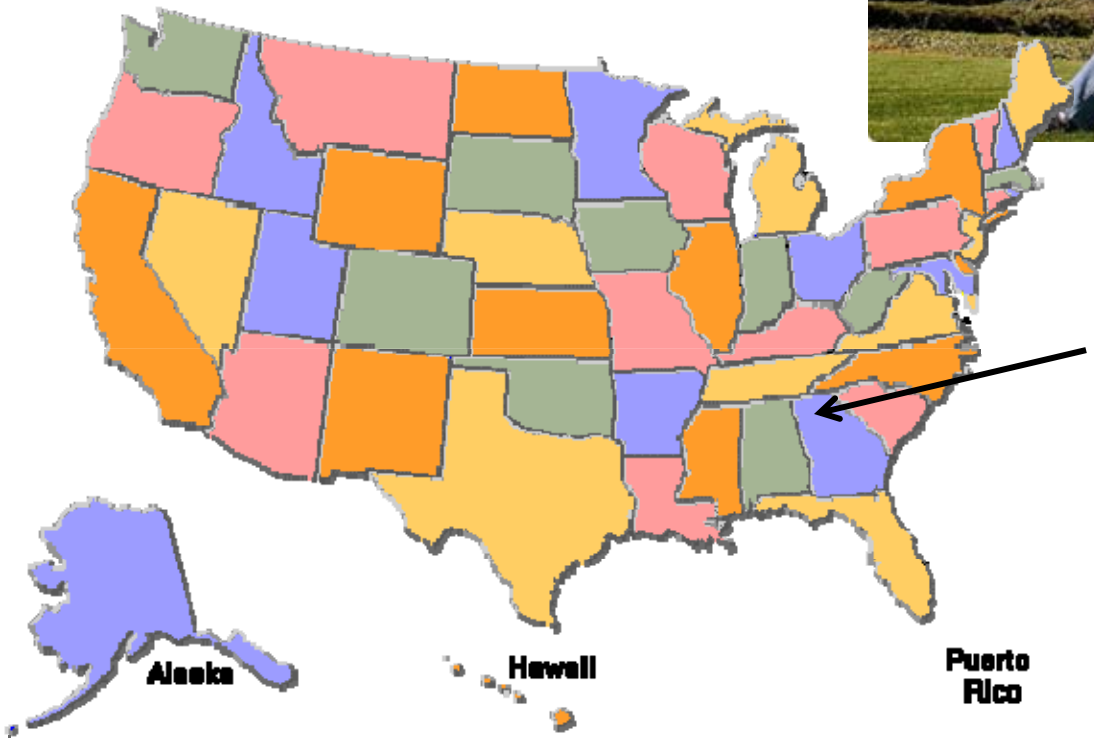
Molecular and Multiscale Modeling in Materials Design

Martha Grover Gallivan
Chemical & Biomolecular Engineering
Georgia Institute of Technology
August 15, 2008

Georgia Institute of Technology

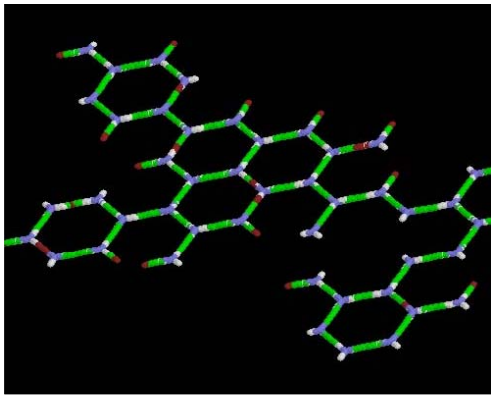
- "Georgia Tech"
- 18,000 students
- 60% engineering
- Founded in 1885

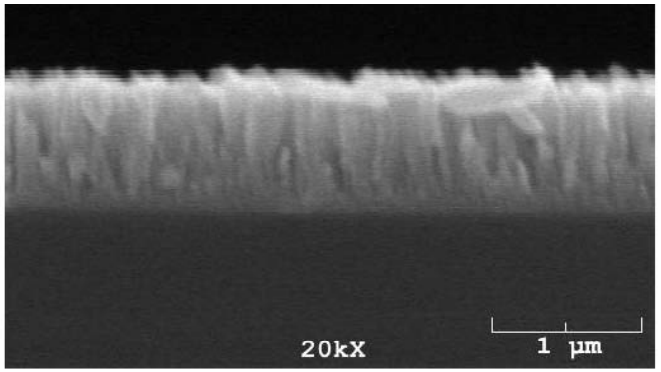


- Atlanta, Georgia
- Home of Coca-Cola and CNN

Outline

- Introduction
- Polymer networks
  - Process modeling for recipe design
- Chemical vapor deposition
  - Recipe design
  - In-situ sensing

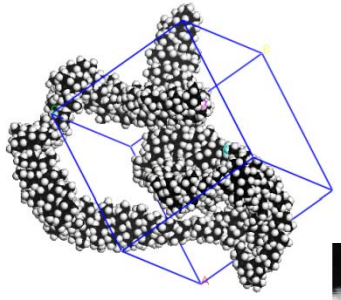




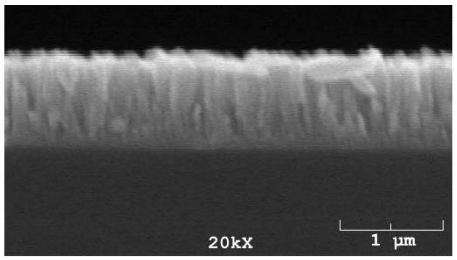
Factors influencing material properties

When chemical composition alone does not determine the material properties, we need to simultaneously consider the chemistry, the process, the material structure, and the resulting properties, ultimately the product.

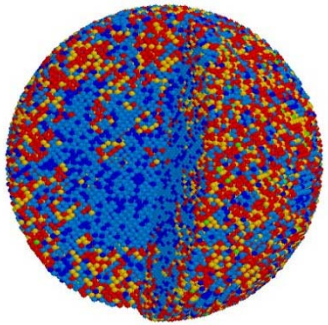
Perfect Crystal



Non-Equilibrium Structures:
- Glass
- Polycrystalline



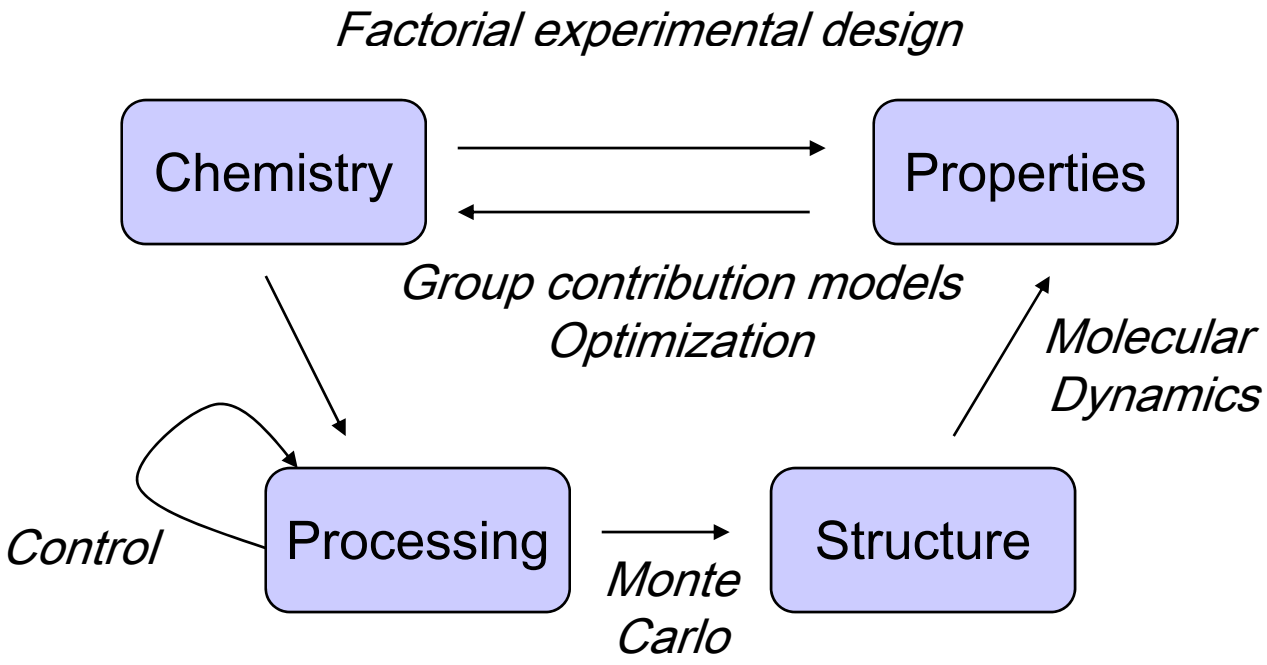
Disordered Liquid



kinetics = dynamics

Process Systems Engineering

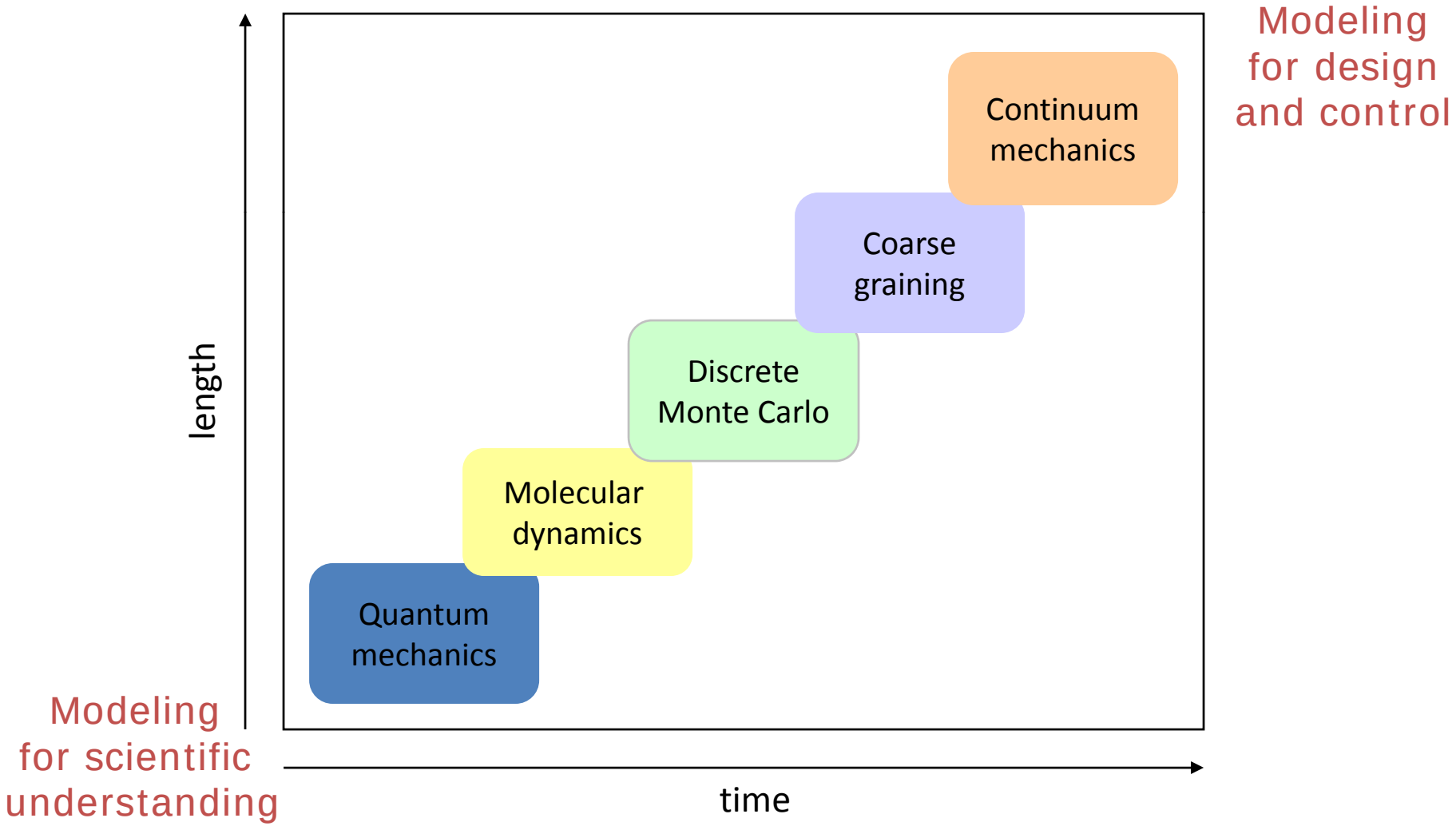
Design, simulation, optimization, and control



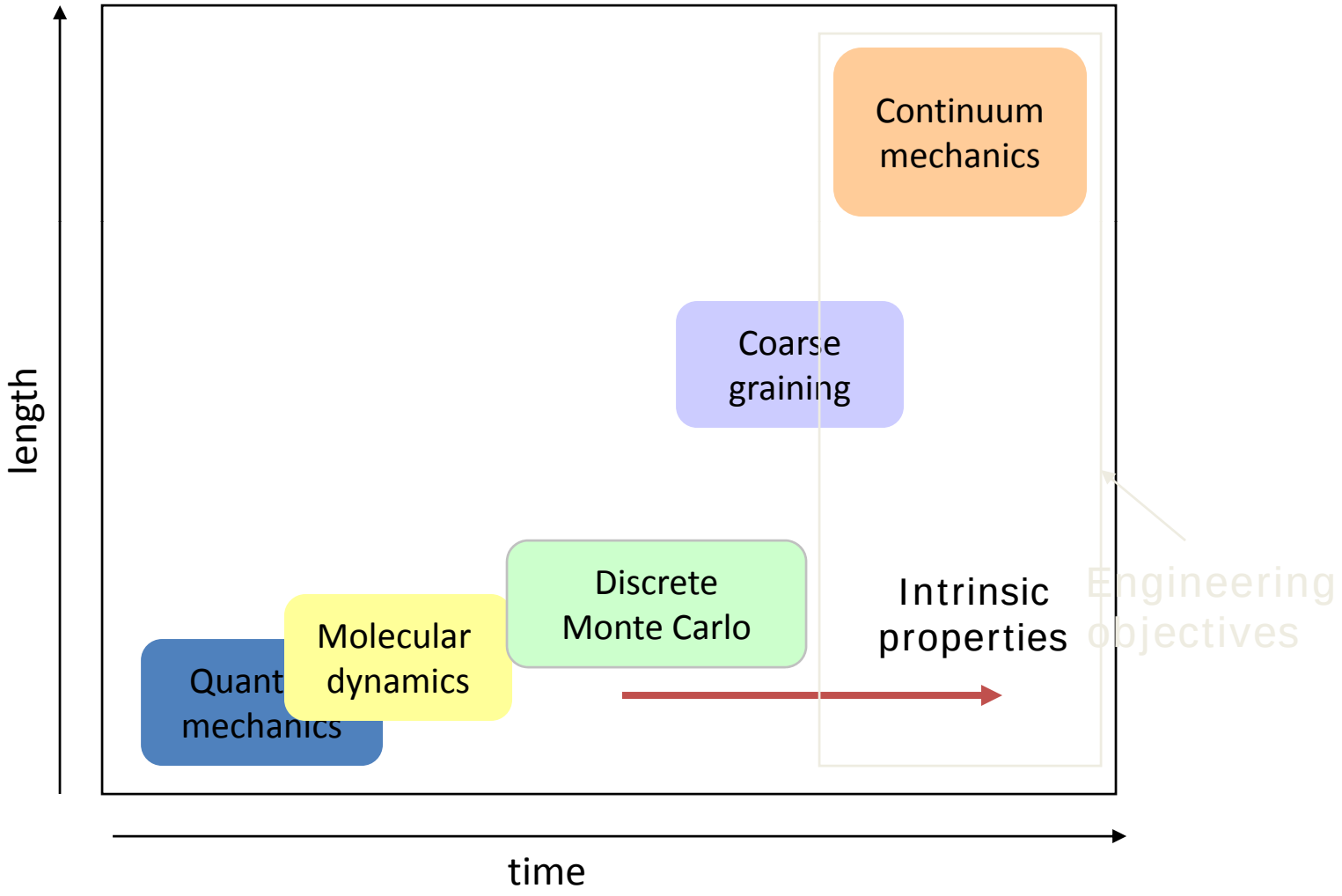
Historically, systems engineering has been applied to systems in which physics and chemistry are well understood. The new challenge is to design systems with only a partial model.

e.g. materials, nanotechnology, synthetic biology

Modeling at across the scales: typical picture



Some modifications to the typical picture

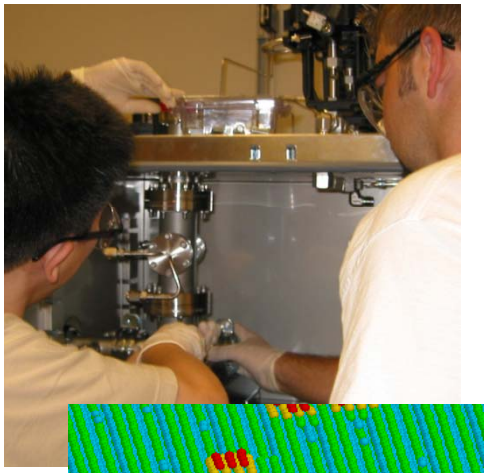


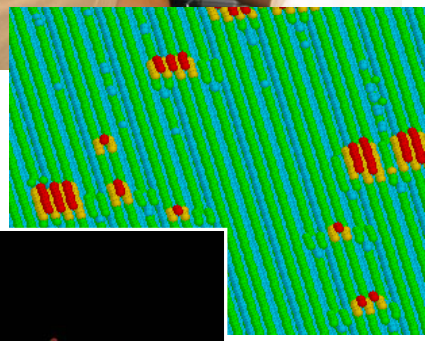
Approach

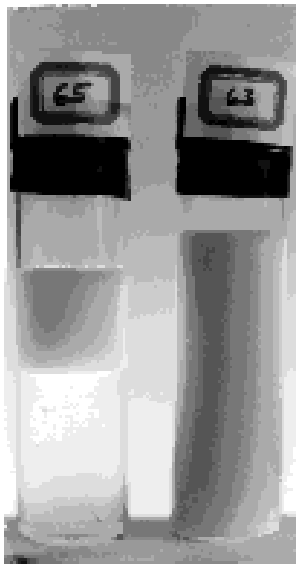
- Barriers to systems engineering in material structure design
  1. Models are not accurate enough
  2. Computational demands of models are too high
  3. In-situ sensing is difficult
- Methodology development
  1. Experimental design
  2. Model reduction
  3. Real-time estimation
- Develop and demonstrate in several specific applications

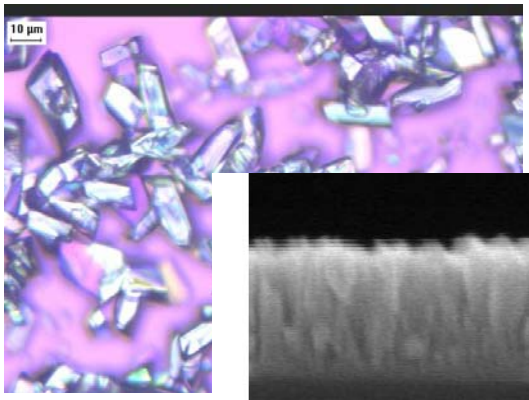
Current practice in materials development

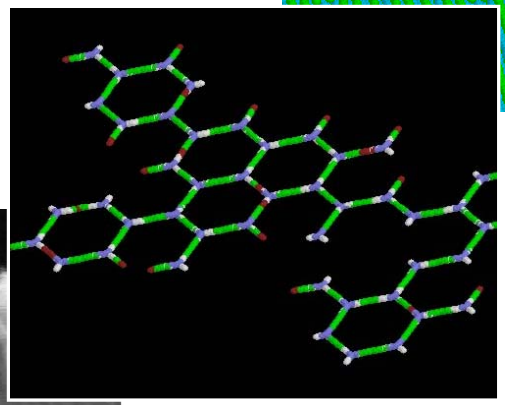
- Design of materials and processes is largely empirical
- Macroscopic models are used in process design, but molecular/microscopic models are not
- Materials properties (advanced materials) require consideration of molecular structure

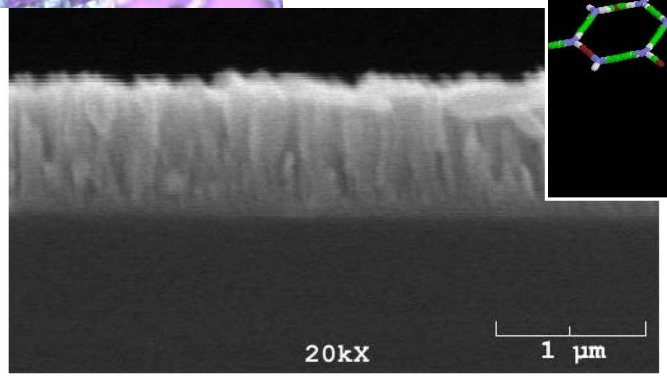


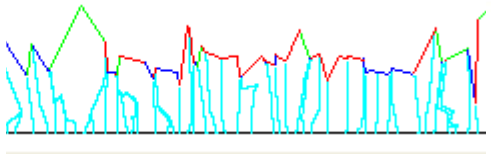


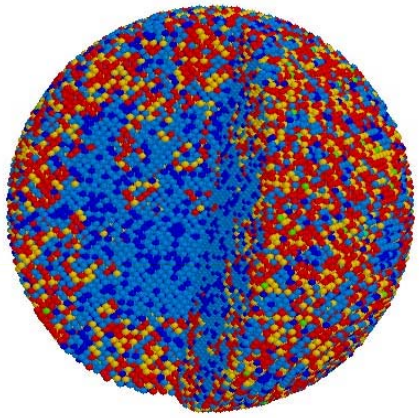












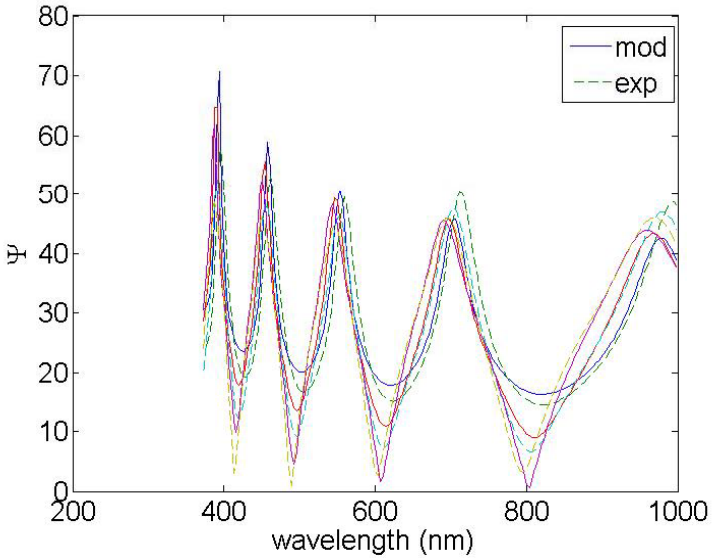
Measurement in materials

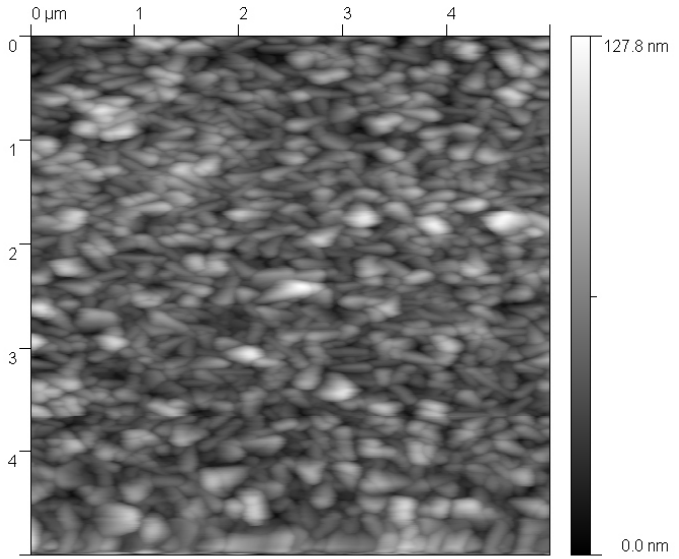
In situ

- Optical: reflection, scattering (UV, Vis, IR)
- Process sensors: temperature, pressure, heat flow (DPC)

Ex situ

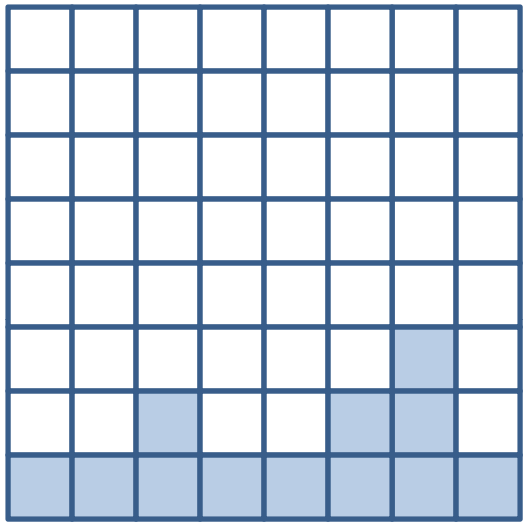
- Optical spectra
- “Images”: AFM, SEM, TEM
- Crystal structure: X-ray diffraction (XRD)
- Size: gel permeation chromotography (GPC), light scattering
- Composition: X-ray photoemission spectroscopy (XPS)





Lattice model with kinetics

- Events
  1. Adsorption: occupy a surface site
  2. Desorption: leave the surface (if have no side neighbors)
- Rates
  1. k_1 = 2 events/s
  2. k_2 = 7 events/s
- Simulate with stochastic simulation algorithm (SSA)



Each lattice site is occupied or not

- $2^{64} = 1.8 \times 10^{19}$

Solid-on-solid assumption

- $9^8 = 4.3 \times 10^7$

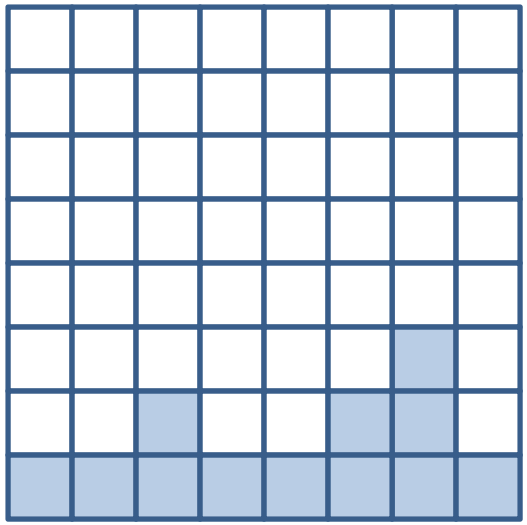
D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *Journal of Physical Chemistry*, **81**(25), 2340-2361 (1977).

Stochastic simulation algorithm

$$\Delta t = \frac{log(x_1)}{\sum_{i=1}^{n_e} k_i N_i}$$

$$\frac{\sum_{j=1}^{p-1} k_j N_j}{\sum_{i=1}^{n_e} k_i N_i} < x_2 \leq \frac{\sum_{j=1}^{p} k_j N_j}{\sum_{i=1}^{n_e} k_i N_i}$$

$$0 < x_1, x_2 < 1$$

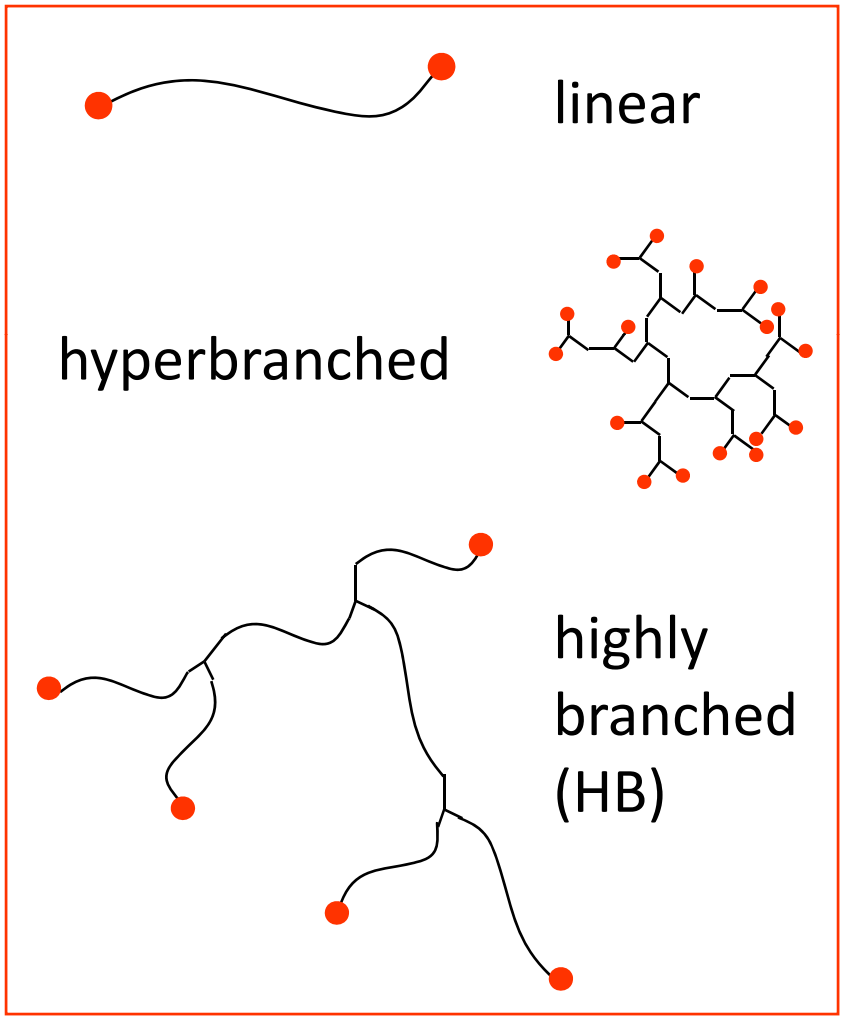


$$\begin{aligned} k_1 &= 2s^{-1} \\ k_2 &= 7s^{-1} \\ N_1 &= 8 \\ N_2 &= 2 \\ k_1 N_1 &= 16s^{-1} \\ k_2 N_2 &= 14s^{-1} \\ k_1 N_1 + k_2 N_2 &= 30s^{-1} \\ x_1 &= 0.95 \\ x_2 &= 0.23 \\ \Delta t &= 0.0017s \\ p &= 2 \end{aligned}$$

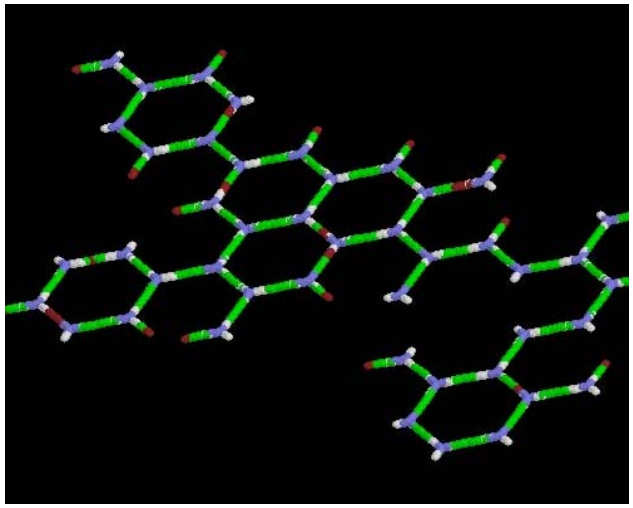
- Remove a surface atom at random (using x_3)
- Recalculate N_i
- Calculate new x_1, x_2, …

Highly branched macromolecules

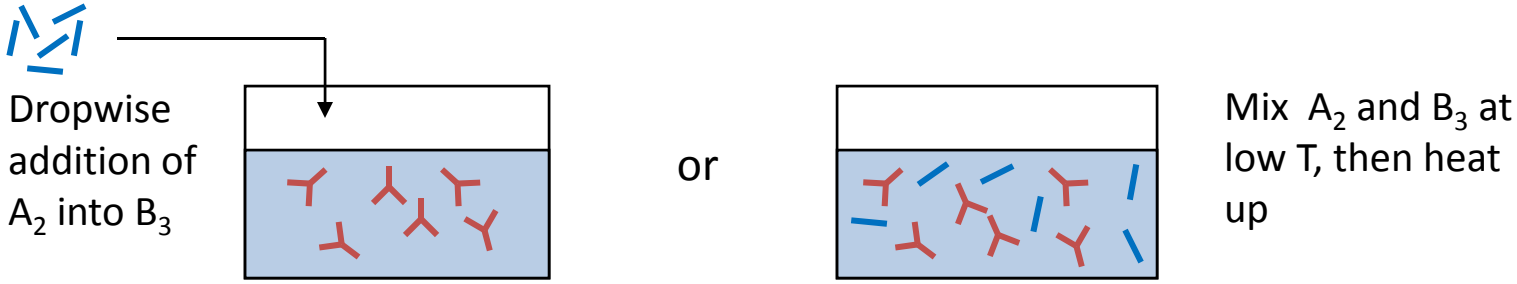
- Increased concentration of chain ends
- Reduced crystallinity
- Increased solubility
- Hydrodynamic volume of highly branched polymers are lower than linear analogs
- Entanglements, leading to improved mechanical properties
- Improved processability



Motivation

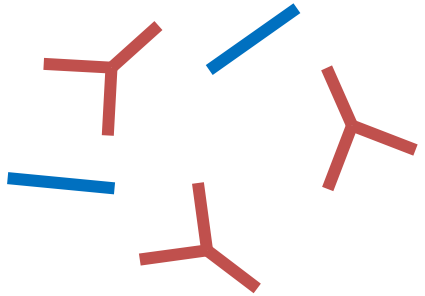


- Relate processing method to molecular structure
  - Promotes better understanding
  - Can be used for design and optimization
- Analytical results require too many limiting assumptions
  - Flory assumes no cyclization
  - Dusek (1994) includes cyclization in a Monte Carlo simulation
- Would like to consider variations in process inputs, for example:



Other possibilities: vary concentration, alternate dropwise addition of A_2 and B_3, heat up during the reaction

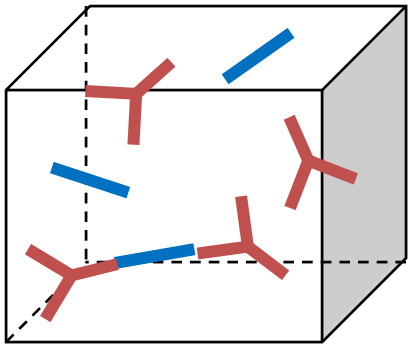
Modeling



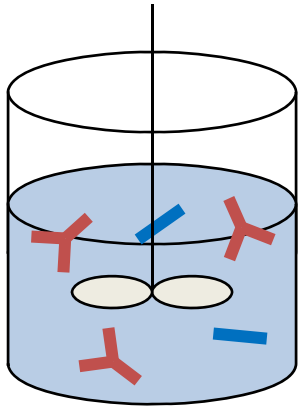
- Want the simplest model to explain the data
- A2 and B3 monomers
  - A and B react, A does not react with A, or B with B
  - Second order kinetics on numbers of A groups and B groups
  - A and B react with one rate if they are in the same molecule, and with another rate if they are in different molecules
- Cyclization parameter: encompasses reactivities, dilution, and mixing

  Cyclization parameter is higher when
  - Per molecule rate of cyclization is higher
  - Concentration in lower
  - Mixing is lower
- View as a free parameter for the system (not for each simulation)
- No sense of spatial location: well-mixed within some “mixing volume”

Simulations



- Kinetic Monte Carlo (SSA)
  - Evolution as a series of discrete events
  - Directly capture the bonding and configuration of each monomer inside the mixing volume
  - Use simple reactivities for now, but easy to add in complexity
  - No spatial positions (just connections)



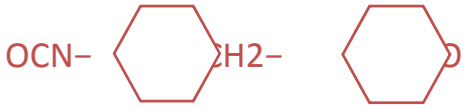
- Comparison to mass action kinetic modeling
  - Captures most of the behavior in the current MC simulations using concentrations
  - Describes the evolution of a single molecular weight (PI=1)
  - Cannot be easily generalized

Another alternative: Population balance model
Can describe MWD, but also difficult to generalize for branching

Experimental Procedure

CH2[OCH2CH(CH3)]x-NH2
|
CH3CH2CCH2[OCH2CH(CH3)]y-NH2
|
CH2[OCH2CH(CH3)]z-NH2

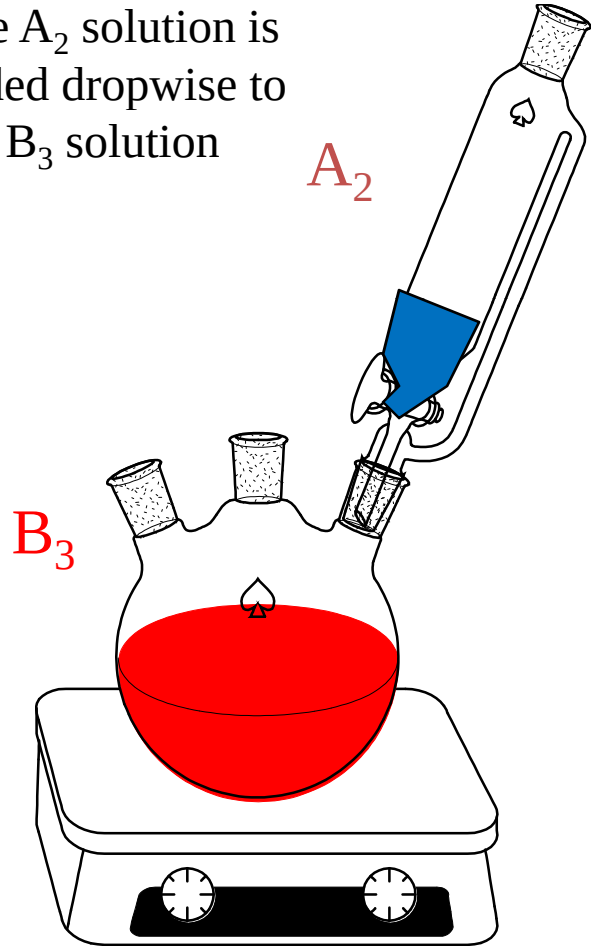
polyoxylalkylenetriamine



Bis(4-isocyanatohexyl)methane

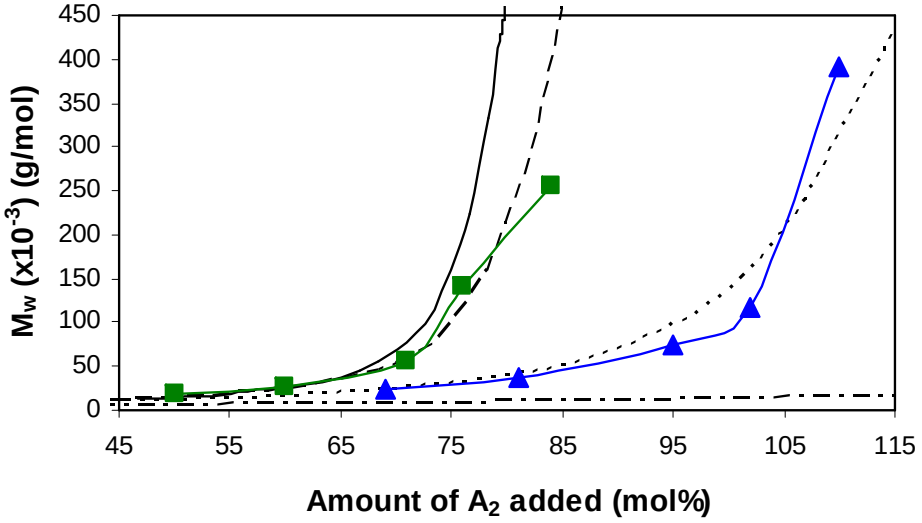
(End capped PTMO)

The A_2 solution is added dropwise to the B_3 solution



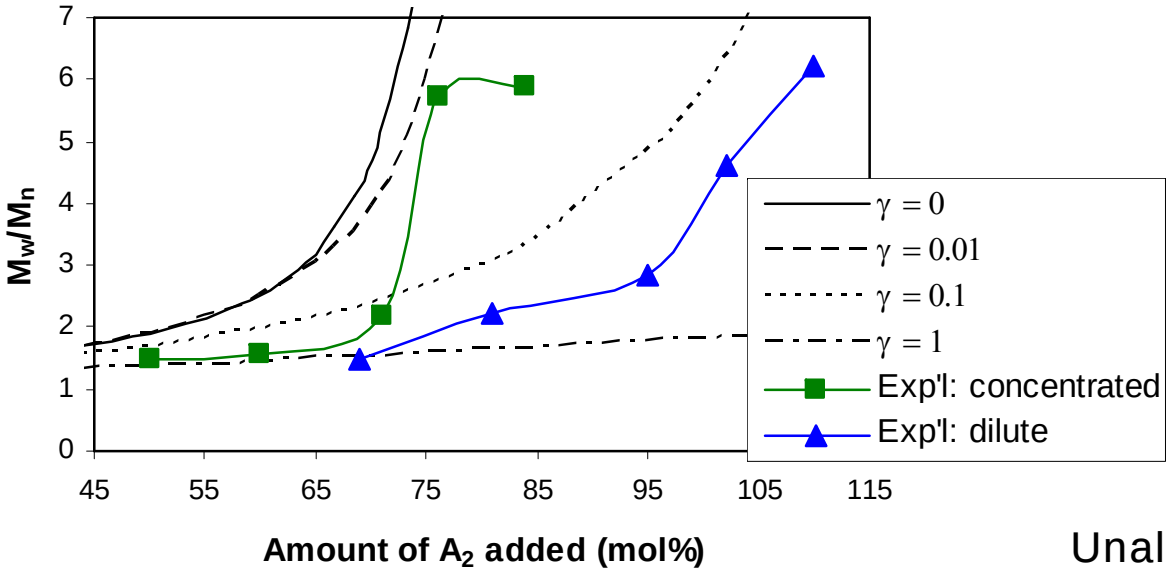
- Polymerization conditions
  - Solvent is isopropyl alcohol
  - Temperature is 23 C
  - Withdraw samples during reaction
- Characterization of molecular weight
  - Size exclusion chromatography (SEC)
  - Laser light scattering (MALLS)

Comparison of molecular weights



Dilution delays the onset of gelation

Distribution of MW is also suppressed by dilution

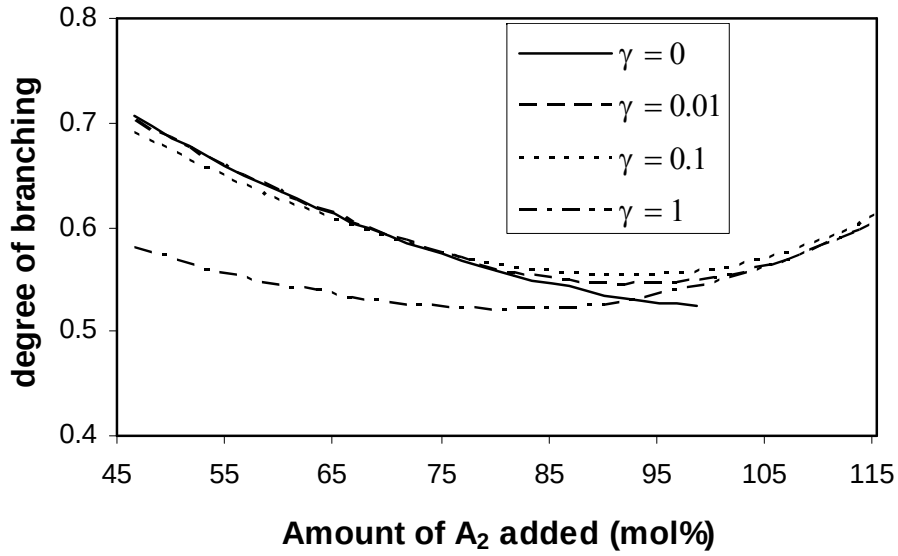


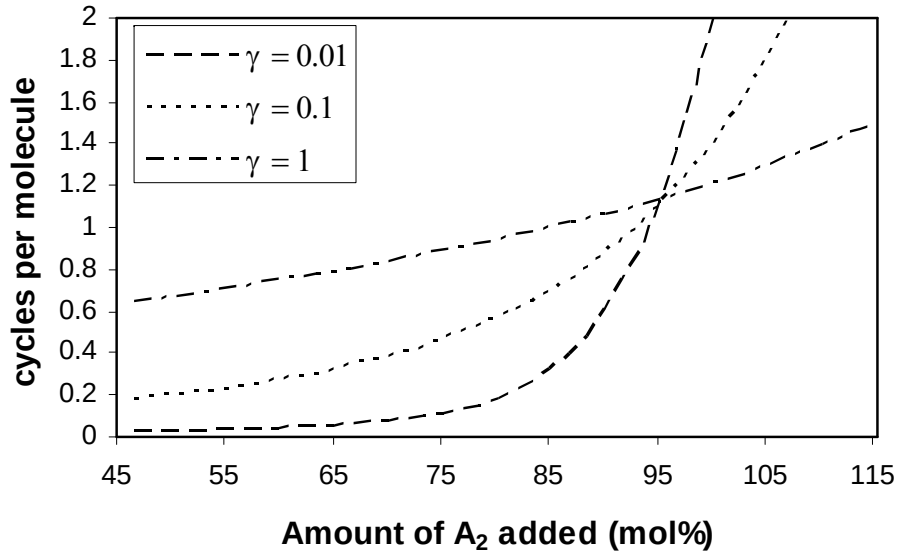
Unal et al, *Polymer* (2005)

Concentration and cyclization ratio

- Concentrations of 10% (dilute) and 25% (concentrated) B_3 solids by weight in IPA
  - Initial concentration ratio of B_3 is 4.1 (assuming additive volumes of B_3 and IPA)
- Approximate ratio of γ from comparison of experiments and simulations is 10 (0.1/0.01)
- For the same kinetics and the same mixing environment, the cyclization ratio γ should be inversely proportional to the concentration of B_3
- Sources of error include the simple kinetics, the uncertainty in mixing environment, and the volume change from A_2 addition

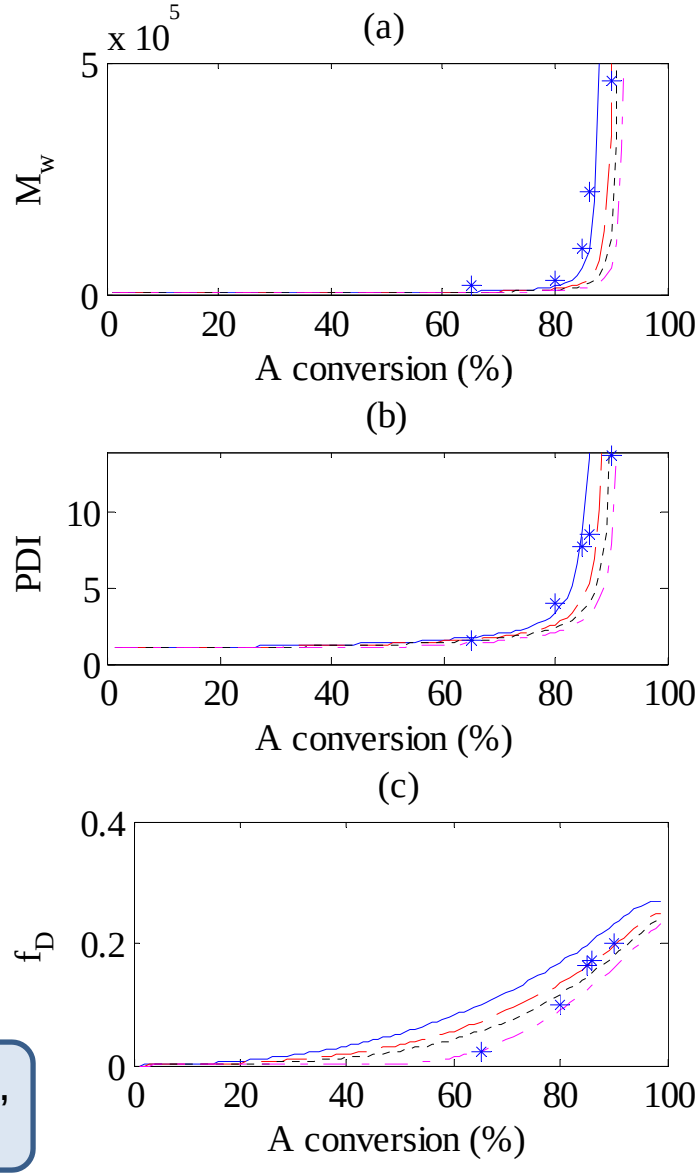
Given the simplicity of the model, this agreement is "good".


degree of branching
0.8
0.7
0.6
0.5
0.4
45
55
65
75
85
95
105
115
Amount of A2 added (mol%)
γ = 0
γ = 0.01
γ = 0.1
γ = 1


cycles per molecule
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
45
55
65
75
85
95
105
115
Amount of A2 added (mol%)
γ = 0.01
γ = 0.1
γ = 1

Modified addition strategy

- A_2+B_3 system
- No solvent → negligible cycle formation
- NMR measurements provide branching structure
  - NMR suggests unequal reactivity of free B_3
- Addition of monofunctional A groups (A_2:B_3:A=1:1:1)
  - Non-intuitive effect
  - Not a robust operating point



Oguz, Unal, Long, and Gallivan, *Macromolecules*, 2007.

What is the state of the system?

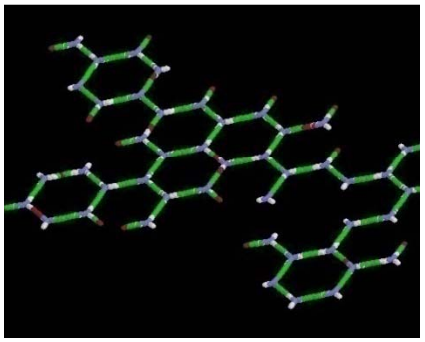
- Molecular simulation is described by a graph
  - Nodes are B_3 monomers
  - Classify each node as dendritic D, linear L, or terminal T
  - Edges are chemical bonds
- If interested in molecular weight distribution, can characterize system by a population balance representation:

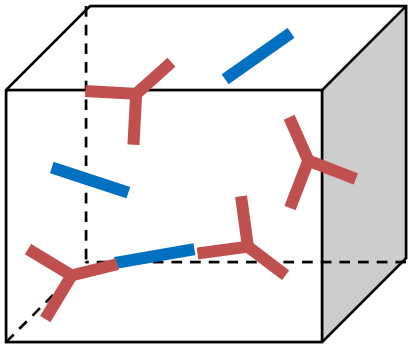
$$P_{D,L,T}(t)$$

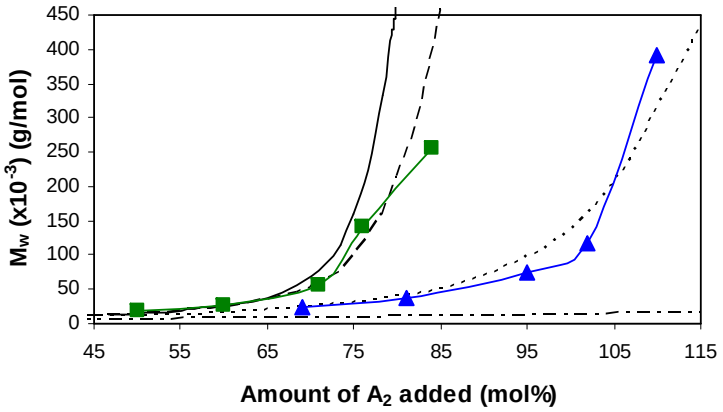
- Molecular weight is D+L+T
- Depending on the kinetics used, reactivity depends on
  - Number of free groups (may be slaved to MW): 2T + L
  - Unequal reactivities and cycle formation complicated the relationship
  - Reactivities could depend on location of group in the graph, but not in our current model

Minimal state depends on the reactivity model used.

Summary





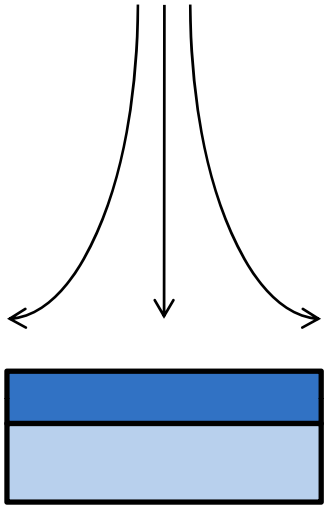


- Good agreement on molecular weight
- Cyclization suppresses molecular weight, and can be promoted using dilution
- Model contains information on molecular structure too
- Need more data to justify a more detailed model
- NMR measurements for molecular structure are underway
- Distance between groups may be considered in the rate of cyclization

Chemical vapor deposition

Commonly used process for depositing thin films

- Thermal CVD
  - Volatile precursor
  - Heated substrate
  - Chemical reactions
  - Formation of solid film
- Case study
  - Yttrium oxide (Y_2O_3) thin films
  - Silicon substrate
  - Polycrystalline



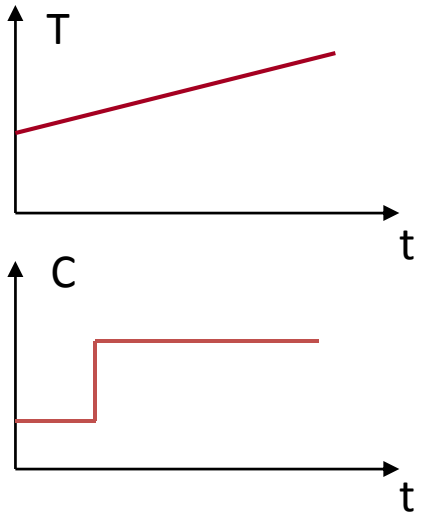
Application	***Microstructure***
Thermal barrier coatings	High grain density, strong
Solid oxide fuel cells	Graded microstructure
Microelectronics gate dielectric	Amorphous
Oxygen sensors	Nanocrystalline

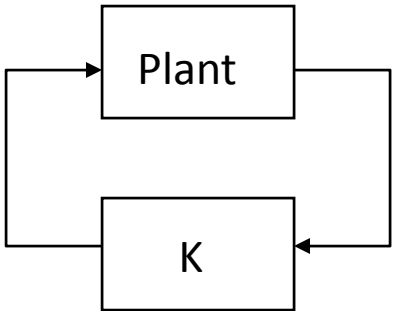
Materials design via process design

Three main steps in the design of a material and process

Should simultaneous consider the entire problem, but can decompose into:

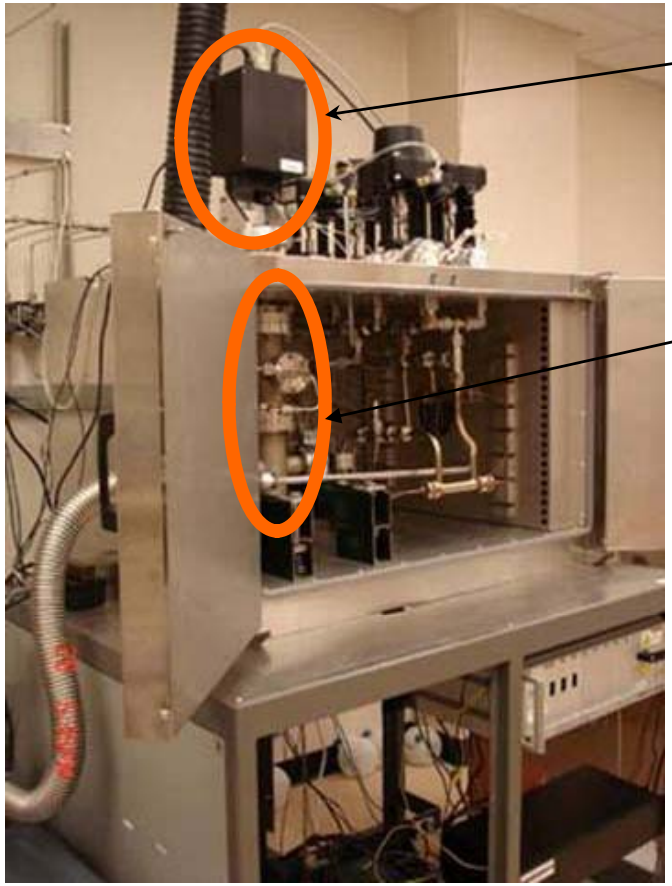
1. Design of hardware (geometry)
2. Design of process settings (open loop control)
3. Correct for disturbances (closed loop control)

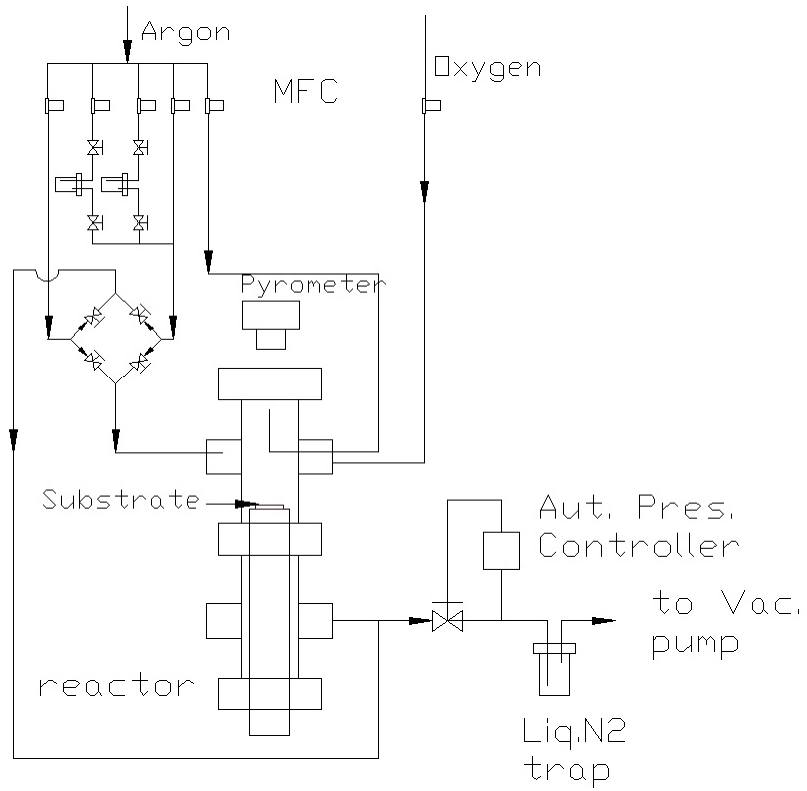




Experimental testbed

Enables case study and demonstration of methods



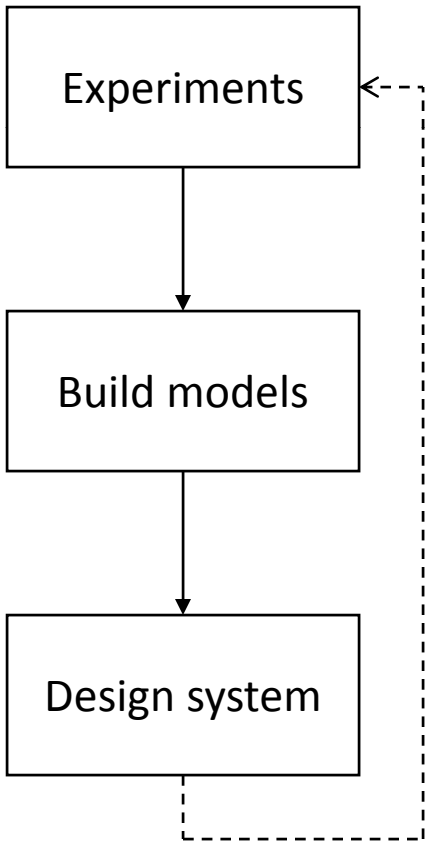


Schematic of CVD testbed

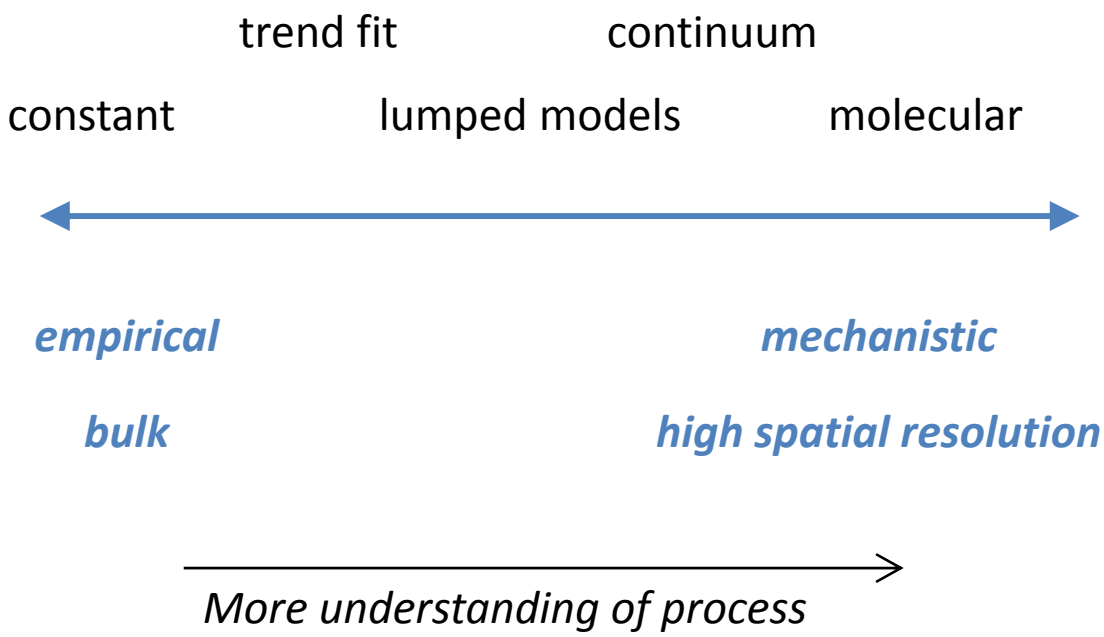
Designing with partial models

Much design work goes on without physics-based models

Ideal case



Models exist at various resolutions:



Best fit parameters

Models often have unknown parameters that can be fit from data

- General case
  - *x* is vector of inputs or "factors"
  - θ is vector of unknown model parameters

$$\hat{y}(x) = f(x, \theta)$$

- Estimate θ from data vectors *(x,y) (*from the *n* experiments)

$$S_r = \sum_{i=1}^{n} (y_i - \hat{y}(x_i, \theta))^2$$

$$\hat{\theta} = \min_{\theta} S_r$$

- When model is linear in the parameters, $\frac{dS_r}{d\theta_j} = 0$ is equivalent to

$$\hat{\theta} = (X^T X)^{-1} X^T y \qquad X_{ij} \equiv \frac{d\hat{y}}{d\theta_j}(x_i)$$

Sensitivity of model to parameters

Uncertainty from the parameters

Uncertainty in the parameters leads to uncertainty in the prediction

- Variance of model error, for the *n* data points

$$\hat{\sigma}_m^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_m(x_i))^2}{\nu_m}$$

$$\nu_m \equiv n - p_m$$

"Degrees of freedom"

- Prediction variance for model *m*

$$\sigma_m(x) \equiv \alpha^m(x)(X^T X)^{-1}\alpha^m(x)\hat{\sigma}_m{}^2$$

$$\alpha^m(x) \equiv \frac{d\hat{y}_m}{d\theta_j}$$

Model discrimination

Goal is to determine which model is "true"

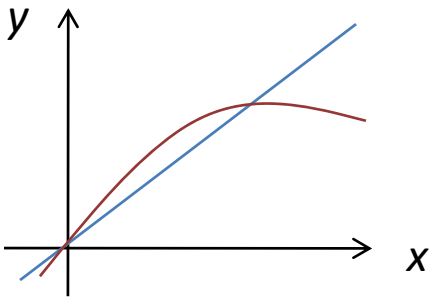
- Bayesian probability
  - Model *m* has p_m parameters
  - A priori probability of model *m* is $P(M_m)$
  - Model is penalized for having more parameters and larger error
  - Number of repetitions of the experiment is v_e

$$P(M_m \mid y, v_e) \propto P(M_m) \times 2^{-p_m/2} \times S_r^{-v_e/2}$$

- Model discrimination
  - Design experiment at settings where model predictions <u>disagree</u> most

$$D_{m,q}(x) = \frac{\left(y_m(x) - y_q(x)\right)^2}{\sigma^2 + \sigma_m^2(x) + \sigma_q^2(x)}$$

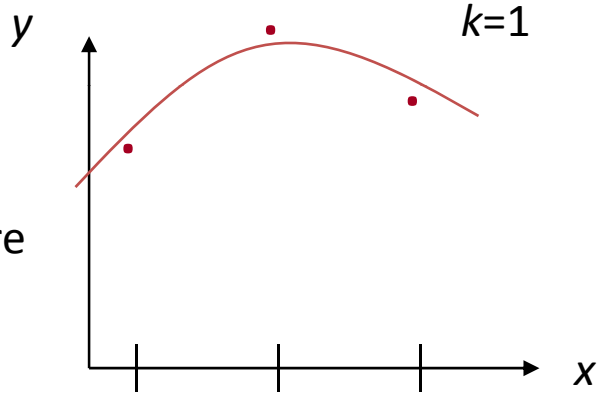
$$x_{\text{exp}} = \max_x D_{m,q}(x)$$

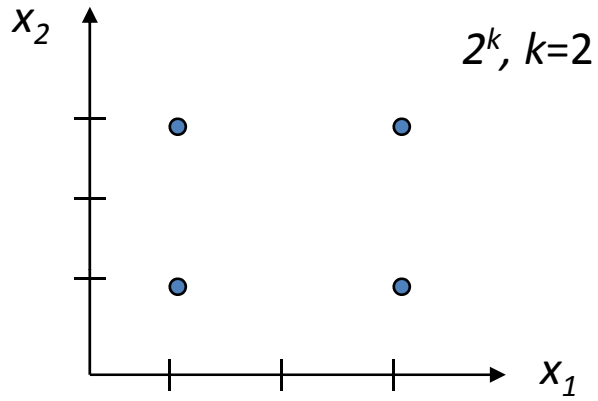


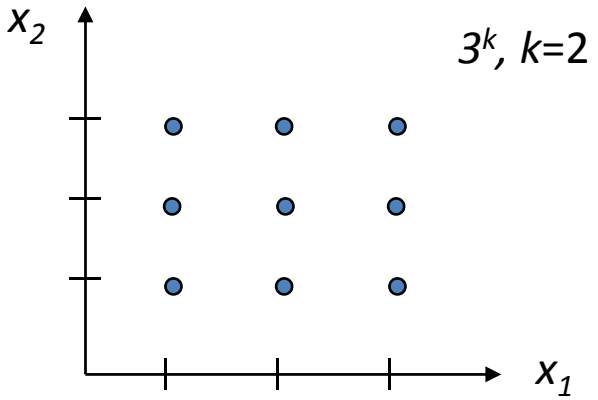
Factorial experimental design

No mechanistic model is required in this approach

- Factorial experimental design
  - Use when experiments are costly and there is little understanding of the process
- Empirical modeling
  - E.g. polynomial
  - Need 3^k for curvature
- Using coded variables
  - More convenient for analysis
  - Under certain model conditions, X^TX is diagonal
  - E.g. $x_{low} = -1$







$$x_c = 2\left(\frac{x - x_{low}}{x_{high} - x_{low}}\right) - 1$$

D. C. Montgomery, Design and Analysis of Experiments, John Wiley (2005).

Example

Need to design experiments based on model structure

- Model $\hat{y}(x) = \theta_1 x^2 + \theta_2 x + \theta_3$ or $\hat{y}(x_c) = \theta_1 x_c^{\,2} + \theta_2 x_c + \theta_3$
- Original variables $x_{low} = 1, x_{high} = 4, x_{mid} = 2.5, x_c = 2\left(\frac{x-1}{3}\right) - 1$

$$-1 \le x_c \le 1$$

- 2^k: $x_{c,1} = -1$, $x_{c,2} = 1$
  - Can't invert
  - $X^T X$ not full rank
  - More parameters than experiments

$$X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}, X^T X = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

- 3^k: $x_{c,1} = -1$, $x_{c,2} = 0$, $x_{c,3} = 1$
  - Full rank
  - Can compute unique θ

$$X = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, X^T X = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 3 \end{bmatrix}$$

Model based experimental design

Is factorial design always "best"?

- Should design experiments to be optimal

- Need to specify what is desired
  - Best parameter estimates
  - Best model discrimination
  - Most helpful in finding maximum of y

- Alphabetic optimal designs
  - Most common is D-optimal
  - Gives most overall information on all parameters

$$x_{\text{exp}} = \max_{x} \left| X^T X \right|$$

  - Consequently, X^TX is easier to invert and parameter estimates are less correlated

Example

- Model $$\hat{y}(x) = \theta_1 x + \theta_2$$

- Design two experiments using the D-optimal method

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} x_1^2 + x_2^2 & x_1 + x_2 \\ x_1 + x_2 & 2 \end{bmatrix}$$

$$\left| X^T X \right| = 2(x_1^2 + x_2^2) - (x_1 + x_2)^2 = x_1^2 + x_2^2 - 2x_1 x_2$$

- To maximize X^TX, x_1 and x_2 should be large and opposite in sign (same as 2^k factorial with coded variables)

Desired features in new approach

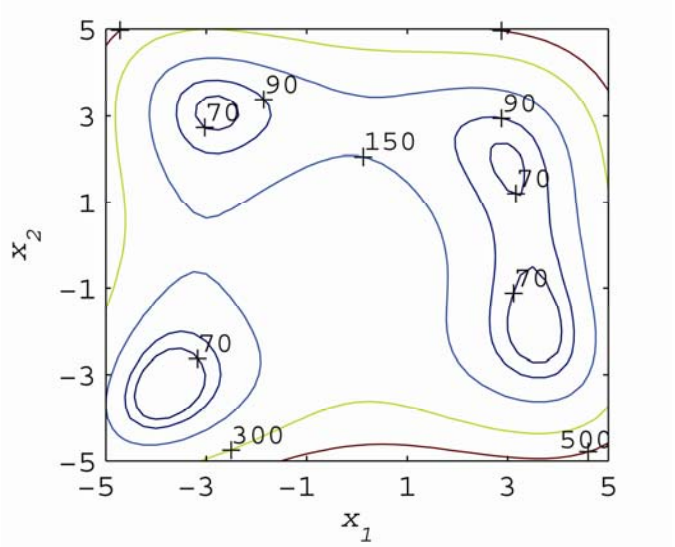
Want to retain the best features of both approaches

1. Mathematical and statistical underpinnings
2. Tractable computation
   - e.g. 1 day
3. Logic consistent with the empirical design approach
   - e.g. include design objectives
4. Transition from low to high resolution models as knowledge is gained
5. Tradeoff between exploration and refinement
   - Global versus local minima

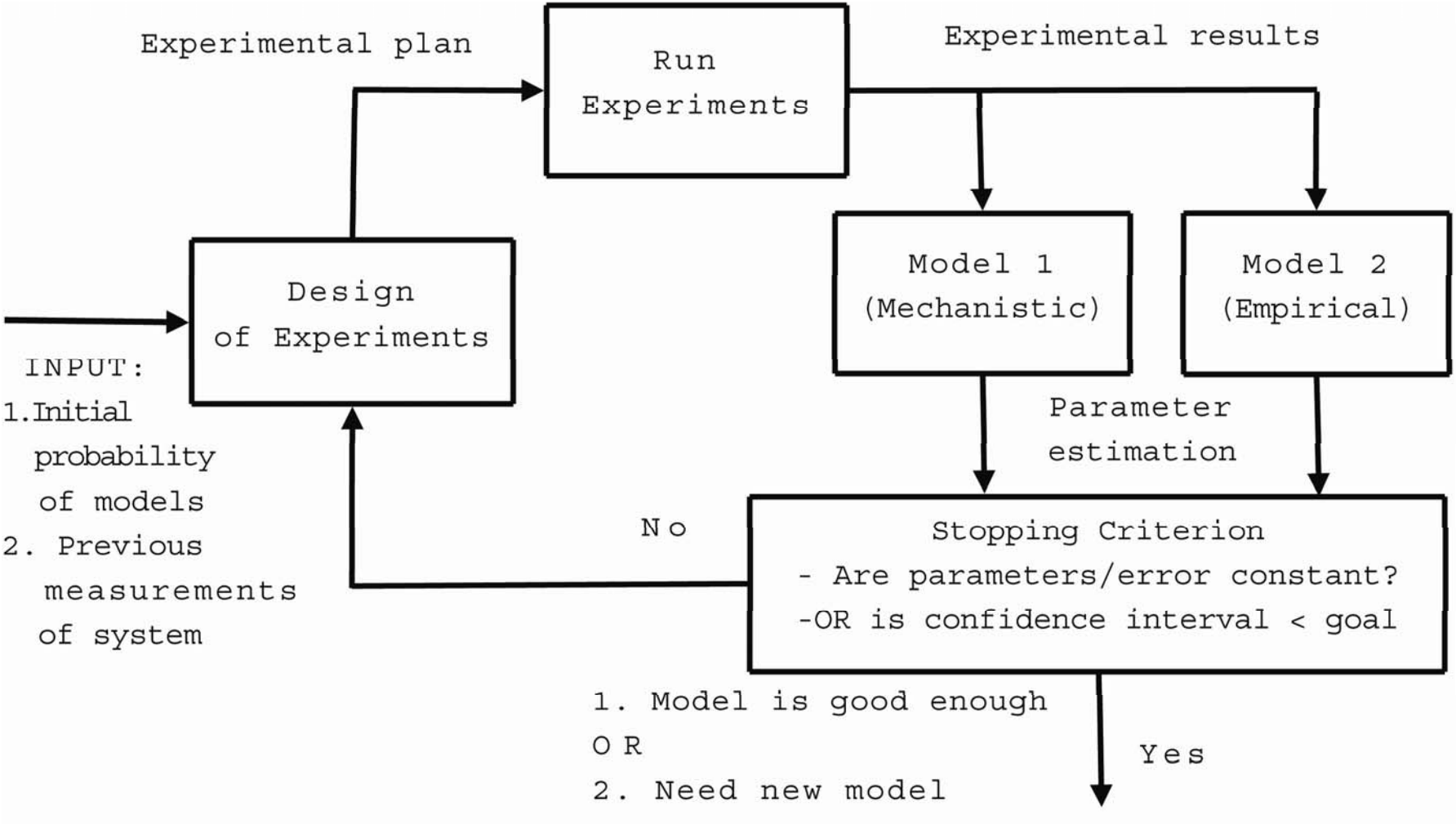
Challenges and open questions

How do we quantify the usefulness of an experiment?

- Experimental design criterion
  - Multiple and competing objectives
- Local minima
  - Parameter estimation
  - Selection of experiments
- Need to avoid repeating experiments
  - Kriging / spatial statistics
  - e.g. batch-wise design: multiple local minima
- Effect of initial experiments
  - Rate of convergence
  - Steady state



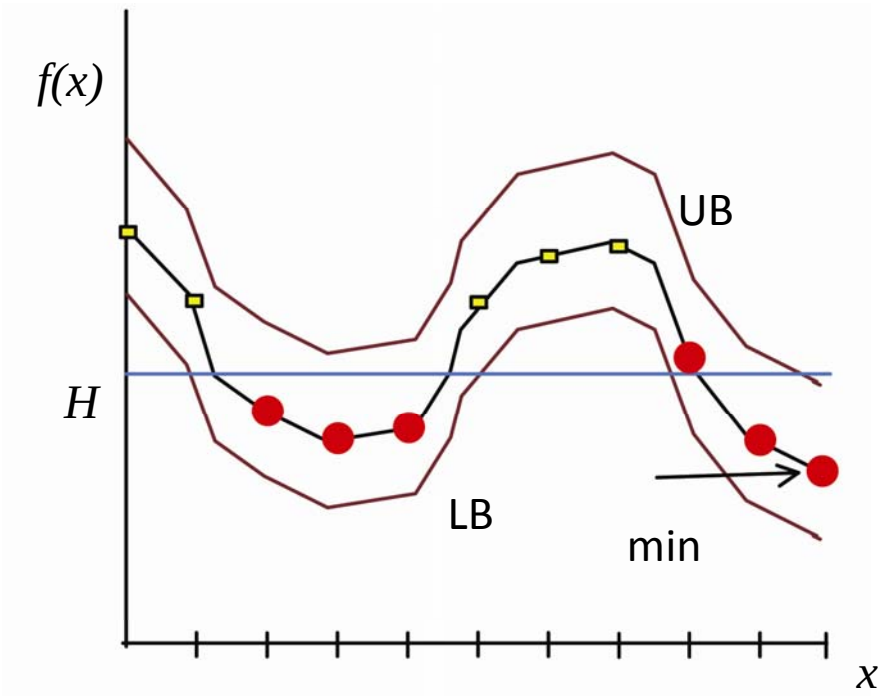
Conceptual Approach



Restrict new experimental points

Only perform experiments at potential optimal points

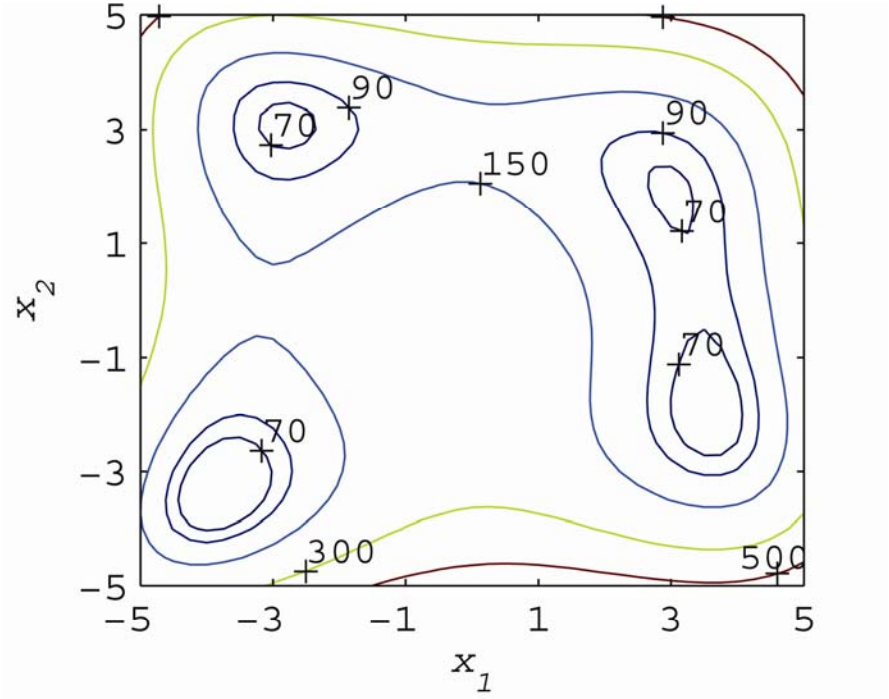
- We don't expect any model to be accurate over the entire domain
- We use confidence intervals to identify regions of potential optima
- The next experiment may only be performed within this region
- Rationale:



By performing experiments at/near the optimum, our models will improve there, and this cycle will converge to the true optimum.

Case study 1: polynomial

Multiple minima in the interior, linear in parameters

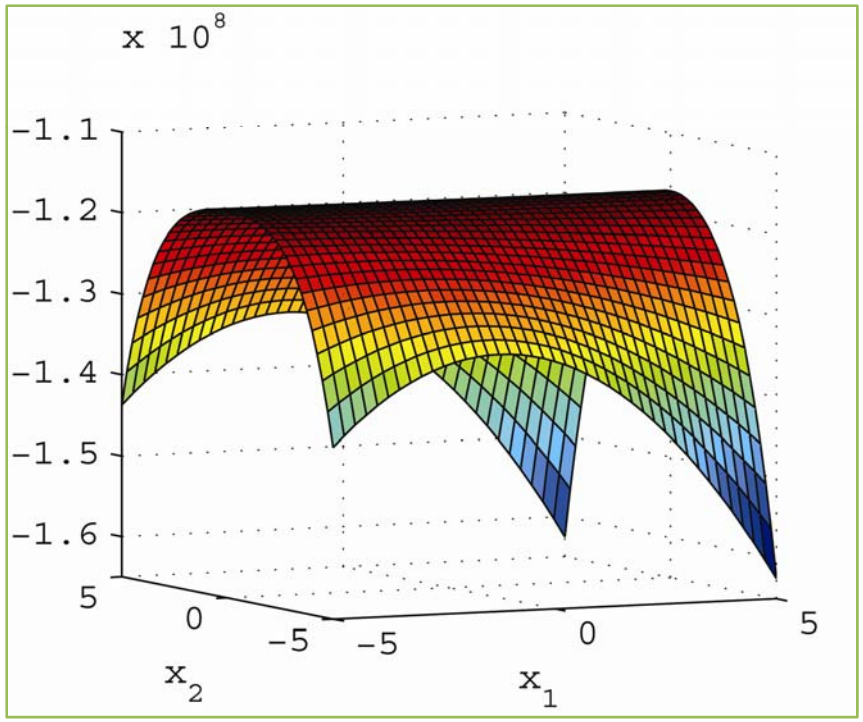


$$
\begin{aligned}
y^{MHF}(x) &= x_1^4 + x_2^4 - 21x_1^2 + 2x_1^2x_2 + 2x_1x_2^2 - 13x_2^2 - 13x_1 - 19x_2 + 227 \\
y_1(x) &= x_1^4 + x_2^4 - 21x_1^2 + 2x_1^2x_2 + \theta_1 x_1x_2^2 - 13x_2^2 + \theta_2 x_2 + \theta_3
\end{aligned}
$$

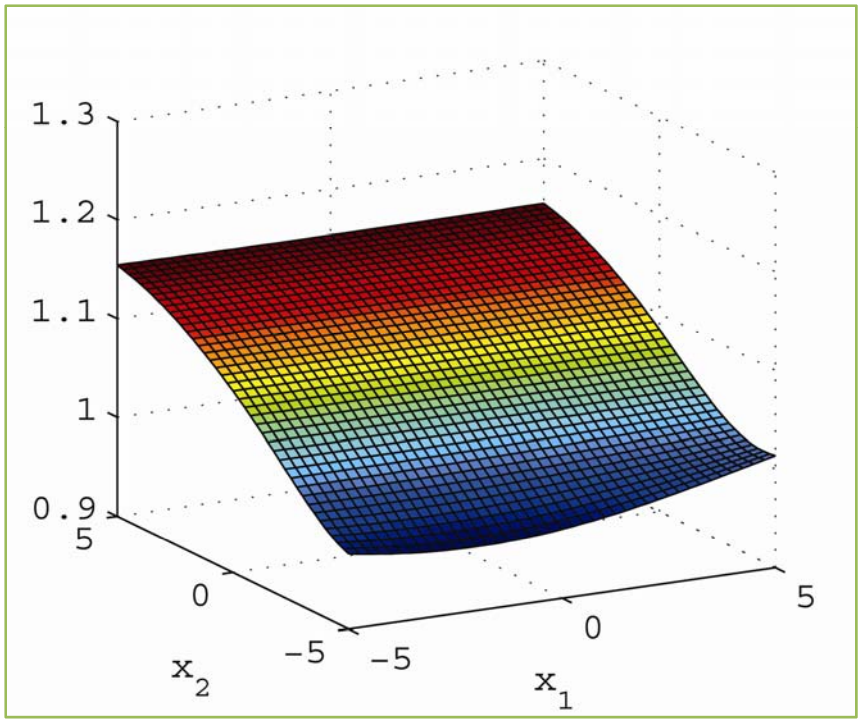
Experimental design surfaces

D-optimal samples at corners, P-optimal along the edge

D-optimal



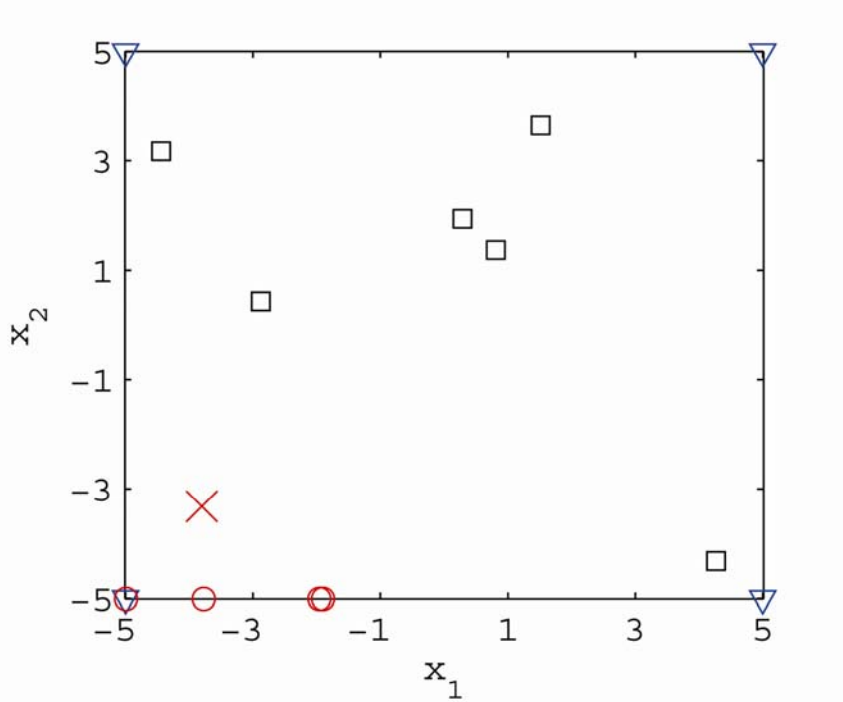
P-optimal



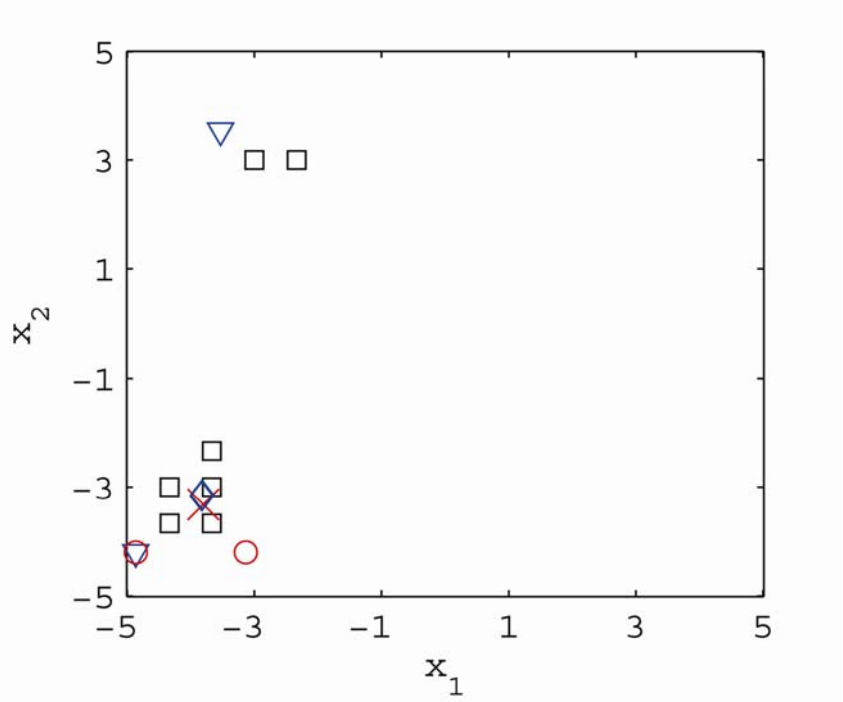
Case study 1: Experimental points

Restriction brings experimental points near global and local optima

With no restrictions



Restricted to potential optima



D-optimal (triangle), P-optimal (circle), random (square), true minimum (x)

Film growth case study

Achieve desired film structure and processing time

- Process inputs
  - temperature *T* (873-1073 K)
  - concentration *C* (0.3-1.5 mol/m^3)
- Goal: achieve desired grain density N_{isl} and time t_f
- Initial experiments: 2^2 factorial + 2 center points

$$
\begin{aligned}
\frac{dN_1}{dt} &= F(1-\kappa) - (\rho+1)K_{nuc}(\eta, T, E_i, E_d, N_1) - K_{agg}(\eta, T, E_i, E_d) \\
\frac{dN_{isl}}{dt} &= K_{nuc}(\eta, T, E_i, N_1) \\
F &= aCexp\left(-\frac{E_a}{RT}\right) \\
f(T, C) &= (t - t_{goal})^2 + 10^6(N_{isl} - N_{goal})^2
\end{aligned}
$$

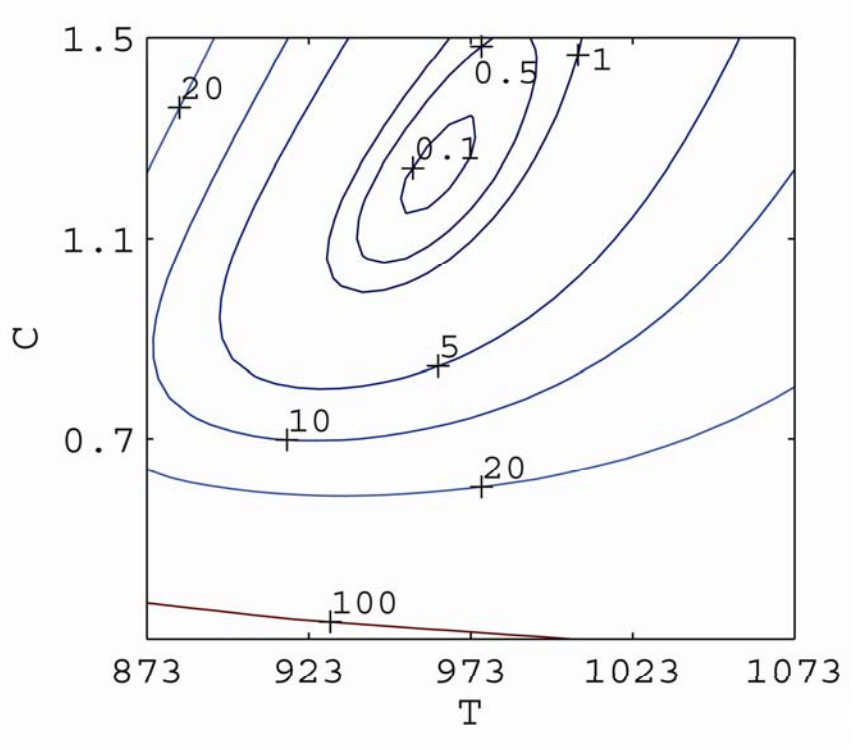
- Four candidate models for flux *F*

Objective function for film growth

A single optimum set of T and C exists in the interior.

Candidate flux models:

$$
\begin{aligned}
F_1 &= \theta_1 \\
F_2 &= \theta_1 C + \theta_2 \\
F_3 &= \theta_1 + \theta_2 T + \theta_3 C \\
F_4 &= 10 C exp\left(\frac{\theta_1}{RT}\right)
\end{aligned}
$$



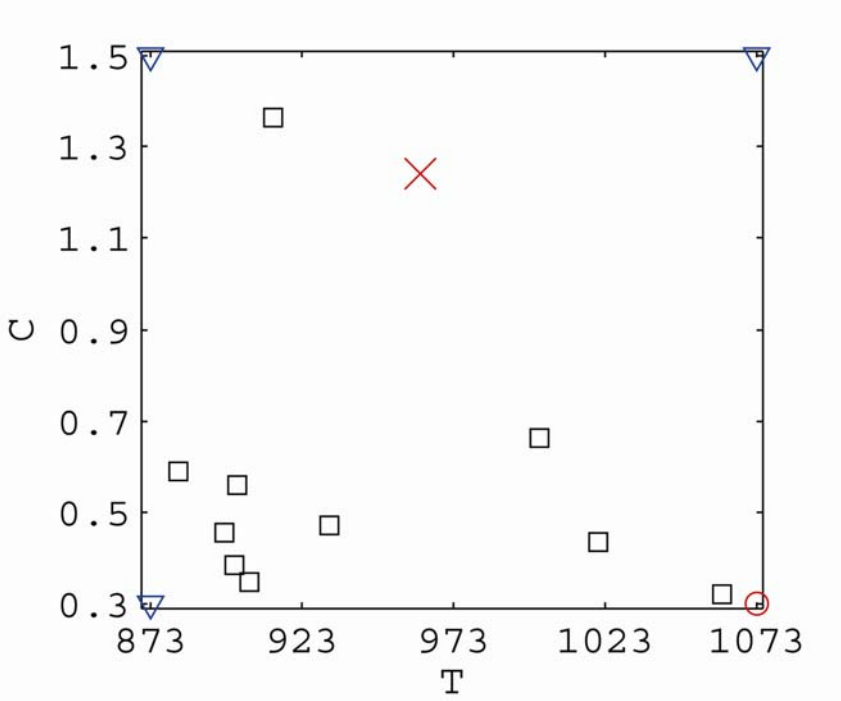
Model selection:

$$
\begin{aligned}
P(M_j|Y, MSE_j) &= P(M_j) \times 2^{-p_j/2} \times MSE_j^{-\nu_e/2} \\
MSE &= \frac{\sum_{i=1}^{n}(y(x_i) - \hat{y}_j(x_i))^2}{n}
\end{aligned}
$$

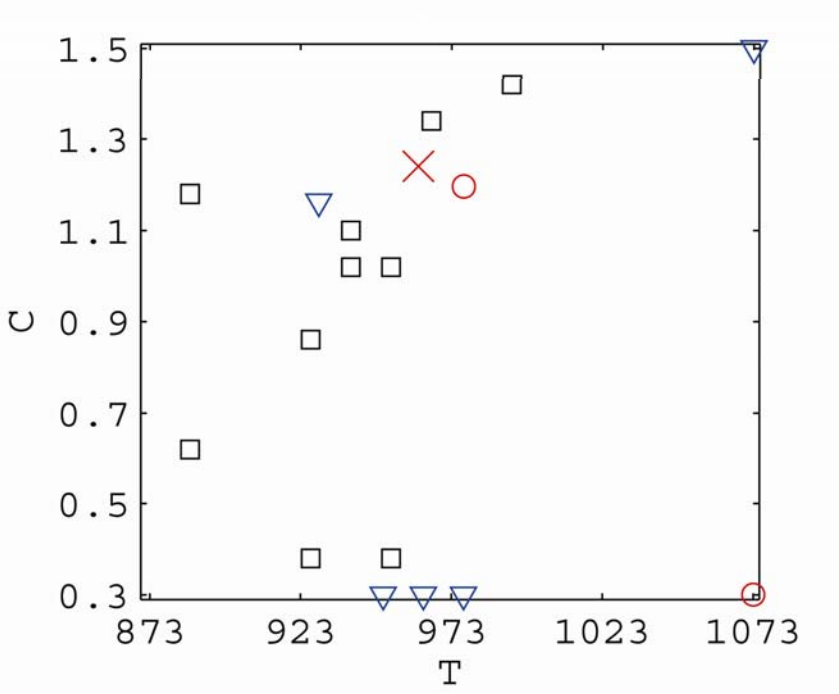
Film growth experimental points

Comparison of D-optimal, P-optimal, and random sampling

With no restrictions



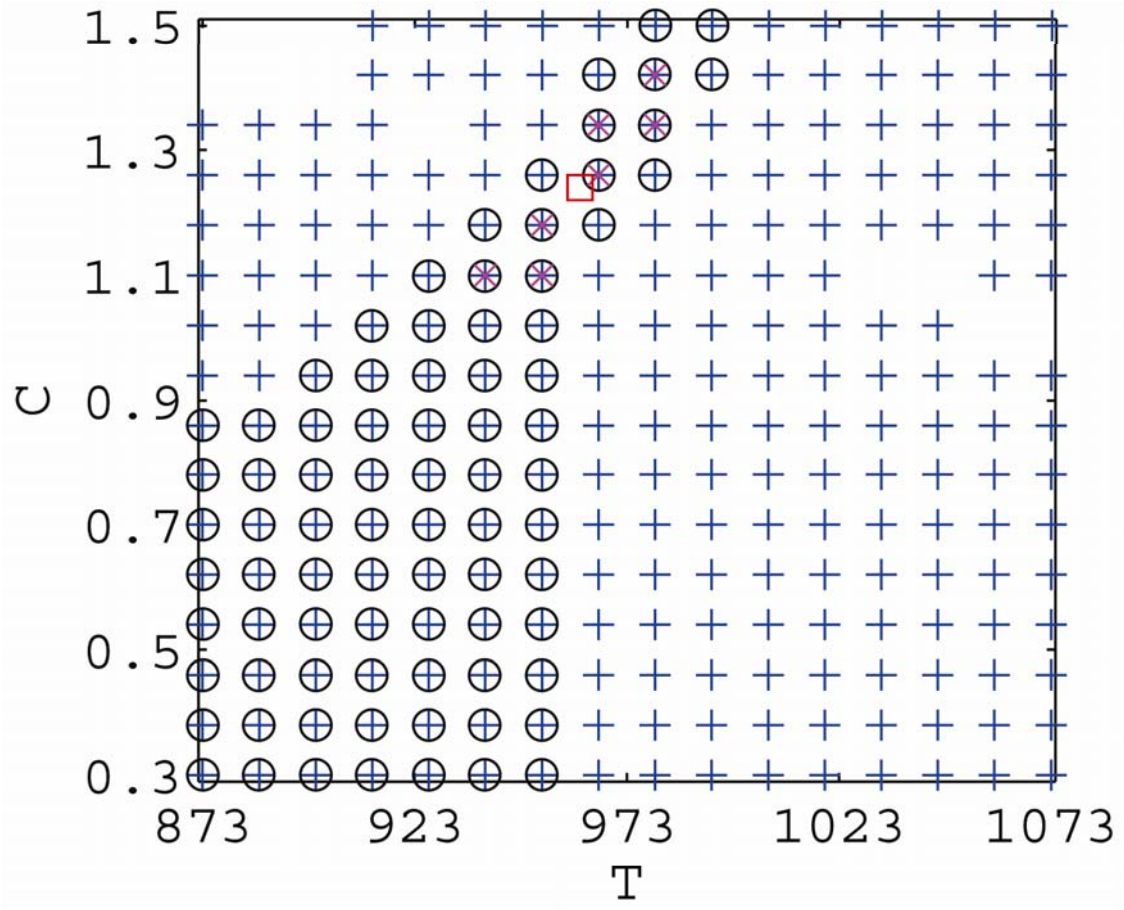
Restricted to potential optima



D-optimal (triangle), P-optimal (circle), random (square), true minimum (x)

Evolution of potential optima

As experiments proceed, region shrinks down near the true optimum



Iteration	
1	+
4	o
7	x

Overall recommendation: Random selection of points from the potential optima. Works as well as the greedy method.

No restriction to potential optima

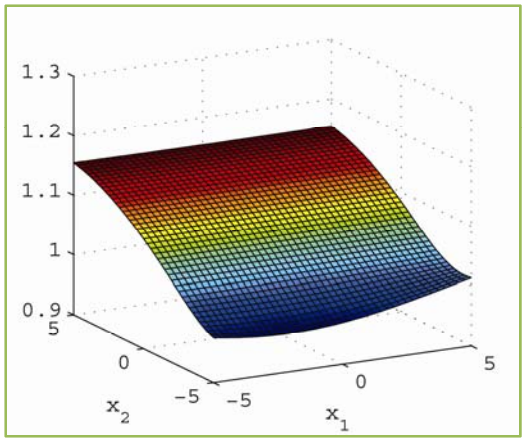
	$\gamma = 5$			$\gamma = 10$			$\gamma = 20$		
	D	P	Rand	D	P	Rand	D	P	Rand
$f_{j^*}(\hat{x})$	0.05	9×10^{-9}	0.02	0.07	5×10^{-8}	0.03	0.02	0.01	0.04
$\sigma_{j^*}(\hat{x}_{j^*})^2$	221	10.8	26.8	304	141	108	754	637	515
$CI(\hat{x}_{j^*})$	7.40	1.95	2.80	9.03	5.24	5.46	16.1	13.6	9.93
MSE_{j^*}	169	7.1	8.48	255	125	33.6	676	579	138
$\hat{\sigma}^2$	675	28.4	30.6	907	401	115	2190	1800	434
iter	8.1	6.7	8.8	8.4	7.2	9.0	7.7	7.4	9.4
$\hat{T}$, % error	1.54	0.47	0.02	2.91	0.46	0.2	1.43	0.05	0.03
$\hat{C}$, % error	9.19	2.85	0.19	17.25	2.82	1.49	8.70	0.35	0.03

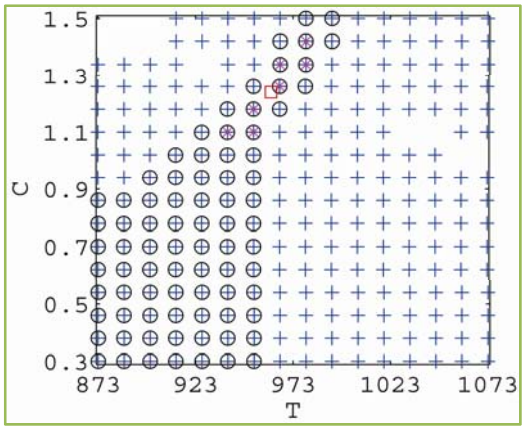
With restriction to potential optima

	$\gamma = 5$			$\gamma = 10$			$\gamma = 20$		
	D	P	Rand	D	P	Rand	D	P	Rand
$f_{j^*}(\hat{x})$	0.12	2×10^{-5}	2×10^{-8}	0.03	2×10^{-3}	2×10^{-8}	0.03	6×10^{-4}	2×10^{-3}
$\sigma_{j^*}(\hat{x}_{j^*})^2$	392	106	21.3	331	150	67.6	885	738	845
$CI(\hat{x}_{j^*})$	3.98	2.62	2.64	10.7	6.11	4.45	18.0	15.4	10.4
MSE_{j^*}	227	65.8	12.5	248	130	37.3	774	677	133
$\hat{\sigma}^2$	933	223	43.1	894	436	120	2440	2110	396
iter	7.4	7.4	8.4	7.3	7.5	9.0	7.5	7.5	9.2
$\hat{T}$, % error	0.03	0.15	0.14	0.31	0.30	0.05	1.19	0.07	0.03
$\hat{C}$, % error	0.36	0.88	1.53	2.13	1.81	0.77	7.47	0.70	0.41

Summary: experimental design

- The proposed approach locates the best design as well as the greedy approach, while also performing exploration for model building.
- A unified approach to experimental design combines aspects of empirical and mechanistic methods, with a consistent rationale.
- Experimental implementation
- Batchwise sequential experimental design may be preferable to sample multiple local minima
- A probabilistic foundation will be explored.





Motivation for *in situ* sensing

Sensing is limited in thin film processing

- *In situ* sensing for thin film deposition is necessary for process monitoring and control
- *In-situ* measurements during process are limited (Edgar et al. 2000)
  - Noisy and inaccurate
  - Expensive and difficult to integrate
  - Difficult to interpret
- Emissivity correcting pyrometer
  - Emission + reflectance measurement
  - One normal incidence view port
  - Currently limited to optically smooth surfaces

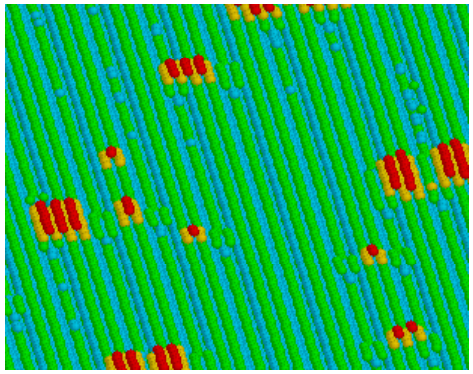
- Goals:
  1. Extend the applicability of a commercially available sensor from smooth surfaces in MBE to rough surfaces in CVD (Breiland 1995)
  2. Use estimation theory to improve measurements for smooth and rough surfaces

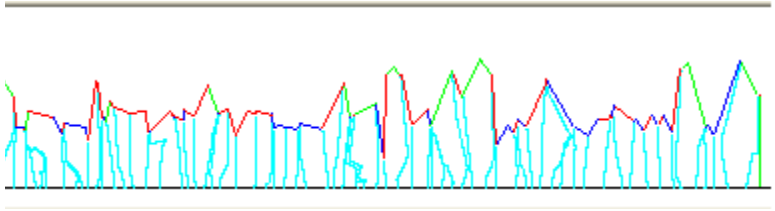
Process model

Key question: What model will be most useful?

- G, G_e, n_1, k_1, n_2, k_2 are expected to drift slowly
  - Model them as parameters
- h and h_e are the time integrals of G and G_e
- Relevant physical phenomena
  - Fluid flow, gas phase reactions
  - Surface chemistry
  - Grain boundary motion
  - Crystal growth and dislocations
  - Chemical composition and incorporation of impurities

$$\begin{bmatrix} h \\ G \\ h_e \\ G_e \\ n_1 \\ k_1 \\ n_2 \\ k_2 \end{bmatrix}_{j+1} = \begin{bmatrix} h + G\Delta t \\ G \\ h_e + G_e\Delta t \\ G_e \\ n_1 \\ k_1 \\ n_2 \\ k_2 \end{bmatrix}_{j}$$





Sensor model

The effective medium model is a common way to model surface microroughness.

$$y = \left| \frac{r_{12} + r_{23}e^{-i2\delta_2} + r_{34}e^{-i2(\delta_2+\delta_3)} + r_{12}r_{23}r_{34}e^{-i2\delta_3}}{1 + r_{12}r_{23}e^{-i2\delta_2} + r_{12}r_{34}e^{-i2(\delta_2+\delta_3)} + r_{23}r_{34}e^{-i2\delta_3}} \right|^2$$

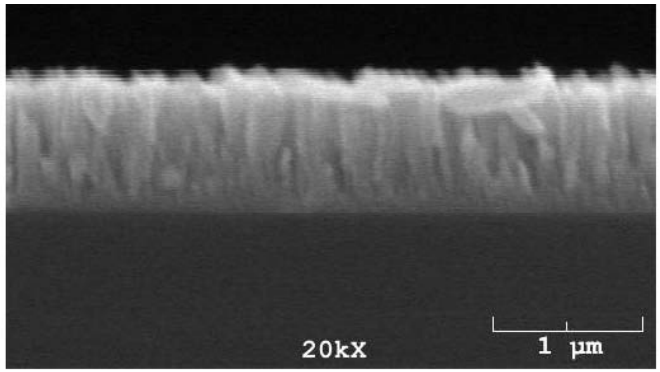
$$r_{ij} = \frac{\hat{n}_i - \hat{n}_j}{\hat{n}_i + \hat{n}_j}$$

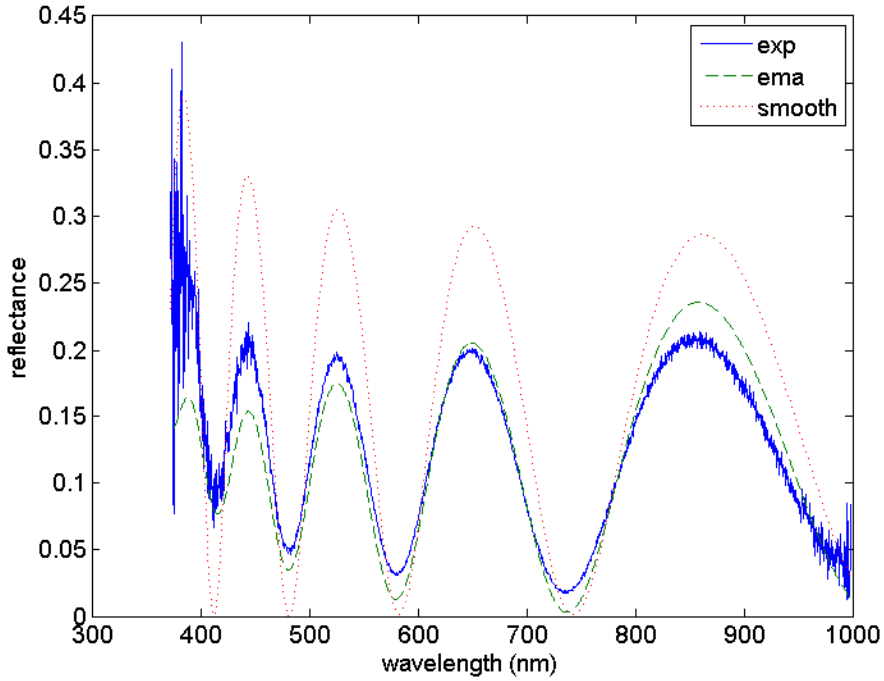
$$\delta_p = 2\pi\hat{n}_p h_p/\lambda$$

$$\hat{n}_e = \frac{\sqrt{\sqrt{\Omega^2 + 8\hat{n}_1^2\hat{n}_3^2} + \Omega}}{2}$$

$$\Omega = (\hat{n}_1^2 + \hat{n}_3^2)/2$$

$$\hat{n} = n - jk$$





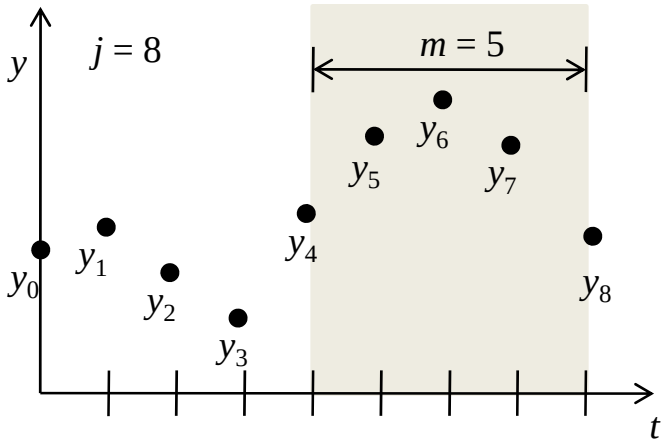
Moving horizon estimation

The general case of least squares estimation

$$\min_{x_{j-m+1},\cdots,x_j} \left[(x^e_{j-m+1})^T P^{-1}_{j-m+1|j-m} x^e_{j-m+1} + \sum_{l=j-m+1}^{j} v_l^T R^{-1} v_l + \sum_{l=j-m+1}^{j-1} w_l^T Q^{-1} w_l \right]$$

such that

$$\begin{aligned} x^e_{j-m+1} &= x_{j-m+1} - x_{j-m+1|j-m} \\ v_l &= y_l - g(x_l) \\ w_l &= x_{l+1} - f(x_l) \end{aligned}$$



Special cases

- RLS
  - $Q = 0$, $P_{j+1|j} \to \infty$
- Modified MHE
  - $Q = 0$

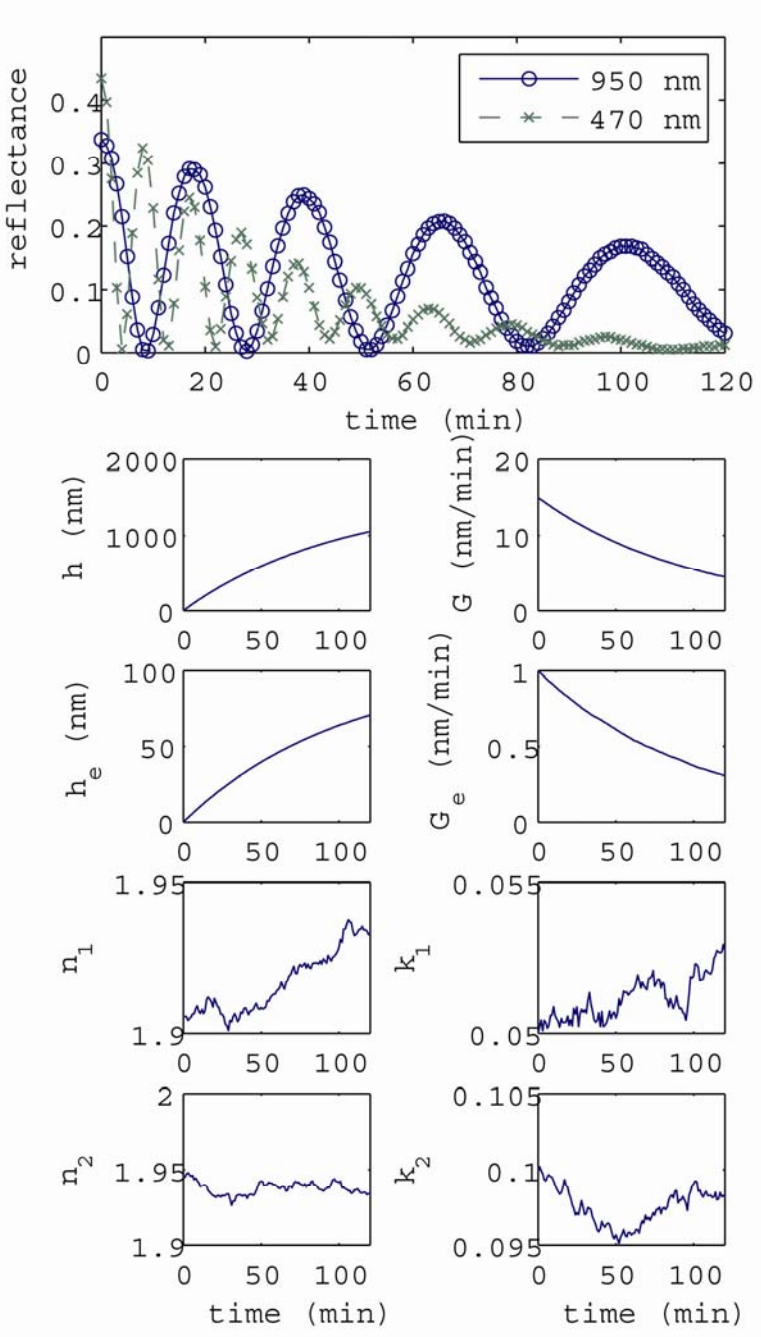
Greatly reduced computational time!

Robertson D.G. *et al.* 1996 *AIChE Journal* (1996) **42**, 2209-2224

Simulated data

Unmodeled drift in thickness and roughness

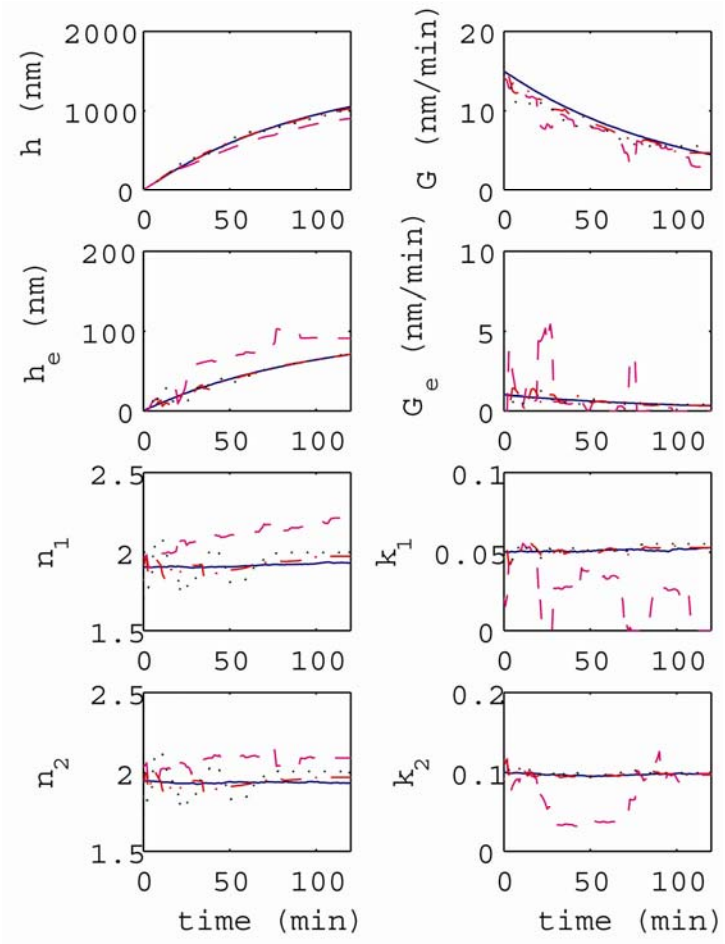
- Construct a simulation that approximates the experimentally observed behavior
- Can better evaluate the estimator, since the states are known
- Growth rate is drifting due to precursor depletion
- Surface is roughening due to grain growth dynamics



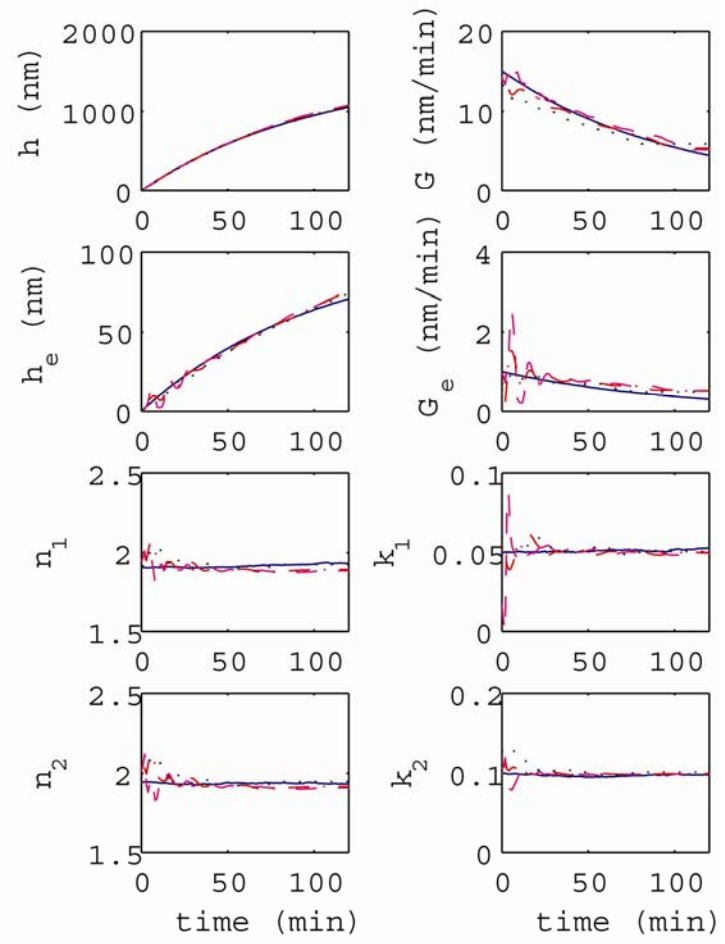
Comparison of methods

mMHE estimate is less oscillatory than RLS

RLS



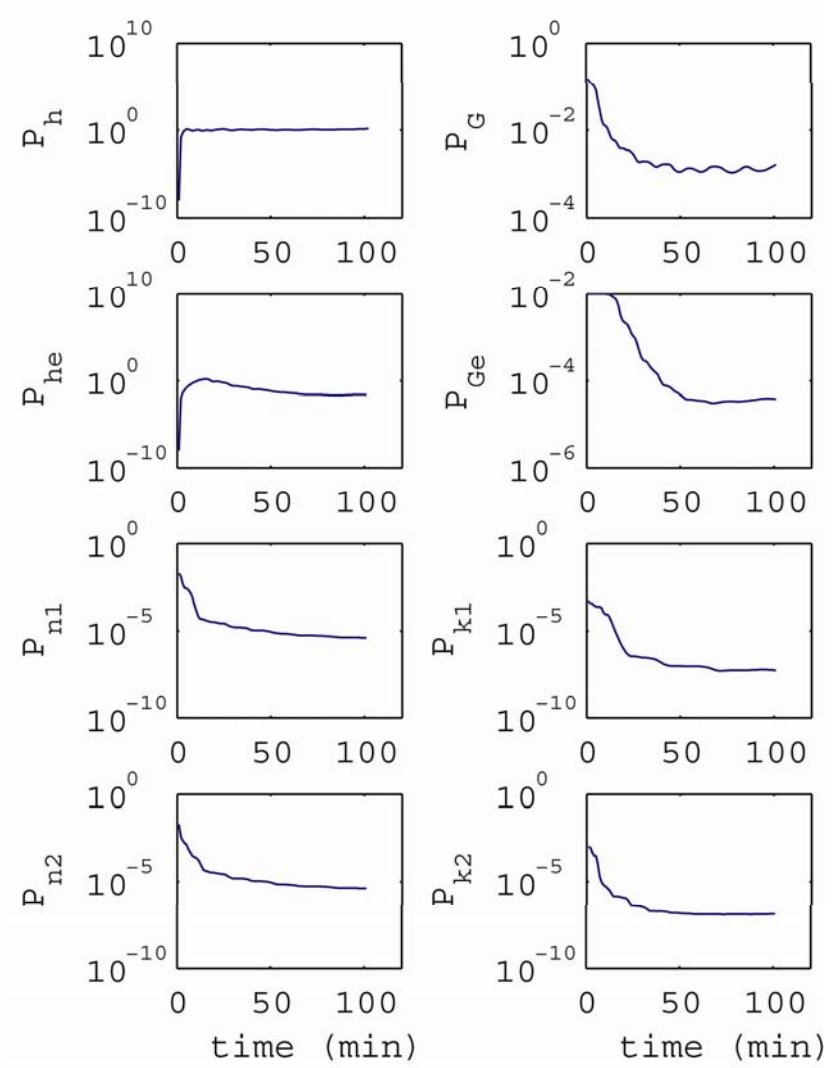
mMHE



mMHE variance

Predicted variance decays, except for the integrating states h and h_e

- Significant pitfall of RLS when states are correlated
- MHE considers this covariance between states
- Highly correlated states
  - Thickness and refractive index
  - Extinction coefficient and roughness



Quantification of performance

mMHE outperforms RLS and can be further optimized

- Compare normalized sum squared error over entire trajectory (average 5 runs)
- mMHE outperforms RLS for all horizon lengths
- EKF can also be an option, but longer horizon can improve estimates
- mMHE improves estimate over RLS except for large R and small Q
  - Process model is the primary error here

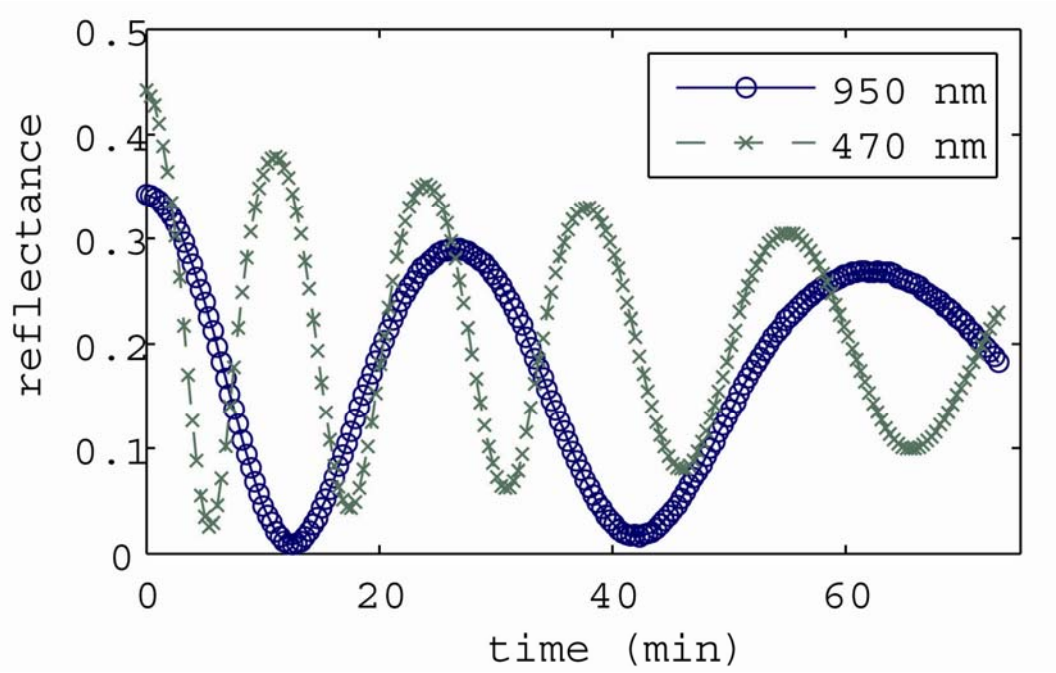
m	EKF	10	20	30	40
RLS	—	25	5.1	5.0	5.5
mMHE	6.0	6.1	4.0	3.4	2.8

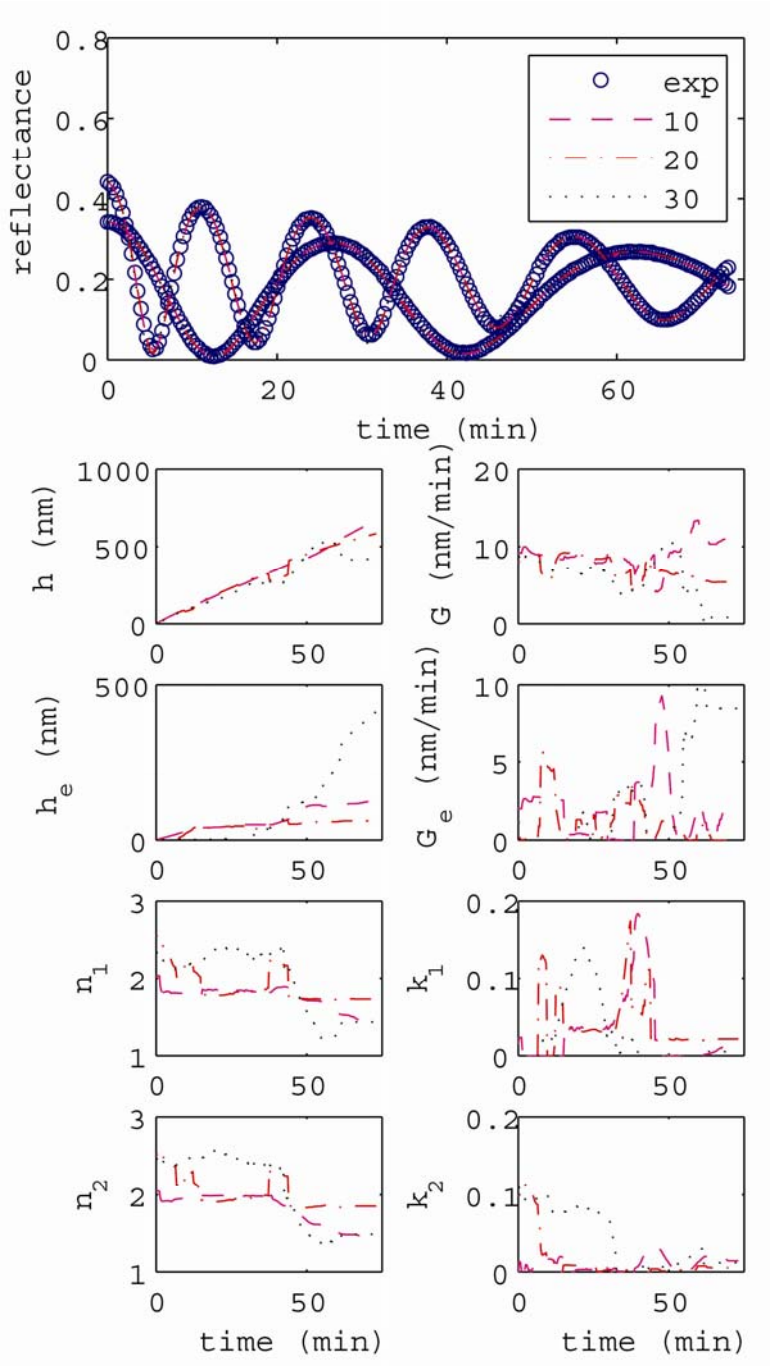
With m = 10

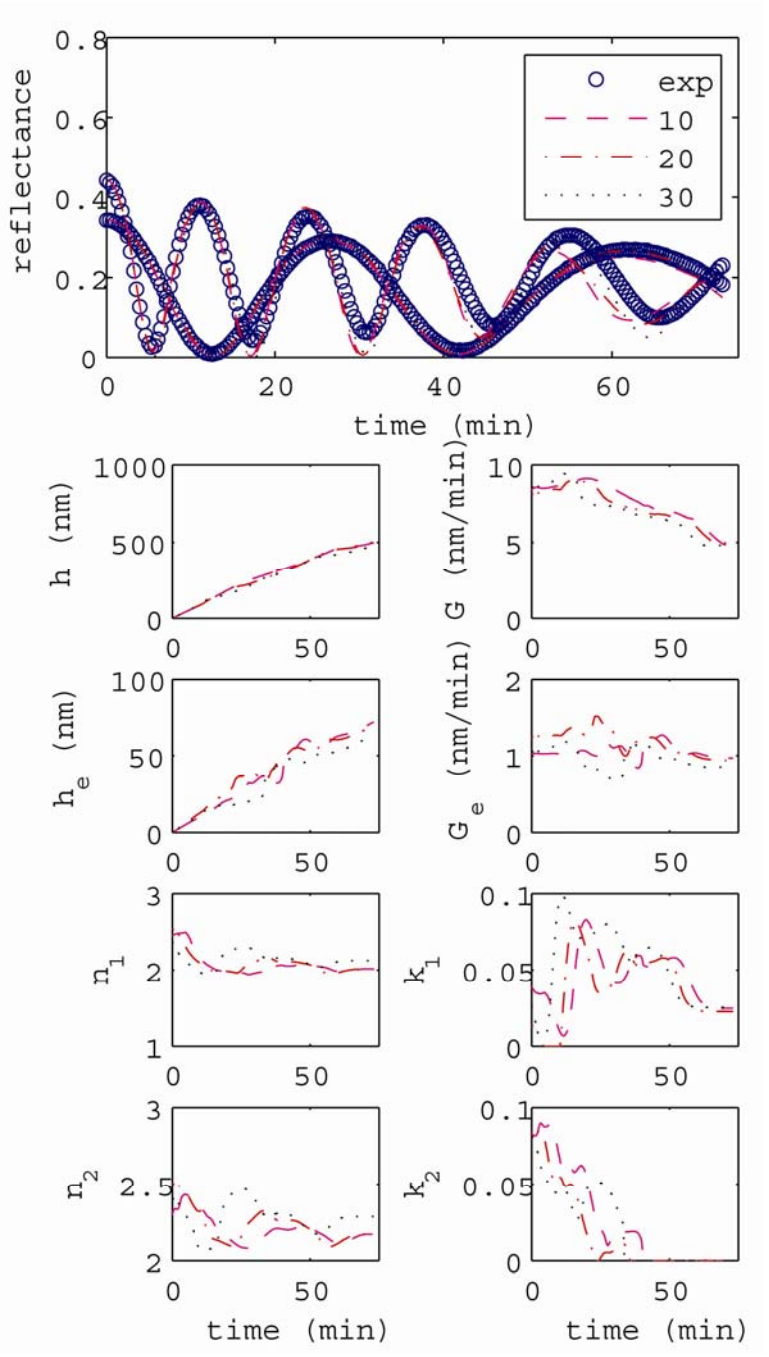
		$0.1P_{1\|0}$	$P_{1\|0}$	$10P_{1\|0}$
$0.1R$	$0.1Q$	5.45	8.98	9.71
	Q	3.02	7.97	9.47
	$10Q$	2.13	4.96	7.43
R	$0.1Q$	5.98	7.96	11.4
	Q	3.52	5.44	9.22
	$10Q$	2.06	3.17	7.97
$10R$	$0.1Q$	30.5	19.8	30.8
	Q	5.42	5.99	8.27
	$10Q$	3.25	3.52	5.45

Experimental data

- Reflectance data indicates a slowly decreasing growth rate




reflectance
exp
10
20
30
time (min)
h (nm)
G (nm/min)
h_e (nm)
G_e (nm/min)
n_1
k_1
n_2
k_2
time (min)
time (min)


reflectance
exp
10
20
30
time (min)
h (nm)
G (nm/min)
h_e (nm)
G_e (nm/min)
n_1
k_1
n_2
k_2
time (min)
time (min)

Comparison to *ex situ* measurements

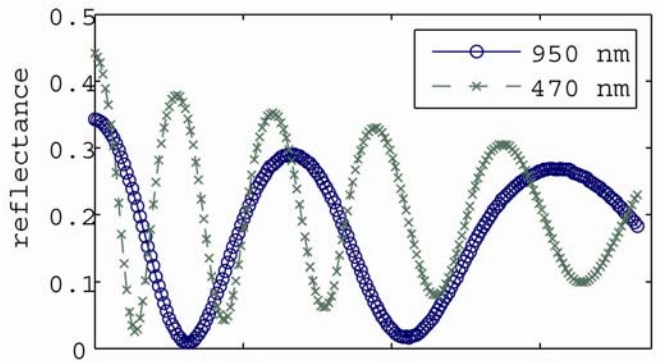
	RLS			mMHE				*ex situ*
	10	20	30	EKF	10	20	30	ellipsometry
h (nm)	605	555	421	490	498	541	477	510.0
h_e (nm)	119	52	413	42	72	69	63	60.4
n_1	1.6	1.9	1.4	2.1	2.0	1.9	2.1	1.8315
k_1	0.02	0.03	0.005	0.05	0.02	0.02	0.02	0.0012
n_2	1.7	2.0	1.5	2.3	2.2	2.0	2.3	1.8828
k_2	0.03	0.02	0.01	0.05	0	0.005	0	0

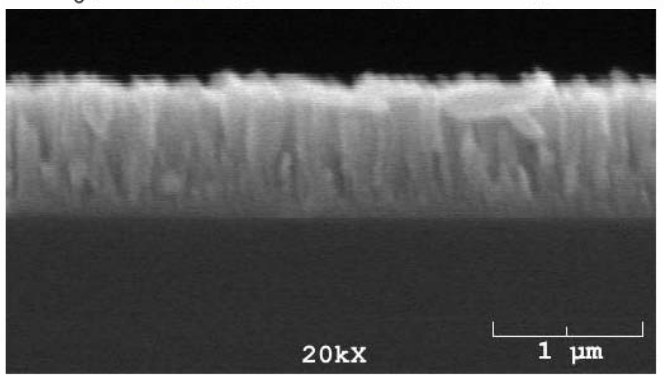
- mMHE matches ex-situ measurements better than RLS
- Extended Kalman filter performs well, except for roughness prediction.

R. Xiong and M. A. Grover, *Journal of Applied Physics* (2008)

Summary: *in situ* sensing

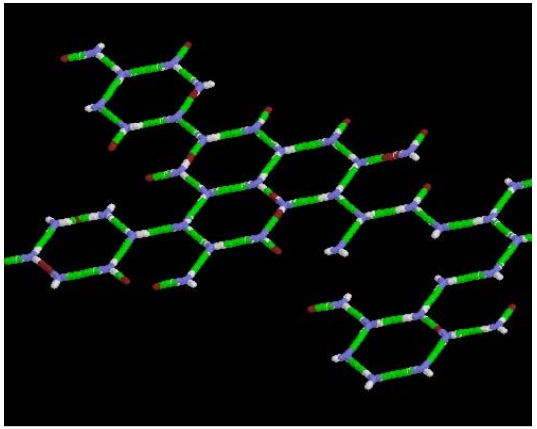
- MHE is the general least squares-based state estimation. RLS, mMHE, and EKF are special cases of MHE.
- mMHE was developed to address the computational issue of MHE. It combines advantages of RLS and MHE.
- In both simulated and experimental processes under different conditions, mMHE consistently yielded better estimates by utilizing
  - process dynamic model
  - sensor model
  - estimates of uncertainties
- mMHE is a robust estimator in terms of tuning matrices and the horizon size.

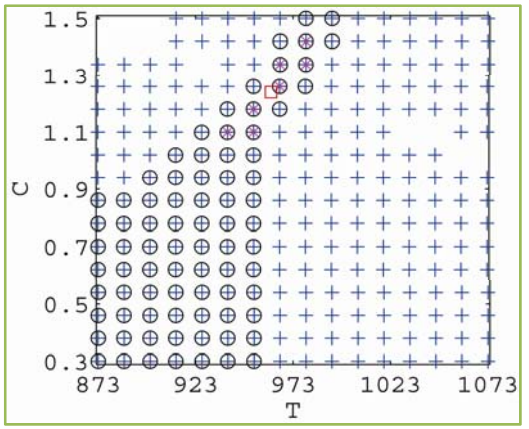




Overall conclusions

- There is a need for modeling in materials design
- Uncertainty in the models must be included
- Need a combination of modeling and statistics, coupled with domain knowledge and experiments
- Many possible areas for impact
  - Nanotechnology
  - Biological systems (systems biology)

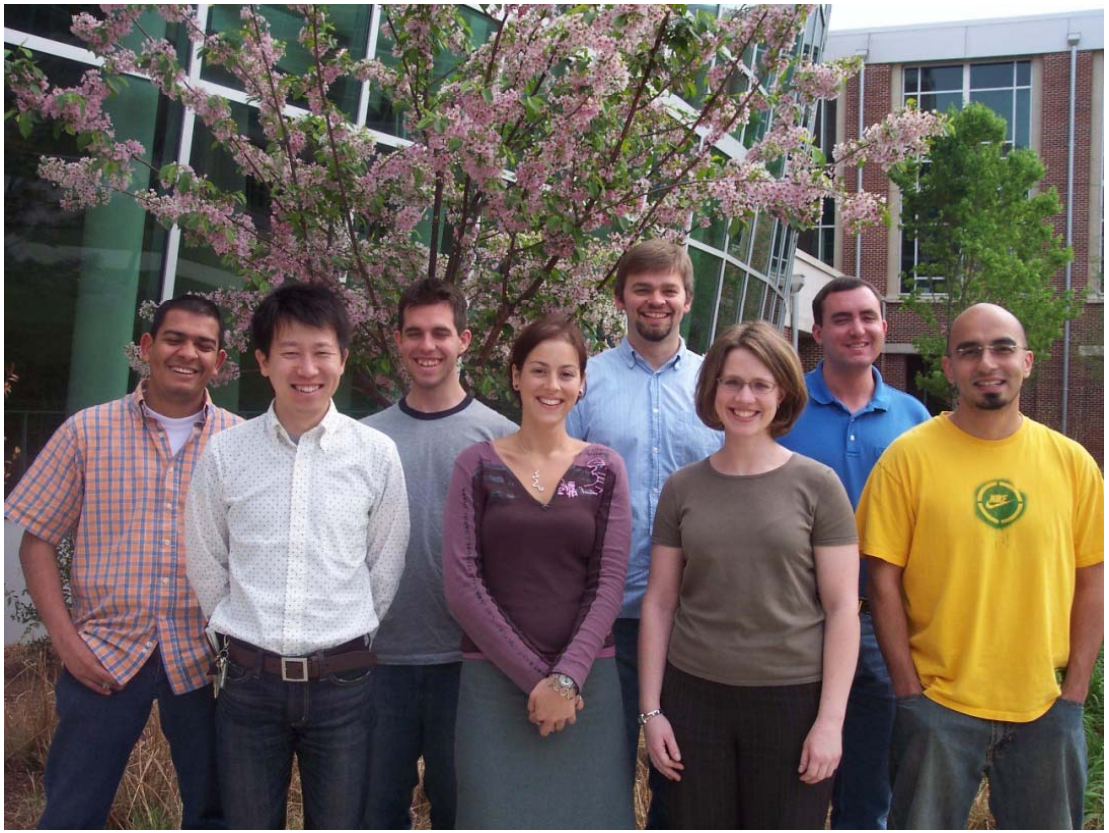




Acknowledgments

- Paul Wissmann
- Rentian Xiong
- Cihan Oguz

- Tim Long
- Iskender Yilgor
- Emel Yilgor



Funding:
NSF CAREER “A Systems Approach to Materials Processing,”
AFOSR, CIBA Vision, United Technologies UTC Power