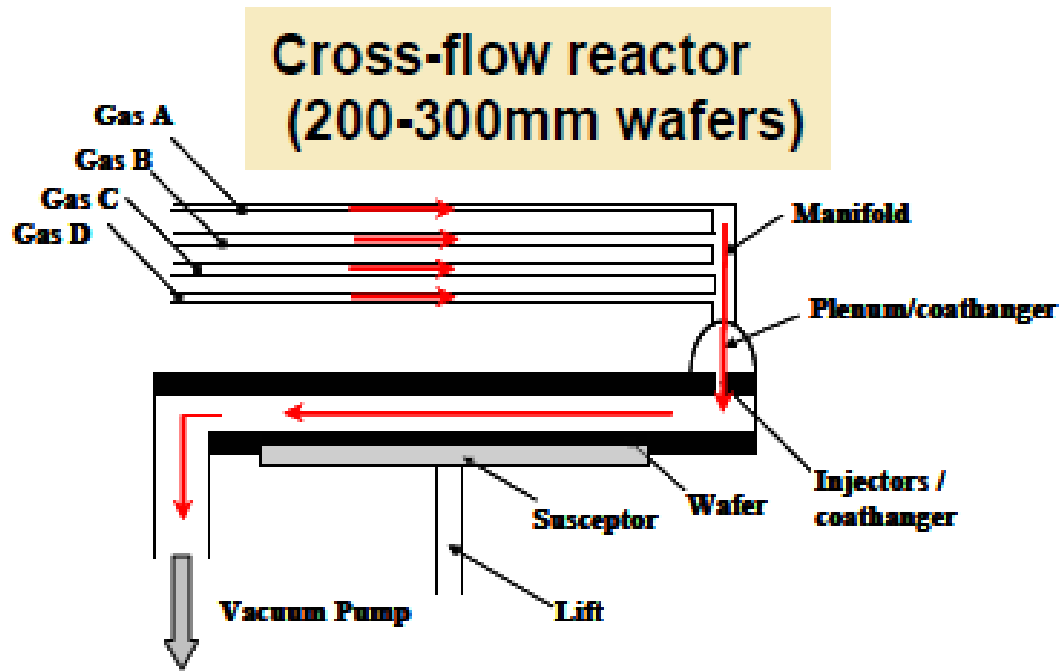


<https://www.dentonvacuum.com/what-is-ion-beam-deposition/>



Ganaie et al: Study of Morphological, Electrical and Optical behaviour of Amorphous Chalcogenide Semiconductor, 2020

ALD reactors (1)

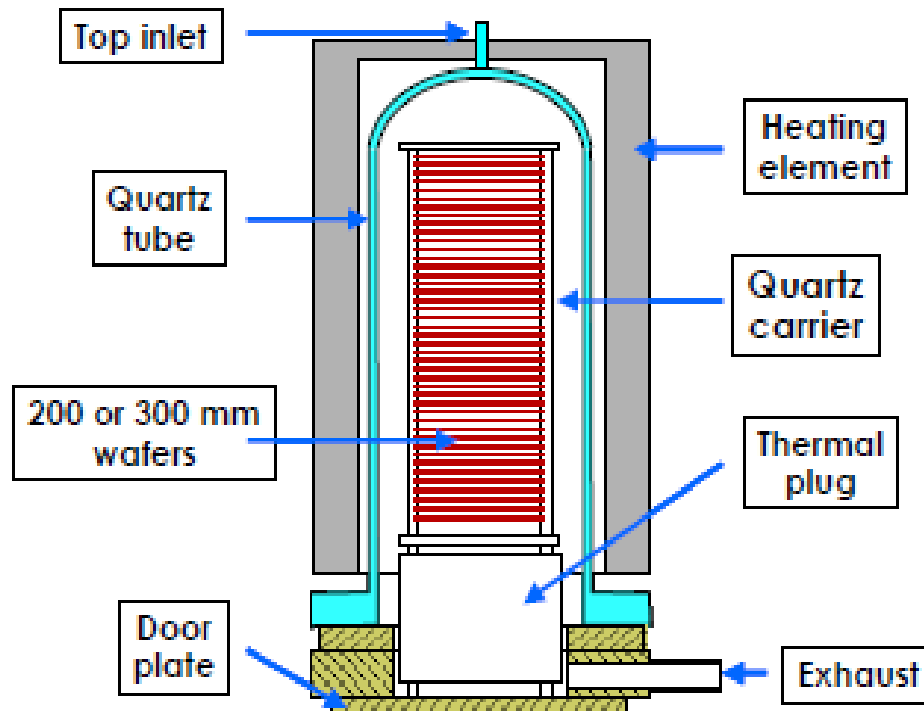


Continuous flow
typical.

Stopped flow:
introduce precursor
pulse, stop pumping;
wait for (expensive)
precursor to diffuse
(into deep cavities).

ALD reactors (2)

Batch reactor (200-300mm wafers)



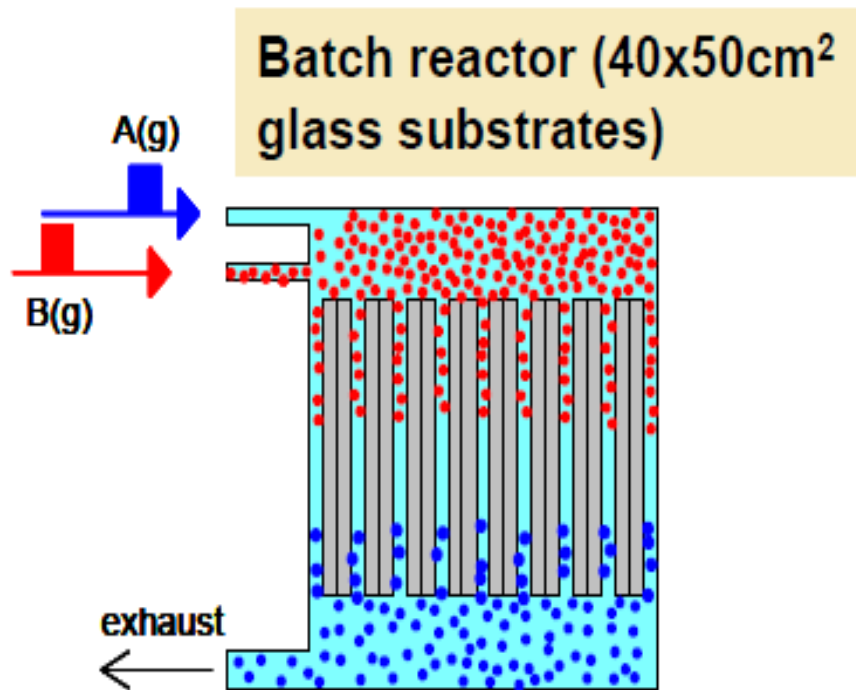
Modern batch reactors, both ALD and CVD, are vertical.

This saves precious cleanroom floor space.

Vertical reactor



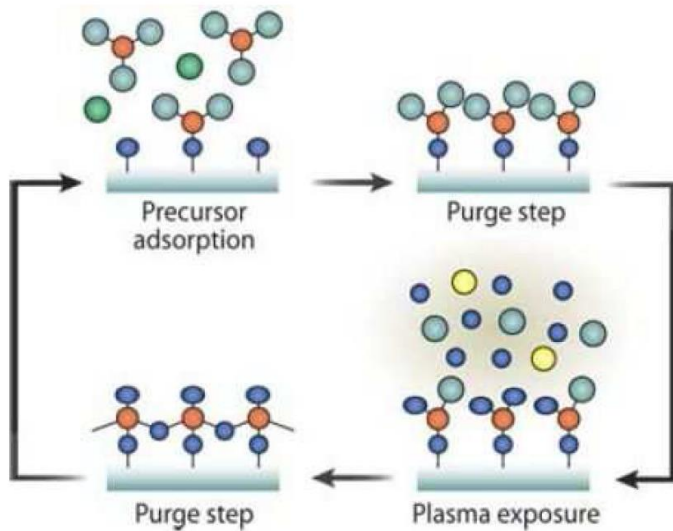
ALD reactors (3)



Because ALD is the ultimate surface reaction controlled deposition technique, we can rely on precursor diffusion in tight spaces between substrates.

For 10 nm thick Al₂O₃ passivation film, 3000 WPH has been shown (500 wafers/batch; 10 min each, deposition time 5 min (ca. 100 cycles = 400 gas pulses/300 seconds)).

Plasma ALD (PEALD)



Precursors can be excited at lower temperature;

Larger choice of precursors; e.g. N_2O as oxidant

Ion bombardment modifies (and damages) film;

Directionality is introduced and ALD excellent conformality is compromised.

Plasma can be ON part of the time during the oxidant pulse; and not at all during metal pulse.

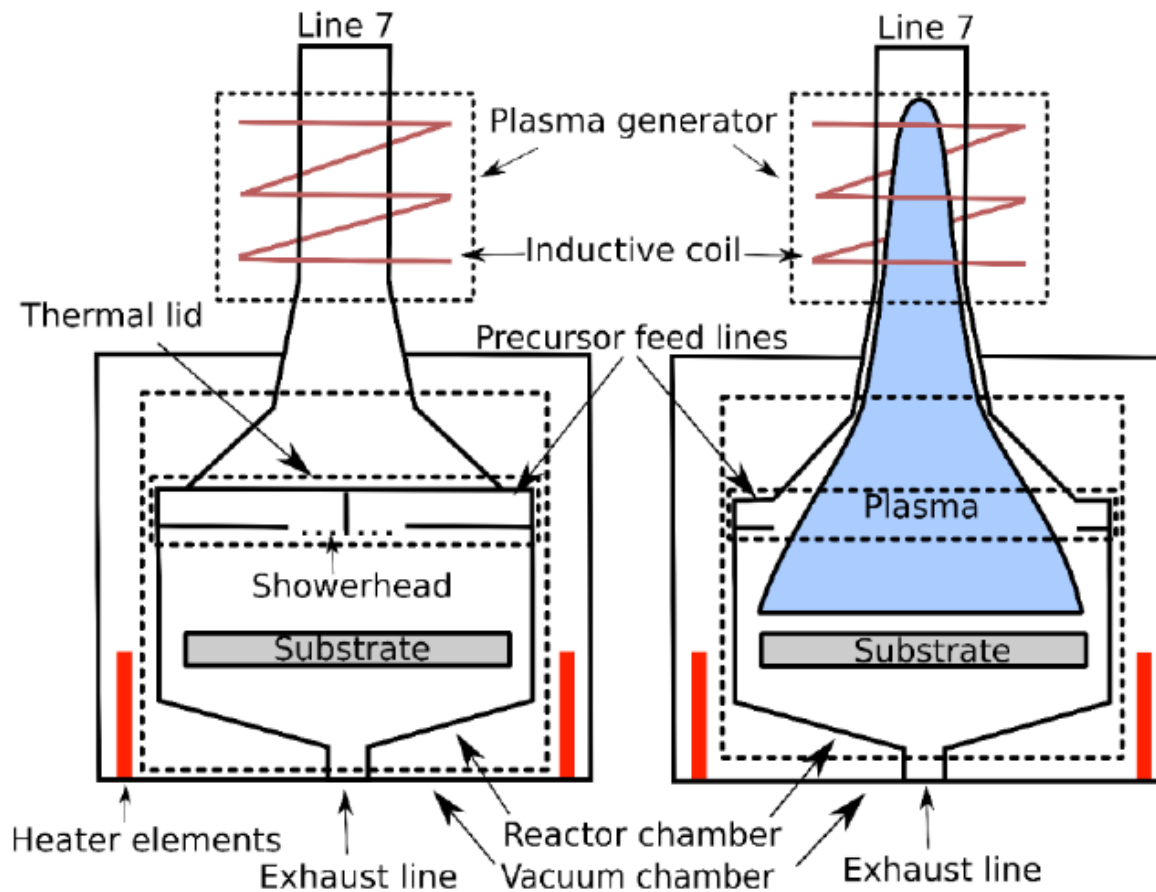
Gyu-Jin Choi, Seong Keun Kim,^a Seok-Jun Won, Hyeong Joon Kim, and Cheol Seong Hwang^{*,z}

Micronova ALD-tools

Picosun R200 Advanced

Thermal mode

Plasma enabled mode

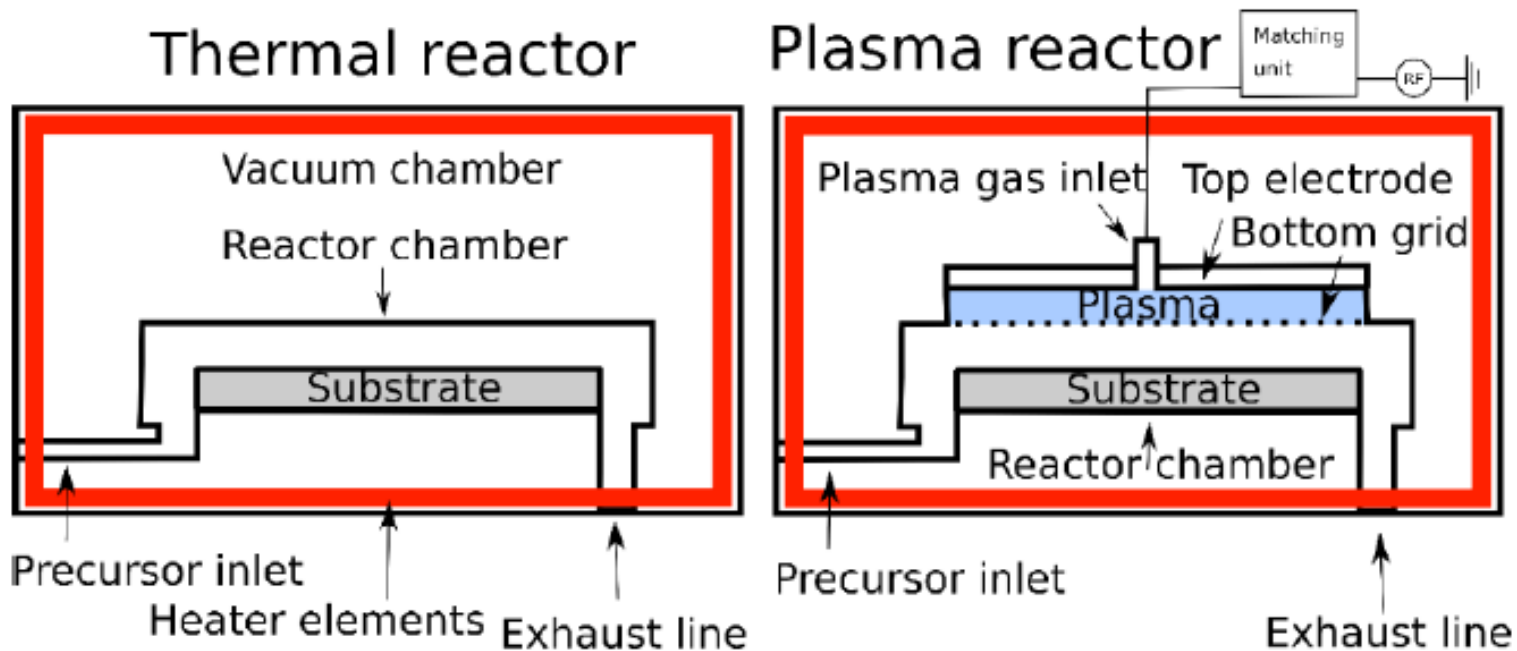


The same tool; yes, but you must make hardware changes when you switch modes.

2.45 GHz

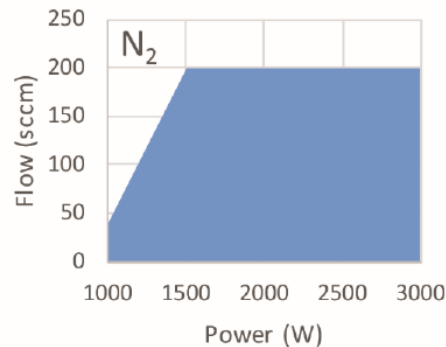
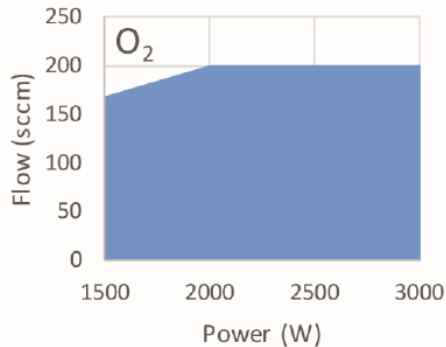
Micronova ALD-tools (2)

Beneq TFS500

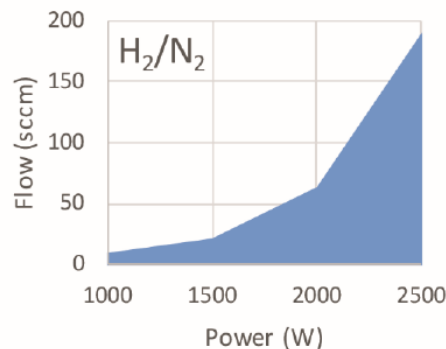
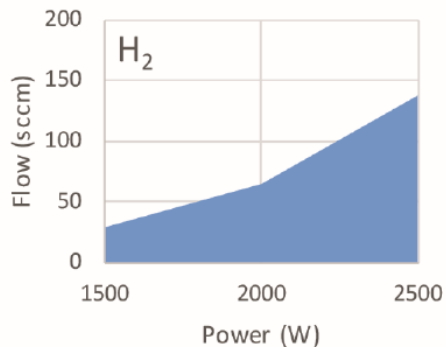


Capacitively coupled plasma (CCP) with showerhead and the freedom to use direct or remote mode with the same plasma head. 13.6 MHz RF power.

Gas flow rate $\leftarrow \rightarrow$ power

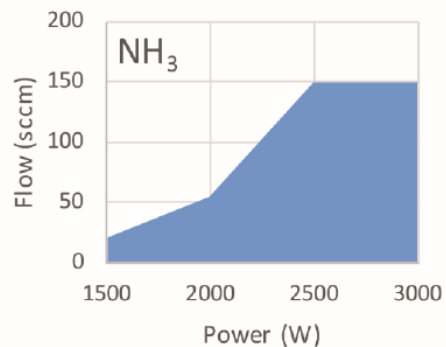


Not all combination of flow and power are available !

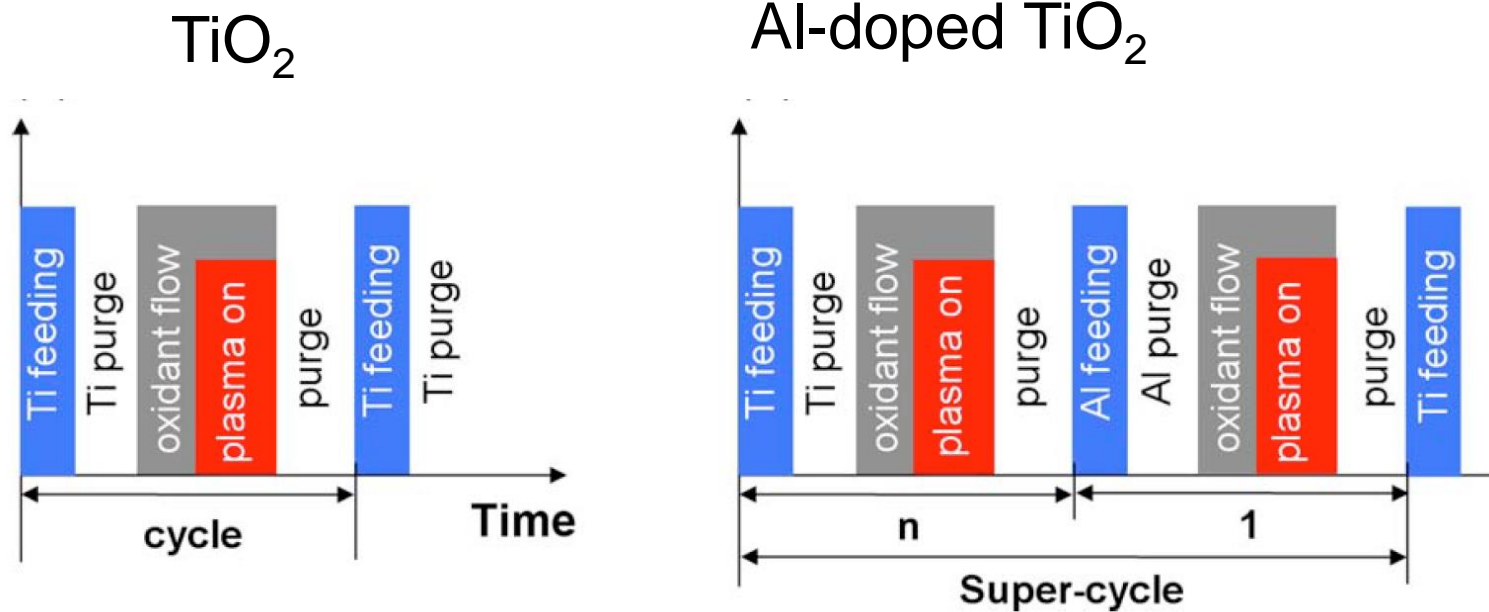


Plasma ignition requires certain pressure, and it may not be reached with given pumps.

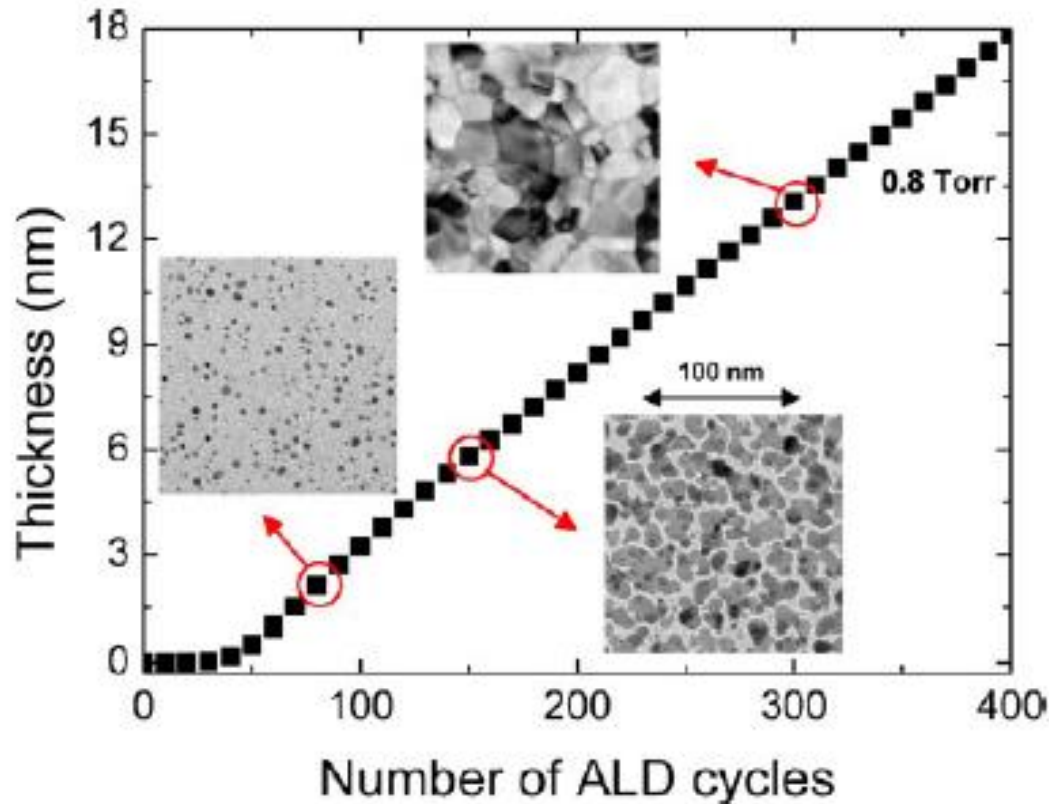
Picosun R200 reactor.



Supercycle PEALD

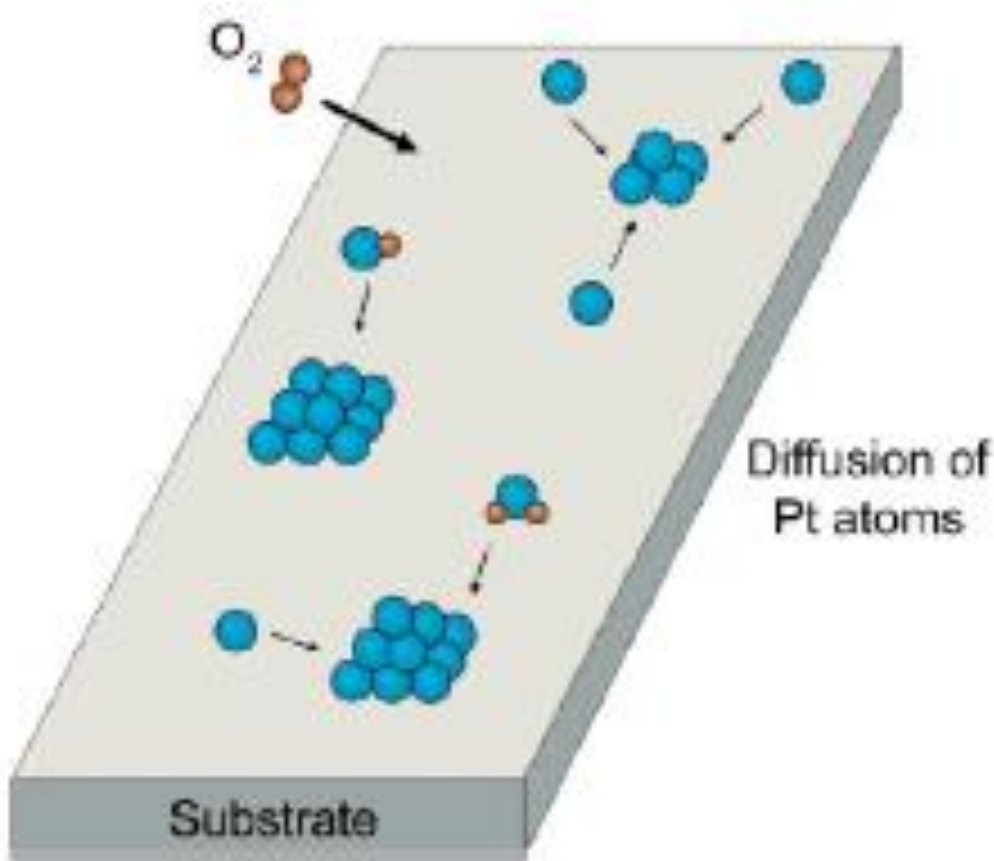


Particle vs. film deposition



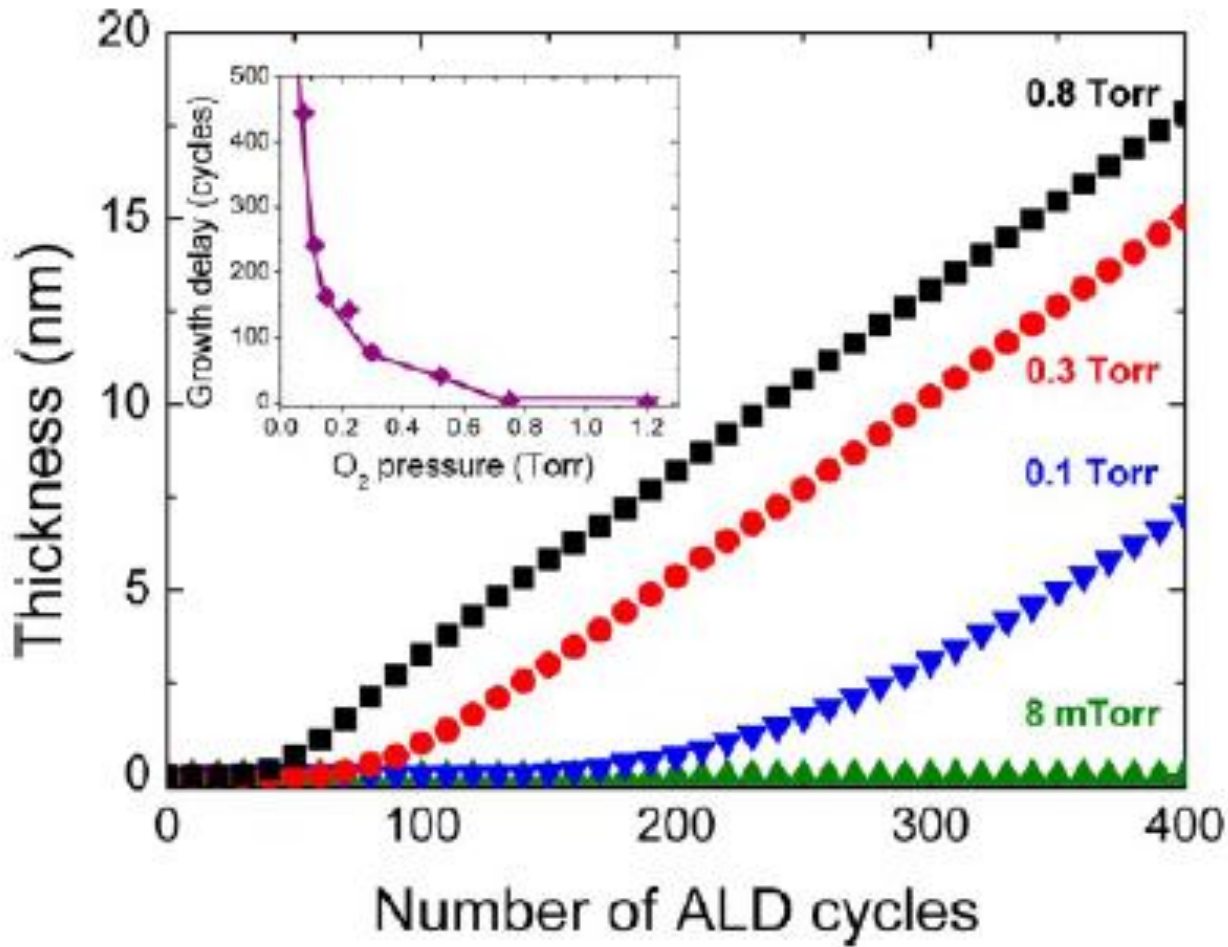
“TEM images in the figure illustrate that Pt ALD nucleation evolves from island growth, via island coalescence, to film closure.”

Pt: oxygen assisted diffusion



“O₂ exposure (when being sufficient) enhances the diffusion of single Pt atoms over the oxide surface leading to aggregation of Pt in metal clusters. The particle ripening (i.e., the formation of clusters) can be employed to prepare nanoparticles or, when increasing the number of cycles, to prepare closed films.”

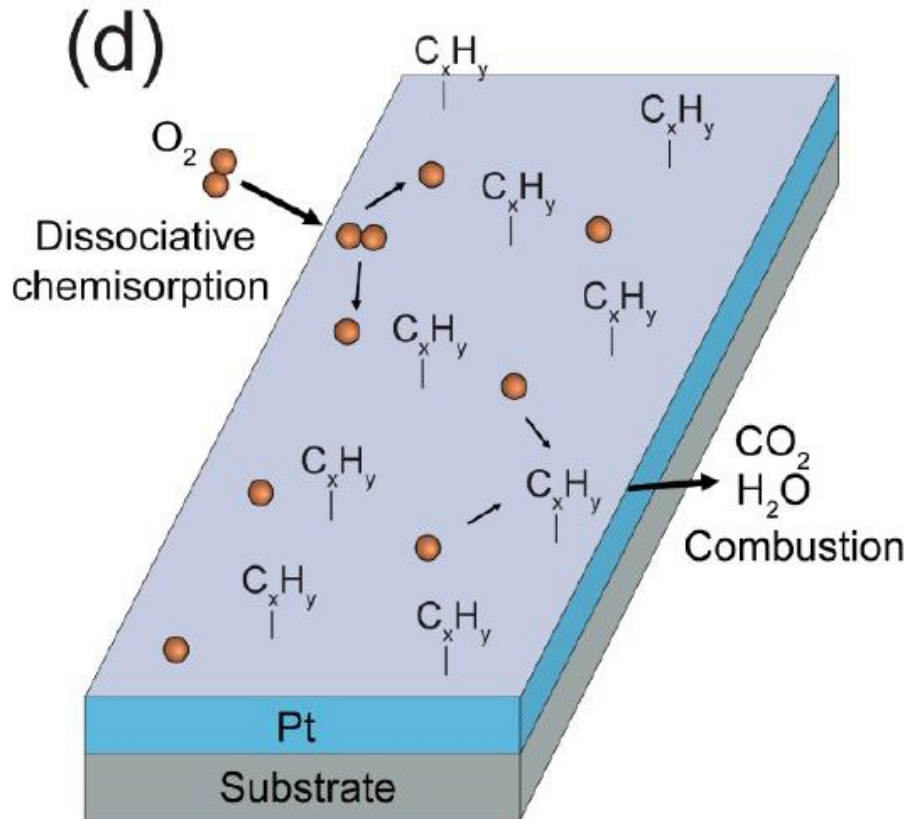
Noble metal ALD: initiation lag



“Thickness as a function of the number of ALD cycles for different O₂ pressures and a 10 s pulse time.

In the inset the growth delay deduced from the nucleation curves is presented as a function of the O₂ pressure.”

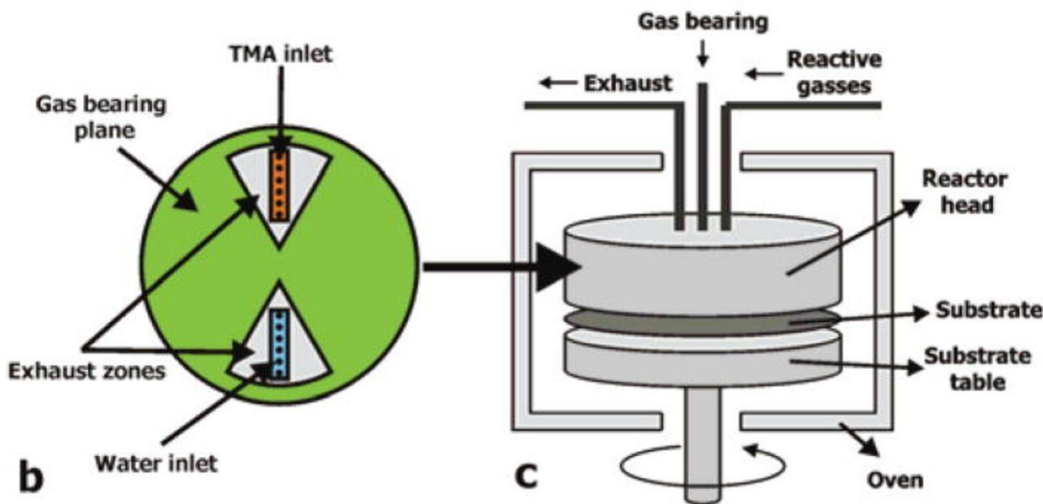
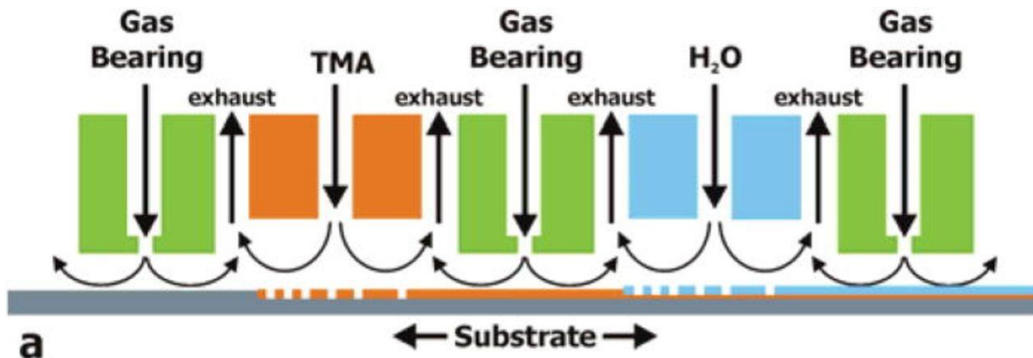
Pt: oxygen improves purity



O_2 dissociatively chemisorbs on the Pt surface, allowing for subsequent combustion of the ligands remaining from the Pt precursor step.

Organic ligands are “burned” away → less carbon residues.

Spatial ALD



Wafer is moving past static gas nozzles.

Linear, back-and-forth and rotating versions exist.

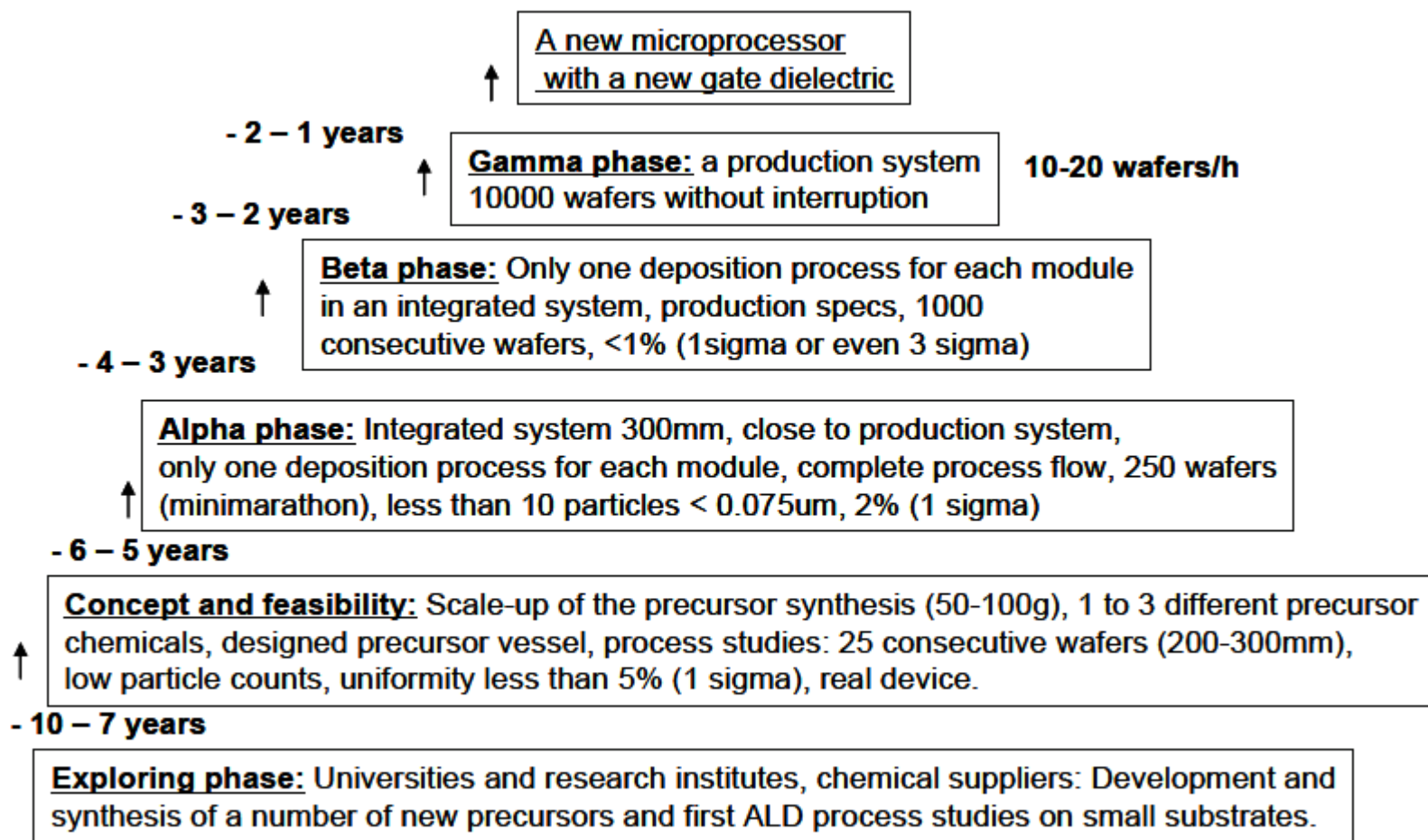
This was an original idea of Tuomo Suntola in 1970's; reinvented in 2000's.

Spatial ALD reactors



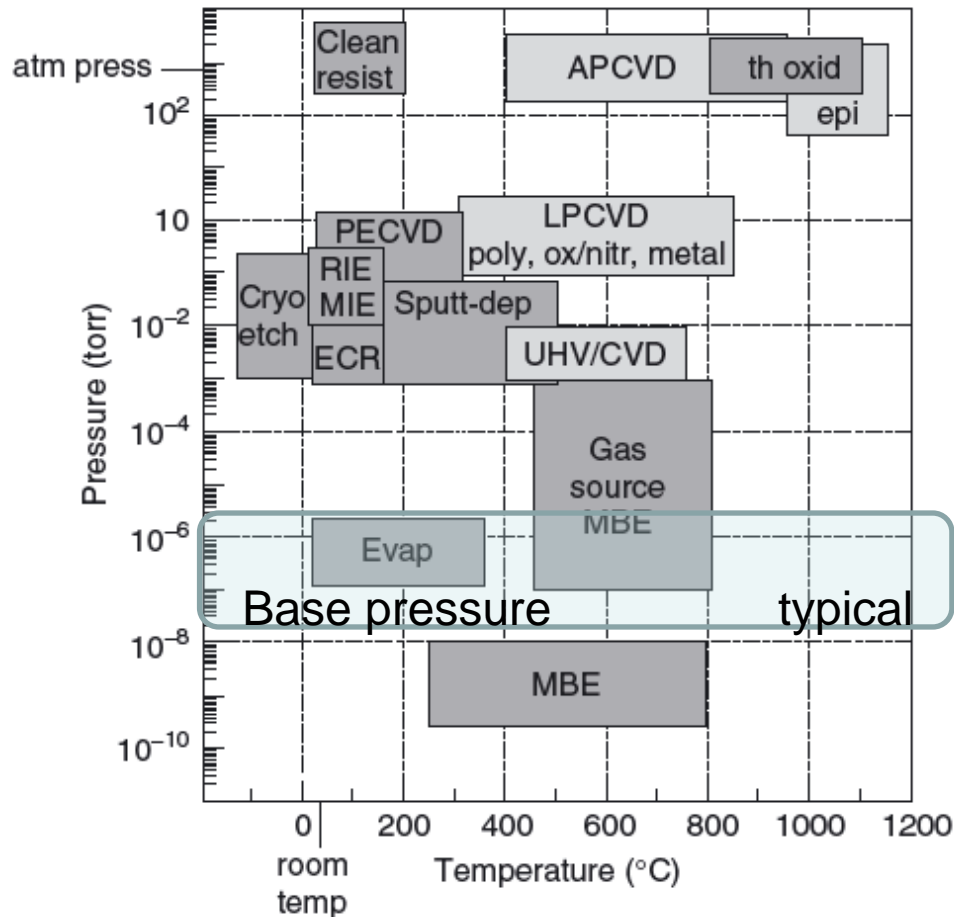
Deposition upto 10 nm Al_2O_3

ALD tool development



Break

Temperature vs. pressure



High vacuum - high temperature is a difficult combo, since stuff evaporates from chamber walls at elevated temperatures.

Base pressure is much lower than process pressure for most processes.

Heating the reactor

Method

resistance heating

induction heating

lamp heating

laser heating

conduction

convection

Equipment

tube furnace

epitaxial reactor

rapid thermal processing RTP

LACVD

hot plates; SW PECVD

resist ovens; gas flow on back

Hot wall vs. cold wall

In hot wall systems all parts are hot → reaction takes place on the walls as well.

At home: oven

Slow ramp rates because huge mass needs to be heated. $\sim 10\text{-}50^\circ\text{C}/\text{min}$

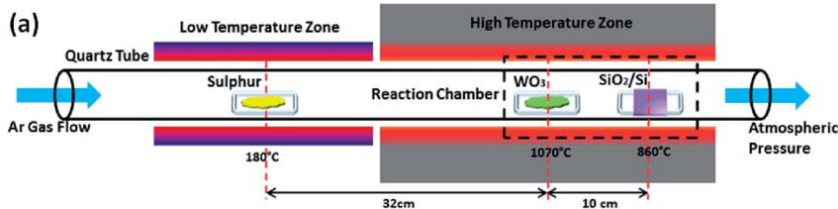
In cold wall systems only the wafer (and susceptor) is heated.

No deposition on walls.

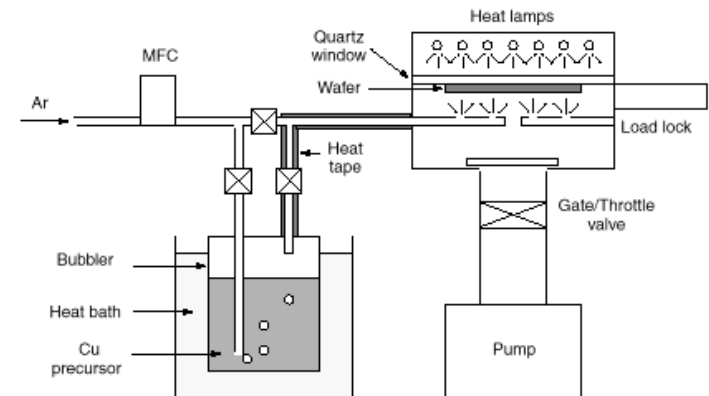
Heating options: lamp; induction, RF.

At home: microwave oven.

Fast ramp rates because small mass to be heated, $\sim 100\text{-}1000^\circ\text{C}/\text{s}$



Rong et al: Controlling sulphur precursor addition for large single crystal domains of WS₂, 2014



Changsup Ryu, PhD thesis, Stanford University, 1998

Batch vs. single wafer

Multiple wafers simultaneously

Exactly the same process for all

Uniformity worse (because of larger area)

Every wafer experiences the same conditions of flow, T, ...

Better uniformity (← smaller area, symmetry)

Need higher rate because otherwise thruput sacrificed



PVA

Samco

Batch vs. single wafer (2)

More expensive to develop a process because you need more wafers.

Nice to do experiments because you will have truly identical wafers (with some uniformity issues).

Simple process development one wafer at a time.

Drifts and variability mean that you never can be 100% sure that even consecutive wafers had exactly the same process.

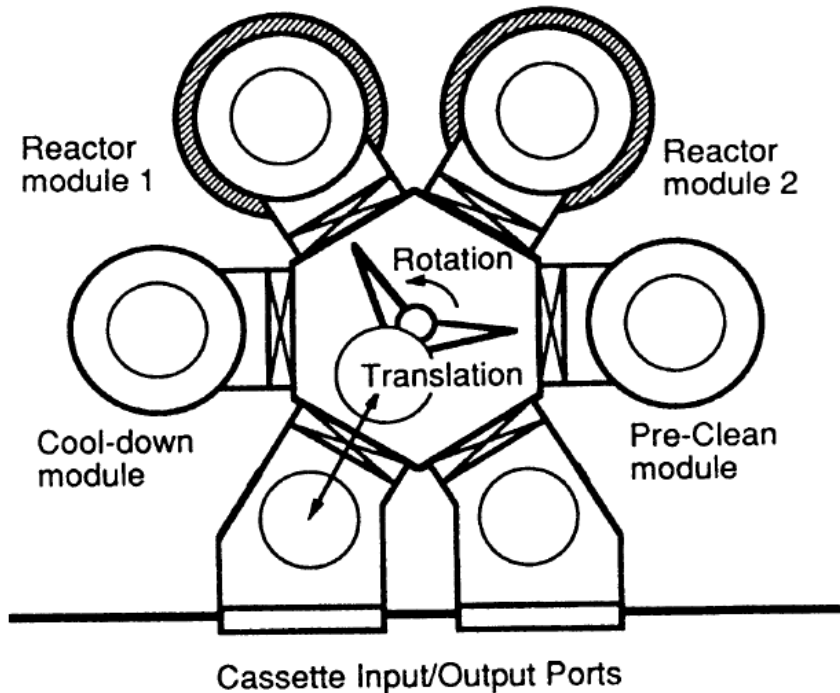


PVA



Samco

Integrated tools



Pressure regimes:

Reactor module 10^{-8} torr

Central handler 10^{-7} torr

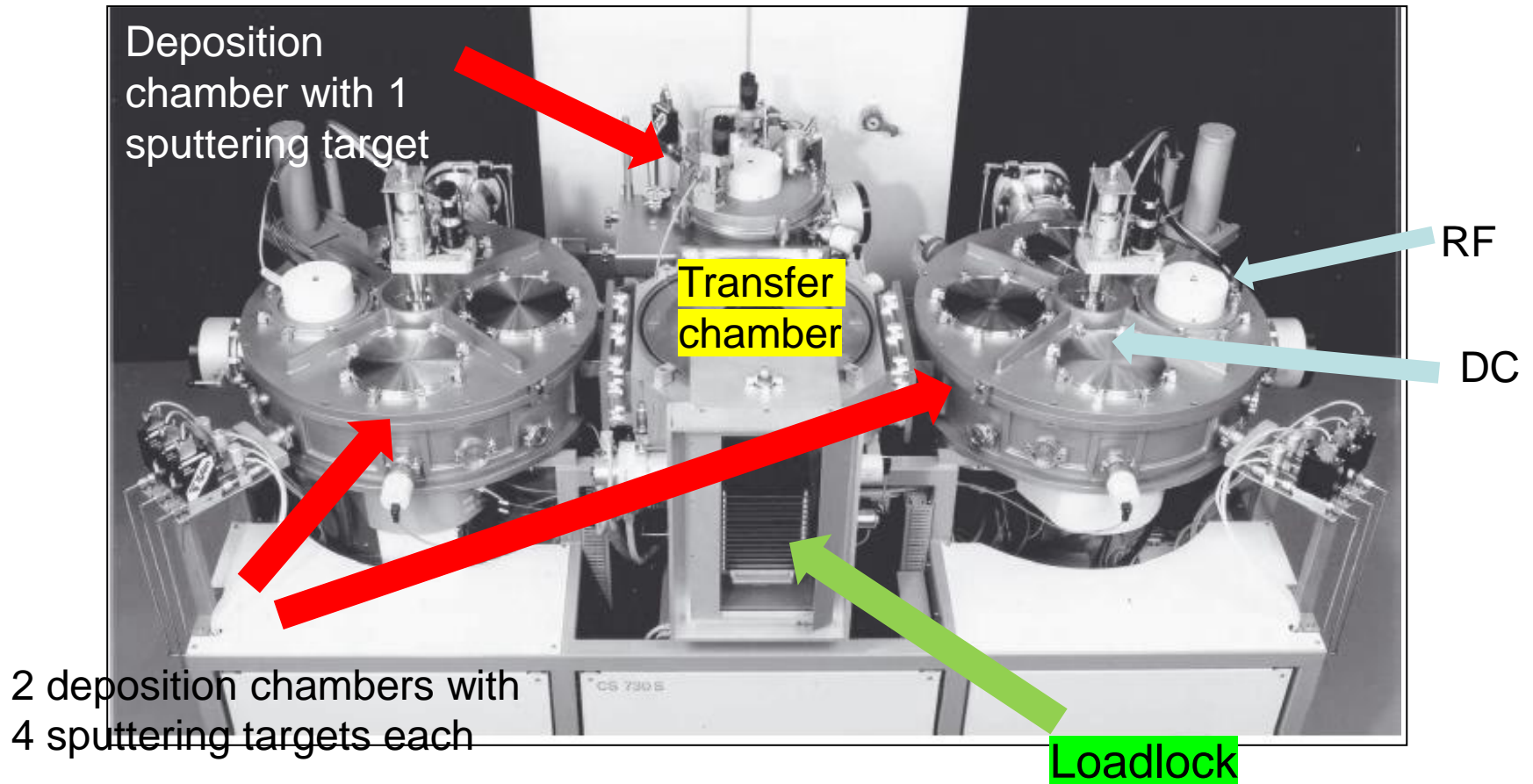
Cool-down/pre-clean 10^{-3} torr

Cassette ports 10^{-3} torr

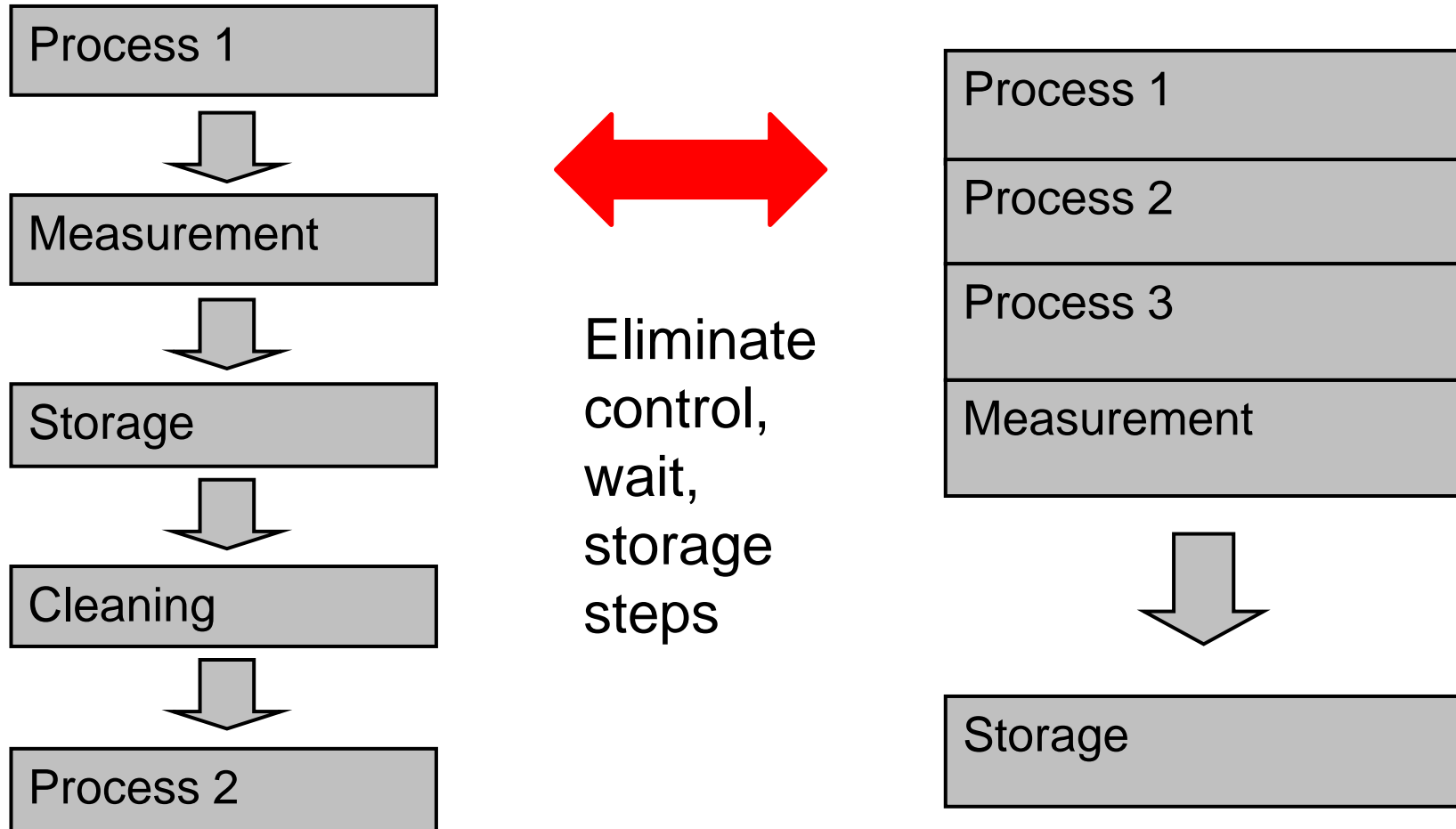
E.g. TiN by reactive sputtering poisons Ti-target; and subsequent Al deposition prone to form AlN if done in same chamber sequentially.

But if two separate chambers → no cross-over effects between steps.

Micronova von Ardenne



Integrated processes



Reactor figures of merit

Uptime/downtime:

Uptime is an overall measure of equipment availability. Uptime is reduced both by scheduled and non-scheduled maintenance. Recalibration/test wafers required to set the process running after a disruption can contribute significantly to downtime. Scheduled system cleaning is often mandatory for deposition equipment, to prevent film flaking from chamber walls.

Utilization

Utilization is a measure of equipment use: actual productive hours of all available hours. Some tools are needed many times in a process, and some may be reserved for some extra special steps only. The latter have low utilization.

Figures of merit (2)

MTTF: mean time to failure

MTBA: mean time between assists

MTBC: mean time before clean

How long will it work before failure ? Do operators need to interfere with its operation ? How often does it have to be cleaned ? These questions are operationalized by MTTF, MTBA and MTBC

MTBC is process dependent: some products tolerate more defects than others, and more relaxed cleaning intervals apply.

FOM (3)

Footprint

How big is it ? Cleanroom space is premium priced: 10000 \$/m² is the price range for a class 1 (Fed. Std.) cleanroom. In most cases, just the front panel of the system is in the cleanroom, the rest of the tool is in the service area which has more relaxed particle cleanliness requirements.

Throughput

How many wafers per hours (WPH) can the system handle ? If film thickness is doubled, deposition time is doubled. Throughput, however, might not change much if overhead (loading, pump down, temperature ramp etc.) is high relative to deposition time.



<https://www.glassdoor.sg/Photos/Lam-Research-Office-Photos-IMG111691.htm>

Cost of ownership (CoO)

	A	B
Purchase price	1 000 000 €	1 500 000 €
Operating costs	250 000 €/yr	120 000 €/yr
5 years costs	1 750 000 €	2 100 000 €

Measurement needs

- in-situ: during wafer processing in the process chamber
- in-line: after wafer processing inside the process tool (e.g. in exit load lock)
- on-line: in the wafer fab by wafer fab personnel
- ex-situ: outside analytical laboratory by expert users

Measurement needs (2)

	R&D	Pilot production	Volume manufacturing
samples	anything	full wafers (monitors)	full wafers (scribe line measurement)
analysis spot time	anything anything	not a concern minutes/hours	test site minutes/seconds

Destructive measurement discards the wafer after measurement →
measurement cost is >wafer cost.

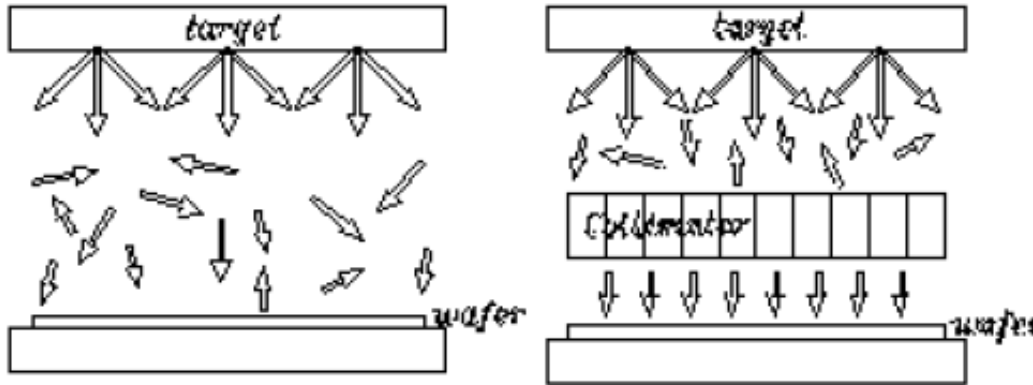
Non-destructive measurement does not destroy the wafer, but maybe it cannot be sold to the customer for cosmetic or reliability reasons.

Non-contact measurement does not physically touch the wafer, e.g. thickness by ellipsometer or resistivity by eddy currents.

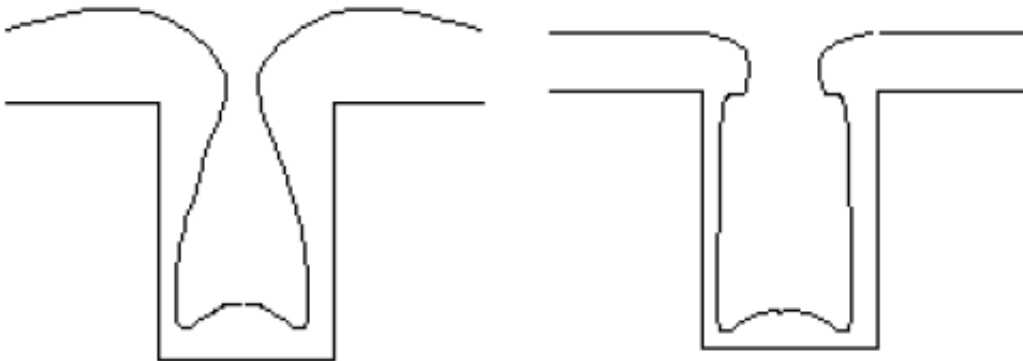
Film characterization needs

- spatial resolution (image spot size)
- depth resolution (concentration profile)
- elemental detection (constituents, impurities)
- structural information (crystal and grain structure)
- dimensional characterization (linewidth, thickness)
- mechanical properties (curvature, stress,...)
- surface properties (roughness, reflectivity,...)
- top view vs. cross sectional imaging
- defects (particles, pinholes,...)

Normal vs. collimated sputtering



Collimator prevents high-angle ions → more directional flux → more uniform step coverage.



Deposition rate reduced because collimator captures some atoms.

Sputtered TiN characterization

Film property	Analytical technique	Collimated TiN	Standard TiN
Thickness	RBS (density = 4.94 g/cm^{-3})	81 nm	161 nm
	TEM cross-section	82 nm	178 nm
Sheet resistance	Four-point probe	13.7 ohm/sq	7.4 ohm/sq
R_s uniformity	Four-point probe	3.3%	5%
Resistivity	R_s by four-point probe	112 $\mu\text{ohm-cm}$	132 $\mu\text{ohm-cm}$
	Thickness by TEM		
Density	Thickness by TEM and RBS	4.88 g/cm^{-3}	4.47 g/cm^{-3}
	Density by RBS	93% of bulk	86% of bulk
Stoichiometry (Ti/N)	RBS	1.31	1.00
Phase (JCPDS card #)	Glancing angle XRD	TiN (38-1420)	TiN (38-1420)
	Electron diffraction	TiN (38-1420)	TiN (38-1420)

Source: Wang, S.-Q. and J. Schlueter (1996).

Sputtered TiN (2)

Film property	Analytical technique	Collimated TiN	Standard TiN
Preferred orientation	$\theta-2\theta$ XRD Electron diffraction	(220)	(220)
Net stress	Wafer curvature	2.7 GPa (tensile)	3.1 GPa (tensile)
Grain structure	Cross-section TEM Plan view TEM	Columnar 2D equiaxial	Columnar 2D equiaxial
Average grain size	TEM	19.2 nm	18.3 nm
Average roughness	AFM	0.43 nm	1.23 nm
Min/max roughness		8 nm	18.7 nm
Specular reflection (% of Si reference)	Scanning UV	248 nm: 142% 365 nm: 55% 440 nm: 57%	145% 95% 123%
Impurities (at. %)	AES	O < 1% C < 0.5%	O < 1% C < 0.5%

Source: Wang, S.-Q. and J. Schlueter (1996).

Deposition rate & thruput

Deposition rate is measured in nm/min.

Thruput is measured in WPH (wafers per hour)

A batch LPCVD polysilicon reactor loads 100 wafers, depo rate is 10 nm/min, which corresponds to time 40 min for 400 nm thick film. Load, ramp etc. take 60 min. → thruput is 60 WPH.

A single wafer PECVD tool deposits silicon at 100 nm/min rate, with load, ramp etc. 1 min/wafer. 400 nm thick film is achieved in 5 min → thruput is 12 WPH.

Cost-of-ownership calculation is needed to see if one is superior to other.

First wafer effect

1st wafer sees dirtier chamber atmosphere than the following wafers, esp. if long time since last deposition.

1st wafer experiences cooler chamber than subsequent wafers (most processes use/release heat)

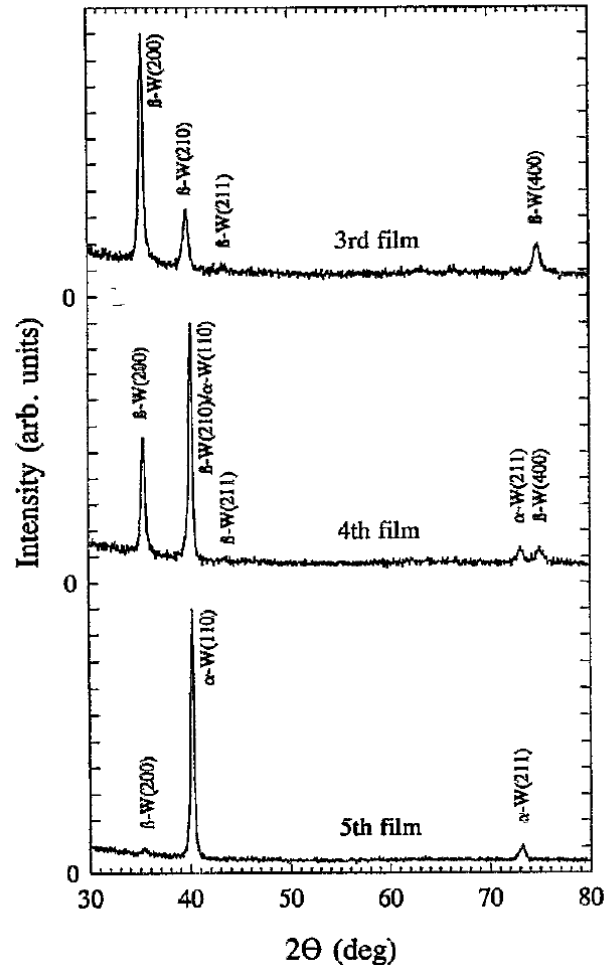
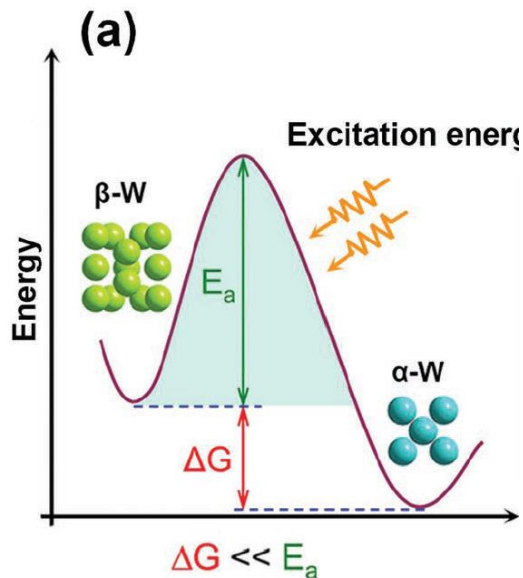
Target may be coated by impurities if it has not been used for a while (pre-sputtering, or evaporation while shutter is in place helps).

1st wafer has shortest time since previous step, later wafers may have adsorbed water and dirt during waiting.

1st wafers see clean chamber after cleaning. It may be that more deposition takes place on reactor walls, and it takes time for wall condition to stabilize.

Oxygen in tungsten film

α -phase is the stable phase, but oxygen stabilizes the metastable β -phase.



8 consecutive sputtering runs:

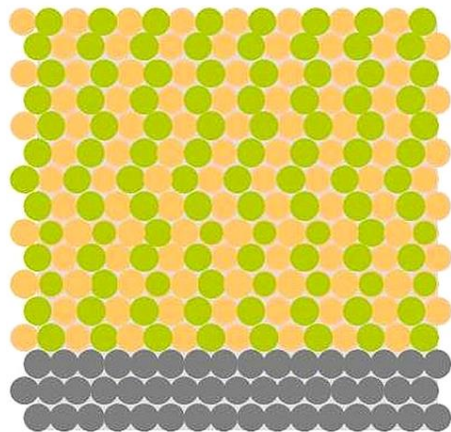
β -phase is seen in runs 1-3, but the α -phase in runs 5-8; with run #4 as a mixture.

Tungsten getters oxygen, and reactor atmosphere gets purified, and later runs contain less of it, leading to α -phase.

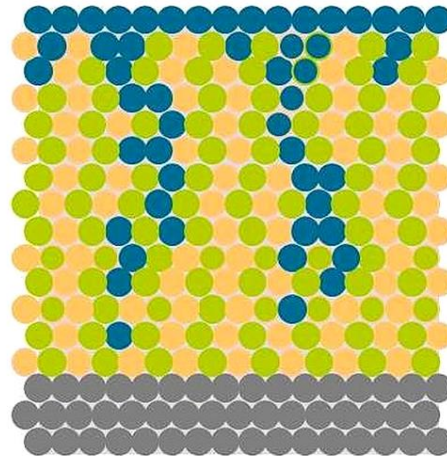
Stable or diffusing during PDA ?



TiN barrier is exposed to air/O₂
→ oxygen diffuses along grain boundaries and reacts to form TiO₂

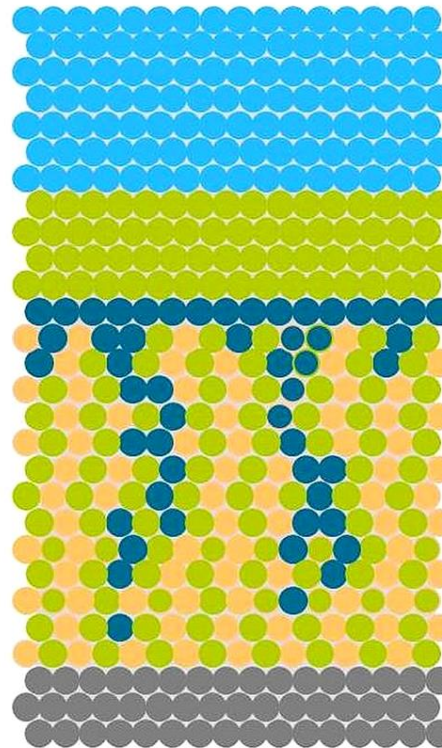


(a) Al + TiN



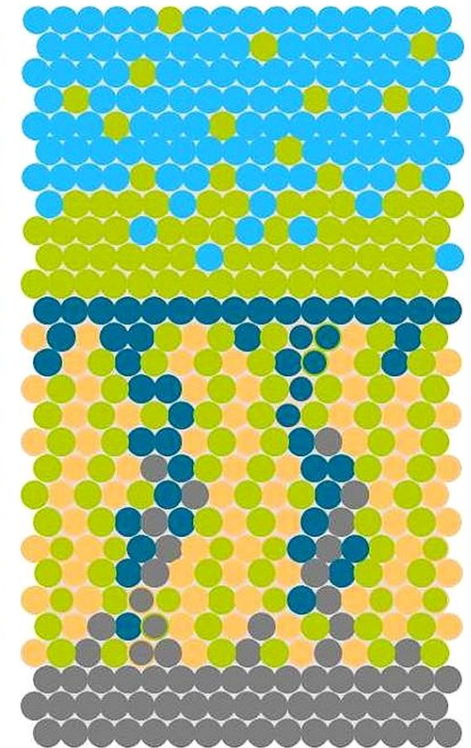
(b) stuffing of grain boundaries

Ti adhesion layer and Ag metal deposited



(c) as deposited

Reduced diffusion due to TiO₂ blocked grain boundaries

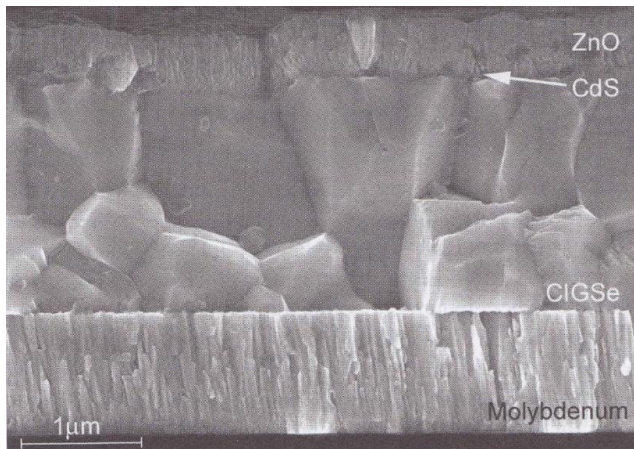


(d) after strong annealing

Vacuum break or not ?

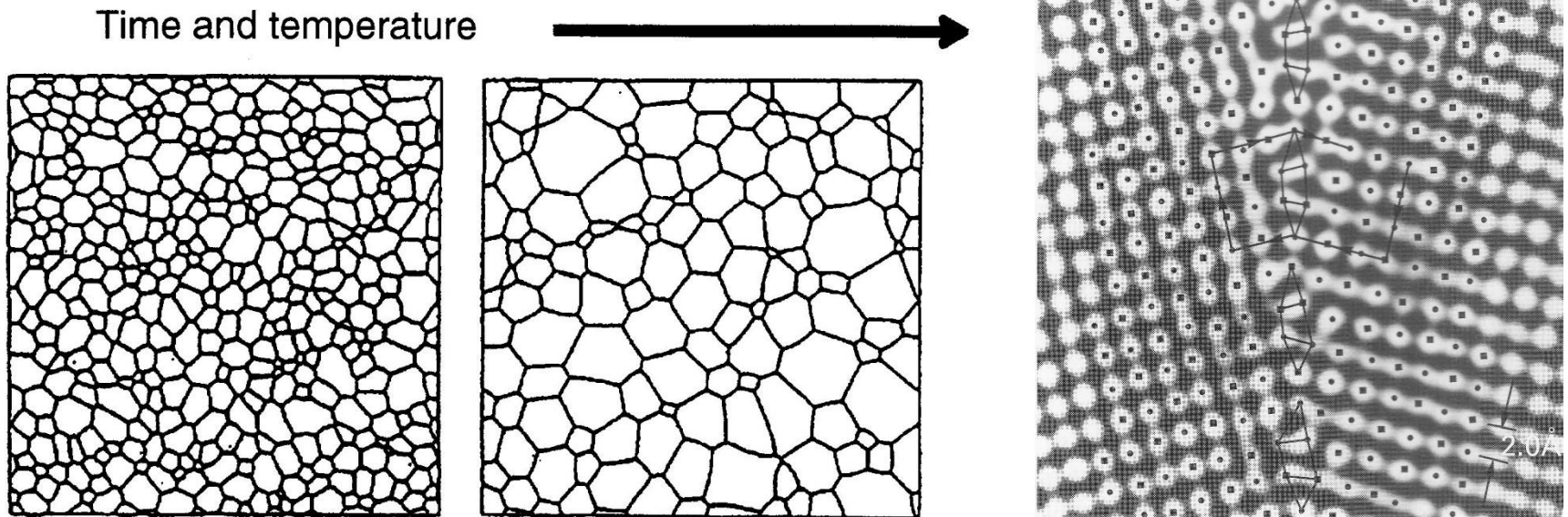
Most often we want to deposit multiple films without vacuum break,
e.g. Ti adhesion layer first and noble metal immediately afterwards → surface is as clean as possible → improved adhesion.

Or, we simply want to improve system thruput, by depositing multiple films immediately after each other:



Stuffing by vacuum break (air exposure) is an exception. Usually we want as clean surfaces as possible.

Grain growth of polycrystalline film



Grain boundaries important because they:

- act as nucleation sites for growth of new phases
- act as sites of enhanced reaction rates
- act as fast diffusion paths
- act as precipitation sites

Modification of amorphous film

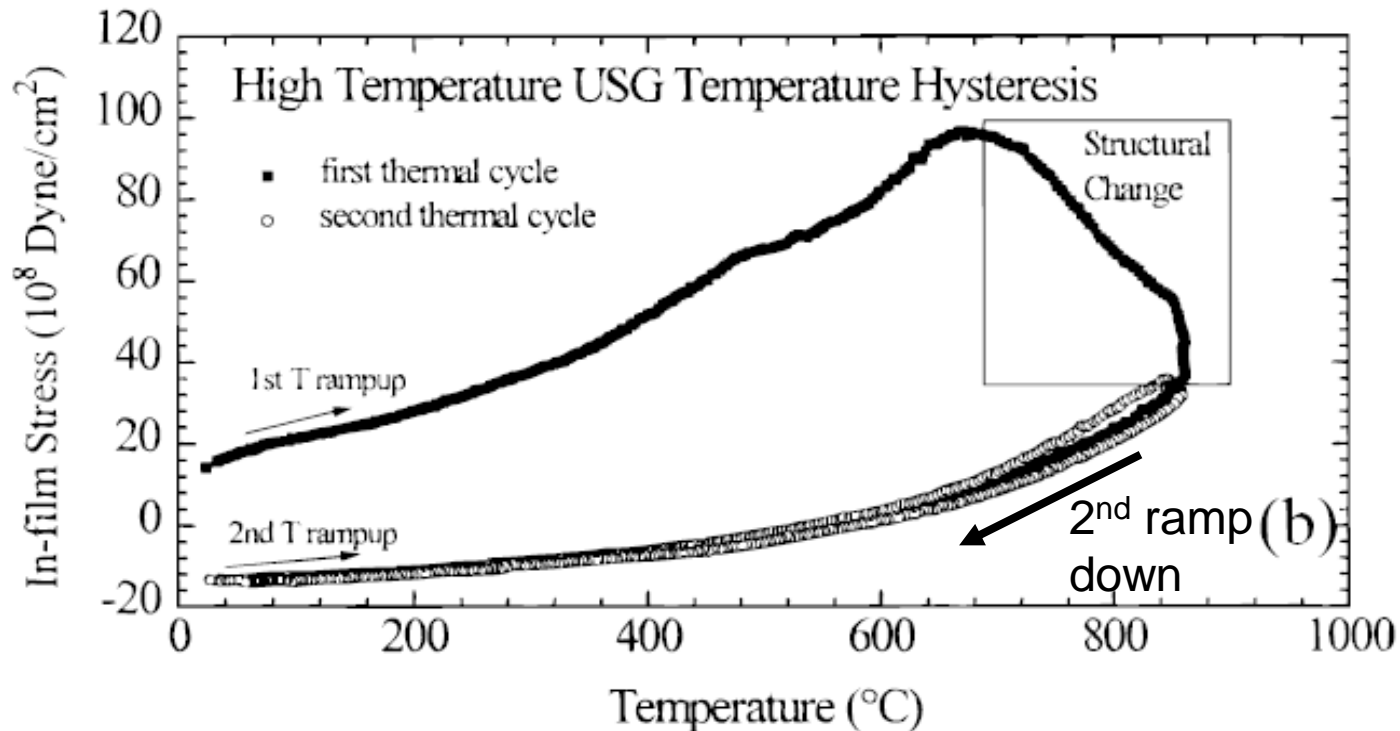
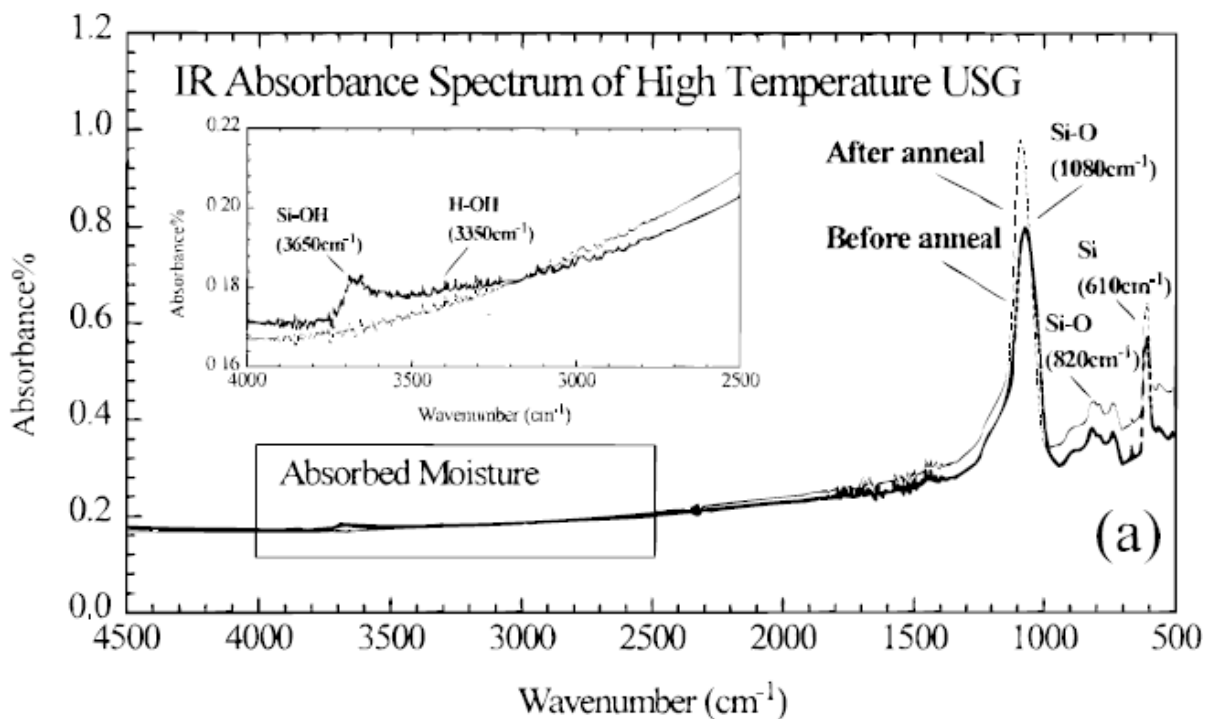


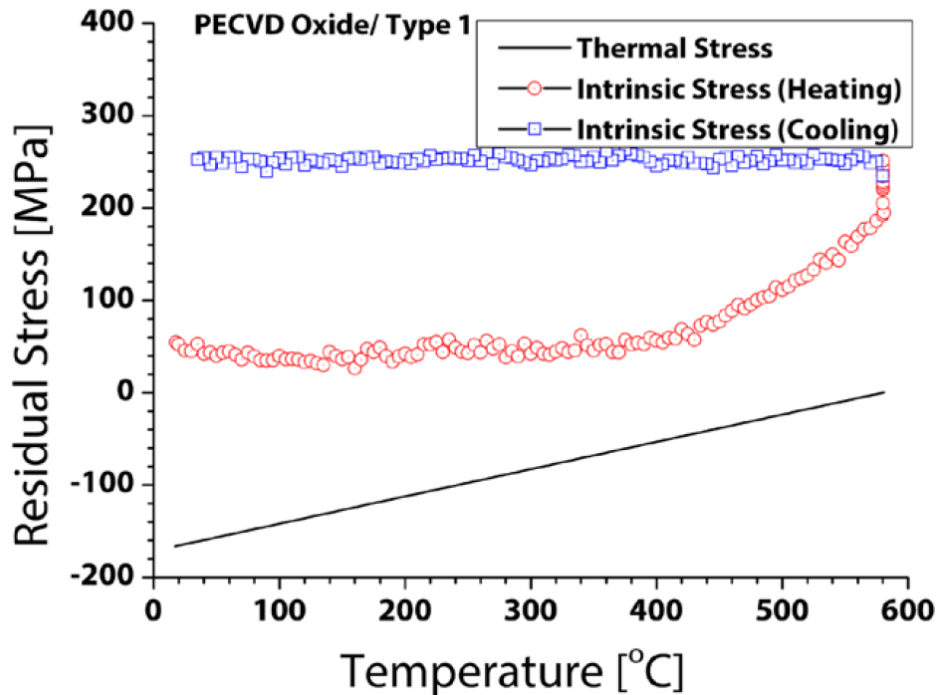
Figure 2. The effect of annealing on USG film properties: (a) film stress temperature variation at two consecutive thermal cycles, showing strong hysteresis for the first cycle only; (b) bonding structure change shown by FTIR, indicating increase in Si–O stretching mode. The deposition temperature is 550°C.

Annealing: IR analysis



- Hydrogen as Si-OH (and other)
- More Si-O bonds after annealing → denser films
- Moisture peak H-OH disappeared

Stress modification by anneal



Thermal stress = extrinsic stress:
comes from mismatch of thermal expansion coefficients.

Intrinsic stress comes from microstructure, voids, pores, dangling bonds, foreign atoms, ... Difficult to get full picture.

Intrinsic stress is modified in PDA.

Thermal stress is always there when temperature changes.

Film quality measures

- uniformity of thickness, refractive index, dielectric constant...
- low impurity levels (esp. mobile impurities)
- low defect levels
- stoichiometric composition
 - ➔ low resistivity, high density, in general properties close to bulk properties
- predictable microstructure (crystallinity/amorphousness, grain size, smoothness...)
 - ➔ known stress state,

PDA: post deposition annealing

Microstructure modified:

- grain growth (polycrystalline), crystallization of amorphous,
- defect elimination: void disappearance, desorption of loosely bound specie (esp. H)
- diffusion and reactions (e.g. Si-H broken, Si-Si formed)

Stress state modified (see CVD/ALD lecture)

Anneal atmosphere affects the result:

- inert (N_2 , Ar)
- oxidative (O_2 , H_2O)
- reductive (H_2 , but in practice N_2/H_2 96%/4%)

Productivity measures

Deposition rate

Thru-put (very different from rate !)

Wide process window (robust against parameter drift)

Yield of precursors (source gases/targets expensive)

Deposition on one side or both sides simultaneously ?

Uniformity across the substrate

Uniformity across the batch

Repeatability run-to-run

Repeatability day-to-day

Use quality

What are the stressors the film is going to experience in use?

- mechanical (bending, contact wear, particle abrasion...)
- chemical (humidity, sea water, acids, solvents...)
- electrochemical (metal-metal contacts, electrolytes, anodic oxidation...)
- electrical (high current density, high E-field...)
- thermal (high-T, low-T, thermal cycling...)
- optical (UV)
- biological (protein adsorption, biofilms,...)