MOCVD PROCESS MODEL FOR DEPOSITION OF COMPLEX OXIDE FERROELECTRIC THIN FILMS

G.S. TOMPA[1], A. COLIBABA-EVULET[1],
J.D. CUCHIARO[1], L.G. PROVOST[1], D. HADNAGY[2],
T. DAVENPORT[2], S. SUN[2], F. CHU[2], G. FOX[2],
R.J. DOPPELHAMMER[3], and G. HEUBNER[1]

[1]Structured Materials Industries, Inc., 120 Centennial Ave.,
Piscataway, NJ 08854-3908
[2]Ramtron Corp., 1850 Ramtron Drive, Colorado Springs, CO 80921
[3]Applied Epi, 4900 Constellation Drive, St. Paul, MN 55127

ABSTRACT

Thin complex oxide films, such as ferroelectrics, pyroelectrics, waveguides, superconductors, MEMS/MOEMS and piezoelectrics are experiencing rapid growth in a wide variety of commercial markets. In particular, ferroelectrics as used in IC cards, embedded memories with micro-controllers, stand-alone memories, and other ASIC applications, require deposition by a technique that is compatible with deep sub-micron advanced geometry IC integration. Metal Organic Chemical Vapor Deposition (MOCVD) is the deposition method of choice for achieving conformal uniform (composition and thickness) contiguous pin hole free thin films over the challenging sub 0.2 micron geometry topology necessary for implementing advanced devices. MOCVD is free of the physical limitations that produce film defects associated with other liquid source and physical vapor deposition techniques. Of the different MOCVD approaches to complex oxide film deposition, we have found rotating disk reactor (RDR) MOCVD to be the most versatile and capable of producing uniform films efficiently. Using computational fluid dynamic methods, a process model for optimizing the multiple process parameters of RDR-MOCVD, single and multiple wafer production tools, has been developed. The models developed are compatible with plasma-assist, single or multiple precursor and graded composition deposition of films. The developed hardware and model have been combined to maximize thin oxide film properties. SMI modeled RDR MOCVD methodology and film results are presented.

Keywords: MOCVD, modeling, oxide, ferroelectric, dielectric, RDR, pyroelectric, waveguides.

INTRODUCTION

Complex oxide films are experiencing rapid growth in a wide variety of commercial markets. Such markets are listed in Table I.   Several deposition techniques have been applied for forming such films, including sputtering [1], sol-gel [2], pulsed laser deposition [3], MIST [4], Molecular Beam Epitaxy [5], CSD [6], and Metal Organic Chemical Vapor Deposition (MOCVD) [7].   Of these deposition techniques, only MOCVD offers the potential for depositing uniform (composition and thickness) contiguous films over planar and complex topologies within advanced manufacturing environments.   However, there are many difficulties in reducing complex oxide MOCVD to practice and to packaging the process in a commercial production tool.

        There are two basic regimes in which to deposit films. The deposition regimes are either kinetically limited or mass transport limited.   Unfortunately, many of the precursors and depositing elements and compounds do not easily allow a single mode of deposition. Typically, for a complex oxide, the various precursors' decomposition temperature offset varies by tens to even hundreds of degrees, while the depositing chemicals may have vapor pressures varying by several orders of magnitude at the deposition temperature. Further, for the films we are depositing and within the constraints applied (such as allowable substrate temperature, deposition rate, reactant residence time, reactant decomposition rate, chemical compatibilities, process gases, and resultant deposited film quality, among others), we generally have chosen not to operate in a very low pressure diffusion furnace arrangement.   Instead, we rely on controlling the thermal and flow dynamics of the process, the partial pressures and distribution of reactants in the vertical reactor in order to achieve uniform high quality film deposition.

        Having defined a general processing phase space, we then evaluated basic reactor geometries, which are shown in Figure 1. The two basic geometries are horizontal and vertical; each geometry must account for thermal buoyancy effects. As indicated in Figure 1, the CVD process can suffer from particle generation or other defects as a result of gas phase pre-reactions, wall coatings, depletion, back diffusion, re-evaporization, recirculation, thermal buoyancy, and so on—all of which must be addressed. In the horizontal reactors, gases

pass over the reactor from one side of the reactor to the other. Unfortunately, these geometries require complex mechanisms to overcome depletion effects.

An alternative deposition design is the vertical reactor. In the vertical reactor gases pass down from the top and over and across the wafer. In this design there are two fortunate advantages that streamline the process optimization. If the vertical geometry is combined with rotation of the substrate, then the reactant concentration can be graded radially to counter reactant depletion and hence maintain compositional and thickness uniformity. There is no provision to accomplish this in a horizontal reactor. A further enabling innovation in a vertical reactor is that the substrate assembly can be rotated at speeds sufficient that the viscous drag generates a pumping action. The pumping action counters thermal buoyancy and hence "pulls" the streamlines down to the surface. The well-defined flow streamlines, combined with the ability to control the reactant distribution, have enabled SMI spin CVD (rotating disk reactor) systems to achieve extremely high quality films in a wide variety of material systems (including oxides, semiconductors, compound semiconductors, nitrides, carbides, diamond, metals, oxides, and so on). Further, thermophoresis at the well-defined thermally graded layer above the wafer advantageously helps to drive small particles from the surface in the RDR.
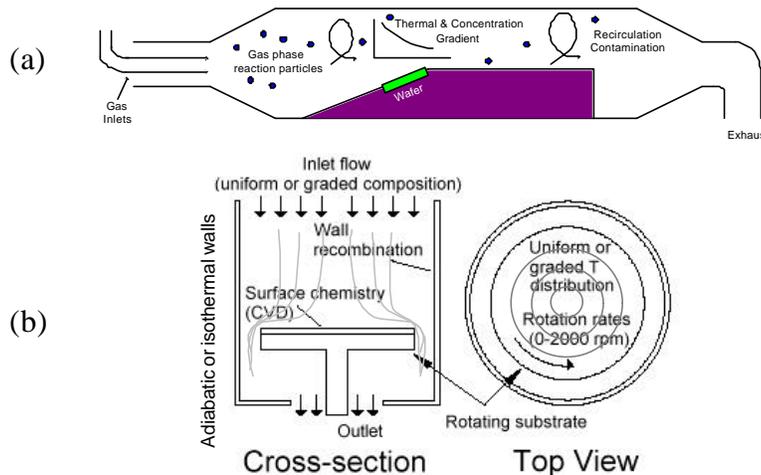
Figure 1. Comparison of fundamental reactor types (a) horizontal tube, and (b) vertical (with or without rotation).

We herein review modeling aspects of RDR MOCVD systems (SMI Spin CVD tools), focusing on achieving high quality ferroelectric oxide films. The review is configured to introduce the reader to the parameters affecting performance and, in a broad fashion, the benefits and limitations in practice. In general, a given set of figures is the distillation of many modelings and not meant to be a complete discussion on any given parameter study set.

Table I. Application areas of complex oxides addressed by the deposition technology reviewed in this paper.

Material/ Application	Description	Material/ Application	Description
Ferroelectrics	Used for non-volatile memory applications	Emitters	Used in a range of light emitting, converting and amplifying devices
Pyroelectrics	Used for optical IR sensing and imaging	Chemical Sensors	Used to sense a variety of chemicals or biologicals
Waveguides	Used for optical communications and spectroscopy	Superconductors	Used in power transmission and high speed devices
Phosphors	Used in producing displays	Piezoelectrics	Used in positioning acoustic and microwave devices
MEMS/MOEMS	Used in an array of positioning, structural and optical applications	Dielectrics	Used in capacitors and microwave devices
Tribological	Used for protecting surfaces	Optical Coatings	Used in anti-reflective, filter, and scratch resistance
Heaters	Used for displays and windows	Other	Complex oxides are used for many additional applications
Transparent Conductive Oxides		Used in windows, displays, heaters, solar cells, thermophotovoltaics, detectors, contacts, and shielding	

MODELING DISCUSSION

Computational fluid dynamical (CFD) modeling codes are used to model the RDR MOCVD deposition process. A variety of meshings, as indicated in Figure 2, are used, depending on what performance aspect is being focused on. In general, the meshing falls into three

categories—rectangular (structured), unstructured (typically using triangular surface elements and tetrahedral cells) and mixing both cell types to simplify portions of the reactor. Figure 2(d) provides an example of a complex non-symmetric CVD reactor meshing.


(a)   (b)

(c)   (d)
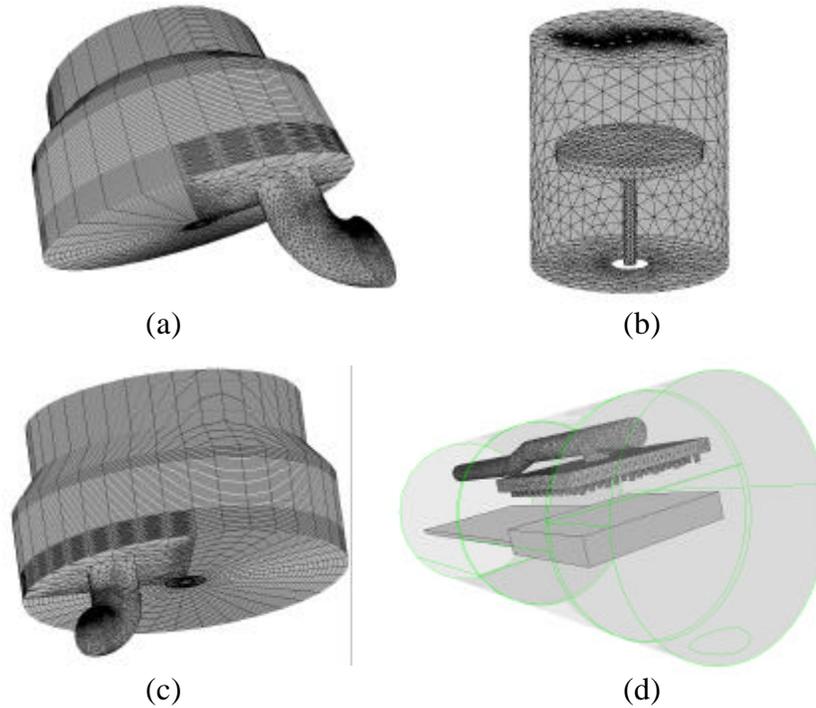
Figure 2. (a) Contoured RDR showing 20,000 rectangular (structured) facelets and 44000 cells, (b) surface grid using triangular elements are the starting point for tetrahedral volume elements, (c) mixed meshing used for optimizing output and modeling time, and (d) complex meshing for an asymmetric horizontal reactor.

2D vs. 3D

Depending on the symmetry of the geometry to be modeled, we may use 2D or 3D models. One of the first examples of a checks and balances criteria applied to accepting or selecting whether or not to use 2D or 3D modeling is shown in Figure 3. The 3D model includes many more cells and therefore requires much greater computation time; however, the added computational time may or may not be warranted to increase precision. If the 2D model can be used, several modelings can be performed in the time it takes to complete a 3D model. This, of course, depends on the symmetry of the problem. In Figure 3 we compare the modeling results of a 2D and a 3D study. In

this symmetric case, we find that there is excellent agreement between the two modelings. Such criteria are used in determining when a series of processes can effectively be modeled in 2D—greatly shortening the development time.
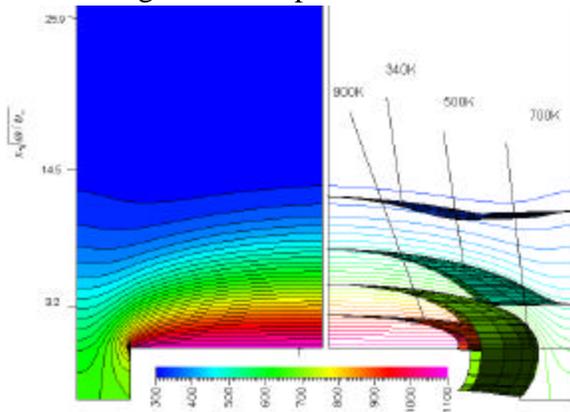
- 3D effects become noticeable when cell size increases or when asymmetries or inclusions modeled are present;
- Disk to cell diameter ratio becomes important;
- Mixed Convection Parameter (MCP) must be maintained low such that forced convection dominates flow.

Rotation Effects

The effect of rotation on gas flow dynamics and, hence, deposited film properties, can be very profound. Figure 4 compares the effects of no rotation and increasing rotation. With no or low rotation rates, the


(a)   (b)   (c)
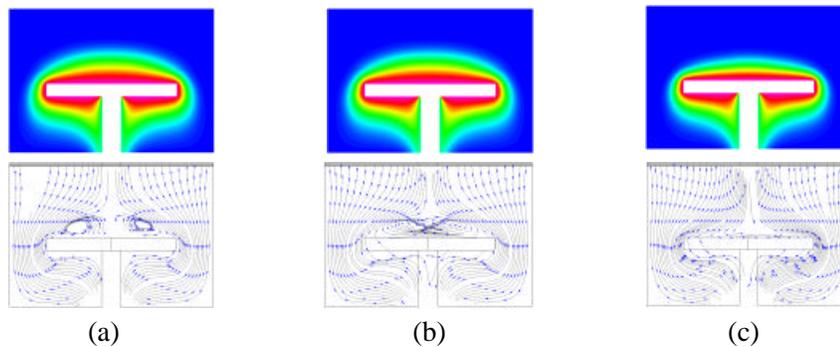
Figure 4. Simple comparison of rotation effects on a 5" platter operated at 20 Torr and (a) 0 rpm, (b) 200 rpm, and (c) 750 rpm; note the lack of recirculation at 750 rpm.

buoyancy effect dominates and an upswelling of flow at the center generates recirculation patterns, resulting in possible particles and loss of film compositional integrity. Rotating the wafer carrier (that can contain multiple substrates) at several hundred rpm is generally sufficient to counter buoyancy effects and generate non-recirculating

streamlines above the deposition plane. In fact, at a given size and ratio of wafer carrier radius to reactor radius, rotation rates, typically through a few thousand rpm, preserve desirable flow patterns. The effect of rotation on the boundary layer thickness is more apparent at low pressures due to transport effect (diffusion).

Flow Rate Effects
Optimally, the inlet flow rate should match the rotation rate of the disk for uniformity of boundary layer thickness and minimization of recirculation. For illustration, using 20 Torr, 750 rpm and 1000 K substrate temperature, one finds that an inlet flow rate of 120 slm, for a 16" platter, can be appropriate. At too low of a flow, recirculation effects are apparent. Significantly increasing the flow serves to further flatten the boundary layer. This is often the approach taken in horizontal systems, e.g. increase the flow to flatten the boundary layer and minimize depletion. Unfortunately, going to very high flow rates necessitates increasing precursor consumption to maintain relative concentrations and hence increasing consumable costs and lowering system efficiency. In general, by using flows slightly greater than those needed to prevent recirculation, consumption costs are easily minimized.


12 slm inlet flow	120 slm inlet flow	1200 slm inlet flow
Recirculation	No Recirculation	Stagnation
(a)	(b)	(c)
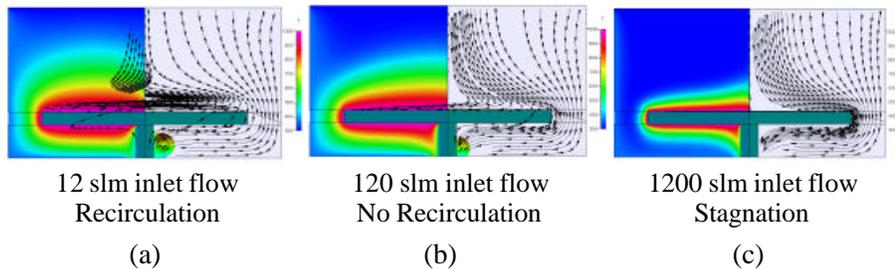
Figure 5. 2D comparison of flow rate effects in a RDR at 20 Torr for a 16" platter: (a) 12 slm, (b) 120 slm, and (c) 1200 slm.

Pressure Effects
Pressure effects tend to be less dramatic. For example, Figure 6 shows the effects of increasing the pressure at a constant inlet flow rate of 116 slm, 750 rpm rotational speed and a temperature of 1000 K. Streamlines, important to deposition, do have dependence on the total pressure in the reactor. The thermal boundary layer becomes much thinner at higher pressure, resulting in a significant temperature gradient increase. While there are advantages with increasing the abruptness of the thermal gradient, higher pressure, however, increases

collision rates and thus increases the potential for gas phase pre-reactions. Fortunately, in the rotating disk reactor with its well-defined streamlines, the more reactive species can often be spatially separated without sacrificing uniformities—even at high pressures. Of course, in any configuration, diffusion can be an issue.


2 Torr
(a)

20 Torr
(b)

200 Torr
(c)
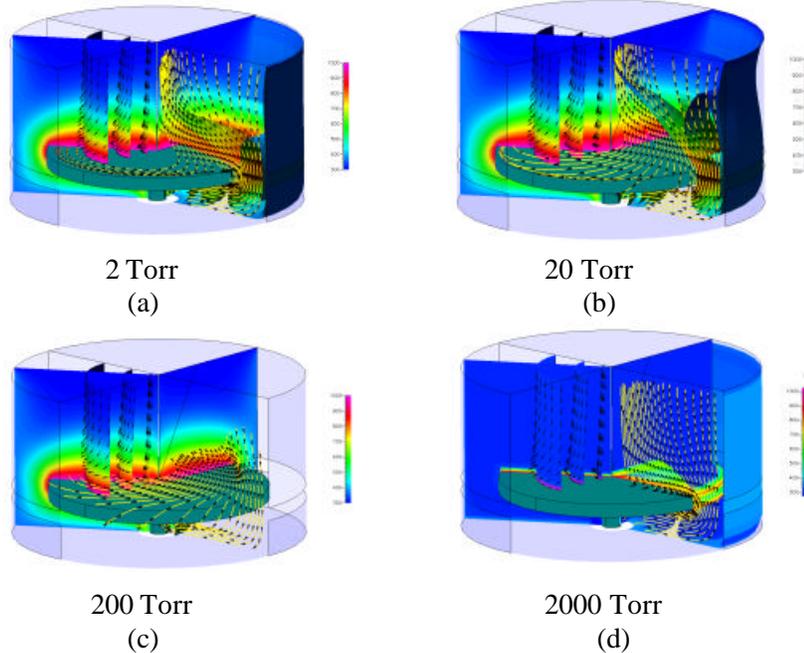
2000 Torr
(d)

Figure 6. Simple comparison of pressure effects in a RDR: (a) 2 Torr, (b) 20 Torr, (c) 200 Torr, and (d) 2000 Torr.

Exhaust Volume Effects
The volume beneath or adjacent to the wafer plane can have significant effects on flow patterns, thermal gradients, and hence, the deposited film. Figure 7 shows the effect of varying the exhaust configuration height while maintaining reactor inlet height and process parameters.   In Fig. 7(a), there is a large exhaust height and streamlines are seen to clearly recirculate below the susceptor plane. One can envision that this in turn may contribute to memory effects or that it may stir up particles. Further, the increased residence time could increase by-product buildup. In Fig. 7(b), the exhaust height has been shortened and the effect is beneficial to performance. Below wafer carrier recirculation has been eliminated. A further extreme is to eliminate below wafer carrier volume entirely and pump to the sides. As seen in Fig. 7(c), side pumping maintains non-recirculating streamlines; however, the side directed flow detrimentally affects the

thermal gradient. It should be remembered that these modelings are parameter snapshots to exemplify extremes and need to be optimized for a given specific set of process parameters. The various parameters can be adjusted to counter any identified effect.
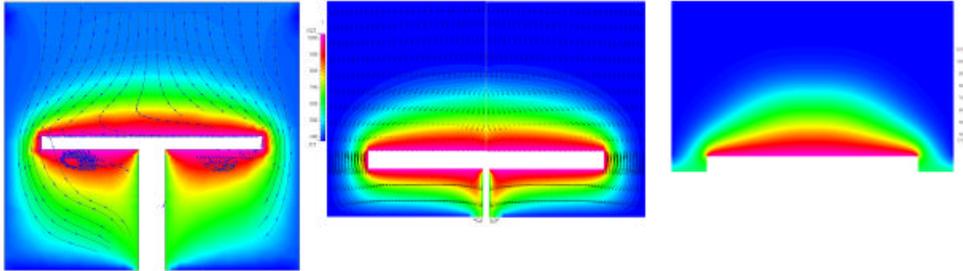
Figure 7.   Shows effects of exhaust location: (a) exhaust height proportional to inlet height, (b) exhaust height half that of (a), and (c) exhaust height zero and exhaust to the side.

Temporal and Spatial Flow Separations
With the need for uniform contiguous nanolayer films, has come great interest in deposition techniques providing films formed essentially

(a)
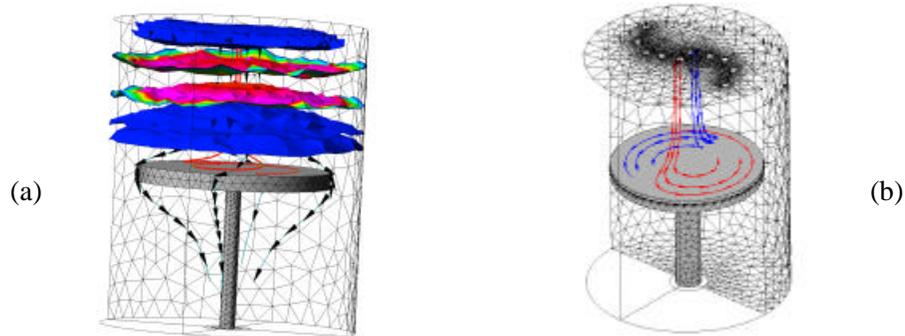
Precursor delivery optimization: ALD simulations can simplify implementation as in this case of sequential precursor "wave" fronts.

(b)

Precursor vapor injection optimization: transport of two (or more) different precursors injected within a uniform flow can be predicted.

Figure 8.   (a) Sequentially injected reactant(s) wavefronts in the uniform carrier flow of a RDR for ALD, and (b) geometrically separated reactant injection within the uniform flow (can also be sequentially pulsed) for ALD.

layer by layer. Of specific interest are Alternating Layer Deposition (ALD) approaches. RDRs are inherently attractive for implementing

ALD. Figures 8(a) and 8(b) show two approaches to achieving ALD that make RDR tools attractive for ALD.

First, because the rotation creates well-defined non-recirculating flows, precursor flow can be switched on and off or cycled through several precursors sequentially within a constantly flowing uniform carrier gas flow, with no "cross talks" between the layers. Fig. 8(a) illustrates "wavefronts" of flow through a reactor. Specifically, each "wavefront plane" is separated vertically by a time difference in that plane of gases introduced into the reactor. While diffusion will slightly spread the flow in time, it is clear that there will be little mixing between alternating layers.

Within the RDR, well-defined non-recirculating flows allow two (or more) flows, injected into a uniform flow, to be well separated within a single reactor to achieve ALD. Fig. 8(b) shows the clear separation of streamlines in a RDR which, allowing for diffusion, can be used as a methodology for ALD.
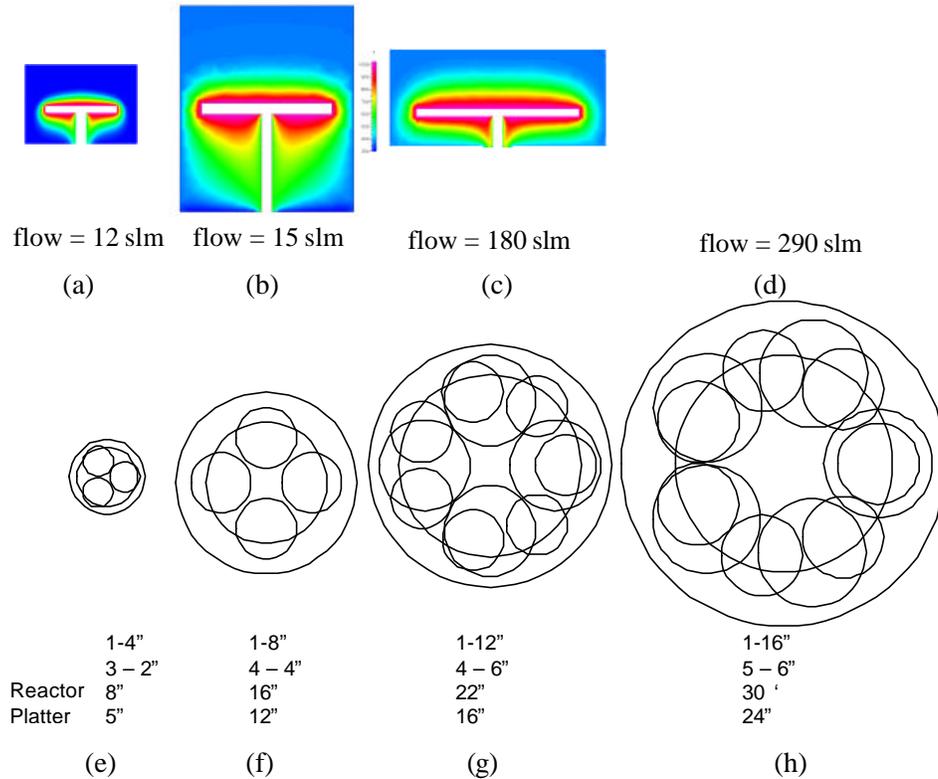
Many RDR attributes can be used to modify and enhance ALD processing. The inlet height can be shortened to minimize diffusion and pre-reaction times. The flow rate and purge times along with other process parameters can be varied to optimize an ALD process. Hence, RDRs are well designed for performing ALD in a multiple of modes.

Scaling and Throughput

Important in selecting a reactor technology is its ability to scale with increasing substrate sizes and throughput demands. In some cases, where deposition times are in minutes, a reactor is generally selected to process a single wafer. However, when film parameters require deposition times with tens to hundreds of minutes, then multiple wafer (batch) depositions become much more advantageous. In Fig. 9 (a-d), we show that RDRs can be simply scaled, while maintaining their performance capabilities. Figure 9 (e-h) depicts correspondingly scaled potential wafer placements.

While we have selected 5", 12", 16", and 24" diameter platters, we could equally have chosen 3", 8", 18", 26", or 36". How large such reactors can be practically scaled and implemented is not precisely known; however, through 26" is certainly practical. Variations in flow, pressure, rpm, or ratios of reactor area to platter area [8] each can be used to further tune the reactor geometry.

5"   11"(or 12")   16"   24"

flow = 12 slm   flow = 15 slm   flow = 180 slm   flow = 290 slm

(a)   (b)   (c)   (d)

	1-4"	1-8"	1-12"	1-16"
	3 – 2"	4 – 4"	4 – 6"	5 – 6"
Reactor	8"	16"	22"	30 '
Platter	5"	12"	16"	24"

(e)   (f)   (g)   (h)
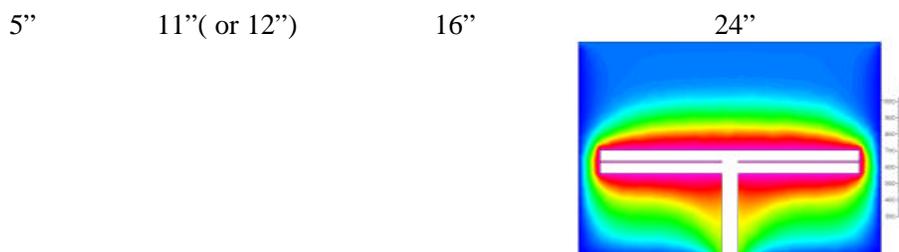
Figure 9. (a-d) shows that scaling through 24" platters can be easily effected. For equivalency, the temperature (1023K) and rotation rate (750 rpm) are kept constant; however, reactor height has been varied to show its effect, and flow is scaled with reactor dimensions. Fig. 9 (e-h) show potential wafer placement configurations for such reactors.

Inclusions

In practice, a reactor may have many additional geometric structures affecting process parameters. Examples are viewports [9], which often include purging flows, filament legs, structural support elements, heat shields, and other probes or injectors. In general, components on the exhaust side can be ignored. Examples of excludable items are filament leads as well as other probes such as a thermocouple or optical fiber. However, if the geometry is large, then great care should be given to modeling and design of such items. Inclusions upstream of the deposition surface have a much greater effect on the depositing film. Extreme examples are shown in Figure 10, where a port into the
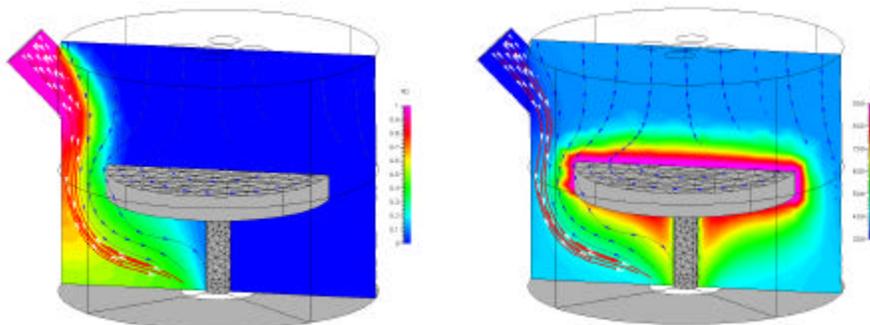
(a) Purge gas molar concentration distribution, (b) Temperature distribution
Port flow density is 1-% of the uniform flow.
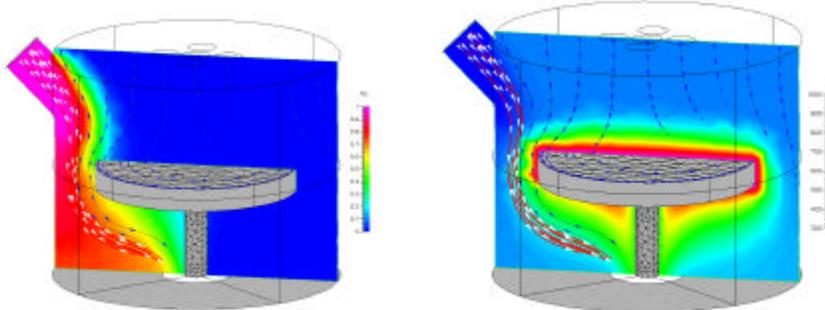
(c) Purge gas molar concentration distribution, (d) Temperature distribution
Port flow density is equal to the uniform flow.
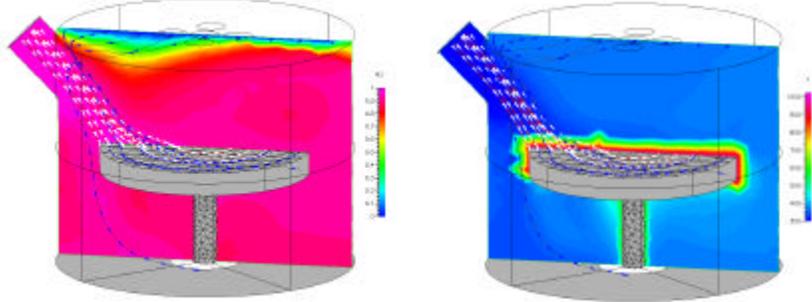
(e) Purge gas molar concentration distribution, (f) Temperature distribution
Port flow density 10x uniform flow.

Figure 10. Example of flow and diffusion effects from a purged port
having 0.1, 1.0 and 10 times the uniform top flow.

reactor is modeled with 0.1, 1.0 and 10 times the uniform top flux
density. Further, the model includes a dissimilar vapor that shows the
effects of diffusion with velocity. At low flow, the effect is negligible;
at matched flow it is still not significant. However, at high flows, the
effect can be quite dramatic. In fact, the use of jetting flows of this

general configuration is the subject of a pending Jet-Spin CVD patent. We believe this effect may have great value in certain future efforts.

Thermal Uniformity

Conjugate thermal modeling – thermal distribution in both gas and Rotating Disk assembly, reactor walls, and so on are extremely important in achieving ultimate reactor performance. Figure 11(a) exemplifies such a modeling. In practice, achieving thermal uniformity is critical to producing films uniform in thickness, composition, and quality. Figure 11(b) shows an example of achievable temperature uniformity. In order to achieve a uniformity across multiple temperatures and under different processing conditions, we use a multi-zone filament with multiple sensing points—the number and their configuration depending on the reactor size.
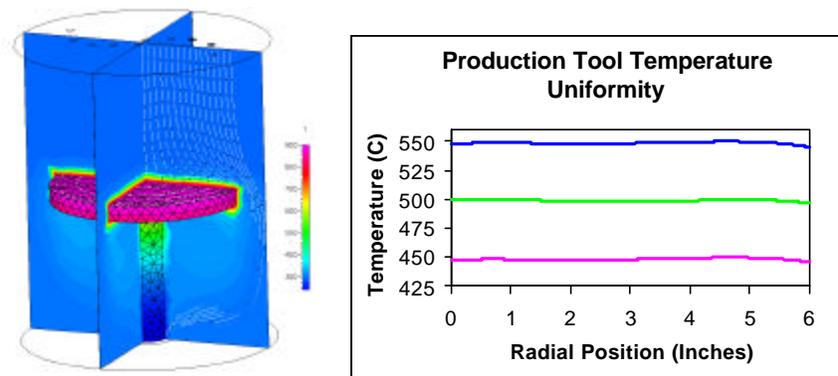
Figure 11. (a) and (b) compare modeled and achieved temperature uniformities, +/-2°C across a 12" platter.

DEPOSITIONS

Apparatus

PZT films were fabricated using a SMI Spin-CVD Metal-Organic Chemical Vapor Deposition (MOCVD) system. Depositions occurred at 525°C temperature, which was sufficient to produce in-situ crystalline thin films on the <111> platinum electrode of the substrates. Our reactor system used is shown schematically in Figure 12 and photographically in Figure 13.
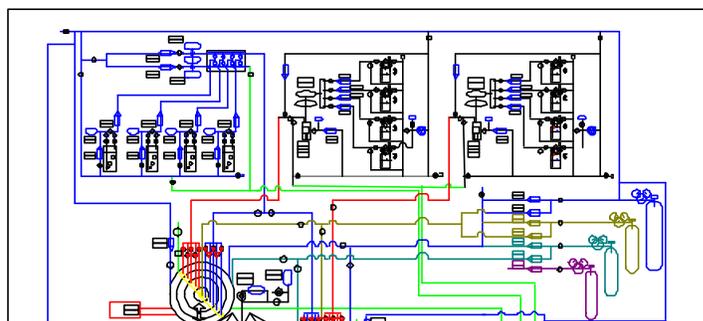
Fig. 12. Example RDR system schematic with conventional bubblers and flash evaporators, representing system used in this effort.


Fig. 13. Photograph of SMI cluster tool with one CVD tool and one etch tool.

The reactor is mounted on a cluster tool wafer exchange platform. Wafers are transferred under low pressure—inserted cold and removed warm. Pressures are controlled from 0.01 to 0.05 atmospheres, flows are a mixture of inert and oxidizing gases, and the rotation rate is typically 830 rpm. A combination of conventional and flash evaporation sources is used to deposit films. In general, deposition rates ranged from 50 nm/min to 200 nm/min.

Results

To characterize the crystalline structure a SCINTAG X-Ray diffractometer (XRD) was used. Figure 14 shows the corresponding XRD pattern of the as deposited PZT thin film on the platinum. The XRD patterns reveal <111> and <200> PZT peaks. The <111> peak splits from the platinum <111> at approximately 38 degrees producing a doublet with platinum. The second observable peak is the <200> and is solitarily at approximately 82.5 degrees. There were no further

observable PZT peaks from the MOCVD deposited films. This corresponds with the PZT being deposited and arranged in an orderly manner because of the high <111> orientation of the underlying platinum. However, XRD shows that there is a distinguishable percentage of the film showing the <200> orientation.
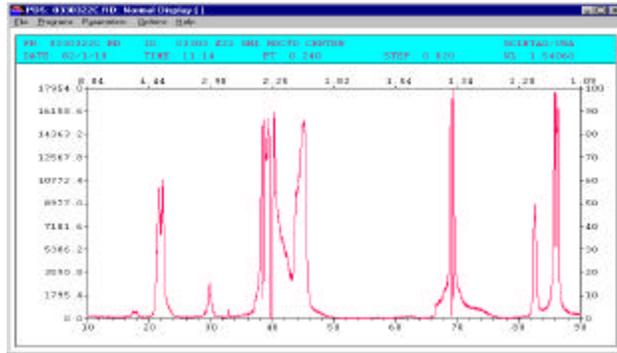
**Figure 14**. XRD pattern for MOCVD deposited PZT.

Since the XRD pattern showed crystalline structure for the MOCVD deposited PZT, the 300nm films were subsequently electroded and processed so that electrical characterization could be performed. A platinum film was sputter deposited and common integrated circuit photo lithography and etch techniques were used to delineate 50X50um capacitor structures. Figure 15(a) shows the hysteresis of the PZT films obtained from incremental voltages applied from 2-6V. The figure shows that a distinguishable hysteresis loop was obtained with the application of 3V, or with an applied field of approximately 100kV/cm^2. The polarization was further characterized by a pulsed measurement. Figure 15(b) shows the plot of polarization vs. applied voltage. The plot shows that the PZT film had little switching below 2V, however saturated at approximately 3.5V. The corresponding field required to saturate the 300nm film was approximately 116 kV/cm^2, or about 2Ec. These properties show reasonable ferroelectric characteristics for an unoptimized in-situ deposited MOCVD PZT film.
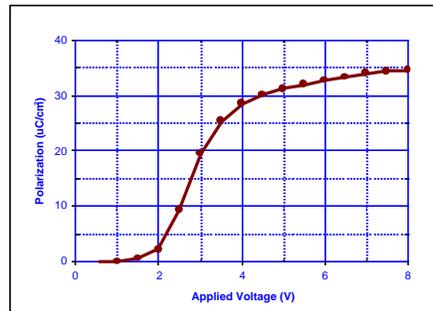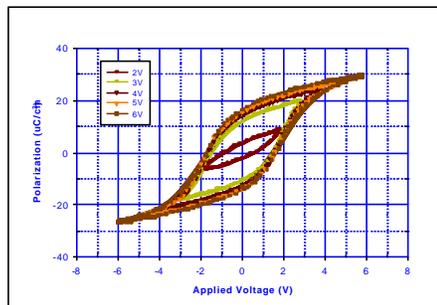
**Figure 15**(a): Measured hysteresis of in-situ crystallized PZT by MOCVD.

Figure 15(b): Measured polarization vs. applied voltage on in-situ crystallized PZT by MOCVD.

Lastly, while modeling has shown how to generate flows that should yield uniform films, it is important to confirm that uniform films can be grown. Figure 16 shows graphically the thickness of films achieved in an SMI Spin CVD reactor. The thickness uniformity is $\pm 1\%$.

PZT Film Thickness +/- 1%
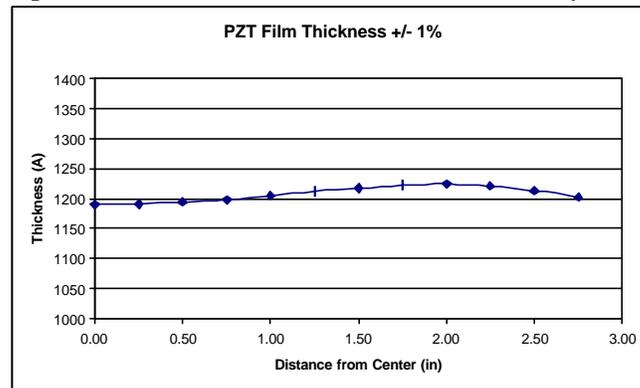
Figure 16. Example 6" PZT film uniformity produced in a rotating disk reactor.

Figure 17(a) shows a photograph of the 6" wafer whose thickness is plotted in Figure 16. The photograph shows a uniform color specular wafer. Figure 17(b) shows an SEM microphotograph of an equivalent film and highlights the grain structure achieved. This film was grown with a $\pm 2°C$ thermal uniformity and using a radial precursor injection manifold.
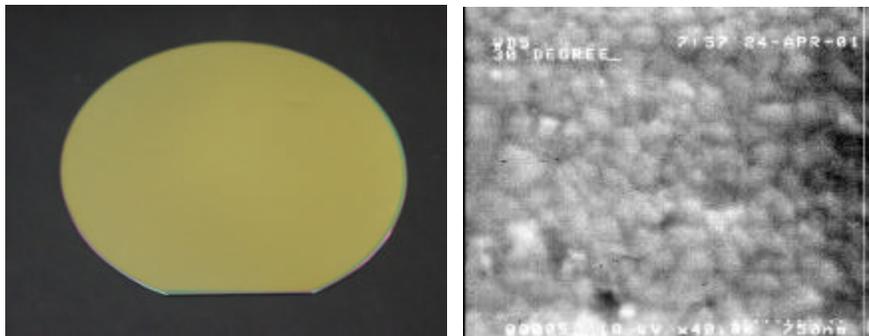
Figure 17. (a) Photograph of a *Spin CVD*^{TM} PZT film deposited on a 6" wafer, showing a uniform color, and (b) SEM microphotograph of an equivalent film, highlighting the grain structure achieved.

CONCLUSIONS

Modeling of MOCVD reactors has proven to be a time and cost saving tool, enabling us to make rapid advances in oxide CVD, especially with relation to optimizing geometric designs and tuning process parameters to efficiently achieve highly uniform high quality films. Modeling has allowed us to scale our 5" test stand to a 12" diameter platter yielding highly uniform 6" films. Modeling has revealed no impediment to scaling to significantly larger deposition areas.

ACKNOWLEDGMENTS

REFERENCES
1.   T.D. Hadnagy, J. of Integrated Ferroelectrics, **18**, 1-17, (1997).
2.   B.M. Melnick, J.D. Cuchiaro, J. Ferroelectrics, **112**, 329-351, (1990).
3.   S. Bhattacharyya and S.B. Krupanidhi, ISIF, **31**, pp. 373-383, (2000).
4.   C.A. Paz deAraujo, J.F. Scott and G. Taylor (Eds.), Ferroelectric Thin Films: Synthesis and Basic Properties, Gordon and Breach Publishers, 1996.
5.   C.A. Paz deAraujo, J.F. Scott and G. Taylor (Eds.), Ferroelectric Thin Films: Synthesis and Basic Properties, Gordon and Breach Publishers, 1996.

6.   N. Pellegri, A. Frattini, C.A. Steren, M.E. Rapp, R. Gil, R. Trbojevich, C.J.R. Gonzalez Oliver, and O. De Sanctis, <u>ISIF</u>, **30**, pp. 111-119, (2000).
7.   P.C. Van Buskirk et al., <u>Integrated Ferroelectrics</u>, **21**, p. 273, (1998).
8.   A.G. Thompson, R.A. Stall, P. Zawadzki and G.H. Evans, <u>J. Elect. Matls</u>., **25**(9), p. 1487 (1996).
9.   A. Prokopenko et al., <u>Mat. Res. Soc. Symp. Proc</u>., **588**, p. 315 (2000).