OAK RIDGE NATIONAL LABORATORY

# *Ellipsometry*
# *Data Analysis: a Tutorial*

**G. E. Jellison, Jr.**

**Solid State Division**

**Oak Ridge National Laboratory**

**WISE 2000**

**University of Michigan**

**May 8-9, 2000**

UT-BATTELLE

# Motivation

**The Opportunity:**

**Spectroscopic Ellipsometry (SE) is sensitive to many parameters of interest to thin-film science, such as**

- **Film thickness**
- **Interfaces**
- **Optical functions (n and k).**

**But**

**SE data is not meaningful by itself.**

**Therefore**

**One must model the near-surface region to get useful information.**

# Outline

**Data representations**

**Calculation of reflection from thin films**

**Modeling of optical functions (n and k)**

**Fitting ellipsometry data**

**Examples**

# Introduction

## What we measure:

**Stokes Vector of a light beam**

$$\mathbf{S} = \begin{pmatrix} I_o \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I_o \\ I_0 - I_{90} \\ I_{45} - I_{-45} \\ I_{rc} - I_{lc} \end{pmatrix}$$

$$I_o \geq (Q^2 + U^2 + V^2)^{1/2}$$

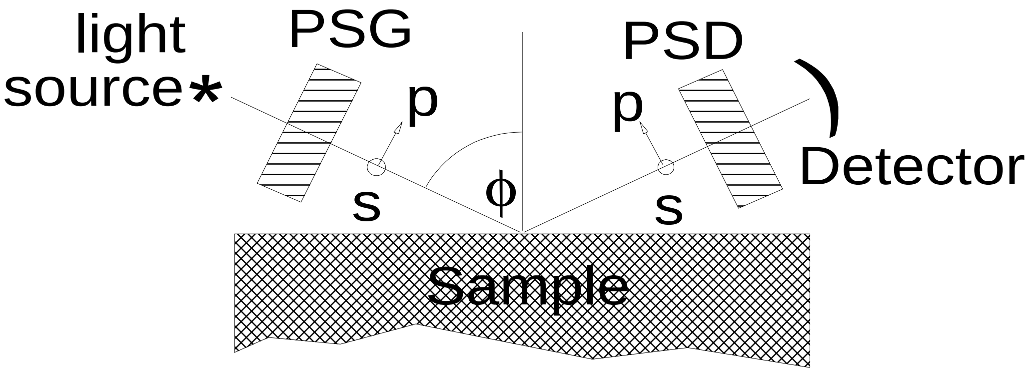
**To transform one Stokes vector to another:**

$$\mathbf{S}_{\text{out}} = \mathbf{M}\,\mathbf{S}_{\text{in}}$$

**M = Mueller matrix (4X4 real)**

# Introduction

## What we measure:



**Intensity of beam at the detector = $S_{PSD}^{T}\ M\ S_{PSG}$**

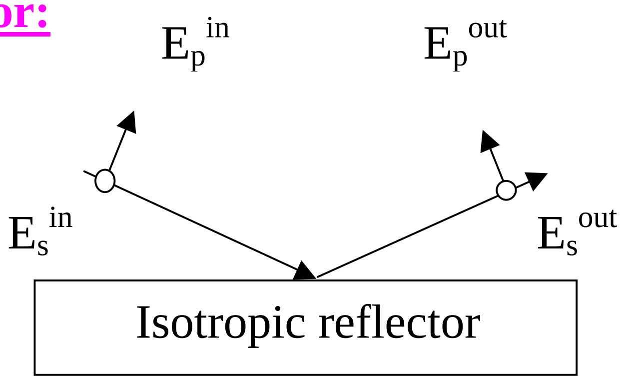
**The matrix M includes:**

**Sample reflection characteristics**

**Intervening optics (windows and lenses)**

# Introduction

**Isotropic Reflector:**



**Mueller Matrix:**

$$\mathbf{M}_{\mathbf{reflector}} = \begin{pmatrix} 1 & -N & 0 & 0 \\ -N & 1 & 0 & 0 \\ 0 & 0 & C & S \\ 0 & 0 & -S & C \end{pmatrix}$$

**$N = \cos(2\psi)$**

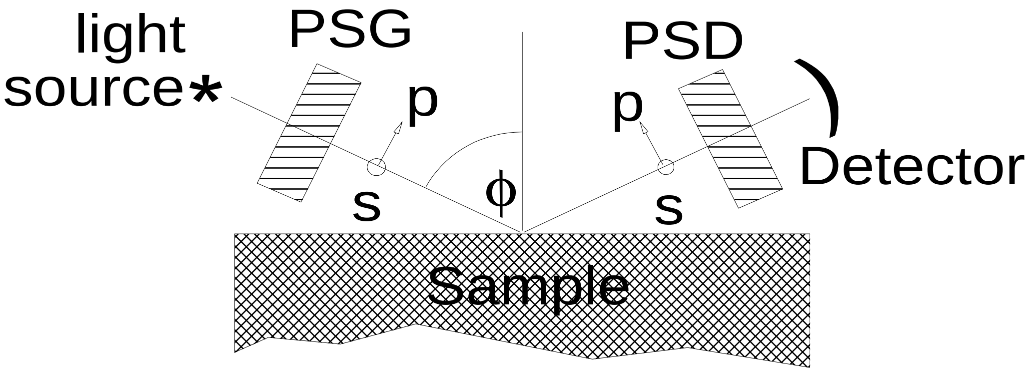
**$S = \sin(2\psi)\sin(\Delta)$**

**$C = \sin(2\psi)\cos(\Delta)$**

$$\mathbf{N^2 + S^2 + C^2 = 1}$$

# Introduction

## What we calculate:



$$\mathbf{J} = \begin{pmatrix} r_{pp} & r_{ps} \\ r_{sp} & r_{ss} \end{pmatrix} = \begin{pmatrix} \hat{E}^{o}_{p} / \hat{E}^{i}_{p} & \hat{E}^{o}_{s} / \hat{E}^{i}_{p} \\ \hat{E}^{o}_{p} / \hat{E}^{i}_{s} & \hat{E}^{o}_{s} / \hat{E}^{i}_{s} \end{pmatrix}$$

$$= r_{ss} \begin{pmatrix} \rho & \rho_{ps} \\ \rho_{sp} & 1 \end{pmatrix} = \begin{pmatrix} \tan(\psi) e^{i\Delta} & \tan(\psi_{ps}) e^{i\Delta_{ps}} \\ \tan(\psi_{sp}) e^{i\Delta_{sp}} & 1 \end{pmatrix}.$$

***There is no difference between Mueller and Jones IF there is no <u>depolarization</u>!***

# Data Representations

## Mueller-Jones matrices:

$$\mathbf{M} = \mathbf{A}\,(\,\mathbf{J} \otimes \mathbf{J}^{*}\,)\,\mathbf{A}^{-1},$$

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{pmatrix}.$$

**Assumes no depolarization!**

$$\rho = \frac{r_p}{r_s} = \frac{E_p^{out} \big/ E_p^{in}}{E_s^{out} \big/ E_s^{in}} = \tan(\psi)e^{i\Delta} = \frac{C + iS}{1 + N}$$

# Data Representations

## Anisotropic samples:

$$\mathbf{M} = \begin{pmatrix} 1 & -N-\alpha_{ps} & C_{sp}+\zeta_1 & S_{sp}+\zeta_2 \\ -N-\alpha_{sp} & 1-\alpha_{sp}-\alpha_{ps} & -C_{sp}+\zeta_1 & -S_{sp}+\zeta_2 \\ C_{ps}+\xi_1 & -C_{ps}+\xi_1 & C+\beta_1 & S+\beta_2 \\ -S_{ps}+\xi_2 & S_{ps}+\xi_2 & -S+\beta_2 & C-\beta_1 \end{pmatrix}$$

$$\rho_{ps} = \frac{C_{ps}+iS_{ps}}{1+N} = \tan(\psi_{ps})e^{i\Delta_{ps}}$$

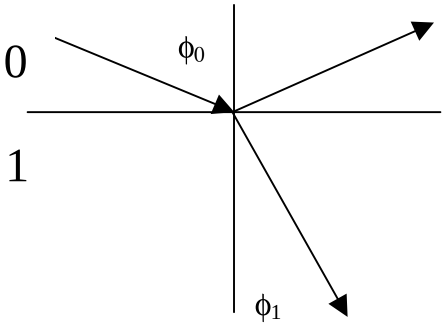
$$\rho_{sp} = \frac{C_{sp}+iS_{sp}}{1+N} = \tan(\psi_{sp})e^{i\Delta_{sp}}$$

$$\boldsymbol{N^2 + S^2 + C^2 + S_{ps}{}^2 + C_{ps}{}^2 + S_{sp}{}^2 + C_{sp}{}^2 = 1}$$

$$\mathbf{J} = r_{ss}\begin{pmatrix} \rho & \rho_{ps} \\ \rho_{sp} & 1 \end{pmatrix} = \begin{pmatrix} \tan(\psi)e^{i\Delta} & \tan(\psi_{ps})e^{i\Delta_{ps}} \\ \tan(\psi_{sp})e^{i\Delta_{sp}} & 1 \end{pmatrix}$$

# Calculation of Reflection Coefficients

## Single Interface: Fresnel Equations (1832):



$$r_{pp} = \frac{\tilde{n}_1 \cos(\phi_0) - \tilde{n}_0 \cos(\phi_1)}{\tilde{n}_1 \cos(\phi_0) + \tilde{n}_0 \cos(\phi_1)}$$

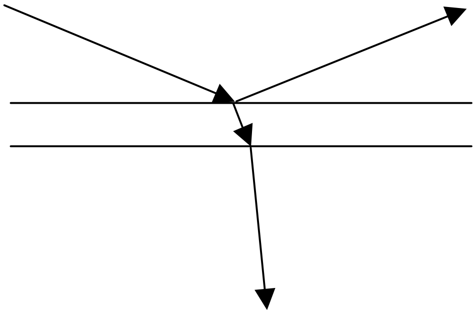
$$r_{ss} = \frac{\tilde{n}_0 \cos(\phi_0) - \tilde{n}_1 \cos(\phi_1)}{\tilde{n}_0 \cos(\phi_0) + \tilde{n}_1 \cos(\phi_1)}$$

## Snell's Law (1621):

$$\xi = \tilde{n}_0 \sin \phi_0 = \tilde{n}_1 \sin \phi_1.$$

# Calculation of Reflection Coefficients

## Two Interfaces: Airy Formula (1833):



$$r_{pp,ss} = \frac{r_{1,pp,ss} + r_{2,pp,ss} \exp(-2ib)}{1 + r_{1,pp,ss} r_{2,pp,ss} \exp(-2ib)}$$

$$b = \frac{2\pi d_f}{\lambda} \tilde{n}_f \cos(\phi_f)$$

# Calculation of Reflection Coefficients

## Three or more interfaces: Abeles Matrices (1950):

**Represent each layer by 2 Abeles matrices:**

$$\mathbf{P}_{j,pp} = \begin{pmatrix} \cos(b_j) & -i\dfrac{\cos(\phi_j)}{\tilde{n}_j}\sin(b_j) \\ i\dfrac{\tilde{n}_j}{\cos(\phi_j)}\sin(b_j) & \cos(b_j) \end{pmatrix}$$

$$\mathbf{P}_{j,ss} = \begin{pmatrix} \cos(b_j) & i\dfrac{\sin(b_j)}{\tilde{n}_j\cos(\phi_j)} \\ i\tilde{n}_j\cos(\phi_j)\sin(b_j) & \cos(b_j) \end{pmatrix}$$

**Matrix multiply:**

$$\mathbf{M}_{pp} = \chi_{0,pp}\Big(\prod_{j=1}^{N}\mathbf{P}_{j,pp}\Big)\chi_{sub,pp}$$

$$\mathbf{M}_{ss} = \chi_{0,ss}\Big(\prod_{j=1}^{N}\mathbf{P}_{j,ss}\Big)\chi_{sub,ss}$$

# Calculation of Reflection Coefficients

## Three or more interfaces: Abeles Matrices:

**Substrate and ambient characteristic matrices:**

$$\chi_{0,pp} = \frac{1}{2}\begin{pmatrix} 1 & \dfrac{\cos(\phi)}{\tilde{n}_0} \\ -1 & \dfrac{\cos(\phi)}{\tilde{n}_0} \end{pmatrix} \qquad \chi_{0,ss} = \frac{1}{2}\begin{pmatrix} 1 & \dfrac{1}{\tilde{n}_0\cos(\phi)} \\ 1 & \dfrac{-1}{\tilde{n}_0\cos(\phi)} \end{pmatrix}$$

$$\chi_{sub,pp} = \begin{pmatrix} \dfrac{\cos(\phi_{sub})}{\tilde{n}_{sub}} & 0 \\ 1 & 0 \end{pmatrix} \qquad \chi_{sub,ss} = \begin{pmatrix} \dfrac{1}{\tilde{n}_{sub}\cos(\phi_{sub})} & 0 \\ 1 & 0 \end{pmatrix}$$

**Final Reflection coefficients:**

$$r_{pp} = \frac{M_{21,pp}}{M_{11,pp}} \qquad r_{ss} = \frac{M_{21,ss}}{M_{11,ss}}$$

# Calculation of Reflection Coefficients

## Anisotropic materials:

**2 2X2 Abeles matrices become one 4X4 Berreman matrix (1972)**

**Reason: s- and p- polarization states are no longer eigenstates of the reflection.**

## Inhomogeneous layers:

**If a layer has a depth-dependent refractive index, there are two options:**

**1) Build up many very thin layers**

**2) Use interpolation approximations**

# Models for dielectric functions

## Tabulated Data Sets:

1) **Usually good for substrates**
2) **Not good for thin films**
3) **Even for substrates: problems**
    a) **surface roughness**
    b) **surface reconstruction**
    c) **surface oxides**
4) **Most tabulated data sets do not include error limits**

***Measured optical functions of silicon depend on the face!***

**[near 4.25 eV = 292 nm, $\varepsilon_2(100)<\varepsilon_2(111)$]**

# Models for dielectric functions

## Lorentz Oscillator Model (1895):

$$\varepsilon(\lambda) = \tilde{n}(\lambda)^2 = 1 + \sum_j \frac{A_j \lambda^2}{\lambda^2 - \lambda_{o,j}^2 + i\zeta_j \lambda}$$

$$\varepsilon(E) = \tilde{n}(E)^2 = 1 + \sum_j \frac{B_j}{E_{o,j}^2 - E^2 + i\Gamma_j E}$$

## Sellmeier Equation:

$$\varepsilon = n^2 = 1 + \sum_j \frac{A_j \lambda^2}{\lambda^2 - \lambda_{o,j}^2}$$

**For Insulators**

## Cauchy (1830):

$$n = B_0 + \sum_j \frac{B_j}{\lambda^{2j}}$$

## Drude (1890):

$$\varepsilon(E) = 1 - \sum_j \frac{B_j}{E} \left( \frac{1}{E - i\Gamma_j} \right)$$

**For Metals**

# Models for dielectric functions

## Amorphous Materials (Tauc-Lorentz):

$$\varepsilon_2(E) = 2n(E)k(E) = \frac{A(E-E_g)^2}{(E^2-E_o^2)^2+\Gamma^2}\,\frac{\Theta(E-E_g)}{E}$$

$$\varepsilon_1(E) = \varepsilon_1(\infty) + \frac{2}{\pi} P \int_{R_g}^{\infty} \frac{\xi\varepsilon_2(\xi)}{\xi^2 - E^2} d\xi$$

**Five parameters:**

| | |
|---|---|
| **$E_g$** | **Band gap** |
| **A** | **Proportional to the matrix element** |
| **$E_0$** | **Central transition energy** |
| **$\Gamma$** | **Broadening parameter** |
| **$\varepsilon_1(\infty)$** | **Normally = 1** |

# Forouhi and Bloomer

(Forouhi and Bloomer *Phys Rev*. B **34**, 7018 (1986).)

## Extinction coefficient:

$$k_{FB}(E) = \frac{A(E - E_g)^2}{E^2 - BE + C}$$

## Refractive index: **(a Hilbert transform)**

$$n_{FB}(E) = n(\infty) + \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{k(\xi) - k(\infty)}{\xi - E} d\xi$$

## Problems:

- **$k_{FB}(E) > 0$ for $E < E_g$. Unphysical.**
- **$k_{FB}(E) \rightarrow$ constant as $E \rightarrow \infty$. Experiment states that $k(E) \rightarrow 0$ as $1/E^3$ as $E \rightarrow \infty$.**
- **Time-reversal symmetry required [$k(-E) = -k(E)$].**
- **Hilbert Transform ≠ Kramers Kronig [$k(\infty) = 0$].**

# Models for dielectric functions

## Models for Crystalline materials:

Critical points, excitons, etc. in the optical spectra make this a very difficult problem!

Collections of Lorentz oscillators:

$$\varepsilon(E) = \varepsilon_o + \sum_j \frac{A_j e^{i\phi_j}}{E - E_j + i\Gamma_j}$$

Can end up with MANY terms

# Models for dielectric functions

## Effective Medium theories:

$$\frac{<\varepsilon> - \varepsilon_h}{<\varepsilon> + \gamma\varepsilon_h} = \sum_j f_j \frac{\varepsilon_j - \varepsilon_h}{\varepsilon_j + \gamma\varepsilon_h}$$

**Choice of host material:**

1) **Lorentz-Lorentz: $\varepsilon_h = 1$**
2) **Maxwell-Garnett: $\varepsilon_h = \varepsilon_1$**
3) **Bruggeman $\varepsilon_h = <\varepsilon>$**

**$\gamma$ Depolarization factor ~2.**

# Fitting Models to Data

## Figure of Merit:

**Experimental quantities:** $\rho_{\mathbf{exp}}(\lambda)$

**Calculated quantities:** $\rho_{\mathbf{calc}}(\lambda, \mathbf{z})$

**z: Vector of parameters to be fit (1 to m)**

**film thicknesses,**

**constituent fractions,**

**parameters of optical function models, etc.**

**Minimize**

$$\chi^2 = \frac{1}{N-m-1}\sum_{j=1}^{N}\frac{[\rho_{\exp}(\lambda_j)-\rho_{calc}(\lambda_j,\mathbf{z})]^2}{\sigma(\lambda_j)^2}$$

**$\sigma(\lambda)$ = random and systematic error**

# Fitting Models to Data

## Calculation Procedure:

1) **<u>Assume</u> a model.**
    a. **Number of layers**
    b. **Layer type (isotropic, anisotropic, graded)**
2) **Determine or <u>parameterize</u> the optical functions of each layer**
3) **Select reasonable starting parameters.**
4) **<u>Fit</u> the data, using a suitable algorithm and *<u>Figure of Merit</u>***
5) **Determine correlated errors in z and cross correlation coefficients**

**If the *<u>Figure of Merit</u>* indicates a bad fit (e.g. $\chi^2>>1$), go back to 1).**

# Fitting Models to Data

**“…fitting of parameters is not the end-all of parameter estimation. To be genuinely useful, a fitting procedure should provide (i) parameters, (ii) error estimates on the parameters, and (iii) a statistical measure of goodness-of-fit. When the third item suggests that the model is an unlikely match to the data, then items (i) and (ii) are probably worthless. Unfortunately, many practitioners of parameter estimation never proceed beyond item (i). They deem a fit acceptable if a graph of data and model “looks good.” This approach is known as *chi-by-eye*. Luckily, its practitioners get what they deserve.”**

**Press, Teukolsky, Vettering, and Flannery, *Numerical Recipes* (2nd ed., Cambridge, 1992), Ch. 15, pg. 650.**

# Fitting Models to Data

## Levenberg-Marquardt algorithm:

**Curvature matrix:**

$$\alpha_{kl} = \frac{1}{2}\frac{\partial^2 \chi^2}{\partial z_k \partial z_l} = \sum_{j=1}^{N} \frac{1}{\sigma(\lambda_j)^2} \frac{\partial \rho_{calc}(\lambda_j, \mathbf{z})}{\partial z_k} \frac{\partial \rho_{calc}(\lambda_j, \mathbf{z})}{\partial z_l}$$

**Inverse of curvature matrix: $\varepsilon = \alpha^{-1}$**

**Error in parameter $z_j = \varepsilon_{jj}$.**

**Cross correlation coefficients: proportional to the elements of $\varepsilon$.**

# Fitting Models to Data

## Meaning of fitted parameters and errors:

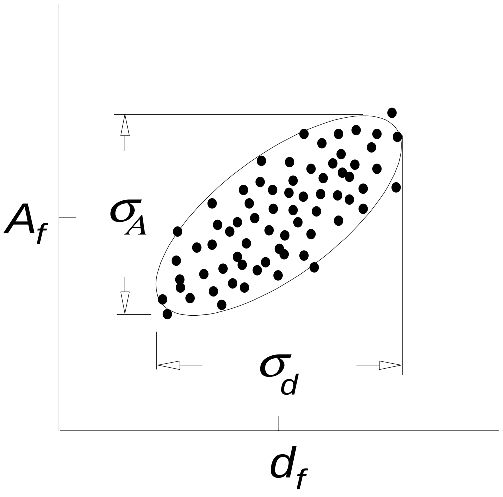
**Assume air/$SiO_2$/Si structure**

**Parameterize $SiO_2$ using Sellmeier dispersion**

**($\lambda_o$=93 nm):**

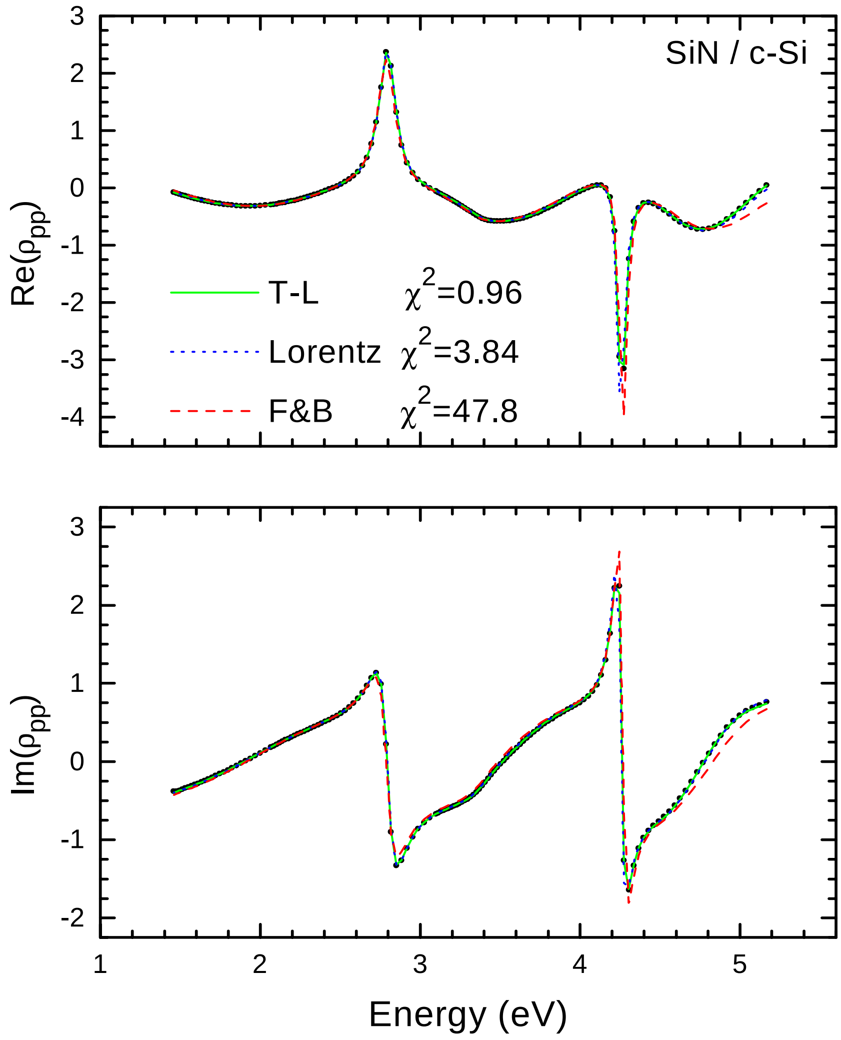
$$\varepsilon = n^2 = 1 + \sum_j \frac{A_j \lambda^2}{\lambda^2 - \lambda_{o,j}^2}$$

**Two fit parameters: $d_f$, $A_f$**



# Fitting Models to Data

## An Example: a-Si$_x$N$_y$:H on silicon:



# Fitting Models to Data

## An Example: a-$Si_xN_y$:H on silicon:

**Model:** 1) air

2) surface roughness

Bruggeman EMA (50% air, 50% a-SiN)

3) a-SiN (3 models)

Lorentz

Forouhi and Bloomer

Tauc-Lorentz

4) interface

Bruggeman EMA (50% Si, 50% a-SiN)

5) silicon

**Fitting parameters:** $d_2$, $d_3$, $d_4$, A, $E_o$, $\Gamma$, $\varepsilon(\infty)$ and

$E_g$ (F&B and T-L)

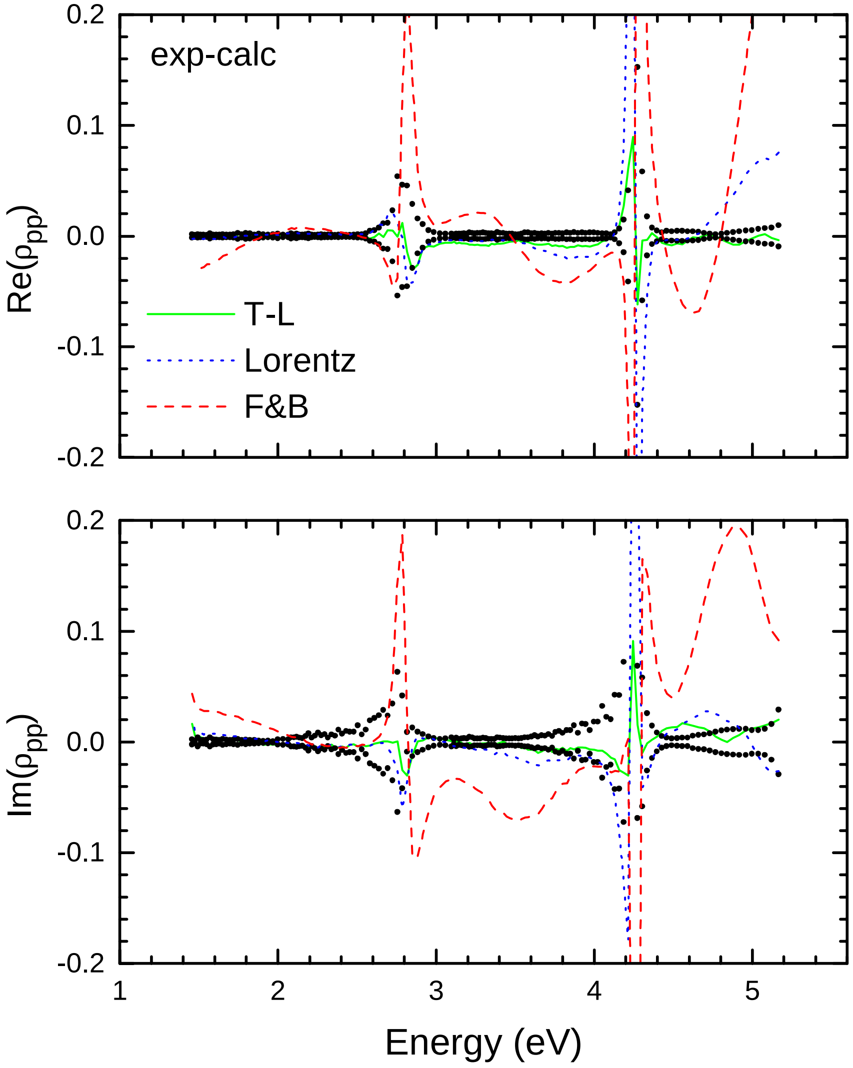
# Fitting Models to Data

## An Example: a-$Si_xN_y$:H on silicon:

| | **Lorentz** | **F&B** | **T-L** |
|---|---|---|---|
| Roughness thick (nm) | 2.1±0.3 | 4.9±0.7 | 1.9±0.3 |
| Film thickness (nm) | 197.8±0.7 | 195.6±1.1 | 198.2±0.4 |
| Interface thick (nm) | 0.6±0.3 | -0.6±0.6 | -0.1±0.2 |
| A | 201.9±4.6 | 4.56±1.9 | 78.4±12.7 |
| $E_o$ (eV) | 9.26±0.05 | 74.4±30.5 | 8.93±0.47 |
| $\Gamma$ (eV) | 0.01±0.02 | 0.74±23.5 | 1.82±0.81 |
| $E_g$ (eV) | ---- | 2.85±0.48 | 4.35±0.09 |
| $\varepsilon(\infty)$ | 1.00±0.02 | 0.93±0.51 | 1.38±0.26 |
| $\chi^2$ | 3.64 | 47.5 | 0.92 |
| Roughness thick (nm) | 2.4±0.3 | 4.7±0.6 | 1.8±0.2 |
| Film thickness (nm) | 198.6±0.4 | 195.3±0.9 | 198.1±0.3 |
| A | 202.5±1.4 | 5.03±0.35 | 97.7±3.2 |
| $E_o$ (eV) | 9.26±0.02 | 70.7±2.2 | 9.61±0.0.03 |
| $\Gamma$ (eV) | 0.01±0.01 | 40.1±10.8 | 3.07±0.33 |
| $E_g$ (eV) | ---- | 2.97±0.27 | 4.44±0.04 |
| $\chi^2$ | 3.75 | 48.0 | 0.96 |

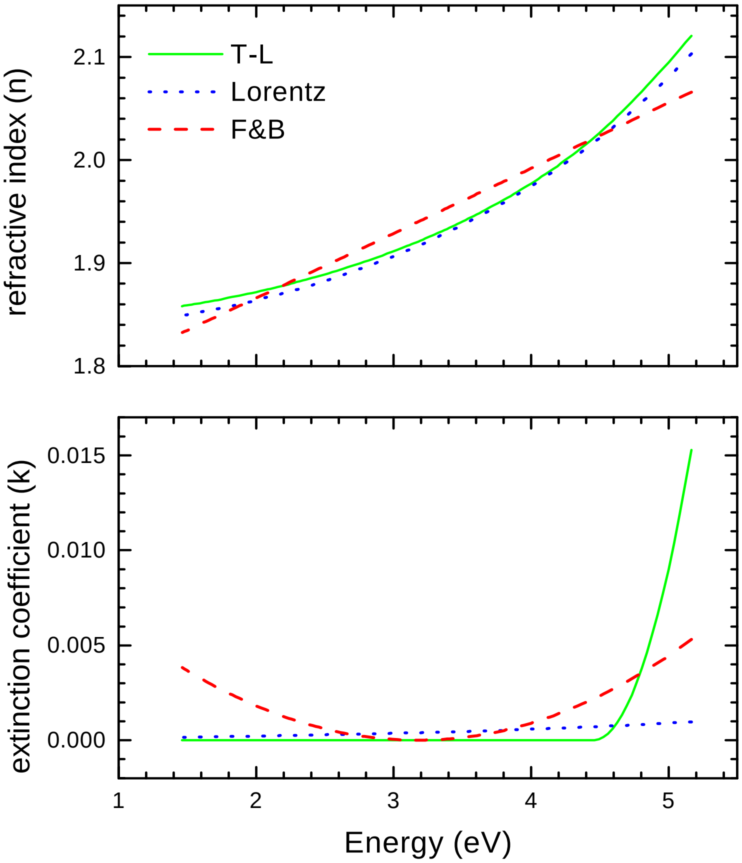
# Fitting Models to Data

## An Example: a-$Si_xN_y$:H on silicon:



# Optical Functions from Ellipsometry

## Optical Functions from Parameterization:



**Error limits:**

**Use the submatrix $\alpha_s$ from the associated fitted parameters.**

# Optical Functions from Ellipsometry

## Newton-Raphson algorithm:

**Solve:**

$$\mathrm{Re}(\rho_{calc}(\lambda, \phi, n_f, k_f, \ldots)) - \mathrm{Re}(\rho_{exp}(\lambda)) = 0$$

$$\mathrm{Im}(\rho_{calc}(\lambda, \phi, n_f, k_f, \ldots)) - \mathrm{Im}(\rho_{exp}(\lambda)) = 0$$

**Jacobian:**

$$\mathbf{J} = \begin{pmatrix} \dfrac{\partial \rho_{re}}{\partial n} & \dfrac{\partial \rho_{re}}{\partial k} \\ \dfrac{\partial \rho_{im}}{\partial n} & \dfrac{\partial \rho_{im}}{\partial k} \end{pmatrix}$$

$n_{new} = n_{old} + \delta n$; **where** $\delta n = -J^{-1} \rho$

**Propagate errors!**

# Optical Functions from Ellipsometry

## Optical functions of semiconductors:

**Dielectric function from air/substrate system:**

$$\varepsilon = \varepsilon_1 + i\varepsilon_2 = \sin^2(\phi)\{1 + [\frac{1-\rho}{1+\rho}]^2 \tan^2(\phi)\}$$

**Only valid if there is no overlayer (almost never true)**

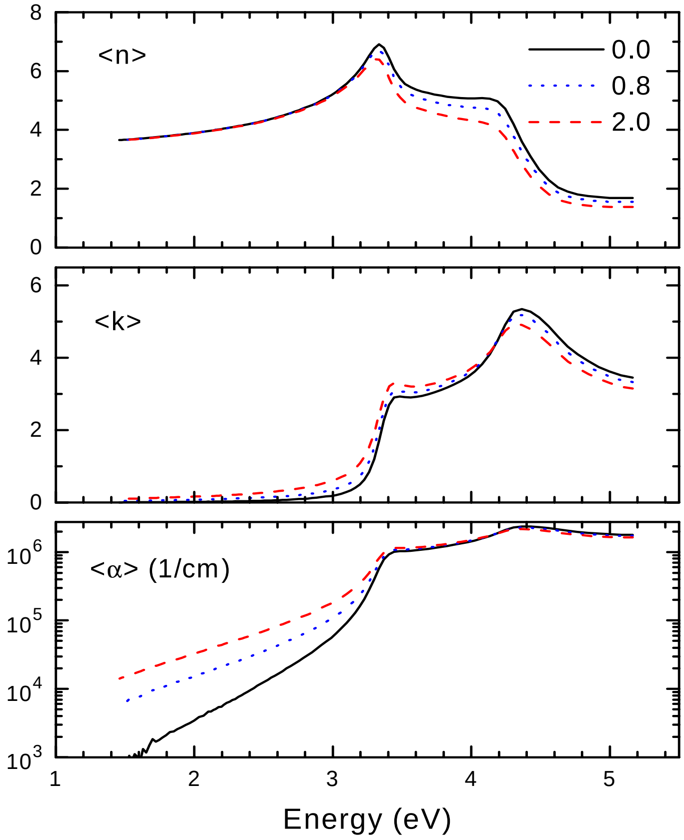
**If there is a thin film, Drude showed:**

$$\Delta = \Delta_o(n_s, k_s) + K \frac{4\pi d}{\lambda} \frac{n_f^2 - 1}{n_f^2}$$

# Optical Functions from Ellipsometry

## Optical functions of semiconductors:

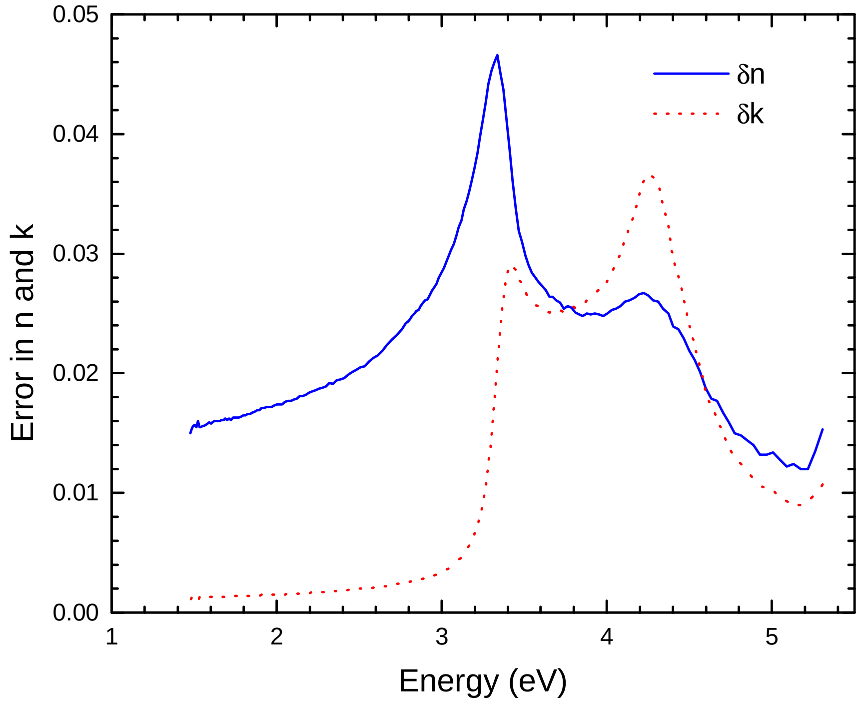
**Pseudo-dielectric functions of silicon with 0, 0.8, and 2.0 nm $SiO_2$ overlayers.**



# Optical Functions from Ellipsometry

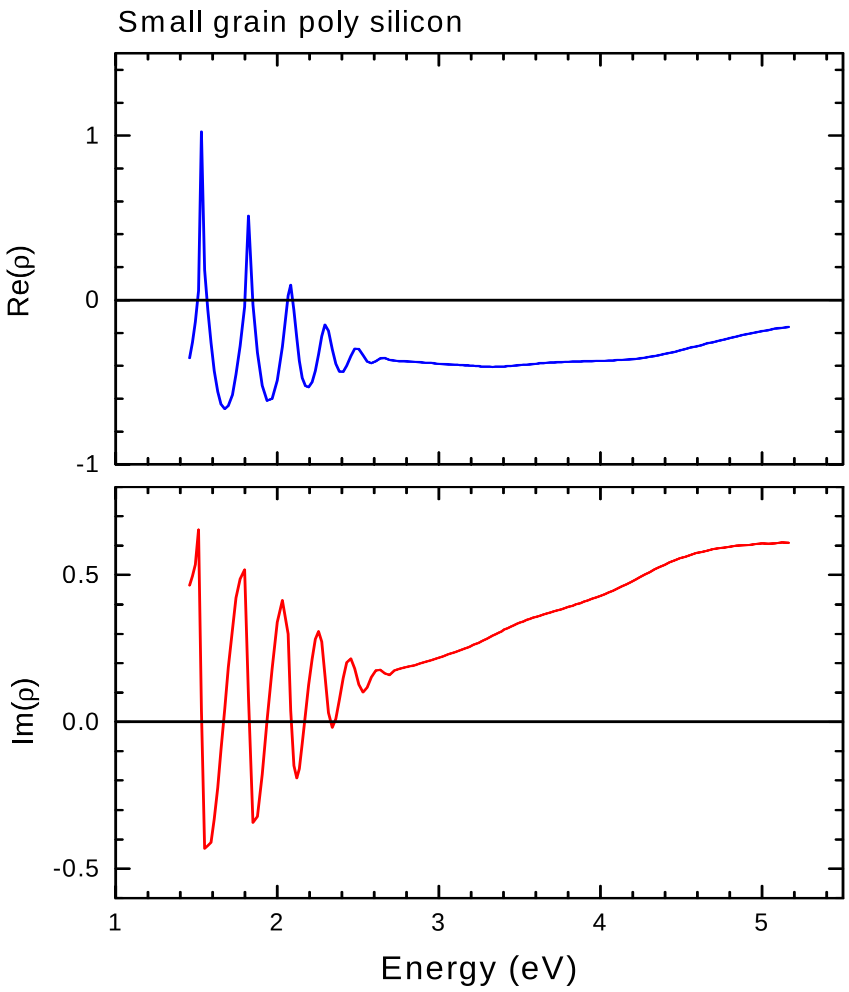
## Optical functions of semiconductors:

### Error limits of the dielectric function of silicon:



# Optical Functions from Ellipsometry

## Optical functions of thin films:



# Optical Functions from Ellipsometry

## Optical functions of thin films:

**Method of analysis:**

**A. Restrict analysis region to interference oscillations.**

**Parameterize the optical functions of the film.**

**1 air**

**2 surface roughness (BEMA)**

**3 T-L model for film**

**4 Lorentz model for a-$SiO_2$**

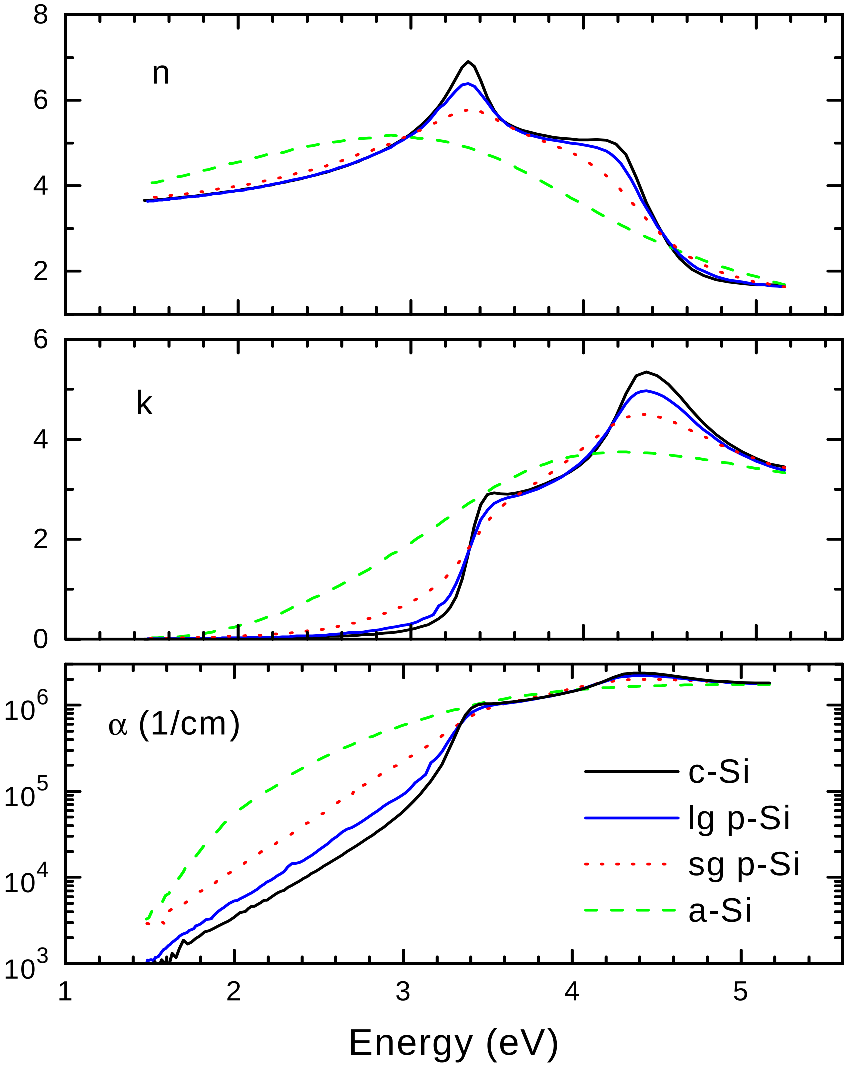
**5 c-Si**

**B. Fit data to determine thicknesses and Lorentz model parameters of a-$SiO_2$.**

**C. Calculate optical functions of thin film using Newton-Raphson.**

# Optical Functions from Ellipsometry

## Optical functions of thin films:



# Parting Thoughts

**SE is a powerful technique, but modeling is critical.**

**Modeling should use an error-based figure of merit**
**Does the model fit the data?**

**Calculate correlated errors and cross-correlation coefficients.**

**When used properly, SE gives very accurate thicknesses and values of the complex dielectric function.**